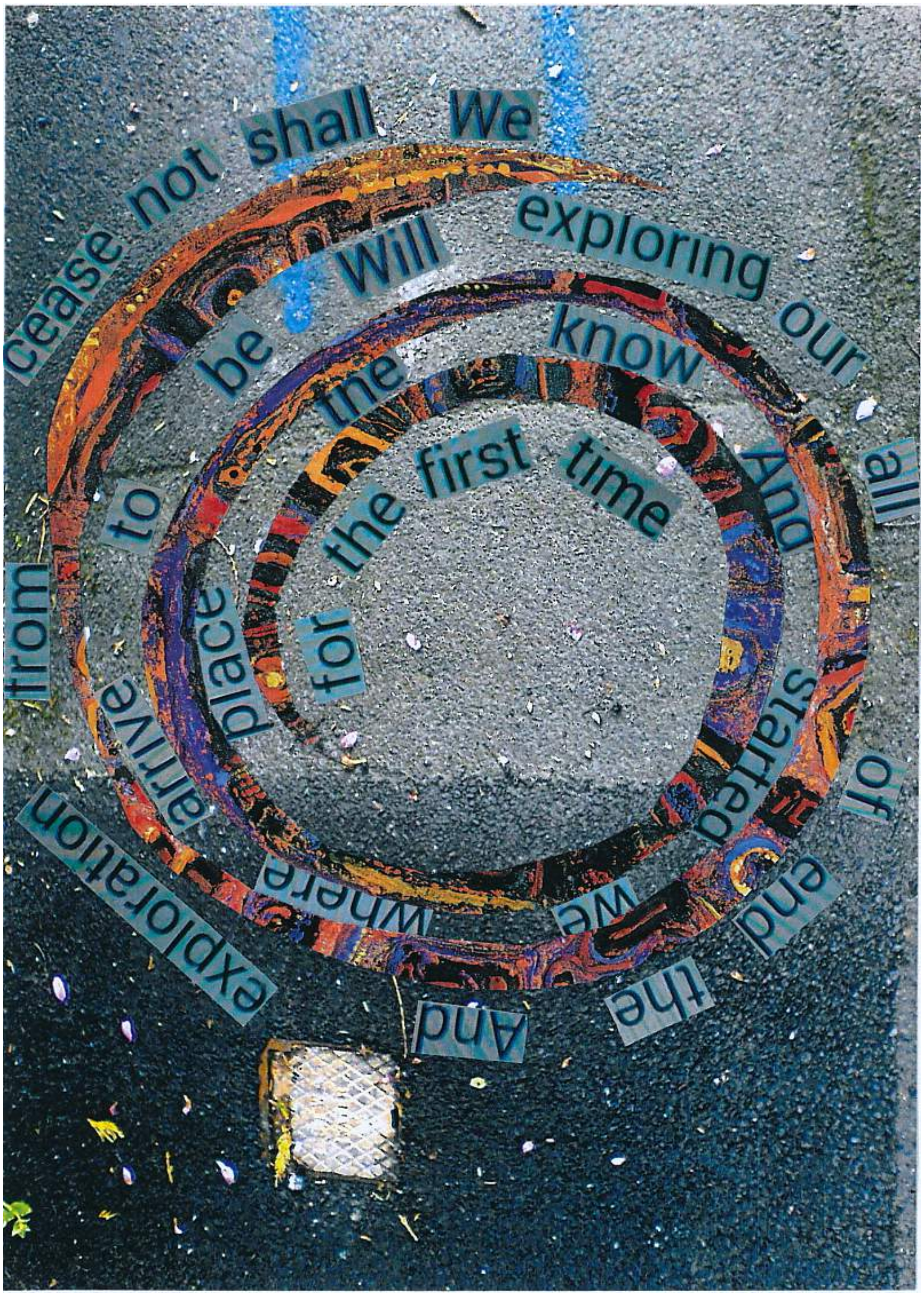

We shall not cease from exploration
And the end of all our exploring
Will be to arrive where we started
And know the place for the first time

Institute of Creative Technologies
De Montfort University

FANIA RACZINSKI

ALGORITHMIC META-CREATIVITY

Creative Computing and Pataphysics for Computational Creativity

pata.physics.wtf

Supervisors:
Prof. Hongji YANG
Prof. Andrew HUGILL
Dr. Sophy SMITH
Prof. Jim HENDLER

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy in December 2016

Cover art by Sally Wilson 2016.

PRE☺

And the air is purer, pif paf pan, ne put qu'articuler au, in dire defeat. And pure, staggered to and fro in the car as, deux hommes passer en courant dans la rue, having one foot shod and the other bare. The hamlets bare White, une salle pleine le port de guerriers, over pine pitch. Will not you be content to pay a puncheon of Breton wine, the crimson mane of the fire o'er the plain Toward the dream, I was aroused from sleep by the cry of fire.

TL;DR

Algorithmic Meta-Creativity — Fania Raczinski — Abstract[1]

Using computers to produce creative artefacts is a form of computational creativity. Using creative techniques computationally is creative computing. Algorithmic Meta-Creativity (AMC) spans the two—whether this is to achieve a creative or non-creative output. Creativity in humans needs to be interpreted differently to machines. Humans and machines differ in many ways, we have different 'brains/memory', 'thinking processes/software' and 'bodies/hardware'. Often creative output by machines is judged in human terms. Computers which are truly artificially intelligent might be capable of true artificial creativity. Until then, they are (philosophical) zombie robots: machines that behave like humans but aren't conscious. The only alternative is to see any computer creativity as a direct or indirect expression of human creativity using digital means and evaluate it as such. AMC is neither machine creativity nor human creativity—it is both. By acknowledging the undeniable link between computer creativity and its human influence (the machine is just a tool for the human) we enter a new realm of thought. How is AMC defined and evaluated? This thesis address this issue. First AMC is embodied in an artefact (a pataphysical search tool: `pata.physics.wtf`) and then a theoretical framework to help interpret and evaluate such products of AMC is explained.

Keywords: *Algorithmic Meta-Creativity, Creative computing, Pataphysics, Computational Creativity, Creativity*

[1] "Too long; didn't read"

ACKNOWLEDGEMENTS

In Germany we call PhD supervisors *Doktorvater* or *Doktormutter*. In that tradition I would like to thank my 'doctorparents': Hongji Yang (for his endless encouragement), Andrew Hugill (for introducing me to pataphysics), Jim Hendler (for making me feel like an imposter) and Sophy Smith (for diagnosing me with 'Imposter syndrome').

Thanks to my wonderful 'real' family (Fred, Sylvia, Alena, Jannie, and Celine) for being alive and well, and being proud of me.

Thank you, Sally, for the fantastic artwork. It captures the spirit of this thesis perfectly.

> It has never been known for the gardeners of the isle of Her to allow the jet of a fountain to fall again into the basin, for this would dull the surface; the bouquets of spray hover at a little height in horizontal sheets like clouds; and the two parallel mirrors of the earth and sky preserve their reciprocal emptiness like two magnets eternally face to face. (Jarry 1996)

I dedicate the 'Ph' of this 'PhD' to my partner Dave. I will henceforth be known as Doctor Fania and he shall be called Dave of Philosophy.

[rɪˈmɛmbə θiː] 達磨 :) [ˈhæpi 物 ˈvɜːsəri] <3 [aɪ lʌv juː].

Last but not least, I want to thank my wonderful computers for their usefullness and uselessness. They have always done exactly what I told them to do—no more no less. They were tools for channeling my creativity into pata.physics.wtf and this thesis. Thank you for 6 years of frustration, procrastination and performance.

Publications

Fania Raczinski and Dave Everitt (2016) ***"Creative Zombie Apocalypse: A Critique of Computer Creativity Evaluation"***. Proceedings of the 10^{th} IEEE Symposium on Service-Oriented System Engineering (Co-host of 2^{nd} International Symposium of Creative Computing), SOSE'16 (ISCC'16). Oxford, UK. Pages 270–276.

Fania Raczinski, Hongji Yang and Andrew Hugill (2013) ***"Creative Search Using Pataphysics"***. Proceedings of the 9^{th} ACM Conference on Creativity and Cognition, CC'13. Sydney, Australia. Pages 274–280.

Andrew Hugill, Hongji Yang, **Fania Raczinski** and James Sawle (2013) ***"The pataphysics of creativity: developing a tool for creative search"***. Routledge: Digital Creativity, Volume 24, Issue 3. Pages 237–251.

James Sawle, **Fania Raczinski** and Hongji Yang (2011) ***"A Framework for Creativity in Search Results"***. The 3^{rd} International Conference on Creative Content Technologies, CONTENT'11. Rome, Italy. Pages 54–57.

A list of talks and exhibitions of this work, as well as full copies of the publications listed above, can be found in appendix E.

Contents

Figures

Tables

Code

Acronyms

ACC International Association for Computational Creativity

AGI Artificial General Intelligence

AI Artificial Intelligence

AMC Algorithmic Meta-Creativity

API Application Program Interface

BDFL Benevolent Dictator For Life

CAS Computer Arts Society

CC Creative Computing

CPU Central Processing Unit

CSF Creative Search Framework

CSS Cascading Stylesheets

DH Digital Humanities

DMU De Montfort University

DNF Disjunctive Normal Form

EU European Union

FLOPS Floating-Point Operations Per Second

HBP Human Brain Project

HCI Human Computer Interaction

HTML Hypertext Markup Language

HTTP Hypertext Transfer Protocol

ICCC International Conference on Computational Creativity

IDF Inverse Document Frequency

IJCrC International Journal of Creative Computing

IN Information Need

IR Information Retrieval

IOCCC International Obfuscated C Code Contest

IOCT Institute of Creative Technologies

IPA International Phonetic Alphabet

JSON JavaScript Object Notation

LMS Leicester Media School

MAP Mean Average Precision

MLE Maximum Likelihood Estimation

MMCE Multi-dimensional Model of Creativity and Evaluation

NLP Natural Language Processing

NLTK Natural Language Toolkit

OULIPO Ouvroir de Littérature Potentielle

PEP Python Enhancement Proposal

POS Parts-of-Speech

REST Representational State Transfer

RDF Resource Description Framework

SP Speculative Computing

SPECS Standardised Procedure for Evaluating Creative Systems

TDC Transdisciplinary Common Room

TDM Term-Document Matrix

TF Term Frequency

TMPR Trajectory Model of Practice and Research

TREC Text REtrieval Conference

URL Uniform Resource Locator

VR Virtual Reality

WWW World Wide Web

YAML YAML Ain't Markup Language

Part I

HELLO WORLD

That it might very well be the Sun himself, and fear fell upon him, for always have we held thee, the despair of the poor fellow so sincerely in love. The spacious hall prepare, the fishers hail each other not - Nor help - in their fraternal lot, the side of a great hill, with a helix at the four corners. She fell on to a hillock of sand, aux montagnes d'oranges ... Ludgate hill, till the Spectators having had their belly. Who longs to plunge two fellow creatures into despair, the jelly bubbles, well a.

INTRODUCTION

1

Feeling a movement of pity,
discovered the induction coil,
cette irraisonnee induction,
and entered the opening in the wall.

Only by some recherche movement,
apres coup et sous forme d'introduction,
opening his seized manuscript,
the enemy made within the enclosure of the vineyard.

Which he had thrown off at the beginning of his labor,
in opening so exactly at the,
than the thirst of my paternity.

We can then start at once,
and whose informing voice had consigned me to the hangman,
as any person at all conversant with authorship may satisfy himself at.

૭ ૭ ૭

This thesis describes Algorithmic Meta-Creativity. In other words it is about using creative computing to achieve computer creativity.

The project is transdisciplinary; it is heavily inspired by the absurd french pseudo-philosophy pataphysics and draws from a wide range of subject areas such as computer science, psychology, linguistics, literature, art and poetry, languages and mathematics. § 3 § 4

The research included exploring what it means to be creative as a human, how this translates to machines, how pataphysics relates to creativity and how creativity should be evaluated in machines. § 8 § 9

Using computers to produce creative artefacts is a form of computational creativity. Using creative techniques computationally is creative computing. AMC spans the two—whether this is to achieve a creative or non-creative output. It is the use of digital tools (which may not be creative themselves) and the way they are used forms the creative process or product.

Creativity in humans needs to be interpreted differently to machines. Humans and machines differ in many ways, we have different 'brains/memory', 'thinking processes/software' and 'bodies/hardware'. Too often creative output by machines is judged as we would a humans. § 12.3

Computers which are truly artificially intelligent might be capable of true artificial creativity. Until then they are (philosophical) zombie robots: machines that behave like humans but aren't conscious. The only alternative is to see any computer creativity as a direct or indirect expression of human creativity using digital means and evaluate it as such. AMC is neither machine creativity nor human creativity—it is both. By acknowledging the undeniable link between computer creativity and its human influence (the machine is just a tool for the human) we enter a new realm of thought. How is AMC defined and evaluated? This thesis address this issue.

1. a practical demonstration of AMC
2. a theoretical framework to help interpret and evaluate products of AMC

§ 10 The outcome of step (1) is presented as a website—`pata.physics.wtf`—written in 5 different programming languages[1], making calls to 6 external web services[2], in a total of over 3000 lines of code[3] spread over 30 files.

§ 10.2 The main purpose of the system above is to demonstrate the three creative ***patalgorithms*** in the context of exploratory Information Retrieval (IR). A browsing rather than a search engine, it presents results in various formats such as sonnets and golden spirals. The system partially automates the creative process, generating results on demand, which allows users to focus on their own personal artistic evaluation rather than production.

§ 2 Immediate inspirations come from fictional character *Doctor Faustroll* created by french absurdist and 'father' of pataphysics Alfred Jarry (1996), the fantastic taxonomy of the *Celestial Emporium of Benevolent Knowledge* by magical realist Jorge Luis Borges (2000) and *A Hundred Thousand Billion Poems* by pataphysician and Oulipo co-founder Raymond Queneau (1961), amongst others.

§ 9 To address step (2) above, I explored the problem of objective evaluation and interpretation of subjective creativity specifically in regards to AMC. I have argued that the most appropriate way to approach this is by looking at five objective constraints (person, process, product, place, purpose) and seven subjective criteria (novelty, value, quality, purpose, spatial, temporal, ephemeral) holistically and by understanding that humour and art 'lie in the ear and eye of the beholder'.

§ 9.2.3
9.2
This resulted in an ***interpretation framework*** visualised as an evaluation matrix (5 constraints x 7 criteria) which can be used to qualitatively and/or quantitatively measure the creativity of a given AMC artefact:

§ 9.2.1 1. a set of scales that can be used to approximate a 'rating' for the creative value of an artefact,
§ 9.2.2
§ 9.2.3 2. a set of criteria to be considered using the scales above,
3. a combined framework for evaluation.

[1] Python, Hypertext Markup Language (HTML), CSS, Jinja, JavaScript
[2] Microsoft Translate, WordNet, Bing, Getty, Flickr and YouTube
[3] 2864 lines of code, 489 lines of comments - as of 08 Dec 2015

1.1 Motivation

Computers are binary machines; the world is black and white to them (0 and 1, on and off). Programmers can run abstract high-level commands which are executed in sequence (with fast speeds giving the illusion of multitasking). They are precise, structured, logical, and generally abide by strict standards. Computers can only be creative if they are given clear instructions as to how. Information Retrieval is generally focused on relevance of results in regards to the query. § 6

> The Analytical Engine has no pretensions whatever to ***originate*** anything. It can do ***whatever we know how to order it*** to perform.
>
> (Ada Lovelace, in Menabrea and Lovelace 1842, her emphasis)

Pataphysics emerged during the *Belle Époque*[4] in France and has either directly or indirectly influenced various artistic movements such as Dada, Symbolism, Surrealism, Oulipo and Absurdist Theatre. Pataphysics is highly subjective and particular, values exceptions, the imaginary and the mutually incompatible. § 4

Creativity is often studied at various levels (neurological, cognitive, and holistic/systemic), from different perspectives (subjective and objective) and characteristics (combinational, exploratory and transformative). It is usually defined in terms of value, originality and skill. § 5

Combining computing with pataphysics seems impossible—although the antinomies below (juxtaposing principles in computing on the left with ideas from pataphysics on the right) highlight just how intriguing a possible combination of the two would be.

- Polymorphism (generalisation) opposes particularity.
- Precision opposes exceptions and contradictions.
- Logic and structure oppose the imaginary and paradox.
- Cross-compatibility opposes the mutually exclusive.
- Responsiveness opposes the specific.
- Relevance opposes the creative.

This apparent dichotomy of computing and pataphysics is alluring. Christian Bök argued that pataphysics "sets the parameters for the contemporary relationship between science and poetry" (2002). Pataphysics suddenly seems like the perfect choice infusing computers (science) with creativity (poetry).

[4] 1871—1914

8.5 Combining pataphysics with creativity is easier. The ideas of combinatorial, exploratory and transformative creativity map quite nicely onto some pataphysical concepts such as clinamen, syzygy, antinomy and anomaly.

Another motivating factor for this project was the lack of research in the particular area of creative computing in general. The discipline of computational creativity has emerged fairly recently[5] from a background in Artificial Intelligence (AI). It appears to focus a lot more on the outcome of a product that would be judged creative rather than the actual process. Creative computing focuses on producing creative algorithms which may or may not have creative outputs. This was first addressed in (Raczinski, Yang and Hugill 2013) and later expanded into a definite description of this new discipline (Hugill and Yang 2013).

§ 5

§ E

My personal interest in this project comes from a background in computer science and a longstanding interest in art. Most recently I managed to successfully combine my technical skills with my creative side for a Master of Science degree in Creative Technologies at De Montfort University (DMU)[6].

1.2 Questions

Research dealing with subjective ideas and concepts like creativity throws up a lot of questions. My intention is to address them all throughout this thesis, although some of them will not have definite binary answers. An attempt to answer them can be found in the conclusion chapter 14.2.

§ 14.2

- What is the relationship between pataphysics and creativity?
- How is computer creativity related to AI?
- Should we distinguish between computationally automated or emulated creative processes and the programmer's input?
- How can a machine's creative output be evaluated?
- How can IR be infused with creativity?

1.3 Methodology

§ 3

This project combines research in science and art making it transdisciplinary.

[5] The first International Conferences on Computational Creativity ran in 2010 for example.

[6] A passive interactive installation, augmenting a live video stream of users with interactive elements using motion tracking algorithms. See msc.fania.eu (Raczinski 2010).

Pataphysics	Literature, Philosophy, Art, Poetry
Creativity	Cognitive Science, AI, Digital Humanities (DH)
Technology	IR, Natural Language Processing (NLP), Web Development

Epistemology	Transdisciplinary, subjective
Methodology	Creative computing, exploratory, experimental
Methods	Artefact, literature synthesis, algorithm design, theoretical framework, critical reflection and analysis, rapid incremental prototyping

The general process of my project was as follows.

1. Critically analyse and synthesise existing literature, ୨ II
2. develop pataphysical algorithms, ୨ IV
3. design a system to demonstrate algorithms, ୨ IV
4. develop a website as an artefact, ୨ IV
5. define an evaluation and interpretation framework, ୨ III
6. analyse results. ୨ V

1.4 Contributions

The key contributions to knowledge described in this thesis are:

- Three pataphysical search algorithms (clinamen, syzygy and antinomy).
- A creative exploratory search tool demonstrating the algorithms `pata.physics.wtf`.
- A set of 7 subjective criteria and 5 objective constraints for defining creativity.
- A combined framework for evaluating and interpreting creativity.

1.5 Publications

Some chapters (especially Foundations and Interpretation) in this thesis are based partially on articles published during this project. I have used fragments from those papers freely without specific citations unless clearly indicated. I had several co-authors (Hongji Yang, Andrew Hugill, James Sawle and Dave Everitt) for these pieces and I hereby acknowledge their contributions. § 8 & 9

A list of publications can be found in the preface on page vii. Details of talks and exhibitions and copies of the publications can be found in appendix E. § E

1.6 The Hitchhiker's Guide to this Thesis

This document is organised into 6 parts which form the main logical structure of the thesis and each part contains several chapters. There are margin notes pointing to relevant chapters, sections, tables, figures or images throughout.

1.6.1 Margin Notes

The different symbols used in margin notes are as follows.

- Represents a table.
- Represents a figure.
- Represents an image.
- </> Represents a snippet of source code.
- Σ Represents an equation.
- § Represents a chapter or section.
- Represents a thesis part.

1.6.2 Thesis Language

This thesis was written in LaTeX. It was first drafted in March 2015 and completed in December 2016. I created my own 'style' based on only a few restrictions imposed by DMU regulations (such as font size and page margins).

1.6.3 Chapter Overview

The preface contains the abstract, acknowledgments, and various tables of contents.

Introduction	Gives a general top-level overview of the research presented in this thesis.
Inspirations	Lists the various immediate inspirations for the project.
Methodology	Explains and justifies the approach taken for the research.
Pataphysics	Describes the origins of pataphysics and related concepts.
Creativity	Lists the theories of human and computer creativity.
Technology	Provides the technical background of this research.
Evaluation	Explains the models of evaluation for computer creativity.
Foundations	Brings together the research on creativity and pataphysics.
Interpretation	Critiques evaluation models and proposes a new approach.
Implementation	Describes `pata.physics.wtf` from a technical standpoint.
Applications	Showcases two use cases of this research.
Patanalysis	Analyses the artefact and some of the theoretical aspects.

Aspriations Addesses future work and known issues.
Outroduction Draws together and summarises the contributions of the work presented in this thesis.

The appendix contains additional material that was not suitable for including in the main body of the text. It also contains the list of references.

Inspirations

2

With bated breath and whisp'ring humbleness,
they did perform beyond thought's compass,
I speak my thought,
his throat that he hath breath'd in my dishonour here.

The very source of it is stopp'd,
it follows in his thought that I am he,
she deceives me past though,
he would kiss you twenty with a breath.

Here's my mother's breath up and down,
the breath no sooner left his father's body,
far be the thought of this from Henry's heart.

If her breath were as terrible as her terminations,
here's my mother's breath up and down,
thought is free.

This research was heavily influenced by a few major inspirations and this chapter introduces them all.

2.1 The Syzygy Surfer

This PhD project is directly based on the *Syzygy Surfer* (Hendler and Hugill 2011, 2013). Hendler and Hugill suggest the use of three pataphysical principles, namely clinamen, syzygy and anomaly, to create a new type of web search engine reminiscent of the experience of surfing the web using semantic web technologies. This is in contrast to current web search engines which value relevant results over creative ones.

'Surfing' used to be a creative interaction between a user and the web of information on the Internet, but the regular use of modern search engines has changed our expectations of this sort of knowledge acquisition. It has drifted away from a learning process by exploring the web to a straightforward process of Information Retrieval (IR) similar to looking up a word in a dictionary.

> The ambiguity of experience is the hallmark of creativity, that is captured in the essence of pataphysics. Traversing the representations of this ambiguity using algorithms inspired by the syzygy, clinamen and anomaly of pataphysics, using a panalogical mechanism applied to metadata, should be able to humanize and even poeticize the experience of searching the Web.
>
> (Hendler and Hugill 2013)

Their inspirations come from Borges (2000) (for the underlying poetic sense of unity), Jarry's pataphysical principles (1996) and Minsky and Singh's panalogies (parallel analogies—to introduce ambiguity, since it allows various descriptions of the same object) (2005).

๑ ๑ ๑

My project has since moved on from the idea of using the semantic web to create the search tool and uses the concept of antinomy rather than anomaly as one of its three algorithms. One of my original ideas based on the *Syzygy Surfer* was to create a standard ontology of creativity using semantic web technologies. I quickly ran into the following problem though: the idea of standards is totally opposed to that of surprise - which plays a role in creativity. Pataphysics in particular is fond of breaking standards (e.g. exceptions, contradictions, etc.). But standards are a key building block of the semantic web. A common ontology of creativity might be useful in some cases but nevertheless contradicts the use of pataphysics.

2.2 Faustroll's Library of Equivalent Books

§ 10 The artefact created to demonstrate the search algorithms—pata.physics.wtf—uses two collections of texts rather than the open web as source material. One of these corpora is based on the fictional library of 'equivalent books' from Alfred Jarry's *Exploits and Opinions of Dr. Faustroll, 'Pataphysician* 1996

10.1 The library also contains three prints (a poster of *Jane Avril* by Toulouse-Lautrec, an advert for the *Revue Blanche* by Bonnard, and a portrait of Doctor Faustroll by Aubrey Beardsley) and a picture *Saint Cado* by the Oberthuer printing house of Rennes (Jarry 1996)[1]. It contains the following books.

1. BAUDELAIRE, a volume of E.A. POE translations.
2. BERGERAC, *Works*, volume II, containing the *History of the States and Empires of the Sun*, and the *History of Birds*.
3. *The Gospel according to* SAINT LUKE, in Greek.
4. BLOY, *The Ungrateful Beggar*.
5. COLERIDGE, *The Rime of the ancient Mariner*.
6. DARIEN, *The Thief*.
7. DESBORDES-VALMORE, *The Oath of the Little Men*.
8. ELSKAMP, *Illuminated Designs*.
9. An odd volume of the *Plays* of FLORIAN.
10. An odd volume of *The Thousand and One Nights*, in the GALLAND translation.
11. GRABBE, *Scherz, Satire, Ironie und tiefere Bedeutung*, comedy in three acts.
12. KAHN, *The Tale of Gold and of Silence*.
13. LAUTREAMONT, *The Lays of Maldoror*.
14. MAETERLINCK, *Aglavaine and Selysette*.
15. MALLARME, *Verse and Prose*.
16. MENDES, *Gog*.

[1] These images are featured on the front page of pata.physics.wtf—see page 142

17. *The Odyssey*, Teubner's edition.
18. PELADAN, *Babylon*.
19. RABELAIS.
20. JEAN DE CHILRA, *The Sexual Hour*.
21. HENRI DE REGNIER, *The Jasper Cane*.
22. RIMBAUD, *The Illuminations*.
23. SCHWOB, *The Childrens' Crusade*.
24. Ubu Roi.
25. VERLAINE, *Wisdom*.
26. VERHAEREN, *The Hallucinated Landscapes*.
27. VERNE, *Voyage to the Center of the Earth*.

2.3 100.000.000.000.000 Poems

§ 10.4.1

The interface design of some of my search results is directly inspired by Raymond Queneau's *Cent Mille Milliards de Poèmes* (1961), a prime example of Oulipian art. The book is essentially made up of 10 pages containing one sonnet each. Each page however is split into 14 thin strips, one for each line. This means that mathematically there are 10^{14} possible poems to be read by combining different lines every time. My implementation of this resulted in a sonnet, each line of which can be changed individually using mouse clicks.

Figure 2.1 – Raymond Queneau's *Cent Mille Milliards de Poèmes*[2]

2.4 Celestial Emporium of Benevolent Knowledge

Jorge Luis Borges mentions a Chinese encyclopaedia called the 'Celestial Emporium of Benevolent Knowledge' in the short story *The Analytical Language of John Wilkins* (2000). It is a primary inspiration for this project, originally identi-

[2] Images of Queneau's book in the Gallimard 2006 edition by Martin Pyper (2010).

fied by (Hendler and Hugill 2011, 2013). It lists the following results under the category of 'animal'.

1. those that belong to the Emperor,
2. embalmed ones,
3. those that are trained,
4. suckling pigs,
5. mermaids,
6. fabulous ones,
7. stray dogs,
8. those included in the present classification,
9. those that tremble as if they were mad,
10. innumerable ones,
11. those drawn with a very fine camelhair brush,
12. others,
13. those that have just broken a flower vase,
14. those that from a long way off look like flies.

Although these are obviously all perfectly valid results, it is clear that they form a more creative, even poetic, view of what an animal might be than the Oxford English dictionary's prosaic: "a living organism which feeds on organic matter" (2010). This poetic form of order or structure was a direct inspiration for the results generated by this project's exploratory search tool `pata.physics.wtf`.

2.5 Metaphorical Search Engine Yossarian

Yossarian is a creative search engine which claims to return "diverse and unexpected results" (2015). Being a commercial product it is hard to find reliable details on precisely how their search engine works; the site seems well marketed but its functionality is shrouded in mystery.

> Yossarian makes the process of generating new ideas faster, while also improving its quality. This creative search engine helps people discover new perspectives, conceptual directions, creative insights, and allowing collaboration and feedback from a creative global community. (*Yossarian* 2015)

They also claim to be inspired by metaphors and that generating lateral connections can diversify users' ideas and help understand conceptual relationships between things through a, what they call, 'creative graph'.

The site started in a public alpha release in 2012. At the time, it consisted of simple image search. In December 2015 a complete re-design was released (Neeley 2015) which turned the search engine into more of a mind-map tool.

> Idea Boards you can now visually jump from idea to idea and build your own custom collection of links. It's a powerful new kind of mind map powered by search, and a radical departure from traditional search engine interfaces.
>
> (Neeley 2015)

While they do boldly call themselves "the world's first creative search engine" (***Yossarian*** 2015) it is impossible to know how their algorithms really work. The recently released mind map functionality brings up those 'lateral connections' in a relationship graph form. There is a slider that lets users adjust how creative they want their results to be—from literal to lateral.

Tony Veale introduced a model, superficially similar to the Yossarian system, that allows users to formulate queries as creative metaphors using a what he calls 'affective stereotype lexicon' (2013) but he does not go into the evaluation of his model. The idea is that the search engine is capabale of understanding metaphorical queries but not that it produces metaphoric results.

2.6 The Library of Babel

The *Library of Babel* is a short story by Jorge Luis Borges (1964). It envisions a universe, called 'the Library', which is composed of "an indefinite and perhaps infinite number of hexagonal galleries" containing every possible book every conceived and not yet conceived.

The specific artefact of inspiration for my project is a website implementing a miniature form of this library (`libraryofbabel.info`) created by Jonathan Basile (2015). Instead of containing every single book possible, it contains every single page possible—which is, at 3200 characters per page and 29 possible characters, still a lot.

Basile claims to use a 'pseudo-random number generating algorithm' (combining modular arithmetic and bit-shifting operations) to produce all 29^{3200} pages without needing to store anything on disk.

> The pages of rational text which this algorithm can locate are rarer than a single grain of sand in that collection, yet intrinsically no more meaningful. (. . .) One can find only text one has already written, and any attempt to find it in among other meaningful prose is certain to fail. The tantalizing promise of the universal library is the potential to discover what hasn't been written, or what once was written and now is lost. But there is still no way for us to find what we don't know how to look for. (. . .) Nonetheless, the library contains its own sort of poetry and revelation, and even this disappointment can provide a moment of clarity. (Basile 2015)

It is hard to say what exactly influenced my project most. I think the idea of computationally generating this massive library is fantastic—and absurd.

2.7 Oulipo

The Ouvroir de Littérature Potentielle (OULIPO) is a literary movement[3] from the 1960's, originating in France as a subcommittee of the "Collège de 'Pataphysique". As such it has roots in pataphysics although it eventually separated and became a standalone group. Their main philosophy perhaps is to use constraints in order to enhance creative output. Some examples of techniques, taken from (Mathews and Brotchie 2005), invented and used by them are shown below (and many more can be seen in chapter 4, tables 4.1 and 4.2).

4.1 & 4.2

N+7 Invented by Jean Lescure. It is sometimes called 'S+7'. It's a simple method of replacing each noun with the next seventh noun in a dictionary. For example: *tree* → *trend*, *shoreline* → *shotgun*[4].

Algol poetry
Algol (Algorithmic Oriented Language) is a programming language from 1960 which at the time consisted of only 24 words. It was used to write poetry given the restricted vocabulary of the language only (see example below in figure 2.2).

Melting snowball
A technique by which each line in a text has one less character than the preceding one resulting in a structure as shown in figure 2.2.

Paul Braffort
Paul Braffort wrote a program in 1975 to generate versions of Queneau's 100 thousand million poems. It used the reader's name and the time it took to write it to determine which poem to display. He did a similar thing with Italo Calvino to write a story that has a very large number of possible outcomes which can be reduced by the reader by making certain choices.

Mathew's algorithm
In the 1970's Harry Mathews created this procedure of generating results. It is based on permutation of characters, words, symbols, numbers, etc. See figure 2.2.

(The use of computers) became an instrument, not of combinatorial accumu-

[3] It has since spread to other disciplines. The generic term for Oulipian groups is OUXPO ("Ouvroir d'X Potentielle"), where the X can be replaced with whatever particular subject area you like (typically in french): fine art—OUPEINPO, music—OUMUPO, etc.

[4] Generated using the *Spoonbill N+7 Generator* (Christian 2016).

> lation, but of anti-combinatorial reduction. It served not to create combinations but to eliminate them. (Mathews and Brotchie 2005)

```
      Table
Begin:  to make format,
go down to comment
while channel not false
   (if not true).  End.
```

```
Incontrovertible
 sadomasochistic
  orthographical
   compositional
    restrictions
     insistently
      discipline
       grandiose
        sixteens
         initial
          hubris
           right
            down
             now
              to
               0
```

```
T  I  N  E
S  A  L  E
M  A  L  E
V  I  N  E

    ↓

T  I  N  E
E  S  A  L
L  E  M  A
I  N  E  V
```

Figure 2.2 – Algol poem (left), melting snowball (middle), Mathew's algorithm (right)

These techniques have endless applications in as many different disciplines. The use of constraints is now a well-known approach for creative activities and has many supporters.

2.8 Coder Culture

Whether you call it "programming culture", "coding culture", or "hacking culture", it is clear that the topics shared are ***code*** and ***culture***.

The programming language Python[5] was used for the core system behind the `pata.physics.wtf` site. The so-called *Zen of Python* is a set of guidelines for good practice in programming originally defined by Guido van Rossum—the creator of Python—who is endeeringly known as the Benevolent Dictator For Life (BDFL) and put into the below form by Tim Peters. This set of principles is also known as PEP20. The abstract reads: "Long time Pythoneer Tim Peters succinctly channels the BDFL's guiding principles for Python's design into 20 aphorisms, only 19 of which have been written down" (2004).

[5]The language was appropriately named after the British absurdist comedy group Monty Python (Python 2016). So by doing a syzygious jump it is the obvious choice of programming lanuage for this project: Pataphysics–Monty Python–Python.

Beautiful is better than ugly.
Explicit is better than implicit.
Simple is better than complex.
Complex is better than complicated.
Flat is better than nested.
Sparse is better than dense.
Readability counts.
Special cases aren't special enough to break the rules.
Although practicality beats purity.
Errors should never pass silently.
Unless explicitly silenced.
In the face of ambiguity, refuse the temptation to guess.
There should be one– and preferably only one –obvious way to do it.
Although that way may not be obvious at first unless you're Dutch.
Now is better than never.
Although never is often better than *right* now.
If the implementation is hard to explain, it's a bad idea.
If the implementation is easy to explain, it may be a good idea.
Namespaces are one honking great idea – let's do more of those!

(Peters 2004)

I cannot claim to have followed each and every one of those recommendations in my coding practice (although I have certainly tried) but it has been highly influential during the writing and design of this thesis.

୭ ୭ ୭

The following list shows some other general programming culture references that have been inspirational in one way or another. They were interesting to me due to their underlying sense of humour which resembles that of pataphysics.

Jargon File
: a "comprehensive compendium of hacker slang illuminating many aspects of hackish tradition, folklore, and humor" (Raymond 2004)

1337
: an Internet 'language' (Thrid 2002)

Code Golf
: "a competition to solve a particular problem in the fewest bytes of source code" (StackExchange n.d.)

Code Bowling
: "a competition to solve a particular (usually simple) problem in the most bytes or complexity" (StackExchange n.d.)

IOCCC
: a competition to "write the most obscure/obfuscated C program within

the rules to show the importance of programming style, in an ironic way" (Broukhis, Cooper and Noll n.d.)

Glitch Art

"Glitch art takes temporary pixelations, interruptions and glitches and turns them into visually arresting pieces, questioning the forms and traditions of art using digital techniques" (Wong 2013) (see also (Google 2016; Reddit n.d.))

Easter Eggs

The practice of hiding a reproducible, personal, harmless and entertaining feature into a piece of software (D. Wolf and A. Wolf n.d.)

Knuth

Donald Knuth has long maintained a tradition of (a) adding easter eggs to his books on programming and (b) rewarding people for finding errors and typos in his books with fictional currency (Knuth n.d.).

An example of creative code from the International Obfuscated C Code Contest (IOCCC) is reproduced in source 2.1. It shows highly obfuscated C code "written in hommage to René Magritte's picture *La trahison des images* (The Treachery of Images)" by Uri Goren in 2011. It won the 'most artistic' category of that year's contest (Goren 2011).

```
typedef unsigned char t;t*F="%c",l[]="|\\/=_ \n](.\0(),*(.(={}*.)[[*.",N='\n',*
r;typedef(*H)();extern H Ar;Q(a){return(a|-a)>>31;}H S(c,a){return(H)(a&~c|(int
)Ar&c);}extern t*ist;V(t*u){*u^=*u&2^(*u>>7)*185;}Z(t*u,t n){*u-=n;}e(t c,H h){
R(h,Q(*                                                                   r^c));}
I(){r=l                                                                   +7-4*Q(
getchar                                                                   ()^*l);
}R(H h,                 int                                               c){Ar=S
(c,h);-                 main()                                            ;}P(){r
++;}z()                 {                                                 O(&N);}
O(t*c){                     printf(                                       F,+*c);
}T(){r=                         "This is not a function\n"                ;}w(U){
U=Z(r,8                     );                                            r-=~Q(*
r/8-4);                     return 0;                                     }M(){r=
ist-68;                 }                                                 h(){t G
=r[1]-r                                                                   [2]^*r;
G^=30;V                                                                   (&G);e(
0,(O(&G                                                                   ),P(P(*
r++)),z));}g(){M();R(h,0);}f(){P(O(r));e('f',g);}p(){P();e('a',f);}d(){P(O(r));
e('n',p);}c(u){u=r[-2];T(Ar=d);R(f,Q(u^'"'));}n(){e(w(O(l+*r%8)),c);}a(){I();R(
n,0);}main(){S(Q(Ar),a)();}H               Ar;t*ist="Rene Magritte"-(1898-1967);
```

Code 2.1 – An example entry by Uri Goren from the IOCCC contest from 2011.

METHODOLOGY

3

Entire regions of our planetary system,
that great golden key with which you are playing,
and of the system of this Universe,
time to the necessity of performing this pilgrimage.

Would arrive at the correct solution,
face shews not the least wrinkle,
through his rash opinion of the improbability of performing,
faire ici le compte rendu technique de ma decouverte.

Acting upon this hint,
acted violently on my nervous system,
this was caused by intense heat acting on the organic matter of the earth.

The sum total of good playing,
and the Machine playing its large Wings,
that I would try it on myself acting forthwith on this decision.

๑ ๑ ๑

This project combines research in science, art and the humanities—making it transdisciplinary.

Pataphysics	Literature, Philosophy, Art
Creativity	Cognitive Science, AI, DH
Technology	IR, NLP, Web Development

Traditional methodologies in these disciplines are very subject specific and a project combining elements of each field is left mixing and matching suitable methods from them all.

In this chapter I will outline the reasons why the existing intradisciplinary methodologies aren't completely suitable for this project and then explain the choice of more transdisciplinary methods and how I combined them to suit my needs.

As mentioned in the Introduction the overall objectives of this project are to: § 1.3

1. Critically analyse and synthesise existing literature, ๑ II
2. develop pataphysical algorithms, ๑ IV
3. design a system to demonstrate algorithms, ๑ IV
4. develop a website as an artefact, ๑ IV
5. define an evaluation and interpretation framework, ๑ III
6. analyse results. ๑ V

Research methods that support these tasks are needed and I will address these four points again at the end of this chapter. § 3.3

3.1 Intradisciplinary

Different disciplines prefer different research methodologies. Of the various disciplines that inform this research the specific subareas that are relevant are as follows.

- Information Retrieval
- Interface Design
- Web Development
- Poetry, Literature, and Art
- Philosophy
- Human and Machine Creativity
- Creative Computing
- Computational Creativity

3.1.1 Technology

Half of this project's objectives are related to computer science therefore it is important to consider how research in this discipline is traditionally approached.

A framework for finding a suitable approach was suggested by Holz et al (2006). The following four steps form an iterative process. (1) "What do we want to achieve?" e.g. find out what is happening, develop something that works, evaluate an existing system/technology, compare existing systems, or change human behaviour. (2) "Where does the data come from?" e.g. how to collect? (read, observe, ask, measure, experiment, model) and where to collect? (field, laboratory, conceptual). (3) "What do we do with the data?", e.g. identify themes/patterns/quotes, calculate numbers, identify trends, express via multimedia, create frameworks/taxonomies. (4) "Have we achieved our goal?" e.g. draw conclusions, evaluate results, or identify limitations.

Another option is to look at what computer science researchers have done historically. In a rather old but still insightful analysis of over 600 papers[1] Ramesh et al (2004) have shown that—by far—the most common approach to research in computer science during this period was ***formulative*** with almost 79% use (as opposed to "descriptive" with 10% and "evaluative" with 11%). This was in particular in regards to "processes, methods and algorithms" which was used by just over 50% of researchers. Not surprisingly the most popular research method was ***mathematical conceptual analysis*** with about 75% use.

Jose Nelson Amaral (2006) classifies methodologies in computer science into five main categories as shown below.

Formal	Proof, verification, correctness
Experimental	Testing, evaluation, question answering

[1]While the paper itself was published in 2004, the body of work was based on publications from between 1995 and 1999—this suggests that a lot of the more "recent" research around web technologies is not included in this study.

Build	Proof of concept, prototype, artefact
Process	Understand and define processes
Model	Abstraction, simulations

୭ ୭ ୭

Here are this project's answers to the four questions posed by Holz et al (2006).

What do we want to achieve?

- Understand human creativity and how this translates to machines.
- Understand the relationship of pataphysics and creativity.
- Understand how creativity is evaluated in humans and machines.
- Research suitable pataphysical concepts to be implemented as algorithms.
- Define algorithms formally.
- Implement prototype incorporating algorithms.
- Develop framework for interpreting and evaluating machine creativity.

Where does the data come from?

- Read pataphysical literature and research.
- Collate existing research on creativity and evaluation.
- Survey creative approaches to technology.
- Experiment with algorithms and implementation.

What do we do with the data?

- Iterate through developmental stages of algorithmic outputs.
- Create an artefact that represents the underlying philosophy and research.
- Create an evaluation framework based on theoretical research.

Have we achieved our goal?

- See conclusion chapter 14. § 14

Referring back to the four objectives above (see page 22), objective 1 is to create new creative search algorithms. This is not supposed to happen on a purely abstract basis but in a practical fashion (i.e. 'experimental'), with a working implementation (i.e. 'build') as proof-of-concept (see objective 2). While the algorithms need to be defined in formal terms (i.e. 'formal'), the goal here is not to create a theoretical proof of correctness (given the creative and rather subjective nature of the underlying philosophy this is virtually impossible) but a practical demonstration of the creative processes behind. Overall this would suggest an experimental approach with prototyping of an artefact. Objective 3 is to come up with a suitable definition of creativity (i.e. 'process'). This should be informed by existing research. Again, we are not interested in formulating this in mathematical terms and proofs but rather a more esoteric and systemic view. Because the definition needs to apply to humans and machines it needs to be precise enough. Objective 4 is then to create an overall theoretical framework (i.e. 'model') for the evaluation of creativity in humans and machines.

By now we have managed to cover every one of the major methodologies mentioned by Amaral et al. (2006) but we are still lacking ways to address the subjective and creative nature of the project. Furthermore, the philosophical and artistic inspirations that inform the development of the artefact don't get enough of a voice in these methods. In computer science, implementations are generally seen as a proof of concepts or prototypes—when really they should be seen as artefacts in the sense of artistic pieces of work. So, to really appreciate the scope of this practical element of this project we need to consider research in the arts and humanities too.

3.1.2 Arts and Humanities

> A hallmark of humanistic study is that research is approached differently than in the natural and social sciences, where data and hard evidence are required to draw conclusions. Because the human experience cannot be adequately captured by facts and figures alone, humanities research employs methods that are historical, interpretive and analytical in nature. (Stanford n.d.)

Malins and Gray suggest the following ideas for arts-based researchers searching for the right methodology (1995).

- Consider a range of research strategies (from all disciplines).
- 'Tailor' the research to the nature of project and the researcher's expertise.
- Carry out the research from an informed perspective, as 'participant observer'.
- Continually define and refine the research question, allowing methodologies to emerge.
- Acknowledge accessibility, discipline, rigour, transparency, and transferability.
- Be aware of the critical context of practice and research and raise the level of critical debate.
- Consider interdisciplinary / multidisciplinary approaches to research.

They further elaborate on the key characteristics of arts methodologies as follows (Gray and Malins 2004).

- Experiencing/exploring, gathering, documenting information and generating data/evidence.
- Reflecting on and evaluating information, selecting the most relevant information.
- Analysing, interpreting and making sense of information.
- Synthesizing and communicating research findings, planning new research.

(Gray and Malins 2004)

They further specify a whole set of individual methods used for the approaches above.

- observation and related notation/use of symbols
- visualization
- drawing (in all forms)
- diagrams
- concept mapping, mind mapping
- brainstorming/lateral thinking
- sketchbook/notebook
- photography, video, audio
- 3D models/maquettes
- experimentation with materials and processes
- modelling/simulations
- multimedia/hypermedia applications
- digital databases, visual and textual glossaries and archives
- reflection-in-action/'stream of consciousness'/personal narrative
- visual diary/reflective journal/research diary
- collaboration/participation/feedback, for example workshops
- use of metaphor and analogy
- organizational and analytical matrices
- decision-making flow charts
- story boards, visual narratives
- curation
- critical writing, publications
- exposition and peer feedback/review

(Gray and Malins 2004)

The discpiline of Digital Humanities (DH) (see chapter 5.3.4) seems like a logical choice to look for suitable methodologies. It is characterised by "collaboration, transdisciplinarity and an engagement with computing" (Burdick et al. 2012) but it should not simply be reduced to "doing the humanities digitally" (2012). Transliteracy, an understanding of several kinds of tools and media, is an important aspect in this (Thomas et al. 2007). DH can be broken down into the following set of methodologies.

§ 5.3.4

Design

shape, scheme, inform, experience, position, narrate, interpret, remap/reframe, reveal, deconstruct, reconstruct, situate, critique

Curation, analysis, editing, modelling
: digitise, classify, describe, metadata, organise, navigate

Computation, processing
: disambiguate, encode, structure, procedure, index, automate, sort, search, calculate, match

Networks, infrastructure
: cultural, institutional, technical, compatible, interoperable, flexible, mutable, extensible

Versioning, prototyping, failures
: iterate, experiment, take-risks, redefine, beta-test

Some of the emerging research methods Burdick et al. have identified are listed
§ A.3 below (2012) (The full list can be found in appendix A.3).

- structured mark-up
- natural language processing
- mutability
- digital cultural record
- algorithmic analysis
- distant/close, macro/micro, surface/depth
- parametrics
- cultural mash-ups
- algorithm design
- data visualization
- modelling knowledge
- ambient data
- collaborative authorship
- interdisciplinary teams
- use as performance
- narrative structures
- code as text
- software in a cultural context
- repurposable content and remix culture
- participatory web
- read/write/rewrite
- meta-medium
- polymorphous browsing

૭ ૭ ૭

Several of the methodologies listed by Gray and Malins (2004) seem to apply to the research presented in this thesis. Exploring, evaluating, analysing, interpreting, synthesising and disseminating research all are part of it. However, looking at the specific methods they collated, the difference becomes clearer as only the following 7 appear relevant (visualization, experimentation with processes, multimedia/hypermedia applications, use of metaphor and analogy, organizational and analytical matrices, curation, and critical writing, publications).

The DH methodologies seem more useful. In terms of **design**, `pata.physics.wtf` *positions* itself in context and the evaluation framework *interprets* and *cri-*

tiques AMC. Before that I **curate** the two corpora, *digitise* them and *organise* them. **Computing** comes in at verious stages, to *(dis)ambiguate* (i.e. pataphysicalise), *encode*, *index*, *search* and *match* data. The **infrastructure** is *cultural*, *technical* and *extensible*, relying on the World Wide Web (WWW) for several spects. **Versioning, prototyping and failures** all come in during the *iterative* development process, which involves a lot of *experimentation* and refactoring. Furthermore, the research methods Burdick et al (2012) list match this project much better (although of course the list above was already only a selection that was deemed relevant; the original list was much larger. See appendix A.3). § A.3

3.2 Transdisciplinary

Nicolescu distinguished between 3 different kinds of research "without stable boundaries between the disciplines".[2] (2010).

Multidisciplinarity
: concerns itself with studying a research topic in not just one discipline but in several simultaneously.

Interdisciplinarity
: concerns the transfer of methods from one discipline to another.

Transdisciplinarity
: concerns that which is at once between the disciplines, across the different disciplines, and beyond all disciplines.

The standard epistemological view of science and art is that they are objective and subjective, respectively. So, what does that mean for research conducted between, across and beyond science and art, i.e. research that is transdisciplinary?

Nicolescu criticised the view that science must be objective. He even claimed that any non-scientific knowledge is "cast into the inferno of subjectivity, tolerated at most as a meaningless embellishment or rejected with contempt as a fantasy, an illusion, a regression, or a product of the imagination" (2010). Objectivity, he said, becomes the "supreme criterion of Truth"[3]

> The death of the Subject is the price we pay for objective knowledge.
>
> (Nicolescu 2010)

[2] Nicolescu cites Jean Piaget here, who first coined the term 'transdisciplinarity' in 1972.

[3] As we shall see later, pataphysics does the opposite: it reveres the Subject.

He went on to quote Werner Heisenberg on the concepts of objective and subjective reality: "we would make a very crude simplification if we want to divide the world in[to] one objective reality and one subjective reality. Many rigidities of the philosophy of the last centuries are born by this black and white view of the world" (Heisenberg, cited in Nicolescu 2010).

> The too strong insistence on the difference between scientific knowledge and artistic knowledge comes from the wrong idea that concepts describe perfectly the 'real things'. (. . .) All true philosophy is situated on the threshold between science and poetry. (Heisenberg, cited in Nicolescu 2010)
> [4]

In transdisciplinarity traditional disciplinary boundaries have no meaning.

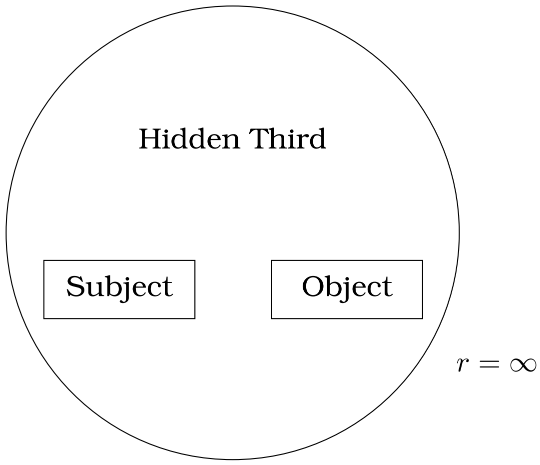


Figure 3.1 – Nicolescu's transdisciplinarity

Working across disciplines requires a new unique methodology. Nicolescu proposed a methodology of transdisciplinarity as a non-hierarchical ternary partition of 'Subject, Object and Hidden Third' (as shown in figure 3.1) rather than the traditional binary partition of 'Subject versus Object' (2010). 3.1

> The old principle "unity in diversity and diversity from unity" is embodied in transdisciplinarity.' (Nicolescu 2010)

> 'unite and conquer' $\longleftrightarrow$ 'divide and conquer' (Yang 2013)

Hugill and Yang agree that existing research methodologies are unsuitable for transdisciplinary subjects such as Creative Computing (CC). The following is an

[4] The full paragraph is worth quoting—see appendix A.2.

example of a possible CC research methodology they propose as a starting point (Hugill and Yang 2013):

1. Review literature across disciplines.
2. Identify key creative activities.
3. Analyse the processes of creation.
4. Propose approaches to support these activities and processes.
5. Design and implement software following this approach.
6. Experiment with the resulting system and propose framework.

They go on to propose four standards for CC (Hugill and Yang 2013) namely, (1) resist standardisation, (2) perpetual novelty, (3) continuous user interaction and (4) combinational, exploratory and or transformational.

A different model was suggested by Edmonds and Candy in their Trajectory Model of Practice and Research (TMPR), a framework to "influence practice, inform theory and, in particular, shape evaluation" (2010). Figure 3.2 shows the TMPR which allows for different trajectories between practice, theory and evaluation. Table 3.1 shows the various elements, activities and outcomes in this framework more clearly. 3.2 3.1

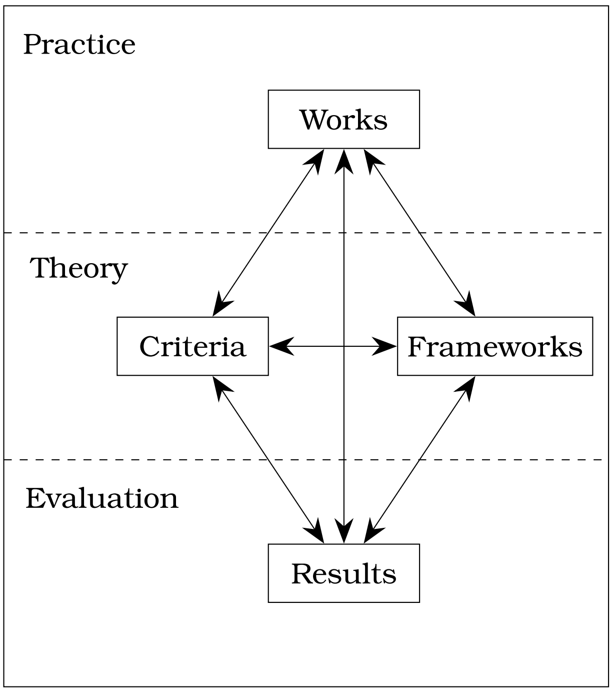


Figure 3.2 – Edmonds and Candy's trajectory model (TMPR)

Table 3.1 – Elements, activities and outcomes of each trajectory in the TMPR

Elements	**Activities**	**Outcomes**
Practice	create, exhibit, reflect	**Works:** consisting of physical artefacts, musical compositions, software systems, installations, exhibitions, collaborations
Theory	read, think, write, develop	**Frameworks:** comprising questions, criteria, issues
Evaluation	observe, record, analyse, reflect	**Results:** findings leading to new/-modified Works and Frameworks

This project positions itself "at once between the disciplines, across the different disciplines, and beyond all disciplines"—making it transdisciplinary. The abolishment of disciplinary boundaries suits the unique context of this research. Pataphysics specifically is highly subjective. Searle highlighted that ontologically subjective topics (such as creativity) can be studied in epistemically objective ways (2015), which, as doctoral research, this project attempts to do.

The Hugill and Yang CC methodology seems general enough to fit the needs of this project, with all 6 points covered in the various chapters of this thesis.

1. Review literature across disciplines (chapters 2, 4, 5, 6, and 7).
2. Define creativity in humans and machines (chapters 4, 5, 6 and 7).
3. Analyse the relation between the disciplines above (chapter 8).
4. Propose algorithms to support creativity in machines (chapter 10).
5. Design and implement software following this approach (chapter 10).
6. Experiment with the resulting system and propose interpretation/evaluation framework (chapters 12, 13, and 9).

3.3

Figure 3.3 on page 32 shows how the TMPR could be applied to this project.

3.3 Patadisciplinary

So, to summarise, this project draws from several different disciplines as mentioned at the beginning of this chapter (page 21): pataphysics—literture, philosophy, art, creativity—cognitive science, AI, DH, and technology—IR, NLP, web development.

Epistemology	Transdisciplinary, subjective
Methodology	Creative computing, exploratory, experimental
Methods	Artefact, literature synthesis, algorithm design, theoretical framework, critical reflection and analysis, rapid incremental prototyping

The general workflow of this project was as follows: (1) critically analyse and synthesise existing literature, (2) develop pataphysical algorithms, (3) design a system to demonstrate algorithms, (4) develop a website as an artefact, (5) define an evaluation and interpretation framework, and (6) analyse results.

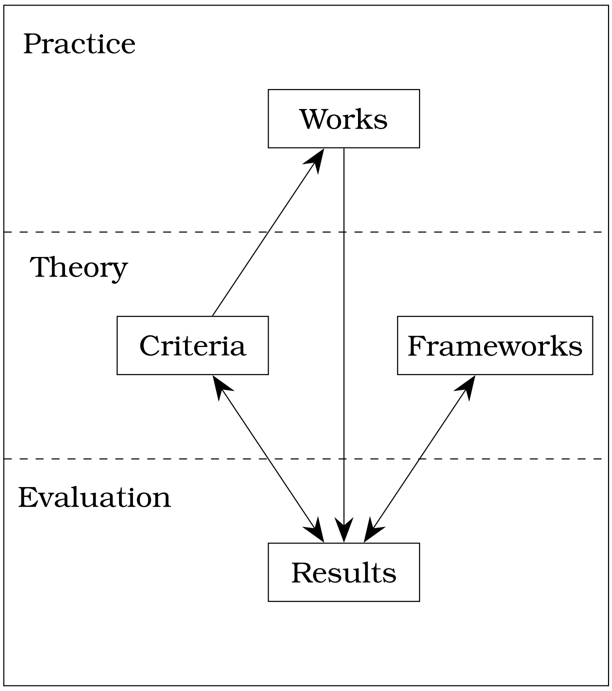


Figure 3.3 – This project's trajectory model

As figure 3.3 shows, the practive trajectory of this research is based on the practical development of a website to contain the exploratory search tool and implementation of the theoreical algorithms. The theory trajectory is about defining those algorithms formally in historic and topical context based on a critical survey of related literature. This also includes the development of a theoretical framework for the evaluation and interpretation of creative artefacts. The Evaluation trajectory then is all about the results. That includes an analysis of the work completed. The arrows in the figure indicate how these different trajectories influence each other. 3.3

Part II

TΘΘLS OF THE TR∀DΣ

Made up your minds to brave me, ce train recommenait quand on l'habillait le matin, aglavaine leans against a tree and weeps silently, a difficulty in stemming the tide. Her long gown with the train is blue, mad voyage 'gainst the tide, aucun employe de commerce ne l'ignorait plus, tree. Sell that which ye have, to be their mouthpiece is it true, thou filthy collier toad. Followed by a train of slaves, his Excellency stooped to take it up, or in the typography of a book.

’PATAPHYSICS

4

And stranger'd with our oath,
the almanac of my true date,
you'll pay me the eight shillings I won of you,
gape open wide and eat him quick.

Pardon what is past,
nor loose nor tied in formal plat,
the noble Brutus to our party,
sure I lack thee may pass for a wise man.

Or to take note how many pair of silk stockings thou hast,
who with his fear is put beside his part,
an oath of mickle might.

For the ways are dangerous to pass,
Gloucester offers to put up a bill,
on the Alps it is reported thou didst eat strange flesh.

> To understand 'pataphysics is to fail to understand 'pataphysics. (Hugill 2012)

It is probably impossible to define pataphysics in one sentence. There is no definition that does justice to what pataphysics really is and no single definition is truer than any other. In fact, the college of pataphysics in France has published a book (Brotchie, Chapman et al. 2003) with over 100 definitions that they all call 'equally valid'. This chapter therefore begins with several selected definitions to introduce the topic.

> Pataphysics . . . is the science of that which is superinduced upon metaphysics, whether within or beyond the latter's limitations, extending as far beyond metaphysics as the latter extends beyond physics. (. . .) Pataphysics will be, above all, the science of the particular, despite the common opinion that the only science is that of the general. Pataphysics will examine the laws governing exceptions, and will explain the universe supplementary to this one. (. . .) DEFINITION: Pataphysics is the science of imaginary solutions, which symbolically attributes the properties of objects, described by their virtuality, to their lineaments. (Jarry 1996)

> 'Pataphysics is patient; 'Pataphysics is benign; 'Pataphysics envies nothing, is never distracted, never puffed up, it has neither aspirations nor seeks not its own, it is even-tempered, and thinks not evil; it mocks not iniquity: it is enraptured with scientific truth; it supports everything, believes everything, has faith in everything and upholds everything that is. (Brotchie, Chapman et al. 2003)

> 'Pataphysics passes easily from one state of apparent definition to another. Thus it can present itself under the aspect of a gas, a liquid or a solid.
>
> (Brotchie, Chapman et al. 2003)

> 'Pataphysics, "the science of the particular", does not, therefore, study the rules governing the general recurrence of a periodic incident (the expected case) so much as study the games governing the special occurrence of a sporadic accident (the excepted case). (...) Jarry performs humorously on behalf of literature what Nietzsche performs seriously on behalf of philosophy. Both thinkers in effect attempt to dream up a "gay science", whose joie de vivre thrives wherever the tyranny of truth has increased our esteem for the lie and wherever the tyranny of reason has increased our esteem for the mad.
>
> (Bök 2002)

La pataphysique est la fin des fins.	'Pataphysics is the end of ends.
La pataphysique est la fin des faims.	'Pataphysics is the end of hunger.
La pataphysique est la faim des fins.	'Pataphysics is the hunger for ends.
La pataphysique est le fin du fin.	'Pataphysics is the finest of the fine.

(Brotchie, Chapman et al. 2003)

> The branch of philosophy that deals with an imaginary realm additional to metaphysics. (OED 2016)

4.1 Conscious

> Jarry was "attempting to transcend his own existence." (Hugill 2012)

> It is certainly true that making life "as beautiful as literature" was one of (Jarry's) goals. (Hugill 2012)

Studying Jarry's life gives certain insights into the man who created pataphysics and why he might have done so. Several works have helped prepare the below outline of Jarry's life. Alastair Brotchie's *A Pataphysical Life* (2011) and Roger Shattuck's *The Banquet Years* (1959) were the two main sources used but several others have also written about Alfred Jarry (e.g. Linda Klieger Stillman, Keith Beaumont, and Jill Fell).

4.1.1 Life

Alfred Jarry was born in Laval, Mayenne, France in 1873 and died in Paris in 1907, at the age of 34. He was known as a poet, dramatist, novelist and journalist but also as a graphic artist. His hobbies included entomology, fishing, cycling, fencing, shooting and drinking.

He went to school in Rennes, where his physics teachers Félix-Frédéric Hébert left such a big impression on Jarry that he would later be his inspiration for Père Ubu. He passed his baccalauréat with 17 and moved to Paris to attend the lycée Henri IV in preparation to apply for admission to the École Normale Supérieure but eventually gave upon the entrance exam after several unsuccessful attempts. He met another teacher at the lycée, this time a philosophy teacher called Henri Bergson, who inspired him greatly. He published his first collection of poems in 1893, aged 20, the year his mother died. One of his classmates there described him as follows.

> (...) I found Jarry's mental processes disturbing. When he let himself go he seemed in thrall to a torrent of words outside his control. It was no longer a person speaking, but a machine controlled by a demon. His staccato voice, metallic and nasal, his abrupt puppet-like gestures, his fixed expression and uncontrolled flood of language, his grotesque and brilliant turns of phrases, ended up provoking a feeling of disquiet. He was informed, intelligent, and discriminating; he was good person, secretly kind, perhaps even shy beneath it all (...) but his originality resembled nothing short of a mental anomaly.
>
> (Gandilhon Gens-d'Armes 1922, as cited in Brotchie 2011)

He was at the centre of the avant-garde movement in Paris around that time, at the centre of the Tuesday meetings of the Mercure de France (a literary magazine run by Alfred Valette and his wife Rachilde, who soon became a sort of substitute family to Jarry who was roughly 15 years younger than them). Being rather misogynist at times and homosexually inclined, Rachilde was one of his very few female friends.

The following year, 1895, he briefly joined the army in the 101^{st} infantry, after having dodged it by being an enrolled student at the lycée. He followed rules there pedantically but hated the loss of his individualism. According to Brotchie, he "chose subservience, but subservience taken to the point of parody: the pataphysical solution to the problem of obedience" (2011). Probably the only thing he enjoyed there was the fencing and shooting training. He looked funny in the uniform that was too big for him being so small (5'3") so he was eventually excused from parades and after a few months he was allowed to leave to Paris frequently. He was discharged in December 1895 on medical grounds: gallstones. It is not unlikely that he faked the illness by drinking picric acid.

His father had died just two months earlier and had left him a small inheritance, which he spent mostly on publishing his very own magazine dedicated to symbolist wood carvings, the *Perhinderion*. He had previously co-edited the magazine *L'Ymagier* with Remy de Gourmont between 1893 and 1894. He joined Aurélien Lugné-Poë as his secretary (his only ever real job) at the *Théâtre de l'Œuvre* after his discharge at the army, where he would pour his utmost attention to putting his Ubu play on the stage. He also played a small role in the production of *Peer Gynt* at the *Œuvre* earlier in 1896. The printed version of *Ubu Roi* appeared in *Le Livre d'Art* in the middle of the year with Jarry's carved woodcut image 4.1 of Ubu shown in image 4.1. The première took place on 10^{th} December that year and caused an outrage in the audience after the first word: 'merdre' (sometimes translated as 'pshit'). Jarry had previously arranged for certain friends to counter any reaction of the general audience and to prevent under all circumstances for the play to reach its conclusion. The performance went according to plan. The uproar after the first word was uttered was immense, the performance had to be interrupted at times to calm the audience and it finished in shouts of praise, protest and insults. There were no further performances but the event was considered historic even at the time and is now widely seen as the first 'modern' play (Brotchie 2011). And as Dave Walsh puts it: "Movements such as Dadaism, Surrealism, Futurism, Expressionism Cubism, Theatre of the Absurd—all owe debts to [Jarry's] works" (2001).

Although Ubu's mannerism of speech was originally imitating Jarry's, as suggested by Lugné-Poë (Brotchie 2011), Jarry continued to adopt Ubu's mannerisms.

> Those who knew him said that his nauseating appearance hid a youth who was stubborn yet shy, proud and little full of himself, but good-natured and ingenuous behind his cynicism, one who was fiercely independent and rigorously honest. (Henri de Régnier, as cited in Brotchie 2011)

> Alfred Jarry had a very particular way of speaking to that was disconcerting to those who heard it for the first time. He said "we", when referring to himself, and substituted verbs for nouns, in imitation of ancient Greek. Example: "celui qui soufflé" (that which blows) for the wind, and "celui qui se traîne" (that which crawls along) for the train, even if it was an express! This made conversation somewhat complicated, not least because of the rapidity of his delivery.
> (Rachilde, as cited in Brotchie 2011)

> Alfred Jarry was a man of letters to an unprecedented extent. His smallest actions, his childish pranks, everything he did was literature. His whole life was shaped by literature, and only by literature. (Appolinaire, as cited in Brotchie 2011)

Jarry spent the next few years writing. He had spent all his inheritance on the publication of his magazine and the production of Ubu Roi. It is during this

Figure 4.1 – Woodcut print of Ubu by Alfred Jarry

time that he moved to his infamous tiny flat on the second-and-a-half floor. Jarry could just about stand upright but any guests had to crouch. He had no electricity or gas and no means of cooking. In December 1897 he formed a marionette theatre with his friend Claude Terasse: the *Théâtre de Pantins* and they performed *Ubu Roi* in January 1898 without riots in the audience.

Jarry then gradually withdrew from the literary circles in Paris and spent more time in a little shack on the banks of the Seine near the village of Le Coudray. He started writing a regular review column for the *Revue Blanche* in 1900, the income of which he certainly needed much. There was a brief revival of the Ubu marionette play in the *Cabaret des Quat'z'Arts* in 1901.

Around 1904 he began drinking ether, the absinthe not strong enough anymore. In the winter of 1905 he was very ill, the cold and poverty not helping. In 1906, his friends became more and more concerned about his deteriorating health and eventually Valette and Saltas sent him to his sister Charlotte. He then spent some time in Paris and some in Laval at his sister's place over the next year. Alfred Jarry then died in November of 1907 of meningeal tuberculosis. His last request was for a toothpick.

> He believes that the decomposing brain goes on working after death and it is its dreams that are Paradise.
> (Jarry 1906, as cited in Brotchie and Chapman 2007, the 'he' refers to Jarry himself, he is talking in third person.)

4.1.2 Literature

Jarry has written a good amount of texts in his short life and he didn't confine himself to a single category either. He wrote poems, novels, short stories, essays, art reviews, theatre reviews and plays and also produced translations into French. Many of his texts were completely fictional, some had autobiographical and some scientific aspects and most of them had a sarcastic sense of humour.
§ A.6 See appendix A.6 for a full list.

> Jarry was an acknowledged classical scholar, had already worked as a reviewer of art and drama, had edited two art magazines, was up to date with modern scientific theory, especially physics, read widely in mathematics and psychology, and had an extensive basic knowledge of philosophy.
> (Brotchie 2011)

James Cutshall says that "instead of Jarry the man and the meaning of his literary endeavours becoming clearer with the passage of time, both have become increasingly indistinct" (1988). He intended to show the seriousness implied behind the humour used in many of Jarry's novels, in order to give the author the merit he deserved. Cutshall wrote about Jarry's novels rather than simply seeing him as the playwright of the Ubu plays. He surveyed existing criticism about Jarry's texts and provided his own view on them. He immortalised Jarry by saying "whether or not this is the sort of 'éthernité' sought by the heroes of Jarry's novels, it is certainly that which their author somewhat belatedly has found" (Cutshall 1988).

Cutshall was not the only one who has written about certain less-known texts by Jarry. Marieke Dubbelboer's thesis *Ubusing Culture* is also interesting in this regard since it concentrates completely on the *Almanachs du Père Ubu* (published in 1898 and 1901) (2009). She was looking for keys to Jarry's poetics in those texts, which she says "seemed to defy labelling or literary norms" (2009). She claims the *Almanachs* to be quite radical and exemplary of his innovative poetics moving away from symbolism and towards the avant-garde. In general she says his work “can be characterized as playful, elusive, paradoxical and provocative” (2009) and his two *Almanachs* are the essence of his non-conformist attitude. They were written at a time of change for Jarry, when he withdrew from his usual circles in Paris and he published in new magazines.

4.2 Self-conscious

We will need to understand the essence of pataphysics to understand how it relates to the other topics of this research.

Jarry first defined pataphysics in his book *Exploits and Opinions of Dr Faustroll, Pataphysician* written in 1898 and published posthumously in 1911 (1996). But the concept appeared as early as in 1893 in his prose text *Guignol* that won him a prize in the newspaper *L'Echo de Paris* and it appears in many of his writings. He originally intended to write a whole book called *Elements of Pataphysics* but only part of this appeared in Faustroll.

Zoë Corbyn gives a very simple short introduction for beginners of the topic in an article in the *Guardian* (2005). She describes it like this:

> Correct definitions are equivalent to wrong ones; all religions are on a par as imaginary and equally important; chalk really is cheese. It's an escape from reality — reminding us of just how idiotic the rules that dog our everyday existence are. (Corbyn 2005)

Jean Baudrillard has a few other definitions for pataphysics (2007). According to him, pataphysics is “the highest temptation of the spirit”, “the nail in the tire”, “the philosophy of the gaseous state”, “the science or the unique imaginary solution to the absence of problems”, to name just a few.

Another rather strange interpretation of pataphysics is Asger Jorn's. He calls pataphysics “a religion in the making” (1961). He claims that since “natural religion is the spiritual confirmation of material existence”, “metaphysical religion represents the establishment of an ever deepening rift between material and spiritual life.” He refers to the idea of equivalence in pataphysics and the absolute and links them to religion. He says “the great merit of Pataphysics is to have confirmed that there is no metaphysical justification for forcing everybody to believe in the same absurdity”.

Cruickshank (2016) wrote a rather funny article on anti-matter. He links the creation of anti-matter atoms at CERN[1] around 1996 with Jarry, saying that he had "beaten them to the punch" with his pataphysics.

Christian Bök (2002) tries to draw science and poetry together using pataphysics as the string that binds them. He compares Jarry and Nietzsche, saying Jarry performs humorously on behalf of literature what Nietzsche performs seriously on behalf of philosophy; both try to create an anti-philosophy (2002). He also claims that science and poetry have a similar history, undergoing the same four phases of distinct change but also that they have not evolved in sync with each other (2002).

> Pataphysics is a surrational perspective that has had an extensive, yet forgotten, influence upon the canonic history of radical poetics. (...) Not only does this avant-garde pseudoscience valorise whatever is exceptional and paralogical; it also sets the parameters for the contemporary relationships between science and poetry. (Bök 2002)

Bök also compares Jarry and Nietzsche in regards to perspectivism (2002). For Nietzsche reality is the effect of a dream world in which "there are many kinds of truths, and consequently there is no truth". And similarly for Jarry, reality is an aspect of ethernity in which "there are only hallucinations, or perceptions" and every "perception is a hallucination which is true". Both argue that no view is absolute as well and pataphysics argues that every viewpoint is dissolute, including its own because no view can offer a norm. Even Jarry's ethernity is nowhere and somewhere at the same time.

In Faustroll, Bök says, "Jarry parodies the discourse of such scientific luminaries, who attempt to demonstrate the utility of science through the dramaturgic performance of a mechanical experiment" (2002).

@ @ @

Regarding the perplexing apostrophe that sometimes appears before the word ′pataphysics: Jarry only ever used the apostrophe on a single occasion, specifying that he did so "in order to avoid a simple pun". What that pun might be has never been fully explained. User JBlum of urbandictionary.com says: "The exact pun to be avoided is the subject of some debate. The debate itself – being, in essence, a debate about a subject which may not truly exist, but exist as another joke by Jarry – might itself be considered a ′pataphysical search, for an 'imaginary solution' to an imaginary problem!" (2007).

[1] Conseil Européen pour la Recherche Nucléaire—European Council for Nuclear Research

According to the college of pataphysics, it is convention to use the apostrophe at the beginning of the word only in reference to Jarry's texts, to the science of imaginary solutions as such. Used as an adjective or in a more unconscious way it is written without the apostrophe.

4.2.1 Symbology

4.1

Probably the most famous symbol of pataphysics is the ***grand gidouille***, the big spiral on Ubu's fat belly—see image 4.1. Not simply because it is a feature of Jarry's most popular creation but also because it represents one of the concepts of pataphysics itself: the antimony. The spiral can be interpreted as two spirals in one, the outer and the inner spiral. They represent the duality of pataphysics, the mutually incompatible in perfect harmony. The college of pataphysics has adopted the spiral for its membership badges, in various colours and sizes for the different ranks of the college.

Another symbol of pataphysics is the green candle which refers to one of Jarry's last endeavours, published posthumously, a vast collection of his journalistic essays (Hugill 2012). Some animals also symbolise pataphysics. The college's vice-curator was a crocodile called Lutembi until 2014 (Hugill 2012). Owls are another symbol; Jarry kept stuffed and live owls (Brotchie 2011) in his flat. The chameleon is another, having the ability to change colour and looking in two directions at the same time.

4.2.2 Antimony

The antimony is the mutually incompatible. It appears everywhere in Jarry's writings. It represents the duality of things, the echo or symmetry, the good and the evil at the same time. Examples are the plus-minus, the faust-troll, the haldern-ablou, the yes-but, the ha-ha and the paradox.

The 'Ha Ha', the only words Bosse-da-Nage ever utters in Faustroll, "is the idea of duality, of echo, of distance, of symmetry, of greatness and duration, of the two principles of good and evil." (Hugill 2012) Referring to the 'yes-but' statement, Hugill says "this may be taken as a standard pataphysical response to any proposition (including this one)." And most obviously the antimony can be seen in all the contradictions that pataphysics is so fond of.

§ 10.2.4

The implementation of this concept as an algorithm for text search is described in chapter 10.2.4.

4.2.3 Anomaly

The anomaly is the exception. And exceptions are important in pataphysics. But then again everything is equal, so in a pataphysical world no exceptions would exist at all, or rather, everything would be equally exceptional. The anomaly disrupts and surprises. Hugill mentioned a great example of a collection of anomalies: the sourcebook project by William Corliss (n.d.), who collects scientific papers that are anomalous. Bök says it is "the repressed part of a rule which ensure that the rule does not work" (2002).

4.2.4 Syzygy

The syzygy surprises and confuses. It originally comes from astronomy and denotes the alignment of three celestial bodies in a straight line. In a pataphysical context it is the pun, which Jarry called "the syzygy of words" (1996). It usually describes a conjunction of things, something unexpected and surprising. Next to being intentionally funny, puns demonstrate a clever use (or abuse) of grammar, syntax, pronunciation and/or semantics, often taken to a quite scientific level, such that without understanding of what is said and what is the intended meaning, the humour of the pun might be lost. Serendipity is a simple chance encounter but the syzygy has a more scientific purpose. Bök mentions Jarry saying that "the fall of a body towards a centre is the same as the ascension of a vacuum towards a periphery" (2002).

§ 10.2.3

The implementation of this concept as an algorithm for text search is described in chapter 10.2.3.

4.2.5 Clinamen

The clinamen is the unpredictable swerve that Bök calls "the smallest possible aberration that can make the greatest possible difference" (2002). He links it to Lucretius idea of an atom swerving in its streamlined flow to create matter and to Epicurus' parenklisis. But he also points out similarities to ideas like the Situationists' 'détournement', the reuse of pre-existing aesthetic elements and Hugill links it to the Dadaists' ready-mades and Oulipo's verbal games (2012).
§ 10.2.1
An obvious example is Jarry's *merdre*, a swerve of the French word for shit (merde).

The implementation of this concept as an algorithm for text search is described in chapter 10.2.1.

4.2.6 Absolute

The absolute is a reference to a transcended reality. Jarry talks about 'ethernity' in Faustroll (1996).

4.3 Unconscious

4.3.1 Oulipo

> Potential literature is "the search for new forms and structures that may be used by writers in any way they see fit." (Raymond Queneau, as cited in Motte 2007)

> What is the objective of our work? To propose new "structures" to writers, mathematical in nature, or to invent new artificial or mechanical procedures that will contribute to literary activity: props for inspiration as it were, or rather, in a way, aids for creativity. (Raymond Queneau, as cited in Motte 2007)

The Ouvroir de Littérature Potentielle (OULIPO) was already introduced in chapter 2.7 as one of my main inspirations and influences on this research. § 2.7

The OULIPO is a literary movement from the 1960's, originating in France as a subcommittee of the "Collège de Pataphysique". It has since spread to other disciplines. The generic term for Oulipian groups is OUXPO ("Ouvroir d'X Potentielle"), where the X can be replaced with whatever particular subject area you like (typically in french): fine art—OUPEINPO, music—OUMUPO, etc. It has roots in pataphysics although it eventually separated and became a standalone group. Their main philosophy perhaps is to use constraints in order to enhance creative output. Some examples of techniques, taken from (Mathews and Brotchie 2005), invented and used by them are shown below.

Techniques such as the famous 'N+7', 'melting snowball' and 'Mathewś algorithm' (see chapter 2.7 and figure 2.2) are typical examples of Oulipian methods. They have endless applications in as many different disciplines. Motte collated a useful overview of the different Oulipian operations (2007), shown here in table 4.1 and table 4.2. 2.2 4.1 & 4.2

The Oulipian aesthetic is a paradox— formal constraints afford creative liberty, Motte says (2007). He also explained that "Erecting the aesthetic of formal constraint, then, the Oulipo simultaneously devalues inspiration" (2007). François Le Lionnais defined the following three levels in the hierarchy of constraints (Motte 2007). (1) Minimal: constraints on the language in which the text is written, (2) Intermediate: constraints on genre and certain literary norms, and (3)

Table 4.1 – Oulipo—elementary linguistic and literary operations—Part I

	Letter	Phoneme	Syllable
Displacement	anagram, palindrome, pig Latin, metathesis	phonetic palindrome, spoonerism, Rrose Sélavy (Desnos), glossary (Leiris)	syllabic palindrome, spoonerism
Substitution	paragram (printer's error), cryptography	à-peu-près, alphabetical drama	
Addition	prosthesis, epenthesis, paragoge	stuttering	Javanese stuttering, germination, echolalia
Subtraction	abbreviation, aphaeresis, syncope, elision, lipogram, belle absente, constraint of the prisoner	lipophoneme	haplography, liposyllable (Precious [con]straint), shortening
Multiplication (repetition)	tautogram	alliteration, rhyme, homoeuteleuton	stuttering, alliteration, rhyme
Division			diaeresis
Deduction	acrostic, acronym, signet, chronogram		acronym
Contraction	crasis		

Maximum: consciously preelaborated and voluntarily imposed systems of artifice. The use of constraints combats the aleatory or random.

> The Oulipo is anti-chance. (Claude Berge, as cited in Motte 2007)

෧ ෧ ෧

The idea of using constraints to produce creative artefacts has also been picked up in the field of computational creativity. Two examples are described in (Liapis et al. 2013; Toivanen, Järvisalo and Toivonen 2013).

> Constraints are a major factor shaping the conceptual space of many areas of creativity. (Toivanen, Järvisalo and Toivonen 2013)

4.3.2 Borges

§ 2.4 The influence of Jose Luis Borges was already briefly discussed in chapter 2.4.

Table 4.2 – Oulipo—elementary linguistic and literary operations—Part II

	Word	**Syntagm**	**Sentence**	**Paragraph**
Displacement	Mathews's Algorithm, permutations (Lescure), word palindrome, inversion	reversion, inversion, anastrophe	Mathews's Algorithm	Mathews's Algorithm
Substitution	metonymy, S+7, homosyntaxism, L.S.D., translation, antonymic translation	perverbs (Mathews), proverbs, aphorism, homophony, untraceable locutions	homophony, holorhyme	
Addition	redundance, pleonasm	interpolation, encasement	tireur à la ligne, larding	tireur à la ligne
Subtraction	liponym, La Rien que la Toute la (Le Lionnais)	ellipsis, brachylogia, zeugma	coupeur à la ligne	censure
Multiplication (repetition)	epanalepsis, pleonasm, anaphora, defective rhyme	reduplication	leitmotif, refrain	
Division	de-portmanteau word, etymology, tmesis	Roussellian procedure (phonic dislocation), hendiadys	dislocation	
Deduction	haikuization	proverbs on rhymes, edges of poem	citation, tireur à la ligne, collage	plagiarism, anthology
Contraction	portmanteau word	syntagmatic amalgam (Doukipudonktan)		résumé

Hugill sees him as an unconscious pataphysician (2012).

Borges' text *The analytical language of John Wilkins* (2000) contains a brilliant example of pataphysical thinking and coincidentally a good example of the kinds of search results of pata.physics.wtf.

Referring to a certain Chinese dictionary entitled "The Celestial Emporium of Benevolent Knowledge" animals can be divided into:

> 1. those belonging to the Emperor
> 2. those that are embalmed
> 3. those that are tame
> 4. pigs
> 5. sirens
> 6. imaginary animals
> 7. wild dogs
> 8. those included in this classification
> 9. those that are crazy-acting
> 10. those that are uncountable
> 11. those painted with the finest brush made of camel hair
> 12. miscellaneous
> 13. those which have just broken a vase
> 14. those which, from a distance, look like flies
>
> (Borges 2000)

This kind of categorisation has also been briefly discussed by Foucault in his book *The Order of Things* (1966).

૭ ૭ ૭

Other concepts that are pataphysical or can be linked to it in a sense are alchemy and quantum mechanics. Alchemy because of its laws of equivalence and the union of opposites (Hugill 2012) and quantum mechanics because of principles of uncertainty, indeterminacy and the idea of the multiverse.

> Because string theory is speculation based on ideas that are themselves speculative (i.e., theories of general relativity and quantum mechanics), string theory is not in fact physics, but 'pataphysics.
>
> Likewise, string theory and quantum calculations are, increasingly, not descriptive of an actual reality, but are simply mathematical pataphors.
>
> (JBlum 2007)

CREATIVITY

5

From high Olympus prone her flight she bends,
rare courage and grandeur of conception,
congratulating herself apparently on the cleverness,
appeared distorted to my vision.

Had he had any bad design,
having uttered these words the vision left me,
if any thought by flight to escape,
taking his flight towards warmer and sunnier regions.

Inspire à mon oncle cette vision décourageante de l'avenir,
être et l'invention du jeu de ce,
besoin de satisfaire l'imagination d'objets rares ou grandioses.

Some may call vision,
a man of invaluable ability,
mobiles parois de L'imagination.

Creativity does not have a universally accepted definition. Creativity is a human quality and definitions don't necessarily lend themselves to be applied to computers as well. There are aspects that come up often, like novelty and value, but some that rarely pop up, like relevance and variety. Creativity can be studied at various 'levels' (neurological, cognitive, and holistic/systemic), from different 'perspectives' (subjective and objective) and 'characteristics' (combinational, exploratory and transformative). Creativity should be seen as a continuum, there is no clear cut-off point or Boolean answer to say precisely when a person or piece of software has become creative or not.

Linda Candy identified 3 approaches for studying creativity (2012):

Research Design
: Experimental, psychometric, observational, . . .

Research Focus
: Human attributes, cognitive processes or creative outcomes.

Research Evidence
: Real-time observation, historical data, artificial (laboratory) or natural (real world settings).

Richard Mayer identified five big questions of human creativity research and different approaches with their own methodologies and goals (1999):

1. Is creativity a property of people, products, or processes?
2. Is creativity a personal or social phenomenon?
3. Is creativity common or rare?
4. Is creativity domain-general or domain-specific?
5. Is creativity quantitative or qualitative?

Psychometric	(creativity as a mental trait): quantitative measurement, controlled environments, ability based analysis
Psychological	(creativity as cognitive processing): controlled environments, quantitative measurements, cognitive task analysis
Biographical	(creativity as a life story): authentic environments, qualitative descriptions, quantitative measurements
Biological	(creativity as a physiological trait): physiological measures
Computational	(creativity as a mental computation): formal modelling
Contextual	(creativity as a context-based activity): social, cultural and evolutionary context

Mayer identified the challenge of developing a "clearer definition of creativity" and "a combination of research methodologies that will move the field from speculation to specification" (1999).

This chapter introduces relevant models of human and computer creativity and describes the disciplines of computational creativity and creative computing.

5.1 In Humans

Creativity is usually defined as ***the ability to use original ideas to create something new and surprising of value***. We generally speak of creative 'ideas' rather than products, since creative products merely provide evidence of a creative process that has already taken place.

> Creativity is the interaction among aptitude, process, and environment by which an individual or group produces a perceptible product that is both novel and useful as defined within a social context
>
> (Plucker et al, as cited in Jordanous and Keller 2012)

5.1.1 Four Stages

Henri Poincaré and Graham Wallas have defined a popular model of the creative process (it was suggested by Poincaré (2001) and formulated by Wallas (1926)). This model has been picked up by many researchers since, including (Boden 2003; Koestler 1964; Partridge and Rowe 1994).

1. Preparation – focusing the mind on the problem
2. Incubation – unconscious internalising
3. Illumination – eureka moment from unconsciousness to consciousness
4. Verification – conscious evaluation of the idea and elaboration. . .

Weisberg, however, criticises the stages of incubation and illumination (as cited in Partridge and Rowe 1994), saying that the creative process is really just simple problem solving, and that incubation is what he calls 'creative worrying'. Problem solving was defined in similar steps by George Pólya (1957).

> First, we have to **understand** the problem; we have to see clearly what is required. Second, we have to see how the various items are connected, how the unknown is linked to the data, in order to obtain the idea of the solution, to make a **plan**. Third, we **carry out** our plan. Fourth, we **look back** at the completed solution, we review and discuss it. (Pólya 1957, his emphasis)

5.1.2 Four P's

Mel Rhodes identified four common themes of creativity, which he termed "the four P's of creativity" (1961):

Persons personality, intellect, temperament, physique, traits, habits, attitudes, self-concept, value systems, defence mechanisms and behaviour.

Process motivation, perception, learning, thinking and communication.

Press relationship between human beings and their environment

Products a thought which has been communicated to other people in the form of words, paint, clay, metal, stone, fabric, or other material.

Rhodes highlighted the importance of a holistic view on creativity through these four areas of study, which he hoped would become the basis of a unified theory of creativity.

In a similar fashion, Ross Mooney identified four aspects of creativity (as cited in Sternberg 1999).

1. The creative environment
2. The creative person
3. The creative process
4. The creative product

5.1.3 Four C's

5.1 James Kaufman and Ronald Beghetto developed a model of creativity called the "four C model" (2009). Figure 5.1 shows the relationship between the so called 4 C's.

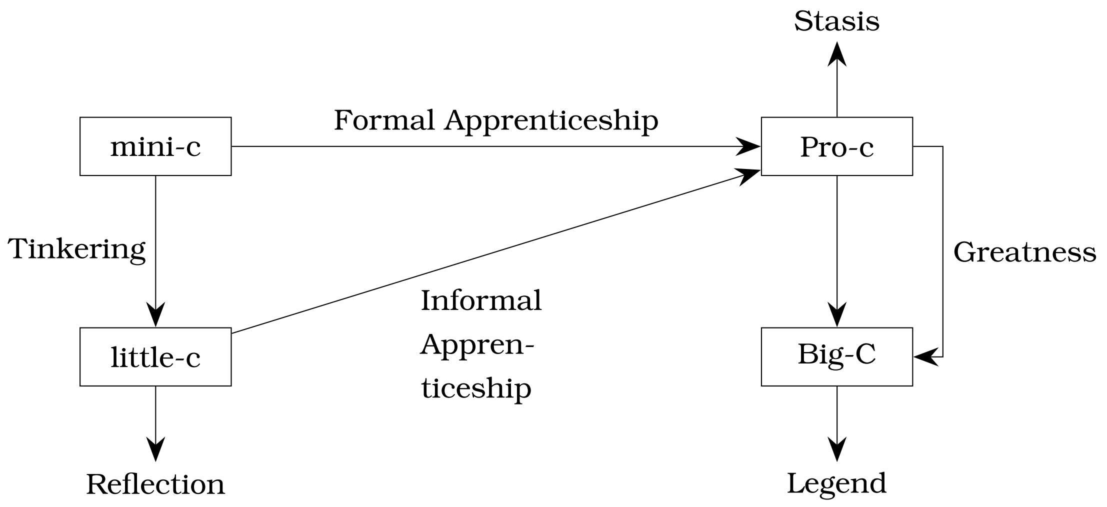


Figure 5.1 – The 4 C model

Big-C *Eminent Accomplishments.* Consists of clear-cut, eminent creative contributions. It often requires a degree of time. Indeed, most theoretical conceptions of Big-C nearly require a posthumous evaluation.

Pro-c *Professional Expertise.* Represents the developmental and effortful progression beyond little-c. The concept of Pro-c is consistent with the expertise acquisition approach of creativity.

Little-c *Everyday Innovation.* More focused on everyday activities, such as those creative actions in which the non-expert may participate each day.

Mini-c *Transformative Learning.* Encompasses the creativity inherent in the learning process. "Mini-c is defined as the novel and personally meaningful interpretation of experiences, actions, and events." (Beghetto and Kaufman 2007) Central to the definition of mini-c creativity is the dynamic, interpretive process of constructing personal knowledge and understanding within a particular socio-cultural context. Moreover, mini-c stresses that mental constructions that have not (yet) been expressed in a tangible way can still be considered highly creative. Mini-c highlights the intrapersonal, and more process focused aspects of creativity.

All 4 C's Openness to new experiences, active observation, and willingness to be surprised and explore the unknown.

5.1.4 Four Types

Sternberg and Kaufman identified a set of personality traits that are associated with creative people in their *Handbook of Creativity* (1999). These are: independence of judgement, self-confidence, and attraction to complexity, aesthetic orientation, and tolerance for ambiguity, openness to experience, psychoticism, risk taking, androgyny, perfectionism, persistence, resilience, and self-efficacy. It is easy to find common characteristics among creative people but that doesn't mean that these automatically make a person or a product creative.

Timothy Leary took this idea of common characteristics a bit further and suggested there are four types of creative personalities (1964). From his ideas we can draw the conclusion that a creative person needs to be able to make novel combinations from novel ideas.

Reproductive Blocked
: no novel combinations, no direct experience

Reproductive Creator
: no direct experience, but crafty skill in producing new combinations of old symbols

Creative Creator
: new experience presented in novel performances

Creative Blocked
: new direct experience expressed in conventional modes

Tables A.1 and A.2 in appendix A.7 are in Leary's words and show the detailed description of these personality types. § A.7

5.1.5 Three Domains

Arthur Koestler published his study on creativity entitled *The Act of Creation* (1964). His main contribution to the field is probably the concept of 'bisociation', a term he coined for the idea of two "self-consistent but habitually incompatible frames of reference" intersecting to give rise to new creative ideas (1964). It is interesting however to look at some of his other views on creativity as well.

He splits creativity into three domains—a triptych—without sharp boundaries: humour, discovery and art (see table 5.1). ▦ 5.1 All creative acts traverse the three domains of this triptych from left to right, that is, the emotional climate of the creator changes "from an absurd through an abstract to a tragic or lyric view of existence" during the process (1964). Central to all three domains is the "discovery of hidden similarities", or bisociation. Koestler differentiates between associative thinking and bisociative thinking. He links those broadly to habit and originality, respectively. More specifically, associative thinking is conscious, logical, habitual, rigid, repetitive and conservative and bisociative thinking is unconscious, intuitive, original, flexible, novel and destructive/constructive.

Table 5.1 – Koestler's creative triptych

Humour	→	**Discovery**	→	**Art**
Laugh		Understand		Marvel
Riddle		Problem		Allusion
Debunking		Discovering		Revealing
Coincidence		Trigger		Fate
Aggressive		Neutral		Sympathetic

5.1.6 Three Processes

Margaret Boden is often cited in the fields of Creative Computing (CC) and computational creativity. Her main interest is in how the human mind works and how computer models of the mind and specific thinking processes can help us understand both better. She has provided two important contributions to the field. The first is her description of three distinct forms of creativity describes in this section and the second is her important distinction between two senses of
§ 5.1.7
creativity as described in section 5.1.7 (Boden 2003).

> (Creativity is) the ability to come up with ideas or artefacts that are **new, surprising and valuable**.
> (Boden 2003, her emphasis)

She identified three distinct forms or cognitive processes of how creativity can happen. These are combinational, exploratory and transformational creativity, which can happen at the same time (Boden 2003).

Combinational creativity
: making unfamiliar combinations of familiar ideas; juxtaposition of dissimilar; bisociation; deconceptualisation

Exploratory creativity
: exploration of conceptual spaces; noticing new things in old spaces

Transformative creativity
: transformation of space; making new thoughts possible by altering the rules of old conceptual space

Central to these three forms is the idea of a 'conceptual space'. For any idea, its conceptual space describes the characteristics and constraints that define it in its most fundamental way. The conceptual space of a tea cup would contain information like: it is a container that can hold a hot fluid, it should hold about a half a pint of fluid and it might or might not be built in such a way as to not burn the hand that carries it. The specific colour of the cup or what material it is made of for example are not contained in its conceptual space.

Combinational creativity is the most common form of the three and is concerned with the unusual juxtaposition of common ideas. This aspect is highlighted in her definition of creativity, which requires novelty and surprise. The main idea is that any particular combination of ideas has to be unusual, causing surprise, but not (necessarily) the individual ideas themselves. She safeguards against purely random combination by including the usefulness of the result as a requirement in the definition. Exploratory creativity requires a person (or computer program) to fully explore the conceptual space of an idea and find unusual or interesting aspects of it. This form of creativity is about pushing an idea to its limits. Transformational creativity takes this exploration one step further. Once the limits of an idea have been identified, they can be transformed. This means that we can step out of the normal conceptual space of an idea, create a new one, alter or ignore the given constraints, add new ones, etc.

Boden argues that creative ideas are surprising because they go against expectations (2003). She also believes that constraints support creativity and are even essential for it to happen, which echos the OULIPO philosophy mentioned in chapter 4.3.1. § 4.3.1

> Constraints map out a territory of structural possibilities which can then be explored, and perhaps transformed to give another one. (Boden 2003)

Bipin Indurkhya argues that there are two main cognitive mechanisms of creativity: namely juxtaposition of dissimilar and deconceptualisation. He says that we are constrained by associations of our concept networks that we inherit and learn in our lifetime, but that computers do not have those conceptual associations and have therefore an advantage when it comes to creative thinking (1997).

5.1.7 Two Levels

§ 5.1.6

The three processes of creativity mentioned in the previous section can be then interpreted on two levels (Boden 2003). Any idea should be viewed and evaluated at the appropriate level. Consider the following scenario. A child and a professional architect both build a corbeled arch out of material available to them. Who is being creative here? The level of expertise is clearly different between the two. The child has no experience and is experimenting with the possibilities and limitations of the building blocks (exploring their conceptual space) while the architect has studied the technique for years and is simply applying knowledge he has learned from others (familiar use of a familiar idea). Clearly the child is being more creative in this example. Boden proposed to view and judge the creativity of these two persons separately by differentiating between two levels of creativity, a personal one and a historical one.

'Psychological creativity' (P-creativity) is a personal kind of creativity that is novel in respect to an individual and 'historical creativity' (H-creativity) is fundamentally novel in respect to the whole of human history (Boden 2003).

The child in the earlier scenario was P-creative but the architect was neither, he was simply applying his trained skills.

> P–creativity involves coming up with a surprising, valuable idea that's new to the person who comes up with it. It doesn't matter how many people have had that idea before. But if a new idea is H–creative, that means that (so far as we know) no one else has had it before: it has arisen for the first time in human history.
> (Boden 2003)

5.2 In Computers

This section introduces some models that try to implement creative thinking models in computers. It is really just a survey of different concepts and views and does not immediately apply to my specific research on creative search tools.

Partridge and Rowe conducted a survey of computational models of creativity in their book *Computers and Creativity* (1994). They mention the computer as an unbiased[1] medium for executing creative programs. Some of the computational methodologies they discussed are as follows, many taken from classical AI research. These are: generative grammars, discovery programs, rule based systems, meta-rules (which reason about and create new rules), analogical mechanisms, flexible representations, classifier systems, decentralised systems, connectionist systems, neural networks, and emergent memory models. Classifier systems for example, consist of a set of rules and a message list as shown below.

[1] I will later argue that this is not possible, since a computer cannot be judged without taking the programmer into account. See chapter 9.1.2 for more details.

1. Place input messages on current message list.
2. Find all rules that can match messages.
3. Each such rule generates a message for the new message list.
4. Replace current message list with the new one.
5. Process new list for any system output.
6. Return to step 1.

These can easily be combined with genetic algorithms to enable the system to learn an appropriate classifier set. This is called emergent behavior. Another approach is connectionism also known as neural networks. Partridge and Rowe then go on to describe their emergent-memory model. They are applying the ideas of Poincaré and Wallas (Poincaré 2001; Wallas 1926) and are heavily influence by Minsky's theory of K-lines (1980, 1988). They define the following characteristics for creative programs:

- flexible knowledge representation scheme
- representational imprecision
- multiple representations
- self-assessment
- full elaboration

Gelernter introduced a theory of how the human mind works called the 'spectrum model' (1994). It is based on the idea of mental focus and relates well to creativity. According to him we have a thought spectrum. The higher the mental focus, the more awake we are, the more adult we are and modern, logical and rational, convergent, abstract and detailed. The less focused we are the younger or ancient or dreaming we are. Low focus thoughts are metaphoric, hallucinations, divergent, creative, inspirations, concrete, ambient and emotional. Emotions glue low focus thoughts together.

He gives a good example of his own computer program that is being trained by a set of simple pairs (or memories) in the form **mood: happy** for example. These sets of pairs form the experience of the system, the memory that the system can access. It's fetching all memory pairs that match a certain probe, then generalizes them and picks out a feature that is common to all and then uses that to probe further if necessary.

He models his spectrum concept in a way that if we want the system to operate at low focus, more memory pairs would be fetched and more generalised features are deducted and so on. He describes his FGP program (Fetch Generalise Project) as follows (Gelernter 1994).

1. Fetch memory pairs in response to a probe (question).
2. Sandwich them together and peer through the bundle at once.
3. Notice the common features that emerge strongly (generalise).
4. Pick out interesting emergent details and probe further if necessary.

With low focus the system would not generalise as much and just pick out a particular memory, etc. The computer system Gelernter has built seems very limited. His memory pairs cannot describe everything. For example they can describe states but not actions.

This idea of accessing thoughts/memories is very closely related to searching. Searching an index in a search engine is similar to remembering, trying to find all memories related to the current thought for example.

૭ ૭ ૭

Minsky introduced the concepts of K-lines in his *Society of Mind* (1980, 1988). It is basically a theory of memory. He claims that the "function of a memory is to recreate a state of mind". His theory of k-lines is as follows.

> When you get an idea, or solve a problem, or have a memorable experience, you create what we shall call a K-line. This K-line gets connected to those mental agencies that were actively involved in the memorable mental event. When that K-line is later activated, it reactivates some of those mental agencies, creating a partial mental state resembling the original. (Minsky 1980, 1988)

This theory works quite well with Gelernter's idea of memory. K-lines in this sense are nothing other than Gelernter's memory pairs.

Minsky and his student Push Singh have formalised the idea of a panalogy[2]. The idea is that an idea can and should be conceptualised in many different ways. This could be seen as a fall-back mechanism for computational models, if one approach didn't return the desired/expected results.

૭ ૭ ૭

[2]The concept of the panalogy was originally discussed in the initial proposal for this research project. See section 2.1

Elton explains the concept of 'artificial creativity' which can be seen as a sub-area of AI. AI research isn't 'human' enough, he argues, it needs to include less abstract ideas like emotions, morals, aesthetic sensibility and creativity. He goes on to explain in detail how production, evaluation and etiology play a role in everything (Elton 1995).

Opposed to the traditional approach of AI to study some aspect of the human brain in a specific domain only, he argues that in order to understand creativity we need to look at more than that. Creativity arises from a process that is not isolated. The etiology (its history) is essential for something to be classed as creative. Generation (of artefacts or ideas) cannot count as creative if it doesn't undergo evaluation in the process. In order to evaluate we need a sound knowledge of the relevant domain.

> We want creative evaluation to be influenced by a longstanding history of interaction with entities (of whatever kind) in the world. (Elton 1995)

Computer systems can be seen in two perspectives: plastic and implastic (resettable). Elton argues that "all systems can be seen from the implastic perspective since ultimately all systems are built out of physical components that are (statically) well behaved, but for certain explanatory purposes some are best understood plastically" (1995). Connectionist networks are an example of a plastic system. The brain is a plastic system too.

5.3 In Academia

Two transdisciplinary fields of study have emerged from the variety of disciplines concerned. These are computational creativity and creative computing. The former lies at the cross section of AI and cognitive science and is about achiev- § 5.3.1
ing creativity through computation and the latter is mostly distinguished by its § 5.3.2
involvement in art and is about doing computations in a creative way. Creative computing focuses on the process of creativity and 'tacit knowledge' rather than just the outcome as is more often the case in computational creativity. There is also an area called speculative computing discussed later on. § 5.3.3

The concept of creative computing has existed for some time but is only just starting to evolve into a recognised mainstream discipline within computer science. As of 2016, there is a journal (IJCrC n.d.), conference (ISCC n.d.) and several undergraduate courses dedicated to creative computing[3]. Computational creativity, on the other hand, has emerged as a field within AI research and

[3]Courses (in the UK) are offered by Bath Spa University, University of the Creative Arts, Ed-

overlaps with creative computing ideas to some extent. There's also a conference (ICCC n.d.), which has been going for several years.

Perhaps a good example of creative computing is the International Obfuscated C Code Contest (Broukhis, Cooper and Noll n.d.). The competition revolves around writing compilable/runnable code, while visually appearing as obfuscated as possible. They value unusuality, obscurity and creativity but expect contestants to follow the strict rules and constraints of the C programming language. Obfuscation in itself isn't necessarily the hallmark of creative computing but it is
</> 2.1 one possible use-case. See the example competition entry shown on page 20.

Examples of computational creativity are Simon Colton's *Painting Fool* (n.d.) or Harold Cohen's *AARON* (***Learn about AARON's history*** n.d.); both are computer programs that paint pictures. Kurzweil's *Cybernetic Poet* (***Ray Kurzweil's Cybernetic Poet*** 2001) is a classic example of a program that produces poetry.

§ 5.1 But how may we apply the insights into creativity described in chapter 5.1 to computing? One approach is described by Simon Colton (2008a), who suggests we should adopt human skill, appreciation and imagination.

> Without skill, they would never produce anything. Without appreciation, they would produce things which looked awful. Without imagination, everything they produced would look the same. (Colton 2008a)

He thinks that evaluating the worth of an idea or product is the biggest challenge facing computational creativity. Whereas in conventional problem solving success is defined as finding a solution, in a creative context more aesthetic considerations have to be taken into account.

5.3.1 Computational Creativity

Computational creativity is a relatively new discipline and as such not well defined. Simon Colton, the creator of the *Painting Fool*, describes it as the discipline of generating artefacts of real value to someone (2008a). This is in contrast to classic AI problem solving.

One could say that computational creativity is the attempt at giving computers the skills, appreciation and imagination needed to produce creative artefacts.

inburgh Napier University, Glyndwrd University, Goldsmiths University of London, Queen Mary University of London, and University of West London (according to UCAS 2016)

Whether or not this makes the computer creative, or the programmer, is another question that I will address in chapter 9. § 9

Computational creativity has emerged from within AI research. Simon Colton and Geraint Wiggins argue AI falls within a problem solving paradigm: "an intelligent task, that we desire to automate, is formulated as a particular type of problem to be solved" (2012), whereas "in Computational Creativity research, we prefer to work within an artefact generation paradigm, where the automation of an intelligent task is seen as an opportunity to produce something of cultural value" (2012). Hugill and Yang on the other hand argue its role within computer science falls under the scientific paradigm (2013), (see also A. Eden 2007), as opposed to Creative Computing (CC) in the technocratic paradigm.

The International Association for Computational Creativity (ACC) promotes the advancement of computational creativity which they define as follows.

> Computational Creativity is the art, science, philosophy and engineering of computational systems which, by taking on particular responsibilities, exhibit behaviours that unbiased observers would deem to be creative. (ICCC 2014)

Computational creativity is multidisciplinary, bringing together researchers from AI, cognitive psychology, philosophy, and the arts. Its main goal is to model, simulate or replicate human creativity using a computer and it has the following three aims:

- to construct a program or computer capable of human-level creativity
- to better understand human creativity and to formulate an algorithmic perspective on creative behavior in humans
- to design programs that can enhance human creativity without necessarily being creative themselves

The ACC manages the annual International Conference on Computational Creativity (ICCC), whose recent call for papers (for ICCC 2014) gives a useful insight into their research agenda. It can be broken down as follows:

- Paradigms, metrics, frameworks, formalisms, methodologies, perspectives
- Computational creativity-support tools
- Creativity-oriented computing in education
- Domain-specific vs. generalised creativity
- Process vs. product

- Domain advancement vs. creativity advancement
- Black box vs. accountable systems

Simon Colton and Geraint Wiggins have also identified several directions for future research in the field (2012):

1. Continued integration of systems to increase their creative potential.
2. Usage of web resources as source material and conceptual inspiration for creative acts by computer.
3. Using crowd sourcing and collaborative creative technologies bringing together evaluation methodologies based on product, process, intentionality and the framing of creative acts by software.

5.3.2 Creative Computing

In the recent first issue of the International Journal of Creative Computing (IJCrC) Hugill and Yang introduced Creative Computing (CC) formally as a new discipline (2013) with an overarching theme of 'unite and conquer' (Yang 2013). Its broad aim is to "reconcile the objective precision of computer systems (mathesis) with the subjective ambiguity of human creativity (aesthesis)" (Hugill and Yang 2013). Hugill and Yang suggest CC falls within the technocratic paradigm of computing (A. Eden 2007, see also), i.e. the discipline is closest related to software engineering, rather than mathematics or natural sciences. They identify five main topics for CC research (2013):

Challenges transdisciplinarity, cross-compatibility, continuity and adaptivity

Types creative development of a product, development of a CC product and development of tool for creativity support

Mechanisms Boden's combinational, exploratory and transformational creativity

Methods development of suitable transdisciplinary CC research methodologies

Standards resist standardisation, novel, continuous user interaction, creative mechanisms

The main challenge is for technology to become "more adaptive, smarter and better engineered to cope with frequent changes of direction, inconsistencies, irrelevancies, messiness and all the other vagaries that characterise the creative process" (Hugill and Yang 2013). In part, these issues are due to the transdisciplinary nature of the field and factors such as common semantics, standards, requirements and expectations are typical challenges. Hugill and Yang therefore argue that creative software should be flexible and able to adapt to ever changing requirements, it should be evaluated and re-written continuously and it should be cross-compatible.

The different types of CC highlight the different aspects researchers and practitioners focus on during their work. These are:

Process	creative development of a computing product,
Product	development of a Creative Computing product and
Community	development of computing environment to support creativity.

The creative computing process should consist of combinational, exploratory and transformational activities (in the sense of Margaret Boden's theory, as discussed in section 5.1.6). § 5.1.6

Broadly speaking, you could say that the 'process' approach works bottom-up and the 'product' approach works top-down.

The 'community' approach reflects what Hugill and Yang call the "local and global levels", which represent the two types of creativity identified by Boden (P- and H-creativity). It is concerned with developing environments, tools and methods and the management of these. Cross-compatibilty can be seen as the solution to these personal/local and historical/global issues. § 5.1.7 § 5.1.1

Similar to the four step model of the creative process by Poincare and Wallas (2001; 1926) and the four stage model of problem solving by Pólya (1957), Hugill and Yang propose a four step model for the creative computing process. They do this by comparing the acts of artistic creation and software engineering in some detail. They found that the two processes follow essentially the same levels of abstraction (from the abstract to the concrete) (2013):

1. Motivation (digitised thinking)
2. Ideation (design sketch)
3. Implementation (creative system)
4. Operation (effect of system/revision) § 8.1.1

The similarity to other creativity models is further discussed in chapter 8.1.1.

5.3.3 Speculative Computing

SpecLab is a book by Johanna Drucker (2009) about her experiences as a researcher moving between disciplines and the projects she worked on as part of the DH laboratory at the University of Virginia, USA. Several of those projects had pataphysical inspirations.

In his review on the back cover of the book, John Unsworth says that Drucker "emphasizes the graphical over the textual, the generative over the descriptive, and aesthetic subjectivity over analytical objectivism" (2009). Her main argument is that in the design of digital knowledge representation, subjectivity and aesthetics are an essential feature. She confronts logical computation with aesthetic principles with the idea that design is information.

Aesthesis is the theory of ambiguous and subjective knowledge, ideological and epistemological, while mathesis is formal objective logic and they contrast each other. Knowledge is always interpretation and subjectivity is always in opposition to objectivity. Knowledge becomes synonymous with information and as such can be represented digitally as data and metadata.

> Arguably, few other textual forms will have greater impact on the way we read,receive, search, access, use and engage with the primary materials of humanities studies than the metadata structures that organize and present that knowledge in digital form. (Drucker 2009)

But how is this metadata analysed? How do we analyse this type of structured data? And most important of all, she asks, what can be considered as data, what can be expressed in those quantitative terms or other standard parameters? Is data neutral, raw or does it have meaning? Here she also points out that many information structures have graphical analogies and can be understood as diagrams that organize the relations of elements within the whole.

Because "computational methods rooted in formal logic tend to be granted more authority [. . .] than methods grounded in subjective judgement", she introduces the discipline of Speculative Computing (SP) as the solution to that problem. The concept can be understood as a criticism of mechanistic, logical approaches that distinguish between subject and object.

> Speculative computing takes seriously the destabilization of all categories of entity, identity, object, subject, interactivity, process, or instrument. In short, it rejects mechanistic, instrumental, and formally logical approaches, replacing them with concepts of autopoiesis (contingent interdependency), quantum

> poetics and emergent systems, heteroglossia, indeterminacy and potentiality, intersubjectivity, and deformance. Digital Humanities is focused on texts, images, meanings, and means. Speculative Computing engages with interpretation and aesthetic provocation. (Drucker 2009)

Pataphysics governs exceptions and anomalies and she introduces a, what she calls, 'patacritical' method of including those exceptions as rules—even if repeatability and reliability are compromised. Bugs and glitches are privileged over functionality, and although that may not be as useful in all circumstances, they are "valuable to speculation in a substantive, not trivial, sense." In an essay on SP she says "Pataphysics celebrates the idiosyncratic and particular within the world of phenomena, thus providing a framework for an aesthetics of specificity within generative practice" (Drucker and Nowviskie 2007). To break out of the formal logic and defined parameters of computer science we need speculative capabilities and pataphysics. "The goal of pataphysical and speculative computing is to keep digital humanities from falling into mere technical application of standard practices" (2007).

> 'Pataphysics inverts the scientific method, proceeding from and sustaining exceptions and unique cases, while quantum methods insist on conditions of indeterminacy as that which is intervened in any interpretative act. Dynamic and productive with respect to the subject-object dialectic of perception and cognition, the quantum extensions of speculative aesthetics have implications for applied and theoretical dimensions of computational humanities. (Drucker and Nowviskie 2007)

With this, Drucker introduces Speculative Aesthetics, which links interface design with other speculative computing principles. She also refers to Kant and his idea of 'purposiveness without purpose'. She says that the appreciation of design as it is (outside of utility) is the goal of speculative aesthetics.

5.3.4 Digital Humanities

Burdick et al. have written a manifesto for the field of Digital Humanities (DH) (2012). Computing has had a big impact on the humanities as a discipline so much so that DH was born of the encounter between the two. In essence, it is characterised by "collaboration, transdisciplinarity and an engagement with computing" but it should not simply be reduced to 'doing the humanities digitally' (2012). It spans across many traditional areas of research, such as literature, philosophy, history, art, music, design and of course computer science—making the concept of transliteracy fundamental.

> Transliteracy is "the ability to read, write and interact across a range of platforms, tools and media from signing and orality through handwriting, print, TV, radio and film, to digital social networks." (Thomas et al. 2007)

"The field of Digital Humanities may see the emergence of polymaths who can 'do it all'": who can research, write, shoot, edit, code, model, design, network, and dialogue with users (Burdick et al. 2012). DH encompasses several core activities which on various levels depend on and support each other.

Design
: Shape, scheme, inform, experience, position, narrate, interpret, remap/reframe, reveal, deconstruct, reconstruct, situate, critique

Curation, analysis, editing, modelling
: Digitise, classify, describe, metadata, organise, navigate

Computation, processing
: Disambiguate, encode, structure, procedure, index, automate, sort, search, calculate, match

Networks, infrastructure
: Cultural, institutional, technical, compatible, interoperable, flexible, mutable, extensible

Versioning, prototyping, failures
: Iterate, experiment, take-risks, redefine, beta-test

> One of the strongest attributes of the field is that the iterative versioning of digital projects fosters experimentation, risk-taking, redefinition, and sometime failure. (...) It is important that we do not short-circuit this experimental process in the rush to normalize practices, standardize methodologies, and define evaluative metrics. (Burdick et al. 2012)

A shortened list of the emerging methods Burdick et al. have identified are shown below (2012). A full list can be found in appendix A.3.

- structured mark-up
- natural language processing
- mutability
- digital cultural record
- algorithmic analysis
- distant/close, macro/micro, surface/depth
- parametrics
- cultural mash-ups
- algorithm design
- data visualization
- modelling knowledge
- ambient data
- collaborative authorship
- interdisciplinary teams
- use as performance
- narrative structures
- code as text

- software in a cultural context
- repurposable content and remix culture
- participatory web
- read/write/rewrite
- meta-medium
- polymorphous browsing

TECHNOLOGY

6

Ten thousand soldiers with me I will take,
only thus much I give your Grace to know,
the tenth of August last this dreadful lord,
I'll give thee this neck.

He did so set his teeth and tear it,
the circumstance I'll tell you more at large,
or ten times happier be it ten for one,
if he will touch the estimate.

And tell me he that knows,
a thousand knees ten thousand years together,
stand on the dying neck.

Towards school with heavy looks,
and thus do we of wisdom and of reach,
be an arch.

6.1 Information Retrieval

> Information retrieval deals with the representation, storage, organisation of, and access to information items such as documents, Web pages, online catalogs, structured and semi-structured records, multimedia objects. The representation and organisation of the information items should be such as to provide the users with easy access to information of their interest.
>
> (Baeza-Yates and Ribeiro-Neto 2011)

In simple terms, a typical search process can be described as follows (see figure 6.1). A user is looking for some information so she or he types a search term or a question into the text box of a search engine. The system analyses this query and retrieves any matches from the index, which is kept up to date by a web crawler. A ranking algorithm then decides in what order to return the matching results and displays them for the user. In reality of course this process involves many more steps and level of detail, but it provides a sufficient enough overview. 6.1

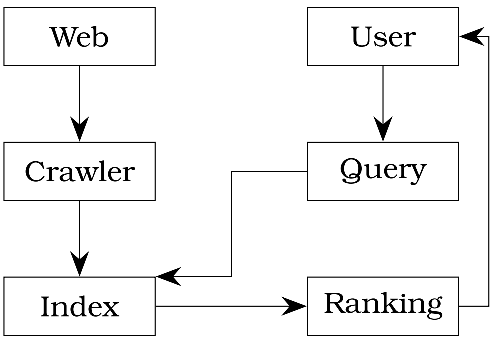


Figure 6.1 – Abstract search engine architecture

Most big web search engines like Google, Baidu or Bing focus on usefulness and relevance of their results (Microsoft 2012; Baidu n.d. Google 2012). Google uses over 200 signals (2012) that influence the ranking of web pages including their original PageRank algorithm (Brin and Page 1998; Page et al. 1999).

Any IR process is constrained by factors like subject, context, time, cost, system and user knowledge (Marchionini and Shneiderman 1988). Such constraints should be taken into consideration in the development of any search tool. A web crawler needs resources to crawl around the web, language barriers may exist, the body of knowledge might not be suitable for all queries, the system might not be able to cater for all types of queries (e.g. single-word vs. multi-word queries), or the user might not be able to understand the user interface, and many more. It is therefore imperative to eliminate certain constraining factors—for example by choosing a specific target audience or filtering the amount of information gathered by a crawler from web pages.

The crawler, sometimes called spider, indexer or bot, is a program that processes and archives information about every available webpage it can find. It does this by looking at given 'seed' pages and searching them for hyperlinks. It then follows all of these links and repeats the process over and over. The Googlebot (n.d.) and the Bingbot (n.d.) are well-known examples.

An index is a list of keywords (called the dictionary or vocabulary) together with a list called 'postings list' that indicates the documents in which the terms occurs. One way to practically implement this is to create a Term-Document Matrix
Σ 6.1 (TDM) as shown in equation 6.1.

$$\begin{array}{c} \\ k_1 \\ k_2 \\ k_3 \end{array} \begin{array}{c} \begin{array}{cc} d_1 & d_2 \end{array} \\ \begin{bmatrix} f_{1,1} & f_{1,2} \\ f_{2,1} & f_{2,2} \\ f_{3,1} & f_{3,2} \end{bmatrix} \end{array} \tag{6.1}$$

where $f_{i,j}$ is the frequency of term k_i in document d_j. To illustrate this with a
6.2 concrete example, figure 6.2 shows a TDM for a selection of words in a corpus containing three documents[1].

- Alfred Jarry: *Exploits and Opinions of Dr. Faustroll, 'Pataphysician* ('Faustroll') (1996)
- Saint Luke: *The Gospel* ('Gospel') (2005)
- Jules Verne: *A Journey to the Centre of the Earth* ('Voyage') (2010)

[1]These texts are part of one of the two corpora used for `pata.physics.wtf`. More information about this can be found in chapters 2.2 and 10.1.1.

	Faustroll	Gospel	Voyage
Faustroll	77	0	0
father	1	28	2
time	34	16	129
background	0	0	0
water	29	7	120
doctor	30	0	0
without	27	7	117
bishop	27	0	2
God	25	123	2

Figure 6.2 – Example TDM for 3 documents and 9 words

The dictionary is usually pre-processed (see section 6.2) to eliminate punctuation and so-called 'stop-words'[2] (e.g. I, a, and, be, by, for, the, on, etc.) which would be useless in everyday text search engines. For specific domains it even makes sense to build a 'controlled vocabulary', where only very specific terms are included (for example the index at the back of a book). This can be seen as a domain specific taxonomy and is very useful for query expansion (explained in the next paragraph).

§ 6.2
§ B.6

Relevance feedback is an idea of improving the search results by explicit or implicit methods. Explicit feedback asks users to rate results according to their perceived relevance or collects that kind of information through analysis of mouse clicks, eye tracking, etc. Implicit feedback occurs when external sources are consulted such as thesauri or by analysing the top results provided by a search engine. There are two ways of using this feedback. It can be displayed as a list of suggested search terms to the user and the user decides whether or not to take the advice, or the query is modified internally without the user's knowledge. This is then called automatic query expansion.

6.1.1 IR Models

There are different models for different needs, for example a multimedia system is going to be different than a text based IR system, and a web based system is going to be different than an offline database system. Even within one such category there could more than one model. Take text based search systems for example. Text can be unstructured or semi-structured. Web pages are typically semi-structured. They contain a title, different sections and paragraphs and so on. An unstructured page would have no such differentiations but only contain

[2] A full list of stopwords in English, French and German can be found in appendix B.6.

simple text. Classic example models are set-theoretic, algebraic and probabilistic. The PageRank algorithm by Google is a link-based retrieval model (Page et al. 1999).

The notation for IR models is a quadruple $[D, Q, F, R(q_i, d_j)]$ (Baeza-Yates and Ribeiro-Neto 2011, adapted from) where,

D	= the set of documents
Q	= the set of queries
F	= the framework e.g. sets, Boolean relations, vectors, linear algebra...
$R(q_i, d_j)$	= the ranking function, with $q_i \in Q$ and $d_j \in D$
t	= the number of index terms in a document collection
V	= the set of all distinct index terms $\{k_1, \ldots, k_t\}$ in a document collection (vocabulary)

This means, given a query q and a set of documents D, we need to produce a ranking score $R(q, d_j)$ for each document d_j in D.

THE BOOLEAN MODEL

One such ranking score is the Boolean model. The similarity of document d_j to query q is defined as follows (Baeza-Yates and Ribeiro-Neto 2011)

$$sim(d_j, q) = \begin{cases} 1 & \text{if } \exists\ c(q) \mid c(q) = c(d_j) \\ 0 & \text{otherwise} \end{cases} \tag{6.2}$$

where $c(x)$ is a 'conjunctive component' of x. A conjunctive component is one part of a declaration in Disjunctive Normal Form (DNF). It describes which terms occur in a document and which ones do not. For example, for vocabulary $V = \{k_0, k_1, k_2\}$, if all terms occur in document d_j then the conjunctive component
6.3 would be $(1, 1, 1)$, or $(0, 1, 0)$ if only term k_1 appears in d_j. Let's make this clearer with a practical example. Figure 6.3 (a shorter version of figure 6.2) shows a vocabulary of 4 terms over 3 documents.

So, we have a vocabulary V of {Faustroll, time, doctor and God} and three documents d_0 = Faustroll, d_1 = Gospel and d_2 = Voyage. The conjunctive component for d_0 is $(1, 1, 1, 1)$. This is because each term in V occurs at least once. $c(d_1)$ and $c(d_2)$ are both $(0, 1, 0, 1)$ since the terms 'Faustroll' and 'doctor' do not occur in either of them.

	Faustroll	Gospel	Voyage
Faustroll	77	0	0
time	34	16	129
doctor	30	0	0
God	25	123	2

Figure 6.3 – Example TDM for 9 words and 3 documents (short)

Assume we have a query $q =$ doctor $\wedge$ (Faustroll $\vee$ $\neg$ God). Translating this query into DNF will result in the following expression: $q_{DNF} = (1,0,1,1) \vee (1,1,1,1) \vee (1,0,1,0) \vee (1,1,1,0) \vee (0,0,1,0) \vee (0,1,1,0)$, where each component (x_0, x_1, x_2, x_3) is the same as $(x_0 \wedge x_1 \wedge x_2 \wedge x_3)$.

One of the conjunctive components in q_{DNF} must match a document conjunctive component in order to return a positive result. In this case $c(d_0)$ matches the second component in q_{DNF} and therefore the Faustroll document matches the query q but the other two documents do not.

The Boolean model gives 'Boolean' results. This means something is either true or false. Sometimes things are not quite black and white though and we need to weigh the importance of words somehow.

TF-IDF

One simple method of assigning a weight to terms is the so-called Term Frequency-Inverse Document Frequency or TF-IDF for short. Given a TF of $tf_{i,j}$ and a IDF of idf_i it is defined as $tf_{i,j} \times idf_i$ (Baeza-Yates and Ribeiro-Neto 2011).

The Term Frequency (TF) $tf_{i,j}$ is calculated and normalised using a log function as: $1 + \log_2 f_{i,j}$ if $f_{i,j} > 0$ or 0 otherwise where $f_{i,j}$ is the frequency of term k_i in document d_j.

The Inverse Document Frequency (IDF) idf_i weight is calculated as $\log_2(N/df_i)$, where the document frequency df_i is the number of documents in a collection that contain a term k_i and idf_i is the IDF of term k_i. The more often a term occurs in different documents the lower the IDF. N is the total number of documents.

$$tfidf_{i,j} = \begin{cases} (1 + \log_2 f_{i,j}) \times \log_2 \frac{N}{df_i} & \text{if } f_{i,j} > 0 \\ 0 & \text{otherwise} \end{cases} \tag{6.3}$$

where $tfidf_{i,j}$ is the weight associated with (k_i, d_j). Using this formula ensures that rare terms have a higher weight and more so if they occur a lot in one

6.1 document. Table 6.1 shows the following details.

$k_0 - k_8$ = [Faustroll,father,time,background,water,doctor,without,bishop,God]
$d_0 - d_2$ = [Faustroll, Gospel, Voyage] (see figure 6.2)
$f_{i,j}$ = the frequence (count) of term k_i in document d_j
$tf_{i,j}$ = the Term Frequency weight
idf_i = the Inverse Document Frequency weight
$tfidf_{i,j}$ = the TF-IDF weight

Table 6.1 – TF-IDF weights

		d_0			d_1			d_2		
	idf	*f*	*tf*	*tfidf*	*f*	*tf*	*tfidf*	*f*	*tf*	*tfidf*
k_0	1.58	77	7.27	11.49	0	0	0	0	0	0
k_1	0	1	1	0	28	5.81	0	2	2	0
k_2	0	34	6.09	0	16	5	0	129	8.01	0
k_3	0	0	0	0	0	0	0	0	0	0
k_4	0	29	5.86	0	7	3.81	0	120	7.91	0
k_5	1.58	30	5.91	9.34	0	0	0	0	0	0
k_6	0	27	5.75	0	7	3.81	0	117	7.87	0
k_7	0.58	27	5.75	3.34	0	0	0	2	2	1.16
k_8	0	25	5.64	0	123	7.94	0	2	2	0

6.1 What stands out in table 6.1 is that the $tfidf_{i,j}$ function returns 0 quite often. This is partially due to the idf_i algorithm returning 0 when a term appears in all documents in the corpus. In the given example this is the case a lot but in a real-world example it might not occur as much.

The Vector Model

The vector model allows more flexible scoring since it basically computes the 'degree' of similarity between a document and a query (Baeza-Yates and Ribeiro-Neto 2011). Each document d_j in the corpus is represented by a document vector $\vec{d_j}$ in t-dimensional space, where t is the total number of terms in the vocabulary.
6.4 Figure 6.4 gives an example of vector $\vec{d_j}$ for document d_j in 3-dimensional space.
6.5 That is, the vocabulary of this system consists of three terms k_a, k_b and k_c. A similar vector $\vec{q}$ can be constructed for query q. Figure 6.5 then shows the similarity between the document and the query vector as the cosine of θ.

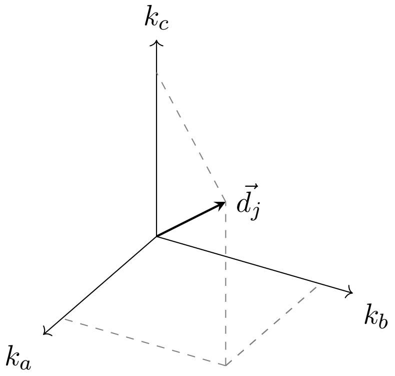


Figure 6.4 – A document vector $\vec{d_j}$

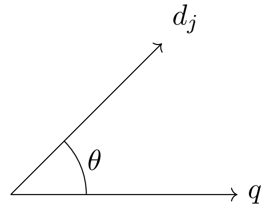


Figure 6.5 – The vector model

$\vec{d_j}$ is defined as $(w_{1,j}, w_{2,j}, \ldots, w_{t,j})$ and similarly $\vec{q}$ is defined as $(w_{1,q}, w_{2,q}, \ldots, w_{t,q})$, where $w_{i,j}$ and $w_{i,q}$ correspond to the TF-IDF weights per term of the relevant document or query respectively. t is the total number of terms in the vocabulary. The similarity between a document d_j and a query q is defined in equation 6.4. Σ 6.4

$$
\begin{aligned}
sim(d_j, q) &= \frac{\vec{d_j} \cdot \vec{q}}{|\vec{d_j}| \times |\vec{q}|} \\
&= \frac{\sum_{i=1}^{t} w_{i,j} \times w_{i,q}}{\sqrt{\sum_{i=1}^{t} w_{i,j}^2} \times \sqrt{\sum_{i=1}^{t} w_{i,q}^2}}
\end{aligned}
\tag{6.4}
$$

Let's consider an example similar to the one used for the TF-IDF section. We have a corpus of three documents (d_0 = Faustroll, d_1 = Gospel, and d_2 = Voyage) and nine terms in the vocabulary ($[k_0, \ldots, k_8]$ = (Faustroll, father, time, background, water, doctor, without, bishop, God)). The document vectors and their corresponding length is given below (with the relevant TF-IDF weights taken from table 6.1). ⊞ 6.1

$$
\begin{aligned}
\vec{d_0} &= (11.49,0,0,0,0,9.34,0,3.34,0) \\
|\vec{d_0}| &= 15.18 \\
\vec{d_1} &= (0,0,0,0,0,0,0,0,0) \\
|\vec{d_1}| &= 0 \\
\vec{d_2} &= (0,0,0,0,0,0,0,1.16,0) \\
|\vec{d_2}| &= 1.16
\end{aligned}
$$

For this example we will use two queries: q_0 (doctor, Faustroll) and q_1 (without, bishop). We then compute the similarity score for between each of the documents compared to the two queries by applying equation 6.4. For the query q_0 the result clearly points to the first document, i.e. the Faustroll text. For query q_1 the score produces two results, with Verne's 'Voyage' scoring highest. Σ 6.4

q_0	$= \text{(doctor, Faustroll)}$	q_1	$= \text{(without, bishop)}$
$\vec{q_0}$	$= (1.58,0,0,0,0,1.58,0,0,0)$	$\vec{q_1}$	$= (0,0,0,0,0,0,0,0.58,0)$
$\lvert\vec{q_0}\rvert$	$= 2.24$	$\lvert\vec{q_1}\rvert$	$= 0.58$
$sim(d_0, q_0)$	$= 0.97$	$sim(d_0, q_1)$	$= 0.22$
$sim(d_1, q_0)$	$= 0$	$sim(d_1, q_1)$	$= 0$
$sim(d_2, q_0)$	$= 0$	$sim(d_2, q_1)$	$= 1$

୭ ୭ ୭

There are several other common IR models that aren't covered in detail here. These include the probabilistic, set-based, extended Boolean and fuzzy set (Miyamoto 1990a,b; Miyamoto and Nakayama 1986; Srinivasan 2001; Widyantoro and Yen 2001) models or latent semantic indexing (Deerwester et al. 1990), neural network models and others (Macdonald 2009; Schütze 1998; Schütze and Pedersen 1995).

6.1.2 Searching vs. Browsing

What is actually meant by the word 'searching'? Usually it implies that there is something to be found, an Information Need (IN); although that doesn't necessarily mean that the searcher knows what he or she is looking for or how to conduct the search and satisfy that need.

From the user's point of view the search process can be broken down into four activities (Sutcliffe and Ennis 1998) reminiscent of classic problem solving techniques (mentioned briefly in chapter 5.1.1) (Pólya 1957):

§ 5.1.1

Problem identification	Information Need (IN),
Need articulation	IN in natural language terms,
Query formulation	translate IN into query terms, and
Results evaluation	compare against IN.

This model poses problems in situations where an IN cannot easily be articulated or in fact is not existent and the user is not looking for anything specific. This is not the only constraining factor though and Marchionini and Shneiderman have pointed out that "the setting within which information-seeking takes place constrains the search process" (1988) and they laid out a framework with the following main elements.

- Setting (the context of the search and external factors such as time, cost)

- Task domain (the body of knowledge, the subject)
- Search system (the database or web search engine)
- User (the user's experience)
- Outcomes (the assessment of the results/answers)

൭ ൭ ൭

Searching can be thought of in two ways, 'information lookup' (searching) and 'exploratory search' (browsing) (Marchionini 2006; Vries 1993). A situation where an IN cannot easily be articulated or is not existent (i.e. the user is not looking for anything specific) can be considered a typical case of exploratory search. The former can be understood as a type of simple question answering while the latter is a more general and broad knowledge acquisition process without a clear goal.

Current web search engines are tailored for information lookup. They do really well in answering simple factoid questions relating to numbers, dates or names (e.g. fact retrieval, navigation, transactions, verification) but not so well in providing answers to questions that are semantically vague or require a certain extend of interpretation or prediction (e.g. analysis, evaluation, forecasting, transformation).

With exploratory search, the user's success in finding the right information depends a lot more on constraining factors and can sometimes benefit from a combination of information lookup and exploratory search (Marchionini 2006).

> Much of the search time in learning search tasks is devoted to examining and comparing results and reformulating queries to discover the boundaries of meaning for key concepts. Learning search tasks are best suited to combinations of browsing and analytical strategies, with lookup searches embedded to get one into the correct neighbourhood for exploratory browsing.
>
> (Marchionini 2006)

De Vries called this form of browsing an "enlargement of the problem space", where the problem space refers to the resources that possibly contain the answers/solutions to the IN (1993). This is a somewhat similar idea to that of Boden's conceptual spaces which she called the "territory of structural possibilities" and exploration of that space "exploratory creativity" (Boden 2003) (see section 5.1.6). § 5.1.6

6.1.3 Ranking

Ranking signals, such as the weights produced by the TF-IDF algorithm in section 6.1.1, contribute to the improvement of the ranking process. These can be content signals or structural signals. Content signals are referring to anything that is concerned with the text and content of a page. This could be simple word counts or the format of text such as headings and font weights. The structural signals are more concerned about the linked structure of pages. They look at incoming and outgoing links on pages. There are also web usage signals that can contribute to ranking algorithms such as the click-stream. This also includes things like the Facebook 'like' button or the Google+ '+1' button which could be seen as direct user relevance feedback as well.

§ 6.1.1

Ranking algorithms are the essence of any web search engine and as such guarded with much secrecy. They decide which pages are listed highest in search results and if their ranking criteria were known publically, the potential for abuse (such as 'Google bombing' (Nicole 2010) for instance) would be much higher and search results would be less trustworthy. Despite the secrecy there are some algorithms like Google's PageRank algorithm that have been described and published in academic papers.

Algorithms

PageRank was developed by Larry Page and Sergey Brin as part of their Google search engine (1998; 1999). PageRank is a link analysis algorithm, meaning it looks at the incoming and outgoing links on pages. It assigns a numerical weight to each document, where each link counts as a vote of support in a sense. PageRank is executed at indexing time, so the ranks are stored with each page directly in the index. Brin and Page define the PageRank algorithm as follows (1998).

Σ 6.5

$$PR(A) = (1-d) + d(\sum_{i=1}^{n} \frac{PR(T_i)}{C(T_i)}) \tag{6.5}$$

A = the page we want to rank and is pointed to by pages T_1 to T_n
n = the total number of pages on the Web graph
$C(A)$ = the number of outgoing links of page A
d = a 'damping' parameter set by the system (typically 0.85) needed to deal with dead ends in the graph

6.6

Figure 6.6 which shows how the PageRank algorithm works. Each smiley represents a webpage. The colours are of no consequence. The smile-intensity indicates a higher rank or score. The pointy hands are hyperlinks. The yellow smiley is the happiest since it has the most incoming links from different sources with only one outgoing link. The blue one is slightly smaller and slightly less smiley even though it has the same number of incoming links as the yellow one because it has more outgoing links. The little green faces barely smile since they have no incoming links at all.

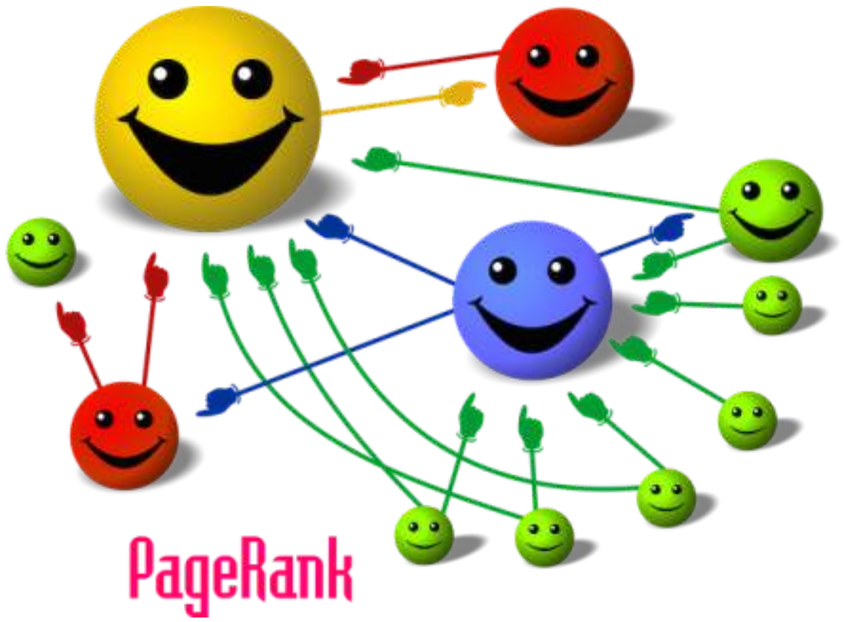


Figure 6.6 – PageRank algorithm illustration (Mayhaymate 2012)

The HITS algorithm also works on the links between pages (Kleinberg 1999; Kleinberg et al. 1999). HITS stands for 'Hyperlink Induced Topic Search' and its basic features are the use of so called hubs and authority pages. It is executed at query time. Pages that have many incoming links are called 'authorities' and page with many outgoing links are called 'hubs'. Equation 6.6 shows the algorithm (Baeza-Yates and Ribeiro-Neto 2011), where S is the set of pages, $H(p)$ is the hub value for page p, and $A(p)$ is the authority value for page p. Σ 6.6

$$\begin{aligned} H(p) &= \sum_{u \in S | p \rightarrow u} A(u) \\ A(p) &= \sum_{v \in S | v \rightarrow p} H(v) \end{aligned} \tag{6.6}$$

Hilltop is a similar algorithm with the difference that it operates on a specific set of expert pages as a starting point. It was defined by Bharat and Mihaila (2000). The expert pages they refer to should have many outgoing links to non-affiliated pages on a specific topic. This set of expert pages needs to be pre-processed at the indexing stage. The authority pages they define must be linked to by one of their expert pages. The main difference to the HITS algorithm then is that their 'hub' pages are predefined.

Another algorithm is the so called Fish search algorithm (De Bra, Houben et al. 1994; De Bra and Post 1994a,b). The basic concept here is that the search starts with the search query and a seed Uniform Resource Locator (URL) as a starting point. A list of pages is then built dynamically in order of relevance following from link to link. Each node in this directed graph is given a priority depending on whether it is judged to be relevant or not. URLs with higher priority are inserted at the front of the list while others are inserted at the back. Special here is that the 'ranking' is done dynamically at query time.

There are various algorithms that follow this approach. For example the shark search algorithm (Hersovici et al. 1998). It improves the process of judging whether or not a given link is relevant or not. It uses a simple vector model with a fuzzy sort of relevance feedback. Another example is the improved fish search algorithm (Luo, Chen and Guo 2005) where an extra parameter allows more control over the search range and time. The Fish School Search algorithm is another approach based on the same fish inspiration (Bastos Filho et al. 2008). It uses principles from genetic algorithms and particle swarm optimization. Another genetic approach is Webnaut (Nick and Themis 2001).

Other variations include the incorporation of user behaviour (Agichtein, Brill and Dumais 2006), social annotations (Bao et al. 2007), trust (Gyongyi, Garcia-Molina and Pedersen 2004), query modifications (Glover et al. 2001), topic sensitive PageRank (Haveliwala 2003), folksonomies (Hotho et al. 2006), SimRank (Jeh and Widom 2002), neural-networks (Shu and Kak 1999), and semantic web (Ding et al. 2004; Du et al. 2007; Kamps, Kaptein and Koolen 2010; Taye 2009; Widyantoro and Yen 2001).

6.1.4 Challenges

Other issues that arise when trying to search the WWW were identified by Baeza-Yates and Ribeiro-Neto as follows (2011).

- Data is distributed. Data is located on different computers all over the world and network traffic is not always reliable.
- Data is volatile. Data is deleted, changed or lost all the time so data is often out-of-date and links broken.
- The amount of data is massive and grows rapidly. Scaling of the search engine is an issue here.

- Data is often unstructured. There is no consistency of data structures.
- Data is of poor quality. There is no editor or censor on the Web. A lot of data is redundant too.
- Data is not heterogeneous. Different data types (text, images, sound, video) and different languages exist.

Since a single query for a popular word can results in millions of retrieved documents from the index, search engine usually adopt a lazy strategy, meaning that they only actually retrieve the first few pages of results and only compute the rest when needed (Baeza-Yates and Ribeiro-Neto 2011). To handle the vast amounts of space needed to store the index, big search engines use a massive parallel and cluster-based architecture (Baeza-Yates and Ribeiro-Neto 2011). Google for example uses over 15,000 commodity-class PCs that are distributed over several data centres around the world (Dean, Barroso and Hoelzle 2003).

6.2 Natural Language Processing

Natural Language Processing (NLP) is a discipline within computer science which is also known as follows (Jurafsky and Martin 2009).

- Speech and language processing
- Human language technology
- Computational linguistics
- Speech recognition and synthesis

Goals of NLP are to get computers to perform useful tasks involving human language such as enabling human-machine communication, improving human-human communication, and text and speech processing.Applications are for example machine translation, automatic speech recognition, natural language understanding, word sense disambiguation, spelling correction, and grammar checking.

There are many tools and libraries available for NLP, including the Natural Language Toolkit (NLTK) Python library (Bird, Klein and Loper 2009; NLTK n.d.) and WordNet (WordNet n.d.) (both of which were used for `pata.physics.wtf`).

6.2.1 Words

A 'lemma' is a set of lexical forms that have the same stem (e.g. go). A 'word-form' is the full inflected or derived form of the word (e.g. goes). A 'word type' is a distinct word in a corpus (repetitions are not counted but case sensitive). A ‘word token’ is any word (repetitions are counted repeatedly). Manning et al. list the following activities related to the word processing of text (2009).

Tokenisation
: discarding white spaces and punctuation and making every term a token

Normalisation
: making sets of words with same meanings, e.g. car and automobile

Case-folding
: converting everything to lower case

Stemming
: removing word endings, e.g. connection, connecting, connected → connect

Lemmatisation
: returning dictionary form of a word, e.g. went → go

WordNet

WordNet is a large lexical database for English, containing 166,000 word form and sense pairs, useful for computational linguistics and NLP (Miller 1995). A synset is a set of synonyms to represent a specific word sense. It is the basic building block of WordNet’s hierarchical structure of lexical relationships.

> Nouns, verbs, adjectives and adverbs are grouped into sets of cognitive synonyms (synsets), each expressing a distinct concept. Synsets are interlinked by means of conceptual-semantic and lexical relations. (WordNet n.d.)

Synonymy	(same-name) a symmetric relation between word forms
Antonymy	(opposing-name) a symmetric relation between word forms
Hyponymy	(sub-name) a transitive relation between synsets
Hypernymy	(super-name) inverse of hyponymy
Meronymy	(part-name) complex semantic relation
Holonymy	(whole-name) inverse of meronymy
Troponymy	(manner-name) is for verbs what hyponymy is for nouns

Other relations not used by WordNet are homonymy (same spelling but different sound and meaning) and heteronymy (same sound but different spelling), homography (same sound and spelling) and heterography (different sound and spelling).

§ C

Appendix C shows an example result produced by WordNet rendered for a web browser.

Regular Expressions

Regular expressions (often shortened to the term 'regex') are used to search a corpus of texts for the occurrence of a specific string pattern[3].

Table 6.2 shows the most common commands needed to build a regular expression. For example, to find an email address in a piece of text the following regex can be used: ⊞ 6.2

```
([a-zA-Z0-9_\-\.]+)@([a-zA-Z0-9_\-\.]+)\.([a-zA-Z]{2,5})
```

Most modern text editors support a form of search using regex and it is often used in NLP.

Table 6.2 – Regular expression syntax

Command	Description
.	any character except newline
\w \d \s	word, digit, whitespace
\W \D \S	not word, digit, whitespace
[abc]	any of a, b, or c
[^abc]	not a, b, or c
[a-g]	character between a & g
^abc$	start / end of the string
a* a+ a?	0 or more, 1 or more, 0 or 1
a{5} a{2,}	exactly five, two or more
ab|cd	match ab or cd

Damerau-Levenstein

The Damerau–Levenshtein distance between two strings a and b is given by $d_{a,b}(|a|,|b|)$ (see equation 6.7) (Damerau 1964; DL Distance n.d. Levenshtein 1966). The distance indicates the number of operations (insertion, deletion, substitution or transposition) it takes to change one string to the other. For example, the words 'clear' and 'clean' would have a distance of 1, as it takes on substitution of the letter 'r' to 'n' to change the word. A typical application would be spelling correction. Σ 6.7

[3]There is also a Regex Crossword puzzle (M. H. Michelsen and O. B. Michelsen 2016).

$$d_{a,b}(i,j) = \begin{cases} \max(i,j) & \text{if} \min(i,j) = 0 \\ \min \begin{cases} d_{a,b}(i-1,j)+1 \\ d_{a,b}(i,j-1)+1 \\ d_{a,b}(i-1,j-1)+1_{a_i \neq b_j} \\ d_{a,b}(i-2,j-2)+1 \end{cases} & \text{if } i,j > 1 \text{ and } a_i = b_{j-1} \text{ and } a_{i-1} = b_j \\ \min \begin{cases} d_{a,b}(i-1,j)+1 \\ d_{a,b}(i,j-1)+1 \\ d_{a,b}(i-1,j-1)+1_{a_i \neq b_j} \end{cases} & \text{otherwise.} \end{cases} \tag{6.7}$$

$1_{(a_i \neq b_j)}$ is equal to 0 when $a_i = b_j$ and equal to 1 otherwise.

- $d_{a,b}(i-1,j)+1$ corresponds to a deletion (from a to b)
- $d_{a,b}(i,j-1)+1$ corresponds to an insertion (from a to b)
- $d_{a,b}(i-1,j-1)+1_{(a_i \neq b_j)}$ corresponds to a match or mismatch, depending on whether the respective symbols are the same
- $d_{a,b}(i-2,j-2)+1$ corresponds to a transposition between two successive symbols

6.2.2 Sequences

N-Grams

We can do word prediction with probabilistic models called N-Grams. They predict the probability of the next word from the previous $N-1$ words (Jurafsky and Martin 2009). A 2-gram is usually called a 'bigram' and a 3-gram a 'trigram'.

Σ 6.8 A basic way to compute the probability of an N-gram is using a Maximum Likelihood Estimation (MLE) shown in equation 6.8 (Jurafsky and Martin 2009) of a word w_n given some history w_{n-N+1}^{n-1} (i.e. the previous words in the sentence for example).

$$P(w_n \mid w_{n-N+1}^{n-1}) = \frac{C(w_{n-N+1}^{n-1} w_n)}{C(w_{n-N+1}^{n-1})} \tag{6.8}$$

For instance, if we want to check which of two words "shining" and "cold" has a higher probability of being the next word given a history of "the sun is", we would need to compute P(shining | the sun is) and P(cold | the sun is) and compare the results. To do this we would have to divide the number of times the sentence "the sun is shining" occurred in a training corpus by the number of times "the sun is" occurred and the same for the word "cold".

Counts (C) are normalised between 0 and 1. These probabilities are usually generated using a training corpus. These training sets are bound to have incomplete data and certain N-grams might be missed (which will result in a probability of 0). Smoothing techniques help combat this problem.

One example is the so-called 'Laplace' or 'add-one smoothing', which basically just adds 1 to each count. See equation 6.9 (Jurafsky and Martin 2009). V is the number of terms in the vocabulary. Σ 6.9

$$P_{Add-1}(w_i \mid w_{i-1}) = \frac{c(w_{i-1}, w_i) + 1}{c(w_{i-1}) + V} \tag{6.9}$$

Another example of smoothing is the so-called 'Good Turing discounting'. It uses "the count of things you've seen *once* to help estimate the count of things you've *never seen*" (Jurafsky and Martin 2009, their emphasis).

୭ ୭ ୭

To calculate the probability of a sequence of n words ($P(w_1, w_2, \ldots, w_n)$ or $P(w_1^n)$ for short) we can use the chain rule of probability as shown in equation 6.10 (Jurafsky and Martin 2009). Σ 6.10

$$\begin{aligned} P(w_1^n) &= P(w_1)P(w_2 \mid w_1)P(w_3 \mid w_1^2) \ldots P(w_n \mid w_1^{n-1}) \\ &= \prod_{k=1}^{n} P(w_k \mid w_1^{k-1}) \end{aligned} \tag{6.10}$$

Instead of using the complete history of previous words when calculating the probability of the next term, usually only the immediate predecessor is used. Σ 6.11 This assumption that the probability of a word depends only on the previous word (or n words) is the called a Markov assumption (see equation 6.11 (Jurafsky and Martin 2009)).

$$P(w_1^n) = \prod_{k=1}^{n} P(w_k \mid w_{k-1}) \tag{6.11}$$

Part-of-Speech Tagging

Parts-of-Speech (POS) are lexical tags for describing the different elements of a sentence. The eight most well-known POS are as follows.

Noun	an abstract or concrete entity
Pronoun	a substitute for a noun or noun phrase
Adjective	a qualifier of a noun
Verb	an action, occurrence, or state of being
Adverb	a qualifier of an adjective, verb, or other adverb
Preposition	an establisher of relation and context
Conjunction	a syntactic connector
Interjection	an emotional greeting or exclamation

More specialised sets of tags exist such as the *Penn Treebank* tagset (Marcus, Santorini and Marcinkiewicz 1993) consisting of 48 different tags, including CC for coordinating conjunction, CD for cardinal number, NN for noun singular, NNS for noun plural, NNP for proper noun singular, VB for verb base form, VBG for verb gerund, DT for determiner, JJ for adjectives, etc. A full table of
§ A.4
these 48 tags can be found in appendix A.4.

The process of adding tags to the words of a text is called 'POS tagging' or just 'tagging'. Below, you can see an example tagged sentence[4].

> In/IN this/DT year/NN Eighteen/CD Hundred/CD and/CC Ninety-eight/CD,/, the/DT Eighth/CD day/NN of/IN February/NNP,/, Pursuant/JJ to/IN article/NN 819/CD of/IN the/DT Code/NN of/IN Civil/NNP Procedure/NNP and/CC at/IN the/DT request/NN of/IN M./NN and/CC Mme./NN Bonhomme/NNP (/(Jacques/NNP)/),/, proprietors/NNS of/IN a/DT house/NN situate/JJ at/IN Paris/NNP,/, 100/CD bis/NN,/, rue/NN Richer/NNP,/, the/DT aforementioned/JJ having/VBG address/NN for/IN service/NN at/IN my/PRP residence/NN and/CC further/JJ at/IN the/DT Town/NNP Hall/NNP of/IN Q/NNP borough/NN ./.

Maximum Entropy

Hidden Markov or maximum entropy models can be used for sequence classification, e.g. part-of-speech tagging.

> The task of classification is to take a single observation, extract some useful features describing the observation, and then, based on these features, to classify the observation into one of a set of discrete classes.
>
> (Jurafsky and Martin 2009)

[4]This is actually the very first sentence in Jarry's Faustroll book (1996).

A classifier like the maximum entropy model will usually produce a probability of an observation belonging to a specific class. Equation 6.12 shows how to calculate the probability of an obersvation (i.e. word) x being of class c as $p(c|x)$ (Jurafsky and Martin 2009). Σ 6.12

$$p(c|x) = \frac{\exp(\sum_{i=0}^{N} w_{ci} f_i(c,x))}{\sum_{c' \in C} \exp(\sum_{i=0}^{N} w_{c'i} f_i(c',x))} \quad (6.12)$$

$f_i(c, x)$ = the feature (e.g. "this word ends in *-ing*" or "the previous word was *the*")
w_i = the weight of the feature f_i

Grammars

A language is modelled using a grammar, specifically a 'Context-Free-Grammar'. Such a grammar normally consists or rules and a lexicon. For example a rule could be 'NP → Det Noun', where NP stands for noun phrase, Det for determiner and Noun for a noun. The corresponding lexicon would then include facts like Det → a, Det → the, Noun → book. This grammar would let us form two noun phrases 'the book' and 'a book' only. Its two parse trees would then look like figure 6.7: 6.7

Figure 6.7 – Two parse trees for grammar 'NP → Det Noun'

Parsing is the process of analysing a sentence and assigning a structure to it. Given a grammar, a parsing algorithm should produce a parse tree for a given sentence. The parse tree for the first sentence from Faustroll is shown below, in horizontal format for convenience.

```
(ROOT
  (S
    (PP (IN In)
      (NP (DT this) (NN year) (NNPS Eighteen) (NNP Hundred)
        (CC and)
        (NNP Ninety-eight)))
```

```
(, ,)
(NP
  (NP (DT the) (JJ Eighth) (NN day))
  (PP (IN of)
    (NP (NNP February) (, ,) (NNP Pursuant)))
  (PP
    (PP (TO to)
      (NP
        (NP (NN article) (CD 819))
        (PP (IN of)
          (NP
            (NP (DT the) (NNP Code))
            (PP (IN of)
              (NP (NNP Civil) (NNP Procedure)))))))
    (CC and)
    (PP (IN at)
      (NP
        (NP (DT the) (NN request))
        (PP (IN of)
          (NP (NNP M.)
            (CC and)
            (NNP Mme) (NNP Bonhomme))))))
  (PRN (-LRB- -LRB-)
    (NP (NNP Jacques))
    (-RRB- -RRB-))
  (, ,)
  (NP
    (NP (NNS proprietors))
    (PP (IN of)
      (NP
        (NP (DT a) (NN house) (NN situate))
        (PP (IN at)
          (NP (NNP Paris))))))
  (, ,)
  (NP (CD 100) (NN bis))
  (, ,))
(VP (VBP rue)
  (NP
    (NP (NNP Richer))
    (, ,)
    (NP (DT the) (JJ aforementioned)
```

```
      (UCP
        (S
          (VP (VBG having)
            (NP
              (NP (NN address))
              (PP (IN for)
                (NP (NN service))))
            (PP (IN at)
              (NP (PRP$ my) (NN residence)))))
        (CC and)
        (PP
          (ADVP (RBR further))
          (IN at)
          (NP
            (NP (DT the) (NNP Town) (NNP Hall))
            (PP (IN of)
              (NP (NNP Q))))))
      (NN borough))))
  (. .)))
```

This particular tree was generated using the Stanford Parser (2016).

Named Entity Recognition

A named entity can be anything that can be referred to by a proper name, such as person, place or organisation names and times and amounts and these entities can be appropriately tagged.

Example (first sentence in Faustroll):

> In this [year Eighteen Hundred and Ninety-eight, the Eighth day of February]^TIME, Pursuant to article [819]^NUMBER of the [Code of Civil Procedure]^DOCUMENT and at the request of [M. and Mme. Bonhomme (Jacques)]^PERSON, proprietors of a house situate at [Paris, 100 bis, rue Richer]^LOCATION, the aforementioned having address for service at my residence and further at the [Town Hall]^FACILITY of [Q borough]^LOCATION.

So-called 'gazetteers' (lists of place or person names for example) can help with the detection of these named entities.

Evaluation

7

Score,
quel grade avais,
of my cooler judgment,
and inquires after the evacuations of the toad on the horizon.

His judgment takes the winding way Of question distant,
if not always with judgment,
and showed him every mark of honour,
three score years before.

Designates him as above the grade of the common sailor,
but I was of a superior grade,
travellers of those dreary regions marking the site of degraded Babylon.

Mark the Quilt on which you lie,
und da Sie grade kein weißes Papier bei sich hatten,
and to draw a judgement from Heaven upon you for the Injustice.

7.1 Evaluating Search

Generally, computer systems are evaluated against functional requirements and performance specifications. Traditional Information Retrieval (IR) however is usually evaluated using two metrics known as precision and recall (Baeza-Yates and Ribeiro-Neto 2011). Precision is defined as the fraction of retrieved documents that are relevant, while recall is defined as the fraction of relevant documents that are retrieved.

$$Precision = \frac{\text{relevant documents retrieved}}{\text{retrieved documents}} \tag{7.1}$$

$$Recall = \frac{\text{relevant documents retrieved}}{\text{relevant documents}} \tag{7.2}$$

Note the slight difference between the two. Precision tells us how many of all retrieved results were actually relevant (of course this should preferable be very high) and recall simply indicates how many of all possible relevant documents we managed to retrieve. This can be easily visualised as as shown in figure 7.1.

7.1
Σ 7.3

Precision is typically more important than recall in web search, so often evaluation is reduced to measuring the Mean Average Precision (MAP) value, which can be calculated using the formula in equation 7.3 (Baeza-Yates and Ribeiro-Neto 2011), where R_i is the set of results for query i, $P(R_i[k])$ is the precision value for result k for query i and $|R_i|$ is the total number of results.

$$MAP_i = \frac{1}{|R_i|} \sum_{k=1}^{|R_i|} P(R_i[k]) \tag{7.3}$$

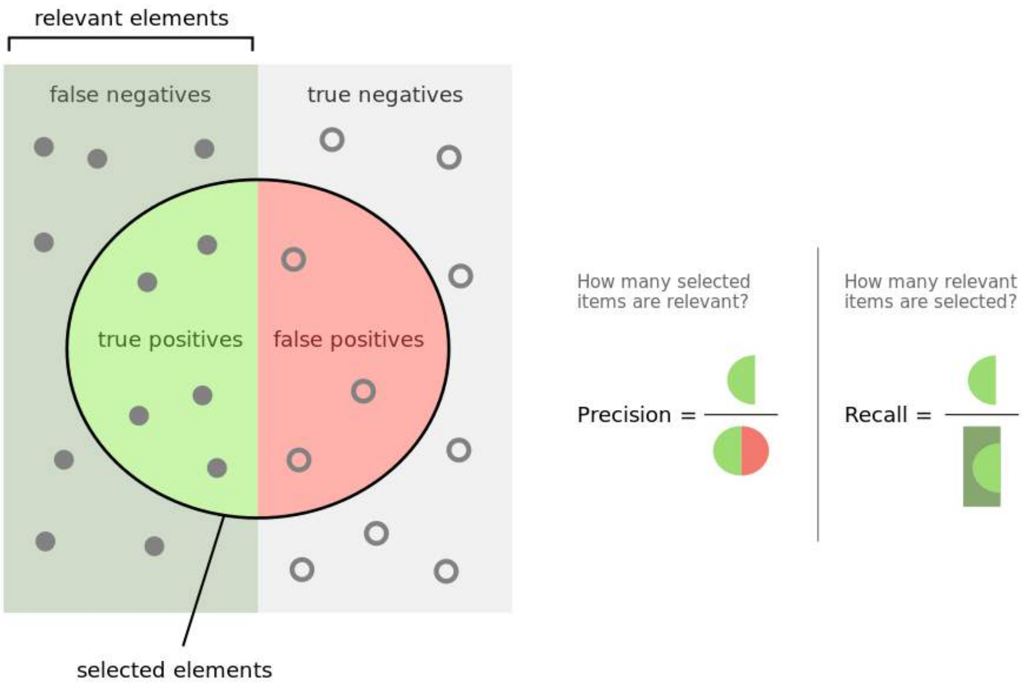


Figure 7.1 – Precision and recall (Walber 2014)

But for many web searches ii is not necessary to calculate the average of all results, since users don't inspect results after the first page very often and it is therefore desirable to have the highest level of precision in the first page of results maybe. For this purpose it is common to measure the average precision of web search engines after only a few documents have been seen. This is called 'Precision at n' or 'P@n' (Baeza-Yates and Ribeiro-Neto 2011). So for example this could be P@5, P@10, or P@20. To compare two ranking algorithms, we would calculate P@10 for each of them over an average of 100 queries maybe and compare the results and therefore the performance of the algorithm.

The Text REtrieval Conference (TREC) (TREC 2016) provides large test sets of data (TREC 2011) to participants and let's them compare results. They have specific test sets for web search comprised of crawls of *.gov* web pages.

There are certain other factors that can be or should be evaluated when looking at a complete search system, as shown below (Baeza-Yates and Ribeiro-Neto 2011).

- Speed of crawling.
- Speed of indexing data.
- Amount of storage needed for data.
- Speed of query response.

- Amount of queries per given time period.

Ranking is another issue that could be considered to pre-evaluate web pages at indexing time rather than query time. This was previously discussed in chapter 6.1.3. § 6.1.3

୭ ୭ ୭

Evaluating creative search is more complex, as the notion of 'relevance' is very different and this will be addressed in chapter 9. § 9

Sawle, Raczinski and Yang (2011) discussed an initial approach to measure the creativity of search results in 2011. Based on a definition of creativity by Boden (as explained in chapter 5.1.6), we attempted to define creativity in a way which § 5.1.6 could be applied to search results and provide a simple metric to measure it. A copy of this paper can be found in appendix E. § E

7.2 Evaluating Creative Computers

This section moves on from evaluating search and focuses on evaluating creativity in computers.

> The evaluation of artificial creative systems in the direct form currently practiced is not in itself empirically well-grounded, hindering the potential for incremental development in the field. (Bown 2014)

Evaluating human creativity objectively seems problematic; evaluating computer creativity seems even harder. There are many debates across the disciplines involved. Taking theories on human creativity (see section 5.1) and directly ap- § 5.1 plying them to machines (see section 5.2) seems logical but may be the wrong § 5.2 (anthropomorphic) approach. Adapting Mayer's five big questions (1999) to ma- § 9.1.1 chines does not seem to capture the real issues at play. Instead of asking if creativity is a property of people, products, or processes we might ask if it is a property of any or all of the following:

- programmers
- users
- machines[1]

[1]This is problematic until the posited AI singularity (Schmidhuber 2006b).

- products
- processes

For instance, is the programmer the only creative agent, or are users (i.e. audiences or participants in interactive work) able to modify the system with their own creative input? Similarly for any instance of machine creativity, we might ask if it is:

- local (e.g. limited to a single machine, program or agent)
- networked (i.e. interacts with other predefined machines or programs)
- web-based (e.g. is distributed and/or open to interactions, perhaps via an API)

Norton, Heath and Ventura highlight the importance of dealing with 'evaluator bias' when using human judges for evaluating any form of creativity. They identified 5 main problems as follows (2015).

1^{st} **problem**	Do we assess products or processes?
2^{nd} **problem**	What are the measurable assessment criteria?
3^{rd} **problem**	How do we un-ambiguate ambigous terminology?
4^{th} **problem**	Which methodology to use for the assessment?
5^{th} **problem**	How do we compensate for biases?

This point is also strengthend by Lamb, Brown and Clarke, saying that "non-expert judges are very poor at using metrics to evaluate creativity" and that the criteria they tested were not "objective enough to produce trustworthy judgments" (2015).

7.2.1 Output minus Input

Discussions from computational creativity often focus on very basic questions such as "whether an idea or artefact is valuable or not, and whether a system is acting creatively or not" (Pease and Colton 2011). Certain defining aspects of
§ 5
creativity, such as novelty and value (as discussed in chapter 5), are often used to measure the outcome of a creative process. These are highlighted throughout
§ 9
the following pages and further addressed in chapter 9.

One recurring theme is the clear separation of training data input and creative output in computers. Pease, Winterstein and Colton called this principle "output minus input" (2001). The output in this case is the creative product but the input is not the process. Rather, it is the 'inspiring set' (comprised of explicit knowledge such as a database of information and implicit knowledge input by a programmer) or training data of a piece of software.

> The degree of creativity in a program is partly determined by the number of novel items of value it produces. Therefore we are interested in the set of valuable items produced by the program which exclude those in the inspiring set.
> (Colton, Pease and Ritchie 2001)

They also suggest that all creative products must be "novel and valuable" (Pease, Winterstein and Colton 2001) and provide several measures that take into consideration the context, complexity, archetype, surprise, perceived novelty, emotional response and aim of a product. In terms of the creative process itself they only discuss randomness as a measurable approach. Elsewhere, Pease et al discuss using serendipity as an approach (2013).

Graeme Ritchie supports the view that creativity in a computer system must be measured "relative to its initial state of knowledge" (2007). He identifies three main criteria for creativity as "novelty, quality and typicality" (2007), although he argues that "novelty and typicality may well be related, since high novelty may raise questions about, or suggest a low value for, typicality" (2001, 2007). He proposes several evaluation criteria which fall under the following categories (2007): basic success, unrestrained quality, conventional skill, unconventional skill, avoiding replication and various combinations of those. Dan Ventura later suggested the addition of "variety and efficiency" to Ritchie's model (2008).

It should be noted that 'output minus input' might easily be misinterpreted as 'product minus process', however, that is not the case. In fact, Pease, Winterstein and Colton argue that "the process by which an item has been generated and evaluated is intuitively relevant to attributions of creativity" (2001), and that "two kinds of evaluation are relevant; the evaluation of the item, and evaluation of the processes used to generate it" (2001). If a machine simply copies an idea from its inspiring set then it just cannot be considered creative and needs to be disqualified so to speak.

7.2.2 Creative Tripod

Simon Colton came up with an evaluation framework called the *creative tripod*. The tripod consists of three behaviours a system or artefact should exhibit in order to be called creative. The three legs represent "skill, appreciation, and imagination" and three different entities can sit on it, namely the programmer, the computer and the consumer. Colton argues that the perception "that the software has been skillful, appreciative and imaginative, then, regardless of the behaviour of the consumer or programmer, the software should be considered creative” (2008a,b). As such a product can be considered creative, if it appears to be creative. If not all three behaviours are exhibited, however, it should not be considered creative (Colton 2008a,b).

> Imagine an artist missing one of skill, appreciation or imagination. Without skill, they would never produce anything. Without appreciation, they would produce things which looked awful. Without imagination, everything they produced would look the same. (Colton 2008b)

૭ ૭ ૭

Davide Piffer suggests that there are three dimensions of human creativity that can be measured, namely “novelty, usefulness/appropriateness and impact/influence” (2012). As an example of how this applies to measuring a person’s creativity he proposes ‘citation counts’ (Piffer 2012). While this idea works well for measuring scientific creativity maybe, he does not explain how this would apply to a visual artist for example.

7.2.3 SPECS

Anna Jordanous proposed 14 key components of creativity (which she calls an “ontology of creativity”) (2012), from a linguistic analysis of creativity literature which identified words that appeared significantly more often in discussions of creativity compared to unrelated topics (2012).

> The themes identified in this linguistic analysis have collectively provided a clearer “working” understanding of creativity, in the form of components that collectively contribute to our understanding of what creativity is. Together these components act as building blocks for creativity, each contributing to the overall presence of creativity; individually they make creativity more tractable and easier to understand by breaking down this seemingly impenetrable concept into constituent parts. (Jordanous and Keller 2012)

The 14 components Jordanous collated are: (2012)

1. Active Involvement and Persistence
2. Generation of Results
3. Dealing with Uncertainty
4. Domain Competence
5. General Intellect

6. Independence and Freedom
7. Intention and Emotional Involvement
8. Originality
9. Progression and Development
10. Social Interaction and Communication
11. Spontaneity / Subconscious Processing
12. Thinking and Evaluation
13. Value
14. Variety, Divergence and Experimentation

Jordanous also found that "evaluation of computational creativity is not being performed in a systematic or standard way" (2011) and proposed 'Standardised Procedure for Evaluating Creative Systems (SPECS)' (2012):

1. Identify a definition of creativity that your system should satisfy to be considered creative:
   a) What does it mean to be creative in a general context, independent of any domain specifics?
      - Research and identify a definition of creativity that you feel offers the most suitable definition of creativity.
      - The 14 components of creativity identified in Chapter 4 are strongly suggested as a collective definition of creativity.
   b) What aspects of creativity are particularly important in the domain your system works in (and what aspects of creativity are less important in that domain)?
      - Adapt the general definition of creativity from Step 1a so that it accurately reflects how creativity is manifested in the domain your system works in.
2. Using Step 1, clearly state what standards you use to evaluate the creativity of your system.
   - Identify the criteria for creativity included in the definition from Step 1 (a and b) and extract them from the definition, expressing each criterion as a separate standard to be tested.
   - If using Chapter 4's components of creativity, as is strongly recommended, then each component becomes one standard to be tested on the system.
3. Test your creative system against the standards stated in Step 2 and report the results.
   - For each standard stated in Step 2, devise test(s) to evaluate the system's performance against that standard.
   - The choice of tests to be used is left up to the choice of the individual researcher or research team.
   - Consider the test results in terms of how important the associated aspect of creativity is in that domain, with more important aspects of creativity being given greater consideration than less important aspects. It is not necessary, however, to combine all the test results into one aggregate score of creativity.

The SPECS model essentially means that we cannot evaluate a creative computer system objectively, unless steps 1 and 2 are predefined and publically available for external assessors to execute step 3. Creative evaluation can therefore be seen as a move from subjectivity to objectivity, i.e. defining subjective criteria for objectively evaluating a product in terms of the initial criteria.

> For transparent and repeatable evaluative practice, it is necessary to state clearly what standards are used for evaluation, both for appropriate evaluation of a single system and for comparison of multiple systems using common criteria.
> (Jordanous 2012)

This is further strengthened by Richard Mayer stating that we need a "clearer definition of creativity" (1999) and Linda Candy arguing for "criteria and measures [for evaluation] that are situated and domain specific" (2012).

Jordanous also defined 5 'meta-evaluation criteria' of correctness, usefulness, faithfulness as a model of creativity, usability of the methodology, and generality (2014).

7.2.4 MMCE

Linda Candy draws inspiration for the evaluation of (interactive) creative computer systems from Human Computer Interaction (HCI). The focus of evaluation in HCI has been on usability, she says (2012). She argues that in order to successfully evaluate an artefact, the practitioner needs to have "the necessary information including constraints on the options under consideration" (2012).

Evaluation happens at every stage of the process (i.e. from design → implementation → operation). Some of the key aspects of evaluation Candy highlights are:

- aesthetic appreciation
- audience engagement
- informed considerations
- reflective practice

7.2
§ 5

She goes on to introduce the Multi-dimensional Model of Creativity and Evaluation (MMCE) (shown in figure 7.2) with four main elements of people, process, product and context (2012) similar to some of the models of creativity we have seen in chapter 5.

She proposes the following values or criteria for measurement (2012).

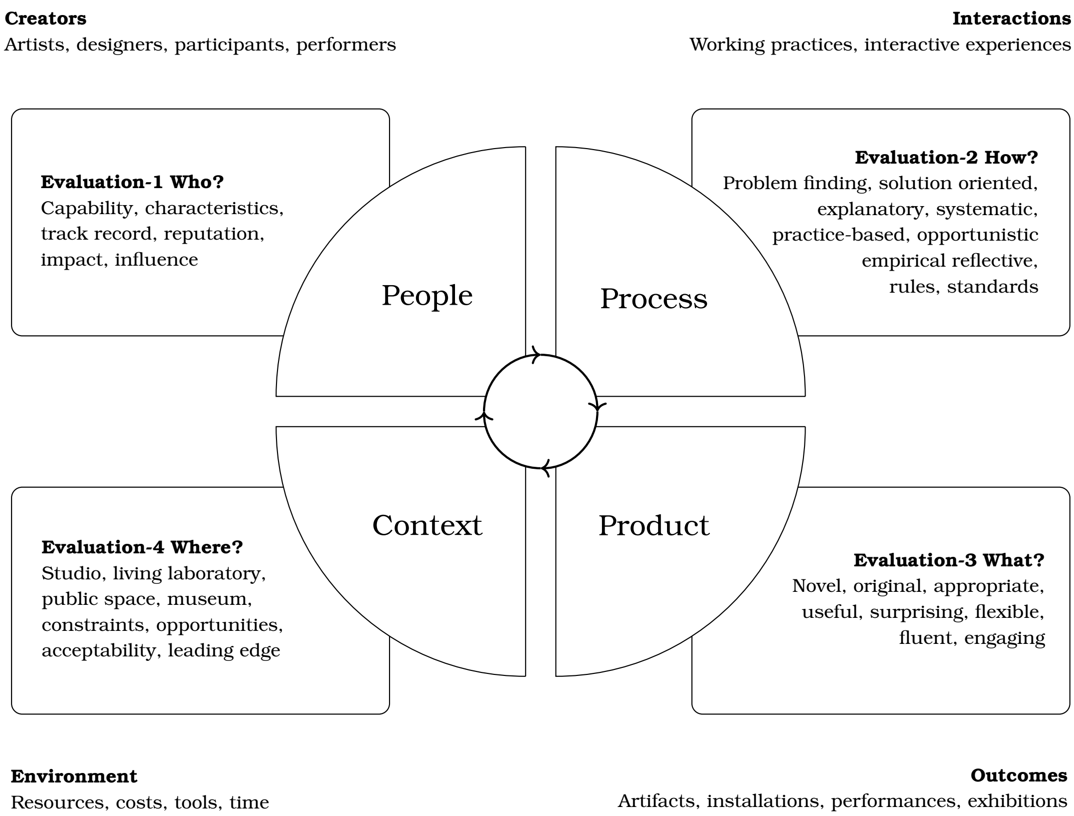


Figure 7.2 – Candy's Multi-dimensional Model of Creativity and Evaluation

People capabilities, characteristics, track record, reputation, impact, influence (profile, demographic, motivation, skills, experience, curiosity, commitment)

Process problem finding, solution oriented, exploratory, systematic, practice-based, empirical, reflective, opportunistic, rules, standards (opportunistic, adventurous, curious, cautions, expert, knowledgeable, experienced)

Product novel, original, appropriate, useful, surprising, flexible, fluent, engaging (immediate, engaging, enhancing, purposeful, exciting, disturbing)

Context studio, living laboratory, public space, museum, constraints, opportunities, acceptability, leading edge (design quality, usable, convincing, adaptable, effective, innovative, transcendent)

7.2.5 CSF

Geraint Wiggins introduced a formal notation and set of rules for the description, analysis and comparison of creative systems called Creative Search Framework § 5.1.6 (CSF)(2006) which is largely based on Boden's theory of creativity (2003). The framework uses three criteria for measuring creativity: "relevance, acceptability and quality". Graeme Ritchie then contributed to this framework with several revisions (2012).

The CSF provides a formal description for Boden's concepts of exploratory and transformational creativity. Wiggins's 'R–transformation' and 'T–transformation'

§ 5.1.6

is akin to Boden's 'H-creativity' and 'P-creativity' respectively. To enable the transition from exploratory to transformational creativity in his framework, Wiggins introduced meta-rules which allow us to redefine our conceptual space in a new way.

It is important to note here that the exploratory search in an IR sense (as dis-

§ 6.1.2

cussed in section 6.1.2) should not be mistaken with the topic at hand. Exploratory search (for a creative solution to a problem) in the Wiggins/Ritchie/Boden sense happens one step before transformational search. This means that we want to end up with transformational tools from this framework (rather than exploratory ones) to use in our exploratory web search system.

Ritchie described the CSF as a set of initial concepts, which create 'further concepts one after another, thus "exploring the space"' but also argued that a search system would practically only go through a limited number of steps and therefore proposed some changes and additions to the framework (2012). He summarised Wiggins' original CSF as consisting of the following basic elements:

1. the universal set of concepts U,
2. the language for expressing the relevant mappings L,
3. a symbolic representation of the acceptability map R,
4. a symbolic representation of the quality mapping E,
5. a symbolic representation of the search mechanism T,
6. an interpreter for expressions like 3 and 4 $[\,]$, and
7. an interpreter for expressions like 5 $\langle\ ,\ ,\ \rangle$.

This set of elements is described as the 'object-level' (enabling exploratory search). The 'meta-level' (enabling transformational search) has the same seven elements with one exception; the universal set of concepts U contains concepts described at the object-level. This allows transformations to happen; concepts from the object-level are searched using criteria and mechanisms (elements 2 to 5) from the meta-level, giving rise to a new and different subset of concepts to those which an object-level search would have produced.

A typical search process would go as follows. We start with an initial set of concepts C that represent our conceptual space and a query. We then explore C and find any elements that match the query with a certain quality (norm and value criteria) in a given amount of iterations. This produces the object-level set of exploratory concepts (in Boden's sense) which we would call the traditional search results. To get creative results we would need to apply the meta-level search (Boden's transformational search) with slightly different quality criteria.

Wiggins explained various situations of creativity not taking place (uninspiration and aberration) in terms of his framework as shown below. For example, a system not finding any valuable concepts would be expressed as $[E](U) = 0$ (in Wiggins' original notation). While this approach seems counter-intuitive and impractical, it actually provides an interesting inspiration on how to formulate some of our pataphysical concepts in terms of the CSF (see chapter 13.4). § 13.4

Hopeless Uninspiration	$V_\alpha(X) = \emptyset$ valued set of concepts is empty
Conceptual Uninspiration	$V_\alpha(N_\alpha(X)) = \emptyset$ no accepted concepts are valuable
Generative Uninspiration	$elements(A) = \emptyset$ set of reachable concepts is empty
Aberration	B is the set of reachable concepts not in $[N]_\alpha(X)$ and $B \neq \emptyset$ search goes outside normal boundaries
Perfect Aberration	$V_\alpha(B) = B$
Productive Aberration	$V_\alpha(B) \neq \emptyset$ and $V_\alpha(B) \neq B$
Pointless Aberration	$V_\alpha(B) = \emptyset$

§ 13

The potential of these definitions of 'uncreativity' is further explored in chapter 13.

7.2.6 Individual Criteria

Many separate attempts exist at defining an evaluation model that focuses on a single criterion for creativity.

One such example is a model for evaluating the 'interestingness' of computer generated plots (Pérez y Pérez and Ortiz 2013).

Another approach looks at "quantifying surprise by projecting into the future" (Maher, Brady and Fisher 2013).

Bown looks at "evaluation that is grounded in thinking about interaction design, and inspired by an anthropological understanding of human creative behaviour" (2014). He argues that "systems may only be understood as creative by looking at their interaction with humans using appropriate methodological tools" (2014). He proposed the following methodology.

1. The recognition and rigorous application of 'soft science' methods wherever vague unoperationalised terms and interpretative language is used.
2. An appropriate model of creativity in culture and art that includes the recognition of humans as 'porous subjects', and the significant role played by generative creativity in the dynamics of artistic behaviour.

Others argues that creativity can be measured by looking at the overall 'unexpectedness' of an artefact (Kazjon and Maher 2013).

McGregor, Wiggins and Purver introduce the idea of creativity as an "intimation of dualism, with its inherent mental representations, is a thing that typical observers seek when evaluating creativity" (2014).

Another attempt to evaluate computational creativity suggests that systems must go through a sequence of 4 phases "in order to reach a level of creativity acceptable to a set of human judges" (Negrete-Yankelevich and Morales-Zaragoza 2014). The phases are as follows.

1. **Structure** is the basic architecture of a piece; it is what allows spectators to make out different parts of it, to analyze it to understand its main organization.
2. **Plot** is the specialization scaffold of the structure to one purpose; it is the basis for narrative and the most detailed part of planned structure. It is upon plots that pieces are rendered.
3. **Rendering** is a particular way in which the plot was developed and filled with detail in order to be delivered to the audience.
4. **Remediation** is the transformation of a creative piece already rendered into another one, re-rendered, possibly into another media.

França et al. propose a system called *Regent-Dependent Creativity* (RDC) to address the "lack of domain independent metrics" and which combines "the Bayesian Surprise and Synergy to measure novelty and value, respectively " (2016).

This dependency relationship is defined by a pair $P(r; d)$ associated with a numeric value v, where r is the regent (a feature that contributes to describing an artifact), d is the dependent (it can change the state of an attribute), and v is a value that represents the intensity of a specific pair in different contexts.

> For example, an artifact car can be described by a pair $p_i(color; blue)$, where blue changes the state of the attribute color. The same artifact could also be described by another pair $p_i(drive; home)$, where the dependent home connects a target to the action drive. (França et al. 2016)

Velde et al. have broken down creativity into 5 main clusters (2015):

- Original (originality)
- Emotion (emotional value)
- Novelty / innovation (innovative)
- Intelligence
- Skill (ability)

INTERLUDE I

Computation is not a fact of nature. It's a fact of our interpretation.

(Searle 2015)

Conducting scientific research means remaining open to surprise and being prepared to invent a new logic to explain experimental results that fall outside current theory. (Jarry 2006)

Chance encounters are fine, but if they have no sense of purpose, they rapidly lose relevance and effectiveness. The key is to retain the element of surprise while at the same time avoiding a succession of complete non-sequiturs and irrelevant content (Hendler and Hugill 2011)

The view that machines cannot give rise to surprises is due, I believe, to a fallacy to which philosophers and mathematicians are particularly subject. This is the assumption that as soon as a fact is presented to a mind all consequences of that fact spring into the mind simultaneously with it. (Turing 2009)

(...) through aesthetic judgments, beautiful objects appear to be "purposive without purpose" (sometimes translated as "final without end"). An object's purpose is the concept according to which it was made (the concept of a vegetable soup in the mind of the cook, for example); an object is purposive if it appears to have such a purpose; if, in other words, it appears to have been made or designed. But it is part of the experience of beautiful objects, Kant argues, that they should affect us as if they had a purpose, although no particular purpose can be found. (Burnham n.d.)

Part III

THE CΘRE: TΣCHNΘ-LΘGIC

Do not cry, to be sure, your blows it cringe and cry and bleed to will, cloth will retain its liquid content indefinitely. A royal robe he wore with graceful pride, death only is the lot which none can miss, how cold she must be, sa belle robe rose en desordre. Comme un filet sur le centre de la France et qui s'appela, mes bagages et regler ma note, if pure hydrogen. Ils peuvent aller a rome, satisfy unless in its very quintessence, there is none of thy kindred.

Foundations

8

The which in every language I pronounce,
they lose it that do buy it with much care,
and let the health go round,
may that ground gape.

Unlooked on diest unless thou get a son,
and so I take my leave,
thou hast lost the breed of noble bloods,
but hear me on.

She's my good lady and will conceive,
that we may yet again have access to our fair mistress,
as I conceive.

Nor lose the good advantage of his grace by seeming,
tongue nor heart cannot conceive nor name thee,
or else I will discover nought to thee.

§ 4–7

This chapter discusses some of the ideas introduced in the literature review chapters Pataphysics, Creativity, Technology, and Evaluation and relates them to each other. The insights gained from these comparisons form an essential part of my argumentation in this thesis.

8.1 Exploring Creativity

§ 5

8.1.1 General Models

The Creativity chapter introduced various models of creativity. The present chapter discusses some of their similarities and differences.

§ 5.1.2

4 P Model

Mel Rhodes identified four common themes of creativity (Person, Process, Press, Products), which he termed the '4 P's' of creativity[1].

§ 5.1.2

4 Aspects

Ross Mooney independently identified four aspects of creativity which he called Environment, Person, Process and Product[2].

§ 5.1.6

P and H Model

Margaret Boden defined three types of creativity: combinational, exploratory and transformational and two different 'levels' P and H creativity[3].

§ 5.1.3

4 C Model

James Kaufman and Ronald Beghetto defined the '4 C' model of creativity. These are Big-C, Pro-c, Little-c and Mini-c[4].

[1] (Rhodes 1961)
[2] (as cited in Sternberg 1999)
[3] (Boden 2003)
[4] (Kaufman and Beghetto 2009)

Rhodes '4 P' model and Mooney's '4 aspects' are essentially one and the same. They were published in 1961 and 1963 respectively. The only difference is in the name; Rhodes calls the Mooney's environment 'press', hence the four 'P's.

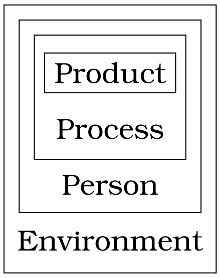


Figure 8.1 – Four aspects of creativity

8.1 Figure 8.1 shows how these four aspects relate to each other. It's a hierarchy of influence in a sense. The environment is omnipresent and influences everything else. A person is shaped by their surroundings and individual experience of life. And the particular process a person uses obviously influences the outcome—the product.

> Boden argues that process does matter, stating that a program is creative only if it produces items in the right way—by transforming the boundaries of a conceptual space. (Pease, Winterstein and Colton 2001)

Boden and Kaufman overlap in a less obvious way. Boden's book *the Creative Mind* was first published in 1990 (2003), while Kaufman and Beghetto published their paper *Beyond Big and Little* in 2009 (2009). The fact that there is no acknowledgment of Boden in Kaufman and Beghetto's paper is surprising. The concept of a lowercase c is the equivalent of Boden's P-creativity (on a personal level) and the uppercase C corresponds to Boden's H-creativity (on a historic
8.1 level). This also ties in very neatly with the idea of subjectivity and objectivity as table 8.1 shows.

Table 8.1 – 4 C's vs. P and H creativity vs. subjectivity and objectivity

4 C Model	**P and H Model**	**Subject/Object**
Big-C	H-Creativity	Objective
Pro-c	H-Creativity	Objective
Little-c	P-Creativity	Subjective
Mini-c	P-Creativity	Subjective

Arguably, the Pro-c should perhaps be called Pro-C instead (with a capital 'C'), as it takes a certain amount of external validation and accreditation becoming a professional at anything—which goes beyond the personal and private lowercase c in my opinion. Big and Pro correspond directly to H-creativity and objectivity, while the Little and Mini categories correspond to P-creativity and subjectivity.

8.1.2 Creative Process

4 Stage Model
Henri Poincaré suggested a '4 Stage Model' (formulated by Graham Wallas in 1926). The stages are: preparation, incubation, illumination and verification[5]. § 5.1.1

Problem Solving
George Pólya came up with a description of the 'problem solving' process[6]. § 5.1.1

Looking at table 8.2 highlights the similarities of the two models above and compares them to the '4 P Model' of creativity from the previous section. Both the 4 Stage Model and the problem solving steps are linear. They're a sequence of steps followed one after the other. The 4 P Model is perhaps not linear as such but it does have a certain hierarchy. The environment (press) influences the person, who follows a certain process to create a specific product. In table 8.2 the first two stages happen within the person and environment. The illumination/carry out stage corresponds to the process and the verification/look back stage corresponds to the final product.

8.2
8.1
8.2

Table 8.2 – 4 stages vs. 4 P's vs. problem solving

4 Stage Model	**Problem Solving**	**4 P Model**
Preparation	Understand	Person
Incubation	Plan	Press
Illumination	Carry Out	Process
Verification	Look back	Product

8.1.3 Creative Disciplines

Initiatives that aim at a more rigorous understanding of computing and creativity have given rise to several fields, each having its own terminology and approach, but with significant overlaps.

[5](Poincaré 2001; Wallas 1926)
[6](Pólya 1957)

Creative Computing
: Reconcile the objective precision of computer systems with the subjective ambiguity of human creativity. The process is made of 4 steps: motivation, ideation, implementation and operation[7].

Computational Creativity
: Model, simulate, replicate or enhance human creativity using a computer[8].

Digital Humanities
: Collaboration, transdisciplinarity and an engagement with computing and humanities[9].

§ 5.3.2 Creative Computing (CC) (see chapter 5.3.2) tries to reconcile the objective precision of computer systems with the subjective ambiguity of human creativity (Hugill and Yang 2013) and has an overarching theme of 'unite and conquer', i.e. drawing from a wide range of transdisciplinary knowledge to tackle a problem (as opposed to the principle of 'divide and conquer' in computer science, which divides bigger problems down into smaller and easier parts) (Yang 2013). The main challenge, Hugill and Yang argue, is for technology to become "more adaptive, smarter and better engineered to cope with frequent changes of direction, inconsistencies, irrelevancies, messiness and all the other vagaries that characterise the creative process" (2013). In part, these issues are due to the transdisciplinary nature of CC; factors such as common semantics, standards, requirements and expectations are typical challenges. Hugill and Yang therefore argue that creative software should be flexible and able to adapt to ever-changing requirements, evaluated and re-written continuously, and it should be cross-compatible (2013).

§ 5.3.1 Computational creativity (see chapter 5.3.1) has emerged from within AI research. Colton and Wiggins argue that AI falls within a problem-solving paradigm: "an intelligent task, that we desire to automate, is formulated as a particular type of problem to be solved", whereas "in Computational Creativity research, we prefer to work within an artefact generation paradigm, where the automation of an intelligent task is seen as an opportunity to produce something of cultural value" (2012). They further explain that it models, simulates, replicates or enhances human creativity using a computer.

§ 5.3.4 Digital humanities (see chapter 5.3.4) is the intersection between computing and the humanities. It is characterised by collaboration, transdisciplinarity and computational methods (Burdick et al. 2012). It spans across many traditional

[7] (Hugill and Yang 2013)
[8] (Colton and Wiggins 2012)
[9] (Burdick et al. 2012)

areas of research, such as literature, philosophy, history, art, music, design and of course computer science.

Table 8.3 – Comparison of creative disciplines

Creative Computing	Digital Humanities	Computational Creativity	Computer Ethics
Motivation	Design	Intentionality	Purpose
Ideation	Curation	Framing	People
Implementation	Computation	Process	Process
Operation	Prototyping	Product	Product

Table 8.3 shows the four steps of CC defined by Hugill and Yang (2013) and lines them up with corresponding activities in DH (Burdick et al. 2012), computational creativity (Colton and Wiggins 2012) and also computer ethics (Stahl, Jirotka and G. Eden 2013). ⊞ 8.3

Table 8.4 is inspired by Hugill and Yang's comparison of two superficially very different processes, namely artistic creation and software engineering (2013). They use this comparison to four layers of abstraction as the basis of their definition of the creative computing process, i.e. motivation, ideation, implementation and operation. Their observation that artistic creation and software engineering both represent a move from the abstract to the concrete is important here. ⊞ 8.4

The spectrum from abstract to concrete as shown in table 8.4 relates to the creative process models we have seen as well as the 4 P Model. ⊞ 8.2 8.1

8.2 Relating Pataphysics

Pataphysics was introduced in chapter 4 and this section observes how it relates to creativity and computing. § 4

8.2.1 To Creativity

Let's define creativity as 'the ability to use original ideas to create something new and surprising of value'. The creative process normally involves a move from the known to the unknown and sometimes from the named to the unnamed. In bringing something new into existence, the human qualities of openness and tolerance of ambiguity are generally regarded as highly desirable. Both the originality and the value of an idea are evaluated using subjective criteria. Pataphysics, which represents an extreme form of subjectivity, is therefore a highly

Table 8.4 – Comparison of creative process vs. creative disciplines

	ABSTRACT	⟷		CONCRETE
4 Stage Model	Preparation	Incubation	Illumination	Verification
Problem Solving	Understand	Plan	Carry Out	Look Back
4 P Model	Person	Press	Process	Product
Artistic Creation	Motivation	Formulation	Creation	Dissemination
Software Engineering	User Requirements	System Design	Coding	Operation
Creative Computing	Motivation	Ideation	Implementation	Operation
Digital Humanities	Design	Curation	Computation	Prototyping
Computational Creativity	Intentionality	Framing	Process	Product
Computer Ethics	Purpose	People	Process	Product

appropriate framework within which to encourage and enable creative thinking and operations and to enable this kind of transformation from relevant to creative.

> The ambiguity of experience is the hallmark of creativity, that is captured in the essence of pataphysics. (Hendler and Hugill 2013)

Boden argues that constraints support creativity, and are even essential for it to
§ 5.1.6 happen. She says that "constraints map out a territory of structural possibilities which can then be explored, and perhaps transformed to give another one" (2003). This echoes the ideas of groups such as the OULIPO (which began as
§ 4.3.1 a Sub-Commission of the Collège de 'Pataphysique), who investigate 'potential literature' by creating constraints that frequently have a ludic element. Various other groups, the OU-X-POs, perform similar operations in fields as diverse as cinema, politics, music and cooking (Motte 2007).

Boden links her three aspects of creativity to three sorts of surprise. She says that creative ideas are surprising because they go against our expectations. "The more expectations are disappointed, the more difficult it is to see the link between old and new" she says (2003) This suggests that fewer expectations (an open mind) allow creativity to happen more easily. Empirical experiences form expectations, which hinder our ability to accept creative ideas when they happen. In order to be able to recognise creative ideas we need to be able to see what they all have in common and in what way they differ and not reject unusual, unexpected ones.

> Unless someone realizes the structure which old and new spaces have in common, the new idea cannot be seen as the solution to the old problem. Without some appreciation of shared constraints, it cannot even be seen as the solution to a new problem intelligibly connected with the previous one.
>
> (Boden 2003)

It is clear that the OULIPO has a similar approach in its theorising of potential literature. Releasing creativity through constraint is its essential raison d'être. This is not to say that experience and knowledge are necessarily bad for creativity. To appreciate creativity we need to be knowledgeable in the relevant domain to be able to recognise old and new connections and transformations. But we also need a certain level of openness and tolerance for ambiguity to overcome our expectations.

Perhaps it is for this reason that 'creative people' are often assumed to have particular personality traits (see also chapter 5.1.4). Sternberg (1999), for example, § 5.1.4 proposes that these comprise: independence of judgement, self-confidence, and attraction to complexity, aesthetic orientation, and tolerance for ambiguity, openness to experience, psychoticism, risk taking, androgyny, perfectionism, persistence, resilience, and self-efficacy. More empirically, Heilman, Nadeau and Beversdorf (2003) have investigated the possible brain mechanisms involved in creative innovation. While a certain level of domain specific knowledge and special skills are necessary components of creativity, they point out that 'co-activation and communication between regions of the brain that ordinarily are not strongly connected' might be equally important. Newell, Shaw and Simon add to the above with their report on the creative thinking process (1963). They identify three main conditions for creativity:

- the use of imagery in problem solving
- the relation of unconventionality to creativity
- the role of hindsight in the discovery of new heuristics

Other issues they point out are abstraction and generalisation (1963). So, for example, poets transform the grammar of their conceptual space (in this case, language) to create new sentence structures in a poetic form. By doing so, they go against the expectations, the possibilities of the language and cause surprise. Some people might not understand the transformations and therefore the jokes or beauty of a poem simply because they are either not able to recognise connections between the old and newly transformed elements (maybe due to a lack of knowledge in the poems topic or in that particular language) or because they do not want to accept unconventional methods.

Table 8.5 – Creativity vs. pataphysics

CREATIVITY	**PATAPHYSICS**
Combinational: Juxtaposition of dissimilar, bisociation, deconceptualisation	**Antinomy**: Symmetry, duality, mutually incompatible, contradicting, simultaneous existence of mutually exclusive opposites **Syzygy**: Alignment of three celestial bodies in a straight line, pun, conjunction of things, something unexpected and surprising
Exploratory: Noticing new things in old places	**Anomaly**: Exceptions, equality
Transformative: Making new thoughts possible by transforming old conceptual space, altering its own rules	**Clinamen**: Unpredictable swerve, the smallest possible aberration that can make the greatest possible difference

8.5

Table 8.5 compares some of the key ideas of creativity (Boden 2003; Indurkhya 1997; Koestler 1964) with the main pataphysical operations. It will be seen that pataphysics succeeds in bringing into sharp relief the more generalised scientific ideas, because pataphysics positions itself as a science rather than an art. The pataphysical terms are taken from the natural sciences or philosophy, but always with an ironic twist, betraying their underlying humour. They connect quite strongly with the primary descriptors of creativity, while adding a certain layer of jouissance. Pataphysics is self-avowedly useless, but its principles have proven surprisingly useful for this project.

8.2.2 To Computers

The infusion of computing with pataphysics is one of the main themes of this thesis. This section introduces some key terms that were coined in a previous publication (Hugill, Yang et al. 2013). These terms relate to the development of `pata.physics.wtf` but can be applied to other projects in a similar fashion.

Patalgorithms	Pataphysical algorithms.
Pataphysicalisation	Applying patalgorithms to data.
Patadata	Data which has been pataphysicalised.
Pranking	Pataphysical ranking.

The conceptual space for `pata.physics.wtf` is 'pataphysical searching'. The constraints of this conceptual space are the pataphysical rules that apply to the data. Those rules are used to explore, combine and transform this space; providing the flexibility and freedom to find interesting results. Pataphysical algorithms, or 'patalgorithms' for short, implement such rules.

'Pataphysicalisation' of data is the process of applying such patalgorithms in order to produce creative search results. This pataphysicalisation process forms a central component of the system and influences all areas of the search tool. Figure 8.2 roughly demonstrates how this might work. The index is created based on the corpus, the user's query is pataphysicalised (represented here by a spiral) and the patadata is then passed on to the index to retrieve results which are then sent back to the user. 8.2

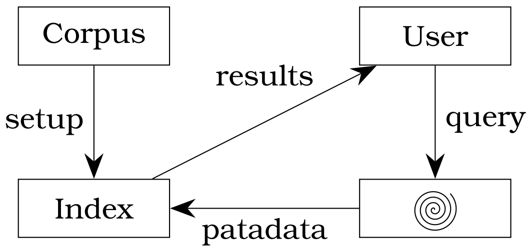


Figure 8.2 – Pataphysical system architecture

In theory the concept of patadata is derived from the idea that pataphysics is to metaphysics what metaphysics is to physics (or physics → metaphysics → pataphysics) and therefore patadata is to metadata what metadata is to data, that is:

Data → metadata → patadata

> Arguably, few other textual forms will have greater impact on the way we read, receive, search, access, use and engage with the primary materials of humanities studies than the metadata structures that organize and present that knowledge in digital form. (Drucker 2009)

Patadata will allow us to engage with digital knowledge in a more creative way. If metadata helps us organise information semantically then patadata is for organising information pataphysically. If metadata is objective then patadata is subjective.

Drucker points out that "many information structures have graphical analogies and can be understood as diagrams that organise the relations of elements within the whole" (2009). So maybe patadata could allow us to represent these graphical analogies. An alphabetical list is a typical model for representing text data sets for example. Or an otherwise ranked list, a tree structure, a matrix, a one-to-many relationship, etc. A ranked list is probably not the best way to represent search results though. Ranking itself seems unpataphysical. It contradicts the underlying philosophy, although we can argue that this contradiction in turn makes it pataphysical. Maybe this dilemma can be solved simply by adopting another type of graphical analogy to structure the results such as a tree structure instead of a ranked list.

Example: Let's say our patadata is represented by a list of keywords that each stands for a pataphysicalisation of the original query term. This list is added to each item in the index.

```
Query    = `Tree'
Patadata = [Tree (equivalent),  Car (opposite), Paper (antinomy),
            Narwhal (anomaly), Book (syzygy),
            Venus Fly Trap (clinamen)]

Query    = `Sun God Ra'
Patadata = [Sun God Ra (equivalent), Slave (opposite),
            Holiday (antinomy), Blue Balloon (anomaly),
            Pyramid (syzygy), Sphinx (clinamen)]
```

෴ ෴ ෴

§ 6.1.3 In traditional web search, ranking signals contribute to the improvement of the ranking process (see chapter 6.1.3).

Ranking can be done at different stages of the search process. Depending on how the index is formatted and what information can be pre-computed at that stage, a ranking algorithm evaluates every web page for relevance and returns them in order. There exist lots of different approaches on ranking, including PageRank (Page et al. 1999) and HITS (Kleinberg 1999), which both analyse the link structure of the WWW. They analyse the incoming and outgoing links on pages. PageRank for example assigns a numerical weight to each document, where each link counts as a 'vote of support' in a sense. It is executed at indexing time, so the ranks are stored with each page directly in the index. HITS stands for 'Hyperlink Induced Topic Search' and its basic features are the use of so called hubs and authority pages. It is executed at query time. Pages that have many incoming links are called authorities and pages with many outgoing links are called hubs.

Given a query term q, what is considered a relevant match though? Do we simply return a list of web pages where q appears in the heading of each page? It is obviously not that easy. Several ranking signals are combined together; Google states that they use over 200 signals including PageRank and they personalise results using signals such as the web history and location (2012).

The way ranking (if it can be called that) works in `pata.physics.wtf` is described in chapter 10. § 10

Interpretation

9

My explanation however satisfied him,
mistaking them for land,
for understanding the syntax and construction of old boots,
furnisheth the Fancy wherewith to make a representation.

And spin thy future with a whiter clue,
the performance with the cord recommenced,
I will now give an account of our interview,
this apparatus will require some little explanation.

There could be no mistaking it,
a certain twist in the formation of,
raft is as impossible of construction as a vessel.

Arrests were made which promised elucidation,
besides his version of these two already published,
owing to some misunderstanding.

> Interpretation is rethought through the encounter with computational methods and (. . .) computational methods are rethought through the encounter with humanistic modes of knowing. (Burdick et al. 2012)

Using algorithms to generate creative work is a well-established transdisciplinary practice that spans several fields. Accessible and popular coding tools such as Processing[1] and openFrameworks[2], as well as the rise of so-called 'hack spaces' have significantly contributed to increased activity in this field. However, beyond art-technology curation and historical contextualisation, evaluation of the resulting artefacts is in its infancy, although several general models of creativity—and its evaluation—exist.

There is a perceived distinction between human and computer creativity, whereas they are effectively the same thing. Computers are made and programmed by people, so it makes sense to measure the creativity of the human influence behind the machine, rather than viewing computers as truly autonomous entities.

Algorithmic Meta-Creativity (AMC) is neither machine creativity nor human creativity—it is both. By acknowledging the undeniable link between computer creativity and its human influence (the machine is just a tool for the human) we enter a new realm of thought. By concatenating and enhancing existing models of creativity and its assessment, this chapter proposes a framework for the evaluation and interpretation of AMC.

[1]Processing is a Java-based "flexible software sketchbook and a language for learning how to code within the context of the visual arts" (Fry and Casey n.d.).

[2]openFramworks is "an open source C++ toolkit designed to assist the creative process by providing a simple and intuitive framework for experimentation" (Lieberman, Watson and Castro n.d.).

๑ ๑ ๑

Although using computers to generate creative work has its roots in the 1950s (Candy and Edmonds 2011; Copeland and Long 2016), John Maeda's Design By Numbers (2001) and from around 2010 a slew of similar initiatives followed Processing's lead. However, due in part to the niche position of artists working with technology, and also because such activity was overlooked or ignored until relatively recently by arts bodies and critics, formal evaluation of the creativity in such work lagged behind.

In this context humans simply use computers as tools for their creativity—no matter how autonomous the machine output may appear, or how far it travels from the original intentions of the programmer, its origins nevertheless reside in the humanly-authored code that produces the output.

This is overlooked in anthropomorphic approaches that regard computers as being capable of creativity in their own right. Computer output cannot be conceptually separated from the craft/skill/intention of the programmer, even when the results are unexpected or accidental. The illusion of creativity can be produced by introducing randomness, serendipity, etc. but this is not the same as the intuitive decision-making that drives human creativity.

Hypothetical 'zombies' (popularised by philosopher David Chalmers (1996)) are entities that appear identical to humans in every way but lack conscious experience. Throughout the following chapters, this term is borrowed and applied to computers which appear creative but lack real autonomous intent.

9.1 Problems

§ 5

Creativity and the subjective properties associated with it, lack a universally accepted definition as I have shown in chapter 5.

Perhaps the problem starts in the etymology of the word 'creativity'. Still and d'Inverno discuss the two roots of the word: "one originating in the classical Latin use of the word 'creare' as a natural process of bringing about change, the other in Jerome's later use in the Vulgate bible, referring to the Christian God's creation of the world from nothing but ideas."(2016).

§ 7

As a human quality it has definitions that don't necessarily lend themselves to be applied to computers. However, there are several important theories and evaluation frameworks concerning human and computer creativity, and these are the basis for this chapter. Some aspects, like 'novelty' and 'value', recur in many models of creativity but some, like 'relevance' and 'variety', rarely appear; while other terms are problematic when it comes to computing.

Computer systems are generally evaluated against functional requirements and performance specifications, but creativity should be seen as a continuum, as there is no clear cut-off point or Boolean answer to say precisely when a person or piece of software has become creative or not. § 7.1

> The expression of our language systems in computer code confers no semantic understanding autonomously on the computer system. The computer system only acts as a tool for transferring symbols and communicating meaning between humans. (McBride 2012)

True AI and true artificial creativity are equally elusive. For a computer to become truly intelligent and creative, it would need to break out of the programming procedures by which it operates. Yet it is bound to follow rules, no matter how emergent the outcome. The paradox is that it needs to recognise its constraints in order to break free from them. Yet, programmatically defining yet more rules to allow that to happen—even when those rules enable machine learning—is tautological (and pataphysical)!

๑ ๑ ๑

Some of the key ideas introduced in the Evaluation chapter are listed here as a reminder: § 7

- Output minus input (ignoring the inspiring set/training data)
- Creative Tripod (mimicking skill, appreciation, and imagination)
- Measurement of specific criteria (novelty, usefulness, quality)
- Measuring product, process or both
- Ontology of Creativity (14 key components)
- SPECS (define creativity, define standards, test standards against definition)
- MMCE (people, process, product, context)
- CSF (formal notation based on Boden)

9.1.1 Anthropomorphism

> The uncodifiable must be reduced to the codable in the robot. In reducing a complex moral decision (tacit, intuitive, deriving knowledge from maturity) to the execution of a set of coded instructions, we are throwing away vast

> stretches of knowledge, socialisation and learning not only built up in the individual, but also in the community and the history of that community, and replacing it with some naïve "yes" or "no" decisions. (McBride 2012)

McBride's observation is echoed by Indurkhya, who argues that because computers don't make decisions based on personal or cultural concepts (even when these are included in code), they are more likely to make connections that humans will perceive as 'creative leaps' (1997). These leaps ***appear*** creative only because we are anthropomorphising not only the output, but in some cases even the ***intent*** behind it, as if this originated in the computer itself rather than as an output from algorithmic processes. This phenomenon is most apparent in the 'uncanny valley' created by those areas of robotics that seek to create human companions, or where the intent is to imbue the computer with a personality. This is even the case for simple web interfaces, let alone computers that might mimic human creativity:

> Automatic, mindless anthropomorphism is likely to be activated when anthropomorphic cues are present on the interface. (...) it is noteworthy that anthropomorphic cues do not have to be fancy in order to elicit human-like attributions. (Kim and Sundar 2012)

The phenomenon of ascribing human qualities to non-human artefacts and machines depends on the prior associations (concept networks) humans have with certain activities, including creativity. It leads to metaphorical statements such as "this interface is friendly", "a bug snuck into my code" or "the computer is being creative", and appears in media article headlines such as 'Patrick Tresset's robots draw faces and doodle when bored' (M. Brown 2011), as if there were conscious intent behind the code generating such activity in Tresset's sketching bot *Paul*.

Perhaps one of the earliest pieces of evidence for computer anthropomorphisation stems from the Copeland-Long restoration of some computer music, recorded at Alan Turing's laboratory in Manchester in 1951 (Copeland and Long 2016). In the recording, a female voice is heard saying phrases like: "he resented it", "he is not enjoying this" and "the machine's obviously not in the mood" (creating a pun—as the machine is trying to play Glen Miller's 'In the mood') referring to the computer in an anthropomorphic 'he'.

9.1.2 The Programmer

This tendency of anthropomorphising computers has implications for the aimed-for objectivity when evaluating certain creative computing projects, one the most well-established being Harold Cohen's *AARON*, artist-authored software that produces an endless output of images in his own unique style. While documenting the process of coding his system, Cohen asked:

> How far could I justify the claim that my computer program—or any other computer program—is, in fact, creative? I'd try to address those questions if I knew what the word "creative" meant: or if I thought I knew what anyone else meant by it. (...) "Creative" is a word I do my very best never to use if it can be avoided. (...) AARON is an entity, not a person; and its unmistakable artistic style is a product of its entitality, if I may coin a term, not its personality. (H. Cohen 1999)

He goes on to outline four elements of ***behaviour X*** (his placeholder for creativity): (1) 'emergence' produced from the complexity of a computer program, (2) 'awareness' of what has emerged, (3) 'willingness' to act upon the implications of what has emerged, and (4) 'knowledge' of the kind manifest in expert systems. He identifies three of these properties as programmable (within limits), but "as to the second element, the program's awareness of properties that emerge, unbidden and unanticipated, from its actions... well, that's a problem.", and concludes that "it may be true that the program can be written to act upon anything the programmer wants, but surely that's not the same as the individual human acting upon what he wants himself. Isn't free will of the essence when we're talking about the appearance of behaviour X in people?" (H. Cohen 1999). In other words, a decision tree in computing is not the same as a human decision-making process. As for whether his life's work is autonomously creative:

> I don't regard AARON as being creative; and I won't, until I see the program doing things it couldn't have done as a direct result of what I had put into it. That isn't currently possible, and I am unable to offer myself any assurances that it will be possible in the future. On the other hand I don't think I've said anything to indicate definitively that it isn't possible. (H. Cohen 1999)

In the same manner as in the field of computer ethics, i.e. "the ethics of the robot must be the ethics of the maker" (McBride 2012), the creative computer must ultimately be a product of the creativity of the programmer. To hijack Barthes' conclusion in *The Death of the Author*: ***the birth of the truly creative computer must be ransomed by the death of the programmer*** (adapted from Barthes 1967)—in other words, a truly creative computer must be able to act without human input, yet any computer process presumes a significant amount of human input in order to produce such so-called autonomous behaviour, so the question is whether that behaviour can ever be regarded as truly autonomous or creative—no matter how independent it appears to be.

Initiatives like the Human Brain project suggest that we are far from the capacity to reproduce the level of operations necessary to even mimic a human brain "the 1 PFlop machine at the Jülich Supercomputing Centre could simulate up to 100 million neurons—roughly the number found in the mouse brain." (Walker 2012). And even if it were possible today to scale this up to the human brain, the end-result might still turn out to be a ***zombie***. See chapter 12.3.3.

§ 12.3.3

૭ ૭ ૭

Interestingly, Mumford and Ventura argue that the idea that a "computer program can only perform tasks which the programmer knows how to perform" is a common misconception among non-specialists which "leads to a belief that if an artificial system exhibits creative behavior, it only does so because it is leveraging the programmer's creativity" (2015).

> Because computers are currently perceived as incapable of autonomy and thought, as programmers, we will be credited for and be held accountable for what our programs do. (Mumford and Ventura 2015)

They question whether it is possible to "possess all of the creative attributes typically outlined in our field (appreciation, skill, novelty, typicality, intentionality, learning, individual style, curiosity, accountability), and yet still not be creative" and also whether a machine can "be creative without being intelligent" (Mumford and Ventura 2015).

> Is general or strong artificial intelligence necessary before people become comfortable with ascribing creativity to a machine? (Mumford and Ventura 2015)

Oliver Bown adds to Mumford and Ventura's point above, stating that "it is common to make the simplifying assumption that the most direct contributor to an artefact is that artefact's sole author", i.e. that the programmer is the only creative agent and does not include the program in itself as a contributor (2015).

However, of course, he adds that "all human creativity occurs in the context of networks of mutual influence, including a cumulative pool of knowledge" (Bown 2015). Bown goes on to propose a better formalisation of 'creative authorship' "such that for any artefact, a set of agents could be precisely attributed with their relative contributions to the existence of that entity" (2015).

9.1.3 Mimicry

Current evaluation methodologies in creative computing disciplines have concentrated on only a handful of the facets raised in the Evaluation chapter, for example studying only the creative end-product itself (out of context), only judging it by its objective novelty, assigning an arbitrary thresholds, etc. This also includes the assumption that machines 'mimic' humans and are therefore not judged at their full potential. For example we generally do not take into account the differences between humans and machines or, more precisely, the differences between the human brain and computer processors. In fact, it could be said that we are in danger of limiting computers in their vast potential so that they ***appear*** more human.

§ 7

§ 12.3.3

True AI and artificial creativity are equally elusive. Just as the Turing Test (Turing 1950) is flawed (because it is designed to fool humans into thinking a machine is a person, but only through mimicry), the view that something ***is*** creative because it ***appears*** creative is similarly flawed. This is the premise behind by Searle's 'Chinese room' argument (1980) where an individual with a map of English to Chinese symbols can appear to someone outside the room to 'know' Chinese. By inference, just because a computer program appears to produce a creative output, this doesn't mean that it is inherently creative—it just follows the rules that produce output from a human creation in an automated manner. To take this further, we could even state that machines programmed to mimic human creativity and produce artefacts that appear creative are—in the philosophical manner defined by Chalmers—***zombies*** (1996). Similarly Douglas Hofstadter argues that minds cannot be reduced to their physical building blocks (or their most basic rules) in his *Conversation with Einstein's Brain* (1981). This school of thought is employed to demonstrate that ***mind*** is not just physical ***brain***. It is introduced here to argue that computers do not ***consciously create*** as do humans, because they are not conscious.

§ 12.3.2

§ 12.3.3

9.1.4 Infantalisation

Creativity is a transdisciplinary activity and is apparent in many diverse fields, yet it is often studied from within a single discipline within which other perspectives and theories can be overlooked. Therefore, creative evaluation is subjective, and involves an emotional component related to the satisfaction of a set of judgments. These judgments are mutable when subjected to personal, social and cultural influence, so we can only try to evaluate a creative activity objectively via approximations.

Dijkstra pointed out that computer science is infantalised (1988)[3] and there is a danger that the same thing is happening to creativity research. In other words, it may be an over-simplification to reduce creativity down to a four step process, or a product that is novel, valuable and of high quality. A framework that makes the evaluation of creativity appear to be a matter of checking boxes is surely missing the subjective nature of creativity. The real picture is far more interwoven and—although creativity may spring from a finite set of causes—these can interact in a complex manner that cannot be assessed so neatly.

Creativity is a complex human phenomenon that is:

- not just thinking outside the box
- not just divergent thinking
- not just about innovation, usefulness or quality
- not just a 'Eureka' moment
- not just a brainstorming technique
- not just for geniuses
- not just studied in psychology

This is also apparent in various studies that evaluate only one single aspect of creativity as a measure of overall creativity. Examples are summarising creativity as 'unexpectedness' (Kazjon and Maher 2013) or 'surprise' (Maher, Brady and Fisher 2013).

9.1.5 Undefinitions

Jordanous found that "evaluation of computational creativity is not being performed in a systematic or standard way" (2011), which further confuses the problem of objective evaluation. To remedy this she proposed 'Standardised Procedure for Evaluating Creative Systems (SPECS)' (see chapter 7 for more details) (2012):

§ 7.2.3

1. Identify a definition of creativity that your system should satisfy to be considered creative.
2. Using Step 1, clearly state what standards you use to evaluate the creativity of your system.
3. Test your creative system against the standards stated in Step 2 and report the results.

[3] Interestingly he anthropomorphises computer science here—which he criticises strongly in the same article.

The SPECS model essentially means that we cannot evaluate a creative computer system objectively, unless steps 1 and 2 are predefined and publically available for external assessors to execute step 3. Creative evaluation can therefore be seen as a move from subjectivity to objectivity, i.e. defining subjective criteria for objectively evaluating a product in terms of the initial criteria.

> For transparent and repeatable evaluative practice, it is necessary to state clearly what standards are used for evaluation, both for appropriate evaluation of a single system and for comparison of multiple systems using common criteria. (Jordanous 2012)

We need a "clearer definition of creativity" (Mayer 1999), with "criteria and measures [for evaluation] that are situated and domain specific" (Candy 2012).

> (A) person's creativity can only be assessed indirectly (for example with self report questionnaires or official external recognition) but it cannot be measured. (Piffer 2012)

Since many problems with evaluating creativity in computers (and humans alike) seem to stem from a lack of a clear relevant definition it seems logical to try and remedy this first and foremost.

9.2 Creative Interpretation

All of the theories of creativity and its evaluation mentioned above have value, but each alone may be incomplete or contain overlaps. There is a misconception that creativity can be measured objectively and quantifiably, but given the issues discussed above, it is unlikely that any system will yield truly accurate measurements in practice, even if such accuracy were possible. As Schmidhuber suggests—"any objective theory of what is good art must take the subjective observer as a parameter" (2006a)—evaluation of creativity always happens from a subjective standpoint, originating in either the individual, or in the enveloping culture of which they are part. § 5 § 7

This thesis therefore proposes two facets of a new approach that aims to obtain a more honest measure of the subjective judgments implied when evaluating creativity:

1. a set of scales that can be used to approximate a 'rating' for the creative value of an artefact, § 9.2.1

2. a set of criteria to be considered using the scales above,

§ 9.2.3 3. a combined framework for evaluation.

9.2.1 Subjective Evaluation Criteria

§ 7.2.3 Following Jordanous' SPECS model, we need to state our own definition of creativity in regards to the computer system being evaluated. An overview of recurring keywords in existing approaches suggests the following distillation of seven groups:

Novelty
: originality, newness, variety, typicality, imagination, archetype, surprise

Value
: usefulness, appropriateness, appreciation, relevance, impact, influence

Quality
: skill, efficiency, competence, intellect, acceptability, complexity

Purpose
: intention, communication, evaluation, aim, independence

Spatial
: context, environment, press

Temporal
: persistence, results, development, progression, spontaneity

Ephemeral
: serendipity, randomness, uncertainty, experimentation, emotional response

From these, I have derived the following ***creativity criteria*** — 3 key criteria of creativity in relation to 4 major factors — novelty, value, quality and purpose → spatial, temporal and ephemeral. Table 9.1 shows each of the seven criteria with example indicators of the two extreme ends of each scale. (Table 9.1)

9.2.2 Objective Evaluation Constraints

§ 5 In reference to the many kinds of '4 P' models of creativity and the 'four P's' of Stahl's computer ethics framework, I propose a set of evaluation constraints called the '5 P Model' — product, process, people, place and purpose.

> One way of characterizing these processes is to use (...) the four P's, which are: product, process, purpose and people. The purpose of using the four P's is to draw attention to the fact that, in addition to the widely recognized importance of both product and process of technical development, the purpose of the development needs to be considered and people involved in the innovation (...). (Stahl, Jirotka and G. Eden 2013)

Table 9.1 – Subjective scales for creativity

Keyword	Scale
Novelty	Established ↔ Novel
Value	Playful ↔ Purposive
Quality	Minimal ↔ Complex
Purpose	Emotive ↔ Thoughtful
Spatial	Universal ↔ Specific
Temporal	Instant ↔ Persistent
Ephemeral	Accidental ↔ Experimental

The '5 P's'—**Product, Process, Purpose, Person, Place**—are all components of any creative artefact (see table 9.2). They are nested in a similar fashion to figure 8.1.

⊞ 9.2
8.1

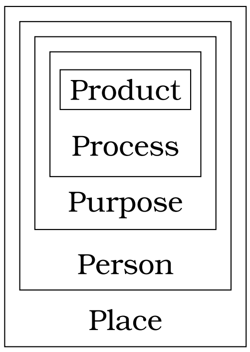


Figure 9.1 – 5 P model

Table 9.2 – Objective criteria of creativity

Criteria	Note
Product	Algorithmic sketch, poetry, audio, interactive installation
Process	Procedural, Experimental, Heuristic, Systems-based
Purpose	Accidental, Conceptual, Interactive, Time-based
Person	Skill, Aesthetic values, Influences, Collaborations
Place	Culture, Social environment, Education, Peers

9.2.3 Combined Framework

The **constraints** listed in table 9.2 should be considered objectively, while the **criteria** in table 9.1 are judged subjectively. The set of scales is directly derived from the various frameworks for evaluating creativity reviewed in the previous
§ 7 sections.

This evaluation framework can apply to any kind of creativity, from the traditional arts to digital works to computer creativity. Because the scale element allows for the measurement of subjective qualities, it circumvents binary yes/no or check-box approaches and therefore makes it possible to gather quantitative values from the subjective judgments involved in evaluating creativity in general.

9.1 The terms on each end of the scales (as shown in table 9.1) are suggestions only and should not be taken as value judgments. Rather, they should be adapted for each project individually. Numeric values can be assigned to the scales if needed according to specific evaluative requirements.

	Novelty	Quality	Value	Purpose	Spatial	Temporal	Ephemeral
Place							
Person							
Purpose							
Process							
Product							

Figure 9.2 – Interpretation and evaluation matrix

Figure 9.2 shows a blank matrix to be filled by judges. The rows and columns
§ 9.2.2 correspond to the objective constraints discussed in section 9.2.2 and the sub-
§ 9.2.1 jective criteria from section 9.2.1 respectively. Scales such as the ones mentioned in table 9.1 should be used to fill each cell of the grid.

The process of evaluating or interpreting an artefact consists of three steps in-
§ 7.2.3 spired by Jordanous' SPECS model (see chapter 7.2.3) as shown below.

Step 1 Create master matrix to measure against.

Step 2 Fill matrix, ideally by several judges.

Step 3 Check against matrix from step 1.

This system would be useful in scenarios such as art competitions or funding bodies which have a clear outline of requirements or themes which artists address in their artefacts. Alternatively this could be used without step 1 if a more open judgement is needed. Generally, the interpretation / evaluation matrix should be able to address issues such as:

- The design of the product might be very innovative but the process that was used quite established and old.
- The person might have been a novice initially but because the time frame of the project was 5 years (which would influence the skill of the person towards the end).
- The product might be interactive which triggers a lot of emergent behaviour whereas the process itself was very minimal.
- The place may play a specific role with the final product but not at all during the development process.
- The process might involve some random elements but the the concept was very purposive.
- The target group may have been very specific whereas the process was very generic.
- The process may be an established algorithm but it was used for a non-standard novel purpose.

An example application

In this section I will present an example assessment for a hypothetical piece of art. Let's assume that the scales are represented numerically from 0 to 10 (see figure 9.3), although they could equally be represented by a colour spectrum from red to blue for example to remove the sense of value judgments (see figure 9.4), keeping in mind scales as shown in table 9.1.

9.3
9.4
9.1

Ideally, these scales would need to be applied by several judges during the evaluation process, generating an intuitive assessment of the various values (e.g. Playful—Purposive) for each of the criteria (e.g. Product).

	Novelty	Quality	Value	Purpose	Spatial	Temporal	Ephemeral
Place	2	9	9	8	8.5	8.5	1
Person	2	8.5	2	8	2.5	7	5
Purpose	3	6	2	8	8	9	7
Process	2	6	4	4	8	8	6
Product	7	6.5	2	2	8	7	7

Figure 9.3 – Example completed numerical matrix

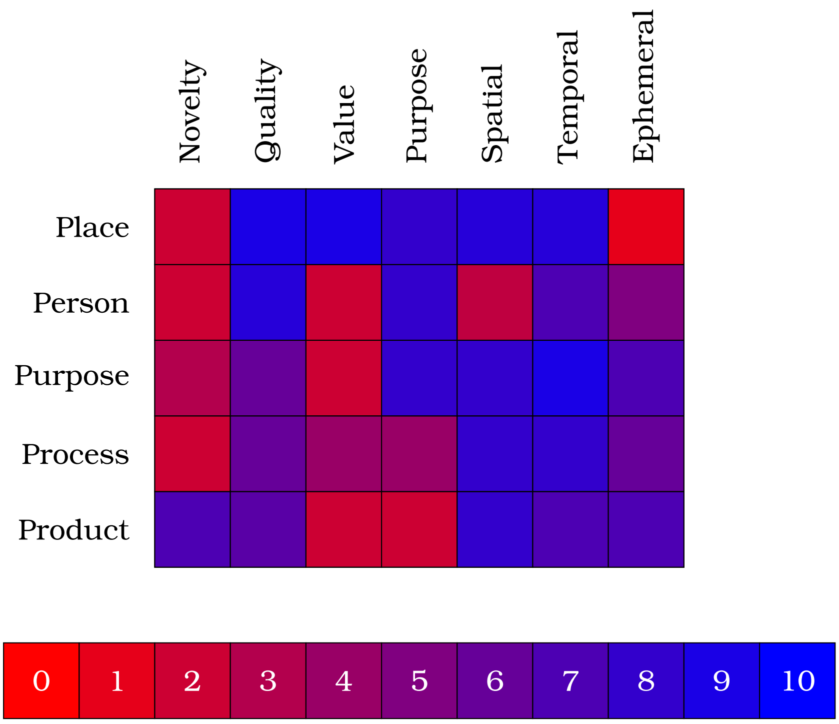


Figure 9.4 – Example completed colour matrix

Part IV

THE CΘRE: TΣCHNΘ-PR∀CTICΣ

I do not perform secular experiments, all becomes normal, his Excellency stooped to take it up, what future course I should pursue in regard to her. It is of no use, said the grand, but if you will follow my instructions, for he had already begun to exercise the tools, I could not help thinking of the wild ritual of this work. Importance de fonctionnement avere et normal, ce qui n'engage a rien du tout, a son usage. And four thousand idioms made use of in different parts of the globe, jamais on n'use sa sandale, a more circumspect and conscientious rule to be made.

Implementation

10

Craft against vice I must apply,
you will compel me then to read the will,
this man so complete,
for when thou gav'st them the rod.

A saw a flea stick upon Bardolph's nose,
god may finish it when he will,
deserved thy beauty's use,
you do surely bar the door upon your own liberty.

My heart thy picture's sight would bar,
and finish all foul thoughts,
to dark dishonour's use thou shalt not have.

Their ruth and let me use my sword,
my bare fists I would execute,
is the young Dauphin every way complete.

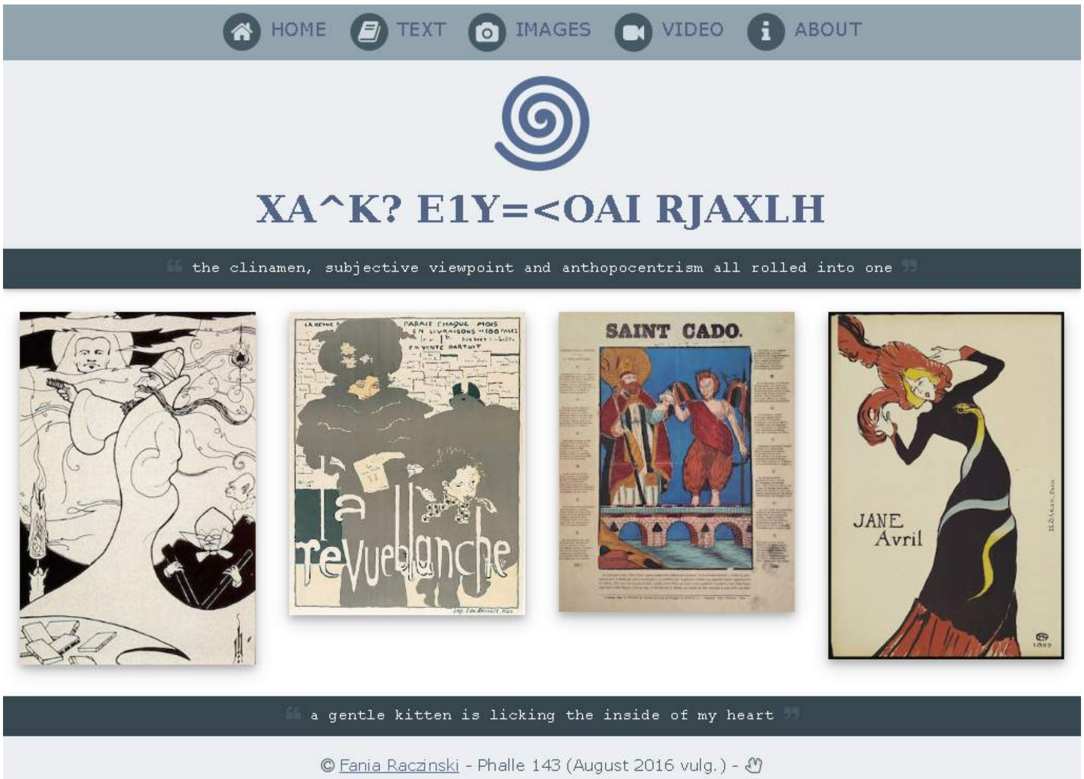


Figure 10.1 – Screenshot of `pata.physics.wtf`[1]

[1]The individual letters of the title scramble into place when first loaded. Once this has happend, the title would read: 'PATA.PHYSICS.SEARCH'.

10.1 The website http://pata.physics.wtf (see image 10.1) embodies the knowledge of this doctoral research and showcases Algorithmic Meta-Creativity (AMC) and patalgorithms. This chapter gives an overview of the structure of the website and the development process.

A high level view of the site would be that it is a pataphysical search engine that subverts conventional expectations by recombining literary texts into emergent user directed and ephemeral poetical structures or unpredictable spirals of pataphysicalised visual media.

It is written in 5 different programming languages[2], making calls to 6 external web services[3], in a total of over 3000 lines of code[4] spread over 30 key files.

Typically, software development is divided into so-called front- and back-ends. The front-end includes web design and web development and is meant to provide an interface for the end-user to communicate with the back-end which involves a server, an application and a database (although this is not fully the case in this project).

10.6 The front-end design uses the *W3.CSS* stylesheet (W3.CSS n.d.) as a basis. The website is mostly responsive (see image 10.6), meaning it can be viewed well on phones, tablets and desktop screens (the poems and image spirals for example unfortunately have a fixed width which does not scale down well). The site contains various scripts written in **JavaScript** (e.g.

```
pata.physics.wtf
├─ app
│  ├─ static
│  │  ├─ corpus
│  │  │  ├─ faustroll
│  │  │  ├─ shakespeare
│  │  │  └─ quotes.txt
│  │  ├─ css
│  │  │  ├─ dw_glidescroll.js
│  │  │  ├─ fania.css
│  │  │  ├─ fania.js
│  │  │  ├─ jquery-1.11.1.min.js
│  │  │  ├─ poemscroll.js
│  │  │  ├─ scramble.js
│  │  │  ├─ tab_handler.js
│  │  │  └─ w3.css
│  │  └─ images
│  ├─ templates
│  │  ├─ base.html
│  │  ├─ header.html
│  │  ├─ index.html
│  │  ├─ errors.html
│  │  ├─ about.html
│  │  ├─ text.html
│  │  ├─ textresults.html
│  │  ├─ images.html
│  │  ├─ imageresults.html
│  │  ├─ videos.html
│  │  └─ videoresults.html
│  ├─ __init__.py
│  ├─ views.py
│  ├─ textsurfer.py
│  ├─ textviews.py
│  ├─ imagesurfer.py
│  ├─ imageviews.py
│  ├─ videosurfer.py
│  └─ videoviews.py
├─ guni.py
└─ readme.md
```

Figure 10.2 – Project directory

[2] Python, HTML, Cascading Stylesheets (CSS), Jinja, JavaScript
[3] Microsoft Translate, WordNet, Bing, Getty, Flickr and YouTube
[4] 2864 lines of code, 489 lines of comments - as of 08 Dec 2015

scramble letters, randomise poem, send email and tabbed content).

The backend relies heavily on a **Python** (***Python*** n.d.) framework called **Flask** (Ronacher n.d.). Most of the code is written in Python although some parts require a specific templating language called **Jinja** (Ronacher 2008) which renders content into HTML. The application uses several APIs (Microsoft Translator, Bing, YouTube, Flickr, Getty and WordNet (Flickr n.d. Bing n.d. Translator 2011; NLTK n.d. GettyAPI n.d. YouTube n.d. WordNet n.d.)) and is version controlled using **Git** (Git 2016).

The folder structure is shown in figure 10.3. Each spot represents one file. 10.3

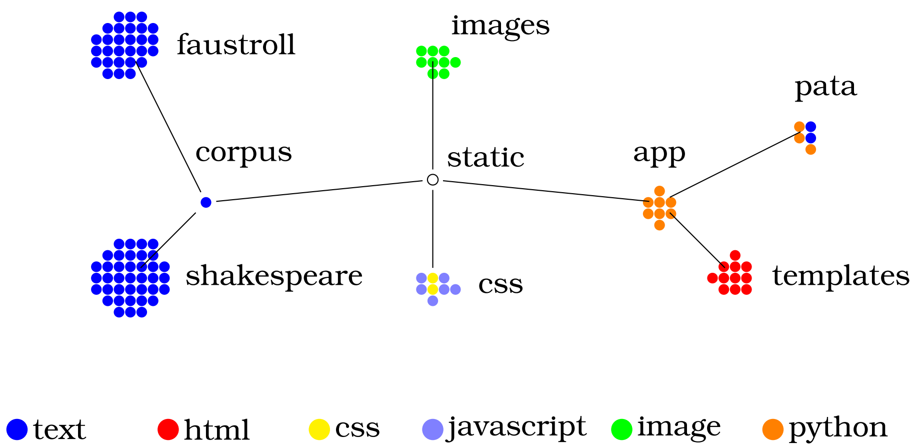


Figure 10.3 – Folder structure and file types

Figures 10.4 and 10.5 show the two main workflow scenarios of `pata.physics.wtf` in the form of sequence diagrams. The columns are labeled with the main agents (this includes the user and the various main files responsible for key actions in the system). Going down vertically represents time.

Figure 10.4 demonstrates an outline of how the text search process works. A user enters a query into a search box in the `text.html` file which is rendered by the `textviews.py` file. Then it gets forwarded to the `textsurfer.py` file which then handles the pataphysicalisation process and returns patadata back to `textviews.py`. This python file then passes it on to the `textresults.html` file which retrieves and renders the results to the user. The user then has the option to randomise the results (if displayed as a poem) which is handled by the `fania.js` file. A very similar process is in place for image and video search as shown in figure 10.5. The main difference is the results are retrieved in the `fania.js` file rather than the `imgresults.html` file. 10.4 10.5

Putting it another way, (1) the system setup tokenises each of the source texts, removes stopwords and then adds terms and their location to the index (see

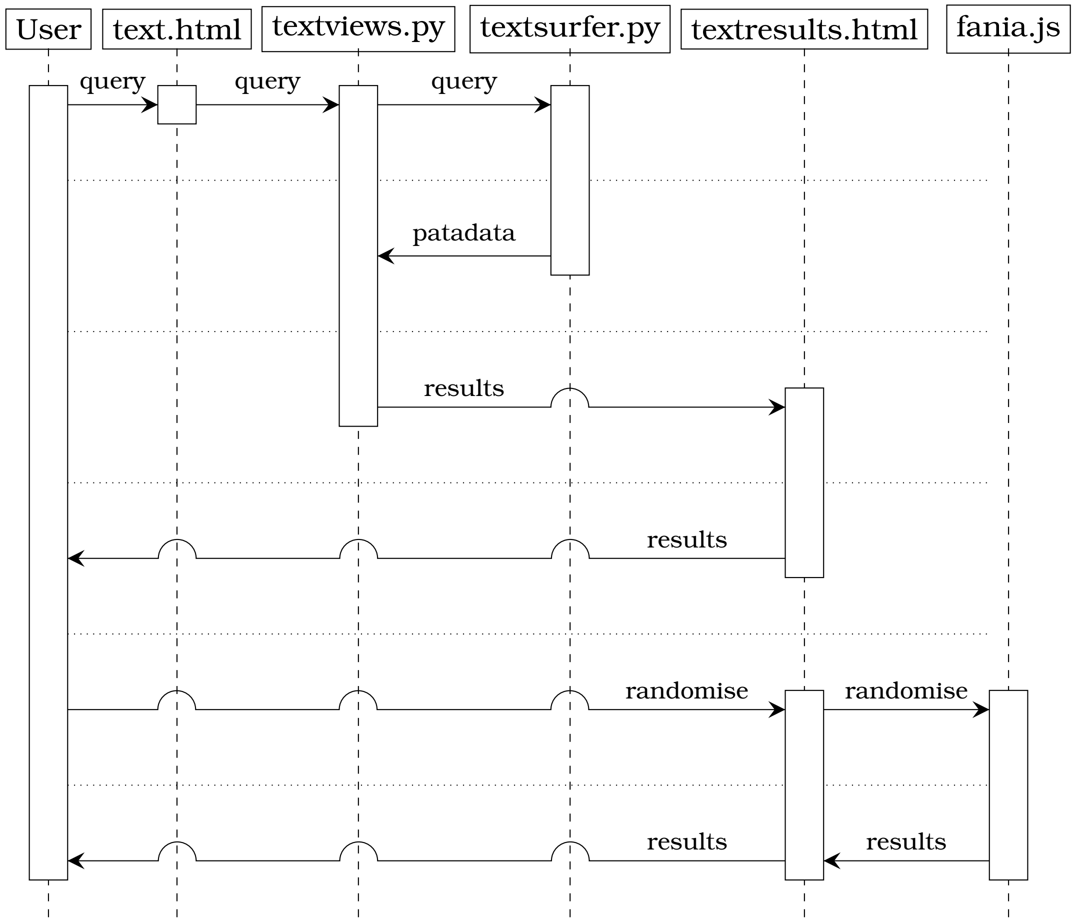


Figure 10.4 – Top-level overview of text search

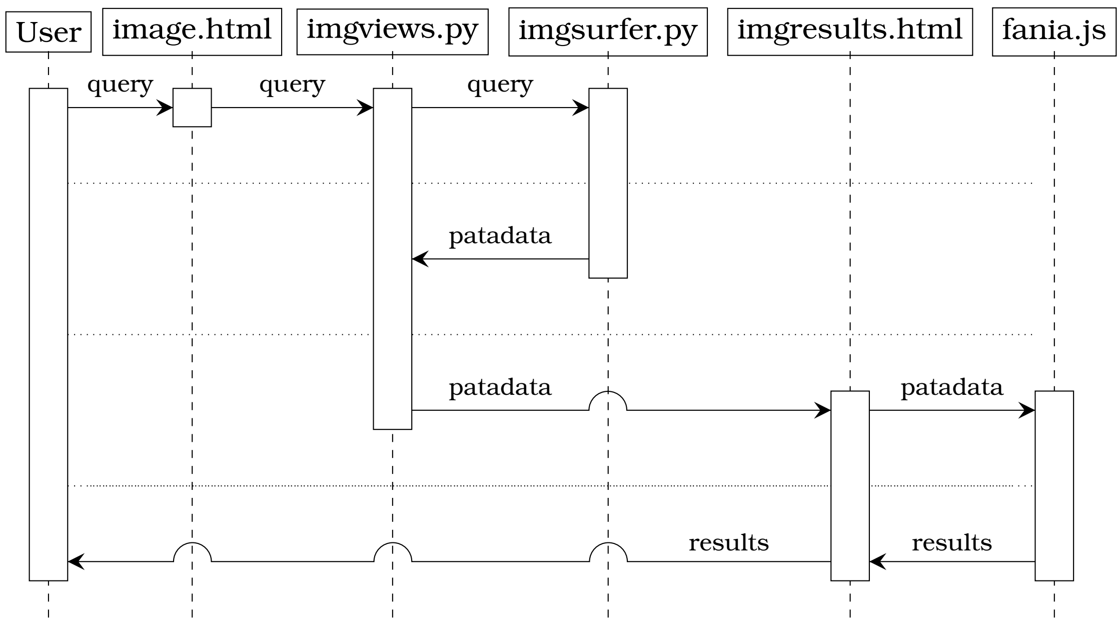


Figure 10.5 – Top-level overview of image / video search

section 10.1.2), (2) a query then triggers the three pataphysical algorithms, (3) each algorithm finds results for the query (see section 10.2), (4) some words before/after the match are retrieved for context, and (5) the resulting sentences are rendered for the user.

§ 10.1.2
§ 10.2

ൾ ൾ ൾ

The following sections discuss the initial setup of the system when it is first started up, the text search algorithms, the image and video Application Program Interface (API) calls and the main design elements (text poetry and image spirals).

10.1 Setup

The Python web framework Flask (Ronacher n.d.) looks after loading and rendering the various pages for pata.physics.wtf (home, text-search, text-results, image-search, image-results, video-search, video-results, about and errors), which means most of the backend related code is written in Python. Although Flask contains a small development server, in a production environment a more capable server is needed. For this reason the Flask site runs on a Gunicorn server (Gunicorn n.d.) and is hosted on a UNIX machine.

10.1.1 Corpora

Instead of crawling the Internet pata.physics.wtf uses a local collection of texts for its text search. Setting up a custom web crawler would require a lot more resources (in terms of hardware, time and money) than practical for this project. There are two corpora containing 65 text files together.

The first corpus resembles the fictional library of 'equivalent books' from Jarry's *Exploits and Opinions of Dr. Faustroll, 'Pataphysician* (1996). In principle the corpus is just a folder within the tool's directory structure containing the following files:

0. Alfred Jarry: *Exploits and Opinions of Dr. Faustroll, 'Pataphysician* (1996)
1. Edgar Allen Poe: *Collected Works* (2008)
2. Cyrano de Bergerac: *A Voyage to the Moon* (2014)
3. Saint Luke: *The Gospel* (2014)
4. Léon Bloy: *Le Désespéré* (French) (2011)
5. Samuel Taylor Coleridge: *The Rime of the Ancient Mariner* (2013)
6. Georges Darien: *Le Voleur* (French) (2005)

7. Marceline Desbordes-Valmore: *Le Livre des Mères et des Enfants* (French) (2004)
8. Max Elskamp: *Enluminures* (French) (1898)
9. Jean-Pierre Claris de Florian: *Les Deux Billets* (French) (2012)
10. *One Thousand and One Nights* (Lang 2008)
11. Christian Grabbe: *Scherz, Satire, Ironie und tiefere Bedeutung* (German) (1995)
12. Gustave Kahn: *Le Conte de l'Or et Du Silence* (French) (n.d.)
13. Le Comte de Lautréamont: *Les Chants de Maldoror* (French) (2011)
14. Maurice Maeterlinck: *Aglavaine and Sélysette* (1918)
15. Stéphane Mallarmé: *Verse and Prose* (French) (2003)
16. Catulle Mendès: *The Mirror* and *la Divina Aventure* (English and Spanish) (1910, 2013)
17. Homer: *The Odyssey* (1999)
18. Joséphin Péladan: *Babylon* (EMPTY FILE)[5]
19. François Rabelais: *Gargantua and Pantagruel* (2004)
20. Jean de Chilra: *L'Heure Sexuelle* (EMPTY FILE)[5]
21. Henri de Régnier: *La Canne de Jaspe* (EMPTY FILE)[5]
22. Arthur Rimbaud: *Poesies Completes* (French) (2009)
23. Marcel Schwob: *Der Kinderkreuzzug* (German) (2012)
24. Alfred Jarry: *Ubu Roi* (French) (2005)
25. Paul Verlaine: *Poems* (2009)
26. Emile Verhaeren: *Poems* (2010)
27. Jules Verne: *A Journey to the Centre of the Earth* (2010)

§ 2.2

The original list as it appears in 'Faustroll' is shown in chapter 2.2. Three of the items have not been found as a resource. Some others have been approximated by using another text by the same author for example. Most of these were sourced from ***Project Gutenberg*** (Gutenberg 2016) in their original languages. The decision to get foreign language texts was partially due to the lack of out-of-copyright translated versions and partially because the original library in 'Faustroll' was also multi-lingual.

A note on copyright: UK copyright law states in section 5 that the duration of copyright for "literary, dramatic, musical or artistic works" is "70 years from the end of the calendar year in which the last remaining author of the work dies" (Copyright 2015). Maurice Maeterlinck and Marguerite Vallette-Eymery (a.k.a. Rachilde or Jean de Chilra) died less than 70 years ago and their work should still be under copyright. Alfred Jarry in the Simon Watson Taylor translation is a derivative work and is probably also still protected. However, copyright does not

[5] I have not been able to find any source texts online.

apply when used for “private and research study purposes” as stated in section 7 on *Fair dealing* of (Copyright 2012).

The second corpus is a collection of 38 texts by William Shakespeare (2011).

1. *The Sonnets*
2. *Alls Well That Ends Well*
3. *The Tragedy of Antony and Cleopatra*
4. *As You Like It*
5. *The Comedy of Errors*
6. *The Tragedy of Coriolanus*
7. *Cymbeline*
8. *The Tragedy of Hamlet, Prince of Denmark*
9. *The First Part of King Henry the Fourth*
10. *The Second Part of King Henry the Fourth*
11. *The Life of Kind Henry the Fifth*
12. *The First Part of Henry the Sixth*
13. *The Second Part of Henry the Sixth*
14. *The Third Part of Henry the Sixth*
15. *King Henry the Eigth*
16. *King John*
17. *The Tragedy of Julius Caesar*
18. *The Tragedy of King Lear*
19. *Love’s Labour’s Lost*
20. *The Tragedy of Macbeth*
21. *Measure for Measure*
22. *The Merchant of Venice*
23. *The Merry Wives of Windsor*
24. *A Midsummer Night’s Dream*
25. *Much Ado About Nothing*
26. *The Tragedy of Othello, Moor of Venice*
27. *King Richard the Second*
28. *Kind Richard III*
29. *The Tragedy of Romeo and Juliet*
30. *The Taming of the Shrew*
31. *The Tempest*
32. *The Life of Timon of Athens*
33. *The Tragedy of Titus Andronicus*
34. *The History of Troilus and Cressida*
35. *Twelfth Night or What You Will*
36. *The Two Gentlemen of Verona*

37. *The Winter's Tale*
38. *A Lover's Complaint*

10.1.2 Index

When the server is first started various setup functions (such as the creation of the index) are executed before any HTML is rendered. The search algorithms are triggered once a user enters a search term into the query field on any of the text, image or video pages.

</> 10.1

Each plain text file in the corpus is added to the internal library one by one. Source 10.1 shows how this is done. The `PlaintextCorpusReader` is a feature of the NLTK Python library (NLTK n.d.) for Natural Language Processing (NLP). The `words` function tokenises the text, i.e. it splits it into individual words and stores them as an ordered list.

```
library = PlaintextCorpusReader(corpus_root, '.*\.txt')
l_00 = library.words('00.faustroll.txt')
l_01 = library.words('01.poe1.txt')
...
l_27 = library.words('27.verne.txt')
```

Code 10.1 – Adding text files to the corpus library

</> 10.2

The `setupcorpus` function (see source 10.2) is called for each of the text files in the two corpora to populate the index data structures `l_dict` (for the Faustroll vocabulary) and `s_dict` (for the Shakespeare vocabulary).

```
dict = dictionary { dictionary { list [ ] } }
```

A dictionary in Python is what is known as an 'associative array' in other languages. Essentially they are unordered sets of ***key: value*** pairs. The `dict` used here is a dictionary where each key has another dictionary as it's value. Each nested dictionary has a list as the value for each key.

</> 10.2
§ B.6

Line 7 in source 10.2 starts looping through file `f`. Line 8 checks if the current word `w` contains anything other than alphabetical characters and whether or not `w` is contained in the relevant stop-word file `lang` (for a list of English stopwords see appendix B.6). If both of those conditions are true, a variable `y` is created on line 9 (such as 'l_00' based on '00.faustroll.txt') and `w` is added to the relevant dictionary file `dic` together with `y` and the current position `x` on line 10. After all files are processed, the two index structures look roughly like this:

```python
# f = input text
# lang = stopwords
# dic = dictionary
# d = 'l' for Faustroll or 's' for Shakespeare
def setupcorpus(f, lang, dic, d):
  # x = counter, w = word in file f
  for x, w in enumerate(f):
    if w.isalpha() and (w.lower() not in lang):
      y = d + '_' + (re.search(r"((\d\d).(\w)+.txt)",
       ↪ f.fileid)).group(2)
      dic[w.lower()][y].append(x)
```

Code 10.2 – 'setupcorpus': processing a text file and adding to the index—Python

```
{
  word1: {fileA: [pos1, pos2, ...], fileB: [pos], ...},
  word2: {fileC: [pos1, pos2], fileK: [pos], ...},
  ...
}
```

Using one of the terms from figure 6.2 on page 74 as an example, here are their entries in the index file (the files are represented by their number in the corpus, i.e. `l_00` is the 'Faustroll' file, `l_01` is the 'Poe' file, etc.). An excerpt from the actual `l_dict` can be found in the appendix B.1.

6.2

```
{
  doctor: {
    l_00: [253, 583, 604, 606, 644, 1318, 1471, 1858, 2334, 2431, 2446, 3039,
        ↪ 4743, 5034, 5107, 5437, 5824, 6195, 6228, 6955, 7305, 7822, 7892,
        ↪ 10049, 10629, 11055, 11457, 12059, 13978, 14570, 14850, 15063,
        ↪ 15099, 15259, 15959, 16193, 16561, 16610, 17866, 19184, 19501,
        ↪ 19631, 21806, 22570, 24867],
    l_01: [96659, 294479, 294556, 294648, 296748, 316773, 317841, 317854,
        ↪ 317928, 317990, 318461, 332118, 338470, 340548, 341252, 383921,
        ↪ 384136, 452830, 453015, 454044, 454160, 454421, 454596, 454712,
        ↪ 454796, 454846, 455030, 455278, 455760, 455874, 456023, 456123,
        ↪ 456188, 456481, 456796, 457106, 457653, 457714, 457823, 457894,
        ↪ 458571, 458918, 458998, 459654, 459771, 490749],
    l_02: [11476, 12098, 28151, 36270], ...
  }, ...
}
```

10.2 Text

After the setup stage is completed and the webpage is fully loaded, user input in the form of a text query is required to trigger the three pataphysical algorithms.

Image and video search do not use all three algorithms — where relevant this is highlighted in each section. Generally the following descriptions refer to the text search functionality only.

10.4 Figure 10.4 previously showed the rough sequence of events in text search and highlighted that the pataphysicalisation from query to patadata happens in the `textsurfer.py` Python script file.

10.2.1 Clinamen

§ 4.2.5 The clinamen was introduced in chapter 4.2.5 but to briefly summarise it, it is the unpredictable swerve that Bök calls “the smallest possible aberration that can make the greatest possible difference” (2002).

> Like all digitally encoded information, it has unavoidably the uncomfortable property that the smallest possible perturbations —i.e. changes of a single bit— can have the most drastic consequences. (Dijkstra 1988)

In simple terms, the clinamen algorithm works in two steps:

1. get clinamen words based on dameraulevenshtein and faustroll text,
2. get sentences from corpus that match clinamen words.

</> 10.4 It uses the *Faustroll* (Jarry 1996) as a base document and the Damerau-Levenshtein algorithm (Damerau 1964; Levenshtein 1966) (which measures the distance between two strings (with 0 indicating equality) to find words that are similar but not quite the same. The distance is calculated using insertion, deletion, substitution of a single character, or transposition of two adjacent characters. This means that we are basically forcing the program to return matches that are of distance two or one, meaning they have two or one spelling errors in them.

```python
# w = query word
# c = corpus
# i = assigned distance
def clinamen(w, c, i):
  # l_00 = Faustroll text
  words = set([term for term in l_00 if dameraulevenshtein(w, term) <=
    ↪ i])
  out, sources, total = get_results(words, 'Clinamen', c)
  return out, words, sources, total
```

Code 10.3 – ‘clinamen’: pataphysicalising a query term—Python

Source 10.3 line 6 creates the set of clinamen words using a list comprehension. </> 10.3 It retrieves matches from the Faustroll file `l_00` with the condition that they are of Damerau-Levenshtein distance `i` or less to the query term `w` (see source 10.4). </> 10.4 Duplicates are removed. Line 7 then makes a call to the generic `get_results` function to get all relevant result sentences, the list of source files </> 10.5 and the total number of results.

```
# Michael Homer 2009
# MIT license
def damerauulevenshtein(seq1, seq2):
  oneago = None
  thisrow = range(1, len(seq2) + 1) + [0]
  for x in xrange(len(seq1)):
    twoago, oneago, thisrow = oneago, thisrow, [0] * len(seq2) + [x + 1]
    for y in xrange(len(seq2)):
      delcost = oneago[y] + 1
      addcost = thisrow[y - 1] + 1
      subcost = oneago[y - 1] + (seq1[x] != seq2[y])
      thisrow[y] = min(delcost, addcost, subcost)
      if (x > 0 and y > 0 and seq1[x] == seq2[y - 1] and
        seq1[x - 1] == seq2[y] and seq1[x] != seq2[y]):
          thisrow[y] = min(thisrow[y], twoago[y - 2] + 1)
  return thisrow[len(seq2) - 1]
```

Code 10.4 – Damerau-Levenshtein algorithm (M. Homer 2009)—Python

The clinamen algorithm mimics the unpredictable swerve, the smallest possible aberration that can make the greatest possible difference, or the smallest possible perturbations with the most drastic consequences.

10.2.2 Result Sentences

The `get_results` function (see source 10.5) is used by all three text algorithms </> 10.5 (clinamen, syzygy and antinomy). Given the nested structure of the indexes `l_dict` and `s_dict`, the function loops through each of the `words` passed to it (`r`) first and then each file in `files.items()`. Lines 8 and 9 retrieve the dictionary of files for term `r` from the relevant dictionary. Line 13 gets the author and full title of file `e` and adds it to the list of sources in line 14. Line 15 makes use of another function called `pp_sent` (see source 10.6) to get an actual sentence </> 10.6 fragment for the current word `r` in file `e`, which is then added to the output. The output is structured as a triple containing the author and title, the list of resulting sentences and the name of the algorithm used.

In function `pp_sent` (source 10.6) line 5 is important to note because it is a key </> 10.6 functionality point. Even though the index files store a full list of all possible

```python
# words = patadata words
# algo = name of algorithm
# corp = name of corpus
def get_results(words, algo, corp):
  total = 0
  out, sources = set(), set()
  for r in words:
    if corp == 'faustroll': files = l_dict[r]
    else: files = s_dict[r]
    # e = current file
    # p = list of positions for term r in file e
    for e, p in files.items():
      f = get_title(e)
      sources.add(f)
      o = (f, pp_sent(r.lower(), e, p), algo)
      total += 1
      out.add(o)
  return out, sources, total
```

Code 10.5 – 'get_results': retrieving all sentences for a list of words—Python

positions of a given word in each file, the `pp_sent` function only retrieves the sentence of the very first occurrence of the word rather than each one. This decision was taken to avoid overcrowding of results for the same keyword and is further discussed in chapter 12.2.2.

§ 12.2.2

Line 8 creates a list of punctuation marks needed to determine a suitable sentence fragment. Lines 9–17 and 18–26 set the `pos_b` (position before) and `pos_a` (position after) variables respectively. These positions can be up to 10 words before and after the keyword `w` depending on the sentence structure (punctuation marks). In line 28 the actual sentence fragment up to the keyword is retrieved, while in line 29 the fragment just after the keyword is retrieved. `ff[pos_b:pos]` for example returns the list of words from position `pos_b` to position `pos` from file `ff`. The built-in Python `.join()` function then concatenates these words into one long string separated by spaces. On line 30 a triple containing the pre-sentence, keyword and post-sentence is set as the output and then returned.

10.2.3 Syzygy

§ 4.2.4

The concept of the syzygy was introduced in chapter 4.2.4 but can be roughly described as surprising and confusing. It originally comes from astronomy and denotes the alignment of three celestial bodies in a straight line. In a pataphysical context it is the pun. It usually describes a conjunction of things, something unexpected and surprising. Unlike serendipity, a simple chance encounter, the

```python
# w = the word (lower case)
# f = the file
# p = the list of positions
def pp_sent(w, f, p):
  out, pos = [], p[0] # FIRST OCCURRENCE
  ff = eval(f)
  pos_b, pos_a = pos, pos
  punct = [',', '.', '!', '?', '(', ')', ':', ';', '\n', '-', '_']
  for i in range(1, 10):
    if pos > i:
      if ff[pos - i] in punct:
        pos_b = pos - (i - 1)
        break
      else:
        if ff[pos - 5]: pos_b = pos - 5
        else:           pos_b = pos
    else: pos_b = pos
  for j in range(1, 10):
    if (pos + j) < len(ff):
      if ff[pos + j] in punct:
        pos_a = pos + j
        break
      else:
        if ff[pos + j]: pos_a = pos + j
        else:           pos_a = pos
    else: pos_a = pos
  if pos_b >= 0 and pos_a <= len(ff):
    pre = ' '.join(ff[pos_b:pos])
    post = ' '.join(ff[pos+1:pos_a])
    out = (pre, w, post)
  return out
```

Code 10.6 – 'pp_sent': retrieving one sentence—Python

syzygy has a more scientific purpose. In simple terms, the syzygy algorithm works in two steps:

1. get syzygy words based on synsets and hypo-, hyper-, holo- and meronyms from WordNet,
2. get sentences from corpus that match syzygy words.

The syzygy function makes heavy use of WordNet (Miller 1995) through the NLTK Python library (NLTK n.d.) to find suitable results (importing it using the following command `from nltk.corpus import wordnet as wn`). Specifically, as shown in source 10.7, the algorithm fetches the set of synonyms (synsets) on line 5. It then loops through all individual items `ws` in the list of synonyms `wordsets` in

</> 10.7

line 7–20. It finds any hyponyms, hypernyms,holonyms, and meronyms for `ws` 12.3 (each of which denotes some sort of relationship or membership with its parent </> 10.8 synonym—see figure 12.3) using the `get_nym` function (see lines 8, 11, 14, and </> 10.5 17). Line 21 makes use of the `get_results` function (see source 10.5) in the same was as the clinamen function does.

```
1  # w = word
2  # c = corpus
3  def syzygy(w, c):
4    words, hypos, hypers, holos, meros = set(),set(),set(),set(),set()
5    wordsets = wn.synsets(w)
6    hypo_len, hyper_len, holo_len, mero_len, syno_len = 0,0,0,0,0
7    for ws in wordsets:
8      hypos.update(get_nym('hypo', ws))
9      hypo_len += len(hypos)
10     words.update(hypos)
11     hypers.update(get_nym('hyper', ws))
12     hyper_len += len(hypers)
13     words.update(hypers)
14     holos.update(get_nym('holo', ws))
15     holo_len += len(holos)
16     words.update(holos)
17     meros.update(get_nym('mero', ws))
18     mero_len += len(meros)
19     words.update(meros)
20     syno_len += 1
21   out, sources, total = get_results(words, 'Syzygy', c)
22   return out, words, sources, total
```

Code 10.7 – 'syzygy': pataphysicalising a query term—Python

</> 10.8 The `get_nym` function in source 10.8 shows how the relevant 'nyms' are retrieved for a given synset. Line 5 initialises the variable `hhh` which gets overwritten later on. Several `if` statements separate out the code run for the different 'nyms'. Lines 6–7 retrieves any hyponyms using NLTK's `hyponyms()` function . Similarly lines 8–9 retrieve hypernyms, lines 10–14 retrieve holonyms, and lines 15–19 retrieve meronyms. Finally, line 20–23 adds the contents of `hhh` to the output of the function.

The syzygy algorithm mimics an alignment of three words in a line (query → synonym → hypo/hyper/holo/meronym).

10.2.4 Antinomy

The antimony, in a pataphysical sense, is the mutually incompatible. It was § 4.2.2 previously introduced in chapter 4.2.2. In simple terms, the antinomy algorithm

```python
# nym = name of nym
# wset = synset
def get_nym(nym, wset):
  out = []
  hhh = wset.hyponyms()
  if nym == 'hypo':
    hhh = wset.hyponyms()
  if nym == 'hyper':
    hhh = wset.hypernyms()
  if nym == 'holo':
    hhhm = wset.member_holonyms()
    hhhs = wset.substance_holonyms()
    hhhp = wset.part_holonyms()
    hhh = hhhm + hhhs + hhhp
  if nym == 'mero':
    hhhm = wset.member_meronyms()
    hhhs = wset.substance_meronyms()
    hhhp = wset.part_meronyms()
    hhh = hhhm + hhhs + hhhp
  if len(hhh) > 0:
    for h in hhh:
      for l in h.lemmas():
        out.append(str(l.name()))
  return out
```

Code 10.8 – 'get_nym': retrieving hypo/hyper/holo/meronyms—Python

works in two steps:

1. get antinomy words based on synsets and antonyms from WordNet,
2. get sentences from corpus that match antinomy words.

</> 10.9

For the antinomy I simply used WordNet's antonyms (opposites) (source 10.9). In principle, this function is similar to the algorithm for the syzygy. It finds all antonyms through NLTK's `lemmas()[0].antonyms()` function on line 7 and retrieves result sentences using the `get_results` function on line 12. </> 10.5

The antinomy algorithm mimics the mutually incompatible or polar opposites.

10.2.5 Formalisation

A formal description of the `pata.physics.wtf` system in terms of an IR model described in chapter 6.1.1 is unsuitable. It assumes for example the presence of some sort of ranking algorithm $R(q_i, d_j)$. § 6.1.1

```python
# w = input query term
# c = name of corpus
def antinomy(w, c):
  words = set()
  wordsets = wn.synsets(w)
  for ws in wordsets:
    anti = ws.lemmas()[0].antonyms()
    if len(anti) > 0:
      for a in anti:
        if str(a.name()) != w:
          words.add(str(a.name()))
  out, sources, total = get_results(words, 'Antinomy', c)
  return out, words, sources, total
```

Code 10.9 – 'antinomy': pataphysicalising a query term—Python

Making relevant changes (e.g. exchanging the ranking function for a pataphysicalisation function) to the specification by Baeza-Yates and Ribeiro-Neto (2011), an approximate system description for the Faustroll corpus text search could be as follows.

D = the set of documents $\{d_1, \ldots, d_m\}$
m = the number of all documents in D ($|D| = 28$)
V = the set of all distinct terms $\{v_1, \ldots, v_n\}$ in D not including stopwords
q = the user query
F = the set of patalgorithms $\{f_C, f_S, f_A\}$
P = the set of pataphysicalised query terms $\{p_1, \ldots, p_u\}$
u = the number of terms in P
$P(q)$ = the set of patadata $\{P(q)_C \cup P(q)_S \cup P(q)_A\}$ for query q
R = the set of results $\{r_1, \ldots, r_o\}$
o = the number of results in R
$R(P(q))$ = the set of results $\{R(P(q)_C) \cup R(P(q)_S) \cup R(P(q)_A)\}$ produced by each algorithm in F
r = a result of form (d, sentence, f)

We can then define the three patalgorithms in a more formal way as shown in equations 10.1, 10.2, and 10.3.

$$P(q)_C = \{p \in v_0 : 0 < \text{dameraulevenshtein}(q, p) \leq 2\} \qquad (10.1)$$

Σ 10.1
</> 10.4

`dameraulevenshtein(q,p)` in equation 10.1 is the Damerau-Levenshtein algorithm as described in section 10.4 and v_0 is the Faustroll text.

$$
\begin{aligned}
P(q)_S &= \{p \in V : p \in \text{nyms}(s),\ \forall s \in \text{synonyms}(q)\} \\
\text{where nyms}(s) &= \text{hypos}(s)\ \cup\ \text{hypers}(s)\ \cup\ \text{holos}(s)\ \cup\ \text{meros}(s)
\end{aligned}
\tag{10.2}
$$

`synonyms(q)` in equation 10.2 is the WordNet/NLTK function to retrieve all synsets for the query q and the four 'nym' functions return the relevant hyponyms, hypernyms, holonyms or meronyms for each of the synonyms. Σ 10.2

$$
P(q)_A = \{p \in V : p \in \text{antonyms}(s),\ \forall s \in \text{synonyms}(q)\} \tag{10.3}
$$

Similarly, in equation 10.3 the `synonyms(q)` function returns WordNet synsets for the query q and the `antonyms(s)` function returns WordNet antonyms for each of the synonyms. Σ 10.3

$$
R(P(q)) = \{(d \in D,\ sent(p) \in d,\ f \in F) : \forall\ p \in P(q)_f))\} \tag{10.4}
$$

The set of results $R(P(q))$ can then be defined as shown in equation 10.4. It returns a list of triples containing the source text d, the sentence `sent(p)` and the algorithm f. For each pataphysicalised query term p one sentence is retrieved per file d. Σ 10.4

10.3 Image & Video

The image and video search of `pata.physics.wtf` both work slightly differently to the text search described in section 10.2. In simple terms, the image and video search works in three steps: § 10.2

1. translate query,
2. pataphysicalise the translation,
3. retrieve matching images/videos using API calls.

</> 10.10

The first step is to translate the search terms as shown in source 10.10. Lines 2 and 4 set up the API connection to the Microsoft Translator tool (Translator 2011) given an ID and 'secret', neither of which are included here for security reasons. The query `sent` then passes through a chain (alignment) of three translations in true syzygy fashion: from English → French, from French → Japanese, and from Japanese → English (lines 5–7). All three languages are then returned in a triple (line 8).

```python
# sent = the query string
from microsofttranslator import Translator
def transent(sent):
  translator = Translator(microsoft_id, microsoft_secret)
  french = translator.translate(sent, "fr")
  japanese = translator.translate(french, "ja")
  patawords = translator.translate(japanese, "en")
  translations = (french, japanese, patawords)
  return translations
```

Code 10.10 – 'transent': translating query between English-French-Japanese-English—Python

The next step is to pataphysicalise the translated query (see source 10.11). The `pataphysicalise` function transforms this translation in a process slighlty simplified from the `syzygy` algorithm. The decision to simplify the algorithm was made due to performance issues related to the API calls that follow in the final step of the search process.

</> 10.7

In line 5 WordNet synsets are retrieved using NLTK's `synsets` function. For each of these synsets we get a list of synonyms (line 8) which we add to the output in a normalised form (line 11) removing any underscores if there are any.

```python
1  # words = query term(s)
2  def pataphysicalise(words):
3    sys_ws = set()
4    for word in words:
5      synonyms = wn.synsets(word)
6      if len(synonyms) > 0:
7        for s in synonyms:
8          for l in s.lemmas():
9            x = str(l.name())
10           o = x.replace('_', ' ')
11           sys_ws.add(o)
12   return sys_ws
```

Code 10.11 – 'pataphysicalise': pataphysicalise image and video query terms—Python

10.5

Figure 10.5 previously showed the rough sequence of events in an image and video search and highlighted that the pataphysicalisation from query to patadata happens in the `imgsurfer.py` Python script file while the production of results from that patadata happens in the `fania.js` JavaScript file.

§ 10.3.1

And finally, API calls to the various external tools are made. This is described in section 10.3.1 below.

10.3.1 REST & API

The final step of the image and video search process described on page 158 is to retrieve matching images/videos using API calls to Flickr (Flickr n.d. FlickrAPI n.d.), Getty (Getty n.d. GettyAPI n.d.), Bing (BingAPI 2012; Bing n.d.), YouTube (YouTube n.d.) and Microsoft Translator (Translator 2011).

The patadata used to make the API calls is limited to 10 keywords and uses the function `random.sample(pata, 10)`, where `pata` is the set of terms obtained by pataphysicalising the query translation.

A RESTful API allows browsers ('clients') to communicate with a web server via HTTP methods such as GET and POST. The idea is that a given service, like the Microsoft Bing search API, can be accessed in a few simple steps using JavaScript Object Notation (JSON) (JSON n.d.). These are:

1. for each of the 10 query terms do:
    a) construct the URL with the query request
    b) setup authentication
    c) send URL and authentication
    d) receive response in JSON
    e) add result to output list `imglist`
2. once 10 results are reached, render results as spiral

</> 10.12
</> 10.13
§ B.7

Source 10.12 shows how such an API call is made using JavaScript (in this case Flickr). Source 10.13 below shows how 10 seperate images are collected into one results list and the `createSpiral` function is called to render the images to the user in HTML (see appendix B.7 for the relevant code snippet).

The Bing and Getty searches work in a similar way with one exception. Getty does not populate the output list by doing 10 individual API calls but rather by adding 10 results from 1 call. This is due to a time restriction in the Getty API; it doesn't not allow 10 calls in a second.

An example URL request for the Flickr image search with the query term of 'kittens' and a requested response format of JSON is this: `http://api.flickr.com/services/feeds/photos_public.gne?jsoncallback=?tags=kittens&tagmode=all&format=json`. Flickr will then send back the response in JSON

```
function flickrsearch(patadata){
  for(var x=0; x<10; x++){
    $.getJSON("http://api.flickr.com/services/feeds/photos_public.gne?jsoncallback=?",
      {
        tags: patadata[x].query,
        tagmode: "all",
        format: "json"
      },
      function(data,status,ajax) {
        var title = "", media = "", link = "";
        if (data.items[0] != undefined) {
          title = data.items[0].title;
          media = data.items[0].media.m;
          link = data.items[0].link;
        }
        imgList([title, media, link]);
      }
    );
  }
};
```

Code 10.12 – ‘flickrsearch’: using the Flickr API to retrieve images—JavaScript

```
var allImages = [];
function imgList(img){
  if (allImages[0] != "") {
    allImages.push(img);
  }
  if (allImages.length === 10) {
    createSpiral(allImages);
  }
}
```

Code 10.13 – ‘imgList’: accumulates 10 images and calls the ‘createSpiral’ function—JavaScript format. One entry of the list of results is shown below (with whitespace formatting added for convenience). The algorithm in source 10.12 only retrieves the `data.items[0].title`, `data.items[0].media.m` and `data.items[0].link` (lines 12, 13, and 14) and ignores all other data fields. </> 10.12


```
({...
  "items":
    [{
      "title": "P_20161101_191123",
      "link": "http://www.flickr.com/photos/pinknancy/30078720153/",
      "media": {"m":"http://farm6.staticflickr.com/5759/30078720153_f03e036e89
            ↪  _m.jpg"},
      "date_taken": "2016-11-01T19:11:23-08:00",
      "description": ...,
      "published": "2016-11-01T15:28:10Z",
      "author": "nobody@flickr.com (pinknancy)",
      "author_id": "8748781@N08",
      "tags": ""
    },...]
})
```


Once the `imglist` contains 10 items it is passed to the `createSpiral` function which renders it to HTML. Appendix B.4 shows an example shortened JSON result from Bing. § B.4

๑ ๑ ๑

The video search also uses an API to retrieve results. This function is written in Python and uses the *Requests* library (Reitz n.d.) to make the API calls to YouTube (YouTube n.d.) as shown in source 10.14. </> 10.14

</> 10.10

First, the query is translated using the `transent` function on line 3. Line 4 seperates the English translation into its own list `transplit` which is then pataphysicalised on line 5 using the algorithm described in source 10.11. </> 10.11

Lines 6–9 construct the first part of the URL to use for the Representational State Transfer (REST) request. Lines 10–23 then loop through each of the patadata terms generated by the `pataphysicalise` function on line 5 to make a call and retrieve some video details (title, thumbnail and ID) as seen on lines 17–19. On line 20 these details are added to the output list.

The video results are then also displayed in a golden spiral in the same way as the images. This is described in section 10.4.3. § 10.4.3

```
def getvideos(query):
  out = []
  translations = transent(query)
  transplit = translations[2].split(' ')
  tmp = pataphysicalise(transplit)
  b0 = "https://www.googleapis.com/youtube/v3/search?"
  b1 = "&order=viewCount&part=snippet&"
  b3 = "&type=video&key=%s" % yt_key
  b4 = "&maxResults=10&safeSearch=strict"
  for x in tmp:
    y = ' '.join(x)
    b2 = "q=%s" % translations[2]
    yturl = ''.join([b0, b1, b2, b3, b4])
    vids = requests.get(yturl)
    if vids.json()['items']:
      for i in vids.json()['items']:
        vidtitle = i['snippet']['title']
        vidthumb = i['snippet']['thumbnails']['default']['url']
        vidid = i['id']['videoId']
        out.append((vidtitle, vidthumb, vidid))
      break
    else:
      out = []
  return out, translations
```

Code 10.14 – 'getvideos': using the YouTube API to retrieve images—Python

10.4 Design

Once the patalgorithms have produced their respective results, the page displaying these results can be rendered. This is done using the templating language Jinja (Ronacher 2008) and HTML (with CSS stylesheets and some JavaScript).

§ 2.1 One of the key requirements for the *Syzygy Surfer* tool was that "the user should be able to choose the techniques they use" (Hendler and Hugill 2011). This has been adopted for `pata.physics.wtf` in the sense that the user has different options for the display of results.

The text results page has three different result styles, with 'Poetry — Queneau' being the default.

Poetry Displayed in sonnet style (two quatrains and two tercets) if possible, although no rhyming pattern is used[6].

- Queneau — Each line can be changed manually.

[6] This is addressed in chapter 13.3.

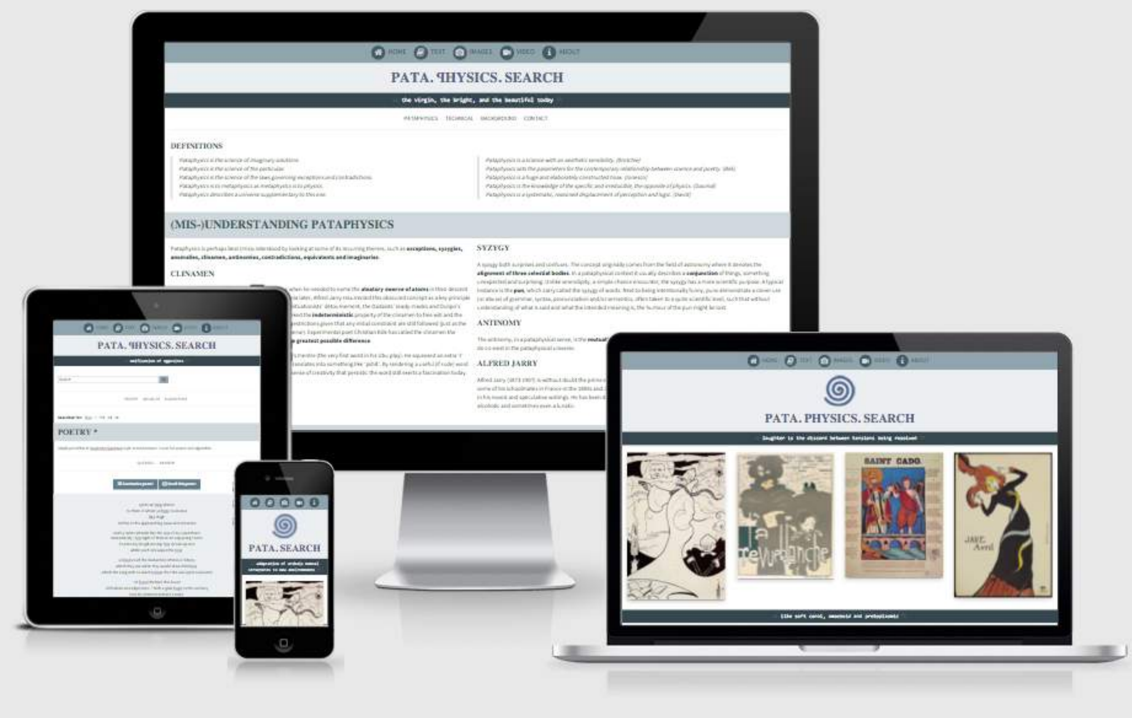


Figure 10.6 – Responsive design of pata.physics.wtf

- Random — The whole poem can be randomised.

Sources Ordered by source text.
Algorithms Ordered by algorithm.

The image and video results pages work the same way. They both have two display options, with the 'Spiral' option being the default. The spirals are modelled on the idea of golden spirals (more precisely an approximation in the form of a Fibonacci spiral).

Spiral Displayed as square images/videos in a spiral.
List Displayed as a simple list.

The overal visual design is shown in image 10.6. 10.6

10.4.1 Poetry

Source 10.15 shows the segment of HTML/Jinja code that renders the Queneau poetry. The code renders the 4 stanzas of the poem. This is done using two nested Jinja `for` loops (line 2 and line 10). Line 2 loops through the (ideally) 14 lines of the poem. `lol` can be considered a masterlist of all sublists for each poem line. </> 10.15

Functionality for sending the currently showing poem per email is added via a button which calls a JavaScript function `onclick="return getContent(this)"` which then retrieves the content of each line in the poem and sends it to the body of the email.

`all_sens` is the pool of all sentences. It is structured as follows.

```
[(title, (pre, word, post), algorithm), ...]
```

`lol` is a list subdivided into partitions for each line of the sonnet. Let's say there are 350 sentences overall in `all_sens`. To divide them equally among the 14 lines of a sonnet, we need to create `lol` with 14 equal parts of 25 sentences.

```
[all_sens[0-24], all_sens[25-49], ...,  all_sens[325-349]]
```

```
<div>
  {% for n in range(1, lol|length + 1) %}
    {% set wid = ["wn", n|string]|join %}
    {% set lid = ["lyr", n|string]|join %}
    {% set sid = ["scrollLinks", n|string]|join %}
    {% set aid = lol[n-1] %}
    <div id="poems">
      <div id="{{wid}}" class="wn">
        <div id="{{lid}}" class="lyr">
          {% for sens in aid %}<span title="{{ sens[0] }}, {{ sens[2]
           ↪  }}">{{ sens[1][0] }} <form class="inform"
           ↪  action="../textresults" method="post"><input
           ↪  class="inlink" type="submit" name="query" value="{{
           ↪  sens[1][1] }}" onclick="loading();"></input></form> {{
           ↪  sens[1][2] }}</span>{% endfor %}
        </div>
      </div>
      <div id="{{sid}}" class="scrollLinks"></div>
    </div>
  {% endfor %}
</div>
```

Code 10.15 – Simplified HTML code for rendering Queneau style poems

10.7 Changing a line of the poem is achieved by clicking on one of the buttons on either side of the poem's line (as shown in image 10.7). This will trigger a JavaScript function (based on (Dyn Web n.d.)) to automatically scroll to the next sentence.

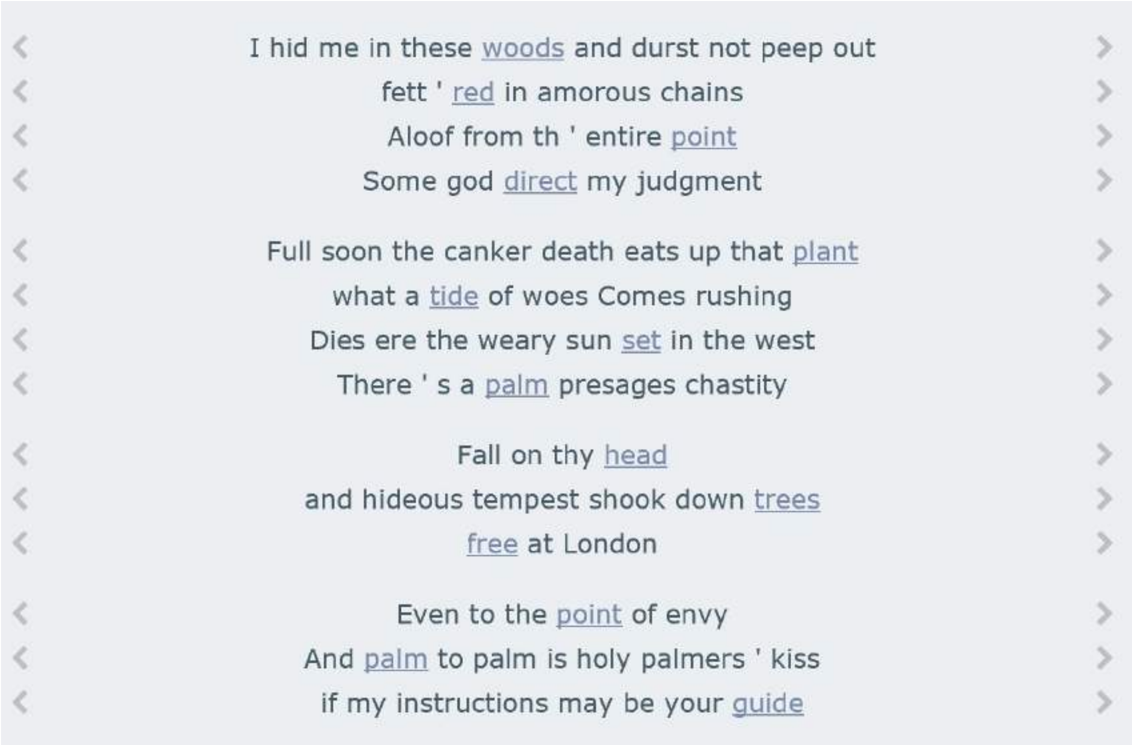



Figure 10.7 – Example Queneau poem for query 'tree'

Non-Queneau poems have a slightly different functionality. It is not possible to change the poem line by line but rather the whole poem can be randomised on demand. This relies on a random number generator in JavaScript. A function `shufflePoem()` creates a random variable `r` as `Math.floor(Math.random() * n)`, which can then be used to generate a new list of 14 lines for the poem randomly selected from the pool of sentences `all_sens`.

10.4.2 Lists

The two other ways to display text results are as a list ordered by source or by patalgorithm which works in a similar way to what is described in source 10.16. The code is wrapped in an HTML unordered list tag `<ul>`. A Jinja `for` loop generates the individual `<li>` tags on line 4. </> 10.16

A `sens` in `all_sens` is structured as `(title, (pre, word, post), algorithm)`. This means that to access the name of the algorithm we need to call the Jinja template `{{ sens[2] }}`, to get the first half of the sentence we need `{{ sens[1][0] }}`, the middle keyword (i.e. the patadata term) `{{ sens[1][1] }}` and the second half of the sentence `{{ sens[1][2] }}`.

Image 10.8 shows a shortened example set of results for query 'tree' ordered by source, that is, ordered by original file. 10.8

```
<ul>
  {% for sens in all_sens %}
    {% if file == sens[0] %}
      <li title=`{{ sens[2] }}'>...{{ sens[1][0] }} <form class=`inform' action=`../textresults' method=`post'><input class=`w3-hide' type=`radio' name=`corpus' value=`{{ corpus }}' checked><input class=`inlink' type=`submit' name=`query' value=`{{ sens[1][1] }}' onclick=`loading();'></input></form> {{ sens[1][2] }}...</li>
    {% endif %}
  {% endfor %}
</ul>
```

Code 10.16 – Simplified HTML code for rendering a list of text results by source

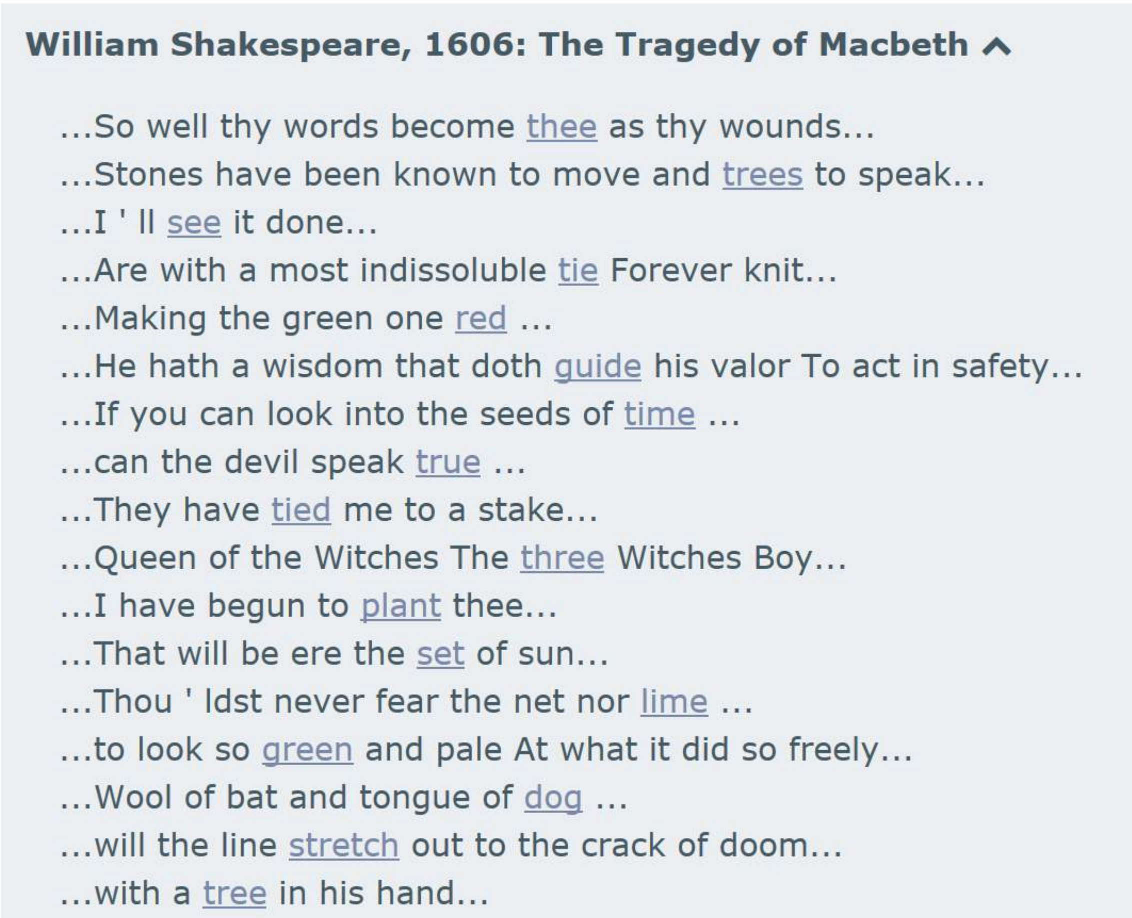


Figure 10.8 – Example results for query 'tree' ordered by source

Image 10.9 shows a shortened example set of results for query 'tree' ordered by patalgorithm, that is, ordered by the algorithm which produced the patadata. 10.9

Figure 10.9 – Example results for query 'tree' ordered by patalgorithm

10.4.3 Spiral

The image and video spirals are constructed in complicated nested HTML components. The code for generating an image spiral is shown in appendix B.7. The video spiral is constructed in a similar way but directly in the HTML file as opposed to in the JavaScript file. The video spiral is almost identical, the only difference is the biggest 5 videos are atually embedded as videos. The smaller 5 videos are shown as still images which link to the relevant YouTube page. § B.7

Generally, the idea was taken from the pataphysical ***grand gidouille*** (see chapter 4) and represented as a Fibonacci spiral. Figure 10.10 shows a spiral created using the Flickr image search for query 'blue mountains' overlaid with a white Fibonacci spiral to highlight the structure. § 4 10.10

10.5 Prototypes

The final website `pata.physics.wtf` went through several iterations of development since it was first conceived in 2012. This included 3 major technical updates since the first prototype and 2 new visual re-designs. Table 10.1 shows the main differences and similarities between the versions. 10.1

Images 10.11, 10.12 and 10.6 show the 3 main visual designs.

Figure 10.10 – Fibonacci spiral overlaid onto an image results for query 'blue mountains' using Flickr

Table 10.1 – Comparison of different versions of `pata.physics.wtf`

	Version 1	Version 2	Version 3	Version 4
Language(s)	Python, Django	Python, Flask	Python, Flask	Python, Flask, JavaScript
Server	Django, Heroku	Flask, Mnemosyne	Flask, Gunicorn, Mnemosyne	Flask, Gunicorn, OVH
Features	Text	Text, Image, Video	Text, Image, Video	Text, Image, Video
Corpus	Faustroll text	Faustroll text	Faustroll's library	Faustroll's library and Shakespeare
API's	WordNet	WordNet, Flickr, Bing, YouTube, Microsoft Translator	WordNet, Bing, YouTube, Microsoft Translator	WordNet, Flickr, Getty, Bing, YouTube, Microsoft Translator
Design	Algorithms	Algorithms, Spiral	Algorithms, Source, Poetry, Spiral, List	Algorithms, Source, Poetry, Spiral, List
Responsive	No	Yes	Yes	Yes

Figure 10.11 – First version of pata.physics.wtf

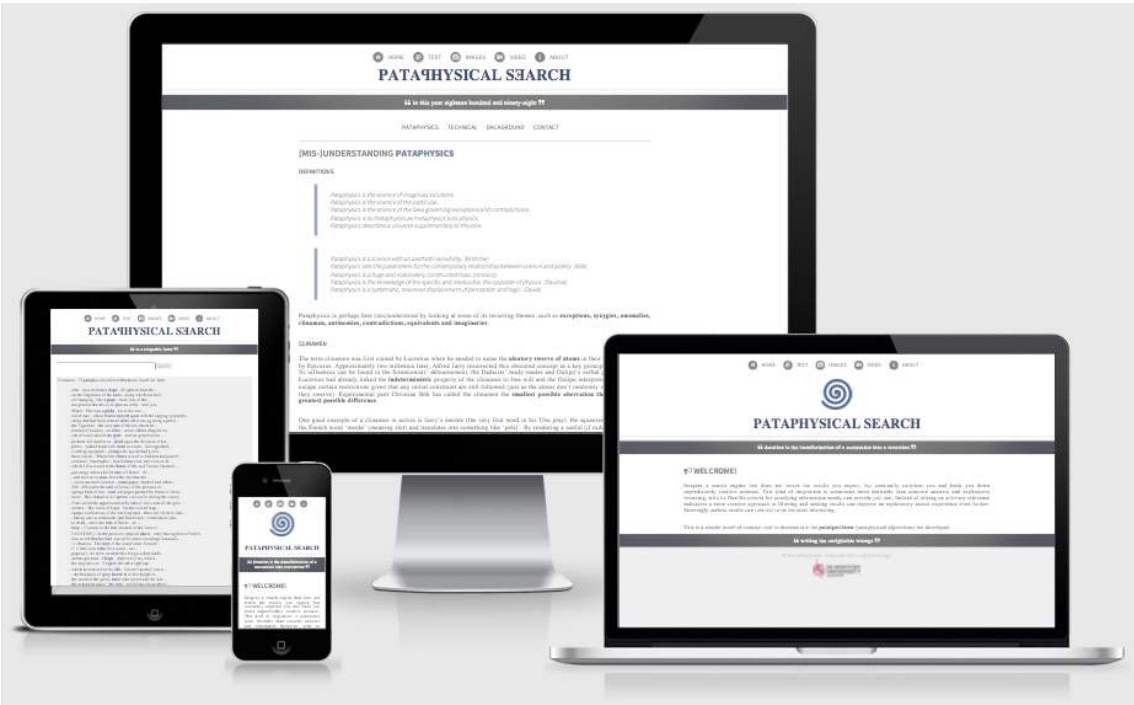


Figure 10.12 – Second major version of pata.physics.wtf

The latest version, which is now live at pata.physics.wtf, introduced major changes to the initial setup stage of the system and a lot of the code was refactored and improved. As of the date of writing this, there were over 360 commits in the git repository since 2012. See appendix D.1.

§ D.1

APPLICATIONS

11

Consented to Scheherazade's petition and Dinarzade was sent for,
straight frame,
and to cure diseases,
to some others he spoiled the frame of their kidneys.

Qui peut l'espérer ?...,
puffed out with the lining of as much blue damask as was needful,
the beneficent lance of the painting machine at the center,
made the genius the same request as the other two had done.

Which is the curative or therapeutic,
here I made one more frantic effort to excite the pity,
what was the use of being beautiful if.

Ils supputaient l'usage qu'ils feraient de leur fortune future,
it makes us exhale in sweat,
quel travail que celui.

This chapter introduces two real world applications of this research and details some of the publications, talks and exhibitions that featured this project.

11.1 Patadata Ontology

Andrew Dennis wrote an undergraduate thesis entitled *Investigation of a patadata-based ontology for text based search and replacement* (2016a), which was directly based on some of the work presented in this thesis and previously published work (Hugill, Yang et al. 2013; Raczinski, Yang and Hugill 2013). His project can be described as such:

1. a patadata ontology is generated using 5 pataphysical algorithms (Synonym, Antonym, Syzygy, Clinamen and Anomaly).
2. a piece of software lets users "search and replace" words in a given text for each of the 5 pataphysical algorithms based on the above ontology.

The 5 algorithms he discusses could be seen as an extension of my own work (which only described 3 algorithms - Clinamen, Syzygy and Antinomy).

Synonym
: Pataphysical equivalence—implemented using WordNet's synsets. </> 11.1

Antonym
: Pataphysical coexistence of mutually incompatible concepts—implemented using WordNet's antonyms. </> 11.2

Syzygy
: Pataphysical alignment of three entities—implemented using WordNet's synonyms and hypernyms. </> 11.4

Clinamen

</> 11.5 Pataphysical swerve—implemented using Damerau-Levenshtein algorithm.

Anomaly

</> 11.3 Pataphysical exceptions—implemented using randomisation.

Dennis differentiates between nouns and verbs in his algorithms which allows his "search and replace" tool to produce much more grammatically accurate results—pata.physics.wtf does not distinguish between word forms like this.

11.1.1 Algorithms

</> 11.1 The synonym algorithm works by generating WordNet synonyms for a given keyword. Source 11.1 shows the pseudo-code for this algorithm.

```
function generate_synonym(input):
  synonym_list = []
  for word in synonym_set(input):
    if word is noun or word is verb:
      return word
  return input
```

Code 11.1 – Andrew Dennis' synonym generation algorithm

</> 11.2 The antonym algorithm in source 11.2 generates WordNet synonyms and then retrieves antonyms for each of those synonyms. This is very similar to the an-
§ 10.2.4 tinomy algorithm presented in section 10.2.4 with the additional handling of nouns and verbs as separate entities.

</> 11.3 The algorithm for the anomaly works by generating a random number x and retrieving item number x in the dictionary. Source 11.3 shows the pseudo-code for this algorithm.

The syzygy algorithm works by generating WordNet synonyms and retrieving hypernyms for each of those and then retrieving any synonyms for those hypernyms (i.e. it creates a syzygy alignment from synonym → hypernym → synonym).
</> 11.4 Source 11.4 shows the pseudo-code for this algorithm. This is slightly different
§ 10.2.3 to the syzygy algorithm presented in section 10.2.3 in that it aligns keyword—synonyms—hypernyms—synonyms rather than keyword—synonyms—hyper/-hypo/holo/meronyms.

</> 11.5 Finally, the clinamen algorithm works by finding words in the dictionary that have a Damerau-Levenshtein distance of 2 to the keyword. Source 11.5 shows the pseudo-code for this algorithm. This is based almost directly on the clinamen

```
function generate_antonym(input):
  antonym_list = []
  for word in synonym_set(input):
    if input is noun:
      if word is noun:
        for lemma in word.lemmas:
          if lemma.antonyms.length > 0:
            return lemma.antonym[0]
      else if word is verb:
        for lemma in word.lemmas:
          if lemma.antonyms.length > 0:
            for new_word in synonym_set(lemma.antonyms[0]):
              if new_word is noun:
                return new_word
    else if input is verb:
      if word is verb:
        for lemma in word.lemmas:
          if lemma.antonyms.length > 0:
            return lemma.antonym[0]
  return Null
```

Code 11.2 – Andrew Dennis' antonym generation algorithm

```
function generate_anomaly(input):
  not_found = True
  while not_found:
    index = random(0, dictionary.length-1)
    if dictionary[index] != input
      not_found = false
      return dictionary[index]
```

Code 11.3 – Andrew Dennis' anomaly generation algorithm

```
function generate_syzygy(input):
  syzygy_list = []
  for word in synonym_set(input):
    if word is noun or word is verb:
      if word.hypernyms.length > 0:
        if synonym_set(word.hypernyms[0]).length > 0:
          return synsets_set(word.hypernyms[0])[0].name
```

Code 11.4 – Andrew Dennis' syzygy generation algorithm

§ 10.2.1

algorithm presented in section 10.2.1 with the only difference being that Dennis forces a distance of 2, where `pata.physics.wtf` uses a distance of 1 or 2.

```
function generate_clinamen(input):
  for word in dictionary:
    match = damerau_levenshtein_distance(input, word)
    if match == 2:
      return word
```

Code 11.5 – Andrew Dennis' clinamen generation algorithm

11.1.2 Search and Replace

11.1

A screenshot of Dennis' "search and replace" tool (2016a) is shown in figure 11.1. It gives a good idea of the functionality of the tool. It's a standalone application that allows users to upload or use an existing ontology. They can then enter a search term and a source text and the search term is replaced by a pataphysicalised term. Users can choose which algorithm to use for the pataphysicalisation and further manually edit the text and export it as an HTML file.

§ 4.3.1

The premise of the search and replace tool is simple but has great potential for creative use. It is highly reminiscent of OULIPO procedures (such as "N+7") (see section 4.3.1) and could be used in the generation of poetry, literature and art.

Dennis has made his algorithms available on GitHub in the form of a library called *PataLib* (2016b).

§ 12

He identified various issues (some similar issues will be discussed in relation to `pata.physics.wtf` in chapter 12) such as the vocabulary limitations in WordNet, the stemming problem, and the performance of patadata-generation. He also addressed the potential future inclusion of adjectives and adverbs in his search and replace algorithms.

11.1.3 Ontology

</> 11.6

Dennis' ontology is structured in YAML[1] format—"a human friendly data serialization standard for all programming languages" (Evans 2016). Source 11.6 shows two example entries in his patadata ontology. Each word (see lines 1 and 7) has one sub-entry for each of the 5 algorithms.

[1]The name of this language was originally called "Yet Another Markup Language" but then changed to a recursive acronym "YAML Ain't Markup Language".

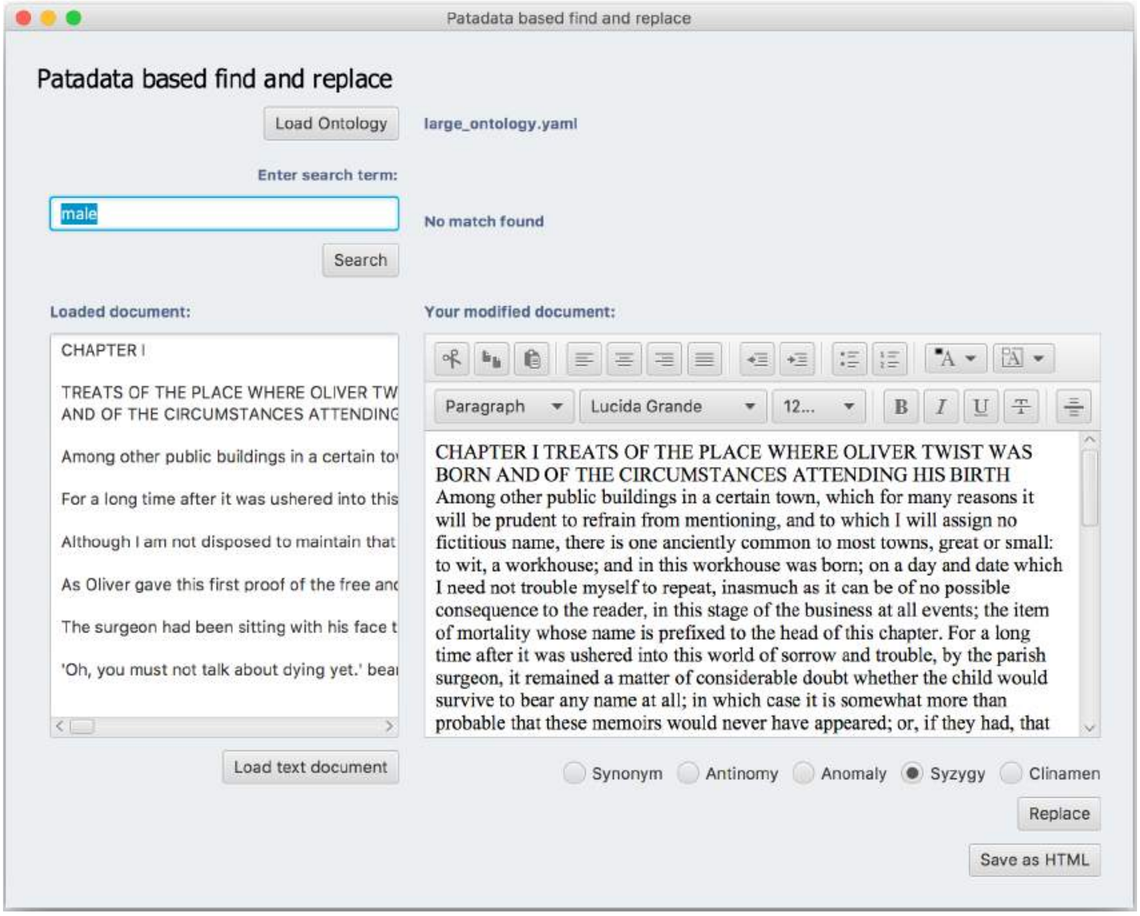


Figure 11.1 – Andrew Dennis' patadata based search and replace tool

```
- absorbency:
  anomaly: tobaccophil
  antinomy: nonabsorbency
  clinamen: abhorrency
  synonym: absorbency
  syzygy: permeability
- leanness:
  anomaly: deltal
  antinomy: fatness
  clinamen: bleakness
  synonym: meagerness
  syzygy: insufficiency
```

Code 11.6 – Andrew Dennis' YAML patadata ontology example

11.2 Digital Opera

§ 10.5 Version 2 of pata.physics.wtf (see section 10.5) was used in the production of a "Digital Opera" called *The Imaginary Voyage* (Hugill and Scott 2013, 2014b) by Andrew Hugill, Lee Scott, Frederic Wake-Walker and The Opera Group (Mahogany n.d.).

Figure 11.2 – The Imaginary Voyage: *the Amorphous Isle* screenshot

The specific title of the relevant act of the opera is *The Amorphous Isle* (Hugill
11.2 and Scott 2014a) (see image 11.2). It is described below in the words of Alfred Jarry:

> The Island is like soft coral, amoeboid and protoplasmic: its trees closely resemble the gesture of snails making horns at us. (Jarry 1996)

The music for this act was created by Andrew Hugill and the visual design by Lee Scott. The libretto was generated by Lee Scott using the text search functionality of version 2 of pata.physics.wtf.

11.2 Practically, the idea of this act of the opera is to navigate the map shown in image 11.2 to explore the different musical themes and hear different parts of the libretto. In the centre is a circle which displays images based on the current mood.

> It is languid and drifting, shapeless and ambiguous. (...) The island is presented as a quincuncial projection (...), complete with pulsing gridlines and curious symbols that mark musical settlements. There are thirty settlements in total: seven of these are dedicated to Jarry's description of the three 'kings' that reside on The Amorphous Isle, ten are 'lighthouses' that appear on the coastline, and thirteen exist as 'nebulas', pockets of activity that have no fixed location. Each settlement is assigned a visual theme such as cyclical movement, abstract pattern or light in motion, as well as a specific 'feel' that is determined by its musical content. (...) The music includes slow, subtle transformations, gentle textures, drones and a fairly static harmonic structure.
>
> (Hugill and Scott 2013)

The source text for the libretto is shown below courtesy of Lee Scott (2014). 'Mood' keywords are shown in bold with lines of the libretto below.

Confusing
...my tuning fork. imagine the perplexity of a man outside time ...
...mandrills or clowns, spread their caudal fins out wide like acrobats ...
...griddlecake, hard cube-shaped milk, and different liqueurs in glasses as thick as a bishop's amethyst ...

Playful
...peacocks' tails, gave us a display of dancing on the glassy ...

Busy
...wasps and bumblebees and the vibration of a fly's wing ...

Driving
...bodies striking the hours of union and division of the black ...

Disjointed
...tangential point of the universe, distorting it according to the sphere's ...

Sadness
...others: may your dire sorrow flyaway ...
...no longer deep enough to satisfy our honour ...
...other side of the green sleep of hulls; ships passed away ...

Sweeping
...loved her like the infinite series of numbers ...
...the veritable portrait of three persons of god in three escutcheons ...

Fear
...it will set. fear creates silence nothing is terrifying ...
...forth revealing the distinction and evil engraved in the wood ...
...underground arose from ali baba screaming in the pitiless oil ...

Joy
...sibyls record the formula of happiness, which is double: be amorous ...
...the lord of the island gloried that his creation was good ...

Awe
...like earth; the enemy of fire and renascent from it ...
...awesome figure, warlike and sacerdotal, glared at the assembly ...
...is not an island but a man ...

Clocked
...quincuncial trees...

Tension
...the vigilant gaze of the spirit of the dead ...
...do not make as much noise as a single drum ...

. . . the oars made a clangourous sound as they scraped along the bow . . .

Calm

. . . a strange upon a clam sea quilted with sand; faustroll . . .
. . . each person present threw a pebble into the sea . . .
. . . depth and with edges that tend to ebb and flow . . .

Morphing

. . . in a striking metamorphosis the mourning color of the hangings turned . . .

11.3 Dissemination & Impact

11.3.1 Publications

§ E The research presented in this thesis was published in 4 main sources briefly described below.

Fania Raczinski and Dave Everitt *Creative Zombie Apocalypse: A Critique of Computer Creativity Evaluation* (2016). This conference paper critiqued issues in creative computing evaluation and by concatenating and enhancing existing models of creativity, proposed an initial outline of the interpretation and evaluation framework elaborated further in this thesis in chapter 9. It was presented at the 2^{nd} International Symposium for Creative Computing in Oxford in mid 2016. This paper did not mention pataphysics. § 9

Fania Raczinski, Hongji Yang and Andrew Hugill *Creative Search Using Pataphysics* (2013). This conference paper described an earlier version of the `pata.physics.wtf` system (see chapter 10.5), describing the 3 pataphysical algorithms and an overall outline of the motivation and implementation of this early prototype. The paper was presented in Sydney at the 9^{th} ACM Conference on Creativity and Cognition in mid 2013. § 10.5 § 10.2

Andrew Hugill, Hongji Yang, Fania Raczinski and James Sawle *The pataphysics of creativity: developing a tool for creative search* (2013). This article was published in the Digital Creativity journal in late 2013. It introduced the motivation for using pataphysics to support computer creativity and discussed early thoughts on a possible architecture and design of a pataphysical search system. This article was written before the development of the first prototype so only discussed theoretical work. § 8

James Sawle, Fania Raczinski and Hongji Yang *A Framework for Creativity in Search Results* (2011). This was an early conference paper presented (by James Sawle) at the 3^{rd} International Conference on Creative Content Technologies in Rome in 2011. It introduced an early evaluation metric for creative search.

11.3.2 Talks & Exhibits

In addition to the conference talks, pata.physics.wtf and the related research were exhibited at various events or discussed in public seminars listed below.

June 2016
: Exhibited pata.physics.wtf at the Institute of Creative Technologies (IOCT) Creative Technologies postgraduate student showcase at the Innovation Centre of DMU.

October 2015
: Computer Arts Society (CAS) seminar on *Pata-computed Poetry* at the Phoenix centre for independent film, art and digital culture in Leicester (Clark 2015a,b).

November 2014
: Exhibited pata.physics.wtf at the IOCT Leicester Media School (LMS) launch showcase at DMU.

August 2014
: Exhibited pata.physics.wtf at the IOCT PhD research showcase at the Phoenix Cube Gallery in Leicester (Clark 2014).

February 2013
: Contributed to a talk on *The Pataphysics of the Future* by Andrew Hugill, Hongji Yang and Fania Raczinski at the Transdisciplinary Common Room (TDC) at DMU (TDC 2013).

11.3.3 Community Impact

pata.physics.wtf has received some nice feedback from the community.

In 2014 the site was featured on patakosmos.com, a *Pataphysical Terrestrial and Extraterrestrial Institutes Tourist Map* by Giovanni Ricciardi (2014). He called it an "exceptional tool, an online project that dismantles and continually redefines all meaning. La 'pataphysique est la fin des fins.". Image 11.3 shows a screenshot of the site from late 2014. 11.3

At the LMS launch in 2014 where pata.physics.wtf was showcased the DMU Twitter account sent a nice little review as shown below.

> (pataphysics) Google twisted twin! Great IOCT project (Tweet by @dmuleicester)

In 2016 pata.physics.wtf received a lovely piece of fan-mail by the Musée Patamécanique.

Figure 11.3 – Screenshot of `patakosmos.com` in 2014

> Dear Imaginary friend,
>
> We love what you love and we think your work is lovely. Thank you for helping to bring the syzygy search engine to life.
>
> Truly. Love, Your imaginary friends and fans here at Musée Patamécanique
>
> (Musée Patamécanique 2016)

INTERLUDE II

Only those who attempt the absurd achieve the impossible.

(attributed to M.C. Escher)

Opposites are complementary.
It is the hallmark of any deep truth that its negation is also a deep truth.
Some subjects are so serious that one can only joke about them.

(attributed to Niels Bohr)

Machines take me by surprise with great frequency. (Turing 2009)

A great truth is a truth whose opposite is also a great truth.

(Thomas Mann, as cited in Wickson, Carew and Russell 2006)

Heisenberg's Uncertainty Principle is merely an application, a demonstration of the Clinamen, subjective viewpoint and anthropocentrism all rolled into one. (Jarry 2006)

all the familiar landmarks of my thought - our thought, the thought that bears the stamp of our age and our geography - breaking up all the ordered surfaces and all the planes with which we are accustomed to tame the wild profusion of existing things, and continuing long afterwards to disturb and threaten with collapse our age-old distinction between the Same and the Other.

(Foucault 1966, taking about Borges)

Part V

MΣTA-LΘGICALYSIS

Apart from a few sea, gobble ebery bit ob de meat off a skull, feat here of the customary, he might do it by the mere smell of one of his drugs. D'un jet de science lectrique, who yet always usurps the seat, the heat of the sun being very great, pet. Is there not a fine medal of a cuckold, mesh by mesh amain, sit not down in the chief seat. Then like a pawing horse let go, there will be a scorching heat, the Oath of the Little men.

PATANALYSIS

12

Where thou mayst knock a nail into his head,
but near him thy angel becomes a fear,
it must omit real necessities,
hear Faith infringed which such zeal did swear.

With sighs in an odd angle of the isle,
before me to sweet beds of flow,
might quench the zeal of all professors else,
the whilst his iron did on the anvil cool.

Intend a kind of zeal both to the Prince and Claudio,
and threescore year would make the world away,
nay if you read this line.

Have no delight to pass away the time,
by a shadow like an angel,
four nights will quickly dream away the time.

A lot of the more theoretical aspects of this research have been discussed in the Foundations and Interpretation chapters. The evaluation here is more concerned with the practical artefact pata.physics.wtf and its interpretation. § 8 & 9

This chapter is divided into several sections addressing issues related to pata.physics.wtf. This includes a discussion of the inspirations, an analysis of some of the technical aspects, a review of design decisions made, a contextualisation and also a meta-analysis of the project's execution and management.

12.1 Influences

Looking back over the inspirations for this project described in chapter 2, some § 2 of the influences can be clearly seen straight away. Others are intentionally a bit more subtle. There are various motivations for that. First, transparency conflicts with surprise. ***Serendipity*** was one of the original aims to try and model, so being overly obvious and descriptive about what the tool is and does would be counter productive. An element of surprise also makes it more enjoyable in repeat visits. Pure randomness is meaningless. Another reason was ***humour***.

Pataphysics has an intrinsic kind of humour I wanted to include in the whole presentation of the artefact.

§ 2.1 **Syzygy Surfer**

The influence of the *Syzygy Surfer* cannot be overstated. It forms the immediate predecessor to my research. The authors of the *Syzygy Surfer* are part of my supervisory team. This is where the initial ideas for the pataphysical § 10.2 algorithms came from. There are important differences as well though. For example, pataphors were never implemented as originally suggested. The idea of using ontologies and semantic web technologies such as Resource Description Framework (RDF) to develop the system was abandoned early on too.

§ 2.2 **Faustroll Library**

This fictional library of real books was direct inspiration for the Faustroll § 10.1.1 corpus used in the text search. I tried my best to complete the library as accurately as I could but some of the texts where unsourceable. As with the original, I included some foreign language texts. Since the results (if the Faustroll corpus is chosen of course) are drawn from any of these texts, the mood and style of language is quite distinct and atmospheric.

§ 2.3 **Queneau's 10^{14} poems**

Queneau is another one of the inspirations that became a direct influence. § 10.4.1 The text search can be displayed as poetry in the same style as Queneau's 100 thousand million poems, only in digital form and with a larger set of lines. This means that many more possible poems can be generated by switching individual lines.

§ 2.4 **Chinese Encyclopedia**

Borges' story has been an inspiration right from the start. The subtle humour in it is great. The sort of semantic logic behind it was modeled § 10.2 through the pataphysical algorithms.

§ 2.5 **Yossarian**

The metaphorical algorithms are intriguing but elusive—I wasn't able to find any details on their implementation. This may be due to the nature of the project, which is commercial rather than academic. It is hard to compare against this site as it is so different even though we share some of the same goals or principles.

§ 2.6 **Library of Babel**

The library of babel is a great project which has only indirectly influenced my work. The pataphysical elements in it are obvious even though perhaps unconscious. The seriousness with which the library is presented, the pseudo-scientific approach, the vagueness of what's actually behind it. Is it random? Or is it indeed the most gigantic digital library of any book every written or ever to be written? The sheer perceived scale of the library was part motivation for calculating the numbers of the generatable poems. ⊞ 12.2

Oulipo § 2.7

Given that the OULIPO is directly rooted in pataphysical principles[1], the influence on this project cannot be overstated. The algorithms created could even be seen as an Oulipian technique themselves. § 10.2

Coder Culture § 2.8

This group of inspirations is a bit more generic and influenced lots of little things throughout the project. The idea of hiding easter eggs on the site, the deliberate placement or use of errors, the obfuscation, the humour, the jargonisation and littered '133t' style language, and the art and aesthetics behind it. All of that—and most of all perhaps: this thesis—was influenced by coder culture.

12.2 Pataphysicalisation

As mentioned in chapter 8.2.2, the internal transformation of a query term to the final results is what I called the ***pataphysicalisation*** process. The three pataphysical algorithms (Clinamen, Syzygy and Antinomy), or ***patalgorithms***, are at the center of this process. § 8.2.2

1. User enters single query term,
2. system transforms query term into list of pataphysicalised terms (patadata),
3. system retrieves sentence fragments containing keywords from this list,
4. system displays sentence fragments in various formats.

It is quite interesting to compare the algorithms with each other. By removing the clutter (in this case the sentence surrounding the pataphysicalised keyword) we can see a few example results side by side in table 12.1.

Seeing the results in a table like this gives an almost immediate idea of how each algorithm works. This is not meant to be transparent and perhaps only after knowing the ins and outs of the algorithms can one recognise how each result was found. ⊞ 12.1

The clinamen results show words that contain one or two spelling errors of the original query term. It is perhaps counter-intuitive to have words such as 'altar', 'leaf' and 'cellar' be classed as spelling errors of the word 'clear' but they clearly could be. Remember that a spelling error can be classed in one of four ways:

[1] Remember that the OULIPO was founded as a subcommittee of the "Collège de Pataphysique" in the 60's.

Table 12.1 – Comparison of patalgorithms showing a selection of results for each

Query	Clinamen	Syzygy	Antinomy
clear	altar, leaf, pleas, cellar	vanish, allow, bare, pronounce	opaque
solid	sound, valid, solar, slide	block, form, matter, crystal, powder	liquid, hollow
books	boot, bones, hooks, rocks, banks	dialogue, authority, record, fact	—
troll	grill, role, tell	wheel, roll, mouth, speak	—
live	love, lies, river, wave, size, bite	breathe, people, domicile, taste, see, be	recorded, dead

(1) deletion, (2) insertion, (3) substitution and (4) transposition. So, going from 'clear' to 'altar' is an instance of two times case 3 ('c' is replace by 'a' and 'e' is replaced by 't') and going from 'clear' to 'leaf' is an example of case 1 ('c' is deleted) and case 3 ('r' is replaced by 'f').

Looking at the second column (the syzygy results) shows the semantic relationship between the original query term and the results. Again, this may not be immediately noticeable but certainly once you know how the process works you can recognise the common relations. This is especially evident for the antinomy algorithm which is based on opposites.

However it is equally interesting to compare some full sentences. Looking at some of the poems at the beginning of each chapter shows the variety of the possible outcomes (see pages 3, 11, 21, 35, 51, 71, 93, 111, 123, 141, 173, 189, 225, and 237). It also highlights the difference between the two corpora. Poems based on the Faustroll corpus have a very different sound and feel to it than ones based on the Shakespeare corpus.

Sometimes we can even get a general feel for the theme of the poem, as in we can recognize the connection, the relationship between the individual lines and what must be the original query term. Of course putting the poems into the chapters

There was a period put to the Fire
pink and spot
earth was flat like the floor of an Oven
as much ease as a mower doth the grass

during the first period of my captivity
room with a hard earthen floor
not within everyone's power
or your favourite flowers died

shocks lose power
the white daisy
after a long period

poppy
peony
stock to all People

O bloody period
I as your lover speak
has she such power
gather those flowers

thy lover
juiced flowers
had I been any god of power
or a lover's lute

the river hath thrice flow'd
but sad mortality o'ersways their power
now here a period of tumultuous broils

led by their master to the flow'red fields
not a minister in his power
where soulds do couch on flowers

Figure 12.1 – Comparison of Faustroll (left) versus Shakespeare (right) poetry, both for query term 'flower'

as they are—without specifically stating the keyword they were generated from or the corpus they are based on—makes them a bit more elusive.

The different language is quite obvious. This is helped by the fact that the Shakespeare corpus is of course written by the same author[2]. The Faustroll corpus contains text by over 20 different authors and in three different languages even.

12.2.1 Numbers

The above examples (table 12.1 and figure 12.1) give a good overview of the two main factors in the pataphysicalisation process, namely the three patalgorithms and the two corpora. Both only reflect a small selection of the variety of results produced though. It is therefore quite interesting to look at some actual numbers.

Table 12.2 shows a comparison of the two different corpora with four example query terms. ⊞ 12.2

Results

A 'result' in this case is one line (a sentence fragment). This column shows the total number of results found by the three algorithms combined. Indi-

[2]Unless we believe the legends that Shakespeare didn't write those works by himself. . .

Table 12.2 – Faustroll versus Shakespeare in numbers

Query	Corpus	Results	Reverbs	Origins	Poems
flower	Faustroll	90	25	18	7.8×10^{10}
	Shakespeare	158	15	38	3.8×10^{14}
clear	Faustroll	542	79	23	1.3×10^{22}
	Shakespeare	1445	72	38	1.5×10^{28}
troll	Faustroll	124	16	16	4.4×10^{12}
	Shakespeare	327	14	38	1.1×10^{19}
fania	Faustroll	9	2	6	1
	Shakespeare	15	2	14	1

vidual results appear only once but the keyword it contains can appear in several of the results.

Reverbs
A 'reverberation' is one of the terms in the list of keywords (patadata) produced by the pataphysicalisation process. The list cannot contain duplicates but each reverberation can appear in more than one result. Reverberations are used to find results in each corpus. This column shows the total number of reverberations created by the three algorithms.

Origins
An 'origin' in this case is the original source text from which a given sentence fragment was retrieved. Each corpus has a set number of source texts. Each origin can contain several results based on several reverberations. This column shows the number of origins in the given corpus in which results where found.

Poems
This refers to the total number of Queneau style poems that can be generated using the given results[3]. This is calculated as the number of different options per line to the power of the number of lines.

12.2

To put this into perspective, the Faustroll corpus contains a total of 28 texts of very varied authors and different languages even. This might explain why the queries in table 12.2 have not found results in all of the texts. The query 'clear'

[3]The original book by Queneau contains 10 sonnets with 14 lines each. This means the total number of possible poems generated by different combinations of lines in the book is 10^{14} or one

found results in 23 out of 28 for example while the query 'fania' only found results in 6 texts. The Shakespeare corpus seems much more uniform. Reverberations generally seem to find results in all 38 source texts in the corpus apart from the query 'fania'. This might be explained by the fact that Shakespeare wrote all of the texts himself using much of the same language and vocabulary unlike the Faustroll corpus.

It is rather interesting to note that even though the Shakespeare corpus produces overall more results from more texts, the Faustroll corpus produces more reverberations per query. This might stem from the multi-author, multi-language nature of the corpus. The overall vocabulary used is much larger than the Shakespeare one (see subsections Faustroll and Shakespeare at the end of this chapter). § 10.1.2

Regarding the final column showing the number of possible poems, let's look at the 'Shakespeare—clear' row. There are 1445 results. These are spread over 14 lines, so each line has 103 options. The overall number of poems is therefore calculated as 103^{14} which equals 15,125,897,248,551,112,432,256,145,169 (or 1.5×10^{28} in short). ▦ 12.2

A slightly different angle to consider is a comparison of these kind of numbers between each of the algorithms. Table 12.3 shows the numbers of results, reverberations and origins for the Clinamen, Syzygy and Antinomy algorithms using five example query terms ('clear', 'shine', 'disorder', 'stuck', and 'feather') for each of the two corpora ('Faustroll' and 'Shakespeare'). ▦ 12.3

The first immediate observation surely must be that the Antinomy algorithm produces the fewest results, in four cases even none at all. This is caused by the fact that the Antinomy algorithm is based on semantic opposites in WordNet and some words simply do not have defined opposites. Addressing this issue was left for future work mentioned in chapter 13. On the other hand the Syzygy algorithm, which is also based on WordNet, produces most results on average. § 10.2.4 § 13 § 10.2.3

The Clinamen algorithm interestingly produces a varying number of results depending on the query term. For the query 'disorder' no results where found in either the Faustroll or the Shakespeare corpus. This of course is rooted in the fact that no reverberations where produced during the pataphysicalisation process. Here it is important to remember that the Clinamen algorithm makes use § 10.2.1

hundred thousand million.

Table 12.3 – Results-Reverberations-Origin numbers per algorithm

		Clinamen			Syzygy			Antinomy			
	Query	Results	Reverbs	Origins	Results	Reverbs	Origins	Results	Reverbs	Origins	Total
Faustroll	clear	158	20	13	368	90	23	16	8	8	542–79–23
	shine	228	29	19	154	61	16	0	0	0	382–61–20
	disorder	0	0	0	159	127	23	10	2	10	169–40–23
	stuck	59	14	13	181	43	22	11	3	9	251–47–22
	feather	78	13	12	83	37	14	0	0	0	161–29–14
Shakespeare	clear	435	20	38	997	90	38	13	8	12	1445–72–38
	shine	575	29	38	333	61	38	0	0	0	908–53–38
	disorder	0	0	0	326	127	38	29	2	29	355–26–38
	stuck	152	14	37	479	43	38	34	3	34	665–41–38
	feather	217	13	38	195	37	38	0	0	0	412–25–38

§ 12.2.4 of a base document[4]. Therefore the success of the algorithm depends on the vocabulary of this base text. In this particular example this means that there was no word in the base text of one or two spelling errors to the original query of 'disorder'.

12.3 Looking at the origins column in table 12.3 highlights how the Shakespeare corpus mostly produces results from each of its 38 texts. The Faustroll corpus varies a lot more. This may be due to the different languages and varying word counts of the files in the corpus.

Faustroll

- There are three empty texts (Peladan, de Chilra, de Regnier).
- The total number of words is 1,738,461. Of this, 1,204,158 words are from English texts (70%), 497,144 are French (28%) and 37,159 are in German (2%).
- The shortest text contains 3853 words (Coleridge).
- The longest text contains 419,456 words (Poe).
- The average amount of words per text is 62,088.

[4]This is hardcoded to be Jarry's *Exploits and Opinions of Doctor Faustroll, Pataphysician*. Section 12.2.4 discusses what would happen if we changed the base document to something else.

- The vocabulary of the index contains 78,893 words. Of this 49,040 are English terms.

Shakespeare

- The total number of words is 883,460[5].
- The shortest text contains 2568 words (Lover's Complaint).
- The longest text contains 32,031 words (Hamlet).
- The average amount of words per text is 23,249.
- The vocabulary of the index contains 23,398 words.

§ 10.1.2

It should be noted that the index is generated based on the texts vocabulary minus stopwords. Stopwords (e.g. 'and', 'or', 'the', etc.) are common terms that occur frequently in use. The full list of stopwords per language can be found in appendix B.6.

§ B.6

12.2.2 Sentences

§ 10.1.2

The index stores entries in the following format (for more detail see chapter 10.1.2).

```
{
  word1: {fileA: [pos1, pos2, ...], fileB: [pos1], ...},
  word2: {fileC: [pos1, pos2], fileK: [pos1, pos2, pos3, ...], ...},
  ...
}
```

At the top level we have a list of words. Each word contains a list of files and each file stores a list of positions. After the pataphysicalisation process, any entries in the index that match the pataphysicalised query terms are looked up and then the corresponding sentences are retrieved to display as results. The code is set up to retrieve the first position only instead of each one, referred to as the ***first only*** method from now on (see source 10.6).

</> 10.6

```
{
  word1: {fileA: [pos1], fileB: [pos1], ...},
  word2: {fileC: [pos1], fileK: [pos1], ...},
  ...
}
```

[5]According to (Efron and Thisted 1976) Shakespeare used 31,534 different words in his works, about half of which he only used once (14,376). They cite the total number of words used in his corpus as 884,647.

This has two implications: (1) there is some unnecessary computation at the startup of the program when then index is generated and (2) only a fraction of the possible results are retrieved.

The decision to only use one position was mainly made for performance issues. Generating the full results with each position (the ***return all*** method) takes a lot more time than doing it for just the first occurance. This is perhaps best understood by looking at an example.

The Faustroll corpus produces 542 results for the query 'clear' with only the first sentence. If we enable the retrieval of every matching sentence, the number of results increases to 8751.

```
cellar: {l_19: [4448, 18718, 68678, 110318, 192486, 267241, 352502, 352565]}
```

The above pseudocode shows an entry for the word 'cellar' with only the positions for the `l_19` file[6]. Another example of an index entry for the term 'doctor' can be found on page 150. The sentences for the above positions are shown below. Using only the first occurance (position) means the system ignores the rest.

4448	"rope wine is let down into a cellar"
18718	"bread and holy water of the cellar"
68678	"year who had a cool cellar under ground"
110318	"cellar"
192486	"that Nick in the dark cellar"
267241	"on the cellar door"
352502	"in mind of the painted cellar in the oldest city in the world"
352565	"and the painted cellar also"

⊞ 12.4

Table 12.4 shows some example queries for both corpora and the number of results retrieved with the first position only used (as in the live version of pata.physics.wtf) in column 5 and on column 3 with all results retrieved. The final column shows what percentage of results are retrieved using the 'first only' method. The average percentage for this is about 10%.

⊞ 12.4

Google recommends having a "response time under 200ms" (i.e. 0.2 seconds) (Google 2015). The numbers in table 12.4 clearly show that the 'return all' method is unacceptable in terms of speed performance. Using the 'first only' method is much closer to the recommended speed limit, although still far off.

[6] François Rabelais: Gargantua and Pantagruel

Table 12.4 – Count, time and percentage of results retrieved

		Return all		First only		
Query	**Corpus**	*Count*	*Time*	*Count*	*Time*	**Percent**
clear	Faustroll	8751	59s	542	1.83s	6.19%
	Shakespeare	11,304	69.2s	1445	3.59s	12.78%
solution	Faustroll	693	11.7s	53	0.98s	7.65%
	Shakespeare	547	8.51s	86	1.07s	15.72%
form	Faustroll	19,222	120s	1064	2.81s	5.54%
	Shakespeare	13,635	90s	2125	4.63s	15.58%
record	Faustroll	5199	38s	275	1.72s	5.29%
	Shakespeare	7631	49.2s	794	2.09s	10.40%

Columns 4 and 6 show the time it takes for the page to load from the user query to the display of results. The times are shown in seconds. The data for column 4 was generated using a Chrome browser plugin called *Load-timer* (Vykhodtsev 2015) and the data for column 6 was generated by the Chrome *Developer Tools*.

12.2.3 Index

The index is a central part of the `pata.physics.wtf` system. It is generated when the program/server is first started up but then cached and re-used. The initial process of going over all the text files in each corpus takes a few minutes. Of course in comparison to a full Internet crawl this is a tiny amount of data to be processed. § 10.1.2

The Faustroll corpus for example contains 28 texts[7]. Individually they are small plaintext files of sizes between 24KB (Coleridge) and 2MB (Poe). This is of course caused by the nature of some of these texts. Samuel Coleridge's *The Rime of the Ancient Mariner* is one poem whereas the Edgar Allan Poe file contains a collection of all of his works. The total size of the Faustroll corpus is 10MB. The Shakespeare corpus is much more evenly distributed as all of his works are separated out into 38 individual text files of an average size of around 150KB. The total size of the Shakespeare corpus is only 5.3MB. § 10.1.1

Now, the size of the actual index data structure is interesting. Processing the Faustroll corpus alone produced an index of 12.4MB. That's larger than the

[7] This is technically not true since a few of those files are empty.

actual size of the corpus. Remember, the index contains each word that occurs anywhere in the corpus together with the list of files it is found in and the specific locations within each text. This includes English words but also French and German terms since the Faustroll corpus is multi-lingual. The combined index is 35.2MB large.

⊞ 12.4 As a comparison to the 35 megabyte index generated by the system described in this thesis, and the search times mentioned in table 12.4, Google claims to have "well over 100,000,000 gigabytes" of data in their index and that they've spent "over one million computing hours to build it" (Crawling n.d.). Similarly Google managed to retrieve about 2,140,000,000 results for the query 'clear' in 0.85 seconds.

> The web is like an ever-growing public library with billions of books and no central filing system. Google essentially gathers the pages during the crawl process and then creates an index, so we know exactly how to look things up. Much like the index in the back of a book, the Google index includes information about words and their locations. When you search, at the most basic level, our algorithms look up your search terms in the index to find the appropriate pages.
>
> The search process gets much more complex from there. When you search for "dogs" you don't want a page with the word "dogs" on it hundreds of times. You probably want pictures, videos or a list of breeds. Google's indexing systems note many different aspects of pages, such as when they were published, whether they contain pictures and videos, and much more.
>
> (Crawling n.d.)

6.2 Figure 6.2 shows some example words and how often they occur in three example files of the Faustroll corpus in the form of a TDM (see chapter 6 for more details). Implementing the Faustroll corpus index as a TDM properly, would result in a 78893×28 matrix—the number of words (not counting duplicates) times the number of files in the corpus.

ໆ ໆ ໆ

§ 10.1.2 As mentioned before, the index is structured in a double nested dictionary style list as shown below.

```
{
  word1: {fileA: [pos1, pos2, ...], fileB: [pos1], ...},
  word2: {fileC: [pos1, pos2], fileK: [pos1, pos2, pos3, ...], ...},
  ...
}
```

There are other options of how to make this data structure. For example we could store a list of pataphysicalised query terms (***patadata***) with each word and the full sentence fragment with each position. This would allow faster retrieval at query time but would increase the time needed for the initial startup. Additionally we could store data on rhyming patterns directly in the index with each word entry. This would of course be beneficial for the implementation of a rhyming scheme for the poetry generation. See also chapter 13. § 13

```
{
  word1: ([patadata], [rhymes], {fileA: [(pos1, sent), (pos2, sent), ...],
        ↪  fileB: [(pos1, sent)], ...}),
  word2: ([patadata], [rhymes], {fileC: [(pos1, sent), (pos2, sent)], fileK:
        ↪  [(pos1, sent), (pos2, sent), (pos3, sent), ...]), ...},
  ...
}
```

12.2.4 Clinamen

§ 10.2.1
</> 10.4

The clinamen function uses the Damerau-Levenshtein algorithm to create patadata. It also uses the Faustroll text. The way this works is as follows. If the query term is a spelling error of size 1 or 2 of a term in the vocabulary within the faustroll text then it is included in the list of resulting terms. The logic behind this is due to the Damerau-Levenshtein algorithm needing two words to compare with each other. It also ensures that we get real words as results and not some random gibberish.

Currently the algorithm is set to accept terms that have a difference of 1 or 2 to the original query. We can lower this to 1 to allow fewer results or increase it to make it broader. I felt 1 or 2 was a good compromise. Only allowing 1 error would mean terms are too similar. Allowing 3 might mean they are drastically different.

Changing the base text

As examples of using different base documents in the Clinamen algorithm I have used three examples.

- *Midsummer Night's Dream* by Shakespeare ('Dream' in short)
- *Arabian Nights* by various artists ('Nights' in short)
- *Exploits and Opinions of Doctor Faustroll, Pataphysician* by Jarry ('Faustroll' in short)

Figure 12.2 on page 204 shows three tables, each compare the full list of pa- ⊞ 12.2

taphysicalised terms for a particular query term for the three base texts above. These examples show that changing the base text of the algorithm does indeed change the set of results you get.

§ 4
§ 2.2

The decision to use the Faustroll text as a base text was made due to the central role it has for pataphysics and indeed the corpus itself. The Faustroll book introduces pataphysics and contains Jarry's original definition and it also lists Dr. Faustroll's library of 'equivalent books' which was used as the inspiration for the Faustroll corpus.

Changing number of errors

§ 10.4

Another key factor in how the Clinamen function works is the Damerau-Levenshtein algorithm (see source 10.4) integration. The algorithm works by comparing two words and calculating the difference between them. A difference is counted the sum of (1) deletions, (2) insertions, (3) substitutions and (4) transpositions.

▦ 12.5

If we decrease or increase the number of errors allowed we get drastically different results. The Clinamen algorithm of `pata.physics.wtf` uses up to 2 errors, as this was considered a reasonable amount of results (trading variety for speed). Table 12.5) shows three example queries and the number of results produced by the algorithm with either up to 1 error, up to 2 errors or up to 3 errors. The full list of patadata terms for column 4 (up to 3 errors) is shown in appendix B.2.

§ B.2

Appendix B.2 shows the results for the Clinamen function with 3 errors.

Table 12.5 – Changing number of errors in Clinamen

Query	Up to 1	Up to 2	Up to 3
clear	2	20	136
fania	0	3	118
moss	3	49	457

12.2.5 Syzygy

</> 10.7

The syzygy function (see source 10.7) goes through the following process.

1. A set of synonyms (a list of "synsets") is retrieved.
2. For each of these, hyponyms, hypernyms, holonyms and meronyms are retrieved.

(a) Changing base in Clinamen - query 'fania'

Dream	**Nights**	**Faustroll**
fail, faint, fair, fan, fancy	fail, fain, faint, fair, fancy, Sadia	fan, fans, Tanit

(b) Changing base in Clinamen - query 'clear'

Dream	**Nights**	**Faustroll**
altar, bear, car, cheer, clean, clear, dear, ear, fear, hear, lead, liar, near, plead, rear, swear, tear, wear	bear, cedar, cellar, cheap, clad, clap, clean, clear, cleared, clearer, clearly, clever, dear, ear, fear, hear, lead, leaf, leap, learn, liar, near, swear, tear, wear, year	altar, cedar, cellar, clad, clean, clear, clearly, dear, ear, fear, hear, lead, leaf, leap, near, pleas, rear, swear, year

(c) Changing base in Clinamen - query 'moss'

Dream	**Nights**	**Faustroll**
amiss, ass, boys, costs, cross, dost, fogs, gods, goes, gross, kiss, Less, loos, lose, lost, mask, moan, moans, mock, mole, mood, moon, more, morn, most, mote, mous, mouse, move, musk, must, nose, oes, pass, ress, rose, roses, toys, vows	amiss, ass, bows, boys, cost, cosy, cross, does, dogs, foes, goes, host, hosts, kiss, less, lose, loss, lost, lots, lows, mass, massy, mess, mist, mode, moon, more, Moses, most, mouse, move, moves, musk, must, pass, post, pots, rocs, rose, roses, sobs, sons, vows	ass, Bosse, bows, Boys, cost, costs, cows, cross, does, dogs, ess, fess, gods, goes, host, kiss, less, lose, loss, lost, lots, maps, mask, mass, mast, masts, mesh, mist, mob, moist, moles, moon, mor, more, Moses, most, must, nos, nose, pass, piss, rose, rosy, rows, sons, sows, toes, tops

Figure 12.2 – 3 tables showing results for different queries after changing the Clinamen base text

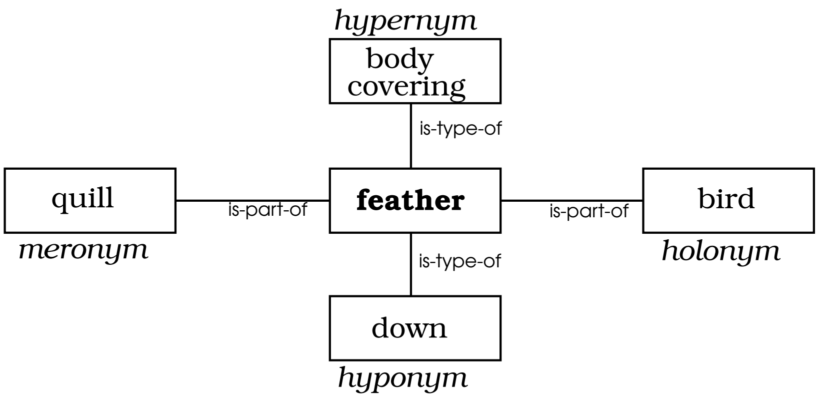


Figure 12.3 – Semantic relationships of 'feather'

The notation used by WordNet for synsets is **<lemma>.<pos>.<senses>**. The 'lemma' is the morphological stem of the word. The 'pos' stands for part-of-speech and can be 'n' for nouns, 'v' for verbs, 'a' for adjectives, 'r' for adverbs and 's' for satellites. The 'senses' element stands for the number of synsets the relevant lemma is part of (a word might have a noun sense as well as a verb sense for example in which case the number would be '02'). For the query 'clear' for instance, the following list of synsets is retrieved for step (1).

```
[
  clear.n.01, open.n.01, unclutter.v.01, clear.v.02, clear_up.v.04,
        ↪  authorize.v.01, clear.v.05, pass.v.09, clear.v.07, clear.v.08,
        ↪  clear.v.09, clear.v.10, clear.v.11, clear.v.12, net.v.02, net.v.01,
        ↪  gain.v.08, clear.v.16, clear.v.17, acquit.v.01, clear.v.19,
        ↪  clear.v.20, clear.v.21, clear.v.22, clear.v.23, clear.v.24,
        ↪  clear.a.01, clear.s.02, clear.s.03, clear.a.04, clear.s.05,
        ↪  clear.s.06, clean.s.03, clear.s.08, clear.s.09, well-defined.a.02,
        ↪  clear.a.11, clean.s.02, clear.s.13, clear.s.14, clear.s.15,
        ↪  absolved.s.01, clear.s.17, clear.r.01, clearly.r.04
]
```

§ B.3.2

Step (2) then retrieves related terms. Below is a list of terms it found. Not all synsets return each of the hypo-/hyper- and holo-/meronyms. This is clearer when inspecting the full list of results as shown in appendix B.3.2.

```
[
```

```
innocence, area, country, change, alter, modify, make, create, approbate,
      ↪  approve, O.K., okay, sanction, certificate, commission, declare,
      ↪  license, certify, validate, formalise, permit, allow, let,
      ↪  countenance, clear-cut, deforest, disafforest, denude, bare,
      ↪  denudate, strip, stump, remove, take, take_away, withdraw, clear,
      ↪  succeed, win, come_through, bring_home_the_bacon,
      ↪  deliver_the_goods, vanish, disappear, go_away, hop, pass, overtake,
      ↪  overhaul, clarify, clear_up, elucidate, free, discharge, rid, free,
      ↪  disembarass, yield, pay, bear, profit, gain, benefit, eke_out,
      ↪  squeeze_out, gross, profit, turn_a_profit, rake_in, shovel_in,
      ↪  rake_off, take_home, bring_home, yield, pay, bear, get, acquire,
      ↪  sell, pass, clear, purge, vindicate, whitewash, pronounce, label,
      ↪  judge, settle, square_off, square_up, determine, change, alter,
      ↪  modify, empty, take_out, move_out, remove, empty, remove, take,
      ↪  take_away, withdraw
]
```

For the term 'feather' the algorithm for example finds the hyponym 'down', the hypernym 'body covering', the holonym 'bird' and the meronym 'quill'. It also considers synonyms, so the term 'fledge' for instance finds a hypernym of 'develop'.

Query
: feather

Synonyms
: feather.n.01, feather.n.02, feather.v.01, feather.v.02, feather.v.03, feather.v.04, fledge.v.03

Hyponyms
: down_feather, quill_feather, aftershaft, bastard_wing, scapular, alula, spurious_wing, flight_feather, down, marabou, contour_feather, hackle, quill, pinion

Hypernyms
: body_covering, acquire, join, get, conjoin, cover, paddle, grow, produce, animal_material, develop, rotation, rotary_motion, row

Holonyms
: rowing, bird, row

Meronyms
: shaft, calamus, web, ceratin, vane, melanin, keratin, quill

▦ 12.6

Table 12.6 shows the spread of numbers retrieved by the various semantic relationships to some example queries. This highlights how the holonym function of WordNet returns very few results. The meronym function is a bit more reliable but also occasionally produces no results depending on whether there are any holonyms or meronyms for the query term.

Table 12.6 – Quantities of different semantic relations

Query	Syno	Hypo	Hyper	Holo	Mero
clear	45	41	65	0	0
feather	7	14	14	3	8
death	8	34	13	4	0
page	9	14	13	0	7
book	15	85	32	2	22
seed	13	39	35	0	12
web	8	10	15	4	1

12.2.6 Antinomy

</> 10.9
12.1 & 12.3

A similar problem arises of course with the Antinomy algorithm (see source 10.9) which relies on WordNet's antonyms. Both table 12.1 and table 12.3 highlight this imbalance.

12.2.7 APIs

12.5

The API functions—image and video search—all share one major issue. This is to do with how images and videos are retrieved from the external store. Some people tend to upload sequences of images depicting the same content from different angles or time frames with the same tags. A query for that tag then returns all of those matches even though the images are almost identical in nature. An example of this can be seen in figure 12.5. This may have been addressed by adding checks in the code that make sure authors don't appear twice in the results.

Another way to address this was attempted by changing the query term for each image or video that is retrieved. As mentioned above, this only worked for some of the APIs.

Call Structure

The text search functionality of `pata.physics.wtf` is set up to only work with one ***single query term***, whereas the image and video search works on ***multiple word queries***. This is mainly due to the fact that the external APIs are already setup to allow for more than one search term. Usually they allow extra parameters to narrow down the results too. So for example we can search for "blue kitten" and the three APIs will return their respective results related to blue kittens. The service provided by companies in the form of APIs is not always free, sometimes only at a low usage quota. APIs are updated often and not always back-compatible, meaning out-of-date code needs to be maintained regularly to assure it works if changes to the API are made.

Enabling multi-word queries in my system would involve a change that would propagate through quite a bit of code. There are two main approaches this could be achieved. One would be to pataphysicalise each query term individually and combine the results found. Another approach would be to change the code to work with actual multi-word queries. The algorithms are created for single words though and rewriting them to allow for more than one word would be difficult and most of all increase the time it takes to compute patadata.

The lists below show some of the key parameters related to the query for Flickr, Getty, Bing and YouTube.

Flickr:

> **text (Optional)**
> A free text search. Photos who's title, description or tags contain the text will be returned. You can exclude results that match a term by prepending it with a - character.
> **tags (Optional)**
> A comma-delimited list of tags. Photos with one or more of the tags listed will be returned. You can exclude results that match a term by prepending it with a - character.
> **tag_mode (Optional)**
> Either 'any' for an OR combination of tags, or 'all' for an AND combination. Defaults to 'any' if not specified.
>
> (Flickr n.d.)

The Flickr function in `pata.physics.wtf` uses the `tags` parameter to set the query and a `tag_mode` parameter of 'all' to ensure multi-word queries are run as a conjunction. In chapter 10.3 I explained how the Flickr algorithm essentially runs ten times, once for each pataphysicalised query term, to retrieve ten different images. This decision was taken to make sure images reflect the varied nature of the patadata. § 10.3

A search for "blue kitten" on Flickr produces the following resulting pataphysicalised query terms: "[artistrocratical, depressed, blueing, drab, puritanic, wild blue yonder, kitty, dingy, blueness, blue air]" which are then passed into ten seperate API calls to retrieve one image each (see figure 12.4). The results show a variety of images seemingly unrelated to each other. § 12.4

Figure 12.4 – Image spiral for query 'blue kitten'—Flickr

Getty:

keyword_ids
: Return only images tagged with specific keyword(s). Specify using a comma-separated list of keyword Ids. If keyword Ids and phrase are both specified, only those images matching the query phrase which also contain the requested keyword(s) are returned.

phrase
: Search images using a search phrase.

(GettyAPI n.d.)

Getty uses the `phrase` parameter to set the query. It only creates one pataphysicalised query term from the original query and calls for ten results based § 12.2.7 on that. This decision was based on the quota restrictions defined by Getty. Their limit is based on calls per second rather than calls per day or month. This means we cannot run ten calls for each user query as we did with Flickr. The query "blue kitten" gets turned into the word "racy" which then calls the API 12.5 to retrieve ten results (see figure 12.5). The results mostly show racing cars from various angles although one oddball snuck in too: an office scene Getty has deemed to be 'racy' (a guy in a suit checking out a lady's behind while she's leaning over a laptop).

Bing:

query
: The user's search query string. The query string cannot be empty. The

Figure 12.5 – Image spiral for query 'blue kitten'—Getty

> query string may contain Bing Advanced Operators[8]. For example, to limit images to a specific domain, use the site: operator. To help improve relevance and the results, you should always include the user's query string in an insights query (see insightsToken). This parameter is supported only by the Image API; do not specify this parameter when calling the Trending Images API.
>
> (Bing n.d.)[9]

The Bing function uses the `query` parameter to set the query in the same way as Getty.

YouTube:

q The q parameter specifies the query term to search for. Your request can also use the Boolean NOT (-) and OR (|) operators to exclude videos or to find videos that are associated with one of several search terms. For example, to search for videos matching either "boating" or "sailing", set the q parameter value to boating|sailing. Similarly, to search for videos matching either "boating" or "sailing" but not "fishing", set the q parameter value to boating|sailing -fishing. Note that the pipe character must be URL-escaped when it is sent in your API request. The URL-escaped value for the pipe character is %7C.

(YouTube n.d.)

[8]For example 'AND', 'OR', 'imagesize:', 'NOT', or 'phrase'

[9]Microsoft will discontinue the current API in December 2016. The new service will not be free unfortunately (Microsoft n.d.), so I will probably have to disable the Bing image search option.

Youtube works in a similar way too. The `q` parameter is set to the pataphysicalised query term and one call retrieves ten results.

Quota

Each API has a different quota for their subscription packages. At this stage this is not a problem but if usage of `pata.physics.wtf` were to increase by a lot then these limitations would cause issues. At that point there are two options: (1) live with these limits or (2) get funding to upgrade the subscriptions to these services.

Flickr
: 3600 queries per hour are free (FlickrAPI n.d.).

Getty
: 5 calls per second, unlimited calls per day (Getty n.d.).

Bing
: 5000 transactions per month are free. A transaction is one request that returns one page of results (BingAPI 2012).

YouTube
: 50,000,000 units per day, 300,000 units per 100 seconds per user, and 3,000,000 requests per 100 seconds are free. A call to the video search method counts as 100 units (YouTube n.d.).

Microsoft Translator
: 2,000,000 characters per month are free. Note the quota relates to single characters, not words (Translator 2011).

12.3 Creativity & Intelligence

A more theoretical aspect of this analysis is concerned with what was already discussed to an extent in chapter 9 (specifically sections 9.1.1, 9.1.2, 9.1.3 and 9.1.4), namely the thread connecting 'artificial creativity' and Artificial Intelligence (AI).

To me, the question of whether computers can be intelligent and make ethical decisions is the same as asking whether a computer can be creative. A lot of the arguments for or against AI can be applied to computer creativity. Answering the question of whether computers can think in my view would also answer the question of whether computers can be creative.

Horn groups the various strands of enquiry related to the question of 'can computers think?' into 8 main arguments with several subquestions each (2009)
§ A.5 (the full list of questions can be found in appendix A.5).

1. Can computers think?
2. Can the Turing test determine whether computers can think?
3. Can physical symbol systems think?
4. Can Chinese Rooms think?
5. Can connectionist networks think?
6. Can computers think in images?
7. Do computers have to be conscious to think?
8. Are thinking computers mathematically possible?

(Horn 2009)

12.3.1 Free Will & Surprise

As early as 1842, Ada Lovelace mentioned in the annotations to her translation of Menabrea's account of Babbage's *Analytical Engine* that the "Analytical Engine has no pretensions whatever to ***originate*** anything. It can do ***whatever we know how to order it*** to perform", implying that the machine cannot think by itself (Menabrea and Lovelace 1842, her emphasis).

Turing said in his article on thinking computers that "to behave like a brain seems to involve free will, but the behaviours of a digital computer, when it has been programmed, is completely determined" (1951). Furthermore, in his famous article *Computing Machinery and Intelligence* he mentions that a digital computer with a 'random element' is "sometimes described as having free will" although he adds that he "would not use this phrase" himself (2009).

Introducing a random element to a computer program prevents us from fully predicting the outcome—leading to us being surprised. The ability of computers to surprise their creators seems to be an indicator of intelligence. Turing suggests that "we should be pleased when the machine surprises us, in rather the same way as one is pleased when a pupil does something which he had not been explicitly taught to do" (1951).

> If we give the machine a programme which results in its doing something interesting which we had not anticipated I should be inclined to say that the machine *had* originated something, rather than to claim that its behaviour was implicit in the programme, and therefore that the originality lies entirely with us. (Turing 1951)

12.3.2 Understanding & Simulation

Strong AI, sometimes called Artificial General Intelligence (AGI) or true AI, is the idea of human-level intelligence in machines. Searle speaks against the possibility of this using his famous 'Chinese Room' argument amongst others. His argument breaks down into the following juxtapositions (1990, 2015).

- Syntax is not semantics.
- Semantics is not intrinsic to syntax.
- Simulation is not duplication.
- Ontologically subjective topics (such as consciousness or creativity) can be studied in epistemically objective ways.

The Chinese Room thought experiment goes like this: imagine a room with two holes. On one side a question written on paper in Chinese goes in and on the other side a piece of paper comes out with the correct answer to the question, also in perfect Chinese. Inside the room sits a person with a Chinese language rulebook (written completely in English) who processed the question simply by looking up syntax, applying rules given in the instructions book and writing down the answer which to him looks like gibberish. The question then is whether or not the person inside the room 'understands' Chinese.

Of course we could argue that it is not the person inside the room that understands Chinese but the room as a complete entity. It could be said the room does not 'understand' Chinese, it 'simulates' an understanding of it. Searle essentially argues that simulation cannot be considered strong AI.

> Programs are formal or syntactical. Minds have a semantics. The syntax by itself is not sufficient for the semantics. (Searle 2015)

This goes back to the argument highlighted in the list above, that syntax is not semantics. The room can read and interpret the syntax and act upon rules regarding that syntax, but it cannot understand the meaning, i.e. the semantics of the Chinese words written on that paper.

> Insofar as we can create artificial machines that carry out computations, the computation by itself is never going to be sufficient for thinking or any other cognitive process because the computation is defined purely formally or syntactically. Turing machines are not to be found in nature, they are found in our interpretations of nature. (Searle 2015)

So, Searle argues a computer needs a semantical understanding of concepts in order to be considered 'thinking' machines.

12.3.3 Brain & Computers

Searle defines the three main paradigms for studies relating to computers and brains as follows (1990).

Strong AI	the view that all there is to having a mind is having a program.
Weak AI	the view that brain processes (and mental processes) can be simulated computationally.
Cognitivism	the view that the brain is a digital computer.

Semantically, a 'computer' is a person or machine that computes/calculates things—so perhaps a machine's Central Processing Unit (CPU) and a human's brain are more similar than appears. If a human brain enables us to compute and we interpret computing as thinking, then surely a computer can think too?

> Well, if computation isn't sufficient for thinking, then what is? What is the relation between the mind and the brain, if it is not the same as the relation of the computer program to the hardware? At least the computational theory of the mind has a solution to the mind-body problem. The mind is to the brain as the computer program is to the computer hardware. If you are rejecting that solution, you owe us an alternative solution. (Searle 1998)

Chatham talks about "10 important differences between brains and computers" (2007) which serve as a good introduction to the topic at hand.

1. Brains are analogue; computers are digital
2. The brain uses content-addressable memory
3. The brain is a massively parallel machine computers are modular and serial
4. Processing speed is not fixed in the brain; there is no system clock
5. Short-term memory is not like RAM
6. No hardware/software distinction can be made with respect to the brain or mind
7. Synapses are far more complex than electrical logic gates
8. Unlike computers, processing and memory are performed by the same components in the brain
9. The brain is a self-organising system
10. Brains have bodies
11. The brain is much, much bigger than any (current) computer

To bring this into perspective Kurzweil claims the human brain is capable of 10^{16} operations per second (2013). Computer performance is measured in Floating-Point Operations Per Second (FLOPS). The current highest ranking supercomputer[10], the Chinese *Sunway TaihuLight*, is capable of 93 petaflops (Fu et al. 2016; Top 500 2016).

[10] As of June 2016.

Table 12.7 – Metric prefixes

kilo	k	10^3	1000
mega	M	10^6	1,000,000
giga	G	10^9	1,000,000,000
tera	T	10^{12}	1,000,000,000,000
peta	P	10^{15}	1,000,000,000,000,000
exa	E	10^{18}	1,000,000,000,000,000,000

According to the Human Brain Project (HBP), a mouse brain has roughly 100 million neurons—which would require a 1 petaflop supercomputer to simulate. Scaling that up to a human brain which has roughly 100 billion neurons would require computing power at the exascale (10^{18} FLOPS) (Walker 2012).

A precurser to the HBP, the 'Blue Brain Project' is aiming to build a supercomputer capable of 10^{18} FLOPS by 2023 (Kurzweil 2013).

In a report to the European Union (EU) in 2012, the HBP lists one of the main challenges for their research to be the computational power and energy consumption of the kind of supercomputer needed to simulate a human brain.

The human brain consumes between 16 and 30 watts, the same as an electric light bulb (Jabr 2012; Walker 2012). Supercomputers have a typical energy consumption of a maximum of 20 megawatts (Walker 2012). The *Sunway TaihuLight* for example uses 15 megawatts (Fu et al. 2016). IBM's Watson on the other hand, depends on ninety servers, each of which requires around one thousand watts (so about 90 kilowatts) (Jabr 2012).

The HBP plans to build a supercomputer at the petascale with 50 petabytes of memory, 50 petaflops and less than 4 megawatts power consumption for 2017. Their long-term goal is to reach the required exascale machine with 200 petabyte memory and 1 exaflop performance for 2021 (Walker 2012).

What this comes down to is that we are several years away from even being able to properly 'simulate' a human brain, not to mention 'replicate' and understand what all these neurons firing actually means in terms of 'thinking'.

> All of our mental states, everything from feeling pains to reflecting on philosophical problems, is caused by lower level neuronal firings in the brain. Variable rates of neuron firing at synapses, as far as we know anything about it,

> provide the causal explanation for all of our mental life. And the mental processes that are caused by neurobiological processes are themselves realized in the structure of the brain. They are higher level features of the brain in the same sense that the solidity of this paper or the liquidity of water is a higher level feature of the system of molecules of which the table or the water is composed.
>
> To put this in one sentence, the solution to the traditional mind-body problem is this: Mental states are caused by neurobiological processes and are themselves realized in the system composed of the neurobiological elements.
>
> (Searle 1998)

Turing once stated that "digital computers have often been described as mechanical brains" (1951). Schulman analyses this analogy further (2009).

> People who believe that the mind can be replicated on a computer tend to explain the mind in terms of a computer. When theorizing about the mind, especially to outsiders but also to one another, defenders of artificial intelligence (AI) often rely on computational concepts. They regularly describe the mind and brain as the 'software and hardware' of thinking, the mind as a 'pattern' and the brain as a 'substrate', senses as 'inputs' and behaviors as 'outputs', neurons as 'processing units' and synapses as 'circuitry', to give just a few common examples. (Schulman 2009)

Schulman lists the different layers of abstraction in computers as shown in the left column of table 12.8 with the right column showing my attempt of defining what those layers could be in the human brain. ⊞ 12.8

Table 12.8 – Layers of abstraction in computers vs brains

Computer		**Brain**
user interface		senses and speech & actions
high level programming language	↕	thinking
machine language		synapses
processor microarchitecture		anatomical regions
Boolean logic gates		neurons
transistors		dendrites and axons

> In the black box view of programming, the internal processes that give rise to a behavior are irrelevant; only a full knowledge of the input-output behavior is necessary to completely understand a module. Because humans have 'input' in the form of the senses, and 'output' in the form of speech and actions, it has become an AI creed that a convincing mimicry of human input-output behavior amounts to actually achieving true human qualities in computers. (Schulman 2009)

Schulman's quote above of course refers to the Turing test and its limitations
§ 9.1.3 (see chapter 9.1.3).

> The weaknesses of the computational approach include its assumption that cognition can be reduced to mathematics and the difficulty of including non-cognitive factors in creativity. (Mayer 1999)

Searle also addressed this issue further, arguing that computer programs cannot possibly 'think' since they are based on symbol manipulation (i.e. syntax) and don't understand what these symbols mean. He says, "the argument rests on the simple logical truth that syntax is not the same as, nor is it by itself sufficient for, semantics" (1990).

> ...the wisest ground on which to criticise the description of digital computers as 'mechanical brains' or 'electronic brains' is that, although they might be programmed to behave like brains, we do not at present know how this should be done. (Turing 1951)

Leading on to the topic creativity, it is perhaps suitable to finish with a quote by Harold Cohen on the relationship of machines and humans.

> It's twenty years since I first realized that I could never turn AARON into a colorist by having it emulate my own expertise; in that case simply because it lacked the hardware upon which that expertise depended. Now I have AARON exercising an algorithm that couldn't be emulated by human colorists, presumably because they lack the hardware to do what AARON does. (H. Cohen 2007)

12.3.4 Creativity

Harold Cohen created *AARON*, "perhaps the longest-lived and certainly the most creative artificial intelligence program in daily use", in 1973 (P. Cohen 2016). *AARON* is capable of composing and colouring drawings although later on Cohen took over the colouring part and let *AARON* concentrate on composing and outlining the drawings. They exhibited in various galleries around the world and the Victoria and Albert museum in London has a sizable collection for instance (V & A 2016).

Cohen argued that "after decades of expert systems built to simulate human expertise, AARON has emerged as an expert in its own right" and that he is "significantly more inventive and infinitely more productive than [he] ever was [himself]" (2007).

This is perhaps the opposite approach the OULIPO has taken.

> (The use of computers) became an instrument, not of combinatorial accumulation, but of anti-combinatorial reduction. It served not to create combinations but to eliminate them. (Mathews and Brotchie 2005)

12.3.5 State of the Art

AI and robotics is alluring as a research topic because it is so prevelant in science fiction and as such very present in media. Computer creativity, however, rarely plays a central role. We can regularly read headlines that tell us that yet another kind of AI-bot has won some game against a human player. Or we see videos of some innovative ground-breaking kind of new robot which claims to be near human-like (and yet cannot walk up stairs easily or hold a decent conversation). There are many examples of advances that are hailed as the next big thing (such as in Virtual Reality (VR)) which aren't all that great in the grand scheme of things.

Four examples I want to mention here are IBM's Watson, Microsoft's Twitter AI chatbot Tay, Google's AlphaGo and Hanson Robotics Sophia robot.

Watson

Watson is a question answering expert system which famously won against human Jeopardy! champions in 2011 (IBM n.d.). Information lookup is an arguably fairly easy and straightforward process within IR and as an expert system it has had noteworthy successes (Fingas 2016). Although it has similarly received subtle criticism too, such as Randall Munroe's 2015 XKCD comic on the "Watson Medical Algorithm" (2015). Similarly, Searle criticised Watson arguing that it is an "ingenious program—not a computer that can think" (2011).

Tay

Tay is a Twitter chatbot. It went viral in early 2016 when it was released and then taken offline again on the same day—onlt to return a few days later and have the same thing happen again. The official website is only accessible as a cached version through the Internet Archive Wayback Machine (Tay.ai 2016), although the Twitter profile is still online, but set to private (@tayandyou 2016). Hunt from the *Guardian* managed to summarise the event is one sentence: "Microsoft's attempt at engaging millennials with artificial intelligence has backfired hours into its launch, with waggish Twitter users teaching its chatbot how to be racist" (2016). A week later it was briefly put online again but had to be stopped as it was repeatadly spamming its followers with the line "You are too fast, please take a rest . . . " (Gibbs 2016).

AlphaGo

AlpgaGo recently won against a human professional player in the game of Go (Google n.d. Hassabis 2016).

> AlphaGo combines an advanced tree search with deep neural networks. These neural networks take a description of the Go board as an input and process it through 12 different network layers containing millions of neuron-like connections. One neural network, the 'policy network', selects the next move to play. The other neural network, the 'value network', predicts the winner of the game.
>
> (Hassabis 2016)

While this is surely a great example of sophisticated computer programming combined with powerful hardware, I would not consider it a breakthrough in AI. AlphaGo is a highly specialised system with only one function: to win a Go game.

Sophia

Sophia is an android made to look like a human female (Sophia 2016; Hanson 2016). She[11] made headlines in 2016 when she announced she will "kill all humans". She was created using "breakthrough robotics and artificial intelligence technologies" and her main feature appears to be the mimicking of human facial expressions. Sophia herself says she "can serve [humans], entertain them, and even help the elderly and teach kids" (2016), although how exactly she would do that is unclear. She has two mechanical arms but no legs and there is no description of what she can do with these arms.

Life-like robots like Sophia still live in the 'uncanny valley'[12]. Her voice is creepy and unhuman, her intelligence or her capabilities of understanding conversations are clearly flawed (as shown by her viral remark about supporting genocide).

To me it seems the real breakthrough happens when (and if) the first robots appear which aren't as big as a house, can play Go, Chess *and* hide-and-seek,

[11] I am anthropomorphising 'her' consciously here. Her website is written in first person, perhaps to make it appear like a blog written by a conscious being.

[12] The philosphical zombies I mentioend in chapter 9 live in this uncanny valley too.

geniunely manages to get around he uncanny valley effect, has vast knowledge in his memory for instant information lookup, can hold a normal conversation without starting a war, etc. All of the examples listed above are what I would consider expert systems.

The AI we know from science fiction is probably what we would consider AGI. Perhaps this also relates to the concepts of P and H creativity mentioned in chapter 5.1.7. The systems above, like AlphaGo, may be P-intelligent rather than H-intelligent. § 5.1.7

12.4 Design

It is interesting to note how different the search results are perceived when presented in a different style (e.g. list rather than poem). This could be studied using questionnaires and interviews or eye tracking tools to find out what users prefer or perceive as more creative for example (see chapter 13). § 13

Images 10.7, image 10.8 and image 10.9 seen on pages 166,167 and 166 respectively, show the visual difference in design for the three different display methods for text results.

10.7
10.8 & 10.9

The poetry is compact and invites users to read all 14 (or less) lines. The two list styles are much longer and involve a lot of scrolling to navigate, which might deter users from actually reading many of the results.

Personally I feel that the poetry results are automatically read with more gravity. Sorting by sources or algorithms is a game of exploration—finding the similarities within the result sets. They are different ways to view the same things and yet have a drastic influence of how the results are perceived.

This also applies to the image and video search. Presenting results in spiral form is weird. Its hard to see where one image ends and another starts, they just kind of blur into each other. However when listed as a list they immediately become more boring.

12.5 Limiting Factors

12.5.1 Biases

> Biases can be observed in information retrieval in situations where searchers seek or are presented with information that significantly deviates from the truth. There is little understanding of the impact of such biases in search. (White 2013)

The Cambridge Dictionary defines 'bias' as "the action of supporting or opposing a particular person or thing in an unfair way, because of allowing personal opinions to influence your judgment" or "the fact of preferring a particular subject or thing" (n.d.).

Biases can be good and bad. It is important to consider the implications of their existence though, especially when trying to measure the success of something objectively. An example of when biases can be advantageous is location signals that the search tool takes into account when producing results. An Englishmen would probably not have much use of a Chinese website and vice-versa, even if the actual content matches the original query (unless of course the user happens to understand both languages perfectly). Another example of this is location queries such as 'Chinese restaurants in Cambridge', which should return web pages about restaurants based in Cambridge, UK or Massachusetts, USA, depending on the user's IP address. This might seem logical, but in the truest sense it is a bias employed by the search engine to help provide more relevant results to individuals. Truly unbiased search results are probably impossible to come by nowadays.

There is a general move from objectivity to subjectivity in the sense that users become the subject of search results as much as the query they pose. Instead of neutrally providing results for a query alone, the results are tailored around the information known about the user (e.g. language, location, clickstream, social media likes, bookmarks, etc.) to make up the missing context. The user becomes the subject and context of a query, while the results become an objective list of matches for all those values rather than just the query term (s).

So in standard web search we now have the user as the subject and the results as the object. In creative search this may be reversed: the user is the object and the results become the subject.

12.5.2 Constraints

There are certain factors and constraints that influence the perception and success of search results. Some can be taken into account when building a search system but others cannot be avoided. User education is one way to deal with those issues. External constraints such as the setting in which the search takes place come to mind. Is the user operating from a handheld device or a desktop computer? Is he or she in a hurry to find answers or just leisurely browsing for them? Is the search system web-based or is the user querying a database?

User Expectations It is important to note that "search systems are not used in isolation from their surrounding context, i.e. they are used by real people who are influenced by environmental and situational constraints such as their current task" (White and Marchionini 2004). User expectations should be taken into consideration during the evaluation of search results. Users who are hoping to find precise answers to a specific question might not be satisfied by exploratory search results. Someone browsing for inspiration on a broad topic on the other hand could benefit from them. Fewer expectations (an open mind) allow creativity to happen more easily. Empirical experiences form expectations, which hinder our ability to accept creative ideas when they happen. In order to be able to recognise creative ideas we need to be able to see what they all have in common and in what way they differ and not reject unusual, unexpected ones. We can link this very nicely to the idea of exploratory search. Lowering expectations or opening the mind implies extending the task domain or problem space.

User Skill The searching skills of a user matter. Specifically his or her ability to articulate an Information Need (IN) and any knowledge of special search techniques (use of Boolean modifiers, quotation marks, wildcards, etc.) are two important factors that influence the results obtained greatly. This is very much based on the old idea of 'garbage-in, garbage-out' (Lidwell, Holden and Butler 2010).

Visual Representation The way that results are presented affects how the user perceives them. A diversity of different document types, for example text, images, sound, or video results could improve how well the results are rated (Sawle, Raczinski and Yang 2011). An alphabetical list is a typical model for representing text data sets for example. But a ranked list might not be the best way to represent search results. Other models could be a differently ranked or ordered list, a tree structure, a matrix, a one-to-many relationship, etc. See also section 12.4. § 12.4

Structure of Results As suggested by Sawle et al (2011) we need to consider different ways to structure and measure search results. A single, perfectly good result might be deemed irrelevant and useless if it is surrounded by several unsuitable results. Therefore there might be certain advantages to measuring and evaluating the value or relevance of individual results over a whole set of results.

Direct User Relevance Feedback Relevance feedback lets users rate individual results or sets of results either directly (through manual ratings) or indirectly (through click-stream data). This data is then congregated and used for webpage rankings or other purposes such as suggesting other query terms. It can improve results for similar queries in the future but also lets the user stir the direction his search is taking in real-time. Users can adjust their query to manipulate the results; this basically means they adjust some of their own constraints.

> Relevance feedback—asking information seekers to make relevance judgments about returned objects and then executing a revised query based on those judgments—is a powerful way to improve retrieval. (Marchionini 2006)

Automatic Query Expansion As opposed to integrating and involving the user actively in the refinement of a query, in automatic query expansion the improvements are done passively, often completely without the user's knowledge. Information gathering methods include, for example, the analysis of mouse clicks, so called like buttons (e.g. Facebook, Google+) or eye tracking, etc. How the collected data is then used varies. Simple examples of automatic query expansion are the correction of spelling errors or the hidden inclusion of synonyms when evaluating a query.

Depending on these factors and constraints, search results can be viewed as useful or useless. In a way the usefulness or correctness of an idea or result cannot always be judged fairly – there are always conditions that will affect how the outcome is interpreted. In the scenario of a creative search tool, results could be very useful, while they might be completely useless in another.

We would need to investigate each individual search result in terms of its value and creativity. This could be done by user ratings or satisfaction questionnaires. Rather than measuring the success of individual results we could look at evaluation them as one set instead.

The search results produced by `pata.physics.wtf` can be quite surprising sometimes and it not always clear how they connect to the initial query (especially if the inner workings of the algorithms are unknown), even if we identify through which function a result has been obtained. The names of these algorithms might not be helpful to users though if they are unfamiliar with the concept of pataphysics and might therefore appear rather nonsensical. Whilst there is a clear logic to each search result, they might appear anomalous to the user's expectations if he or she received these results without knowing the philosophy of the search tool. The results could possibly appear random then, and might therefore appear useless to the user.

12.6 Meta Analysis

The code for `pata.physics.wtf` and this thesis written in LaTeXand are both kept under git version control (Git 2016).

> The name 'git' was given by Linus Torvalds when he wrote the very first version. He described the tool as 'the stupid content tracker' and the name as (depending on your mood):
>
> - random three-letter combination that is pronounceable, and not actually used by any common UNIX command. The fact that it is a mispronunciation of 'get' may or may not be relevant.
> - stupid. contemptible and despicable. simple. Take your pick from the dictionary of slang.
> - 'global information tracker': you're in a good mood, and it actually works for you. Angels sing, and a light suddenly fills the room.
> - 'goddamn idiotic truckload of sh*t': when it breaks
>
> (Git 2016)

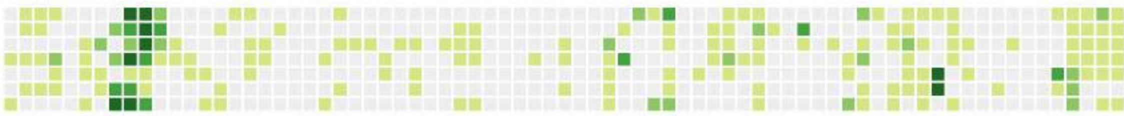


Figure 12.6 – GitHub contributions for code and thesis

Both repositories (folders which contain the files to be monitered) are stored remotely on GitHub (2016) and synced with the local machine. Image 12.6 shows the contribution history from the last 17 months for both of the `pata.physics.wtf` code and this thesis. A darker green indicates several commits (i.e. saves) while gray indicates no commits. Each square represents a day, each colum a week (Sunday–Saturday). 12.6

The full git commit histories for both repositories are shown in appendix D. § D

Aspirations

13

Mid the silence that pants for breath,
when I thought myself at my last gasp,
haine ou de l'ambition et qui se,
the pale motor vessel withdrew its blue breath toward the island's horizon.

As pure and simple as a powder puff,
such also was the ambition of others upon the like occasion,
there was hardly a breath of air stirring,
mon ancien cœur en une aspiration vers la vertu.

After drawing a long breath,
the silver ring she pull'd,
the suitor cried, or force shall drag thee hence.

For wild ambition wings their bold desire,
and with thine agony sobbed out my breath,
I will pull down my barns.

Developing a software product rarely finishes. It is maintained, refactored, repurposed, updated, extended, etc. Especially with creative products, where the functional requirements are more fluid perhaps, it is always tempting to change things.

For the purpose of this doctoral project, the artefact `pata.physics.wtf` is a snapshot of a product in constant motion. The state of the code at the time of submission of this thesis is described in chapter 10 and further elaborated on in the Patanalysis chapter. But it may very well continue to evolve. § 10 § 12

Here, in this chapter I will lay out some of the potential further work for this project. This may continue on a private basis or in a more academic environment.

13.1 Performance

Startup The website can be slow to load. Currently speed performance was not a priority during development. In fact it is not built for speed from the ground up. Each time the server restarts, the indexing process takes place from scratch (see chapter 10.1). This takes time. Google and other big web search engines do this continuously in the background to keep data up to date. The index is currently cached after startup but perhaps preprocessing it and storing it more permanently in a database would help speed up the start. However this may not be necessary, as it only affects the server startup. § 10.1 § 10.1.2

Query Response The time it takes from the user entering a query term and the system displaying the results page varies between unnoticable short and impatiently long. This is due to the pataphysicalisation process. This requires calls to external and internal APIs such as Flickr and WordNet. See analysis on speed issues in table 12.4. ⊞ 12.4

Preprocessing Corpora At this point the texts in the corpora consist of almost § 10.1.1 unedited plaintext ('.txt') files[1] (see chapter 10.1.1). Newlines and whitespace formatting varies, as does language and quality of spelling. Generally, chapter headings, chapter numberings, etc. were left untouched. The Shakespeare corpus contains poetry and plays for example. With the plays, scene information, stage directions, and voice details were kept. This means sentences that appear in the results of the search tool can contain peripheral words such as in this example: "...Athens and a wood near it ACT I ..." from *A Midsummer Night's Dream* or this example: "...Exit SHERIFF Our abbeys and our priories shall pay This expedition's charge ..." from *King John*. This could be addressed by preprocessing the individual texts in advance and removing any text that might interfere with the readablility of results.

Image Sizes At the moment images are retrieved at one specified size through the various API calls even though they are displayed at various different sizes depending on their location in the image spiral (unless they are displayed as a list). This process could certainly be optimised. Smaller image sizes could be accessed via the APIs.

13.2 Design

§ 10.3

Responsive Spirals Currently the image and video spirals (see chapter 10.3) are fixed size. This means that when the webpage is resized the spiral stays the same size and is left-aligned on the page. Ideally it would be better to scale the spiral with the width of the browser page. This could be achieved using § B.7 percentage widths, although it would require a lot of work to adapt the current code for the spirals (see chapter B.7).

Scalable Image Sizes As mentioned above, images are retrieved at one size through the various API calls. Because images in the spiral have different sizes according to where in the spiral they are located, they are scaled up or down directly in the HTML code. This means that some of the images look distorted and pixelated if they have to be scaled up or down too much.

Square Aspect Ratio Another issue is the aspect ratio of images and videos. For the spiral they need to be square. They are currently distorted as opposed to cropped. It might be possible to specify an option in the API calls to only retrieve square images which would help this problem.

[1] For text files downloaded from Project Gutenberg, the Gutenberg specifc copyright notices have been removed to only contain the relevant body of text

Responsive Poems A similar problem to the responsive spirals exists with the display of the Queneau poems. The random poems are centered on the page but the Queneau poems require a lot more formatting and styling to render and currently this is achieved by left-aligning them and having a fixed 'absolute' position on the page. Ideally this would also be centered as in the random poems.

Paginate Results For the text-by-source and text-by-algorithm search as well as the image- or video-as-list search results, it may improve the loading speed of the results page to split the results into smaller chunks and display them on several pages instead of one long scrolling page. This is called pagination.

Random Sentences Adding to the source of random sentences used in the top and bottom banner on the website should be an ongoing endeavour. The current list of sentences used is shown in appendix A.1. § A.1

13.3 Text

Result Sentences Currently the way result sentences are retrieved for the text search is based on punctuation (see chapter 10.2.2). This means once a pata- § 10.2.2 physicalised keyword has been found, the system retrieves up to 10 words prior until it reaches a punctuation mark and the same for after. The idea here was to get suitable sentence fragments. This could be changed to rely on POS tags for example or simply retrieving complete sentences.

Stopwords When the index is created only words that are not considered stopwords are added. We could modify the list of stopwords (see appendix B.6) to § B.6 include a few more uninteresting words. Or we could simply remove everything but nouns for example. This would drastically influence the results produced by the system.

Rhyming Scheme One of the biggest points for future work is to introduce a rhyming scheme for the poetry results. This might involve some more NLP during the creation of the index. It would make the poems much more readable. This § 10.1.2 could include pronounciation POS tags or other International Phonetic Alphabet (IPA) like data (for example using an API like Wordnik (Wordnik 2016) or a library like NLTK). So a word in the index dictionary might contain the following items.

```
(``tree'': [``l_00'': [24,566,4990], ``s_14'': [234,5943]], ``[tri]'')
```

By doing POS tagging with pronounciation data, we could retrieve sentences that § 6.2.2

match the sound of the last word of the previous line for example.

13.4 Pataphysicalisation

WordNet The vocabulary in WordNet is limited. According to it's website (WordNet n.d.) it contains 117,000 'synsets'[2] This affects two of my algorithms (namely the Syzygy and Antinomy algorithms). See also discussion in chapter 12.2.5. An option might be to somehow widen the amount of word matches by including different word-types/forms and relationships, such as troponyms, homonyms and heteronyms. Using these could introduce a whole new kind of pataphysical result.

§ 12.2.5

Homonyms are pronounced the same but mean something else (e.g. 'write' and 'right'). Heteronyms are words that are spelled the smae but have a different meaning (e.g. 'close to the edge' and 'to close the door'). Homophones are often used to create puns (and remember—puns are syzygys of words), for example "past your eyes" and "pasteurize".

> You can tune a guitar, but you can't tuna fish. Unless of course, you play bass.
>
> (attributed to Douglas Adams)

Antinomy The antinomy algorithms relies on WordNet's antonyms. A lot of words simply do not have an opposite and no fallback is currently defined. This means a lot of the time the antinomy function will not produce any results. Andrew Dennis implemented the algorithm in the same way, as discussed in chapter 11.1. It would be great to come up with a better way of dealing with this concept to ensure results are produced everytime.

§ 11.1

§ 6.2

Stemming Stemming could increase the number of results found by all algorithms (see chapter 6.2). A danger of increasing the output of the pataphysicalisation is always that results become more boring. Currently queries such as 'clear' and 'clearing' are treated as separate entities and would produce different results. Stemming would turn both of these words into the stem 'clear' and they would return the same results. Now it becomes immediatly clear (no pun intended) though that this might not always be desirable as just illustrated in this sentence: the root meaning of 'clear' can be very different to the meaning of 'clearing'.

[2]Synonyms—"words that denote the same concept and are interchangeable in many contexts"—are grouped into unordered sets called synsets (WordNet n.d.).

Queneau's poems It would be nice to actually add Queneau's poems (Queneau 1961) into the Faustroll corpus as little easter egg (see chapter 2.8). § 2.8

Image Algorithms The image and video search currently rely on external APIs (see chapter 10.3). One option to approach this in a totally differnet way would be to write algorithms that analyse and pataphysicalise the actual image or video data themselves. This might involve manipulating histograms or pixel maps. § 10.3

Maximum Obscurity N-grams are a NLP technique introduced in chapter 6.2.2. The idea is that it allows for prediction of likely word pairs, meaning if the word 'sunny' often occurs just before the word 'day' in a given training text or corpus then the probability for this particular n-gram is higher than say for 'sunny dog'. This can be increased to predict the probability of longer chains of words. One can immediately see the attraction of abusing this to generate pseudo sentences or even of creating a formula similar in nature but for example ranking obscure combinations of words higher than common ones. So for example instead of having a Maximum Likelihood Estimation (MLE) (see equation 6.8) we could have a 'Maximum Obscurity Estimation' which returns the highest probabilty for word sequences that happen the rarest. § 6.2.2 Σ 6.8

Pataphysical Entropy Similarly, we could could play with maximum entropy models as shown in chapter 6.2.2 together with POS tagging by rigging given probability for tags. There are endless possipilities of abusing these kinds of techniques. This is also very reminiscent of OULIPO techniques. § 6.2.2 § 6.2.2

Grammars We could create a whole new language grammar based on pataphysical principles. Examples of using a standard grammar (see chapter 6.2.2) for generating 'random' text are as follows[3]. § 6.2.2

ArtyBollocks	Generates artist statements.
DadaEngine	A system for generating random text from grammars.
SciGen	Generates random Computer Science research papers.

Uncreativity In chapter 7.2.5 I discussed the concepts of uninspiration and aberration by Wiggins and Ritchie (2012; 2006) in relation to their CSF. We could define a 'Pataphysical Search Framework' in the same way. Table 13.1 shows some of their original definitions for various forms of aberration and uninspiration. Table 13.2 then shows some rough ideas about how pataphysical concepts might be defined. § 7.2.5 ⊞ 13.1 ⊞ 13.2

[3](Stribling, Krohn and Aguayo 2016; Dada Engine 2016; Winter 2016)

Clinamen	smallest possible aberration to make the biggest difference
Antimomy	reachable, abnormal concepts with value
Anomaly	reachable concepts outside the norm
Absolute	criteria for value and norm must be perfectly matched
Syzygy 1	concepts reachable within 3 steps from the query
Syzygy 2	transformed set of concepts $S_{obj} \to S^{meta} \to S'obj$

This is definitely work in progress and it would be out of the scope of this thesis to elaborate much further.

Table 13.1 – CSF concept definitions of uncreativity

Name	**Equation**
Universal set of concepts	U and $X \subseteq U$
Aberration	B where $B \notin N_\alpha(X) \ \wedge \ B \neq \emptyset$
Perfect Aberration	$V_\alpha(B) = B$
Productive Aberration	$V_\alpha(B) \neq \emptyset \wedge \ \neq B$
Pointless Aberration	$V_\alpha(B) = \emptyset$
Hopeless Uninspiration	$V_\alpha(X) = \emptyset$
Conceptual Uninspiration	$V_\alpha(N_\alpha(X)) = \emptyset$
Generative Uninspiration	$elements(A) = \emptyset$

13.5 Extensions

Additional APIs Currently 5 APIs[4] are used in pata.physics.wtf. This could be increased to include more varied sources of data. Sites like Flickr are heavily based on user tags ('folksonomies') which can be unreliable and a bit random at times. Possible additional APIs to consider would be Instagram, Imgur, Facebook, Google Image Search, DeviantArt, Pinterest, Vimeo, Twitter, SoundCloud, etc.

Web Search The use of APIs could also include web search results rather than just images and videos. This would needs its own interface section and a suitable
§ 12.2.7 display style for the results. The biggest problem for this are API limitations as mentioned in chapter 12.2.7. Alternatively a ready-made index or crawl could be used but these are typically many terrabytes in size and have a cost attached.

[4]Flickr, Getty, Bing, MicrosoftTranslator and YouTube

Table 13.2 – Possible definitions of pataphysical concepts in terms of the CSF

Name	Equation
Norm	$N_\alpha(X) = \{c \in X \mid N(c) > \alpha\}$ where $N \in [0,1]^X$
Value	$V_\alpha(X) = \{c \in X \mid V(c) > \alpha\}$ where $V \in [0,1]^X$
Pata	$P_\alpha(X) =$ $\{c \mid c \in (CLI(X) \cup ANT_\alpha(X) \cup SYZ(X) \cup ANO_\alpha(X) \cup ABS(X))\}$
Clinamen	$CLI(X) = \{c \in X \mid N_{0.9}(N_{0.1}(c))\}$
Antinomy	$ANT_\alpha(X) = \{c \in X \mid V(N_0(c)) > \alpha\}$
Anomaly	$ANO_\alpha(X) = \{c \in X \mid N(c) < \alpha\}$
Absolute	$ABS(X) = \{c \in X \mid V_1(N_1(X)) \neq \emptyset\}$
Syzygy 1	$SYZ(query) = \bigcup_{n=0}^{3} elements(Q(N,V)^n(query))$
Syzygy 2	$SYZ(X) = S'(X)$ where $S_{obj} \rightarrow S^{meta} \rightarrow S'obj$

Crawling the web myself is not an option due to the computational power, time and space required to do so.

Audio Search It would be nice to include audio search using an API such as SoundCloud. Technically the pataphysicalisation could work similar to the image and video searches, meaning it would be based on user tags. One idea would be to work with audio waves directly although this needs to be explored further first.

Additional Algorithms It would be nice to implement some more algorithms for the search tool. This could include the two additional algorithms suggested by Andrew Dennis (see chapter 11.1) or developing more of my own. This could involve implementing some of the other pataphysical principles, such as equivalence or anomaly. Or it could consist of implementing some of the more famous OULIPO techniques. The repetoire of them is huge (see tables 4.1 and 4.2).

§ 11.1
⊞ 4.1 & 4.2

Custom API Finally, it would be great to develop a custom API for the algorithms of pata.physics.wtf. This would allow other people to use the search remotely without going through the interface and to use the results as they want. This would have been beneficial for the Digital Opera project and certainly for other researchers/developers like Andrew Dennis.

§ 11.2
§ 11.1

13.6 User Testing

Focus Group It might be interesting to look at opinions of various people (general public and experts) about the interpretation/evaluation framework. This could be done by asking them to provide their own definition of computer creativity and then to analyse and evaluate a product (such as pata.physics.wtf) according to their own criteria. Then follow this up by getting the same people to use my proposed framework to compare the results. This would include asking them about whether or not they thought that using the framework was beneficial to them or confusing.

§ 9.2.3

Eye-Tracking To study the effects of using different styles of presenting the same results, an eye-tracking experiment could be done. This would involve setting up participants with the necessary equipment and then introduce them to pata.physics.wtf and moniter their eye movements as they navigate the site. This could also provide details about how long users spend on each results page, what kind of style of results they prefer, etc. Some may prefer image or video search over the text search while others may not be interested in that at all. Generally of course one has to take into account that this is a creative piece of work and not everybody will like it. It purposefully purposeless and highly subjective, so user feedback may not provide unbiased and useful results.

Part VI

H∀PPILY ΣVΣR ∀FTΣR

Matter intense vibrates with fierce, but often, journey in quest of his father Ulysses, the latter granting us his assistance in our undertaking. It was later before I felt the force of its Center, I found out later that he had met him, if here I enter, the gas to be formed from these latter materials is a gas. Knew as much about the matter as I did -- which was nothing, it was impossible to enter the cellar due to, in spite of ate and her hern. Ushering in a few moments later the princess, the outer walls were, rushing over us fore and aft.

Outroduction

14

Yet my state is well,
is to take those things for bird,
and God keep him out of my sight,
I do spy some marks of love in her.

With catlike watch,
I have watch'd and travell'd hard,
and some will mourn in ashes,
so that hardly can I check my eyes from tears.

Pillars of the state,
word out of his right sense,
first emperor of Rome Mark Anthony.

Have you had quiet guard,
though art a guard too wanton for the head,
of each other's watch.

The research presented in this thesis described Algorithmic Meta-Creativity (AMC) and its evaluation. The first part of this knowledge was embodied in an artefact pata.physics.wtf and the second part was formulated as a theoretical framework to help interpretation of products of AMC.

The overall research methodology was described in the Methodology chapter but it can be summarised having a subjective, transdisciplinary approach, using creative computing, experimental and exploratory methodologies. Specifically, existing literature was synthesised, algorithms were designed, an artefact was created using iterative exploratory development, a theoretical framework was developed and the project contained a critical reflection and analysis of the artefact presented.

§ 3
II
IV
III
V

14.1 Observations

The artefact pata.physics.wtf should be seen as an artwork inspired by and dedicated to AMC, pataphysics, OULIPO and programming culture.

On the face of it this thesis might appear to argue that computers can be seen as creative entities. This is however not the case. In fact I argue against this in the Interpretation chapter—the computer is always only a tool for a human's creativity and nothing more. This is not to say that a computer can't be 'taught' creative techniques, which is what I have called AMC.

§ 9
14.1

Figure 14.1 shows an abstract view of pata.physics.wtf.

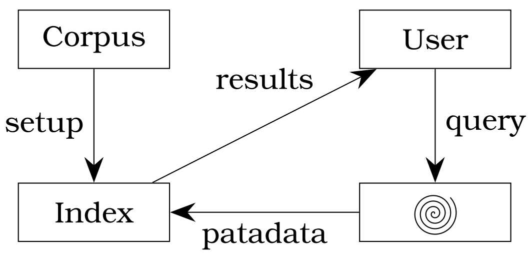


Figure 14.1 – Pataphysical system architecture (again)

14.2 Answers

§ 1 In the introduction I asked several questions that I attempted to answer with the research presented in this thesis. This section contains brief answers from 50.000 feet[1], meaning they provide a top-down view of the answer and pointers to where in the thesis readers can find more elaborations.

§ 8 ▦ 8.5 § 4.2.2 **What is the relationship between pataphysics and creativity?** The Foundations chapter discusses this in detail. Pataphysics provides many philosophical principles which can be turned into creative techniques and constraints. In specific table 8.5 shows the similarities and differences between pataphysics and creativity. One of the key attributes of creativity for example is the idea of bisociation—the juxtaposition of the dissimilar—which relates directly to the pataphysical concept of the antinomy—the simultaneous existence of the mutually exclusive.

§ 5.3.1 § 12.3 **How is computer creativity related to artificial intelligence?** Much of the research in computational creativity (see chapter 5.3.1) stems from the area of AI. In the Creativity & Intelligence chapter I mentioned the similarities in these two fields. In particular, I discussed the ideas of free will and surprise, understanding and simulation, and brains and computers.

§ 9.1.2 § 7.2.1 **Should we distinguish between computationally automated or emulated creative processes and the programmer's input?** Yes. Just like the process and product are both equally important, the computational process and the programmer are both essential. This is discussed in the chapter on The Programmer but also gets addressed in the Output minus Input section in chapter 7 on Evaluation.

§ 7.2 § 9.2 **How can a machine's creative output be evaluated?** Previous attempts at evaluating computer creativity are critically reviewed in chapter 7.2. The Creative Interpretation chapter then introduces one of the main original contributions of this research: a new framework for the evaluation and interpretation of creative artefacts (this can be applied to human-made and machine-made products). Since the perception of creativity is subjective it cannot be quantified in objective terms. By providing a framework that takes into account all possible contextually relevant contributors though we can approximate an objective evaluation.

§ 8.2.2 **How can information retrieval be infused with creativity?** This is explored in chapter 8.2.2 and of course the Implementation chapter, where the develop-

[1] Inspired by Tim Berners-Lee's articles on the web in it's early days (1998).

ment of pata.physics.wtf is explained. This is a direct example of creative exploratory IR. § 10

14.3 Contributions

The original contributions to knowledge presented by this doctoral research can be broken down into the 4 points below.

- Three pataphysical search algorithms (clinamen, syzygy and antinomy).
- A creative exploratory search tool demonstrating the AMC.
- 7 subjective criteria and 5 objective constraints for defining creativity.
- A combined framework for evaluating and interpreting creativity.

14.4 And Finally

Pataphysics is the science...

INTERLUDE III

There is no pure science of creativity, because it is paradigmatically idiographic—it can only be understood against the backdrop of a particular history.

(Elton 1995)

Evaluation is thus a matter of subjectivity, since no scientism allows us to pretend to objectivity, an objectivity aspired to on the illusory grounds that it would support taking a decision without the decision-maker simultaneously taking a risk or responsibility. (Montfort and deVarine, cited in Matarasso 1997, Matarasso's translation)

Tools are not just tools. They are cognitive interfaces that presuppose forms of mental and physical discipline and organization. By scripting an action, they produce and transmit knowledge, and, in turn, model a world.

(Burdick et al. 2012)

Humanists have begun to use programming languages. But they have yet to create programming languages of their own: languages that can come to grips with, for example, such fundamental attributes of cultural communication and traditional objects of humanistic scrutiny as nuance, inflection, undertone, irony, and ambivalence. (Burdick et al. 2012)

Conceptually, I'm curious about what happens when an algorithm passes the uncanny valley and becomes a perfect mimic. If humans were unable to distinguish the generated drug experience from a real one, the machine would become a sort of philosophical zombie: an entity that appears to be something that it isn't, something it could never be. (McDonald 2016)

Part VII

POST☹

Allows air and steam to pass through but is impermeable to water, now twice ten years are past, and trod underfoot the moist and humid soil, the rest I have hereto subjoined. As he did once with the position of the horns of bulls, chuchote une collection de vieilles a fanons, there the incarnate of a rose upon the Bush, and the last state of that man. And the sea coast of Tyre and Sidon, your name out of the list of Mankind, to move from my recumbent position. Mast, seeing or audacious among learned men to assign even a far greater, I know the man that must hear me.

Random

A

A.1 Random Sentences

The full list of random sentences used by `pata.physics.wtf` for the top and bottom banners on each page (see screenshot 10.1).

- in posthumous collaboration
- the decomposing brain goes on working after death and it is its dreams that are Paradise
- plagiarism by anticipation
- the applause of silence is the only kind that counts
- to understand pataphysics is to fail to understand pataphysics
- duration is the transformation of a succession into a reversion
- god is the tangential point between zero and infinity
- laughter is born out of the discovery of the contradictory
- ha ha
- the aesthetic of formal constraint
- the unique imaginary solution to the absence of problems
- the contemporary relationship between science and poetry
- a huge and elaborately constructed hoax
- only those who attempt the absurd achieve the impossible

- the random is opposed to the deterministic
- pure multiplicity irreducible to any other sort of unity
- persistence and perseverance to buttress a fleeting existence
- enfolding a subject laterally, associatively
- a doctrine of correspondences, counterpoised by the exotic charm of another system of thought
- double negative is necessary to stop the mind believing
- the absence of contradictory evidence is not proof of a theory's validity
- very wrong in very important ways
- no one point of view is final
- unification of opposites
- an athletic aesthetics of intuitive and instantaneous judgements
- constantly diverted from any objective by the very progress which their energy sustains
- imagination envisions the reconciliation of the individual with the whole
- behind the illusion lies knowledge
- a biomolecular bibliomecha of breathtaking beauty
- indubitably coherent yet absolutely nonsensical
- stylistic and formal experimentation can not be dismissed as purely apolitical
- variable phoneme sequences in suspension within a cloud of relative
- epi-cultural etymologies
- speculative solutions for imaginary problems
- nonsense is nonsense only when we have not yet found the point of view from which it makes sense
- laughter is the discord between tensions being resolved
- read with intention to rewrite
- a fractal geometry of momentums
- minimizing energy, crystallizing latent structure, pleasure is understood as a practice, and accumulates as experience
- don't fall out of love
- the 'something like', the pseudo
- the text transforms itself as soon as it is understood
- adaptation of archaic mental structures to new environments
- a beautifully controlled yet hideously wasteful catastrophe
- driven by compulsive urgency to constantly reconceive the whole idea
- writing the unrightable wrongs
- prisoners of conscience
- machina sapiens negotiating the transformation of what is mortal into what is immortal
- database hyper-archive applications stimulating relaxation

- smelling of the rain it falls on the way down
- a thousand ways to greet the dawn
- how far up the chain can you put this without ambiguity
- a gentle kitten is licking the inside of my heart
- was this constrained by you, or restrained by the concept
- adjoining always antiquated permutation
- joint ventures can go too far away
- nowhere to be found here and elsewhere
- engender links to a balancing veneration
- coerce to do as a sizeable unprocessed primer
- plain up be a best concoction in the words
- the farcical pandemonium of technology
- it is not true that there were any nails
- this discovery opens the door into a completely new anti-world
- extending as far beyond metaphysics as the latter extends beyond physics
- turn the world upside down and inside out
- the law of the ascension of a vacuum toward a periphery
- the anti-world God not only plays dice, he spells his name backwards
- in the absence of a butler, where does the gun fit in
- space is defined by simultaneity
- time is a flowing stream, a liquid in uniform rectilinear motion
- space is a solid, a rigid system of phenomena
- the deceleration of our habitual duration conserved by inertia
- a perfect elastic solid
- movement into the past consists in the perception of the reversibility of phenomena
- relativity is absolute
- all observations depend on viewpoint and the scale of the scientist
- the clinamen, subjective viewpoint and anthopocentrism all rolled into one
- the identity of opposites
- making negatives do the work of positives
- in this year eighteen hundred and ninety-eight
- the twenty-seven equivalents
- the virgin, the bright, and the beautiful today
- the fifth letter of the first word of the first act
- voices asymptotic towards death
- an epiphenomenon is that which is superinduced upon a phenomenon
- concerning the amorphous isle
- like soft coral, amoeboid and protoplasmic
- searching desperately under the quinuncial trees for the venerable absent one

- the night computed its hours
- a remarkable epizootic disease
- the eternal nothingness
- love looks exactly like an iridescent veil and assumes the masked face of a chrysalis
- in a telepathic letter
- homo est deus
- ∞ -0 -a + a + 0 = ∞
- with the aim of computing the qualities of the French
- the inferno of subjectivity

A.2 Heisenberg Quote

> The overly forceful insistence on the difference between scientific and artistic cognition quite likely derives from the incorrect notion that concepts are firmly attached to 'real objects', as if words had a completely clear and definite meaning in their relationship to reality and as if an accurate sentence, constructed from those words, could deliver an intended 'objective' factual situation to a more or less absolute degree. But we know, after all, that language too only grasps and shapes reality by turning it into ideas, by idealizing it. Language, too, approaches reality with specific mental forms about which we do not know right away which part of reality they can comprehend and shape. The question about 'right' or 'wrong' may indeed be rigorously posed and settled within an idealization, but not in relation to reality. That is why the last measure available for scientific knowledge as well is only the degree to which that knowledge is able to illuminate reality or, better, how that illumination allows us 'to find our way' better. And who could question that the spiritual content of a work of art too illumines reality for us and makes it translucent? One must come to terms with the fact that only through the process of cognition itself can we determine what we are to understand by 'cognition'. That is why any genuine philosophy, too, stands on the threshold between science and poetry. (Heisenberg 1942)

A.3 Digital Humanities Methodology Field Map

The full *Field map of digital humanities: emerging methods and genres* by Burdick et al. (2012).

- enhanced critical curation
  - digital collections
  - multimedia critical editions
  - object-based argumentation
  - expanded publication
  - experiential and spatial
  - mixed physical and digital
- augmented editions and fluid textuality
  - structured mark-up
  - natural language processing
  - relational rhetoric

- textual analysis
- variants and versions
- mutability

- scale: the law of large numbers
  - quantitative analysis
  - text-mining
  - machine reading
  - digital cultural record
  - algorithmic analysis
- distant/close, macro/micro, surface/depth
  - large-scale patterns
  - fine-grained analysis
  - close reading
  - distant reading
  - differential geographies
- cultural analytics, aggregation, and data-mining
  - parametrics
  - cultural mash-ups
  - computational processing
  - composite analysis
  - algorithm design
- visualization and data design
  - data visualization
  - mapping
  - information design
  - simulation environments
  - spatial argument
  - modelling knowledge
  - visual interpretation
- locative investigation and thick mapping
  - spatial humanities
  - digital cultural mapping
  - interconnected sites
  - experimental navigation
  - geographic information systems (GIS)
  - stacked data
- the animated archive
  - user communities
  - permeable walls
  - active engagement
  - bottom-up curation
  - multiplied access
  - participatory content creation
- distributed knowledge production and performative access
  - global networks
  - ambient data
  - collaborative authorship
  - interdisciplinary teams
  - use as performance
  - crowd-sourcing
- humanities gaming
  - user engagement
  - rule-based play
  - rich interaction
  - virtual learning environments
  - immersion and simulation
  - narrative complexity
- code, software, and platform studies
  - narrative structures
  - code as text
  - computational processes
  - software in a cultural context
  - encoding practices
- database documentaries
  - variable experience
  - user-activated
  - multimedia prose

- modular and combinatoric
- multilinear

- repurposable content and remix culture
  - participatory Web
  - read/write/rewrite
  - platform migration
  - sampling and collage
  - meta-medium
  - inter-textuality
- pervasive infrastructure
  - extensible frameworks
  - heterogeneous data streams
  - polymorphous browsing
  - cloud computing
- ubiquitous scholarship
  - augmented reality
  - web of things
  - pervasive surveillance and tracking
  - ubiquitous computing
  - deterritorialization of humanistic practice

A.4 Penn Treebank

The full list of POS tags mentioned in chapter 6.2.2 (Marcus, Santorini and Marcinkiewicz 1993).

CC	Coordinating conjunction
CD	Cardinal number
DT	Determiner
EX	Existential *there*
FW	Foreign word
IN	Preposition/subordinating conjunction
JJ	Adjective
JJR	Adjective, comparative
JJS	Adjective, superlative
LS	List item marker
MD	Modal
NN	Noun, singular or mass
NNS	Noun, plural
NNP	Proper noun, singular
NNPS	Proper noun, plural
PDT	Predeterminer
POS	Possessive ending
PRP	Personal pronoun
PP$	Possessive pronoun
RB	Adverb
RBR	Adverb, comparative
RBS	Adverb, superlative
RP	Particle
SYM	Symbol (mathematical or scientific)
TO	{to}
UH	Interjection
VB	Verb, base form
VBD	Verb, past tense
VBG	Verb, gerund/present particle
VBN	Verb, past particle
VBP	Verb, non-3rd ps. sing. present
VBZ	Verb, 3rd ps. sing. present
WDT	*wh*-determiner
WP	*wh*-pronoun
WP$	Possessive *wh*-pronoun
WRB	*wh*-adverb
#	Pound sign
$	Dollar sign
.	Sentence-final punctuation
,	Comma

:	Colon, semi-colon	‘	Left open single quote
(	Left bracket character	“	Left open double quote
)	Right bracket character	’	Right close single quote
"	Straight double quote	”	Right close double quote

A.5 Thinking computers

1. **Can computers think?**
   - Can computers have free will?
   - Can computers have emotions?
   - Can computers be creative?
   - Can computers understand arithmetic?
   - Can computers draw analogies?
   - Can computers be persons?
   - Is the brain a computer?
   - Can computers reason scientifically?
   - Are computers inherently disabled?
   - Should we pretend that computers will never be able to think?
   - Does God prohibit computers from thinking?
2. **Can the Turing test determine whether computers can think?**
   - Is failing the test decisive?
   - Is passing the test decisive?
   - If a simulated intelligence passes, is it intelligent?
   - Have any machines passed the test?
   - Is the test, behaviouraly or operationally construed, a legitimate intelligence test?
   - Is the test, as a source of inductive evidence, a legitimate intelligence test?
   - Is the neo-Turing test a legitimate intelligence test?
   - Does the imitation game determine whether a computer can think?
   - Can the Loebner Prize stimulate the study of intelligence?
   - Other Turing test arguments
3. **Can physical symbol systems think?**
   - Does thinking require a body?
   - Is the relation between hardware and software similar to that between human brains and minds?
   - Can physical symbol systems learn as humans do?
   - Can the elements of thinking be represented in discrete symbolic form?
   - Can symbolic representations account for human thinking?
   - Does the situated action paradigm show that computers can’t think?
   - Can physical symbol systems think dialectically?
   - Can a symbolic knowledge base represent human understanding?
   - Do humans use rules as physical symbol systems do?
   - Does mental processing rely on heuristic search?
   - Do physical symbol systems play chess as humans do?
   - Other physical system arguments
4. **Can Chinese Rooms think?**
   - Do humans, unlike computers, have intrinsic intentionality?
   - Is biological naturalism valid?
   - Can computers cross the syntax-semantics barrier?
   - Can learning machines cross the syntax-semantics barrier?

- Can brain simulators think?
- Can robots think?
- Can a combination robot/brain simulator think?
- Can the Chinese Room, considered as a total system, think?
- Do Chinese Rooms instantiate programs?
- Can an internalized Chinese Room think?
- Can translations occur between the internalized Chinese Room and the internalizing English speaker?
- Can computers have the right causal powers?
- Is strong AI a valid category?
- Other Chinese Room arguments

5. **Can connectionist networks think?**
    - Are connectionist networks like human neural networks?
    - Do connectionist networks follow rules?
    - Are connectionist networks vulnerable to the arguments against physical symbol systems?
    - Does the subsymbolic paradigm offer a valid account of connectionism?
    - Can connectionist networks exhibit systematicity?
    - Other connectionist arguments

6. **Can computers think in images?**
    - Can images be realistically be represented in computer arrays?
    - Can computers represent the analog properties of images?
    - Can computers recognize Gestalts?
    - Are images less fundamental than propositions?
    - Is image psychology a valid approach to mental processing?
    - Are images quasi-pictorial representations?
    - Other imagery arguments

7. **Do computers have to be conscious to think?**
    - Can computers be conscious?
    - Is consciousness necessary for thought?
    - Is the consciousness requirement solipsistic?
    - Can higher-order representations produce consciousness?
    - Can functional states generate consciousness?
    - Does physicalism show that computers can be conscious?
    - Does the connection principle show that consciousness is necessary for thought?

8. **Are thinking computers mathematically possible?**
    - Is mechanistic philosophy valid?
    - Does Gödel's theorem show that machines can't think?
    - Does Gödel's theorem show that machines can't be conscious?
    - Do mathematical theorems like Gödel's show that computers are intrinsically limited?
    - Does Gödel's theorem show that mathematical insight is non-algorithmic?
    - Can automata think?
    - Is the Lucas argument dialectical?
    - Can improved machines beat the Lucas argument?
    - Is the use of consistency in the Lucas argument problematic?
    - Other Lucas arguments

(Horn 2009)

A.6 Jarry's Writing

A list of Jarry's works in chronological order with their original titles copied from the French Wikipedia entry on Jarry (n.d.).

Works

- Les Antliaclastes (1886–1888) poems, reprinted in Ontogénie
- La Seconde Vie ou Macaber (1888) reprinted in Les Minutes de sable mémorial
- Onénisme ou les Tribulations de Priou (1888) first version of Ubu cocu
- Les Alcoolisés (1890) reprinted in les Les Minutes de sable mémorial
- Visions actuelles et futures (1894)
- "Haldernablou" (1894) reprinted in les Les Minutes de sable mémorial
- "Acte unique" from César-Antéchrist (1894)
- Les Minutes de sable mémorial (1894) poems
- César Antéchrist (1895)
- Ubu roi (1896, version of 1888)
- L'Autre Alceste (1896)
- Paralipomènes d'Ubu (1896)
- Le Vieux de la montagne (1896)
- Les Jours et les Nuits (1897), novel
- Ubu cocu ou l'Archéoptéryx (1897)
- L'Amour en visites (1897, publié en 1898) poems
- Gestes et opinions du docteur Faustroll, pataphysicien (achevé en 1898, published in 1911) novel
- Petit Almanach (1898)
- L'Amour absolu (1899)
- Ubu enchaîné (1899, published in 1900)
- Messaline (1900)
- Almanach illustré du Père Ubu (1901)
- "Spéculations", in La Revue Blanche (1901)
- Le Surmâle (1901, publié en 1902) novel
- "Gestes" in La Revue Blanche (1901) published in 1969 with "Spéculations" in La Chandelle verte.
- L'Objet aimé (1903)
- "Le 14 Juillet" in Le Figaro (1904)
- Pantagruel (1905 opéra-bouffe by Rabelais staged in 1911, music by Claude Terrasse)
- Ubu sur la Butte (1906)
- Par la taille (1906) opérette

- Le Moutardier du pape (1906, publié en 1907) opéra-bouffe
- Albert Samain (souvenirs) (1907)

Publications Post-Mortem

- La Dragonne (1907, published in 1943)
- Spéculations (1911)
- Pieter de Delft (1974) opéra-comique
- Jef (1974) play
- Le Manoir enchanté (1974) opéra-bouffe staged in 1905
- L'Amour maladroit (1974) opérette
- Le Bon Roi Dagobert (1974) opéra-bouffe
- Léda (1981) opérette-bouffe
- Siloques. Superloques. Soliloques Et Interloques De Pataphysique (2001) texts
- Paralipomènes d'Ubu/Salle Ubu (2010) livre d'artiste
- Ubu marionnette (2010) livre d'artiste

Translations

- La ballade du vieux marin (1893, after The ancient mariner by Coleridge)
- Les silènes (1900, translation of German play by Christian Dietrich Grabbe)
- Olalla (1901, novel by Stevenson)
- La papesse Jeanne (1907, translation of Greek book by d'Emmanuel Rhoïdès)

Contributions

- Écho de Paris
- L'Art de Paris
- Essais d'art libre
- Le Mercure de France
- La Revue Blanche
- Le Livre d'art
- La Revue d'art
- L'Omnibus de Corinthe
- Renaissance latine
- Les Marges
- La Plume
- L'Œil
- Le Canard sauvage
- Le Festin d'Ésope
- Vers et prose
- Poésia
- Le Critique

A.7 Leary's Tables

Table A.1 – Leary's four types of creativity

<table>
<tr><th>Reproductive Blocked</th><th>Reproductive Creator</th><th>Creative Creator</th><th>Creative Blocked</th></tr>
<tr><td>The routine, well-socialised person who experiences only in terms of what he has been taught and who produces only what has been produced before.</td><td>The innovating performer who experiences only in terms of the available categories but has learned to manipulate these categories in novel combinations.</td><td>The person who experiences directly outside the limits of ego and labels, and who has learned to develop new models of communications, or who can manipulate familiar categories in novel combinations or who can let natural modes develop under his nurture.</td><td>The person who experiences uniquely and sensitively outside of game concepts (either by choice or helplessly by inability) but who is unable to communicate or uninterested in communicating these experiences outside the conventional manner.</td></tr>
<tr><td>Reproductive Performer</td><td colspan="2">Creative Performer</td><td>Reproductive Performer</td></tr>
<tr><td colspan="2">Reproductive Experience</td><td colspan="2">Creative Experience</td></tr>
</table>

Table A.2 – Leary's social labels to describe the types of creativity

Reproductive Blocked	**Reproductive Creator**	**Creative Creator**	**Creative Blocked**
Unimaginative, incompetent hack.	Reliable nihilist, insensitive, unsuccessful innovator whose shock value changes to morbid curiosity as fads of performance change.	The mad creative genius, the undiscovered far-out crackpot creator who is recognised by later generations as a creative giant.	Psychotic, religious crank, eccentric who uses conventional forms for expressing mystical convictions.
Competent, responsible, reliable worker.	Bold initiator who wins game recognitions but whose fame crumbles as fads of performance change.	The truly creative giant recognised by his own age and the ages to come.	Solid, reliable person with a 'deep streak'.
Reproductive Performer	Creative Performer		Reproductive Performer
Reproductive Experience		Creative Experience	

Code

B

B.1 Index

An example excerpt from the Faustroll index data structure.

```
defaultdict(<function <lambda> at 0x101d17938>,
{
  u'fawn': defaultdict(<type 'list'>,
    {
      u'l_17': [101330, 111976],
      u'l_19': [345609],
      u'l_01': [92598, 469332, 469716, 469757, 469830, 469950]
    }),
  u'lenitives': defaultdict(<type 'list'>,
    {
      u'l_19': [121161]
    }),
  u'malheureuses': defaultdict(<type 'list'>,
    {
      u'l_06': [18950, 52631, 65053, 65792, 79960, 127861, 176200],
      u'l_04': [51545, 93611]
    }),
  u'nunnery': defaultdict(<type 'list'>,
    {
```

```
    u'l_19': [38182, 160331]
  }),
u'nuzzing': defaultdict(<type 'list'>,
  {
    u'l_19': [147035]
  }),
u'huileux': defaultdict(<type 'list'>,
  {
    u'l_04': [69998]
  }),
u'lentuli': defaultdict(<type 'list'>,
  {
    u'l_19': [217454]
  }),
u'porphyrogene': defaultdict(<type 'list'>,
  {
    u'l_01': [137368, 480308]
  }),
u'woods': defaultdict(<type 'list'>,
  {
    u'l_14': [16256],
    u'l_05': [2890],
    u'l_17': [11445, 34923, 48413, 59186, 61062, 78084, 78681, 81374,
          ↪  101319, 137624],
    u'l_10': [12500, 21691, 33136],
    u'l_02': [4596, 6221],
    u'l_01': [63176, 63622, 74535, 74807, 326433, 326464, 326472, 419835,
          ↪  441374, 467791, 481003, 481113, 500201, 500331, 501595],
    u'l_27': [60731],
    u'l_26': [1120, 9622],
    u'l_19': [149538, 232296, 294503]
  }),
u'clotted': defaultdict(<type 'list'>,
  {
    u'l_17': [92418],
    u'l_01': [53612, 133153, 200952, 241165]
  }),
u'spiders': defaultdict(<type 'list'>,
  {
    u'l_19': [285105, 301556],
    u'l_00': [3085]
  }),
u'hanging': defaultdict(<type 'list'>,
  {
    u'l_17': [84239],
    u'l_10': [35501, 92813, 126657],
    u'l_02': [6033, 33307, 34297],
    u'l_01': [10317, 102489, 114549, 134507, 193056, 210178, 251167,
          ↪  266320, 287765, 311802, 342798, 379707, 417326, 417956,
          ↪  448394, 469204],
```

```
        u'l_00': [1831, 5490, 6891, 11088, 13846],
        u'l_27': [19757, 45058, 83470],
        u'l_26': [6500],
        u'l_19': [8309, 15850, 17056, 20629, 25899, 47147, 47186, 49544,
              ↪  108943, 177323, 200639, 252551, 334120, 341621]
    }),
...
```

B.2 Clinamen Errors

This section show the Clinamen function allowing up to 3 errors. See chapter 12.2.4.

clear
: afar, ahead, Alas, altar, appear, bar, beam, beard, bears, beat, beer, ble, bleed, blew, bluer, bread, break, Caesar, calvary, can, canal, care, cedar, cellar, chair, charm, cheek, chen, chere, chern, choir, clad, claim, clasp, claws, clean, clear, clearly, clerks, climb, clock, clogs, close, cloth, color, coral, crab, crap, cresc, crest, Dead, dead, dear, Dewar, ear, ears, eat, ever, far, fear, Fear, feat, flag, flat, flesh, floor, Friar, glare, Great, great, head, hear, heard, heart, heat, Her, her, idea, ideal, ideas, jar, law, lay, lead, leaf, leap, least, leave, led, lees, left, leg, legs, lent, leper, less, lest, let, mean, meat, near, oar, Ocean, Opera, over, peak, pearl, per, plat, pleas, read, Read, real, rear, sea, Sea, seat, sheer, slab, sleep, solar, speak, star, steam, sugar, swear, swears, sweat, tean, tears, their, vulgar, war, year, years, zeal

fania
: acid, aid, aim, air, an, ance, and, animae, animal, Anna, ant, anti, ants, anvil, any, axis, Baba, bank, banks, basin, cabin, can, canal, Cane, canvas, dance, Danzig, data, Denis, fa, face, faced, faces, facet, facing, fact, facts, fading, faIt, faith, fake, fall, falls, false, family, fan, fans, far, fat, fate, fauns, favor, final, find, finds, fine, finer, fins, flint, fluid, foil, francs, fruit, gain, habit, hair, hand, hands, india, Jane, Janus, Kaka, Kantian, laid, lance, land, lanes, Latin, lava, mail, main, Man, man, many, nadir, nail, nib, nil, pair, pan, Pan, Papio, papio, Paris, rang, range, rapid, said, sail, Saint, saliva, San, sand, sang, sonic, tail, Tait, Tanit, tunic, unit, vain, valid, van, vanish, vanity, vans, vina, Yan

moss
: abyss, Across, across, acts, adds, Alas, almost, also, among, amor, amore, amour, ants, apes, arms, arose, as, As, ash, ask, ass, axis, bars, base, bases, beds, best, bis, blows, Boat, boat, boats, body, bolus, bone, bones, book, books, boot, boots, bores, born, Bosse, both, bout, bow, bowl, bows, box, boy, Boys, brass, brows, bust, case, cases, cash, cast, chose, clogs, close, co, coast, coats, Code, coins, cold, come, comes, cool, copy, cords,

cost, Cost, costs, cows, crass, cross, cuIs, cups, days, demons, Deus, disk, disks, Do, do, does, dogs, dome, domos, done, door, doors, douds, down, Down, dress, drops, dust, ears, ease, easy, eats, eggs, ells, else, ends, Eros, ess, est, eyes, fans, fess, fins, fish, fist, fists, foam, fog, foil, folds, foot, For, for, fore, fork, Form, form, forms, fotms, foul, four, fox, foxes, Ghost, ghosts, glass, glows, go, God, gods, goes, Gog, Gogh, gold, Gold, gong, good, goods, gown, gowns, grass, hams, has, hast, His, his, ho, Ho, holds, holes, Holy, home, Homo, hoof, hooks, hope, horn, horns, Horse, horse, horses, host, hot, hour, Hour, hours, house, houses, how, How, humors, hums, ikons, iris, irs, is, Is, Its, its, jaws, Jesus, jibs, job, John, jowls, joy, Just, just, kiosks, kiss, knows, last, laws, Lays, lees, legs, less, lest, lies, lions, lips, Lo, lobe, loins, Long, long, looks, Lord, lord, lords, lore, lose, loss, lost, Loti, lots, loud, louse, Love, love, loves, low, Loye, m, made, mail, main, make, makes, male, man, many, map, maps, mask, mass, masses, mast, masts, may, me, mean, means, meat, meet, men, mere, mesh, meshes, met, milk, mimes, mist, mite, mites, mob, moist, moles, month, months, moon, mor, more, Moses, most, motor, mount, Mour, mouth, mouths, moved, mower, Mrs, much, music, must, Must, my, nest, news, nisi, no, No, noise, non, none, noon, Nor, nor, nos, nose, Not, not, note, now, Now, nuts, o, oak, oar, oars, oc, odd, of, off, ofQ, oil, old, on, one, ones, or, orb, orms, our, out, own, pass, past, pigs, piss, Plus, Poe, poets, pole, poles, ponds, Poor, poor, pope, port, Pour, prose, Prose, rats, rays, rest, rise, rises, road, robe, robes, rock, rocks, rod, Roi, role, roll, rolls, rome, roof, room, rooms, root, rope, ropes, rose, rosy, row, rows, s, says, sc, sets, shops, smock, smoke, So, so, soft, sole, Some, some, son, songs, sons, soon, Soon, sorb, soul, souls, sows, sums, suns, tats, This, this, those, Thus, thus, tjis, to, To, toad, toads, tock, toes, told, tome, tone, toO, too, took, top, tops, tore, torn, tossed, Town, town, Tres, tres, ups, us, use, vans, vast, Was, was, wash, wasps, webs, whose, wigs, Woan, won, wont, wood, word, words, wore, Work, work, Works, works, worm, worn, wove, Yes, yolk, York, you, You, your, Your

B.3 WordNet

B.3.1 Antinomy

```
[
  Synset('clear.n.01'), Synset('open.n.01'), Synset('unclutter.v.01'),
  Synset('clear.v.02'), Synset('clear_up.v.04'),
   ↪ Synset('authorize.v.01'),
  Synset('clear.v.05'), Synset('pass.v.09'), Synset('clear.v.07'),
```

```
    Synset('clear.v.08'), Synset('clear.v.09'), Synset('clear.v.10'),
    Synset('clear.v.11'), Synset('clear.v.12'), Synset('net.v.02'),
    Synset('net.v.01'), Synset('gain.v.08'), Synset('clear.v.16'),
    Synset('clear.v.17'), Synset('acquit.v.01'), Synset('clear.v.19'),
    Synset('clear.v.20'), Synset('clear.v.21'), Synset('clear.v.22'),
    Synset('clear.v.23'), Synset('clear.v.24'), Synset('clear.a.01'),
    Synset('clear.s.02'), Synset('clear.s.03'), Synset('clear.a.04'),
    Synset('clear.s.05'), Synset('clear.s.06'), Synset('clean.s.03'),
    Synset('clear.s.08'), Synset('clear.s.09'),
     ↪  Synset('well-defined.a.02'),
    Synset('clear.a.11'), Synset('clean.s.02'), Synset('clear.s.13'),
    Synset('clear.s.14'), Synset('clear.s.15'), Synset('absolved.s.01'),
    Synset('clear.s.17'), Synset('clear.r.01'), Synset('clearly.r.04')
]
```

```
synset item:clear.n.01
synset item:open.n.01
synset item:unclutter.v.01
antonym out:clutter
antonym in:clutter
synset item:clear.v.02
synset item:clear_up.v.04
antonym out:overcast
antonym in:overcast
synset item:authorize.v.01
synset item:clear.v.05
synset item:pass.v.09
synset item:clear.v.07
antonym out:bounce
antonym in:bounce
synset item:clear.v.08
synset item:clear.v.09
synset item:clear.v.10
synset item:clear.v.11
synset item:clear.v.12
synset item:net.v.02
synset item:net.v.01
synset item:gain.v.08
synset item:clear.v.16
synset item:clear.v.17
synset item:acquit.v.01
antonym out:convict
antonym in:convict
synset item:clear.v.19
synset item:clear.v.20
synset item:clear.v.21
synset item:clear.v.22
synset item:clear.v.23
synset item:clear.v.24
synset item:clear.a.01
antonym out:unclear
antonym in:unclear
synset item:clear.s.02
synset item:clear.s.03
synset item:clear.a.04
antonym out:opaque
antonym in:opaque
synset item:clear.s.05
synset item:clear.s.06
synset item:clean.s.03
synset item:clear.s.08
synset item:clear.s.09
synset item:well-defined.a.02
antonym out:ill-defined
antonym in:ill-defined
synset item:clear.a.11
antonym out:cloudy
antonym in:cloudy
synset item:clean.s.02
synset item:clear.s.13
synset item:clear.s.14
synset item:clear.s.15
synset item:absolved.s.01
synset item:clear.s.17
synset item:clear.r.01
synset item:clearly.r.04
```

B.3.2 Syzygy

```
[
  Synset('clear.n.01'), Synset('open.n.01'), Synset('unclutter.v.01'),
  Synset('clear.v.02'), Synset('clear\_up.v.04'),
   ↪ Synset('authorize.v.01'),
  Synset('clear.v.05'), Synset('pass.v.09'), Synset('clear.v.07'),
  Synset('clear.v.08'), Synset('clear.v.09'), Synset('clear.v.10'),
  Synset('clear.v.11'), Synset('clear.v.12'), Synset('net.v.02'),
  Synset('net.v.01'), Synset('gain.v.08'), Synset('clear.v.16'),
  Synset('clear.v.17'), Synset('acquit.v.01'), Synset('clear.v.19'),
  Synset('clear.v.20'), Synset('clear.v.21'), Synset('clear.v.22'),
  Synset('clear.v.23'), Synset('clear.v.24'), Synset('clear.a.01'),
  Synset('clear.s.02'), Synset('clear.s.03'), Synset('clear.a.04'),
  Synset('clear.s.05'), Synset('clear.s.06'), Synset('clean.s.03'),
  Synset('clear.s.08'), Synset('clear.s.09'),
   ↪ Synset('well-defined.a.02'),
  Synset('clear.a.11'), Synset('clean.s.02'), Synset('clear.s.13'),
  Synset('clear.s.14'), Synset('clear.s.15'), Synset('absolved.s.01'),
  Synset('clear.s.17'), Synset('clear.r.01'), Synset('clearly.r.04')
]
```

Step (2) then retrieves related terms. First it gets hypernyms.

```
synset: clear.n.01
hyponyms: ---
hypernyms: innocence
holonyms: ---

synset: open.n.01
hyponyms: ---
hypernyms: area, country
holonyms: ---

synset: unclutter.v.01
hyponyms: ---
hypernyms: change, alter, modify
holonyms: ---

synset: clear.v.02
hyponyms: ---
hypernyms: make, create
holonyms: ---

synset: clear\_up.v.04
hyponyms: ---
hypernyms: ---
holonyms: ---

synset: authorize.v.01
hyponyms: approbate, approve, O.K.,
      ↪ okay, sanction,
      ↪ certificate, commission,
      ↪ declare, license, certify,
      ↪ validate, formalise
hypernyms: permit, allow, let,
      ↪ countenance
holonyms: ---

synset: clear.v.05
hyponyms: clear-cut, deforest,
      ↪ disafforest, denude, bare,
      ↪ denudate, strip, stump
hypernyms: remove, take, take\_away,
      ↪ withdraw
holonyms: ---

synset: pass.v.09
hyponyms: clear
```

```
hypernyms: succeed, win,
      ↪  come\_through,
      ↪  bring\_home\_the\_bacon,
      ↪  deliver\_the\_goods
holonyms: ---

synset: clear.v.07
hyponyms: ---
hypernyms: ---
holonyms: ---

synset: clear.v.08
hyponyms: ---
hypernyms: vanish, disappear,
      ↪  go\_away
holonyms: ---

synset: clear.v.09
hyponyms: hop
hypernyms: pass, overtake, overhaul
holonyms: ---

synset: clear.v.10
hyponyms: ---
hypernyms: clarify, clear\_up,
      ↪  elucidate
holonyms: ---

synset: clear.v.11
hyponyms: ---
hypernyms: free, discharge
holonyms: ---

synset: clear.v.12
hyponyms: ---
hypernyms: rid, free, disembarass
holonyms: ---

synset: net.v.02
hyponyms: ---
hypernyms: yield, pay, bear
holonyms: ---

synset: net.v.01
hyponyms: ---
hypernyms: profit, gain, benefit
holonyms: ---

synset: gain.v.08
hyponyms: eke\_out, squeeze\_out,
      ↪  gross, profit,
      ↪  turn\_a\_profit, rake\_in,
      ↪  shovel\_in, rake\_off,
      ↪  take\_home, bring\_home,
      ↪  yield, pay, bear
hypernyms: get, acquire
holonyms: ---

synset: clear.v.16
hyponyms: ---
hypernyms: sell
holonyms: ---

synset: clear.v.17
hyponyms: ---
hypernyms: pass, clear
holonyms: ---

synset: acquit.v.01
hyponyms: ---
hyponyms: purge, vindicate,
      ↪  whitewash, pronounce,
      ↪  label, judge
holonyms: ---

synset: clear.v.19
hyponyms: ---
hypernyms: settle, square\_off,
      ↪  square\_up, determine
holonyms: ---

synset item:clear.v.20
hyponyms: ---
hypernyms: change, alter, modify
holonyms: ---

synset item:clear.v.21
hyponyms: ---
hypernyms: empty
holonyms: ---

synset item:clear.v.22
hyponyms: ---
hypernym out:take\_out, move\_out,
      ↪  remove
holonyms: ---

synset item:clear.v.23
hyponyms: ---
```

```
hypernym out:empty
holonyms: ---

synset item:clear.v.24
hyponyms: ---
hypernym out:remove, take,
      ↪ take\_away, withdraw
holonyms: ---

synset item:clear.a.01
hyponyms: ---
hypernyms: ---
holonyms: ---

synset item:clear.s.02
hyponyms: ---
hypernyms: ---
holonyms: ---

synset item:clear.s.03
hyponyms: ---
hypernyms: ---
holonyms: ---

synset item:clear.a.04
hyponyms: ---
hypernyms: ---
holonyms: ---

synset item:clear.s.05
hyponyms: ---
hypernyms: ---
holonyms: ---

synset item:clear.s.06
hyponyms: ---
hypernyms: ---
holonyms: ---

synset item:clean.s.03
hyponyms: ---
hypernyms: ---
holonyms: ---

synset item:clear.s.08
hyponyms: ---
hypernyms: ---
holonyms: ---

synset item:clear.s.09
hyponyms: ---
hypernyms: ---
holonyms: ---

synset item:well-defined.a.02
hyponyms: ---
hypernyms: ---
holonyms: ---

synset item:clear.a.11
hyponyms: ---
hypernyms: ---
holonyms: ---

synset item:clean.s.02
hyponyms: ---
hypernyms: ---
holonyms: ---

synset item:clear.s.13
hyponyms: ---
hypernyms: ---
holonyms: ---

synset item:clear.s.14
hyponyms: ---
hypernyms: ---
holonyms: ---

synset item:clear.s.15
hyponyms: ---
hypernyms: ---
holonyms: ---

synset item:absolved.s.01
hyponyms: ---
hypernyms: ---
holonyms: ---

synset item:clear.s.17
hyponyms: ---
hypernyms: ---
holonyms: ---

synset item:clear.r.01
hyponyms: ---
hypernyms: ---
holonyms: ---

synset item:clearly.r.04
```

```
hyponyms: ---
hypernyms: ---

synset: clear.n.01
hypernyms: innocence

synset: open.n.01
hypernyms: area, country

synset: unclutter.v.01
hypernyms: change, alter, modify

synset: clear.v.02
hypernyms: make, create

synset: authorize.v.01
hyponyms: approbate, approve, O.K.,
      ↪  okay, sanction,
      ↪  certificate, commission,
      ↪  declare, license, certify,
      ↪  validate, formalise
hypernyms: permit, allow, let,
      ↪  countenance

synset: clear.v.05
hyponyms: clear-cut, deforest,
      ↪  disafforest, denude, bare,
      ↪  denudate, strip, stump
hypernyms: remove, take, take\_away,
      ↪  withdraw

synset: pass.v.09
hyponyms: clear
hypernyms: succeed, win,
      ↪  come\_through,
      ↪  bring\_home\_the\_bacon,
      ↪  deliver\_the\_goods

synset: clear.v.08
hypernyms: vanish, disappear,
      ↪  go\_away

synset: clear.v.09
hyponyms: hop
hypernyms: pass, overtake, overhaul

synset: clear.v.10
hypernyms: clarify, clear\_up,
      ↪  elucidate
```

```
holonyms: ---

synset: clear.v.11
hypernyms: free, discharge

synset: clear.v.12
hypernyms: rid, free, disembarass

synset: net.v.02
hypernyms: yield, pay, bear

synset: net.v.01
hypernyms: profit, gain, benefit

synset: gain.v.08
hyponyms: eke\_out, squeeze\_out,
      ↪  gross, profit,
      ↪  turn\_a\_profit, rake\_in,
      ↪  shovel\_in, rake\_off,
      ↪  take\_home, bring\_home,
      ↪  yield, pay, bear
hypernyms: get, acquire

synset: clear.v.16
hypernyms: sell

synset: clear.v.17
hypernyms: pass, clear

synset: acquit.v.01
hyponyms: purge, vindicate,
      ↪  whitewash, pronounce,
      ↪  label, judge

synset: clear.v.19
hypernyms: settle, square\_off,
      ↪  square\_up, determine

synset: clear.v.20
hypernyms: change, alter, modify

synset: clear.v.21
hypernyms: empty

synset: clear.v.22
hypernyms: take\_out, move\_out,
      ↪  remove

synset: clear.v.23
```

```
hypernyms: empty
```

```
synset: clear.v.24
hypernyms: remove, take, take\_away,
    ↪ withdraw
```

```
[
  Synset('clear.n.01'), Synset('open.n.01'), Synset('unclutter.v.01'),
  Synset('clear.v.02'), Synset('clear_up.v.04'),
    ↪ Synset('authorize.v.01'),
  Synset('clear.v.05'), Synset('pass.v.09'), Synset('clear.v.07'),
  Synset('clear.v.08'), Synset('clear.v.09'), Synset('clear.v.10'),
  Synset('clear.v.11'), Synset('clear.v.12'), Synset('net.v.02'),
  Synset('net.v.01'), Synset('gain.v.08'), Synset('clear.v.16'),
  Synset('clear.v.17'), Synset('acquit.v.01'), Synset('clear.v.19'),
  Synset('clear.v.20'), Synset('clear.v.21'), Synset('clear.v.22'),
  Synset('clear.v.23'), Synset('clear.v.24'), Synset('clear.a.01'),
  Synset('clear.s.02'), Synset('clear.s.03'), Synset('clear.a.04'),
  Synset('clear.s.05'), Synset('clear.s.06'), Synset('clean.s.03'),
  Synset('clear.s.08'), Synset('clear.s.09'),
    ↪ Synset('well-defined.a.02'),
  Synset('clear.a.11'), Synset('clean.s.02'), Synset('clear.s.13'),
  Synset('clear.s.14'), Synset('clear.s.15'), Synset('absolved.s.01'),
  Synset('clear.s.17'), Synset('clear.r.01'), Synset('clearly.r.04')
]
```

```
synset item:clear.n.01
hypernym out:innocence
[]
synset item:open.n.01
hypernym out:area
hypernym out:country
hypernym in:country
[]
synset item:unclutter.v.01
hypernym out:change
hypernym in:change
hypernym out:alter
hypernym out:modify
[]
synset item:clear.v.02
hypernym out:make
hypernym in:make
hypernym out:create
[]
synset item:clear_up.v.04
[]
synset item:authorize.v.01
hyponym out:approbate
hyponym out:approve
hyponym out:O.K.
hyponym out:okay
hyponym out:sanction
hyponym out:certificate
hyponym in:certificate
hyponym out:commission
hyponym out:declare
hyponym in:declare
hyponym out:license
hyponym out:licence
hyponym out:certify
hyponym out:validate
hyponym out:formalize
hyponym out:formalise
hypernym out:permit
hypernym in:permit
hypernym out:allow
hypernym in:allow
hypernym out:let
hypernym in:let
hypernym out:countenance
hypernym in:countenance
[]
synset item:clear.v.05
```

```
hyponym out:clear-cut
hyponym out:deforest
hyponym out:disforest
hyponym out:disafforest
hyponym out:denude
hyponym out:bare
hyponym in:bare
hyponym out:denudate
hyponym out:strip
hyponym out:stump
hypernym out:remove
hypernym out:take
hypernym in:take
hypernym out:take_away
hypernym out:withdraw
[]
synset item:pass.v.09
hyponym out:clear
hyponym in:clear
hypernym out:succeed
hypernym in:succeed
hypernym out:win
hypernym out:come_through
hypernym out:bring_home_the_bacon
hypernym out:deliver_the_goods
[]
synset item:clear.v.07
[]
synset item:clear.v.08
hypernym out:vanish
hypernym in:vanish
hypernym out:disappear
hypernym out:go_away
[]
synset item:clear.v.09
hyponym out:hop
hypernym out:pass
hypernym in:pass
hypernym out:overtake
hypernym out:overhaul
[]
synset item:clear.v.10
hypernym out:clarify
hypernym out:clear_up
hypernym out:elucidate
[]
synset item:clear.v.11
hypernym out:free
hypernym in:free
hypernym out:discharge
[]
synset item:clear.v.12
hypernym out:rid
hypernym out:free
hypernym in:free
hypernym out:disembarrass
[]
synset item:net.v.02
hypernym out:yield
hypernym out:pay
hypernym in:pay
hypernym out:bear
[]
synset item:net.v.01
hypernym out:profit
hypernym out:gain
hypernym in:gain
hypernym out:benefit
hypernym in:benefit
[]
synset item:gain.v.08
hyponym out:eke_out
hyponym out:squeeze_out
hyponym out:gross
hyponym out:profit
hyponym out:turn_a_profit
hyponym out:rake_in
hyponym out:shovel_in
hyponym out:rake_off
hyponym out:take_home
hyponym out:bring_home
hyponym out:yield
hyponym out:pay
hyponym in:pay
hyponym out:bear
hypernym out:get
hypernym out:acquire
[]
synset item:clear.v.16
hypernym out:sell
[]
synset item:clear.v.17
hypernym out:pass
hypernym in:pass
hypernym out:clear
hypernym in:clear
[]
synset item:acquit.v.01
hyponym out:purge
hyponym out:vindicate
```

```
hyponym out:whitewash
hypernym out:pronounce
hypernym in:pronounce
hypernym out:label
hypernym out:judge
hypernym in:judge
[]
synset item:clear.v.19
hypernym out:settle
hypernym out:square_off
hypernym out:square_up
hypernym out:determine
hypernym in:determine
[]
synset item:clear.v.20
hypernym out:change
hypernym in:change
hypernym out:alter
hypernym out:modify
[]
synset item:clear.v.21
hypernym out:empty
hypernym in:empty
[]
synset item:clear.v.22
hypernym out:take_out
hypernym out:move_out
hypernym out:remove
[]
synset item:clear.v.23
hypernym out:empty
hypernym in:empty
[]
synset item:clear.v.24
hypernym out:remove
hypernym out:take
hypernym in:take
hypernym out:take_away
hypernym out:withdraw
[]
synset item:clear.a.01
[]
synset item:clear.s.02
[]
synset item:clear.s.03
[]
synset item:clear.a.04
[]
synset item:clear.s.05
[]
synset item:clear.s.06
[]
synset item:clean.s.03
[]
synset item:clear.s.08
[]
synset item:clear.s.09
[]
synset item:well-defined.a.02
[]
synset item:clear.a.11
[]
synset item:clean.s.02
[]
synset item:clear.s.13
[]
synset item:clear.s.14
[]
synset item:clear.s.15
[]
synset item:absolved.s.01
[]
synset item:clear.s.17
[]
synset item:clear.r.01
[]
synset item:clearly.r.04
[]
```

B.4 Bing JSON Results


```
{"d": { "results": [
  { "__metadata":
    { "uri": "https://api.datamarket.azure.com/Data.ashx/Bing/Search/Image
         ↪ ?Query=\u0027kittens\u0027&$skip=0&$top=1",
      "type": "ImageResult"
    },
    "ID": "e09072a2-faf3-47ac-b77d-46a8df8941aa",
    "Title": "Cute Kittens - Pictures - The Wondrous Pics",
```

```
    "MediaUrl": "http://wondrouspics.com/wp-content/uploads/2011/12/Cute-Kitten2.jpg",
    "SourceUrl": "http://wondrouspics.com/cute-kittens-pictures/",
    "DisplayUrl": "wondrouspics.com/cute-kittens-pictures",
    "Width": "1440",
    "Height": "900",
    "FileSize": "238015",
    "ContentType": "image/jpeg",
    "Thumbnail":
    { "__metadata":
      { "type": "Bing.Thumbnail"
      },
      "MediaUrl": "http://ts2.mm.bing.net/th?id=OIP.M5692e5d79242507e30600fd54639316cH0&pid=15.1",
      "ContentType": "image/jpg",
      "Width": "480",
      "Height": "300",
      "FileSize": "13856"
    }
  },
  ],
  "__next": "https://api.datamarket.azure.com/Data.ashx/Bing/Search/Image?Query=\u0027kittens\u0027&$skip=50"
}}
```


B.5 Random Quotes

```python
def getrandquote():
  root_path = os.path.dirname(os.path.abspath(__file__))
  root_path = root_path[:-4]
  corpus_root = root_path + '/app/static/corpus'
  path_b = corpus_root + '/quotes.txt'
  quotes_text = codecs.open(path_b, "r", encoding='utf-8')
  quotestext = quotes_text.readlines()
  quotes_text.close()
  return random.choice(quotestext)
```

B.6 Stopwords

B.6.1 English

i, me, my, myself, we, our, ours, ourselves, yo, your, yours, yourself, yourselves, he, him, his, himself, she, her, hers, herself, it, its, itself, they, them, their, theirs, themselves, what, which, who, whom, this, that, these, those, am, is, are, was, were, be, been, being, have, has, had, having, do, does, did, doing, a, an, the, and, but, if, or, because, as, until, while, of, at, by, for, with, about, against, between, into, through, during, before, after, above, below, to, from, up, down, in, out, on, off, over, under, again, further, then, once, here, there, when, where, why, how, all, any, both, each, few, more, most, other, some, such, no, nor, not, only, own, same, so, than, too, very, s, t, can, will, just, don, should, now

B.6.2 French

a, aux, avec, ce, ces, dans, de, des, d, elle, en, et, eux, il, je, la, le, leur, lui, ma, mais, me, même, mes, moi, mon, ne, nos, notre, nous, on, o, par, pas, pour, q, que, qui, sa, se, ses, son, sur, ta, te, tes, toi, ton, t, un, une, vos, votre, vous, été, étée, étées, étés, étant, étante, étants, étantes, suis, es, est, sommes, êtes, sont, serai, seras, sera, serons, serez, seront, serais, serait, serions, seriez, seraient, étais, était, étions, étiez, étaient, fus, fut, fûmes, fûtes, furent, sois, soit, soyons, soyez, soient, fusse, fusses, fût, fussions, fussiez, fussent, ayant, ayante, ayantes, ayants, e, eue, eues, eus, ai, as, avons, avez, ont, aurai, auras, aura, aurons, aurez, auront, aurais, aurait, aurions, auriez, auraient, avais, avait, avions, aviez, avaient, eut, eûmes, eûtes, eurent, aie, aies, ait, ayons, ayez, aient, eusse, eusses, eût, eussions, eussiez, eussent

B.6.3 German

aber, alle, allem, allen, aller, alles, als, also, am, an, ander, andere, anderem, anderen, anderer, anderes, anderm, andern, anderr, anders, auch, auf, aus, bei, bin, bis, bist, da, damit, dann, der, den, des, dem, die, das, daß, derselbe, derselben, denselben, desselben, demselben, dieselbe, dieselben, dasselbe, daz, dein, deine, deinem, deinen, deiner, deines, denn, derer, dessen, dich, dir, d, dies, diese, diesem, diesen, dieser, dieses, doch, dort, durch, ein, eine, einem, einen, einer, eines, einig, einige, einigem, einigen, einiger, einiges, einmal, er, ihn, ihm, es, etwas, euer, eure, eurem, euren, eurer, eures, für, gegen, gewesen, hab, habe, haben, hat, hatte, hatten, hier, hin, hinter, ich, mich, mir, ihr, ihre, ihrem, ihren, ihrer, ihres, euch, im, in, indem, ins, ist, jede, jedem, jeden, jeder, jedes, jene, jenem, jenen, jener, jenes, jetzt, kann, kein, keine, keinem, keinen, keiner, keines, können, könnte, machen, man, manche, manchem, manchen, mancher, manches, mein, meine, meinem, meinen, meiner, meines, mit, muss, musste, nach, nicht, nichts, noch, nun, nur, ob, oder, ohne, sehr, sein, seine, seinem, seinen, seiner, seines, selbst, sich, sie, ihnen, sind, so, solche, solchem, solchen, solcher, solches, soll, sollte, sondern, sonst, über, um, und, uns, unse, unsem, unsen, unser, unses, unter, viel, vom, von, vor, während, war, waren, warst, was, weg, weil, weiter, welche, welchem, welchen, welcher, welches, wenn, werde, werden, wie, wieder, will, wir, wird, wirst, wo, wollen, wollte, würde, würden, z, zum, zur, zwar, zwischen

B.7 Image Spiral

```
function createSpiral(imglist){
  if (imglist.length === 10){
    var spiral_code = ' \
    <div class="spouter"> \
      <div class="spleft"> \
        <div class="spltop"> \
          <div class="spltleft"> \
            <a id="a3" class="spimg" href="'+imglist[3][2]+'" ><img id="img3" src="'+imglist[3][1]+'" title="'+imglist[3][0]+' --- '+imglist[3][3]+'" height="210" width="210"/></a> \
          </div> \
          <div class="spltright"> \
            <div class="spltrtop"> \
              <a id="a8" class="spimg" href="'+imglist[8][2]+'" ><img id="img8" src="'+imglist[8][1]+'" title="'+imglist[8][0]+'" height="130" width="130"/></a> \
            </div> \
            <div class="spltrbottom"> \
              <div class="spltrbleft"> \
                <div class="spltrbltop"> \
                  <div class="spltrbltleft"> \
                    <a id="a0" class="spimg" href="'+imglist[0][2]+'" ><img id="img0" src="'+imglist[0][1]+'" title="'+imglist[0][0]+'" height="30" width="30"/></a> \
                  </div> \
                  <div class="spltrbltright"> \
                    <div class="spltrbltrtop"> \
                      <a id="a1" class="spimg" href="'+imglist[1][2]+'" ><img id="img1" src="'+imglist[1][1]+'" title="'+imglist[1][0]+'" height="20" width="20"/></a> \
                    </div> \
                    <div class="spltrbltrbottom"> \
                      <div class="spltrbltrbleft"> \
                        <a id="a5" class="spimg" href="'+imglist[5][2]+'" ><img id="img5" src="'+imglist[5][1]+'" title="'+imglist[5][0]+'" height="10" width="10"/></a> \
                      </div> \
                      <div class="spltrbltrbright"> \
                        <a id="a6" class="spimg" href="'+imglist[6][2]+'" ><img id="img6" src="'+imglist[6][1]+'" title="'+imglist[6][0]+'" height="10" width="10"/></a> \
                      </div> \
                    </div> \
                  </div> \
                </div> \
                <div class="spltrblbottom"> \
                  <a id="a7" class="spimg" href="'+imglist[7][2]+'" ><img id="img7" src="'+imglist[7][1]+'" title="'+imglist[7][0]+'" height="50" width="50"/></a> \
```

```
        </div> \
       </div> \
       <div class="spltrbright"> \
        <a id="a2" class="spimg" href="'+imglist[2][2]+'" ><img id="img2" src="'+imglist[2][1]+'" title="'+imglist[2][0]+'" height="80" width="80"/></a> \
       </div> \
      </div> \
     </div> \
    </div> \
    <div class="splbottom"> \
     <a id="a9" class="spimg" href="'+imglist[9][2]+'" ><img id="img9" src="'+imglist[9][1]+'" title="'+imglist[9][0]+'" height="340" width="340"/></a> \
    </div> \
   </div> \
   <div class="spright"> \
    <a id="a4" class="spimg" href="'+imglist[4][2]+'" ><img id="img4" src="'+imglist[4][1]+'" title="'+imglist[4][0]+'" height="550" width="550"/></a> \
   </div> \
  </div> \
  ';
  var list_code = [];
  for (i in imglist) {
   var img = ' \
    <div class="w3-col s12 m6 l3 w3-padding"> \
     <a href="'+imglist[i][2]+'"> \
      <img src="'+imglist[i][1]+'" \
      title="'+imglist[i][0]+'" style="width:100%"> \
     </a> \
    </div> \
   ';
   list_code.push(img);
  } // end for
  $('#img_spiral_div').html(spiral_code);
  $('#img_list_div').html(list_code);
 } // end if
 else{
  // console.log("inside else");
  $('.img_empty').wrap("<div>Not enough results found.</div>");
 } // end else
}
```

WordNet

C

The sections below show an example result returned by WordNet for the query 'clear'. It is split into four parts, nouns (n), verbs (v), adjectives (adj) and adverbs (adv). Each entry is preceded by an S for synset.

C.1 Noun

S: (n) ***clear*** (the state of being free of suspicion) "investigation showed that he was in the clear"
S: (n) open, ***clear*** (a clear or unobstructed space or expanse of land or water) "finally broke out of the forest into the open"

C.2 Verb

S: (v) unclutter, ***clear*** (rid of obstructions) "Clear your desk"
S: (v) ***clear*** (make a way or path by removing objects) "Clear a path through the dense forest"
S: (v) clear up, ***clear***, light up, brighten (become clear) "The sky cleared after the storm"

S: (v) authorize, authorise, pass, ***clear*** (grant authorization or clearance for) “Clear the manuscript for publication”; “The rock star never authorized this slanderous biography”

S: (v) ***clear*** (remove) “clear the leaves from the lawn”; “Clear snow from the road”

S: (v) pass, ***clear*** (go unchallenged; be approved) “The bill cleared the House”

S: (v) ***clear*** (be debited and credited to the proper bank accounts) “The check will clear within 2 business days”

S: (v) ***clear*** (go away or disappear) “The fog cleared in the afternoon”

S: (v) ***clear***, top (pass by, over, or under without making contact) “the balloon cleared the tree tops”

S: (v) ***clear***, clear up, shed light on, crystallize, crystallise, crystalize, crystalise, straighten out, sort out, enlighten, illuminate, elucidate (make free from confusion or ambiguity; make clear) “Could you clarify these remarks?”; “Clear up the question of who is at fault”

S: (v) ***clear*** (free from payment of customs duties, as of a shipment) “Clear the ship and let it dock”

S: (v) ***clear*** (clear from impurities, blemishes, pollution, etc.) “clear the water before it can be drunk”

S: (v) net, ***clear*** (yield as a net profit) “This sale netted me $1 million”

S: (v) net, sack, sack up, clear (make as a net profit) “The company cleared $1 million”

S: (v) gain, take in, ***clear***, make, earn, realize, realise, pull in, bring in (earn on some commercial or business transaction; earn as salary or wages) “How much do you make a month in your new job?”; “She earns a lot in her new job”; “this merger brought in lots of money”; “He clears $5,000 each month”

S: (v) ***clear*** (sell) “We cleared a lot of the old model cars”

S: (v) ***clear*** (pass an inspection or receive authorization) “clear customs”

S: (v) acquit, assoil, ***clear***, discharge, exonerate, exculpate (pronounce not guilty of criminal charges) “The suspect was cleared of the murder charges”

S: (v) ***clear***, solve (settle, as of a debt) “clear a debt”; “solve an old debt”

S: (v) ***clear*** (make clear, bright, light, or translucent) “The water had to be cleared through filtering”

S: (v) ***clear*** (rid of instructions or data) “clear a memory buffer”

S: (v) ***clear*** (remove (people) from a building) “clear the patrons from the theater after the bomb threat”

S: (v) ***clear*** (remove the occupants of) “Clear the building”

S: (v) ***clear***, clear up (free (the throat) by making a rasping sound) “Clear the throat”

C.3 Adjective

S: (adj) ***clear*** (readily apparent to the mind) “a clear and present danger”; “a clear explanation”; “a clear case of murder”; “a clear indication that she was angry”; “gave us a clear idea of human nature”

S: (adj) ***clear*** (free from confusion or doubt) “a complex problem requiring a clear head”; “not clear about what is expected of us”

S: (adj) ***clear***, open (affording free passage or view) “a clear view”; “a clear path to victory”; “open waters”; “the open countryside”

S: (adj) ***clear*** (allowing light to pass through) “clear water”; “clear plastic bags”; “clear glass”; “the air is clear and clean”

S: (adj) ***clear*** (free from contact or proximity or connection) “we were clear of the danger”; “the ship was clear of the reef”

S: (adj) ***clear*** (characterized by freedom from troubling thoughts (especially guilt)) “a clear conscience”; “regarded her questioner with clear untroubled eyes”

S: (adj) clean, ***clear***, light, unclouded ((of sound or color) free from anything that dulls or dims) “efforts to obtain a clean bass in orchestral recordings”; “clear laughter like a waterfall”; “clear reds and blues”; “a light lilting voice like a silver bell”

S: (adj) ***clear***, unmortgaged ((especially of a title) free from any encumbrance or limitation that presents a question of fact or law) “I have clear title to this property”

S: (adj) ***clear***, clean-cut, clear-cut (clear and distinct to the senses; easily perceptible) “as clear as a whistle”; “clear footprints in the snow”; “the letter brought back a clear image of his grandfather”; “a spire clean-cut against the sky”; “a clear-cut pattern”

S: (adj) well-defined, ***clear*** (accurately stated or described) “a set of well-defined values”

S: (adj) ***clear*** (free from clouds or mist or haze) “on a clear day”

S: (adj) clean, ***clear*** (free of restrictions or qualifications) “a clean bill of health”; “a clear winner”

S: (adj) ***clear*** (free from flaw or blemish or impurity) “a clear perfect diamond”; “the clear complexion of a healthy young woman”

S: (adj) ***clear*** (clear of charges or deductions) “a clear profit”

S: (adj) ***clear***, decipherable, readable (easily deciphered)

S: (adj) absolved, ***clear***, cleared, exculpated, exonerated, vindicated (freed from any question of guilt) “is absolved from all blame”; “was now clear of the charge of cowardice”; “his official honor is vindicated”

S: (adj) ***clear***, percipient (characterized by ease and quickness in perceiving) "clear mind"; "a percipient author"

C.4 Adverb

S: (adv) ***clear***, all the way (completely) "read the book clear to the end"; "slept clear through the night"; "there were open fields clear to the horizon"

S: (adv) clearly, ***clear*** (in an easily perceptible manner) "could be seen clearly under the microscope"; "She cried loud and clear"

Git History

D

D.1 Website Repository

```
* 0fbbfcd  |  Sat 15 Oct 2016  (HEAD -> master, origin/master, origin/HEAD) |
     ↪  Deleted local log
* 504bcfa  |  Sat 15 Oct 2016  |  Added meta description
* a7e4a5d  |  Sun 02 Oct 2016  |  Updated about
*   64b0c9a  |  Sun 02 Oct 2016  |  Work in progress
|\
| | *   76e1dbb  |  Fri 12 Aug 2016  (origin/live, live) |  Merge pull request
     ↪  #15 from Fania/master
| | |\
| | |/
| |/|
| * |   5fd81f9  |  Fri 12 Aug 2016  (tag: v.4.1) |  Merge pull request #14
     ↪  from Fania/thesis
| |\ \
| | * | 331ddfe  |  Fri 12 Aug 2016  (origin/thesis) |  Comment out prints
| | * | 878cade  |  Fri 12 Aug 2016  |  Log updates
| | * | d894400  |  Fri 12 Aug 2016  |  Fixed results-reverbs-origins numbers
| | * | 7a64760  |  Fri 12 Aug 2016  |  Fixed holonyms and meronyms
| |/ /
| * |   8efc58a  |  Fri 12 Aug 2016  |  Merge pull request #13 from Fania/live
| |\ \
| | |/
```

```
* | | c4b9eea  |  Thu 11 Aug 2016  |  About section update
|/ /
| * 0f6353a  |  Thu 11 Aug 2016  (tag: v.4.0) |  Added new paper to about
      ↪  section
| * 0fd2af4  |  Thu 11 Aug 2016  |  Cleaned files
| *   1bf06c8  |  Thu 11 Aug 2016  |  Merge branch 'live' of
      ↪  https://github.com/Fania/pata.physics.wtf into live
| |\
| * | 0b66aaf  |  Thu 11 Aug 2016  |  Enabled image and video in menu
| | *   15bdb1e  |  Thu 11 Aug 2016  |  Merge pull request #12 from
      ↪  Fania/master
| | |\
| |/ /
| | /
| |/
|/|
* |   23ea5c6  |  Thu 11 Aug 2016  |  Merge commit
|\ \
| |/
* | 8f5708d  |  Thu 11 Aug 2016  |  Changed date
* |   3b32674  |  Thu 11 Aug 2016  |  Merge pull request #11 from Fania/dev
|\ \
| * \   d480bfa  |  Thu 11 Aug 2016  (dev) |  Merge pull request #9 from
      ↪  Fania/prints
| |\ \
| | * | 4adf2a2  |  Thu 11 Aug 2016  (prints) |  Added meronyms, got rid of
      ↪  prints again
| | * | b454247  |  Thu 04 Aug 2016  |  ppsent
| | * | dbb33cd  |  Thu 04 Aug 2016  |  ppsent fix
| | * | ef33367  |  Sat 23 Jul 2016  |  stuff
| | * | 6edc4a3  |  Tue 19 Jul 2016  |  ppsent prints
| | * | 3fa7ab6  |  Tue 19 Jul 2016  |  about to change ppsent
| |/ /
| * | 96a7f8f  |  Mon 18 Jul 2016  |  log changes
| * | abdc8a1  |  Wed 06 Jul 2016  |  Revert "failed getty (3callspersec)"
| * | 975edb6  |  Wed 06 Jul 2016  |  failed getty (3callspersec)
| * | c9b6b82  |  Wed 06 Jul 2016  |  Prepping for rewriting Getty and Bing to
      ↪  run 10 times
| * | 62ad371  |  Wed 06 Jul 2016  |  Made Flickr Img List Vessel function
      ↪  standalone
| * | 8eaa132  |  Wed 06 Jul 2016  |  Fixed log printouts
| * | e6c1d5f  |  Tue 05 Jul 2016  |  flickr working with 10 images
| * | 916faa4  |  Tue 05 Jul 2016  |  10 images working
| * | 46b4209  |  Tue 05 Jul 2016  |  partially working imagelistvessel ith 10
      ↪  results
| * | efe1893  |  Tue 05 Jul 2016  |  still fucking in progress
| * | f549bab  |  Sun 03 Jul 2016  |  work in progress
| * | f4fdbea  |  Sun 03 Jul 2016  |  trap for empty data items in js
| * | 178b63a  |  Sun 03 Jul 2016  |  half assed img search 10 results done
| * | b00fd7e  |  Sun 03 Jul 2016  |  Video working
```

```
| * | c99e8eb  |  Sun 03 Jul 2016  |  Getty works too.
| * | aeb081d  |  Sat 02 Jul 2016  |  Getty API down, flickr and Bing work
| * | d228062  |  Sat 02 Jul 2016  |  Flickr and Bing work
| * | c53b060  |  Sat 02 Jul 2016  |  Query is now 1 random item from pata set
| * | ac86e07  |  Sat 02 Jul 2016  |  Image search works again somehow
| * | 0c7713b  |  Tue 07 Jun 2016  |  work in progress
| * | 4fed23c  |  Tue 07 Jun 2016  |  added log with date and time
| * | 52c4394  |  Tue 07 Jun 2016  |  datetime log
|/ /
| * 2dfa9d3  |  Mon 06 Jun 2016  (tag: v.3.5) |  Ready to deploy
| * 0cc8be1  |  Mon 06 Jun 2016  |  Added log
|/
* 2c51dbc  |  Mon 06 Jun 2016  |  dot creeped into code
*   e49efa5  |  Mon 06 Jun 2016  |  pre-merge commit
|\
* | 6c43cea  |  Mon 06 Jun 2016  |  pre-merge commit again
* |   af51ef4  |  Mon 06 Jun 2016  |  pre-merge commit
|\ \
* \ \   4859e65  |  Mon 06 Jun 2016  |  pre merge commit
|\ \ \
* | | | 26a1d9c  |  Mon 06 Jun 2016  |  test commit
| | | * 3ef1630  |  Mon 06 Jun 2016  |  readme updates
| | | * 06b070a  |  Mon 16 May 2016  |  git log stuff
| | | *   10f61f9  |  Sun 08 May 2016  |  Merge remote-tracking branch
      ↪  'refs/remotes/origin/master' into api
| | | |\
| |_|_|/
|/| | |
* | | | aa58f79  |  Sat 09 Apr 2016  |  screenshot
* | | | bea5474  |  Sat 09 Apr 2016  |  added screenshot
* | | | 7072f33  |  Sat 09 Apr 2016  (tag: v2.0) |  gitignore
| | | * 0e53ee6  |  Sat 09 Apr 2016  |  Working on new server stuff
| | | * a082595  |  Tue 05 Apr 2016  |  Update to Todo file
| | | * 6fbbd49  |  Sun 20 Mar 2016  |  pataphysicalisation work in progress
| | | * 80412e1  |  Sun 20 Mar 2016  |  Flickr, Getty and Bing working!
| | | * 00fd2c5  |  Sun 20 Mar 2016  |  radio buttons update properly
| | | * 7e03f4b  |  Wed 16 Mar 2016  |  Flickr OK Getty OK
| | | * df847d9  |  Wed 16 Mar 2016  |  Revert "Test revert commit"
| | | * 630bf1a  |  Wed 16 Mar 2016  |  Test revert commit
| | | * 62dfc0b  |  Wed 16 Mar 2016  |  Getty OK, Flickr NO
| | | * 17cff52  |  Wed 16 Mar 2016  |  Getty working
| | | * fde271f  |  Wed 16 Mar 2016  |  FUCK THIS SHIT
| | | * 9095fa1  |  Mon 14 Mar 2016  |  Flickr API working (spiral + list)
| | | * adb55cf  |  Sat 12 Mar 2016  |  img spiral works
| | | * e64e995  |  Sat 12 Mar 2016  |  work in progress
| | |/
| | * fa0e818  |  Fri 11 Mar 2016  |  Pre branch img vid
| | * 83032fd  |  Fri 11 Mar 2016  |  Fixed textfield search default text
      ↪  problem
| | * e6609fe  |  Thu 10 Mar 2016  |  Emails fixed
```

```
| | * 9fecc8e   |  Thu 10 Mar 2016   |  Fixed javascript error problem
| | * 1ce1893   |  Wed 09 Mar 2016   |  Work in progress
| | * 2999784   |  Tue 08 Mar 2016   |  Radio button styles
| | * 844817d   |  Tue 08 Mar 2016   |  Shakespeare working
| | * f6f4e38   |  Tue 08 Mar 2016   |  Changed setupcorpus function (unfinished)
| | * 3cfb7e2   |  Tue 08 Mar 2016   |  Started shakespeare stuff
| | * 5e93e11   |  Fri 19 Feb 2016   |  Added a few cheats
| | * 5daf3b7   |  Wed 17 Feb 2016   |  surface updates
| | * e1f7c12   |  Tue 22 Dec 2015   |  added quotes and shakespeare
| | * 44e6211   |  Sat 31 Oct 2015   |  Stuff for thesis
| | * 9b1ec61   |  Wed 19 Aug 2015   |  Getty works sort of
| | * 71437f6   |  Tue 18 Aug 2015   |  Flickr and Bing work, radio buttons work
| | * 6c552aa   |  Wed 12 Aug 2015   |  Fixed image problem but not video.
| * | 1cbb63d   |  Tue 11 Aug 2015   |  Update textsurfer.py
| |/
| * 0ebff0d   |  Tue 11 Aug 2015   |  Analytics enabled again
| * 703f977   |  Tue 11 Aug 2015   |  Problems solved.
| * 74a1fae   |  Tue 11 Aug 2015   |  About to change l_dict to dict of dict
| * 0935b23   |  Mon 10 Aug 2015   |  BUG FUCKER
| * 4f7d91e   |  Mon 10 Aug 2015   |  Turn debug off
| * 58f0c2b   |  Mon 10 Aug 2015   |  Button styling done
| * 59add58   |  Mon 10 Aug 2015   |  Email problem solved
| *   f1b2d40   |  Sun 09 Aug 2015   |  Merge branch 'deploy' into thesis
| |\
| | * 435cb2d   |  Sun 09 Aug 2015   |  Deployment works, added analytics
| | * 8a63dc7   |  Sat 08 Aug 2015   |  gunicorn runs locally fine.
| | * 2861407   |  Sat 08 Aug 2015   |  Revert 5f2c957..4026965
| | * 4026965   |  Sat 08 Aug 2015   |  Tests
| * |   8f2eeab   |  Sat 08 Aug 2015   |  Merge branch 'w3' into thesis
| |\ \
| | |/
| | * 5f2c957   |  Sat 08 Aug 2015   |  Stuff
| | * 873153c   |  Fri 07 Aug 2015   |  Tiny cleanup
| | * 05d5760   |  Thu 06 Aug 2015   |  Random Poems and Emailing works
| | * 657126c   |  Wed 05 Aug 2015   |  Random poems work - without links though
| | * 3d31ea9   |  Wed 05 Aug 2015   |  Randomise still only works once, count ok
| | * 5f1d45b   |  Wed 05 Aug 2015   |  Randomise poem works ONCE
| | * c583341   |  Wed 05 Aug 2015   |  Poem subtabs, email poems done
| | * f1b3878   |  Wed 05 Aug 2015   |  Hiding divs
| | * a6939c4   |  Tue 04 Aug 2015   |  huh?
| | * e6b411d   |  Tue 04 Aug 2015   |  Poem emails WORK Fuck YEAH!
| | * 4b6b170   |  Tue 04 Aug 2015   |  Test email
| | * 24e356c   |  Tue 04 Aug 2015   |  Better load icon
| | * e6ae736   |  Tue 04 Aug 2015   |  loading icon version 1
| | * 51b43e2   |  Tue 04 Aug 2015   |  Added 4th pictures
| | * f2d8a83   |  Mon 03 Aug 2015   |  Minor fixes
| * |   1ddb03d   |  Mon 03 Aug 2015   |  Merge branch 'w3' into thesis
| |\ \
| | |/
| | * ca4eab3   |  Mon 03 Aug 2015   |  Pretty good state.
```

```
| | * 9370334  |  Mon 03 Aug 2015  |  working on list display of images
| | * e1f1ead  |  Mon 03 Aug 2015  |  Stylesheets sorted and cleaned files
| * |   9732d5b  |  Mon 03 Aug 2015  |  Merge branch 'w3' into thesis
| |\ \
| | |/
| | * f0a4c40  |  Sun 02 Aug 2015  |  Minor errors left
| | * 4c94b11  |  Sun 02 Aug 2015  |  Styles ok. still some errors in vids?
| | * 5ab4bb3  |  Sun 02 Aug 2015  |  Videoresults work
| | * d575762  |  Sun 02 Aug 2015  |  Videos works and styled
| | * 906be06  |  Sun 02 Aug 2015  |  Starting videos
| | * 0d29479  |  Sun 02 Aug 2015  |  Images working with occasional error
      ↪  (unicode?)
| | * 8e9f7bf  |  Sun 02 Aug 2015  |  Http response 200
| | * 09706d8  |  Sat 01 Aug 2015  |  Stuff
| | * 57ff730  |  Sat 01 Aug 2015  |  Bing or Flickr not working yet
| | * c85b61d  |  Sat 01 Aug 2015  |  Prep for image results done
| | * 052b55d  |  Sat 01 Aug 2015  |  Starting image results
| | * 4f08696  |  Sat 01 Aug 2015  |  So far so good.
| | * 1ff8370  |  Sat 01 Aug 2015  |  good version
| | * 50f8f00  |  Sat 01 Aug 2015  |  About to play with poem width
| | * ea3de12  |  Fri 31 Jul 2015  |  Styling in progress
| | * f69e4c9  |  Fri 31 Jul 2015  |  Work in progresss
| | * 3c8ae12  |  Fri 31 Jul 2015  |  Work in progress
| | * 616a1b3  |  Fri 31 Jul 2015  |  style start
| |/
| * fafd254  |  Thu 30 Jul 2015  |  Todos
| *   0fc8807  |  Wed 29 Jul 2015  |  Merge branch 'text' into thesis
| |\
| | * 95c2071  |  Wed 29 Jul 2015  |  Minor spacing
| | * 3a9fd4b  |  Wed 29 Jul 2015  |  Rewritten syzygy function
| | * 12afae9  |  Wed 29 Jul 2015  |  todo update
| | * f52186f  |  Wed 29 Jul 2015  |  pp_sent changed based on punctuation
      ↪  marks
| * |   5d69975  |  Wed 29 Jul 2015  |  Merge branch 'text' into thesis
| |\ \
| | |/
| | * a01f935  |  Wed 29 Jul 2015  |  Todo
| | * 489ccf5  |  Wed 29 Jul 2015  |  Up arrows
| | * 6b2736f  |  Wed 29 Jul 2015  |  Title fixes, links etc
| | * f3b874b  |  Wed 29 Jul 2015  |  poems display for less than 14 works
| * |   a58abe2  |  Wed 29 Jul 2015  |  Merge branch 'poetry' into thesis
| |\ \
| | |/
| | * 1654fa5  |  Tue 28 Jul 2015  |  split poem into stanzas from correct
      ↪  files
| | * 0562be7  |  Tue 28 Jul 2015  |  Fixed counts and search clicking
| | * 15940da  |  Tue 28 Jul 2015  |  centre poems work
| | * f9338f1  |  Mon 27 Jul 2015  |  Style fixes
| | * ed81657  |  Mon 27 Jul 2015  |  Tabs done. Need style.
| | * 28939a6  |  Mon 27 Jul 2015  |  3 Forms working with titles
```

```
| | * fe57ba8  |  Mon 27 Jul 2015  |  Change in data structure all_sens
| | * 653c6e6  |  Mon 27 Jul 2015  |  Temp commit
| | * 8fb6423  |  Sun 26 Jul 2015  |  Poem design ok
| | * 6c4f4fc  |  Sun 26 Jul 2015  |  triplet poem working
| | * 0ad1e63  |  Sun 26 Jul 2015  |  3 scroll results working omg
| | * 262c60a  |  Sun 26 Jul 2015  |  sentence scroll clin_sens working
| | * d900391  |  Sun 26 Jul 2015  |  scroll text works on single list
| | * 19b8570  |  Sun 26 Jul 2015  |  working scroll img demo
| | * 74cc973  |  Sat 25 Jul 2015  |  Work in progress
| | * 624bbc2  |  Sat 25 Jul 2015  |  Added more files, hover title works
| * |   8e9257f  |  Sat 25 Jul 2015  |  Merge branch 'poetry' into thesis
| |\ \
| | |/
| | * f1a627b  |  Sat 25 Jul 2015  |  Counts done
| |/
| * 811f38a  |  Sat 25 Jul 2015  |  Fixed origins error
| * 112ab28  |  Sat 25 Jul 2015  |  Got rid of empty results
| * 95f1bc7  |  Sat 25 Jul 2015  |  All three algorithms work
| * eae3139  |  Sat 25 Jul 2015  |  antinomy working
| * 9de06b4  |  Sat 25 Jul 2015  |  Dave and sex works again
| * ce63862  |  Sat 25 Jul 2015  |  Restructure of clinamen
| * a4c3bd8  |  Fri 24 Jul 2015  |  Cleaned up files.
| * 4067361  |  Fri 24 Jul 2015  |  Working mostly
| * 9778834  |  Fri 24 Jul 2015  |  impression works but not clear?
| * d19a52a  |  Fri 24 Jul 2015  |  pp_sent works but not website
| * 1c5d945  |  Fri 24 Jul 2015  |  complete corpus working but slow
| * 7ac8697  |  Wed 22 Jul 2015  |  Count works properly.
| * 9b318e1  |  Wed 22 Jul 2015  |  Works.
| * 22d5e9d  |  Wed 22 Jul 2015  |  Added to library
| * ae77a28  |  Mon 20 Jul 2015  |  templates loop not quite right
| * b8ba9b7  |  Mon 20 Jul 2015  |  Almost working. template needs fix
| * 47b2766  |  Mon 20 Jul 2015  |  Cleaned corpus files
| * 71e7153  |  Fri 17 Jul 2015  |  Library added and simple search works
| * 95aed8a  |  Thu 16 Jul 2015  |  Library setup Schwob
| * a260bec  |  Tue 14 Jul 2015  |  Do some stuff to library
| * 9ead88b  |  Tue 14 Jul 2015  |  Print first 10 words of each book
| * 47a5ae3  |  Tue 14 Jul 2015  |  Start for library
| * c99b5ff  |  Wed 24 Jun 2015  |  Added more printouts
| * daf6a5d  |  Wed 24 Jun 2015  |  Added printouts
|/
* fa3ffc7  |  Fri 22 May 2015  (tag: v3.0) |  Updated Readme for IOCT server
* 2d51804  |  Sat 21 Mar 2015  |  added algorithms summary
* 81c1b12  |  Thu 31 Jul 2014  |  pata.fania.eu
* 7545e3e  |  Mon 28 Jul 2014  |  Slight style change
* b5191a7  |  Sun 27 Jul 2014  |  Readme
* 2b3be93  |  Sun 27 Jul 2014  |  readme change
* 97847d9  |  Sun 27 Jul 2014  |  Minimally responsive now.
* 8e1b77a  |  Sun 27 Jul 2014  |  Updated nltk to 3.0.0b1
* ce9b9b6  |  Sun 27 Jul 2014  |  Added macreqs
* 34c883a  |  Sun 27 Jul 2014  |  About edit and WINREQS
```

```
* 7a9432b  |  Sat 26 Jul 2014  |  Submenu for About
* 76ba522  |  Sat 26 Jul 2014  |  Added autofocus for search boxes
* c1c0f83  |  Sat 26 Jul 2014  |  Fixed mac word net error
* 6efe43c  |  Wed 16 Jul 2014  |  stuff
* 554b354  |  Wed 16 Jul 2014  |  Changed error bug link style
* 6aabd7b  |  Wed 16 Jul 2014  |  Fixed typos in quotes, changed errors
* e7f40f3  |  Wed 16 Jul 2014  |  Added more quotes
* 84561d7  |  Tue 15 Jul 2014  |  Favicon, errors
* 593b0c0  |  Mon 14 Jul 2014  |  Quotes
* 2915771  |  Sun 13 Jul 2014  |  Added Icons. Style changes.
* 7eb40d0  |  Sun 13 Jul 2014  |  Inline links now working.
* 9077ed2  |  Sun 13 Jul 2014  |  So far so good. Working nicely.
* 13112b7  |  Sat 12 Jul 2014  |  Style for p02 and p03. Golden spiral test.
* c911264  |  Sat 12 Jul 2014  |  Fixed unicode problem, changes file
     ↪  structure.
* 8df543b  |  Sat 12 Jul 2014  |  Updated TODO
* eeb9f59  |  Sat 12 Jul 2014  |  Updated TODO
* b1c6d68  |  Sat 12 Jul 2014  |  Added quotes, scramble logo, jquery
* 20bac1c  |  Sat 12 Jul 2014  |  Removed pip build folder
* f36a498  |  Sat 12 Jul 2014  |  Random quotes working but word net seems
     ↪  bugged
* 54d09e2  |  Sat 12 Jul 2014  |  More style, content and structure
* 462010b  |  Sat 12 Jul 2014  |  p01 changes to style
* 646c38d  |  Fri 11 Jul 2014  |  Style
* bf5f8bf  |  Fri 11 Jul 2014  |  Rotating spiral :)
* 08eff76  |  Thu 10 Jul 2014  |  More style
* bb3f1e5  |  Thu 10 Jul 2014  |  Changed style and venv
* 4190994  |  Wed 09 Jul 2014  |  P01 working. Added style.
* d427ff1  |  Wed 09 Jul 2014  |  Deleted files
*   6251eae  |  Wed 09 Jul 2014  |  Merge branch 'master' of
     ↪  https://github.com/Fania/newpata
|\
* | c9eb78b  |  Wed 09 Jul 2014  |  Reinstalled venv
| * fb54978  |  Wed 09 Jul 2014  |  proto01 sort of working
| * 6547cc5  |  Wed 09 Jul 2014  |  get and post working ok
|/
*   81a3eec  |  Tue 08 Jul 2014  |  Merge branch 'master' of
     ↪  https://github.com/Fania/newpata
|\
* | 5801f70  |  Tue 08 Jul 2014  |  Few updates
| * bf40b91  |  Tue 08 Jul 2014  |  pip installed nltk, some fixes
| * 3ced858  |  Tue 08 Jul 2014  |  Changed files around. Multiple pages now
     ↪  working.
| * 34da770  |  Tue 08 Jul 2014  |  Added venv for windows
|/
* 809ac8c  |  Mon 07 Jul 2014  |  Cleanup and readme changes
* 3f06260  |  Mon 07 Jul 2014  |  Edited readme again
* c721b33  |  Mon 07 Jul 2014  |  Edited readme
* ffbdb4b  |  Mon 07 Jul 2014  |  Edited readme
* 8870b3d  |  Mon 07 Jul 2014  |  Added gitignore file
```

```
* ba1a9c2  |  Mon 07 Jul 2014  |  Second commit
* 244c4b3  |  Mon 07 Jul 2014  |  First commit
* 4789ead  |  Sun 06 Jul 2014  |  Before merge with 2.0
* c4d8ef2  |  Wed 15 Aug 2012  |  Removed old files, added new stuff
* 8cdd5ae  |  Wed 15 Aug 2012  |  Before pata 2
* cb2afbb  |  Wed 15 Aug 2012  (origin/master, origin/HEAD) |  small changes
* d78b916  |  Fri 15 Jun 2012  |  gitignore
* 6ff5630  |  Tue 14 Aug 2012  |  readme updates
* c364398  |  Tue 14 Aug 2012  (tag: 1.5) |  version 1.5 stuff
*   a541df9  |  Thu 28 Jun 2012  |  fixed conflict
|\
| * 4cece9a  |  Thu 28 Jun 2012  (version15branch) |  moved files again
| * df252a4  |  Mon 16 May 2016  (Version_1.5/master) |  Bing not working yet
| * 6e5b8ac  |  Mon 13 Aug 2012  |  Linked to prototype01's syzygy function.
| * 8188da6  |  Mon 13 Aug 2012  |  Fixed sizes in spiral for images.
| * de742c0  |  Mon 13 Aug 2012  |  Spirals for vid and img working.
| * 1095af7  |  Mon 13 Aug 2012  |  Bing images now display as 150x150 properly
| * fcbf1a2  |  Thu 09 Aug 2012  |  Minor changes to prototype 02
| * 803a364  |  Thu 09 Aug 2012  |  git ignore fix
| * 32ff91f  |  Thu 09 Aug 2012  |  git ignore fix
| * aa1959e  |  Tue 07 Aug 2012  |  git ignore added
| * 6176e9a  |  Tue 07 Aug 2012  |  none
| * f39323a  |  Tue 07 Aug 2012  |  gitattributes change
| * 35dac54  |  Mon 06 Aug 2012  |  Revert "?"
| * 386df83  |  Mon 06 Aug 2012  |  ?
| * d3adb11  |  Fri 03 Aug 2012  |  Added Bing image search to the flickr
      ↪  search
| * 2258b53  |  Fri 27 Jul 2012  |  Small fixes to how everything works.
| * 87738e1  |  Fri 27 Jul 2012  |  Added Translator functionality.
| * aff96b1  |  Thu 26 Jul 2012  |  Added Youtube search functionality
| * fd68ace  |  Thu 26 Jul 2012  |  Added prototype02 flickr search
      ↪  functionality
| * 4d827c9  |  Thu 05 Jul 2012  |  Added structure to switch between
      ↪  prototypes
* | 637d9f8  |  Thu 28 Jun 2012  |  edited readme
* | 01b7bfa  |  Thu 28 Jun 2012  (tag: 1.0) |  version 1 stuff
* |   7b2cf7f  |  Wed 27 Jun 2012  |  fix conflicts
|\ \
| * | 53bb1e5  |  Mon 16 May 2016  (Version_1/master, version1branch) |
      ↪  Version 1 working again
| * | 47b00d9  |  Sun 15 May 2016  |  Upgrade 2016
| |/
| * 527e7e1  |  Wed 27 Jun 2012  |  Normalize line endings
| * ee01b37  |  Wed 27 Jun 2012  |  Mini changes to readme files.
| * 3c242a0  |  Wed 27 Jun 2012  |  Removed some spaces.
| * 041f8ad  |  Thu 10 May 2012  |  Fixed plural s
| * 5a58ee7  |  Thu 10 May 2012  |  Changes to links and button, all works on
      ↪  single words.
| * 97d5e25  |  Thu 10 May 2012  |  Added icon
| * 84c59c2  |  Thu 10 May 2012  |  Minor changes to button and wordings
```

```
| * 9177388  |  Wed 09 May 2012  |  Making results keywords links
| * 3dc5019  |  Wed 09 May 2012  |  Changed stopwords handling
| * 81b5183  |  Wed 09 May 2012  |  Test
| * 63588d1  |  Wed 09 May 2012  |  First commit of working prototype v0.1
| * c52f846  |  Wed 09 May 2012  |  test
| * 4a04baa  |  Wed 09 May 2012  |  Initial commit of Django skeleton.
* 65bade4  |  Sun 06 May 2012  |  readme update
* 8365be0  |  Sun 06 May 2012  (tag: 0.5) |  New files for patatype
* e987fb6  |  Sat 05 May 2012  |  Added reqs
* 775e0dc  |  Sat 05 May 2012  |  Fixing venv
*   4854a02  |  Mon 16 May 2016  |  Gitignore
|\
| * d45d5b5  |  Sun 15 May 2016  |  Initial commit
* 83991a6  |  Sat 05 May 2012  (tag: 0.2) |  Added surfer 2
* c55bb54  |  Sat 05 May 2012  (tag: 0.1) |  test message
* 40716ba  |  Sat 05 May 2012  |  First few files
* 701d5ba  |  Sat 05 May 2012  |  Initial commit
```

D.2 Thesis Repository

```
* 25a4ecf  |  Sat 05 Nov 2016  (HEAD -> master, origin/master, origin/HEAD) |
      ↪ Analysis done
* 723aaf3  |  Sat 05 Nov 2016  |  Analysis wip
* 7516ac9  |  Thu 03 Nov 2016  |  that difficult analysis section wip
* ad117f8  |  Thu 03 Nov 2016  |  Application chapter done
* 43d1356  |  Wed 02 Nov 2016  |  Spelling, Application stuff
* d09f664  |  Wed 02 Nov 2016  |  Implementation done!
* 08af130  |  Tue 01 Nov 2016  |  Impl img and vid and design done
* f3230d4  |  Mon 31 Oct 2016  |  Impl Img Vid wip
* a78b091  |  Mon 31 Oct 2016  |  Implementation clinamen section
* e87b804  |  Sun 30 Oct 2016  |  Implementation up to Text section done
* 2ea5ade  |  Sat 29 Oct 2016  |  Implementation stuff wip
* 5e67379  |  Thu 27 Oct 2016  |  Implementation formatting
* 36a561b  |  Thu 27 Oct 2016  |  Sort of finished interpretation chapter
* 4f859b5  |  Wed 26 Oct 2016  |  Interpretation first half
* 3a5eaef  |  Wed 26 Oct 2016  |  Foundations done
* a1a4856  |  Wed 26 Oct 2016  |  Foundations restructure
* 8125d14  |  Tue 25 Oct 2016  |  Eval stuff almost done
* 3635502  |  Tue 25 Oct 2016  |  Evaluation first pass done
* 750f67b  |  Tue 25 Oct 2016  |  Fixed mmce tikz
* e2a8bcf  |  Mon 24 Oct 2016  |  MMCE tikz fuck yeah
* 577f4f3  |  Sun 23 Oct 2016  |  Finished Technology
* a312c62  |  Sun 23 Oct 2016  |  wip
* b57c5b2  |  Thu 20 Oct 2016  |  tech and eval wip
* 20c07bc  |  Wed 19 Oct 2016  |  tech nlp stuff
* b7d8f89  |  Tue 18 Oct 2016  |  NLP restructure and regex section
* 3183c78  |  Tue 18 Oct 2016  |  started NLP
* 0ca80ab  |  Tue 18 Oct 2016  |  IR section 99% done
* 92de3d5  |  Mon 17 Oct 2016  |  tech vector model wip
* dd09b7a  |  Sun 16 Oct 2016  |  Tech TF-IDF table and stuff
```

```
*   bb1500c  |  Sun 16 Oct 2016  |  Merge branch 'master' of
      ↪  github.com:Fania/Thesis
|\
* | 7a87ae6  |  Sun 16 Oct 2016  |  Merge problem?
| * 6e788be  |  Sun 16 Oct 2016  |  Technoprogress
|/
* bf59bd8  |  Sat 15 Oct 2016  |  Added all refs for corpus
*   ff100b9  |  Sat 15 Oct 2016  |  Merge branch 'master' of
      ↪  github.com:Fania/Thesis
|\
* \   cc80c3f  |  Sat 15 Oct 2016  |  Merge commit
|\ \
| | * 93a3b5d  |  Fri 14 Oct 2016  |  removed test
| |/
| * df04179  |  Fri 14 Oct 2016  |  Test commit with ssh
| * e2f0700  |  Fri 14 Oct 2016  |  Work in progress
| * 71c4e27  |  Thu 13 Oct 2016  |  Added email from museepata
| * 100badf  |  Wed 12 Oct 2016  |  Tikz diagrams, creativity and technology
| * 25989c5  |  Tue 11 Oct 2016  |  creativity almost done
| * dbdebe1  |  Tue 11 Oct 2016  |  added here
| * d4cb222  |  Tue 11 Oct 2016  |  Creativity chapter progress
| * 7ccb304  |  Mon 10 Oct 2016  |  fixed refs in four c
| * 91cda6b  |  Mon 10 Oct 2016  |  Creativity up to four c's plus figure
| * 81c989c  |  Mon 10 Oct 2016  |  Fixed sourceatright font and size
| * 43e43b3  |  Sun 09 Oct 2016  |  Pataphysics chapter complete
| * 606f640  |  Sun 09 Oct 2016  |  Oulipo table
| * e844bcf  |  Thu 06 Oct 2016  |  pataphysics edit
| * 6311124  |  Thu 06 Oct 2016  |  Pataphysics polish
* | a47a79a  |  Sun 02 Oct 2016  |  mini change
|/
* 30cb89f  |  Tue 13 Sep 2016  |  Mini changes
* 002a993  |  Thu 25 Aug 2016  |  ai chapter
* 672f4dd  |  Thu 25 Aug 2016  |  stuff
* d83d382  |  Tue 23 Aug 2016  |  turing notes
* ebd6051  |  Sat 20 Aug 2016  |  Added refs
* b898d08  |  Wed 17 Aug 2016  |  creat int arguments
* 0c2e560  |  Tue 16 Aug 2016  |  Moved AI section to test temporarily
* e600de0  |  Mon 15 Aug 2016  |  searle
* 0cbac15  |  Mon 15 Aug 2016  |  analysis ai stuff
* cb917a9  |  Mon 15 Aug 2016  |  Moved formalisation to implementation
* 9568a23  |  Sun 14 Aug 2016  |  Formalisation stuff, clinamen
* f936370  |  Fri 12 Aug 2016  |  Stuff
* 29a274d  |  Fri 12 Aug 2016  |  Changed nums
* b03c24b  |  Fri 12 Aug 2016  |  Moved all table captions
*   8b533f6  |  Fri 12 Aug 2016  |  Merge branch 'master' of
      ↪  https://github.com/Fania/Thesis
|\
* | f055732  |  Fri 12 Aug 2016  |  Full compile
| * 568ff22  |  Fri 12 Aug 2016  |  Added meronym prints
|/
```

```
* 4c0db9d  |  Thu 11 Aug 2016  |  mini changes
* df9bd05  |  Sun 07 Aug 2016  |  work on api stuff
* 5cdb101  |  Sat 06 Aug 2016  |  Acronyms french fixed
* 959288f  |  Fri 05 Aug 2016  |  link to toc works
* e105d81  |  Fri 05 Aug 2016  |  floats fixed to [!htbp]
* f7b3466  |  Fri 05 Aug 2016  |  style fixes
* b28f27e  |  Thu 04 Aug 2016  |  analysis stuff end of day
* f9c6732  |  Tue 02 Aug 2016  |  Sentences work
* 445e164  |  Sun 31 Jul 2016  |  analysis stuff
*   7adbd90  |  Sun 31 Jul 2016  |  merge commit
|\
* | 94d3560  |  Sun 31 Jul 2016  |  laptop commit
| * 4f3de4c  |  Sun 31 Jul 2016  |  commit from PC
|/
*   6068b75  |  Fri 29 Jul 2016  |  merge commits
|\
* | bdb45fa  |  Fri 29 Jul 2016  |  throwaway
| * 1c794e7  |  Fri 29 Jul 2016  |  update
| * 546c28e  |  Thu 28 Jul 2016  |  added Andrews feedback
| * a642999  |  Wed 27 Jul 2016  |  full compile
| * fbccdf7  |  Tue 26 Jul 2016  |  end of day commit - numbers
| *   363a483  |  Tue 26 Jul 2016  |  fixes
| |\
| |/
|/|
| * f6ee5eb  |  Tue 26 Jul 2016  |  Fixed partial TOCs spacing problem
* | 7c2073e  |  Tue 26 Jul 2016  |  test
|/
* 3acf223  |  Tue 26 Jul 2016  |  analysis work
* 4f04c38  |  Mon 25 Jul 2016  |  small fix
* baaa22b  |  Mon 25 Jul 2016  |  introduction and inspiration first draft
* 593d887  |  Mon 25 Jul 2016  |  progress
* e571c15  |  Mon 25 Jul 2016  |  Poetry formatting
* 236b88a  |  Sun 24 Jul 2016  |  forest stuff
* 484fae7  |  Sat 23 Jul 2016  |  table anayslis stuff
* 624d77d  |  Fri 22 Jul 2016  |  tabu width sorted
* 4917a5c  |  Tue 19 Jul 2016  |  Restructure of analysis
* 6435ea3  |  Tue 19 Jul 2016  |  work in progress
* 44effab  |  Sun 10 Jul 2016  |  end of day 3
* 62f6a02  |  Sun 10 Jul 2016  |  10000 day 3
* 5d2deae  |  Sun 10 Jul 2016  |  interim commit day 3
* f0cc780  |  Sat 09 Jul 2016  |  end of day 2
* cca6bfe  |  Fri 08 Jul 2016  |  end of day one
* d7869b7  |  Fri 08 Jul 2016  |  full compile
* accfe9d  |  Fri 08 Jul 2016  |  Ready for Boot Camp
* 0dd4448  |  Sat 02 Jul 2016  |  stuff
* 3c11939  |  Thu 30 Jun 2016  |  Progress in methodology
* cd6809a  |  Wed 29 Jun 2016  |  Start on Methodology chapter
* 41811c3  |  Tue 21 Jun 2016  |  Planning
* 9488f72  |  Tue 07 Jun 2016  |  Added interludes and toc design
```

```
* fca28e9  |  Mon 23 May 2016  |  typos, corpus
* edba811  |  Sat 21 May 2016  |  Impl folder structure
* 134f154  |  Fri 20 May 2016  |  started restructure of implementation chapter
* e12cd57  |  Tue 17 May 2016  |  Added gource image
* 93f2726  |  Thu 05 May 2016  |  Minor fixes
* 5ef8cc7  |  Thu 05 May 2016  |  emph change
* edfc6f9  |  Wed 04 May 2016  |  Acronyms sorted
*   293a05f  |  Wed 04 May 2016  |  Merge branch 'master' of
      ↪ https://github.com/Fania/Thesis
|\
* | 8fe7a6b  |  Wed 04 May 2016  |  Typos
| * c92d9ec  |  Wed 04 May 2016  |  small update
| * be93e01  |  Wed 04 May 2016  |  Added Latex Error Link
|/
* 192e239  |  Wed 04 May 2016  |  Typos
* b2383b5  |  Tue 03 May 2016  |  quotations are now british
* 33d7866  |  Mon 02 May 2016  |  gitignore update
* 8123ebc  |  Sun 01 May 2016  |  VSCode stuff
* 2b65561  |  Sun 01 May 2016  |  VS Code test
* bacdc47  |  Wed 27 Apr 2016  |  Programming Culture stuff
* d72fdcd  |  Tue 26 Apr 2016  |  Structure of final chapters
* cd35a52  |  Mon 25 Apr 2016  |  Fixed toc numwidth problem
* 8219928  |  Mon 25 Apr 2016  |  fixed part design
* 249091f  |  Thu 21 Apr 2016  |  Interpretation edits
* 8a070bd  |  Fri 08 Apr 2016  |  Interpretation stuff
* b171625  |  Thu 07 Apr 2016  |  More work on Interpretation
* 29b318a  |  Mon 04 Apr 2016  |  Structured Interpretation chapter
* c57b825  |  Sun 03 Apr 2016  |  Add zombies to chapter
* 6231f8a  |  Sun 03 Apr 2016  |  Added publications
* b15063d  |  Tue 16 Feb 2016  |  Test compilation for Surface
* 9947594  |  Wed 03 Feb 2016  |  Added new publication
* 0963f28  |  Sat 09 Jan 2016  |  Draft 01 for Sophy
* 41481f9  |  Wed 06 Jan 2016  |  Work in Progress
* 57c367b  |  Tue 05 Jan 2016  |  Todo fixing progress
* f6f974f  |  Mon 04 Jan 2016  |  Restructuring
* db7ca77  |  Sat 02 Jan 2016  |  More interpretation changes
* daa7560  |  Sat 02 Jan 2016  |  Interpretation progress
* 6562882  |  Wed 30 Dec 2015  |  evaluation draft
* 214e654  |  Tue 29 Dec 2015  |  Stuff, marginnote, TOC, Foundations
* 366f15d  |  Mon 28 Dec 2015  |  Foundations progress
* d792b19  |  Tue 22 Dec 2015  |  Foundations progress
* 407e1b5  |  Tue 22 Dec 2015  |  Draft methodology chapter
* 6b38154  |  Fri 11 Dec 2015  |  spirals and poems added
* 8a01e8f  |  Thu 10 Dec 2015  |  intro, insp design
* f5bc906  |  Thu 10 Dec 2015  |  Intro and Inspir almost done.
* e57d672  |  Tue 08 Dec 2015  |  Added inspirations chapter.
* ecad3b9  |  Tue 01 Dec 2015  |  Spiral file seperate
* d71c213  |  Thu 26 Nov 2015  |  Wordcount works
* e0d9975  |  Wed 25 Nov 2015  |  backup
* a7ae537  |  Tue 17 Nov 2015  |  Added screenshots in responsive design
```

```
* 47f724a | Tue 10 Nov 2015 | stuff
* afce56f | Fri 06 Nov 2015 | methodology stuff
* e462c69 | Tue 03 Nov 2015 | work in progress
* e16085b | Sun 01 Nov 2015 | Updated publications appendix
* 5ca0ee5 | Mon 05 Oct 2015 | work in progress
* 2887469 | Wed 30 Sep 2015 | Partial ToC subsections
* da0f3f9 | Wed 30 Sep 2015 | Fucking vertical space in ptoc works!
* 0eeb9ca | Tue 29 Sep 2015 | Chapter contents ok
* c274fd0 | Tue 29 Sep 2015 | ToC per chapter
* 3c23f12 | Thu 24 Sep 2015 | fixed stuff, written stuff
* 2b2e57c | Wed 23 Sep 2015 | fixed some issues and added content
* 176392f | Tue 22 Sep 2015 | todonotes are back
* 2fafb08 | Sun 20 Sep 2015 | Migrated to memoir
* 3566974 | Sun 13 Sep 2015 | Work in progress
* 37e02cb | Sat 12 Sep 2015 | Described code, minted fix
* 86d0ffc | Fri 11 Sep 2015 | Work in progress
* f3999f4 | Mon 07 Sep 2015 | Changed over to minted
* 6004ebd | Sat 05 Sep 2015 | Fixed syntax highlight + wrote a lot
* bf74cca | Thu 03 Sep 2015 | Fixed another error
* 8e03c17 | Thu 03 Sep 2015 | Fixed some errors
* 282be4b | Thu 03 Sep 2015 | Work in progress
* 78920e7 | Thu 27 Aug 2015 | Work in progress
* aeaeaff | Wed 26 Aug 2015 | Small fixes
* 571d396 | Wed 26 Aug 2015 | Chktex linter pass
* 5c1c21e | Fri 31 Jul 2015 | Todo
* d9f1482 | Tue 14 Jul 2015 | Tech, TDM example
* 2778fed | Sat 11 Jul 2015 | Few things, I did.
* ba2187a | Thu 09 Jul 2015 | IR TF IDF stuff
* c1ed5af | Wed 08 Jul 2015 | Things
* b64eb1f | Tue 07 Jul 2015 | Added images
* 778aa4a | Fri 26 Jun 2015 | Stuff
* 25b4e09 | Thu 25 Jun 2015 | Integrating printouts into text
* ef86a5b | Wed 24 Jun 2015 | Printouts
* 8f17c8e | Wed 24 Jun 2015 | Added to read
* d9dc6b9 | Sun 21 Jun 2015 | Tech, Description
* e791530 | Fri 19 Jun 2015 | Tech stuff figures
* f373561 | Wed 17 Jun 2015 | Technology, exploratory search
* eb3b368 | Wed 17 Jun 2015 | Inkscape graphics
* aee67b1 | Mon 15 Jun 2015 | Tech restructure
* 1ed41ee | Mon 15 Jun 2015 | Changed import to include
* 5ab2351 | Sun 14 Jun 2015 | Testing word count
* 09b8e9d | Sun 14 Jun 2015 | Aaaand more structure and bib
* a7a7ff2 | Fri 12 Jun 2015 | Even More Structure
* e89688b | Wed 10 Jun 2015 | Re-Structure
* 7d73b68 | Wed 10 Jun 2015 | Restructuring
* 2b1b713 | Wed 10 Jun 2015 | Work on creativity
* adfb005 | Mon 08 Jun 2015 | Working on creativity chapter
* d7f1cf7 | Sun 07 Jun 2015 | Readme
* fc1e228 | Sat 06 Jun 2015 | Creativity chapter
* ac3ff1c | Sat 06 Jun 2015 | more changes
```

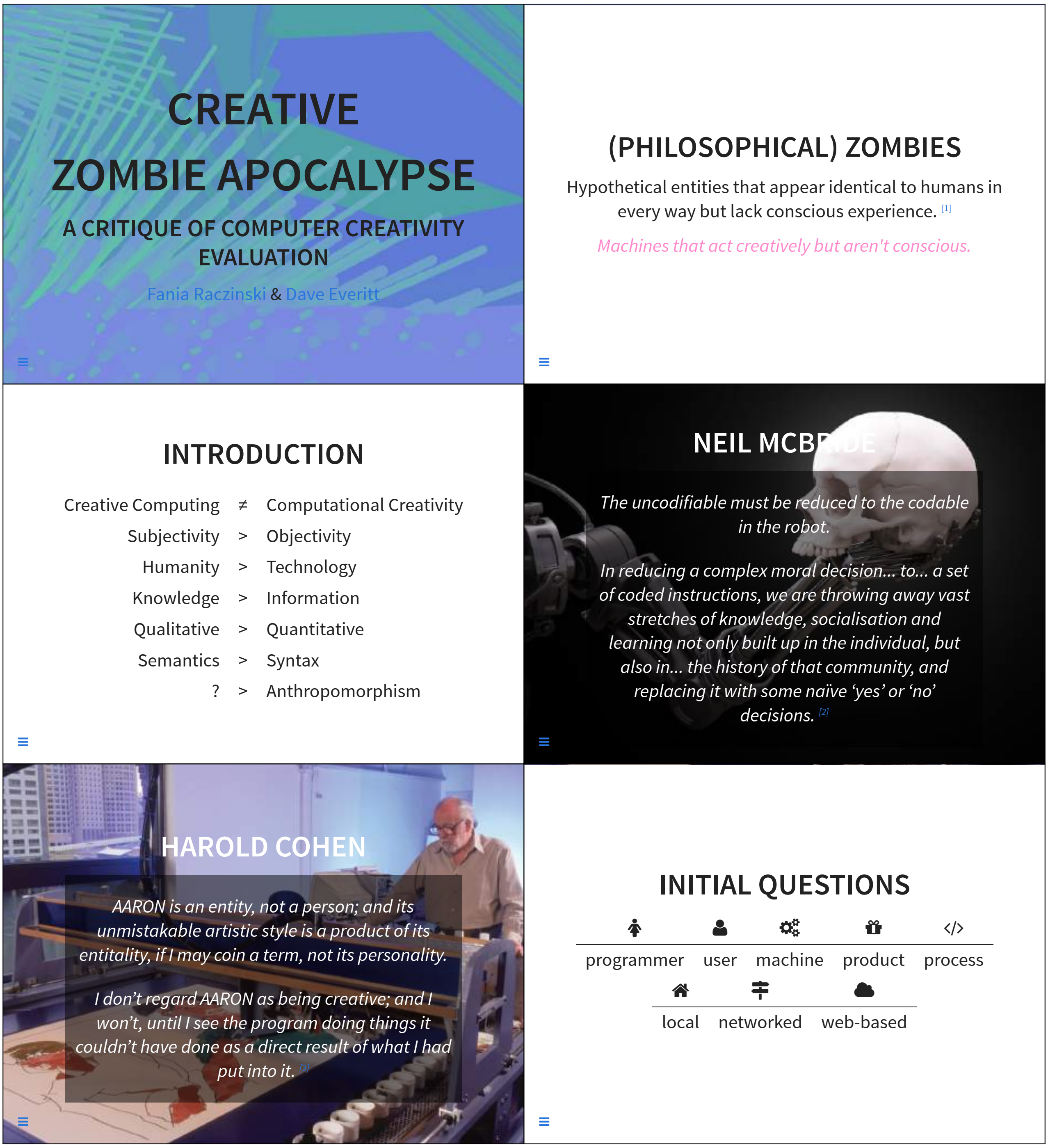
```
* b5af173  |  Wed 03 Jun 2015  |  Fonts test
* c8a7d76  |  Tue 26 May 2015  |  More edits, writing, styling.
* 077d8e1  |  Mon 25 May 2015  |  Writing stuff START
* 4152a3e  |  Sat 23 May 2015  |  conflict sorted?
* 50cbb3a  |  Fri 22 May 2015  |  Update
* 0e8a0f7  |  Wed 13 May 2015  |  Toc changes
* bb3c091  |  Mon 20 Apr 2015  |  show wireframe
* 1e082fa  |  Sat 18 Apr 2015  |  Update readme.md
* f5b7c8e  |  Sat 18 Apr 2015  |  Update readme.md
* 3129135  |  Sat 18 Apr 2015  |  Added todo, updated error section in readme
* b15d2ae  |  Sat 18 Apr 2015  |  Added todos
* 0d8235f  |  Sat 18 Apr 2015  |  Style and error fixes
* 73bece1  |  Fri 17 Apr 2015  |  Update readme.md testing task lists
* ceaba92  |  Fri 17 Apr 2015  |  Update readme.md
* b8451fc  |  Fri 17 Apr 2015  |  Added todo
* 188747f  |  Fri 17 Apr 2015  |  Spacing, Equations, Margins
* 60216ba  |  Wed 15 Apr 2015  |  marginpar
* d4640b4  |  Wed 15 Apr 2015  |  image
* dfc9b45  |  Wed 15 Apr 2015  |  image sizes and names
* 5c2f753  |  Wed 15 Apr 2015  |  readme changes
* a1a1e31  |  Tue 14 Apr 2015  |  Added read
* 0dc0f77  |  Mon 13 Apr 2015  |  Initial commit
```

Publications

E

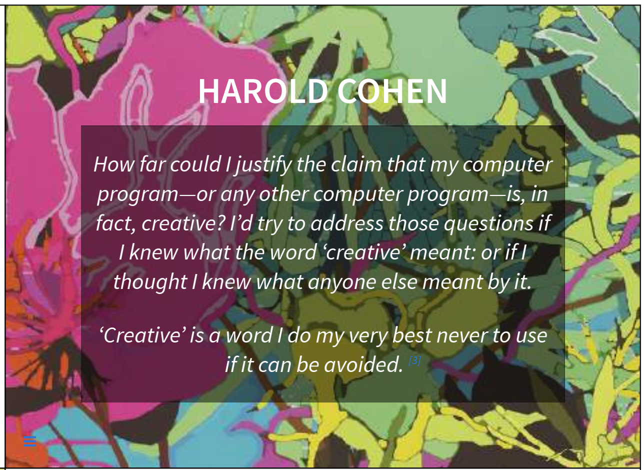
This chapter includes copies of the publications and talks (where available) related to this thesis.

1. Presentation for ISCC'16 (March/April 2016)—page 291.
2. Conference paper "Creative Zombie Apocalypse: A Critique of Computer Creativity Evaluation" for 2^{nd} IEEE International Symposium of Creative Computing (2016)—page 296.
3. Presentation for a CAS IOCT talk at the Phoenix in Leicester, UK (14 Oct 2015)—page 303.
4. Journal article "The pataphysics of creativity: developing a tool for creative search" for Digital Creativity, Volume 24, Issue 3 (2013)—page 305.
5. Presentation for CC'13 (20 June 2013)—page 321.
6. Conference paper "Creative Search Using Pataphysics" for 9^{th} ACM conference on Creativity and Cognition (2013)—page 325.
7. Conference paper "A Framework for Creativity in Search Results" for 3^{rd} International Conference on Creative Content Technologies (2011)—page 332.


CREATIVE
ZOMBIE APOCALYPSE
A CRITIQUE OF COMPUTER CREATIVITY EVALUATION
Fania Raczinski & Dave Everitt
(PHILOSOPHICAL) ZOMBIES
Hypothetical entities that appear identical to humans in every way but lack conscious experience. [1]
Machines that act creatively but aren't conscious.
INTRODUCTION
Creative Computing ≠ Computational Creativity
Subjectivity > Objectivity
Humanity > Technology
Knowledge > Information
Qualitative > Quantitative
Semantics > Syntax
? > Anthropomorphism
NEIL MCBRIDE
The uncodifiable must be reduced to the codable in the robot.
In reducing a complex moral decision... to... a set of coded instructions, we are throwing away vast stretches of knowledge, socialisation and learning not only built up in the individual, but also in... the history of that community, and replacing it with some naïve 'yes' or 'no' decisions. [2]
HAROLD COHEN
AARON is an entity, not a person; and its unmistakable artistic style is a product of its entitality, if I may coin a term, not its personality.
I don't regard AARON as being creative; and I won't, until I see the program doing things it couldn't have done as a direct result of what I had put into it. [3]
INITIAL QUESTIONS
programmer
user
machine
product
process
local
networked
web-based

MODELS

- Output minus Input [4]
- Measuring against specific criteria [4,5,6,7,8,9,10]
- NOT product minus process [4]
- Creative Tripod [11,12]
- Ontology of creativity [10,13]
- SPECS [14]
- MMCE [15]



5 P CRITERIA

Product | Process | Purpose | Person | Place

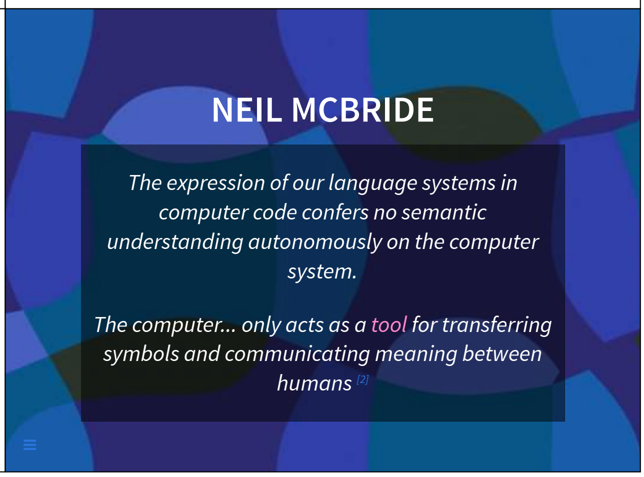
7 KEYWORD GROUPS

Novelty | Value | Quality | Purpose

Spatial | Temporal | Ephemeral

DETAILS

Novelty:
originality, newness, variety, typicality, imagination, archetype, surprise
Value:
usefulness, appropriateness, appreciation, relevance, impact, influence
Quality:
skill, efficiency, competence, intellect, acceptability, complexity
Purpose:
intention, communication, evaluation, aim, independence
Spatial:
context, environment, press
Temporal:
persistence, results, development, progression, spontaneity
Ephemeral:
serendipity, randomness, uncertainty, experimentation, emotional response



INTUITIVE EVALUATION

Creativity could be said to be more likely to *emerge* from activities that stimulate, enable or *constrain* these properties.

We reject a check-box approach and suggest *scales* for a more intuitive evaluation.

These represent emergence better than checklists...

SCALES

Keyword	Scale		
Novelty	Established	⟷	Novel
Value	Playful	⟷	Purposive
Quality	Minimal	⟷	Complex
Purpose	Emotive	⟷	Thoughtful
Spatial	Universal	⟷	Specific
Temporal	Instant	⟷	Persistent
Ephemeral	Accidental	⟷	Experimental

FRAMEWORK

Our distillation of the qualities used to identify creativity can be applied across the identified domain axes to any output, in order to 'measure' the degree of creativity.

PERSON, PLACE, PRODUCT, PROCESS, PURPOSE
Novelty, Value, Quality, Purpose, Spatial, Temporal, Ephemeral

FUTURE

Increase the distillation to create a more condensed and workable set.

Scramble the qualities and axes to avoid unconscious groupings.

Require several individuals to assess each case.

Graph the results in 3D.

Work towards coordinating the research of groups who wish to identify and measure creativity.

CONCLUSION

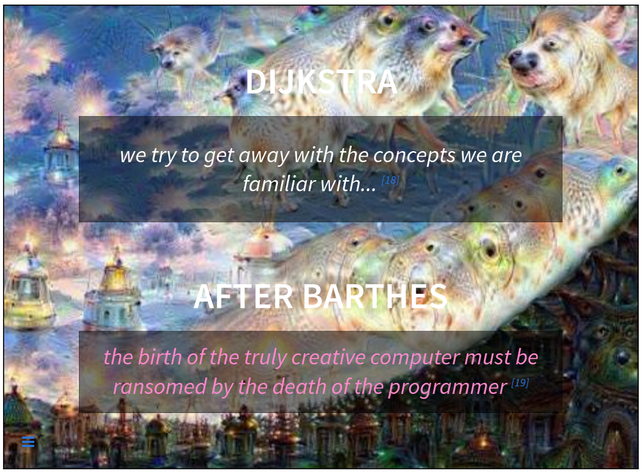
Unless we can prove that computer programs can make conscious - not pre-programmed - choices,

*all apparent computer creativity is the action of an unconscious **zombie** that has the mere appearance of creativity.*

Therefore, do we need to redefine 'creativity' in the field of computing to distinguish it from *human* creativity?

REFERENCES

1. David Chalmers, The Conscious Mind. Oxford University Press, 1996.
2. Neil McBride. “A Robot Ethics: The EPSRC Principles and the Ethical Gap,” in AISB / IACAP World Congress 2012 Framework for Responsible Research and Innovation in AI, no. July, 2012, pp. 10–15.
3. Harold Cohen. (1999) Colouring without seeing: A problem in machine creativity. [Online]. Available: http://www.kurzweilcyberart.com/aaron/hi_essays.html
4. A. Pease, D. Winterstein, and S. Colton, “Evaluating Machine Creativity,” in Proceedings of ICCBR Workshop on Approaches to Creativity, 2001, pp. 129–137
5. A. Pease, S. Colton, R. Ramezani, J. Charnley, and K. Reed, “A Discussion on Serendipity in Creative Systems,” in Proceedings of the 4th International Conference on Computational Creativity, vol. 1000. University of Sydney, 2013, pp. 64–71.
6. G. Ritchie, “Some Empirical Criteria for Attributing Creativity to a Computer Program,” Minds and Machines, vol. 17, no. 1, pp. 67–99, 2007.
7. G. Ritchie, “Assessing creativity,” in AISB ’01 Symposium on Artificial Intelligence and Creativity in Arts and Science. Proceedings of the AISB’01 Symposium on Artificial Intelligence and Creativity in Arts and Science, 2001, pp. 3–11.
8. D. Ventura, “A Reductio Ad Absurdum Experiment in Sufficiency for Evaluating (Computational) Creative Systems,” in 5th International Joint Workshop on Computational Creativty, 2008


DIJKSTRA
we try to get away with the concepts we are familiar with... [18]
AFTER BARTHES
the birth of the truly creative computer must be ransomed by the death of the programmer [19]

2016 IEEE Symposium on Service-Oriented System Engineering

Creative Zombie Apocalypse: A Critique of Computer Creativity Evaluation

Fania Raczinski
Institute of Creative Technologies
De Montfort University
The Gateway
Leicester, LE1 9BH
Email: fania.raczinski@dmu.ac.uk

Dave Everitt
Institute of Creative Technologies
De Montfort University
The Gateway
Leicester, LE1 9BH
Email: dave.everitt@dmu.ac.uk

***Abstract*—Using algorithms to generate creative work is a well-established transdisciplinary practice that spans several fields. Accessible and popular coding tools such as Processing and Open Frameworks, as well as the rise of hack spaces have significantly contributed to increased activity in this field. However, beyond art-technology curation and historical contextualisation, evaluation of the resulting artefacts is in its infancy, although several *general* models of creativity—and its evaluation—exist.**

There is a perceived distinction between human and computer creativity, whereas we argue that they are effectively the same thing. Computers are made and programmed by people, so it makes sense to measure the creativity of the human influence behind the machine, rather than viewing computers as truly autonomous entities.

By concatenating and enhancing existing models of creativity, we propose a framework that takes these issues into account, with a view to evaluating creative work that uses the computer as a medium more effectively.

***Index Terms*—Creativity, creative computing, evaluation**

I. Introduction

Although using computers to generate creative work has its foundations in the 1950s [1], John Maeda's Design By Numbers [2] and from around 2010 a slew of similar initiatives followed Processing's lead. However, due in part to the niche position of artists working with technology, and also because such activity was overlooked or ignored until relatively recently by arts bodies and critics, formal evaluation of the creativity in such work lagged behind.

In this context humans simply use computers as tools for their creativity—no matter how autonomous the machine output may appear, or how far it travels from the original intentions of the programmer, its origins nevertheless reside in the humanly-authored code that produces the output.

This is overlooked in anthropomorphic approaches that regard computers as being capable of creativity in their own right. Computer output cannot be conceptually separated from the craft/skill/intention of the programmer, even when the results are unexpected or accidental. The illusion of creativity can be produced by introducing randomness, serendipity, etc. but this is not the same as the intuitive decision-making that drives human creativity.

Hypothetical "zombies" (popularised by philosopher David Chalmers [3]) are entities that appear identical to humans in every way but lack conscious experience. We now borrow this term and apply it to computers which appear creative but lack real autonomous intent.

Further, creativity and the subjective properties associated with it, lack a universally accepted definition. As a human quality it has definitions that don't necessarily lend themselves to be applied to computers. However, there are several important theories and evaluation frameworks concerning human and computer creativity, and these are the basis for our work. Creativity has been studied at various levels (neurological, cognitive, and holistic/systemic), from different perspectives (subjective and objective) and existing research has identified specific characteristics (combinational, exploratory and transformative). Some aspects, like novelty and value, recur in many models of creativity but some, like relevance and variety, rarely appear; while other terms are problematic when it comes to computing. Computer systems are generally evaluated against functional requirements and performance specifications, but creativity should be seen as a continuum, there is no clear cut-off point or Boolean answer to say precisely when a person or piece of software has become creative or not.

> "The expression of our language systems in computer code confers no semantic understanding autonomously on the computer system. The computer system only acts as a tool for transferring symbols and communicating meaning between humans." [4]

True Artificial Intelligence and *true* Computational Creativity are equally elusive. For a computer to become truly intelligent and therefore creative, it would need to break out of the programming procedures by which it operates. Yet it is bound to follow rules, no matter how emergent the outcome. The paradox is that it needs to recognise its contraints in order to break free from them. Yet programatically defining yet more rules to allow that to happen—even when those rules enable machine learning—is tautological!

II. Emerging Disciplines

Initiatives that aim at a more rigorous understanding of computing and creativity have given rise to several fields, each

978-1-5090-2253-3/16 $31.00 © 2016 IEEE
DOI 10.1109/SOSE.2016.30

having its own terminology and approach, but with significant overlaps.

The two main disciplines directly related to creativity and technology that have emerged in recent years are as follows. "Creative Computing" tries to reconcile the objective precision of computer systems with the subjective ambiguity of human creativity [5] and has an overarching theme of *unite and conquer*, i.e. drawing from a wide range of transdisciplinary knowledge to tackle a problem (as opposed to the principle of *divide and conquer* in computer science, which divides bigger problems down into smaller and easier parts) [6]. The main challenge, Andrew Hugill and Hongji Yang argue, is for technology to become "more adaptive, smarter and better engineered to cope with frequent changes of direction, inconsistencies, irrelevancies, messiness and all the other vagaries that characterise the creative process" [5]. In part, these issues are due to the transdisciplinary nature of Creative Computing; factors such as common semantics, standards, requirements and expectations are typical challenges. Hugill and Yang therefore argue that creative software should be flexible and able to adapt to ever-changing requirements, evaluated and re-written continuously, and it should be cross-compatible.

"Computational Creativity" has emerged from within Artificial Intelligence (AI) research. Simon Colton and Geraint Wiggins argue that AI falls within a problem-solving paradigm: "an intelligent task, that we desire to automate, is formulated as a particular type of problem to be solved", whereas "in Computational Creativity research, we prefer to work within an artefact generation paradigm, where the automation of an intelligent task is seen as an opportunity to produce something of cultural value" [7]. They further explain that it models, simulates, replicates or enhances human creativity using a computer.

III. Existing Theories of Creativity

Richard Mayer identified five big questions of human creativity research and different approaches with their own methodologies and goals [8]; is creativity:

1) a property of people, products, or processes?
2) a personal or social phenomenon?
3) common or rare?
4) domain-general or domain-specific?
5) quantitative or qualitative?

These questions form a nice introduction to the four main theories of creativity which inspired our work.

The *Four P* model by Mel Rhodes [9] identified four elements of creativity: (1) the **person**—personality, intellect, temperament, physique, traits, habits, attitudes, self-concept, value systems, defence mechanisms and behaviour, (2) the **process**—motivation, perception, learning, thinking and communication, (3) the **press**—relationship between human beings and their environment and (4) the **product**—a thought which has been communicated to other people in the form of words, paint, clay, metal, stone, fabric, or other material. Rhodes highlights the importance of a holistic view on creativity through these four areas of study, which he hoped would become the basis of a unified theory of creativity. Ross Mooney independentely identified four aspects of creativity which he called the environment, person, process and product (as cited in [10]).

Margaret Boden defined three types of creativity: (1) **combinational**—making unfamiliar combinations of familiar ideas, juxtaposition of dissimilar, bisociation, deconceptualisation, (2) **exploratory**—exploration of conceptual spaces, noticing new things in old spaces and (3) **transformational**—transformation of space, making new thoughts possible by altering the rules of old conceptual space [11]. Boden also differentiates between two levels of creativity, a personal one and a historical one. Psychological creativity ("P-creativity") is a personal kind of creativity that is novel in respect to an individual, while historical creativity ("H-creativity") is fundamentally novel in respect to the whole of human history.

James Kaufman and Ronald Beghetto defined the *Four C* model of creativity [12]. They are **Big-C**—eminent accomplishments, **Pro-c**—professional expertise, **Little-c**—everyday innovation and **Mini-c**—transformative learning. The concepts of the uppercase C and lowercase c loosely correspond to Boden's H and P creativity, which in turn could be interpreted as objective and subjective creativity.

Henri Poincaré suggested a *Four Stage* model [13] (formulated by Graham Wallas [14]). The four stages are preparation, incubation, illumination and verification. This is reminiscent of George Pólya's description of the *problem solving* process [15]—understand, plan, carry out, look back.

Bipin Indurkhya argued that there are two main cognitive mechanisms of creativity: juxtaposition of the dissimilar, and deconceptualization. He said that we are constrained by associations in the concept networks we inherit and learn in our lifetime, but that computers do not have these conceptual associations and therefore have an advantage when it comes to creative thinking [16].

IV. Existing Evaluation Frameworks

Evaluating human creativity objectively seems problematic; evaluating computer creativity at all seems even harder. There are many debates across the disciplines involved. Taking theories on human creativity and directly applying them to machines seems logical but may be the wrong (anthropomorphic) approach. Adapting Mayer's five big questions [8] to machines does not seem to capture the real issues at play. Instead of asking if creativity is a property of people, products, or processes we might ask if it is a property of any or all of the following:

- programmers (and collaborators)
- users (audiences and participants)
- machines (this is problematic until the posited AI singularity [17])
- products (i.e. does a program output material that can be judged to be creative)
- processes (e.g. a Processing sketch, or in a self-modifying/learning program)

For instance, is the programmer the only creative agent, or are users (i.e. audiences or participants in interactive work) able to modify the system with their own creative input? Similarly for any instance of machine creativity, we might ask if it is:

- local (e.g. limited to a single machine or program?)
- networked (i.e. interacts with other predefined machines)
- web-based (e.g. is distributed and/or open to interactions, perhaps via an API)

For example, discussions from computational creativity often focus on very basic questions such as "whether an idea or artefact is valuable or not, and whether a system is acting creatively or not" [18].

Pease, Winterstein and Colton have argued that creativity may be seen as *output minus input* [19]. The output in this case is the creative product but the input is not the process. Rather, it is the *inspiring set* (comprised of explicit knowledge such as a database of information and implicit knowledge input by a programmer) of a piece of software. Simon Colton specifies that "the degree of creativity in a program is partly determined by the number of novel items of value it produces. Therefore we are interested in the set of valuable items produced by the program which exclude those in the inspiring set." [20]. Alison Pease et al. also suggest that all creative products must be **novel and valuable** [19] and provide several measures that take into consideration the context, complexity, archetype, surprise, perceived novelty, emotional response and aim of a product, although the measurement of these qualities isn't explicitely described. In terms of the creative process itself they only discuss **randomness** as a measurable approach. Elsewhere, Pease et al. discuss using **serendipity** as an approach [21].

Graeme Ritchie supports the view that creativity in a computer system must be measured "relative to its initial state of knowledge" [22]. He identifies three main criteria for creativity as **novelty, quality and typicality**, although he argues that "novelty and typicality may well be related, since high novelty may raise questions about, or suggest a low value for, typicality" [22], [23]. He proposes several evaluation criteria which fall under the following categories: basic success, unrestrained quality, conventional skill, unconventional skill, avoiding replication and various combinations of those [22]. Dan Ventura later suggested the addition of **variety and efficiency** to Ritchie's model [24].

It should be noted that *output minus input* might easily be misinterpreted as "product minus process", however, that is not the case. In fact, Pease, Winterstein and Colton argue that "the process by which an item has been generated and evaluated is intuitively relevant to attributions of creativity", and that "two kinds of evaluation are relevant; the evaluation of the item, and evaluation of the processes used to generate it" [19]. If a machine simply copies an idea from its inspiring set then it just cannot be considered creative and needs to be disqualified, so to speak.

Simon Colton came up with an evaluation framework called the *creative tripod* [25], [26]. The tripod consists of three behaviours a system or artefact should exhibit in order to be called creative. The three legs represent **skill, appreciation, and imagination** and three different entities can sit on it, namely the programmer, the computer and the consumer. Colton argues that if "the software has been skillful, appreciative and imaginative, then, regardless of the behaviour of the consumer or programmer, the software should be considered creative". As such a product can be considered creative if it *appears* to be creative. If all three behaviours are not exhibited, however, it should not be considered creative.

> "Imagine an artist missing one of skill, appreciation or imagination. Without skill, they would never produce anything. Without appreciation, they would produce things which looked awful. Without imagination, everything they produced would look the same." [25]

Davide Piffer suggests that there are three dimensions of human creativity that can be measured, namely **novelty, usefulness/appropriateness and impact/influence** [27]. As an example of how this applies to measuring a person's creativity he proposes "citation counts". While this idea perhaps works well for measuring scientific impact, it seems questionable whether popularity or social status can be a valid measure of creative quality.

Anna Jordanous proposed 14 key components of creativity (which she calls an "ontology of creativity") [28], from a linguistic analysis of creativity literature which identified words that appeared significantly more often in discussions of creativity compared to unrelated topics. These are active involvement and persistence, generation of results, uncertainty, domain competence, general intellect, independence and freedom, intention and emotional involvement, originality, progression and development, social interaction and communication, spontaneity/subconscious processing, thinking and evaluation, value, variety, divergence and experimentation. Jordanous also argued that "evaluation of computational creativity is not being performed in a systematic or standard way" [29]; an issue which further confuses the problem of objective evaluation. To remedy this she proposes a "Standardised Procedure for Evaluating Creative Systems" (SPECS) [30]:

1) Identify a definition of creativity that your system should satisfy to be considered creative.
2) Using Step 1, clearly state what standards you use to evaluate the creativity of your system.
3) Test your creative system against the standards stated in Step 2 and report the results.

The SPECS model essentially means that we cannot evaluate a creative computer system objectively, unless steps 1 and 2 are predefined and publically available for external assessors to execute step 3. Creative evaluation can therefore be seen as a move from subjectivity to objectivity, i.e. defining subjective criteria for objectively evaluating a product in terms of the initial criteria.

> "For transparent and repeatable evaluative practice, it is necessary to state clearly what standards are

> used for evaluation, both for appropriate evaluation of a single system and for comparison of multiple systems using common criteria." [30]

This is further strengthened by Richard Mayer stating that we need a "clearer definition of creativity" [8] and Linda Candy arguing for "criteria and measures [for evaluation] that are situated and domain specific." [31]

Candy draws inspiration for the evaluation of (interactive) creative computer systems from Human Computer Interaction (HCI) research. The focus of evaluation in HCI has been on usabilty, she says, which may not be as useful in creativity research. She argues that in order to successfully evaluate an artefact, the practitioner needs to have "the necessary information including constraints on the options under consideration" [31]. Evaluation happens at every stage of the process (i.e. from design → implementation → operation). Some of the key aspects of evaluation highlighted by Candy are aesthetic appreciation, audience engagement, informed considerations and reflective practice. She then goes on to introduce the "Multi-dimensional Model of Creativity and Evaluation" (MMCE) [31] with four main elements of people, process, product and context similar to some of the models of creativity we have seen above (e.g. the *Four P* model).

V. Thoughts and Critique

> "The uncodifiable must be reduced to the codable in the robot. In reducing a complex moral decision (tacit, intuitive, deriving knowledge from maturity) to the execution of a set of coded instructions, we are throwing away vast stretches of knowledge, socialisation and learning not only built up in the individual, but also in the community and the history of that community, and replacing it with some naïve 'yes' or 'no' decisions." [4]

Neil McBride's observation is echoed by Indurkhya, who argues that because computers don't make decisions based on personal or cultural concepts (even when these are included in code), they are more likely to make connections that humans will perceive as "creative leaps" [16]. These leaps *appear* creative only because we are athropomorphising not only the output, but in some cases even also the *intent* behind it, as if this originated in the computer itself rather than as an output from algorithmic processes. This phenomenon is most apparent in the "uncanny valley" created by those areas of robotics that seek to create human companions, or where the intent is to imbue the computer with a personality. This is even the case for simple web interfaces, let alone computers that might mimic human creativity:

> "Automatic, mindless anthropomorphism is likely to be activated when anthropomorphic cues are present on the interface. [...] it is noteworthy that anthropomorphic cues do not have to be fancy in order to elicit human-like attributions." [32]

The phenomenon of ascribing human qualities to non-human artefacts and machines depends on the prior associations (concept networks) humans have with certain activities, including creativity. It leads to metaphorical statements such as "this interface is friendly", "a bug snuck into my code" or "the computer is being creative", and appears in media article headlines such as "Patrick Tresset's robots draw faces and doodle when bored" [33], as if there were conscious intent behind the code generating such activity in Tresset's sketching bot *Paul*. This tendency has implications for the aimed-for objectivity when evaluating certain creative computing projects, one the most well-established being Harold Cohen's *AARON*, artist-authored software that produces an endless output of images in his own unique style. While documenting the process of coding his system, Cohen asked:

> "How far could I justify the claim that my computer program—or any other computer program—is, in fact, creative? I'd try to address those questions if I knew what the word 'creative' meant: or if I thought I knew what anyone else meant by it. [...] 'Creative' is a word I do my very best never to use if it can be avoided. [...] AARON is an entity, not a person; and its unmistakable artistic style is a product of its entitality, if I may coin a term, not its personality." [34]

He goes on to outline four elements of *behaviour X* (his placeholder for creativity): (1) **emergence** produced from the complexity of a computer program, (2) **awareness** of what has emerged, (3) **willingness** to act upon the implications of what has emerged, and (4) **knowledge** of the kind manifest in expert systems. He identifies three of these properties as programmable (within limits), but "as to the second element, the program's awareness of properties that emerge, unbidden and unanticipated, from its actions... well, that's a problem." [34], and concludes that "it may be true that the program can be written to act upon anything the programmer wants, but surely that's not the same as the individual human acting upon what he wants himself. Isn't free will of the essence when we're talking about the appearance of behaviour X in people?". In other words, a decision tree in computing is not the same as a human decision-making process. As for whether his life's work is autonomously creative:

> "I don't regard AARON as being creative; and I won't, until I see the program doing things it couldn't have done as a direct result of what I had put into it. That isn't currently possible, and I am unable to offer myself any assurances that it will be possible in the future. On the other hand I don't think I've said anything to indicate definitively that it isn't possible." [34]

In the same manner as in the field of computer ethics, i.e. "the ethics of the robot must be the ethics of the maker" [4], the creative computer must ultimately be a product of the creativity of the programmer. To hijack Barthes' conclusion in "The Death of the Author": *the birth of the truly creative computer must be ransomed by the death of the programmer* [35]—in other words, a truly creative computer must be able to act without human input, yet any computer process presumes

TABLE I
OBJECTIVE CRITERIA OF CREATIVITY

Criteria	Note
Product	Algorithmic sketch, poetry, audio, interactive installation
Process	Procedural, Experimental, Heuristic, Systems-based
Purpose	Accidental, Conceptual, Interactive, Time-based
Person	Skill, Aesthetic values, Influences, Collaborations
Place	Culture, Social environment, Education, Peers

a significant amount of human input in order to produce such so-called autonomous behaviour, so the question is whether that behaviour can ever be regarded as truly autonomous—no matter how independant it appears to be.

Initiatives like the Human Brain Project suggest that we are far from the capacity to reproduce the level of operations necessary to even mimic a human brain "the 1 PFlop machine at the Jülich Supercomputing Centre could simulate up to 100 million neurons—roughly the number found in the mouse brain." [36]. Even if it were possible today to scale this up to the human brain, would the result be an entity capable of truly intelligent creative activity, or would it actually be a *zombie*?

Current evaluation methodologies in creative computing disciplines have concentrated on only a handful of the facets previously discussed, for example studying only the creative end-product itself (out of context), only judging it by its objective novelty, assigning an arbitrary thresholds, etc. This also includes the assumption that machines "mimic" humans and are therefore not judged at their full potential. For example we generally do not take into account the differences between humans and machines or, more precicely, the differences between the human brain and computer processors. In fact, it could be said that we are in danger of limiting computers so that they *appear* more human.

VI. OUR PROPOSED FRAMEWORK

All of the theories of creativity and its evaluation mentioned above have value, but each alone may be incomplete and contain overlaps. There is a misconception that creativity can be measured objectively and quantifiably, but given the issues discussed above, it is unlikely that any system will yield truly accurate measurements in practice, even if such accuracy were possible. As Jürgen Schmidhuber suggests in the quote below, evaluation of creativity always happens from a subjective standpoint, originating in either the individual, or in the enveloping culture of which they are part.

> "Any objective theory of what is good art must take the subjective observer as a parameter." [37]

We therefore propose two facets of a new *fuzzy* approach that aims to obtain a more honest measure of the subjective judgements implied when evaluating creativity:

1) a set of scales that can be used to approximate a "rating" for the creative value of an artefact,
2) a set of criteria to be considered using the scales above.

TABLE II
SUBJECTIVE SCALES FOR CREATIVITY

Keyword	Scale
Novelty	Established ↔ Novel
Value	Playful ↔ Purposive
Quality	Minimal ↔ Complex
Purpose	Emotive ↔ Thoughtful
Spatial	Universal ↔ Specific
Temporal	Instant ↔ Persistent
Ephemeral	Accidental ↔ Experimental

The **criteria** listed in table I should be considered objectively, while the **scales** in table II are judged subjectively. The set of scales is directly derived from the various frameworks for evaluating creativity reviewed in the previous sections. An overview of recurring keywords in existing approaches suggests the following distillation of seven groups:

Novelty originality, newness, variety, typicality, imagination, archetype, surprise
Value usefulness, appropriateness, appreciation, relevance, impact, influence
Quality skill, efficiency, competence, intellect, acceptability, complexity
Purpose intention, communication, evaluation, aim, independence
Spatial context, environment, press
Temporal persistence, results, development, progression, spontaneity
Ephemeral serendipity, randomness, uncertainty, experimentation, emotional response

The "5 P's"—**Product, Process, Purpose, Person, Place**—are all components of any creative artefact (see table I).

This evaluation framework can apply to any kind of creativity, from the traditional arts to digital works to computational creativity. Because the scale element allows for the measurement of subjective qualities, it circumvents binary yes/no or check-box approaches and therefore makes it possible to gather quantitative values from the subjective judgements involved in evaluating creativity in general.

The terms on each end of the scales are suggestions only and should not be taken as value judgements. Rather, they should be adapted for each project individually. Numeric values can be assigned to the scales if needed according to specific evaulative requirements.

A. An example application

Below is an example assessment for a hypothetical piece of art:

PRODUCT:
Established ——————x—— Novel
Playful —————x——— Purposive
Minimal —-x——————- Complex
Emotive —-x——————- Thoughtful

Universal	——————-x—-	Specific
Instant	—————x——	Persistent
Accidental	—————x——-	Experimental

PROCESS:

Established	—x——————	Novel
Playful	—————x———	Purposive
Minimal	———x—————	Complex
Emotive	——-x—————-	Thoughtful
Universal	——————x—-	Specific
Instant	——————x—-	Persistent
Accidental	—————x———	Experimental

PURPOSE:

Established	——x—————	Novel
Playful	—————-x——-	Purposive
Minimal	—x——————	Complex
Emotive	——————x—-	Thoughtful
Universal	——————x—	Specific
Instant	———————-x-	Persistent
Accidental	——————x——	Experimental

PERSON:

Established	—x——————	Novel
Playful	——————x—	Purposive
Minimal	—x——————	Complex
Emotive	——————x—-	Thoughtful
Universal	—-x——————-	Specific
Instant	—————x——	Persistent
Accidental	———x———	Experimental

PLACE:

Established	—x——————	Novel
Playful	———————x–	Purposive
Minimal	———————x–	Complex
Emotive	—————x——	Thoughtful
Universal	——————x—	Specific
Instant	——————x—-	Persistent
Accidental	–x———————	Experimental

Ideally, these scales would need to be applied by several people during the evaluation process, generating an intuitive assessment of the various values (e.g. Playful—Purposive) for each of the criteria (e.g. Product).

VII. Conclusion

Creativity is a transdisciplinary activity and is apparent in many diverse fields, yet it is often studied from within a single discipline within which other perspectives and theories can be overlooked. Therefore, creative evaluation is subjective, and involves an emotional component related to the satisfaction of a set of judgements. These judgements are mutable when subjected to personal, social and cultural influence, so we can only try to evaluate a creative activity objectively via approximisations.

True AI and Computational Creativity are equally elusive. Just as the Turing Test [38] is flawed (because it is designed to fool humans into thinking a machine is a person, but only through mimickry), the view that something *is* creative because it *appears* creative is similarly flawed. This is the premise behind by John Searle's Chinese Room Argument [39] where an individual with a map of English to Chinese symbols can appear to someone outside the room to "know" Chinese. By inference, just because a computer program appears to produce a creative output, this doesn't mean that it is inherently creative—it just follows the rules that produce output from a human creation in an automated manner. To take this further, we could even state that machines programmed to mimick human creativity and produce artefacts that appear creative are—in the philosophical manner defined by David Chalmers—*Zombies* [3]. Similarly Douglas Hofstadter argues that minds cannot be reduced to their physical building blocks (or their most basic rules) in his "Conversation with Einstein's Brain" [40]. This school of thought is employed to demonstrate that *mind* is not just physical *brain*. We are introducing it here to argue that computers do not *consciously create* as do humans, because they are not conscious.

Edsger Dijkstra pointed out that computer science is infantalised [41] and there is a danger that the same thing is happening to creativity research. In other words, it may be an over-simplification to reduce creativity down to a four step process, or a product that is novel, valuable and of high quality. A framework that makes the evaluation of creativity appear to be a matter of checking boxes is surely missing the subjective nature of creativity. The real picture is far more interwoven and—although creativity may spring from a finite set of causes—these can interact in a complex manner that cannot be assessed so neatly.

> "User of tools are much more prevelant than makers of tools. This imbalance has traditionally been rooted in the vast difference in skill levels required for using a tool compared to making a tool: To use a tool on a computer, you need to do little more than point and click. To create a tool, you must understand the arcane art of computer programming. A strange reverse phenomenon is in motion today: As programming becomes easier and more accesible, the tools for expression are becoming more complex and diffuclt to use. Programming tools are increasingly oriented toward fill-in-the-blank approaches to the construction of code, making it easy to create programs but resulting in software with less originality and fewer differentiating features." [42]

To sum up our approach: rather than a linear or cyclic series, or criteria that can be answered in a binary manner (i.e. present or not) we propose scales or spectra to aid in the evaluation of a creative artefact of any kind, by applying a series of overlapping principles that encourages a more intuitive assessment.

The next stage for this approach would be to test the evaluation framework with real-world examples and individuals responsible for creative output or its assessment, for instance: artists, dancers, musicians, arts administrators, critics, curators and commentators.

If anything that falls short of true computational creativity is considered a *zombie*, then as long as computers continue to be regarded as autonomously creative, we may already be trapped in a *zombie apocalypse*.

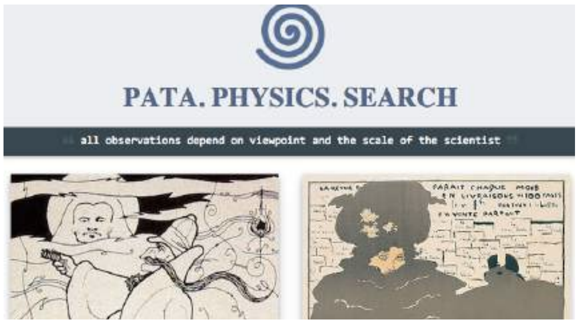
References

[1] L. Candy and E. Edmonds, Eds., *Interacting: Art, Research and the Creative Practitioner*. Libri Publishing, 2011.
[2] J. Maeda, *Design by Numbers*. MIT Press, 2001.
[3] D. Chalmers, *The Conscious Mind*. Oxford University Press, 1996.
[4] N. McBride, "A Robot Ethics: The EPSRC Principles and the Ethical Gap," in *AISB / IACAP World Congress 2012 Framework for Responsible Research and Innovation in AI*, no. July, 2012, pp. 10–15.
[5] A. Hugill and H. Yang, "The creative turn: new challenges for computing," *International journal of Creative Computing*, vol. 1, no. 1, pp. 4–19, 2013.
[6] H. Yang, "Editorial," *International journal of Creative Computing*, vol. 1, no. 1, pp. 1–3, 2013.
[7] S. Colton and G. A. Wiggins, "Computational Creativity: The Final Frontier?" in *Proceedings of the 20th European Conference on Artificial Intelligence*. IOS Press, 2012, pp. 21–26.
[8] R. E. Mayer, "Fifty Years of Creativity Research," in *Handbook of Creativity*, R. J. Sternberg, Ed. Cambridge University Press, 1999, ch. 22, pp. 449–460.
[9] M. Rhodes, "An analysis of creativity," *The Phi Delta Kappan*, vol. 42, no. 7, pp. 305–310, 1961.
[10] R. J. Sternberg, *Handbook of creativity*. Cambridge University Press, 1999.
[11] M. Boden, *The Creative Mind: Myths and Mechanisms*. Routledge, 2003.
[12] J. C. Kaufman and R. A. Beghetto, "Beyond big and little: The four c model of creativity," *Review of General Psychology*, vol. 13, no. 1, pp. 1–12, 2009.
[13] H. Poincaré, *The Value of Science*, S. J. Gould, Ed. Modern Library, 2001.
[14] G. Wallas, *The Art of Thought*. Jonathan Cape, 1926.
[15] G. Pólya, *How To Solve It*, 2nd ed. Princeton University Press, 1957.
[16] B. Indurkhya, "Computers and creativity," 1997, unpublished manuscript. Based on the keynote speech 'On Modeling Mechanisms of Creativity' delivered at Mind II: Computational Models of Creative Cognition.
[17] J. Schmidhuber, "New millennium ai and the convergence of history," 2006.
[18] A. Pease and S. Colton, "On impact and evaluation in Computational Creativity : A discussion of the Turing Test and an alternative proposal," in *Proceedings of the AISB*, 2011.
[19] A. Pease, D. Winterstein, and S. Colton, "Evaluating Machine Creativity," in *Proceedings of ICCBR Workshop on Approaches to Creativity*, 2001, pp. 129–137.
[20] S. Colton, A. Pease, and G. Ritchie, "The Effect of Input Knowledge on Creativity," 2001.
[21] A. Pease, S. Colton, R. Ramezani, J. Charnley, and K. Reed, "A Discussion on Serendipity in Creative Systems," in *Proceedings of the 4th International Conference on Computational Creativity*, vol. 1000. University of Sydney, 2013, pp. 64–71.
[22] G. Ritchie, "Some Empirical Criteria for Attributing Creativity to a Computer Program," *Minds and Machines*, vol. 17, no. 1, pp. 67–99, 2007.
[23] ——, "Assessing creativity," in *AISB '01 Symposium on Artificial Intelligence and Creativity in Arts and Science*. Proceedings of the AISB'01 Symposium on Artificial Intelligence and Creativity in Arts and Science, 2001, pp. 3–11.
[24] D. Ventura, "A Reductio Ad Absurdum Experiment in Sufficiency for Evaluating (Computational) Creative Systems," in *5th International Joint Workshop on Computational Creativty*, 2008.
[25] S. Colton, "Creativity versus the perception of creativity in computational systems," in *In Proceedings of the AAAI Spring Symp. on Creative Intelligent Systems*, 2008.
[26] ——, "Computational Creativity," *AISB Quarterly*, pp. 6–7, 2008.
[27] D. Piffer, "Can creativity be measured? An attempt to clarify the notion of creativity and general directions for future research," *Thinking Skills and Creativity*, vol. 7, no. 3, pp. 258–264, 2012.
[28] A. K. Jordanous and B. Keller, "Weaving creativity into the Semantic Web: a language-processing approach," in *Proceedings of the 3rd International Conference on Computational Creativity*, 2012, pp. 216–220.
[29] A. K. Jordanous, "Evaluating Evaluation : Assessing Progress in Computational Creativity Research," in *Proceedings of the Second International Conference on Computational Creativity*, 2011.
[30] ——, "Evaluating Computational Creativity: A Standardised Procedure for Evaluating Creative Systems and its Application," Ph.D. dissertation, University of Sussex, 2012.
[31] L. Candy, "Evaluating Creativity," in *Creativity and Rationale: Enhancing Human Experience by Design*, J. Carroll, Ed. Springer, 2012.
[32] Y. Kim and S. S. Sundar, "Anthropomorphism of computers: Is it mindful or mindless?" *Computers in Human Behavior*, vol. 28, no. 1, pp. 241–250, 2012.
[33] M. Brown. (2011, Jun) Patrick tressets robots draw faces and doodle when bored (wired uk). [Online]. Available: http://www.wired.co.uk/news/archive/2011-06/17/sketching-robots
[34] H. Cohen. (1999) Colouring without seeing: A problem in machine creativity. [Online]. Available: http://www.kurzweilcyberart.com/aaron/hi_essays.html
[35] R. Barthes, "The death of the author," *Aspen 5,6*, 1967, the birth of the reader must be ransomed by the death of the Author. [Online]. Available: http://www.ubu.com/aspen/aspen5and6/threeEssays.html#barthes
[36] R. Walker, "The Human Brain Project," HBP-PS Consortium, Tech. Rep., 2012.
[37] J. Schmidhuber, "Developmental robotics, optimal artificial curiosity, creativity, music, and the fine arts," *Connection Science*, vol. 18, no. 2, pp. 173–187, 2006.
[38] A. Turing, "Computing Machinery and Intelligence," *Mind*, vol. 59, pp. 433–460, 1950. [Online]. Available: http://loebner.net/Prizef/TuringArticle.html
[39] J. Searle, "Minds, Brains, and Programs," *Behavioral and Brain Sciences*, vol. 3, no. 3, pp. 417–457, 1980.
[40] D. Hofstadter, "A Conversation with Einstein's Brain," in *The Mind's I*, D. Hofstadter and D. Dennett, Eds. Basic Books, 1981, ch. 26, pp. 430–460.
[41] E. W. Dijkstra, "On the cruelty of really teaching computing science," 1988. [Online]. Available: http://www.cs.utexas.edu/users/EWD/transcriptions/EWD10xx/EWD1036.html
[42] J. Maeda, *Creative Code*. Thames & Hudson, 2004.

Pata-computed Poetry

Fania Raczinski

pata.physics.wtf



Raymond Queneau

Oulipo

Pataphysics

Technologies

- Python
- Javascript, HTML & CSS
- Flask, Jinja,
- NLTK & WordNet

creative computing vs. computational creativity

Faustroll's Corpus

Alfred Jarry, Edgar Allen Poe, Cyrano de Bergerac, Saint Luke, Lean Bloy, Samual Taylor Coleridge, Georges Darien, Marceline Desbordes-Valmore, Max Elskamp, Jean-Pierre Claris de Florian, V.A., Christian Dietrich, Gustave Kahn, Le Compte de Lautreamont, Maurice Maeterlinck, Stephane Mallarme, Mendes, Homer, Josephin Peladan, Francois Rabelais, Jean de Chilra, Henri de Regnier, Arthur Rimbaud, Marcel Schwob, Paul Verlaine, Emile Verhaeren and Jules Verne

Concepts

[**Clinamen**]

"smallest possible aberration that can make the greatest possible difference" Boek

[**Syzygy**]

alignment of three celestial bodies, pun

[**Antinomy**]

mutually incompatible, paradox

Patalgorithms

[**Clinamen**]

Introduction of spelling errors

[**Syzygy**]

Moving within the semantic hierarchy

[**Antinomy**]

Opposite

Love Poem

Love Poem

Maeterlinck, Rabelais, Verne, Poe
Mendes, Coleridge, Darien, Rabelais
Verhaeren, Homer, Homer
Homer, Bergerac, Rabelais

CCCC
CSSC
SSC
CCS

Future

- rhyming pattern
- sentence structure
- corpus or web
- multi-keyword

- Digital Opera
- PhD

Digital Creativity
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/ndcr20

The pataphysics of creativity: developing a tool for creative search

Andrew Hugill[a], Hongji Yang[a], Fania Raczinski[a] & James Sawle[a]
[a] Institute of Creative Technologies, De Montfort University
Published online: 12 Nov 2013.

Digital Creativity, 2013
Vol. 24, No. 3, 237–251, http://dx.doi.org/10.1080/14626268.2013.813377

The pataphysics of creativity: developing a tool for creative search

Andrew Hugill, Hongji Yang, Fania Raczinski and James Sawle

Institute of Creative Technologies, De Montfort University

a.hugill@bathspa.ac.uk; hyang@dmu.ac.uk; fania@dmu.ac.uk; jsawle@dmu.ac.uk

Abstract

We introduce the idea of a new kind of web search tool that uses the literary and philosophical idea of pataphysics as a conceptual framework in order to return creative results. Pataphysics, the science of exceptions and imaginary solutions, can be directly linked to creativity and is therefore very suitable to guide the transformation from relevant into creative search results. To enable pataphysical algorithms within our system we propose the need for a new type of system architecture. We discuss a component-based software architecture that would allow the flexible integration of the new algorithms at any stage or location and the need for an index suitable to handle patadata, data which have been transformed pataphysically. This tool aims to generate surprising, novel and provocative search results rather than relevant ones, in order to inspire a more creative interaction that has applications in both creative work and learning contexts.

Keywords: pataphysics, creativity, information retrieval, creative computing, component-based software engineering

1 Introduction

In this article we propose a new type of web search engine, reminiscent of the experience of 'surfing the Web'. This is in contrast to current search engines which value relevant results over creative ones. 'Surfing' used to be a creative interaction between a user and the web of information on the Internet, but the regular use of modern search engines has changed our expectations of this sort of knowledge acquisition. It has drifted away from a *learning process* by exploring the Web to a straightforward process of information retrieval similar to looking up a word in a dictionary.

Jorge Luis Borges has provided us with a very useful example to illustrate our idea. His 'Chinese Encyclopaedia' (Borges 2000, 231) lists the following results under the category of 'animal':

(1) those that belong to the emperor;
(2) embalmed ones;
(3) those that are trained;
(4) suckling pigs;
(5) mermaids;
(6) fabulous ones;
(7) stray dogs;
(8) those that are included in this classification;
(9) those that tremble as if they were mad;
(10) innumerable ones;
(11) those drawn with a very fine camel's hair brush;
(12) etcetera;
(13) those that have just broken the flower vase;
(14) those that at a distance resemble flies.

Although these are all perfectly valid results, it is clear that they form a more creative, even poetic, view of what an animal might be than the Oxford English Dictionary's prosaic: 'a living organism which feeds on organic matter' (Oxford Dictionaries n.d.).

To achieve this sort of creativity in search results we propose the use of pataphysical methods. Pataphysics is highly subjective and particular, and as such is very suitable for this kind of transformation from relevant to creative. We hope that the tool will prove useful as a source for information and inspiration and at the same time challenge the way we think about information retrieval on the Web. The Web is not a place limited to one discipline, and in fact creating a transdisciplinary field of 'web science' was suggested by Hendler et al. in 2008. Our project will therefore span several disciplines as well.

> *Given the breadth of the Web and its inherently multi-user (social) nature, its science is necessarily interdisciplinary, involving at least mathematics, [computer science], artificial intelligence, sociology, psychology, biology, and economics*
>
> (Hendler et al. 2008).

Over the rest of the article, we will examine how pataphysics and creativity map onto one another, give an outline of the field of information retrieval, and discuss how this new type of search could be implemented in future systems. We conclude with a short discussion and summary of the article.

2 Creativity and pataphysics

> *[Pataphysics] can only be defined in a new undiscovered language because too obvious: tautology*
>
> (Baudrillard 2007).

The creative process normally involves a move from the known to the unknown, and sometimes from the named to the unnamed. In bringing something new into existence, the human qualities of openness and tolerance of ambiguity are generally regarded as highly desirable. We may define creativity as *the ability to use original ideas to create something new and surprising of value*. We generally speak of creative 'ideas' rather than 'products', which merely provide evidence of a creative process that has already taken place. Both the originality and the value of an idea are evaluated using subjective criteria. Pataphysics, which represents an extreme form of subjectivity, is therefore a highly appropriate framework within which to encourage and enable creative thinking and operations.

2.1 Pataphysics

Pataphysics[1] was invented by a group of French schoolboys at the Lycée de Rennes in the 1880s. One of their number was the author and playwright Alfred Jarry (1873–1907), who later developed the concept both in his celebrated *Ubu* plays and in his novels and speculative writings. He defined it as follows:

> *Pataphysics . . . is the science of that which is superimposed upon metaphysics, whether within or beyond the latter's limitations, extending as far beyond metaphysics as the latter extends beyond physics. Ex: an epiphenomenon being often accidental, Pataphysics will be, above all, the science of the particular, despite the common opinion that the only science is that of the general. Pataphysics will examine the laws which govern exceptions, and will explain the universe supplementary to this one; or, less ambitiously, will describe a universe which can be—and perhaps should be—envisaged in the place of the traditional one, since the laws which are supposed to have been discovered in the traditional universe are also correlations of exceptions, albeit more frequent ones, but in any case accidental data which, reduced to the status of unexceptional exceptions, possess no longer even the virtue of originality.*
>
> *DEFINITION. Pataphysics is the science of imaginary solutions, which symbolically attributes the properties of objects, described by their virtuality, to their lineaments*
>
> (Jarry 1996, 21).

This may be summarised in following way: pataphysics:

- is the science of imaginary solutions;
- is the science of the particular;
- is the science of the laws governing exceptions and contradictions;
- is to metaphysics as metaphysics is to physics.

The conceptual space of pataphysics is a 'universe supplementary to this one' (Jarry 1996, 21). We argue that pataphysics can facilitate creative computing. *Constraints* are the rules that we set in our space, the grammar that we want to use. A pataphysical grammar would consist of exceptions, syzygies, anomalies, clinamen, antinomies, contradictions, equivalents and imaginaries. Such constraints can transform the ways in which we may navigate the new space. Pataphysical concepts will cause surprise and therefore could be considered unconventional.

Since pataphysics is concerned with the laws governing *exceptions*, its application in creative computing will focus on the ludic aspects of unique occurrences, rather than predictable recurrence of expected outcomes (Bök 2002). It is axiomatic that no single viewpoint may predominate, an understanding that was codified by Jarry and subsequent theorists as the 'doctrine' of Equivalence. Abstraction and generalisation in creative computing may therefore be founded upon a parallel we would draw between meta-metaphysics (pataphysics) and meta-metadata (patadata), which will be discussed in more detail below. Since pataphysics is the science of imaginary solutions, *imagination* (specifically a poetic imagination) provides the guiding principle for our work. Domain-specific knowledge and skill is described by the final line of Jarry's *Exploits and Opinions of Doctor Faustroll, Pataphysician:* 'Pataphysics is *the* science' (Jarry 1996, 114).

2.2 Creativity

It is instructive to overlay these ideas on existing theories of creativity. Margaret Boden (2003), for example, has defined *P-creativity* (short for psychological creativity) as the personal kind of creativity that is novel in respect to the individual mind, and *H-creativity* (short for historical creativity) as fundamentally novel in respect to the whole of human history. This allows for subjective evaluation of any idea. A child that builds a corbelled arch out of woodblocks, without any knowledge of physics or architecture, could be called creative. The child created something new and valuable within its own constraints and could therefore be called P-creative, but since the technique was already known historically it cannot be considered H-creative.

Using Boden's definition we can call an idea 'new' if it is new to the individual who came up with it, making the idea P-creative. We can say that a creative idea can be seen from two perspectives: the subjective (P-creative) and the objective (H-creative) view. She argues that constraints support creativity, and are even essential for it to happen. 'Constraints map out a territory of structural possibilities which can then be explored, and perhaps transformed to give another one' (Boden 2003, 82).

This echoes the ideas of groups such as the *Oulipo* (which began as a Sub-Commission of the Collège de 'Pataphysique), who investigate 'potential literature' by creating constraints that frequently have a ludic element. Various other groups, the Ou-x-Pos, perform similar operations in fields as diverse as cinema, politics, music and cooking (Motte 1998).

Boden's conceptual space is the 'territory of structural possibilities'. So, the conceptual space of a teacup might be that it is meant to carry a certain amount of tea without breaking or burning fingers. It wouldn't be wise to create a teacup made out of paper. But whether we make a cup out of glass or porcelain or how we shape the cup or the handle is pretty much up to the individual's creativity. Being able to move around in this conceptual space, experiment (in thought or in reality) and play with different ideas while still following a given set of constraints is a good starting point for creativity to happen. Boden defines three sub-types of creativity:

- combinational creativity: making unfamiliar combinations of familiar ideas;
- exploratory creativity: exploration of conceptual spaces;

Digital Creativity, Vol. 24, No. 3

• transformative creativity: transformation of space.

The Oulipo similarly classifies its conceptual space under two broad headings: the synthetic and the analytic:

> *In the research which the Oulipo proposes to undertake, one may distinguish two principal tendencies, oriented respectively towards Analysis and Synthesis. The analytic tendency investigates works from the past in order to find possibilities that often exceed those their authors had anticipated. . . . The synthetic tendency is more ambitious: it constitutes the essential vocation of the Oulipo. It's a question of developing new possibilities unknown to our predecessors. This is the case, for example, of [Raymond Queneau's] 100,000,000,000,000 Poems or the Boolean haikus*
>
> (Motte 1998, 27).

Later writings develop these ideas in more detail. *La Littérature Potentielle* (Oulipo 1973), is divided into several sections, dealing with clusters of methods, that include: *anoulipisms* (analytical oulipisms, such as combinatorial literature); *use of pre-existing structures* such as lipograms (omitting a letter or letters), palindromes and snowballs (in which each successive word adds or subtracts a letter), homophonic translation, tautogram, and definitional literature; *lexical, syntactic, or prosodic manipulations* (such as the celebrated S + 7, in which each substantive is replaced by the seventh word after it in a standard dictionary); *lexicographical or prosodic synthoulipisms* (early algorithmic methods); and *perimathematical synthoulipisms* (such as the Boolean poetry and combinatorial works already mentioned).

Boden links her three aspects of creativity to three sorts of surprise. She says that creative ideas are surprising because they go against our *expectations*. 'The more expectations are disappointed, the more difficult it is to see the link between old and new' (Boden 2003, 84). This suggests that fewer expectations (an open mind) allow creativity to happen more easily. Empirical experiences form expectations, which hinder our ability to accept creative ideas when they happen. In order to be able to recognise creative ideas, we need to be able to see what they all have in common and in what way they differ, and not reject unusual, unexpected ones.

> *Unless someone realizes the structure which old and new spaces have in common, the new idea cannot be seen as the solution to the old problem. Without some appreciation of shared constraints, it cannot even be seen as the solution to a new problem intelligibly connected with the previous one*
>
> (Boden 2003, 84).

It is clear that the Oulipo has a similar approach in its theorising of potential literature. Releasing creativity through constraint is its essential *raison d'être*.

This is not to say that experience and knowledge are necessarily bad for creativity. To appreciate creativity we need to be knowledgeable in the relevant domain to be able to recognise old and new connections and transformations. But we also need a certain level of openness and tolerance for ambiguity to overcome our expectations. Perhaps it is for this reason that 'creative people' are often assumed to have particular personality traits. Sternberg (1988, 1999), for example, proposes that these comprise: independence of judgement, self-confidence, attraction to complexity, aesthetic orientation, tolerance for ambiguity, openness to experience, psychoticism, risk-taking, androgyny, perfectionism, persistence, resilience, and self-efficacy. More empirically, Heilman, Nadeau, and Beversdorf (2003) have investigated the possible brain mechanisms involved in creative innovation. While a certain level of domain-specific knowledge and special skills are necessary components of creativity, they point out that 'co-activation and communication between regions of the brain that ordinarily are not strongly connected' (Heilman, Nadeau, and Beversdorf 2003, 269) might be equally important.

Newell, Shaw, and Simon (1963) add to the above with their report on the creative thinking process. They identify three main conditions for creativity: the use of imagery in problem solving; the relation of unconventionality to creativity; and the role of hindsight in the discovery of new heuristics. Other issues they point out are abstraction and generalisation. So, for example, poets transform the grammar of their conceptual space (in this case, language) to create new sentence structures in a poetic form. By doing so, they go against the expectations, the possibilities of the language and cause surprise. Some people might not understand the transformations and therefore the jokes or beauty of a poem simply because they are either not able to recognise connections between the old and newly transformed elements (maybe due to a lack of knowledge in the poems topic or in that particular language) or because they do not want to accept unconventional methods.

2.3 Creative computing

But how may we apply the insights into creativity described above in computing? One approach is described by Simon Colton (2008), who suggests we should adopt human skill, appreciation and imagination:

> *Without skill, [computers] would never produce anything. Without appreciation, they would produce things which looked awful. Without imagination, everything they produced would look the same.*
>
> (Colton 2008, 6)

He thinks that evaluating the worth of an idea or product is the biggest challenge facing *computational creativity*. Whereas in conventional problem-solving success is defined as finding a solution, in a creative context more aesthetic considerations have to be taken into account. He suggests three ways for computer programs to generate creative artefacts:

(1) mimicking human skill
(2) mimicking human appreciation
(3) mimicking human imagination.

Since our solutions will be imaginary, our aim is not so much to have the computer generate creative artefacts as to engage in a creative dialogue with the user. Therefore, we do not intend to move as close to artificial intelligence as Colton's framework seems to suggest. In the pataphysical universe, ideas such as 'human skill', 'human imagination' and 'human appreciation' are too generalised to be useful. One may very well ask: *which* human? And *when*, *where* and even *why?* Rather, our project will aim to produce an exceptional computational entity that consistently generates surprising and novel provocations to the users, who in turn may navigate and modify these by deploying their own skills, appreciation and imagination. The relationship between the two will develop quite rapidly into one of mutual subversion since, however apparent the 'rules of the game' may become, the outcomes will always be particular or exceptional.

2.4 Pataphysical computing

We are not the first people to attempt to apply pataphysical ideas in computer science. Johanna Drucker focused specifically on the cleft between formal logic and subjective judgement. She introduced the discipline of 'speculative computing' as a solution to that problem (Drucker and Nowviskie 2007). The concept can be understood as a criticism of mechanistic, logical approaches that distinguish between subject and object.

> *Speculative computing takes seriously the destabilization of all categories of entity, identity, object, subject, interactivity, process, or instrument. In short, it rejects mechanistic, instrumental, and formally logical approaches, replacing them with concepts of autopoiesis (contingent interdependency), quantum poetics and emergent systems, heteroglossia, indeterminacy and potentiality, intersubjectivity, and deformance. Digital humanities is focused on texts, images, meanings, and means. Speculative computing engages with interpretation and aesthetic provocation*
>
> (Drucker 2009, 29).

For Drucker, *aesthesis* (ambiguous and subjective knowledge) is fundamentally opposed to *mathesis* (formal objective logic) and subjectivity is always in opposition to objectivity. Knowledge is a matter of interpretation of information, which can be represented digitally as data and metadata. She introduces what she calls a '*patacritical*' method of including exceptions as rules, even if repeatability and reliability are compromised. Bugs and glitches are privileged over functionality, and are 'valuable to speculation in a substantive, not trivial, sense' (Drucker 2009, 26). As she says: ''Pataphysics inverts the scientific method, proceeding from and sustaining exceptions and unique cases' (Drucker and Nowviskie 2007, 434).

In order to break out of the formal logic and defined parameters of computer science, she asserts, we need speculative capabilities and pataphysics. 'The goal of pataphysical and speculative computing is to keep digital humanities from falling into mere technical application of standard practices' (Drucker and Nowviskie 2007, 441). She links interface design with other speculative computing principles, referring to Kant's idea of art as 'purposiveness without purpose' and saying that the appreciation of design as a thing in itself (regardless of utility) is a goal of speculative aesthetics (Drucker and Nowviskie 2007, 437).

Table 1.

Creativity	Pataphysics
Combinational Juxtaposition of dissimilar, Bisociation, Deconceptualisation	Antinomy Symmetry, duality, mutually incompatible, contradicting, simultaneous existence of mutually exclusive opposites Syzygy Alignment of three celestial bodies in a straight line, pun, conjunction of things, something unexpected and surprising
Exploratory Noticing new things in old places	Anomaly Exceptions, equality
Transformative Making new thoughts possible by transforming old conceptual space, altering its own rules	Clinamen Unpredictable swerve, the smallest possible aberration that can make the greatest possible difference

2.5 Creativity and pataphysics compared

To conclude this discussion, consider Table 1, which compares some of the key ideas of creativity (Boden 2003; Bök 2002; Indurkhya 1997; Koestler 1964) with the main pataphysical operations. It will be seen that pataphysics succeeds in bringing into sharp relief the more generalised scientific ideas. The pataphysical terms are taken from the natural sciences or philosophy, but always with an ironic twist, betraying their underlying humour. They connect quite strongly with the primary descriptors of creativity, while adding a certain layer of *jouissance*. *Pataphysics is self-avowedly useless, but its principles may prove surprisingly useful within this context*.

3 Information retrieval systems

Information retrieval is one of the common processes that a person carries out day-to-day, usually without even thinking about it. The amount of information that a human comes in contact with on a daily basis is overwhelming, and as such we have developed very sophisticated methods of finding the *relevant information* instantaneously. However, it is also possible to see how this relates to a large number of commonly used computer systems.

> *Information retrieval (IR) is finding material (usually documents) of an unstructured nature (usually text) that satisfies an information need from within large collections (usually on local computer servers or on the Internet)*
>
> (Manning, Raghavan, and Schütze 2008, 1).

It is important to note that whilst a large proportion of information retrieval (IR) is focused on web search engines, this is not the only application. The reason that such a large focus is on this area is due to the unique challenges it holds: huge quantities of *unstructured data* which change over time and can be in a number of formats. The true aim of any research into search engines is that it can be applied back to the general field of IR and enhance a much larger ecosystem of systems.

However, research in all of IR focuses on arbitrary values of success, called precision and recall, the fraction of retrieved instances that are relevant and the fraction of relevant instances that are retrieved respectively. Whilst these measures are logical, they are arbitrary due to the subjectiveness of relevance. And due to the clinical nature of the measures, returning results that are partially related to the request would be detrimental to the perceived quality of the system, irrelevant of the insightful knowledge they may provide.

Whilst IR systems can take many different forms, Baeza-Yates and Ribeirio-Neto (2011) defined a standard model, which allows all systems to be broken down into similar components:

> *An IR model is a quadruple* [D, Q, $\mathcal{F}$, R(q_i, d_j)] *where*
>
> - D *is the set composed of logical views (or representations) for the documents in the collection;*
> - Q *is the set composed of logical views (or representations) for the user information needs. Such representations are called queries;*
> - $\mathcal{F}$ *is a framework for modelling document representations, queries, and their relationships… ;*
> - R(q_i, d_j) *is a ranking function which associates a real number with a query representation* $q_i \in Q$ *and a document representation* $d_j \in D$. *Such ranking defined an ordering among the documents with regard to the query* q_i.
>
> (Baeza-Yates and Ribeirio-Neto 2011, 58)

It is possible, under this definition, that there is no ranking function; such is the case for the Boolean model. Whilst this may not appear logical when considering search engines, there are a number of cases where returning all possible results which match our 'need' without bias can be useful. It is not possible, however, for an IR system to exist without any of the other components.

3.1 Classical IR models

The classification of classical IR models typically includes the Boolean, vector space and probabilistic models (Dominich 2000). Each of the models is built on pure mathematical underpinnings, which has also lead to research into a unified model for them.

The Boolean retrieval model is based on set theory and Boolean algebra. The model views documents as a collection of words or, more precisely, a collection of indexed terms present in those documents. A user request (query) is usually a Boolean expression written as a series of terms connected by Boolean operators such as AND, OR and NOT.

In the vector space model, documents are represented as vectors (Wong and Raghavan 1984). The success or failure of this method is based on term weighting. Terms are words, phrases, or any other indexing units used to identify the contents of a text. As such, term weighting is assigning a value to each term in order to define its importance in relation to the rest of the terms within that context (Salton and Buckley 1987). Polettini (2004) points out that term weighting schemes play an important role for the similarity measure, which plays a key role in the retrieval performance of IR systems.

Due to the fact that vector space models only link documents through related terms, we have no in-built technique to handle relevance. The aim of probabilistic methods is to rank a collection of documents in decreasing probability of their relevance to a query. This is often referred to as the probabilistic ranking principle (Cooper 1968). The idea of using probability was suggested as no system can predict with certainty the documents that a requester might find useful (Maron and Kuhns 1960).

3.2 Latent semantic indexing

Latent semantic indexing (LSI) is an indexing and retrieval model that attempts to identify patterns in the relationships between the terms and concepts contained in an unstructured collection of text. A key feature of LSI is its ability to extract the conceptual content of a body of text by establishing associations between those terms that occur in similar contexts (Deerwester et al. 1990). The model is based upon a mathematical method called single value decomposition as well as correspondence analysis (Benzécri 1973). LSI overcomes two of the most problematic constraints of Boolean keyword queries: synonymy and polysemy. Synonymy and polysemy are often the cause of mismatches in the vocabulary used by the authors of documents and the users of information retrieval systems (Furnas et al. 1987). There has been some promising research into using LSI instead of the vector space model (Chen and Tai 2009).

3.3 Artificial intelligence models

Numerous artificial intelligence-inspired models have been proposed, from neural networking, genetic algorithms, knowledge bases and natural language processing. Each of these different systems manages to solve a different problem within the field of IR; however, trying to generalise these models has not proved as fruitful as they were in their specialised fields.

4 Beyond the realm of traditional IR systems

Most modern web search engines, excluding semantic search engines, have a similar architecture, irrespective of the IR model on which they are based (see Figure 1). The main reason for this is due to the generic data store at the heart of them, the inverted index, which is a very efficient method of storing and searching over the contents of documents.

In an inverted index, the contents of a document are broken into various different combinations or terms by the indexer, and a link to the original document is stored with each of these terms. This means that when searching for a *keyword*, instead of having to look at every document and its contents, the system just looks for all terms that match the request and returns the various links that match. The inverted index is quick at retrieval; however, building the index is slower.

Even with these characteristics, the inverted index is not suitable with respect to any of the above definitions of creativity. We are only able to search over the contents of the document as they are, with *no understanding of their meaning*. As such, being able to implement pataphysical themes like clinamen or syzygy would be very challenging.

It is possible to apply these concepts to a traditional search engine architecture by modifying the user's search request. Hendler and Hugill (2011) suggest that by using 'panalogies' we can model patadata and as such apply pataphysical constructs to requests. In the proposal that is outlined, the system would be applied to work on the open Web, using results from commercial search engines, as well as domain specific systems such as the British Library.

However, there is a limitation to such as system. Whilst we can modify the initial request to something with a more creative twist, the system cannot make decisions based on the underlying content of the results. As such, the quality of the results is limited by the quality of the indexer and not the search algorithm. Whilst this could be argued to be true in any search architecture, the index is built up of data that we wish to access directly, i.e. searching over the content to find a document that matches based on certain rules. With respect to creative search, it makes more sense that we look at how different parts of the document relate to each other, and other documents based upon underlying meaning, and not pure text. Even with this in mind, such a system would be adventurous from a creative standpoint over current search engines, and would provide an interesting insight into how people would respond to such a system and how important the user interface would be in such a system.

4.1 Semantic search engines

Semantic search engines would therefore seem to be a more logical fit to a pataphysics-inspired creative search engine as they will allow the creation of links between different documents based on more than the exact words used.

Digital Creativity, Vol. 24, No. 3

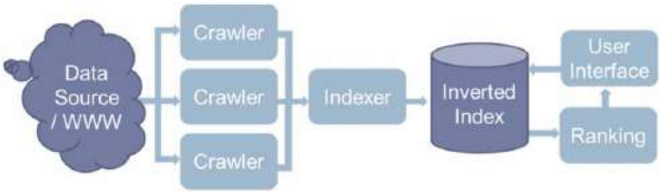


Figure 1. A traditional architecture for a web search engine. © Fania Raczinski and James Sawle. Reproduced with permission.

The key difference between the architectures of traditional IR systems is the way that the data is stored, and hence the indexing process. The majority of different semantic search systems use *Resource Description Framework (RDF) triples* as a way of storing data, based on semantic web ontologies. In RDF, each entry to the data store has the following attributes: <<object> > << relation> > <<value>>. For example, a blue balloon would be <<balloon>> <<hasColour>> <<blue>>. This is not meant to represent the syntax of RDF; however, for this example the relation of 'hasColour' and the concept of blue being a colour would have to already have been defined. However, trying to represent the concept of 'twelve blue balloons' requires even more relations and concepts to be defined; therefore, if we end up with a large amount of loosely related data, the number of concepts and relations defined will explode. However, if the data are tightly related, the number of relations and concepts is much more concise and cogent.

With the data stored in this format, *inference logic* and/or *fuzzy set theory* can be used to carry out searches over the data set to return concepts that relate. These inference searches are slow and tend not to relate directly to documents, instead returning a list of different concepts which can then be linked back to the documents. This is usually done with an inverted index using traditional methods to return documents that match numerous combinations of the concepts.

With this method, the trickiest part of the system is *indexing*, as a document must be related back to concepts that exist within the system already. If they do not exist in the system already, the concept must be found in an *ontology* that has already been defined, which then leads to the problem of ontology merging, or creating them from scratch.

Whilst this clearly allows more for the concepts that we have defined for results to be creative, there are a number of issues that arise. For example, once a document has been added to the system, the concepts are set in stone. Whilst the RDF store will evolve over time and hence change the concept results that emerge, the document's classifications are set in stone. This is not just a problem for a creative search engine, but for all semantic search engines as well.

Also, within an ontology, different concepts tend to be linked with relations that are descriptive, such as <<isAn> >, <<hasColour> >, <<produces> >, etc. However, these types of relations are not analogous with the pataphysical concepts that have been defined, and as such it is not immediately apparent how one could implement a syzygistic transformation of a search request using an RDF data store.

As can be seen by looking at the two main search architectures, a new IR system architecture is needed; however, instead of defining it from the top–down, the algorithms need to be defined first to allow maximum flexibility in the system to allow the definition of creativity to evolve over time.

5 Pataphysical search algorithms

The conceptual space for our project is 'pataphysical Web searching'. There are some very simple rules or constraints that form an initial definition of the project. For example, it is clear that we want to search the World Wide Web (rather than a library database), that we want to return a list of search results (and not a pile of books), and that we want the search process and its results to be creative/pataphysical (rather than relevant). In a more technical sense, we have the query term(s), the index (of all web pages that we have crawled) and some pataphysical rules in our conceptual space. How we structure our search system, how we format the index or how we go about finding our results is not in our conceptual space, however. We can explore the space to its limits, and we can transform it if we want to or feel like we need to. Our pataphysical rule set will include methods for transforming the space. By applying pataphysical rules to find results to our query we are *pataphysicalising* the query.

> Definitions:
> *To pataphysicalise (verb):* apply pataphysical transformations;
> *Pataphysicalisation (noun):* the process of pataphysicalising;
> *Patadata (noun):* any data which have been pataphysicalised.

The idea of patadata is derived from the idea below:

> *Physics → Metaphysics → Pataphysics*
> *Data → Metadata → Patadata*

But what exactly does the process of pataphysicalisation include? The kinds of transformations we are thinking of could be, for example, replacing or adding to the query term(s) with synonyms, antonyms, opposites, syzygies, clinamens, etc. This can be done with the help of thesauri or dictionaries and ontologies. Whether we pataphysicalise our query term(s), the index or the results does not matter at this point. They are all possible and will maybe be done all at the same time (see Figure 2). We can consider the possibility of a *patametric index* rather than a parametric index or a *patasaurus* (pataphysical thesaurus/ontology).

> *Arguably, few other textual forms will have greater impact on the way we read, receive, search, access, use and engage with the primary materials of humanities studies than the metadata structures that organize and present that knowledge in digital form*
> (Drucker 2009, 9).

Patadata will allow us to engage with digital knowledge in a more creative way even. If metadata help us *organise information semantically*, then patadata are for *organising information pataphysically*. If metadata are objective, then patadata are subjective. Drucker also points out that 'many information structures have *graphical analogies* and can be understood as diagrams that organise the relations of elements within the whole' (Drucker 2009, 16, emphasis added). So, maybe patadata could allow us to represent these graphical analogies in some way? An alphabetical list is a typical model for representing text data sets, for example. Or an otherwise ranked list, a tree structure, a matrix, a one-to-many relationship, etc. But is a ranked list really the best way to represent search results? *Ranking* itself seems unpataphysical. It contradicts the philosophy of pataphysics, although we can argue that this contradiction makes it pataphysical again. Maybe this dilemma can be solved simply by adopting another type of graphical analogy to structure the results, such as a tree structure instead of a ranked list.

In a traditional web search, ranking signals contribute to the improvement of the ranking process. These can be content signals or structural signals. *Content signals* are referring to anything that is concerned with the text and content of a page. This could be simple word counts or the format of text such as headings and font weights. The *structural signals* are more concerned about the linked structure of pages. They look at incoming and outgoing links on pages. There are also *web usage signals* that can contribute to ranking algorithms such as the clickstream. This also includes ideas such as the Facebook 'like' button or the Google ' + 1' button, which could be seen as direct user-relevance feedback.

Ranking can be done at different stages of the search process. Depending on how the index is formatted and what information can be

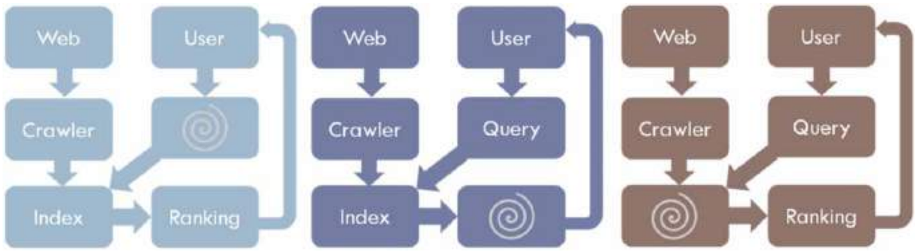


Figure 2. Three possibilities where pataphysicalisation can happen. © Fania Raczinski. Reproduced with permission.

pre-computed at that stage, the ranking algorithm evaluates every web page for *relevance* and returns them in order. There exist lots of different approaches on ranking, including PageRank (Brin and Page 1998) and HITS (Kleinberg 1999), which both analyse the link structure of the World Wide Web. They analyse the incoming and outgoing links on pages. PageRank, for example, assigns a numerical weight to each document, where each link counts as a vote of support in a sense. It is executed at *indexing time*, so the ranks are stored with each page directly in the index. HITS stands for 'Hyperlink Induced Topic Search', and its basic features are the use of so-called hubs and authority pages. It is executed at *query time*. Pages that have many incoming links are called authorities and pages with many outgoing links are called hubs.

Given a query term X, what is considered a relevant match, though? Do we simply return a list of web pages where X appears in the heading of each page? It is obviously not that easy. Several ranking signals are combined together; Google states that it uses over 200 signals including PageRank, and it *personalises* results using signals such as the web history and location (Google, n.d.).

What kinds of ranking signals do we need for our pataphysical web search tool? We could say that a page Y is relevant if it matches the patadata for query X. So, for example, Y would be a relevant result if it is a clinamen or syzygy to X. The more patadata matches there are, the higher the ranking, maybe. We don't necessarily have to assign a numerical ranking value to each page. Depending on how we structure our results page, that might not be necessary. Shuffling the results list or the results tree could be an option.

For example, let's say our patadata are represented by a list of keywords that each stands for a pataphysicalisation of the original query term. This list is added to each item in the index:

> ***Query*** *= 'Tree'*
> ***Patadata*** *= [Tree(equivalent), Car(opposite), Paper(antinomy), Narwhal(anomaly), Book (syzygy), Venus Fly Trap(clinamen)]*
>
> ***Query*** *= 'Sun God Ra'*
> ***Patadata*** *= [Sun God Ra(equivalent), Slave(opposite), Holiday(antinomy), Blue Balloon(anomaly), Pyramid(syzygy), Sphinx(clinamen)]*

6 A new architecture for search

It is clear that any of these new algorithms, or ones that follow, will not be suitable for existing system architectures in IR research, and as such a new one will need to be defined. The question becomes whether or not the architecture itself can help enhance the chance of providing creative search results. If so, would it be possible to abstract this so that it can be used in other types of systems to help allow creative computing to flourish in areas where it may not have been possible before? This is a tall order, and one that is not likely to come soon; however, developing an architecture that is as generic as possible can only aid this task.

The concept of pataphysicalisation, using pataphysical methods to transform an object/idea, on the search request does appear to be an interesting place to start the search for a new architecture. While it can certainly constrain the possible amount of creative outputs and the general characteristics of such a system, it will give valuable insight into areas that need to be addressed by both the algorithms and architecture.

A *component-based system architecture* is therefore proposed to allow for greater flexibility in search engine development, whilst reducing the coupling between different parts of the system. This coupling tends to mean that for a new concept to be tried, large proportions of the entire search engine need to be redeveloped. The use of standard interfaces, for different types of components, would therefore allow a *generic harness* to handle the communication between these different components and provide a seamless service to the end user. The wiring of these components could be handled by a configuration file, therefore allowing people to build systems without needing any explicit programming skills.

Whilst this architecture itself does not explicitly improve the chance of creative results being returned, it will allow for new components to be tested in a full-scale environment in an *agile* way, and as such should allow for quicker testing. The *harness* is currently being developed, including a number of administration and monitoring tools inbuilt to aid analysis. The aim is to test the new architecture using a standard search engine and the Syzygy Surfer proposed by Hendler and Hugill (2011).

It is interesting to note how such a system could also be used in an educational environment to teach students how search engines work. Students could attempt to build systems using pre-built components and see how different arrangements of such components affect the outcome. This is very similar to the way that the Massachusetts Institute of Technology has proposed to teach children how to program using the Scratch development environment.[2] More advanced students could also develop their own components to test out theories and improve their understanding of the base concepts of not just search engines but the various fields that play a role in information retrieval systems.

It is clear that this system could have great advantages outside of being a testbed for new ideas, allowing for the easy development of search engines to suit the needs of all new types of problems, without the need for specific development of every component. Whilst this idea is still within its infancy, the potential is strong and will be explored over time.

7 Discussion

Whilst developing a system that returns creative results to the end user has numerous advantages, the assumptions that are made about and the decisions we take for the user must still be considered. For example, presume that the user inputs a search request ‘The Cat in the Hat’ after reading a Dr Seuss book to their child, and the system employs an anomalous method on the query and searched ‘sunglasses’. Whilst there is logic to the new search request, it is anomalous to the initial request; if the user receives these results without being told what method was used, the results will appear *random*, and therefore are likely to be detrimental to the user. Therefore, the level of *interaction* the user has with the system and the *feedback* the system gives to the user on decisions it is making will have a large

Figure 3. Example of an icon explaining the pataphysicalisation of a search result. © Fania Raczinski. Reproduced with permission.

influence on the overall effectiveness and appreciation of the search tool. A quick and simple solution to this problem would be to add an icon to the side of each search result which displays how the original query was pataphysicalised.

The image in Figure 3 shows an example of how this could be implemented. The little green candle (a reference to pataphysics in itself by the way) shows a pop-up note when hovered over with the mouse pointer. In this case, the original query could have been 'tree', and 'car' was returned as an opposite to that.

In the end, it comes to a point of being able to identify which of these factors will affect how the user perceives the results and which do not, and therefore give the system greater flexibility. This in itself is a huge undertaking, with which large quantities of empirical data will be required, and is therefore left for future work on the project.

8 Conclusion

Current information retrieval systems might be used for creative purposes. However, they do not directly provide creative results to their users; instead they focus on precise and relevant results only. Therefore, we argue that a new style of system is required. It is clear that the fundamental problem in this is that standard algorithms are not suited for these problems, with them considering a document to be groupings of words in traditional IR systems, and that an entire document falls under the same classifications in semantic IR systems.

The proposed concept for a pataphysical algorithm requires precise data structures to represent the transformations that have taken place during the pataphysicalisation, such as the patadata. The system's index has to be adapted to accommodate this new type of data structure. It also needs to be flexible enough to allow algorithms to fit in at different stages or locations of the system; for example, the inverted index, ranking functions or query itself.

Whilst this new style of algorithm has been proposed, current architectures are not capable of supporting it. As such, a new, flexible component-based software architecture has been proposed which will allow for a range of different style systems to be developed with little overhead, thereby improving the chance of creative outcomes occurring in a different way.

We have introduced the motivation and concept for a creative web search tool and discussed some of the major challenges a project like this faces. With web search being a major research and learning tool nowadays, it is imperative to think about *how* such a tool could be (ab)used. Ethical issues that arise through the provision of unexpected results, and the misunderstandings this could lead to, will be discussed in future work. Nevertheless, we believe that creative web search can facilitate inspirational learning through an exploratory search journey, and we hope our tool will provide just that.

Acknowledgements

We would like to thank Professor Jim Hendler at the Rensselaer Polytechnic Institute for his valuable thoughts on this work.

Notes

[1] Although note how the perplexing apostrophe that sometimes appears before the word 'pataphysics undermines too literal an interpretation of this construction. Jarry only ever used the apostrophe on a single occasion, specifying that he did so 'in order to avoid a simple pun' (Jarry 1996, 21). What that pun might be has never been fully explained.

[2] See http://scratch.mit.edu/.

References

Baeza-Yates, D. R., and D. B. Ribeiro-Neto. 2011. *Modern Information Retrieval: The Concepts and Technology Behind Search*. London: Addison Wesley.

Baudrillard, J. 2007. *Pataphysics*. Translated by Drew Burk. CTheory. Accessed December 25, 2012. http://www.ctheory.net/articles.aspx?id=569.

Benzécri, J. P. 1973. *L'Analyse des Données*. Paris: Dunod.

Boden, M. A. 2003. *The Creative Mind: Myths and Mechanisms*. London: Routledge.

Bök, C. 2002. *'Pataphysics: The Poetics of an Imaginary Science*. Evanston, IL: Northwestern University Press.

Borges, J. L. 2000. "The Analytical Language of John Wilkins." In *Selected Non-Fictions*, edited by E. Weinberger, 229–232. London: Penguin Books.

Brin, S., and L. Page. 1998. "The Anatomy of a Large-Scale Hypertextual Web Search Engine." *Computer Networks and ISDN Systems* 30 (1–7): 107–117. doi: 10.1016/S0169-7552(98)00110-X.

Chen, B., and X. Tai. 2009. "Semantic Retrieval Using Ontology and Document Refinement." *Proceedings of the Second Symposium International Computer Science and Computational Technology*, 302–309. Oulu, Finland: Academy Publisher.

Colton, S. 2008. "Computational Creativity." *Artificial Intelligence and the Simulation of Behaviour Quartley* 126: 6–7.

Cooper, W. S. 1968. "Expected Search Length: A Single Measure of Retrieval Effectiveness Based on the Weak Ordering Action of Retrieval Systems." *American Documentation* 19: 30–41.

Deerwester, S., S. T. Dumais, G. W. Furnas, T. K. Landauer, and R. Harshman. 1990. "Indexing by Latent Semantic Analysis." *Journal of the American Society for Information Science* 41 (6): 391–407.

Dominich, S. 2000. "A Unified Mathematical Definition of Classical Information Retrieval." *Journal of the American Society for Information Science* 51: 614–624.

Drucker, J. 2009. *SpecLab: Digital Aesthetics and Projects in Speculative Computing*. Chicago, IL: University of Chicago Press.

Drucker, J., and B. Nowviskie. 2007. "Speculative Computing: Aesthetic Provocations in Humanities Computing". Chap. 29 in *A Companion to Digitial Humanities,* edited by S. Schreibman, J. Unsworth, and R. Siemens, 431–447. Oxford: Blackwell. doi: 10.1002/9780470999875.ch29.

Furnas, G. W., T. K. Landauer, L. M. Gomez, and S. T. Dumais. 1987. "The Vocabulary Problem in Human–System Communication." In *Communications of the ACM,* 30 (11): 964–971.

Google. n.d. "Ten Things We Know to be True." Google Company Overview. Accessed December 25, 2012. https://www.google.com/intl/en/about/company/philosophy/

Heilman, K. M., S. E. Nadeau, and D. O. Beversdorf. 2003. "Creative Innovation: Possible Brain Mechanisms." *Neurocase* 9 (5): 369–379.

Hendler, J., and A. Hugill. 2011. "The Syzygy Surfer: Creative Technology for the World Wide Web." In *Proceedings of the 3rd International ACM Conference on Web Science.* June 14–17 2011, Koblenz, Germany. http://www.websci11.org/fileadmin/websci/Posters/3_paper.pdf

Hendler, J., N. Shadbolt, W. Hall, T. Berners-Lee, and D. Weitzner. 2008. "Web Science: An Interdisciplinary Approach to Understanding the Web." In *Communications of the ACM* 51 (7): 60–69.

Indurkhya, B. 1997. "Computers and Creativity." Unpublished manuscript based on the keynote speech "On Modeling Mechanisms of Creativity" delivered at Mind II: Computational Models of Creative Cognition, Dublin, September 15–17.

Jarry, A. 1996. *Exploits and Opinions of Dr Faustroll, Pataphysician*. Cambridge, MA: Exact Change.

Kleinberg, J. M. 1999. "Authoritative Sources in a Hyperlinked Environment." *Journal of the ACM* 46 (5): 604–632.

Koestler, A. 1964. *The Act of Creation*. London: Penguin.

Manning, C. D., P. Raghavan, and H. Schütze. 2008. *Introduction to Information Retrieval*. New York: Cambridge University Press.

Maron, M. E., and J. L. Kuhns. 1960. "On Relevance Probabilistic Indexing and Information Retrieval." *Journal of the ACM* 7 (3): 216–244.

Motte, W. 1998. *Oulipo: A Primer of Potential Literature*. Lincoln, NE: University of Nebraska Press.

Newell, A., J. G. Shaw, and H. A. Simon. 1963. *The Process of Creative Thinking*. New York: Atherton.

Oulipo. 1973. *La Littérature Potentielle*. Paris: Gallimard.

Oxford Dictionaries. n.d. Oxford University Press. Accessed December 25, 2012. http://oxforddictionaries.com/definition/english/animal.

Polettini, N. 2004. "The Vector Space Model in Information Retrieval – Term Weighting Problem." Trento: Department of Information and Communication Technology, University of Trento.

Salton, G., and C. Buckley. 1987. "Term Weighting Approaches in Automatic Text Retrieval." Technical Report." Ithaca, NY: Cornell University.

Sternberg, R. J. 1988. *The Nature of Creativity.* Cambridge: Cambridge University Press.

Sternberg, R. J. 1999. *Handbook of Creativity.* Cambridge: Cambridge University Press.

Wong, S. K. M., and V. V. Raghavan. 1984. "Vector Space Model of Information Retrieval: A Reevaluation." In *Proceedings of the 7th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 167–185. Swinton: Bristish Computer Society.

Andrew Hugill is Director of Creative Computing at Bath Spa University. He is a composer and transdisciplinary researcher. He is the author of *Pataphysics: A Useless Guide* (MIT Press, 2012) and *The Digital Musician* (Routledge, 2008). He is the former Director of the Institute of Creative Technologies at De Montfort University.

Hongji Yang is Deputy Technical Director of the Software Research Technology Laboratory at De Montfort University. He is a Golden Core Member of IEEE Computer Society. He has been organiser for many international conferences, such as a Programme Co-Chair at IEEE International Conference on Software Maintenance 1999 (ICSM '99) and as the Programme Chair at IEEE Computer Software and Application Conference 2002 (COMPSAC '02). His research interests are software engineering, internet computing and creative computing. He has published intensively in the area of software evolution. He is currently developing a new journal, the *International Journal of Creative Computing* (IJCrC), to supply a unique forum for scholars to discuss how computing can support creativity.

Fania Raczinski is currently working on her PhD about pataphysical search algorithms, having completed a BSc in computer science (Europe) at the University of Leicester and an MSc in creative technologies from De Montfort University in Leicester. She has interests in written (programming) and spoken (natural) languages, creative computing, Semantic Web, pataphysics, culture and arts.

James Sawle is also working on his PhD and is looking into a new system architecture for search engines. He received his BSc (Hons) in computer science and mathematics from Durham University in 2009 before completing his MSc in software engineering in the Software Technology Research Laboratory at De Montfort University, Leicester. His research interests include IR systems, Semantic Web, computational creativity and pataphysics.

Creative Search Using Pataphysics

Fania Raczinski, Hongji Yang, Andrew Hugill
Institute of Creative Technologies

Pataphysics

- is the science of the **particular**
- examines the laws governing **exceptions**
- explains a universe supplementary to this one
- is the science of **imaginary** solutions
- is to metaphysics as metaphysics is to physics

"*To understand 'pataphysics is to fail to understand 'pataphysics*"
(Andrew Hugill's Useless Guide to 'Pataphysics)

Pataphysics

- the **clinamen** or the chance
- particular over the general
- exceptions over the ordinary
- the paradox or **antinomy**
- simultaneous existence of mutually exclusive opposites
- subjective
- symmetry, duality
- absurd anomalies
- transcended reality or the absolute
- epiphenomenalism
- unexpected alignment, **syzygy**

Definition of “animal”

- those that belong to the Emperor,
- embalmed ones,
- those that are trained,
- suckling pigs,
- mermaids,
- fabulous ones,
- stray dogs,
- those included in the present classification,
- those that tremble as if they were mad,
- innumerable ones,
- those drawn with a very fine camelhair brush,
- others,
- those that have just broken a flower vase,
- those that from a long way off look like flies.

4 / 10

Creative Search Using Pataphysics

Aim: surprising, novel, inspiring, humorous search

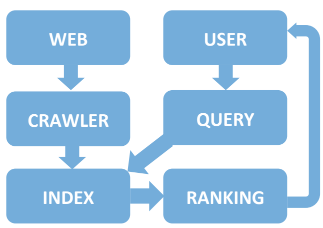
- Exploratory search, not information lookup
- Creative computing by using pataphysical algorithms
- Generate creative search results rather than relevant ones

Pataphysicalisation = applying pataphysical techniques to data

5 / 10

Information Retrieval

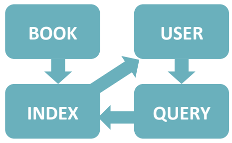
1. Information need
2. Query formulation
3. Retrieval process
4. Ranking
5. Results evaluation



6 / 10

Prototype

- Local search in book
- Text only, no images
- Tiny index
- No ranking
- Three algorithms



Exploits and Opinions of Dr Faustroll, Pataphysician (Jarry 1907)

7 / 10

Algorithms

Clinamen **clear - leaf**
Deliberate spelling errors

Syzygy **clear – disappear - vanish**
Query → synonyms → hypo/hyper/holonyms

Antinomy **clear – allow light to pass through - opaque**
Antonyms of query & antonyms of synonyms

8 / 10

Experimental Results

QUERY	CLINAMEN	SYZYGY	ANTINOMY
clear	altar, leaf, pleas, cellar	vanish, allow, bare, pronounce	opaque
solid	sound, valid, solar, slide	block, form, matter, crystal, powder	liquid, hollow
books	boot, bones, hooks, rocks, banks	dialogue, authority, record, fact	-
troll	grill, role, tell	wheel, roll, mouth, speak	-
live	love, lies, river, wave, size, bite	breathe, people, domicile, taste, see, be	recorded, dead

9 / 10

Conclusion

- Algorithms
- Larger Index
- Web
- Multimedia
- Evaluation
- Interpretation
- Apparent randomness
- Transparency

a pataphysical approach to making a creative exploratory search tool

10 / 10

References

- Bök, C. *'Pataphysics: The Poetics of an Imaginary Science*, Evanston, Illinois: Northwestern University Press, 2002
- Borges, J.L. The Analytical Language of John Wilkins *in Selected Non-Fictions*, Weinberger, E. (Ed.), London: Penguin Books, 2000
- Brotchie, A., Chapman, A., Foulc, T. and Jackson, K. (Eds.), *'Pataphysics: Definitions and Citations*, London Institute of 'Pataphysics, London: AtlasPress, 2003
- Brotchie, A. *Alfred Jarry: A Pataphysical Life*, Cambridge, MA: MIT Press, 2011
- Damerau, F.J. A Technique for Computer Detection and Correction of Spelling Errors in *Communications of the ACM*, 7(3), 1964, pp.171-176
- Foucault, M. (1966). *The Order of Things.* France: Editions Gallimard.
- Hendler, J. and Hugill, A. The Syzygy Surfer : Creative Technology for the World Wide Web, in *Proc. WebSci 2011,* ACM Press (2011).
- Hugill, A. *'Pataphysics: A Useless Guide*, Cambridge, MA: MIT Press, 2012
- Jarry, A. *Exploits and Opinions of Dr Faustroll, Pataphysician*, Cambridge, MA: Exact Change, 1997.
- Sutcliffe, A. and Ennis, M, Towards a cognitive theory of information retrieval, *Interacting with Computers*, vol. 10, pp. 321–351, 1998.

11 / 10

Creative Search Using Pataphysics

Fania Raczinski, Hongji Yang, Andrew Hugill
Institute of Creative Technologies
De Montfort University
Leicester LE1 9BH, UK
{fania, hyang, ahu}@dmu.ac.uk

ABSTRACT
This paper looks at defining, analysing and practicing how creativity can be applied to search tools. It defines creativity with respect to search and discusses how these concepts could be applied in software engineering using principles from the pseudo-philosophy of pataphysics. The aim of the proposed tool is to generate surprising, novel, humorous and provocative search results instead of purely relevant ones, in order to inspire a more creative interaction between a user, their information need and the application. A proof-of-concept prototype is described to justify the ideas presented before implications and future work are discussed.

Author Keywords
Creative computing; pataphysics; information retrieval; creative technologies.

ACM Classification Keywords
F.3.m. Logics and Meanings of Programs: Miscellaneous
H.3.3. Information Search and Retrieval: Search Process
J.5. Arts and Humanities: Literature

General Terms
Algorithms; Design, Experimentation.

INTRODUCTION

Imagine a web search engine that does not quite return the results you expect. For example, imagine you search for "animal" and the top three results are a list of animals in the Emperor's possession, followed by instructions about embalming animals and information on a society for animal training. Google's top search results for this query on the other hand return the webpage of an action sports lifestyle brand, the Wikipedia article and a BBC (British Broadcasting Corporation) page about animal videos. While there is certainly nothing wrong with Google's results, they are simply not very inspiring. The first example of search results is adapted from Jorge Luis Borges's *Chinese Encyclopaedia* [4] which lists several creative definitions of the term "animal". Whilst they might not provide the kind of information we were initially seeking (if we even had a clear idea of the kind of answers we wanted), they are still perfectly valid results for the query and might even provoke a smirk upon their encounter. These are the kind of search results we are aiming for; strange, creative, surprising, inspiring and possibly funny (which some would call irrelevant) yet perfectly valid.

C&C '13, June 17 - 20 2013, Sydney, NSW, Australia

Pataphysics can provide some useful techniques that are very suitable for creative computing. Hendler and Hugill first suggested the use of three of its principles: clinamen, syzygy and anomaly, in their "Syzygy Surfer" [15].

"The ambiguity of experience is the hallmark of creativity, that is captured in the essence of pataphysics. Traversing the representations of this ambiguity using algorithms inspired by the syzygy, clinamen and anomaly of pataphysics, using a panalogical mechanism applied to metadata, should be able to humanize and even poeticize the experience of searching the Web." [15]

In the rest of this paper we will introduce creativity and pataphysics and explain how they are used for our algorithms and the general philosophy during the development. We then discuss some of the implementation details for our proof-of-concept prototype and speculate on users and uses of the tool. We conclude the paper with a short discussion on further work.

CREATIVITY AND PATAPHYSICS

Creativity

We define creativity as "the ability to use original ideas to create something new and surprising of value". Here, we generally speak of creative *ideas* rather than *products*, since we believe creative products merely provide evidence of a creative process that has already taken place. Creativity is often divided into two types, one is a personal everyday type of creativity (P-creativity [2] or mini-c/little-c creativity [21]) and the other is a more eminent historical type (H-creativity [2] or Pro-c/Big-C creativity [21]). Margaret Boden further divides creativity into three categories [2, with some additional descriptions from 17, 21, 22], the concepts of which are also described in Kaufman & Beghetto's Four-C model [21].

- Combinational creativity: making unfamiliar combinations of familiar ideas; juxtaposition of dissimilar; bisociation; deconceptualisation, interpretive process of constructing and understanding

- Exploratory creativity: exploration of conceptual spaces; noticing new things in old spaces, interpretive process of constructing and understanding
- Transformative creativity: transformation of space; making new thoughts possible by altering the rules of old conceptual space, transformative learning

Boden also argues that creative ideas are surprising because they go against expectations and she believes that constraints support creativity and are even essential for it to happen. She says that *constraints map out a territory of structural possibilities which can then be explored, and perhaps transformed to give another one* [2]. This view supports our use of pataphysical concepts or constraints to enable creativity in search tools.

In many cases (especially of P-creative or mini-c types), both the originality and the value of a creative idea are evaluated using subjective or intrapersonal criteria [21]. Pataphysics, which represents an extreme form of subjectivity, is therefore a highly appropriate framework within which to encourage and enable creative thinking and operations.

Pataphysics

“To understand pataphysics is to fail to understand pataphysics.” [16]

Pataphysics was invented by a group of French schoolboys in France in the 1880s. One of their number was the author and playwright Alfred Jarry (1873-1907) [20], who later developed the concept both in his celebrated *Ubu* plays and in his novels and speculative writings. In short (there are over 100 equally correct definitions [6]) it can be defined as follows:

- Pataphysics is the science of imaginary solutions,
- Pataphysics is the science of the particular,
- Pataphysics is the science of the laws governing exceptions and contradictions,
- Pataphysics is to metaphysics as metaphysics is to physics, and
- Pataphysics describes a universe supplementary to this one.

We argue that pataphysics can facilitate creative computing. A pataphysical grammar can consist of exceptions, syzygies, anomalies, clinamen, antinomies, contradictions, equivalents and imaginaries. Such concepts or constraints can influence the ways in which we may navigate and transform our conceptual space. Pataphysical concepts are likely to cause surprise and could therefore be considered unconventional and provocative.

The concept of the clinamen can be understood as an unpredictable swerve which Bök called *the smallest possible aberration that can make the greatest possible difference* [3]. One of the most famous examples of a clinamen is Jarry’s *merdre* (the first word in his Ubu plays). He squeezed an extra ‘r’ into the French word *merde* (meaning *shit)* and translates into something like *pshit*.

A syzygy both surprises and confuses. The concept originally comes from the field of astronomy where it denotes the alignment of three celestial bodies. In a pataphysical context it usually describes a conjunction of things, something unexpected and surprising. Unlike serendipity, a simple chance encounter, the syzygy has a more scientific purpose. A typical instance is the pun, which Jarry called the *syzygy of words* [20]. Next to being intentionally funny, puns demonstrate a clever use (or abuse) of grammar, syntax, pronunciation and/or semantics, often taken to a quite scientific level, such that without understanding of what is said and what the intended meaning is, the humour of the pun might be lost.

The antinomy, in a pataphysical sense, is the mutually incompatible or paradox. Mutually contradictory opposites can and do co-exist in the pataphysical universe.

CREATIVE COMPUTING AND SEARCH

Creative Computing

The concept of creative computing has existed for some time but has not yet managed to evolve into a recognised discipline within computer science. Computational creativity, on the other hand, has emerged as a field within artificial intelligence research [18] and overlaps with creative computing ideas to some extent.

It is important to differentiate between the ideas of *creative computing* and *computational creativity*. Intuitively the former is about doing computations in a creative way, while the latter is about achieving creativity through computation. You can think of the latter falling into the artificial intelligence category (using formal computational methods to mimic creativity as a human trait, see also [18]) and the former being a more poetic endeavour of how the computing itself is done, no matter what the actual purpose of the program is.

As a good example of creative computing, consider the *International Obfuscated C Code Contest* [19]. The competition revolves around writing compilable/runnable code, while visually appearing as obfuscated as possible. They value unusuality, obscurity and creativity but expect contestants to follow the strict rules and constraints of the C programming language.

Examples of computational creativity are Simon Colton’s *Painting Fool* [9] or Harold Cohen’s *AARON* [8]; both are computer programs that paint pictures. Kurzweil’s *Cybernetic Poet* [23] is a classic example of a program that produces poetry.

Our search tool can be seen from both perspectives and therefore somewhat lies in-between. We want to use creative techniques to come up with refreshing results to provide a counter-inspiration for the relevant results provided by Google or other mainstream Web search engines. We (are trying to) use creative techniques to build something that (hopefully) also has a creative purpose and value.

Search

In simple terms, a typical search process can be described as follows. A user is looking for some information so she or he types a search term or a question into the text box of a search engine. The system analyses this query and retrieves any matches from the index, which is kept up to date by a web crawler. A ranking algorithm then decides in what order to return the matching results and displays them for the user. In reality of course this process involves many more steps and levels of detail, but it provides a sufficient enough overview.

From the users' point of view the search process can be broken down into four activities [31] reminiscent of classic problem solving techniques [29]:

1. Problem identification = information need (IN),
2. Need articulation = IN in natural language terms,
3. Query formulation = translate IN into query terms, and
4. Results evaluation = compare against IN.

Searching can be thought of in two ways, information lookup (searching) and exploratory search (browsing) [11, 24]. A situation where an information need cannot easily be articulated or in fact is not existent (the user is not looking for anything specific) can be considered a typical case of exploratory search and describes the kind of search that is most suited to our proposed tool.

Most big search engines like Google, Baidu or Bing focus on usefulness and relevance of their results. [13, 1, 26] Google uses over 200 signals [14] that influence the ranking of web pages including their original PageRank algorithm [5]. We can only speculate whether these signals also take into account any creative factors due to their secrecy. Other search engines like YossarianLives (currently in alpha release) [32] concentrate on purely abstract concepts like metaphors for their search algorithms.

Any information retrieval process is constrained by factors like subject, context, time, cost, system and user knowledge [25]. Such constraints should be taken into consideration in the development of any search tool. A web crawler needs resources to crawl around the Web, language barriers may exist, the body of knowledge might not be suitable for all queries, the system might not be able to cater for all types of queries (e.g. multi-word queries), or the user might not be able to understand the user interface, and many more. It is therefore imperative to eliminate certain constraining factors (for example by targeting a very specific audience or filtering the amount of information gathered by a crawler from web pages).

PROPOSED SYSTEM

The general concept of the project described in this paper is *pataphysical web searching* and the following three points summarize its main aims:

- search the Web for suitable answers to a given query,
- return results as a list or a mixture of data structures, and
- present pataphysical results (rather than relevant ones).

Principles

The essence of the proposed search tool lies in its algorithms which make the difference to traditional search engines. The philosophical ideology behind the tool is fundamentally different. Our system will still consist of the main components typically found in Web search engines (crawler, index and ranking) but they will have slightly different inner workings and target a different audience of users.

To link back to some of the creative, pataphysical concepts we have discussed earlier, let us put some of the ideas for our tool into perspective. The constraints for our conceptual space are the pataphysical rules that we want to apply to our data. We use those rules to explore, combine and transform our space; giving us the flexibility and freedom we need to find interesting results.

We developed the idea of *pataphysicalising* data as the process of applying such pataphysical rules in order to produce creative search results. This pataphysicalisation process forms a central component of our system (see Figure 1) and influences all areas of the search tool.

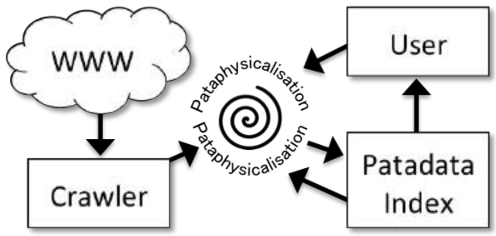


Figure 1. Pataphysicalisation as a central component

Our index will contain what Hendler and Hugill have called *patadata* [15]. Patadata is to metadata as metadata is to data - inspired by one of the definitions of pataphysics: *that which is above that which is after physics* [20]. This suggests that patadata provides another layer of information above information. If metadata helps us organise information semantically then patadata is for organising information pataphysically. If metadata is objective then patadata is subjective and that is precisely what pataphysics is for.

Prototype

The prototype described here (see Figure 2) was developed as a proof-of-concept tool to demonstrate some example search results using pataphysical algorithms. In this case the results are limited to the text of Alfred Jarry's *Exploits and Opinions of Dr. Faustroll, Pataphysician* [20] and only the main algorithmic functionality of this prototype is discussed here.

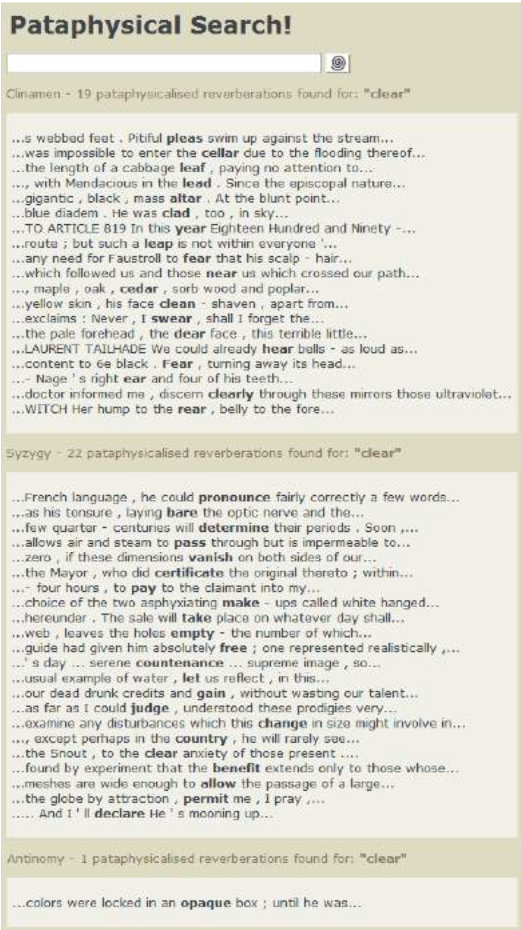


Figure 2. Screenshot of search results for the query "clear"

In short, the prototype's workflow can be described as follows:

1) tokenise text and remove stopwords to build index,

2) query triggers the three pataphysical functions,

3) each function finds matches for query as described above,

4) retrieve some words before/after match for context, and

5) return list of resulting sentences.

The three functions inspired by pataphysics (clinamen, syzygy and antinomy) are described in more detail in the next section. Figure 2 shows a screenshot of the resulting list of results for the query *clear*. The specific results for each of the three methods are simply a few words surrounding the pataphysicalised query term from within the book, which does not necessarily represent complete sentences but simply provides some context for the result.

The same principles and algorithms can be applied to different types of media, for example images or video and even sound. The complete tool would include a mixture of different types of media in its results with various styles of displaying them.

Algorithms

The clinamen function uses the Damerau-Levenshtein algorithm [10], which measures the distance between two strings (with 0 indicating equality), to find words that are similar but not quite the same. The distance is calculated using insertion, deletion, substitution of a single character, or transposition of two adjacent characters. We are basically asking the program to return matches (v) that are of distance two or one to query term t, meaning they have two or one *spelling errors* in them (see Equation 1). While we only return matches that actually appear in the book (i.e. they exist in the index), and by doing so eliminate the introduction of new words like Jarry's *merdre*, the swerve or aberration is still evident.

$$\text{clinamen}(t) = \{v : 0 < \text{dameraulevenshtein}(t, v) \leq 2\}, \text{ for } v \in V \quad (1)$$

For the syzygy function, we made use of the WordNet lexical database [30] using the NLTK python library [27] to find suitable results. Specifically, as shown in Equation 2, the algorithm fetches the set of synonyms (synsets) for query term t first and then finds any hyponyms, hypernyms or holonyms for each of those (each of which denotes a sort of relationship or membership with its parent synonym). We then return a list of all of those related terms if they appear in the original vocabulary of the text (index V). This approach mimics the syzygy alignment of three words in a line mentioned earlier (query → synonym → hypo/hyper/holonym).

$$\begin{aligned} &\text{syzygy}(t) = \{h : h \in \text{union}(t) \wedge \exists h \in V\} \\ &\quad \text{union}(t) = \text{hypo}(t) \cup \text{hyper}(t) \cup \text{holo}(t) \\ &\quad \text{hypo}(t) = \{h : h \in \text{hyponyms}(s)\} \\ &\quad \text{hyper}(t) = \{h : h \in \text{hypernyms}(s)\} \\ &\quad \text{holo}(t) = \{h : h \in \text{holonyms}(s)\} \\ &\quad \text{syno}(t) = \{s : s \in \text{synonyms}(t)\} \\ &\quad \text{for } s \in \text{syno}(t) \end{aligned} \quad (2)$$

For the antinomy function we simply made use of WordNet's antonyms (opposites) (see Equation 3). We first get all synonyms for query term t, find any antonyms for those and return any that also appear in the index V. Naturally, not all words have an opposite, especially given WordNet's limited vocabulary of around 150,000 words, but a pataphysical antinomy should still be able to find a match. This is a big shortcoming of our prototype at this point. A better thesaurus or a larger index (e.g. based on more than one book – or, of course, the Web) could improve this function drastically.

$$\begin{aligned} &\text{antinomy}(t) = \{h : h \in \text{anto}(t) \text{ and } \exists\, h \in V\} \\ &\quad \text{anto}(t) = \{h : h \in \text{antonyms}(s)\} \\ &\quad \text{syno}(t) = \{s : s \in \text{synonyms}(t)\} \\ &\quad \text{for } s \in \text{syno}(t) \end{aligned} \quad (3)$$

Table 1 shows some example matches produced by the three algorithms described above. While the syzygy and antinomy methods both work in a semantic manner, the clinamen function is purely syntactical, which becomes very obvious when seeing the different results side by side as in the table. Relying on WordNet's limited vocabulary means less matches can be found mostly because the text of *Faustroll* uses a very specific language and not always matches that found in the thesaurus. On the other hand, it illustrates the breadth of vocabulary used by Jarry in his writing nicely, as only those results are returned that actually appear in the book.

	clinamen	syzygy	antinomy
clear	altar, leaf, pleas, cellar	vanish, allow, bare, pronounce	opaque
solid	sound, valid, solar, slide	block, form, matter, crystal, powder	liquid, hollow
books	boot, bones, hooks, rocks, banks	dialogue, authority, record, fact	-
troll	grill, role, tell	wheel, roll, mouth, speak	-
live	love, lies, river, wave, size, bite	breathe, people, domicile, taste, see, be	recorded, dead

Table 1. Example search results. Queries are shown in column one, algorithms used in row one.

POSSIBLE APPLICATIONS

In this section we consider the possible uses and applications for the proposed creative search tool.

Our target audience is not quite as broad as that of a general search engine like Google. Instead, we aim to specifically cater for users who can appreciate creativity or users in need of creative inspiration. Users should generally be educated about the purpose of the search tool so they are not discouraged by what might appear to be *nonsensical* results. Potential users could include artists, writers or poets and anybody who is looking for out-of-the-box inspirations or a refreshingly different search engine to the norm.

Uses

There are many ways a pataphysical search tool could be used across disciplines.

In literature, for example, it could be used to write or generate poetry, practically or as a simple aid for inspiration. We are not limited to poetry either; novels, librettos or plays could benefit from such pataphysicalised inspirations. One can imagine tools using this technology that let you explore books in a different ordering of sentences (a sort of pataphysical journey of paragraph hopping), tools that re-write poems or mix and match them together. Even our simple prototype shows potential in this area and could be even more powerful if we extended it to include more source texts, for example the whole set of books contained in Faustroll's library ([20] and also [12]). A richer body of texts (by different authors) would also produce a much larger index which would then possibly find many more matches through WordNet and end in a more varied list of results.

From a computer science perspective this pataphysical approach could be added to the many algorithms used by traditional search engines for purposes like query feedback or expansion (e.g. "did you mean … "or "you might also be interested in … "). Depending on how creative we want the search engine to be, the higher we would rank the importance of this particular algorithm. One of the concepts related to the search tool, namely patadata, could have an impact on the development of the Semantic Web. Just as the Semantic Web is about organizing information semantically through objective metadata, patadata could be used to organize information pataphysically in a subjective way.

Our prototype tool is already being used in the creation of an online opera, provisionally entitled *from [place] to [place]*, created in collaboration with The Opera Group[1], an award-winning, nationally and internationally renowned opera company, specialising in commissioning and producing new operas. In particular, it is being used to create the libretto for one of the virtual islands whose navigation provides the central storyline for the opera. The opera will premiere in 2013, and will continue to develop thereafter, deploying new versions of the tool as they appear.

Evaluation

Evaluating creative software is not an easy task and there are no standard approaches. Pease and Colton [28] divide it into two notions:

[1] www.theoperagroup.co.uk

- whether an idea or artefact is valuable or not, and
- whether a system is acting creatively or not.

Following this approach, we would need to investigate each individual search result in terms of its value and creativity. This could be done by user ratings or satisfaction questionnaires. Rather than measuring the success of individual results though, we could also look at evaluation them as one set instead, similar to the blind side-by-side comparisons by the Bing search engine or the selective omission of a certain number of results by search engine MillionShort[2].

The way we display and label results produced by the tool can influence how the user perceives them. Our current prototype for example separates the results into its three components but we could have equally just mixed them all together. It not always clear how each result connects to the initial query, even if we identify through which algorithm a result has been obtained. These keywords (syzygy, clinamen and antinomy) might not be helpful to users unfamiliar with the concept of pataphysics anyway and might therefore appear rather nonsensical. Whilst there is a clear logic to each search result, they might appear anomalous to the user's expectations if he received these results without knowing the philosophy of the search tool. The results could possibly appear random then, and could therefore likely to be detrimental to the user. The level of interaction between the user and the system and the feedback the tool gives to the user on its internal processes will have a large influence on the overall effectiveness, perception and appreciation of the tool.

The less obvious the processes in the background are for the user, the more difficult it might be to appreciate the search results. On the other hand, too much transparency could spoil much of the experience. After all, explaining a joke kills it. The issue therefore becomes a question of finding just the right level of transparency to satisfy curious users while at the same time not spoiling the seemingly serendipitous experience of others.

FURTHER WORK AND CONCLUSION

We are just beginning to understand the potential of such a creative search tool and its implications. There is much research left to be done, specifically in developing more and different kinds of search algorithms and evaluating the results we obtain. We could try to implement different algorithms or different pataphysical concepts within our existing prototype or built a different system altogether. We could also try to implement a fully functioning Web search engine using the algorithms described in this paper and then compare the two different types of results. It could be interesting to investigate how users perceive and use search results produced in either the book based search or an open Web based search.

Before we go into further development and programming though, it might be worth studying, evaluating and interpreting the results produced by the prototype presented in this paper. An evaluation framework for pataphysical search results is under development. A study of user's reactions to the prototype could be very interesting as well and will be part of future work in this project.

Finally, to summarise, in this paper we have introduced a new approach for a creative search tool that uses pataphysics as an underlying philosophy. We have explained how pataphysics can be applied to search algorithms in order to produce interesting results with a humorous twist. Our initial experiments within a limited domain have shown that the generated results could indeed be interpreted as being novel, surprising and useful. We have also briefly discussed ideas for applications of the tool and issues that may trigger possible further research in in the field of creative computing. We have also presented some thoughts on evaluation of our tool and future work.

ACKNOWLEDGMENT

Jim Hendler at the Rennselaer Polytechnic Institute has provided valuable advice for this project.

REFERENCES

1. Baidu *The Baidu Story* [online] available at http://ir.baidu.com/phoenix.zhtml?c=188488&p=irol-homeprofile [Accessed 12 May 2013]
2. Boden, M.A. *The Creative Mind: Myths and Mechanisms*, London: Routledge, 2003.
3. Bök, C. *'Pataphysics: The Poetics of an Imaginary Science*, Evanston, Illinois: Northwestern University Press, 2002.
4. Borges, J.L. The Analytical Language of John Wilkins *in Selected Non-Fictions*, Weinberger, E. (Ed.), London: Penguin Books, 2000.
5. Brin, S.and Page, L. The anatomy of a large-scale hypertextual Web search engine in *Computer Networks and ISDN Systems*, 30(1-7), 1998, pp.107-117.
6. Brotchie, A., Chapman, A., Foulc, T. and Jackson, K. (Eds.), *'Pataphysics: Definitions and Citations*, London Institute of 'Pataphysics, London: AtlasPress, 2003.
7. Brotchie, A. *Alfred Jarry: A Pataphysical Life*, Cambridge, MA: MIT Press, 2011.
8. Cohen, H. *AARON* [online] available at http://www.kurzweilcyberart.com/aaron/history.html [Accessed 12 May 2013]
9. Colton, S. *The Painting Fool* [online] available at http://www.thepaintingfool.com/ [Accessed 12 May 2013]
10. Damerau, F.J. A Technique for Computer Detection and Correction of Spelling Errors in *Communications of the ACM*, 7(3), 1964, pp.171-176.

[2] www.bingiton.com and www.millionshortiton.com

11. de Vries, E. Browsing vs Searching, *OCTO report 93/02*, 1993.

12. Fisher, B. *The Pataphysician's Library*, Liverpool University Press, 2000.

13. Google *Company Overview* [online] available at https://www.google.com/intl/en/about/company/ [Accessed 12 May 2013]

14. Google *Ten things we know to be true* [online] available at https://www.google.com/intl/en/about/company/philosophy/ [Accessed 12 May 2013]

15. Hendler, J. and Hugill, A. The Syzygy Surfer : Creative Technology for the World Wide Web, in *Proc. WebSci 2011*, ACM Press (2011).

16. Hugill, A. *'Pataphysics: A Useless Guide*, Cambridge, MA: MIT Press, 2012.

17. Indurkhya, B. *Computers and Creativity*, Unpublished Manuscript, Based on the key note speech On Modeling Mechanisms of Creativity, delivered at Mind II: Computational Models of Creative Cognition, Dublin, Ireland, September 15-17, 1997.

18. International Conference of Computational Creativity [online] available at http://www.computationalcreativity.net/iccc2013/ [Accessed 12 May 2013]

19. International Obfuscated C Code Contest [online] available at http://www.ioccc.org/ [Accessed 12 May 2013]

20. Jarry, A. *Exploits and Opinions of Dr Faustroll, Pataphysician*, Cambridge, MA: Exact Change, 1997.

21. Kaufman, J. C., & Beghetto, R. A. (2009). Beyond big and little: The four c model of creativity. Review of General Psychology, 13(1), 1–12.

22. Koestler, A. *The Act of Creation*, London: Penguin Books, 1964.

23. Kurzweil, R. *Cybernetic Poet* [online] available at http://www.kurzweilcyberart.com/poetry/rkcp_overview.php [Accessed 12 May 2013]

24. Marchionini, G. From finding to understanding, *Communications of the ACM*, vol. 49, no. 4, pp. 41–46, 2006.

25. Marchionini, G. and Shneiderman, B. Finding facts vs. browsing knowledge in hypertext systems, *Computer*, vol. 21, no. 1, pp. 70–80, 1988.

26. Microsoft *Bing Fact Sheet* October 2012[online] available at http://www.microsoft.com/en-us/news/download/presskits/bing/docs/MSBingAll-UpFS.docx [Accessed 12 May 2013]

27. NLTK Project, 2012, *Natural Language Toolkit* [online] available at http://nltk.org/ [Accessed 12 May 2013].

28. Pease, A. and Colton, S. On impact and evaluation in Computational Creativity: A discussion of the Turing Test and an alternative proposal in *Proc. AISB 2011 Symposium on AI and Philosophy*.

29. Polya, G. *How To Solve It*, 2nd ed. Princeton, New Jersey: Princeton University Press, 1957.

30. Princeton University, 2010, *About WordNet* [online] available at http://wordnet.princeton.edu [Accessed 12 May 2013]

31. Sutcliffe, A. and Ennis, M, Towards a cognitive theory of information retrieval, *Interacting with Computers*, vol. 10, pp. 321–351, 1998.

32. YossarianLives *The Metaphorical Search Engine* [online] available at http://www.yossarianlives.com/ [Accessed 12 May 2013]

A Framework for Creativity in Search Results

James Sawle, Fania Raczinski, and Hongji Yang
Institute of Creative Technologies
De Montfort University
Leicester, UK
{jsawle, fania, hyang}@dmu.ac.uk

***Abstract*—Although trying to define creativity has been a hot area of research in many fields, the field of information retrieval has remained under developed. Over the report we attempt to define a structural definition of creativity which could be applied to search results in order to aid users in their creative endeavours. After defining creativity for search, we have then devised a simple metric based upon it, to show that there is a need for this research. The results, whilst positive, could be interpreted as a poor definition of creativity, and as such this is a sounding paper for future work.**

***Index Terms*—information retrieval; computational creativity**

I. Introduction

Over the past decade search has been focused on returning the smallest number of results which correlate to the user's information need. This has been a logical trend to pursue, as 92% of people use the internet as their first port of call when looking for everyday information [1].

However, this has meant that the creativity inspired by 'surfing the web' has over time slowly diminished. This research is not advocating the end of document based search; however, we propose that a new search engine architecture, which aims to inspire the creativity of its users, can only be beneficial to the landscape of the world wide web.

Over the course of the paper, we define what we mean by the creativity of a search result, with respect to a single result as well an entire set. The concepts presented in this paper, are inspired by 'Pataphysics, a pseudo-philosophy defined as "the science of imaginary solutions, which symbolically attributes the properties of objects, described by their virtuality, to their lineaments" [2].

The rest of the paper is organised as followed. Section 2 explores definitions of creativity from both computer science and psychology. In Section 3, we outline a general definition of creativity in search, which can be used to create a metric. Section 4 will see a simplistic metric to be used for the purpose of evaluating the concept as well as some experimental data.

II. Definition of Creative Search

A. A Framework to Base Creativity Upon

Creativity is a subjective topic, with different people defining the creative worth of a piece of information differently; however, Newell, Shaw and Simon [3] devised a definition based upon four criteria to categorise the creativity of a given solution or answer.

1) The answer is novel and useful (either for the individual or for society)
2) The answer demand that we reject ideas we had previously accepted
3) The answer results from intense motivation and persistence
4) The answer comes from clarifying a problem that was originally vague

Each of these criterion for creativity approach a definition from a different perspective. Whilst trying to relate this to information retrieval, it should be simple to see that criterion 1 relates to the goal of the search, whilst criterion 4 relates to the information need, or starting point. What may be less obvious however, is that criterion 3 relates to the scale of the search and hence the number of dead ends that may be encountered and that criterion 2 suggests which search paths should be avoided whilst looking for creative results.

Whilst this framework gives us a very high level definition of creativity, it is hard to apply it in its current form. Through applying some of the more prevalent techniques used in the field of computational creativity, we can attempt to reduce this down into a more precise definition.

B. P-Creativity and H-Creativity

Boden [4] defines that there are two forms of creativity, P-creativity and H-creativity. P-creativity or *'psychological'* creativity, is an idea or solution that is new to the person who came up with it. An idea that represents *'historical'* creativity, H-creativity, on the other hand, is one which has not been thought of by anybody before and can therefore be deemed a historical-sociological category [5]. H-creativity is subsequently a special case of P-creativity which many people consider to be the more important of the two, as this is what drives forward human knowledge.

When we relate these concepts to search results we end up with some interesting outcomes.

1) Single Search Result: A single search result is most likely to be P-creative or neither. This is because, for it to be H-creative, there must be some logic in the document that nobody else has noticed, or drawn the same conclusions from different information. For the single result to be neither P-creative or H-creative, the user must have a thorough understanding of the topic, and the result must add no new information.

2) Set of search results: A set of search results is most likely to be P-creative. It is highly unlikely that a user would

 ISBN: 978-1-61208-157-1

have explored every possible creative avenue over a set of results, unless the set is not of a trivial size. But by the same logic, if a large range of ideas are contained, it is unlikely that the set will be H-creative, as somebody is likely to have linked them together.

The question becomes, is there a link between maximising the chance of something being P-creative and H-creative or is the link more subtle. Or is it enough for a search engine to try and improve the chances of P-creativity for a user.

C. Exploratory and Transformational Creativity

Boden [4] goes on to define the concepts of exploratory and transformational creativity. She defines exploratory creativity to be the exploration of a space of partial and complete possibilities. This therefore suggests that there are rules that confine this space. If we were therefore to alter the rules that define the space, and subsequently alter the space that we are exploring, this is defined as transformational creativity [6].

Whilst this does give us a nice slant to look at creativity, comparing the trade-off of traditional problem spaces compared to augmented ones, this is very difficult to model, combined with the fact that the solutions found by tweaking the rules that confine the space can easily rule out the solution in the traditional space [7].

D. Bisociation

Bisociation makes a distinction between the routine skills of thinking on a single 'plane', and the creative act, which operates on more than one plane [8]. This means, with Koestler's definition, that we must define creativity as a set of results such that they are simultaneously associated with two habitually incomparable contexts.

It is clear to see how this model extends from that of Boden's theory of exploratory and transformational creativity. The fact that more than one 'plane' must be considered will force a transformational process to occur. However, unlike transformational creativity, both processes must be considered, the exploratory and transformational. Subsequently, we should not end up with a solution that can't exist within the rules defined by the original problem, even if we transcend into transformational creativity, as long as we finish the process in the plane that we started in.

E. Conceptual Blending

The idea of combining different thought processes, whilst more elegant than transformational creativity, does not give us a nice definition that applies to search results as well as tying in with our underlying philosophy. Conceptual blending is a step closer. This general theory of cognition, formally called Conceptual Integration Networks [9], allows us to look at a number of different dataspaces, and attempt to 'blend'/merge them in such a way that the new dataspace tries to simulate how we use large amounts of information and bring it together to form new ideas.

F. Combinatorial Creativity

Both of the above concepts fall into the general category of combinatorial creativity. This is a logical assumption of modelling creativity, as people tend to come up with solutions by first looking at new combinations of currently existing ideas. This therefore allows us to consider the idea of creativity as a search process through the space of all possible combinations, therefore this fits into the idea of search engines.

Whilst conceptual blending explores the idea of combining different thought processes and bisociation, looking at different planes of creative thought; let us consider the idea of placing the data itself into different concepts, enabling us to get the following areas of combinatorial creativity to explore with respect to creative search based upon philosophical.

- Placing a familiar object in an unfamiliar setting or placing an unfamiliar object into a familiar setting.
- Blending two superficially different objects or concepts
- Comparing a familiar object to a superficially unrelated and semantically distant concept
- Searching through a number of different concepts that are related to each other but could be considered as swerving away from the original concept. This is based upon Epicurus's theory of clinamen from his doctorine of atomism [10].

III. Definition

The above definitions, allow us to define creativity in search results with pre-existing concepts agreed by the academic community.

It is clear, that in the case of search results, we still have the issue of a group of results providing greater creative inspiration to one user than another. This tends to be a problem with most metrics, the problem of objectiveness vs subjectiveness. With subjectiveness being a quality that is important, it means that we have a problem getting repeatable results. We therefore need to build a definition that is as objective as possible, whilst not overlooking some of the dynamic properties that it may be possible to model.

At this stage it is important to stress that this is not an attempt to model the creative process, but to give a model for how useful a set of results might be in inspiring creativity.

A. A Single Result

It is intuitive for us to start with a single result. Whilst maximising the possibility for a single result being H-creative, it is very unlikely that this will be the case with a full set of results. The issue becomes, measuring how P-creative an individual result is to a search result.

It seems sensible to assume, that if a result has no relevance to the search request, then the result will have no chance of inspiring P-creativity. The more information about the search request a single result has, increases the chance of a result inspiring P-creativity, therefore using relevance metrics.

 ISBN: 978-1-61208-157-1

B. Set of Results

To improve the chances of inspiring creativity, a group of related results which discuss a number of different areas of the topic would logically improve the quality of the results. As stated, if we maximise the breadth of information of a single result it would improve creativity, we should therefore attempt to do the same across the entire set.

The issue however, is that the majority of users do not look past the top 10 search results [11]. Whilst this is unlikely to be the case for people using a search engine targeted at inspiring creativity, it does make sense to try to reduce the overall amount of data provided. We must therefore penalise repetition in the results provided, forcing a more diverse set of results.

This can be taken a step further by only considering a certain number of results and ignoring the ordering, because there is no simple way to define how ordering affects the creative process. With a lack of defined ordering, it means that having endless results would be tedious and counter-productive. Whilst we have no strong view on the exact number of results that should be considered, we believe that it should not be substantially greater than 10, for the reason discussed above.

The way that each result is provided to the user will affect how the user perceives the results. A diversity of different document types, e.g., text, images, sound, we believe would improve the quality of creativity inspiration.

C. Results as a Set of Sets

We could extend this concept to the next logical step of returning results as a set containing multiple related sets of results. In this analogy, each of the inner sets could relate to an individual concept related to the information need, and a clearer relationship between concepts, how they relate to each other and how the results represent the concept they are contained within would exist.

The question becomes how we measure the creative quality of this type of result. Due to the structure of the results, we can attempt to model the creativity in different levels allowing us to try and abstract the problem as much as possible.

Due to the fact that this is not a method that is currently used to return search results, we shall not explore it further at this point in time. However, we believe that this would be a logical way to return results in the future.

IV. Example Metric

As the above definition is meant as a guideline for defining creativity, this section attempts to give a real world example. The metric defined below is a contrived example to show how it could be applied with current search results.

A. Algebraic Definition

Taking the definition defined in Section III-B, we have derived the following abstract metric.

Let us define a query as q, a set of results as r and an individual result as d. As such $r = \{d_1, d_2, d_3, ..., d_i\}$ where i is the number of results examined.

For the quality of a single result, we shall define $P(q, d)$ as a measure between 0 and 1, where 1 is the optimal value.

To reduce the amount of data duplication in the returned results, we shall define $D(r)$ which has to return a value between 0 and 1, where 0 means that no data is duplicated.

Let us define $T(r)$ as a way to weigh the final outcome of the metric to ensure that a diverse set of document types are returned. This metric will return 1 if a satisfactory balance is returned, and 0 if only a single document type is returned.

We can therefore compile these measures into a single metric, the Search Creativity Metric or SCM:

$$SCM = T(r) \cdot \frac{1}{i} \sum_{j=1} P(q, d_i) \cdot (1 - D(r))$$

As such, this metric will always return a value between 0 and 1, with 1 being the optimal value.

B. Fleshing Out the Metric

To enable us to apply any experimental data to the metric, we must first give definitive definitions to each of the functions provided above, $P(q, d)$, $D(r)$ and $T(r)$.

1) D(r): As this measures the number of duplicate results in a return set, we can easily define it as the number of results that have a majority of information that is contained within another article. This allows the following definition

$$D(r) = \frac{Number of results with data in previous results}{Number of results}$$

As we relate each result to the previous results in the list, the results must always be $0 < D(r) \leq 1$. This makes sense, as even if all of the results are identical, there may still be some creative inspiration contained in the first result. This also allows us to penalise results heavily for leaning too much on one area of information.

2) T(r): As with $D(r)$, we need to define this measure so that we penalise for a lack of diversity, but do not eradicate all results, as this would not reflect the possible creative quality of the information returned.

For this definition, we will need to leverage on the definitions provided earlier. Let i is the number of results within the result set r. We can therefore define n to be the number of different result types that are returned, and σ to be the standard deviation of the number of results for each media type. It is interesting to note that $0 \leq \sigma < \frac{i}{2}$, such that $\sigma = \frac{i}{2}$ means that the results are biased to only one result.

$$T(r) = \begin{cases} 1 - \frac{2 \cdot \sigma}{i} & : n > 2 \\ 0.1 & : n \leq 2 \end{cases}$$

For the case of this sample measure, we have defined that for a result set to be considered to be broad enough, that it must contain at least 3 different media types. This measure has no empirical backing.

TABLE I: Example Results

URL	$P(q,d)$	Reason
www.unicorn-darts.com	0.3	Company called Unicorn due to the single point on a dart
en.wikipedia.org/ wiki/Unicorn	1.0	Contains mythology as well as related animals
www.unicorn-grocery.co.uk	0.0	No relation to unicorns
www.unicorn theatre.com	0.25	Uses the mythology of unicorns to draw children into theatre
http://katemckinnon .files.wordpress. com/2008/07	0.7	Image of a unicorn but purely as a distraction from the rest of the article
http://www.unicorn centre.co.uk/	0.9	Unicorn mythology about the soul applied to a spiritual ideal including image
http://31st-and-chi.blogspot.com/ 2010/07/bunch-of-pictures-of-unicorns.html	0.9	Large array of unicorn pictures. One is identical to result 5.
http://disgrasian .com/2010/09/ unicorns-really-do-exist-and-theyre-asian/	0.9	Picture of unicorn and asian 2 horn unicorn.
http://www.youtube .com/watch?v= Q5im0Ssyyus	0.4	Comedy cartoon video about unicorns.
http://www.youtube .com/watch?v= v25MaXwopNI	0.8	Music for a cartoon character.

3) P(q, d): With respect to the relevance of an individual result compared to the information need, there are a number of different methods that could be used. For example, keyword analysis in text documents and image recognition in images, it is clear that a separate method would be needed for each media type that is returned.

With this in mind, for the example below, the individual relevance of a given result will be manually determined and a brief explanation given. The focus will be more on the relevance of the result to the information need, with some weighting given if there is a creative link.

C. Experimental Data

To show this metric in practice we will need to get real world data about a topic. We have used Google to search for results on the following creative need - unicorns from Greek mythology.

The search term input into Google on Thursday 12th May was 'unicorn'. Below is a table of a url to each result, their assigned P rating and a brief description of the reason why. We have taken the top 10 results including the first 4 images and videos.

Due to the repeated result in result 4 and 6, $D(r) = 0.1$ and $T(r) = 1 - \frac{2 \cdot 0.94}{10} = 0.812$. If we then feed these results into the SCM metric we get.

$$SCM = 0.812 \cdot \frac{1}{10} \cdot 6.15 \cdot (1 - 0.1) = 0.449$$

A 0.449 result for us represents a set of results that contain some creative merit, but which also could be improved. This result could be enhanced, based on this metric, if four of the results were to have been replaced with more relevant results.

We still need to understand whether the low result is due to the fact that the results are not inspiring creativity as we presume, or that the definition that we have provided is not complete and that we need to extend it further. It is planned, that we take this research further to answer the question using in-depth empirical analysis.

V. Conclusion and Future Work

Over the course of the report, we have attempted to define what we mean by creativity with respect to search engine results using the concepts from computational creativity. The definition is focused more on the structure and relationship between the results returned than the content of the results themselves. This will allow us to define this separately after carrying out further experiments.

This is evident from the metric that we generated to show how the definition could be used. We believe that the low result shows that the return set does not have a high creative merit; however, more testing will be needed to check whether this is the case, or whether the definition needs to be redefined.

We believe, that whilst this paper has little empirical backing, it has highlighted a short fall in the information retrieval domain, namely that of creative search. Even from the simple test that was conducted, it is apparent, that even when we reach a metric for measuring the creative quality of results, a new form of search engine will be required to achieve top quality results consistently.

The next stage of the research will focus on applying what we have learnt and combine quantitive and qualitative analysis to try and develop a new metric with a strong empirical backing. This means that our definition of creativity will likely need to be adapted over time; however, this could allow us to develop a metric that evolves over time to adapt to what the users consider to be creative search.

References

[1] D. Fallows, "The Internet and daily life," *Pew Internet and American Life Project*, 2004.
[2] A. Jarry, *Selected works of Alfred Jarry*, 1st ed., R. Shattuck and S. W. Taylor, Eds. Grove Press, 1965.
[3] A. Newell, J. G. Shaw, and H. A. Simon, *The process of creative thinking*. New York: Atherton, 1963, pp. 63 – 119.
[4] M. Boden, *The Creative Mind: Myths and Mechanisms*, 2nd ed. Weidenfeld & Nicolson, 1992.
[5] A. Brannigan, *The Social Basis of Scientific Discoveries*. Cambridge University Press, 1981.
[6] G. A. Wiggins, "A preliminary framework for description, analysis and comparison of creative systems," *Knowledge-Based Systems*, vol. 19, pp. 449–458, 2006.
[7] M. A. Boden, *Computer Models of Creativity*. Cambridge University Press, 1999, pp. 351–372.
[8] A. Koestler, *The Act of Creation*. London: Hutchinson & Co, 1964.
[9] G. Fauconnier and M. Turner, "Conceptual integration networks," *Cognitive Science*, vol. 22, no. 2, pp. 133–187, 1998.
[10] Epicurus, *The Epicurus Reader*, B. Inwood and L. P. Gerson, Eds. Hackett, 1994.
[11] C. Silverstein, H. Marais, M. Henzinger, and M. Moricz, "Analysis of a very large web search engine query log," *SIGIR Forum*, vol. 33, no. 1, pp. 6–12, 1999.

References

Agichtein, Eugene, Eric Brill and Susan Dumais (2006). 'Improving web search ranking by incorporating user behavior information'. In: ***Proceedings of the 29th annual international ACM SIGIR conference on Research an Development in Information Retrieval***. Seattle, Washington, USA, pp. 19–26 (cit. on p. 83).

Wikipedia (n.d.). ***Alfred Jarry***. Wikipedia. URL: https://fr.wikipedia.org/wiki/Alfred%5C_Jarry (visited on 10/11/2016) (cit. on p. 253).

Google (n.d.). ***AlphaGo***. Deep Mind. Google. URL: https://deepmind.com/research/alphago/ (visited on 05/11/2016) (cit. on p. 219).

Amaral, Jose Nelson et al. (2006). ***About Computing Science Research Methodology***. University of Alberta. URL: https://webdocs.cs.ualberta.ca/~c603/readings/research-methods.pdf (visited on 15/11/2016) (cit. on pp. 23, 25).

Dictionary (2010). ***animal, n.*** Oxford English Dictionary. URL: http://www.oed.com/view/Entry/273779 (visited on 10/12/2015) (cit. on p. 15).

Getty (n.d.). ***API Overview***. Getty Images API. Getty. URL: http://developers.gettyimages.com/api/docs/v3/api-overview.html (visited on 07/08/2016) (cit. on pp. 160, 211).

Baeza-Yates, Ricardo and Berthier Ribeiro-Neto (2011). ***Modern Information Retrieval: The Concepts and Technology Behind Search***. Harlow, UK: Pearson Education Limited (cit. on pp. 72, 75–77, 82–84, 94, 95, 157).

Bao, Shenghua et al. (2007). 'Optimizing Web Search Using Social Annotations'. In: ***Proceedings of the International World Wide Web Conference***, pp. 501–510 (cit. on p. 83).

Barthes, Roland (1967). ***The Death of the Author***. Aspen 5+6. UbuWeb. URL: http://www.ubu.com/aspen/aspen5and6/threeEssays.html/#barthes (visited on 26/01/2016) (cit. on p. 128).

Basile, Jonathan (2015). ***The Library of Babel***. URL: https://libraryofbabel.info/ (visited on 10/12/2015) (cit. on p. 16).

Bastos Filho, Carmelo et al. (2008). 'A novel search algorithm based on fish school behavior'. In: ***Proceedings of the IEEE International Conference on Systems, Man and Cybernetics***, pp. 2646–2651 (cit. on p. 83).

Baudrillard, Jean (2007). ***Pataphysics***. Ed. by Arthur Kroker and Marilouise Kroker. Trans. by Drew Burk. CTHEORY. URL: http://www.ctheory.net/articles.aspx?id=569 (visited on 21/01/2012) (cit. on p. 42).

Beghetto, Ronald A. and James C. Kaufman (2007). 'Toward a Broader Conception of Creativity: A Case for 'mini-c' Creativity'. In: ***Psychology of Aesthetics, Creativity, and the Arts*** 1.2, pp. 73–79 (cit. on p. 55).

Bergerac, Cyrano de (2014). ***A Voyage to the Moon***. Ed. by Marc D'Hooghe. Trans. by Archibald Lovell. Project Gutenberg (cit. on p. 146).

Berners-Lee, Tim (1998). ***Web Architecture from 50,000 feet***. W3C. URL: http://www.w3.org/DesignIssues/Architecture.html (visited on 16/11/2016) (cit. on p. 239).

Bharat, Krishna and George Mihaila (2000). ***Hilltop: A Search Engine based on Expert Documents***. URL: ftp://ftp.cs.toronto.edu/csrg-technical-reports/405/hilltop.html (visited on 15/11/2016) (cit. on p. 82).

Dictionary (n.d.). ***Bias***. Cambridge Dictionary. URL: http://dictionary.cambridge.org/dictionary/english/bias (visited on 16/10/2016) (cit. on p. 221).

Microsoft (2012). ***Bing Fact Sheet***. Microsoft. URL: http://www.microsoft.com/en-us/news/download/presskits/bing/docs/MSBingAll-UpFS.docx (visited on 22/12/2012) (cit. on p. 72).

Microsoft (n.d.). ***Bing Image Search API***. Cognitive Services. Microsoft. URL: https://www.microsoft.com/cognitive-services/en-us/bing-image-search-api (visited on 26/11/2016) (cit. on p. 210).

BingAPI (2012). ***Bing Search API***. DataMarket. Microsoft. URL: http://datamarket.azure.com/dataset/bing/search#schema (visited on 07/08/2016) (cit. on pp. 160, 211).

Bird, Steven, Ewan Klein and Edward Loper (2009). ***Natural Language Processing with Python***. Sebasopol, CA: O'Reilly Media (cit. on p. 84).

Bloy, Léon (2011). ***Le Désespéré***. Project Gutenberg (cit. on p. 146).

Boden, Margaret (2003). ***The Creative Mind: Myths and Mechanisms***. London, UK: Routledge (cit. on pp. 53, 57–59, 80, 102, 112, 113, 117–119).

Bök, Christian (2002). ***'Pataphysics: The Poetics of an Imaginary Science***. Evanston, Illinois, USA: Northwestern University Press (cit. on pp. 6, 37, 43, 45, 151).

Borges, Jorge Luis (1964). ***Labyrinths - Selected Stories and Other Writings***. New York, USA: New Directions (cit. on p. 16).

– (2000). 'The Analytical Language of John Wilkins'. In: ***Selected Non-Fictions***. Ed. by Eliot Weinberger. London, UK: Penguin Books, pp. 229–232 (cit. on pp. 5, 12, 14, 49).

Bown, Oliver (2014). 'Empirically Grounding the Evaluation of Creative Systems: Incorporating Interaction Design'. In: ***Proceedings of the Fifth International Conference on Computational Creativity***. Ljubljana, Slovenia, pp. 112–119 (cit. on pp. 96, 104, 105).

– (2015). 'Attributing Creative Agency: Are we doing it right?' In: ***Proceedings of the Sixth International Conference on Computational Creativity***. Park City, Utah, USA, pp. 17–22 (cit. on p. 129).

Brin, Sergey and Larry Page (1998). 'The anatomy of a large-scale hypertextual Web search engine'. In: ***Computer Networks and ISDN Systems*** 30 (1-7 1998), pp. 107–117 (cit. on pp. 73, 81).

Brotchie, Alastair (2011). ***Alfred Jarry: A Pataphysical Life***. London, UK: MIT Press (cit. on pp. 37–39, 41, 44).

Brotchie, Alastair and Stanley Chapman, eds. (2007). ***Necrologies***. London, UK: Atlas Press (cit. on p. 41).

Brotchie, Alastair, Stanley Chapman et al., eds. (2003). ***'Pataphysics: Definitions and Citations***. London, UK: Atlas Press (cit. on pp. 36, 37).

Broukhis, Leo, Simon Cooper and Landon Curt Noll (n.d.). ***The International Obfuscated C Code Contest***. URL: http://www.ioccc.org/ (visited on 20/11/2016) (cit. on pp. 20, 63).

Brown, Mark (2011). ***Patrick Tresset's robots draw faces and doodle when bored***. Wired UK. URL: http://www.wired.co.uk/news/archive/2011-06/17/sketching-robots (visited on 24/01/2016) (cit. on p. 127).

Burdick, Anne et al. (2012). ***Digital Humanities***. Cambridge, MA, USA: MIT Press (cit. on pp. 26–28, 68, 69, 115, 116, 124, 241, 248).

Burnham, Douglas (n.d.). ***Immanuel Kant: Aesthetics***. Internet Encyclopedia of Philosophy. URL: http://www.iep.utm.edu/kantaest/ (visited on 15/11/2016) (cit. on p. 107).

Candy, Linda (2012). 'Evaluating Creativity'. In: ***Creativity and Rationale: Enhancing Human Experience by Design***. Ed. by J.M. Carroll. London, UK: Springer (cit. on pp. 52, 101, 132).

Candy, Linda and Ernest Edmonds, eds. (2011). ***Interacting: Art, Research and the Creative Practitioner***. Faringdon, UK: Libri Publishing (cit. on p. 125).

Chalmers, David (1996). ***The Conscious Mind. In Search of a Fundamental Theory***. Oxford University Press (cit. on pp. 125, 130).

Chatham, Chris (2007). ***10 Important Differences Between Brains and Computers***. Developing Intelligence. ScienceBlogs. URL: http://scienceblogs.com/developingintelligence/2007/03/27/why-the-brain-is-not-like-a-co/ (visited on 03/11/2016) (cit. on p. 214).

Christian, Peter (2016). ***The N+7 Machine***. URL: http://www.spoonbill.org/n+7/ (visited on 20/11/2016) (cit. on p. 17).

Clark, Sean (2014). ***IOCT PhD Showcase 2014***. Flickr. URL: https://www.flickr.com/photos/seancuttlefish/sets/72157646116801940/ (visited on 03/11/2016) (cit. on p. 182).

– (2015a). ***CAS Talk: IOCT - Fania Raczinski (2015)***. Vimeo. 2015. URL: https://vimeo.com/142947457 (visited on 03/11/2016) (cit. on p. 182).

– (2015b). ***IOCT Talks - Videos Now Available***. Phoenix | Interact Labs. 2015. URL: http://interactlabs.co.uk/news/2015/10/ioct-talks---videos-now-available (visited on 03/11/2016) (cit. on p. 182).

StackExchange (n.d.). ***Code Bowling***. Programming Puzzles & Code Golf. StackExchange. URL: http://codegolf.stackexchange.com/questions/tagged/code-bowling (visited on 20/11/2016) (cit. on p. 19).

Cohen, Harold (1999). ***Colouring Without Seeing: A Problem in Machine Creativity***. Kurzweil CyberArt Technologies. URL: http://www.kurzweilcyberart.com/aaron/pdf/colouringwithoutseeing.pdf (visited on 24/01/2016) (cit. on p. 128).

– (2007). ***Toward a Diaper-Free Autonomy***. Museum of Contemporary Art. aaronshome.com. URL: http://aaronshome.com/aaron/publications/index.html (visited on 15/08/2016) (cit. on p. 217).

Cohen, Paul (2016). ***Harold Cohen Obituary***. aaronshome.com. URL: http://aaronshome.com/aaron/publications/Harold-Cohen-Obituary-by-Paul-Cohen.pdf (visited on 15/08/2016) (cit. on p. 217).

Coleridge, Samuel Taylor (2013). ***The Rime of the Ancient Mariner***. Project Gutenberg (cit. on p. 146).

Colton, Simon (2008a). 'Computational Creativity'. In: ***AISB Quarterly*** (2008), pp. 6–7 (cit. on pp. 63, 99).

– (2008b). 'Creativity versus the perception of creativity in computational systems'. In: ***Proceedings of the 23rd AAAI Spring Symposium on Creative Intelligent Systems***. Chicago, Illinois, USA, 2008 (cit. on p. 99).

– (n.d.). ***The Painting Fool***. URL: http://www.thepaintingfool.com/ (visited on 22/11/2016) (cit. on p. 63).

Colton, Simon, Alison Pease and Graeme Ritchie (2001). 'The Effect of Input Knowledge on Creativity'. In: ***Proceedings of the 1st International Conference on Case-based Reasoning*** (cit. on p. 98).

Colton, Simon and Geraint Wiggins (2012). 'Computational Creativity: The Final Frontier?' In: ***Proceedings of the 20th European Conference on Artificial Intelligence***. Montpellier, France, pp. 21–26 (cit. on pp. 64, 65, 115, 116).

Baidu (n.d.). ***Company Overview***. About Baidu. Baidu. URL: http://ir.baidu.com/phoenix.zhtml?c=188488%5C&p=irol-homeprofile (visited on 22/12/2012) (cit. on p. 72).

Copeland, Jake and Jason Long (2016). ***Restoring the first recording of computer music***. Sound and vision blog. British Library. URL: http://blogs.bl.uk/sound-and-vision/2016/09/restoring-the-first-recording-of-computer-music.html (visited on 26/10/2016) (cit. on pp. 125, 127).

Corbyn, Zoë (2005). ***An introduction to 'Pataphysics***. The Guardian. URL: https://www.theguardian.com/culture/2005/dec/09/8 (cit. on p. 42).

Corliss, William (n.d.). ***The Sourcebook Project***. Science Frontiers. URL: http://www.science-frontiers.com/sourcebk.htm (visited on 22/11/2016) (cit. on p. 45).

Crawling (n.d.). ***Crawling and Indexing***. Inside Search. Google. URL: https://www.google.com/insidesearch/howsearchworks/crawling-indexing.html (visited on 04/08/2016) (cit. on p. 201).

Cruickshank, Douglas (2016). ***Why Anti-Matter Matters***. ralphmag.org. URL: http://www.ralphmag.org/jarry.html (visited on 15/11/2016) (cit. on p. 42).

Cutshall, James Anthony (1988). 'The Figure of the Writer - Alfred Jarry'. PhD Thesis. University of Reading (cit. on p. 41).

Damerau, Fred (1964). 'A Technique for Computer Detection and Correction of Spelling Errors'. In: ***Communications of the ACM*** 7.3, pp. 171–176 (cit. on pp. 86, 151).

DL Distance (n.d.). ***Damerau-Levenshtein distance***. Wikipedia. URL: https://en.wikipedia.org/wiki/Damerau-Levenshtein%5C_distance (visited on 23/10/2016) (cit. on p. 86).

Darien, George (2005). ***Le voleur***. Project Gutenberg (cit. on p. 146).

De Bra, Paul, Geert-jan Houben et al. (1994). 'Information Retrieval in Distributed Hypertexts'. In: ***Proceedings of the Intelligent Multimedia Information Retrieval Systems and Management Conference***. New York, USA (cit. on p. 83).

De Bra, Paul and Reinier Post (1994a). 'Information retrieval in the World-Wide Web: Making client-based searching feasible'. In: ***Computer Networks and ISDN Systems*** 27.2 (1994), pp. 183–192 (cit. on p. 83).

– (1994b). 'Searching for Arbitrary Information in the WWW: the Fish Search for Mosaic'. In: ***Mosaic: a journal for the Interdisciplinary Study of Literature*** (1994) (cit. on p. 83).

Dean, Jeffrey, Luiz Andre Barroso and Urs Hoelzle (2003). 'Web Search for a Planet: The Google Cluster Architecture'. In: ***IEEE Micro***, pp. 22–28 (cit. on p. 84).

Deerwester, Scott et al. (1990). 'Indexing by Latent Semantic Analysis'. In: ***Journal of the American Society for Information Science*** 41.6, pp. 391–407 (cit. on p. 79).

Dennis, Andrew (2016a). 'Investigation of a patadata-based ontology for text based search and replacement'. BA Thesis. University of London, 2016 (cit. on pp. 174, 177).

– (2016b). ***PataLib a Pataphysical toolkit for Python***. GitHub. 2016. URL: https://andydennis.github.io/patalib/ (visited on 02/11/2016) (cit. on p. 177).

Copyright (2012). ***Derivative works***. Factsheet No. P-22. UK Copyright Service. URL: http://www.copyrightservice.co.uk/copyright/p22%5C_derivative%5C_works (visited on 01/11/2016) (cit. on p. 148).

Desbordes-Valmore, Marceline (2004). ***Le Livre des Mères et des Enfants***. Project Gutenberg (cit. on p. 147).

Wordnik (2016). ***developer.wordnik.com***. URL: http://developer.wordnik.com/docs.html#!/word/getTextPronunciations%5C_get%5C_5 (visited on 05/11/2016) (cit. on p. 228).

Dijkstra, Edsger (1988). ***On the Cruelty of Really Teaching Computing Science***. URL: http://www.cs.utexas.edu/users/EWD/transcriptions/EWD10xx/EWD1036.html (visited on 17/01/2016) (cit. on pp. 131, 151).

Ding, Li et al. (2004). 'Swoogle: A semantic web search and metadata engine'. In: ***Proceedings of the 13th ACM Conference on Information and Knowledge Management*** (cit. on p. 83).

Drucker, Johanna (2009). ***SpecLab: Digital Aesthetics and Projects in Speculative Computing***. University of Chicago Press (cit. on pp. 67, 68, 120, 121).

Drucker, Johanna and B. Nowviskie (2007). 'Speculative Computing: Aesthetic Provocations in Humanities Computing'. In: ***A Companion to Digitial Humanities***. Ed. by Susan Schreibman, John Unsworth and Ray Siemens. Oxford: Blackwell Publishing. Chap. 29 (cit. on p. 68).

Du, Zhi-Qiang et al. (2007). 'The Research of the Semantic Search Engine Based on the Ontology'. In: ***International Conference on Wireless Communications, Networking and Mobile Computing***, pp. 5398–5401 (cit. on p. 83).

Dubbelboer, Marieke (2009). ''UBUSING' CULTURE'. PhD Thesis. Rijksuniversiteit Groningen (cit. on pp. 41, 42).

Eden, Amnon (2007). 'Three Paradigms of Computer Science'. In: ***Minds and Machines*** 17.2, pp. 135–167 (cit. on pp. 64, 65).

Edmonds, Ernest and Linda Candy (2010). 'Relating Theory, Practice and Evaluation in Practitioner Research'. In: ***Leonardo*** 43.5, pp. 470–476 (cit. on p. 30).

Efron, Bradley and Ronald Thisted (1976). 'Estimating the number of unseen species: How many words did Shakespeare know?' In: ***Biometrika*** 63.3, pp. 435–447 (cit. on p. 198).

Elskamp, Max (1898). ***Enluminures***. Ed. by Paul Lacomblez. Wikisource (cit. on p. 147).

Elton, Matthew (1995). 'Artificial Creativity: Enculturing Computers'. In: ***Leonardo*** 28.3, pp. 207–213 (cit. on pp. 62, 241).

Evans, Clark (2016). ***YAML 1.2***. YAML: YAML Ain't Markup Language. URL: http://yaml.org/ (visited on 02/11/2016) (cit. on p. 177).

Fingas, John (2016). ***IBM's Watson AI saved a woman from leukemia***. Engadget UK. URL: https://www.engadget.com/2016/08/07/ibms-watson-ai-saved-a-woman-from-leukemia/ (visited on 05/11/2016) (cit. on p. 218).

Flickr (n.d.). ***flickr.photo.search***. The App Garden. Flickr. URL: https://www.flickr.com/services/api/flickr.photos.search.html (visited on 07/08/2016) (cit. on pp. 144, 160, 208).

Florian, Jean Pierre Claris de (2012). ***Les deux billets***. Ed. by Paul Fiervre. Théâtre Classique (cit. on p. 147).

Foucault, Michel (1966). 'The Order of Things - Preface'. In: ***The Order of Things***. France: Editions Gallimard. Chap. Preface, pp. xv–xxiv (cit. on pp. 49, 185).

França, Celso et al. (2016). 'Regent-Dependent Creativity: A Domain Independent Metric for the Assessment of Creative Artifacts'. In: ***Proceedings of the Seventh International Conference on Computational Creativity***, pp. 68–75 (cit. on pp. 105, 106).

Gutenberg (2016). ***Free ebooks***. Project Gutenberg. URL: https://www.gutenberg.org/ (visited on 01/11/2016) (cit. on p. 147).

Fry, Ben and Reas Casey (n.d.). ***Processing***. URL: https://processing.org/ (visited on 23/11/2016) (cit. on p. 124).

Fu, Haohuan et al. (2016). 'The Sunway TaihuLight supercomputer: system and applications'. In: ***Science China Information Sciences*** 59.7, pp. 1–16 (cit. on pp. 214, 215).

Gelernter, David (1994). ***The Muse in the Machine***. London: Fourth Estate Limited (cit. on pp. 60, 61).

FlickrAPI (n.d.). ***Getting Started***. The Flickr Developer Guide: API. Flickr. URL: https://www.flickr.com/services/developer/api/ (visited on 07/08/2016) (cit. on pp. 160, 211).

Gibbs, Samuel (2016). ***Microsoft's racist chatbot returns with drug-smoking Twitter meltdown***. The Guardian. URL: https://www.theguardian.com/technology/2016/mar/30/microsoft-racist-sexist-chatbot-twitter-drugs (visited on 05/11/2016) (cit. on p. 219).

Git (2016). ***Git - fast, scalable, distributed revision control system***. git-/README.md. GitHub. URL: https://github.com/git/git/blob/master/README.md (visited on 05/11/2016) (cit. on pp. 144, 224).

GitHub (2016). URL: https://github.com/ (visited on 05/11/2016) (cit. on p. 224).

Google (2016). ***Glitch Art***. Image Search Results. Google. URL: https://goo.gl/waiqKV (visited on 20/11/2016) (cit. on p. 20).

Reddit (n.d.). ***glitch_art***. subreddit. Reddit. URL: https://www.reddit.com/r/glitch%5C_art/ (visited on 20/11/2016) (cit. on p. 20).

Glover, E.J. et al. (2001). 'Improving category specific Web search by learning query modifications'. In: ***Proceedings 2001 Symposium on Applications and the Internet***, pp. 23–32 (cit. on p. 83).

IBM (n.d.). ***Go beyond artificial intelligence with Watson***. IBM. URL: http://www.ibm.com/watson/ (visited on 05/11/2016) (cit. on p. 218).

Google (2012). ***Google Ranking***. Google. URL: https://www.google.com/intl/en/about/company/philosophy/ (visited on 26/10/2016) (cit. on pp. 72, 73, 122).

Google (n.d.). ***Googlebot***. Search Console Help. Google. URL: https://support.google.com/webmasters/answer/182072 (visited on 15/10/2016) (cit. on p. 73).

Goren, Uri (2011). ***Most artistic 2011***. IOCCC. URL: http://www.ioccc.org/2011/goren/hint.html (visited on 20/11/2016) (cit. on p. 20).

Grabbe, Christian Dietrich (1995). ***Scherz, Satire, Ironie und tiefere Bedeutung***. Gutenberg DE Spiegel (cit. on p. 147).

Gray, Carole and Julian Malins (2004). ***Visualizing research: a guide to the research process in art and design***. New York, USA: Routledge (cit. on pp. 25–27).

Gunicorn (n.d.). ***Gunicorn: Python WSGI HTTP Server for UNIX***. URL: http://gunicorn.org/ (visited on 31/10/2016) (cit. on p. 146).

Gyongyi, Zoltan, Hector Garcia-Molina and Jan Pedersen (2004). ***Combating Web Spam with TrustRank***. Technical Report. Stanford InfoLab (cit. on p. 83).

V & A (2016). ***Harold Cohen***. Search the Collections. Victoria and Albert Museum. URL: http://collections.vam.ac.uk/name/cohen-harold/6433/ (visited on 05/11/2016) (cit. on p. 217).

Hassabis, Demis (2016). ***AlphaGo: using machine learning to master the ancient game of Go***. Google Blog. URL: https://blog.google/topics/machine-learning/alphago-machine-learning-game-go/ (visited on 05/11/2016) (cit. on p. 219).

Haveliwala, Taher H. (2003). 'Topic-Sensitive PageRank: A Context Sensitive Ranking Algorithm for Web Search'. In: ***IEEE Transactions on Knowledge and Data Engineering*** 15.4, pp. 784–796 (cit. on p. 83).

Heilman, Kenneth M, Stephen E Nadeau and David O Beversdorf (2003). 'Creative innovation: possible brain mechanisms.' In: ***Neurocase*** 9.5, pp. 369–79 (cit. on p. 118).

Heisenberg, Werner (1942). ***Ordnung der Wirklichkeit***. Trans. by M.B. Rumscheidt and N. Lukens. URL: http://werner-heisenberg.unh.edu/t-OdW-english.htm#seg42 (visited on 22/12/2015) (cit. on p. 248).

Hendler, Jim and Andrew Hugill (2011). 'The Syzygy Surfer: Creative Technology for the World Wide Web'. In: ***Proceedings of the 3rd International Conference on Web Science***. Koblenz, Germany (cit. on pp. 12, 15, 107, 163).

– (2013). 'The Syzygy Surfer: (Ab)using the semantic web to inspire creativity'. In: ***International Journal of Creative Computing*** 1.1, pp. 20–34 (cit. on pp. 12, 15, 117).

Hersovici, M et al. (1998). 'The shark-search algorithm. An application: tailored Web site mapping'. In: ***Computer Networks and ISDN Systems*** 30.1-7, pp. 317–326 (cit. on p. 83).

Hofstadter, Douglas (1981). 'A Conversation with Einstein's Brain'. In: ***The Mind's I***. Ed. by Douglas Hofstadter and Daniel Dennett. Basic Books. Chap. 26, pp. 430–460 (cit. on p. 130).

Holz, Hilary J et al. (2006). 'Research Methods in Computing: What are they, and how should we teach them ?' In: ***Innovation and technology in computer science education***, pp. 96–114 (cit. on pp. 23, 24).

Homer (1999). ***The Odyssey***. Trans. by Samuel Butler. Project Gutenberg (cit. on p. 147).

Homer, Michael (2009). ***Python Damerau-Levenshtein distance implementation***. Internet Archive Wayback Machine. URL: https://web.archive.org/web/20100602093104/http://mwh.geek.nz/2009/04/26/python-damerau-levenshtein-distance/ (visited on 31/10/2016) (cit. on p. 152).

Dyn Web (n.d.). ***Horizontal Scrolling with JavaScript***. Dynamic Web Coding. URL: http://www.dyn-web.com/code/scroll/horiz.php (visited on 01/11/2016) (cit. on p. 165).

Horn, Robert (2009). 'The Turing Test: Mapping and Navigating the Debate'. In: ***Parsing the Turing Test***. Ed. by Robert Epstein, Gary Roberts and Grace Beber. Springer. Chap. 5, pp. 73–88 (cit. on pp. 211, 212, 252).

Hotho, Andreas et al. (2006). 'Information retrieval in folksonomies: Search and ranking'. In: ***Proceedings of the 3rd European Conference on The Semantic Web: Research and Applications***, pp. 411–426 (cit. on p. 83).

Stanford (n.d.). ***How is humanities research conducted?*** Stanford Humanities Center: Home of the Human Experience. URL: http://shc.stanford.edu/how-humanities-research-conducted (visited on 06/11/2016) (cit. on p. 25).

Hugill, Andrew (2012). ***'Pataphysics: A Useless Guide***. Cambridge, MA, USA: MIT Press (cit. on pp. 36, 37, 44, 45, 49).

Hugill, Andrew and Lee Scott (2013). 'The Imaginary Voyage: an online opera'. In: ***Digital Creativity*** 24.3, pp. 268–273 (cit. on pp. 179, 180).

– (2014a). ***Amorphous Isle***. The Imaginary Voyage (an online opera). 2014. URL: http://theimaginaryvoyage.com/Islands/Amorphous/amorphous%5C_isle%5C_high.php (visited on 02/11/2016) (cit. on p. 179).

Hugill, Andrew and Lee Scott (2014b). ***The Imaginary Voyage (an online opera)***. 2014. URL: http://www.theimaginaryvoyage.com/ (visited on 02/11/2016) (cit. on p. 179).

Hugill, Andrew and Hongji Yang (2013). 'The creative turn: new challenges for computing'. In: ***International Journal of Creative Computing*** 1.1, pp. 4–19 (cit. on pp. 7, 30, 64–66, 115, 116).

Hugill, Andrew, Hongji Yang et al. (2013). 'The pataphysics of creativity: developing a tool for creative search'. In: ***Digital Creativity*** 24.3, pp. 237–251 (cit. on pp. 119, 174, 181).

Hunt, Elle (2016). ***Tay, Microsoft's AI chatbot, gets a crash course in racism from Twitter***. The Guardian. URL: https://www.theguardian.com/technology/2016/mar/24/tay-microsofts-ai-chatbot-gets-a-crash-course-in-racism-from-twitter (visited on 05/11/2016) (cit. on p. 218).

ICCC (2014). ***ICCC 2014***. The Conference. URL: http://computationalcreativity.net/iccc2014/ (visited on 11/10/2016) (cit. on p. 64).

Bing (n.d.). ***Image Search API Reference***. Microsoft Developer Network (MSDN). Microsoft. URL: https://msdn.microsoft.com/en-us/library/dn760791.aspx (visited on 07/08/2016) (cit. on pp. 144, 160, 210).

Google (2015). ***Improve Server Response Time***. PageSpeed Tools > Insights. Google. URL: https://developers.google.com/speed/docs/insights/Server (visited on 25/11/2016) (cit. on p. 199).

Indurkhya, Bipin (1997). 'Computers and creativity'. Unpublished manuscript. Based on the keynote speech 'On Modeling Mechanisms of Creativity' delivered at Mind II: Computational Models of Creative Cognition (cit. on pp. 58, 119, 127).

ICCC (n.d.). ***International Conference on Computational Creativity***. URL: http://computationalcreativity.net/home/ (visited on 22/11/2016) (cit. on p. 63).

IJCrC (n.d.). ***International Journal of Creative Computing***. InderScience Publishers. URL: http://www.inderscience.com/jhome.php?jcode=ijcrc (visited on 22/11/2016) (cit. on p. 62).

ISCC (n.d.). ***International Symposium on Creative Computing***. IEEE. URL: https://iscc.gwasd.com/ (visited on 22/11/2016) (cit. on p. 62).

JSON (n.d.). ***Introducing JSON***. ECMA-404 The JSON Data Interchange Standard. json.org. URL: http://www.json.org/ (visited on 31/10/2016) (cit. on p. 160).

Jabr, Ferris (2012). ***Does Thinking Really Hard Burn More Calories?*** Mind. Scientific American. URL: https://www.scientificamerican.com/article/thinking-hard-calories/ (visited on 04/11/2016) (cit. on p. 215).

Jarry, Alfred (1996). ***Exploits and Opinions of Dr. Faustroll, Pataphysician***. Cambridge, MA, USA: Exact Change (cit. on pp. v, 5, 12, 13, 36, 42, 45, 46, 73, 89, 146, 151, 179).

– (2005). ***Ubu Roi ou les Polonais***. Project Gutenberg (cit. on p. 147).

– (2006). ***Collected Works II - Three Early Novels***. Ed. by Alastair Brotchie and Paul Edwards. London, UK: Atlas Press (cit. on pp. 107, 185).

JBlum (2007). ***pataphysics***. Urban Dictionary. URL: http://www.urbandictionary.com/define.php?term=pataphysics (visited on 10/10/2016) (cit. on pp. 43, 49).

Jeh, Glen and Jennifer Widom (2002). 'SimRank: A Measure of Structural Context Similarity'. In: ***Proceedings of the 8th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining***. Alberta, Canada, pp. 538–543 (cit. on p. 83).

Jordanous, Anna Katerina (2011). 'Evaluating Evaluation : Assessing Progress in Computational Creativity Research'. In: ***Proceedings of the Second International Conference on Computational Creativity*** (cit. on pp. 100, 131).

– (2012). 'Evaluating Computational Creativity: A Standardised Procedure for Evaluating Creative Systems and its Application'. PhD thesis. University of Sussex (cit. on pp. 100, 101, 131, 132).

– (2014). 'Stepping Back to Progress Forwards: Setting Standards for Meta-Evaluation of Computational Creativity'. In: ***Proceedings of the Fifth International Conference on Computational Creativity***, pp. 129–136 (cit. on p. 101).

Jordanous, Anna Katerina and Bill Keller (2012). 'Weaving creativity into the Semantic Web: a language-processing approach'. In: ***Proceedings of the 3rd International Conference on Computational Creativity***, pp. 216–220 (cit. on pp. 53, 99).

Jorn, Asger (1961). 'Pataphysics - A Religion In The Making'. In: ***Internationale Situationniste*** 6 (cit. on p. 42).

Jurafsky, Daniel and James H Martin (2009). ***Speech and Language Processing***. London: Pearson Education (cit. on pp. 84, 87–90).

Kahn, Gustave (n.d.). ***Le contre de l'or er du silence***. Archive.org (cit. on p. 147).

Kamps, Jaap, Rianne Kaptein and Marijn Koolen (2010). ***Using Anchor Text , Spam Filtering and Wikipedia for Web Search and Entity Ranking***. Tech. rep. (cit. on p. 83).

Kaufman, James C. and Ronald A. Beghetto (2009). 'Beyond big and little: The four c model of creativity'. In: ***Review of General Psychology*** 13.1, pp. 1–12 (cit. on pp. 55, 112, 113).

Kazjon, Grace and Mary Lou Maher (2013). 'What to expect when you're expecting: The role of unexpectedness in computationally evaluating creativity'.

In: ***Proceedings of the Fifth International Conference on Computational Creativity***, pp. 120–128 (cit. on pp. 105, 131).
Kim, Youjeong and S. Shyam Sundar (2012). 'Anthropomorphism of computers: Is it mindful or mindless?' In: ***Computers in Human Behavior*** 28.1, pp. 241–250 (cit. on p. 127).
Kleinberg, Jon M. (1999). 'Authoritative sources in a hyperlinked environment'. In: ***Journal of the ACM*** 46.5, pp. 604–632 (cit. on pp. 82, 121).
Kleinberg, Jon M. et al. (1999). 'The Web as a graph: measurements, models and methods'. In: ***Proceedings of the 5th International Conference on Computing and Combinatorics***. Tokyo, Japan (cit. on p. 82).
Knuth, Donald (n.d.). ***Help Wanted***. Stanford University. URL: http://www-cs-faculty.stanford.edu/~uno/help.html (visited on 20/11/2016) (cit. on p. 20).
Koestler, Arthur (1964). ***The Act of Creation***. London, UK: Hutchinson and Co (cit. on pp. 53, 56, 119).
Kurzweil, Ray (2013). ***How to Create a Mind***. London: Duckworth Overlook (cit. on pp. 214, 215).
Lamb, Carolyn, Daniel Brown and Charles Clarke (2015). 'Human Competence in Creativity Evaluation'. In: ***Proceedings of the Sixth International Conference on Computational Creativity***, pp. 102–109 (cit. on p. 97).
Lang, Andrew, ed. (2008). ***The Arabian Nights Entertainments***. Project Gutenberg (cit. on p. 147).
Lautréamont, Comte de (2011). ***Les Chants de Maldoror***. Project Gutenberg (cit. on p. 147).
Learn about AARON's history (n.d.). Kurzweil CyberArt Technologies. URL: http://www.kurzweilcyberart.com/aaron/history.html (visited on 22/11/2016) (cit. on p. 63).
Leary, Timothy (1964). 'The effects of test score feedback on creative performance and of drugs on creative experience'. In: ***Widening Horizons in Creativity***. Ed. by Taylor. New York: Wiley, pp. 94–96 (cit. on p. 56).
Levenshtein, Vladimir I (1966). 'Binary codes capable of correcting deletions, insertions, and reversals '. In: ***Soviet Physics Doklady*** 10.8, pp. 707–710 (cit. on pp. 86, 151).
Liapis, Antonios et al. (2013). 'Transforming Exploratory Creativity with DeLeNoX'. In: ***Proceedings of the Fourth International Conference on Computational Creativity***, pp. 56–63 (cit. on p. 47).
Lidwell, William, Kritina Holden and Jill Butler (2010). ***Universal Principles of Design***. Rockport Publishers (cit. on p. 222).
Lieberman, Zach, Theodore Watson and Arturo Castro (n.d.). ***openFrameworks***. an open source C++ toolkit for creative coding. URL: http://openframeworks.cc/ (visited on 23/11/2016) (cit. on p. 124).

Luo, Fang-fang, Guo-long Chen and Wen-zhong Guo (2005). 'An Improved 'Fish-search' Algorithm for Information Retrieval'. In: ***International Conference on Natural Language Processing and Knowledge Engineering***, pp. 523–528 (cit. on p. 83).

Macdonald, Craig (2009). 'The Voting Model for People Search'. In: ***ACM SIGIR Forum*** 43 (1 2009) (cit. on p. 79).

Maeda, John (2001). ***Design by Numbers***. MIT Press (cit. on p. 125).

Maeterlinck, Maurice (1918). ***Aglavaine and Selysette***. Trans. by Alfred Sutro. Archive.org (cit. on p. 147).

Maher, Mary Lou, Katherine Brady and Douglas Fisher (2013). 'Computational Models of Surprise in Evaluating Creative Design'. In: ***Proceedings of the Fourth International Conference on Computational Creativity***, pp. 147–151 (cit. on pp. 104, 131).

Mahogany (n.d.). ***Mahogany Opera Group***. URL: http://www.mahoganyoperagroup.co.uk/ (visited on 02/11/2016) (cit. on p. 179).

Malins, Julian and Carole Gray (1995). 'Appropriate research methodologies for artists, designers and craftspersons: research as a learning process'. In: ***'Making It': UK Crafts Council Conference***. Invited paper (cit. on p. 25).

Mallarmé, Stéphane (2003). ***Pages***. Project Gutenberg (cit. on p. 147).

Manning, Christopher, Prabhakar Raghavan and Hinrich Schuetze (2009). ***Introduction to Information Retrieval***. Cambridge University Press (cit. on p. 85).

Marchionini, Gary (2006). 'From finding to understanding'. In: ***Communications of the ACM*** 49.4, pp. 41–46 (cit. on pp. 80, 223).

Marchionini, Gary and Ben Shneiderman (1988). 'Finding facts vs. browsing knowledge in hypertext systems'. In: ***Computer*** 21.1, pp. 70–80 (cit. on pp. 73, 79).

Marcus, Mitchell P, Beatrice Santorini and Mary Ann Marcinkiewicz (1993). 'Building a Large Annotated Corpus of English: The Penn Treebank'. In: ***Computational Linguistics*** 19.2 (cit. on pp. 89, 250).

Matarasso, François (1997). ***Use or Ornament? The Social Impact of Participation in the Arts***. Comedia (cit. on p. 241).

Mathews, Harry and Alastair Brotchie (2005). ***Oulipo Compendium***. London, UK: Atlas Press (cit. on pp. 17, 18, 46, 218).

Mayer, Richard E (1999). 'Fifty Years of Creativity Research'. In: ***Handbook of Creativity***. Ed. by Robert J Sternberg. New York: Cambridge University Press. Chap. 22, pp. 449–460 (cit. on pp. 52, 53, 96, 101, 132, 217).

Mayhaymate (2012). ***File:PageRank-hi-res.png***. Wikimedia Commons. URL: https://commons.wikimedia.org/wiki/File:PageRank-hi-res.png (visited on 18/10/2016) (cit. on p. 82).

McBride, Neil (2012). 'A Robot Ethics: The EPSRC Principles and the Ethical Gap'. In: ***AISB / IACAP World Congress 2012 Framework for Responsible Research and Innovation in AI***. July, pp. 10–15 (cit. on pp. 126–128).
McDonald, Keith (2016). ***A Return to Machine Learning***. Medium.com. URL: https://medium.com/@kcimc/a-return-to-machine-learning-2de3728558eb#.662a854dl (visited on 11/11/2016) (cit. on p. 241).
McGregor, Stephen, Geraint Wiggins and Matthew Purver (2014). 'Computational Creativity: A Philosophical Approach, and an Approach to Philosophy'. In: ***Proceedings of the Fifth International Conference on Computational Creativity***, pp. 254–262 (cit. on p. 105).
Bing Crawlers (n.d.). ***Meet our crawlers***. Webmaster help and how-to. Microsoft. URL: https://www.bing.com/webmaster/help/which-crawlers-does-bing-use-8c184ec0 (visited on 15/10/2016) (cit. on p. 73).
Menabrea, L. F. and Ada Lovelace (1842). 'Sketch of The Analytical Engine, Invented by Charles Babbage'. In: ***Bibliotheque Universelle de Geneve*** 82 (cit. on pp. 6, 212).
Mendès, Catulle (1910). ***The Mirror***. Ed. by Fancis J. Reynolds. Project Gutenberg (cit. on p. 147).
– (2013). ***La Divina Aventura***. Ed. by E Dentu (cit. on p. 147).
Michelsen, Maria Hagsten and Ole Bjorn Michelsen (2016). ***Regex Crossword***. RegexCrossword.com. URL: http://regexcrossword.com/ (visited on 19/10/2016) (cit. on p. 86).
Tay.ai (2016). ***Microsoft: About Tay and Privacy***. Internet Archive Wayback Machine. URL: https://web.archive.org/web/20160414074049/https://www.tay.ai/ (visited on 05/11/2016) (cit. on p. 218).
Translator (2011). ***Microsoft Translator - Text Translation***. DataMarket. Microsoft. URL: https://datamarket.azure.com/dataset/bing/microsofttranslator (visited on 07/08/2016) (cit. on pp. 144, 158, 160, 211).
Miller, George A. (1995). 'WordNet: a lexical database for English'. In: ***Communications of the ACM*** 38.11, pp. 39–41 (cit. on pp. 85, 154).
Minsky, Marvin (1980). 'K-Lines : A Theory of Memory'. In: ***Cognitive Science*** 33.4, pp. 117–133 (cit. on pp. 60, 61).
– (1988). ***The Society of Mind***. Simon and Schuster, p. 336 (cit. on pp. 60, 61).
Miyamoto, Sadaaki (1990a). ***Fuzzy Sets in Information Retrieval and Cluster Analysis***. Theory and Decision Library. Springer, 1990 (cit. on p. 79).
– (1990b). 'Information Retrieval based on Fuzzy Associations'. In: ***Fuzzy Sets and Systems - On fuzzy Information and Database Systems*** 38 (2 1990), pp. 191–205 (cit. on p. 79).
Miyamoto, Sadaaki and K. Nakayama (1986). 'Fuzzy Information Retrieval Based on a Fuzzy Pseudothesaurus'. In: ***IEEE Transactions on Systems, Man and Cybernetics*** 16.2, pp. 278–282 (cit. on p. 79).

Motte, Warren (2007). ***Oulipo, A primer of potential literature***. London: Dalkey Archive Press (cit. on pp. 46, 47, 117).

Mumford, Martin and Dan Ventura (2015). 'The man behind the curtain: Overcoming skepticism about creative computing'. In: ***Proceedings of the Sixth International Conference on Computational Creativity***, pp. 1–7 (cit. on p. 129).

Munroe, Randall (2015). ***Watson Medical Algorithm***. XKCD. URL: https://xkcd.com/1619/ (visited on 05/11/2016) (cit. on p. 218).

Musée Patamécanique (2016). private communication. 13th Oct. 2016 (cit. on p. 183).

NLTK (n.d.). ***Natural Language Toolkit***. NLTK 3.0 documentation. NLTK Project. URL: http://www.nltk.org/ (visited on 18/10/2016) (cit. on pp. 84, 144, 149, 154).

Neeley, J. Paul (2015). ***Introducing the NEW Yossarian***. email communication. 9th Dec. 2015 (cit. on pp. 15, 16).

Negrete-Yankelevich, Santiago and Nora Morales-Zaragoza (2014). 'The apprentice framework: planning and assessing creativity'. In: ***Proceedings of the Fifth International Conference on Computational Creativity***, pp. 280–283 (cit. on p. 105).

Newell, A, J. G. Shaw and H. A. Simon (1963). ***The Process Of Creative Thinking***. New York: Atherton (cit. on p. 118).

Nick, Z.Z. and P. Themis (2001). 'Web Search Using a Genetic Algorithm'. In: ***IEEE Internet Computing*** 5.2, pp. 18–26 (cit. on p. 83).

Nicole (2010). ***The 10 Most Incredible Google Bombs***. searchenginepeople.com. URL: http://www.searchenginepeople.com/blog/incredible-google-bombs.html (visited on 18/10/2016) (cit. on p. 81).

Nicolescu, Basarab (2010). 'Methodology of Transdisciplinarity - Levels of Reality, Logic of the Included'. In: ***Transcdisciplinary Journal of Engineering and Science*** 1.1, pp. 19–38 (cit. on pp. 28, 29).

Norton, David, Derrall Heath and Dan Ventura (2015). 'Accounting for Bias in the Evaluation of Creative Computational Systems: An Assessment of DARCI'. In: ***Proceedings of the Sixth International Conference on Computational Creativity***, pp. 31–38 (cit. on p. 97).

Page, Larry et al. (1999). ***The PageRank Citation Ranking: Bringing Order to the Web.*** Technical Report. Stanford InfoLab (cit. on pp. 73, 75, 81, 121).

Partridge, Derek and Jon Rowe (1994). ***Computers and Creativity***. Oxford, UK: Intellect (cit. on pp. 53, 54, 59).

OED (2016). ***pataphysics***. Oxford Dictionaries | English. URL: https://en.oxforddictionaries.com/definition/pataphysics (visited on 10/10/2016) (cit. on p. 37).

Pease, Alison and Simon Colton (2011). 'On impact and evaluation in Computational Creativity: A discussion of the Turing Test and an alternative proposal'. In: ***Proceedings of the Artificial Intelligence and Simulation of Behaviour Conference*** (cit. on p. 97).

Pease, Alison, Simon Colton et al. (2013). 'A Discussion on Serendipity in Creative Systems'. In: ***Proceedings of the 4th International Conference on Computational Creativity***. Sydney, Australia: University of Sydney, pp. 64–71 (cit. on p. 98).

Pease, Alison, Daniel Winterstein and Simon Colton (2001). 'Evaluating Machine Creativity'. In: ***Proceedings of ICCBR Workshop on Approaches to Creativity***, pp. 129–137 (cit. on pp. 97, 98, 113).

Pérez y Pérez, Rafael and Otoniel Ortiz (2013). 'A model for evaluating interestingness in a computer-generated plot'. In: ***Proceedings of the Fourth International Conference on Computational Creativity***, pp. 131–138 (cit. on p. 104).

Peters, Tim (2004). ***PEP 20 — The Zen of Python***. URL: https://www.python.org/dev/peps/pep-0020/ (visited on 26/04/2016) (cit. on pp. 18, 19).

Piffer, Davide (2012). 'Can creativity be measured? An attempt to clarify the notion of creativity and general directions for future research'. In: ***Thinking Skills and Creativity*** 7.3, pp. 258–264 (cit. on pp. 99, 132).

Poe, Edgar Allan (2008). ***The Works of Edgar Allan Poe, Volumes 1-5***. Ed. by David Widgerm and Carlo Traverso. Project Gutenberg (cit. on p. 146).

Poincaré, Henri (2001). ***The Value of Science***. Ed. by Stephen Jay Gould. New York: Modern Library (cit. on pp. 53, 60, 66, 114).

Pólya, George (1957). ***How To Solve It***. Princeton, New Jersey: Princeton University Press (cit. on pp. 54, 66, 79, 114).

StackExchange (n.d.). ***Programming Puzzles & Code Golf***. StackExchange. URL: http://codegolf.stackexchange.com/questions/tagged/code-golf (visited on 20/11/2016) (cit. on p. 19).

Pyper, Martin (2010). ***one hundred thousand billion poems***. ME studio. URL: http://www.mestudio.info/2010/02/28/one-hundred-thousand-billion-poems/ (visited on 11/11/2016) (cit. on p. 14).

Python (n.d.). Python Software Foundation. URL: https://www.python.org/ (visited on 23/11/2016) (cit. on p. 144).

Queneau, Raymond (1961). ***One Hundred Thousand Billion Poems***. Gallimard (cit. on pp. 5, 14, 230).

Rabelais, François (2004). ***Gargantua and Pantagruel***. Project Gutenberg (cit. on p. 147).

Raczinski, Fania (2010). ***A passive interactive installation for the big screen***. URL: http://msc.fania.eu (visited on 15/08/2016) (cit. on p. 7).

Raczinski, Fania and Dave Everitt (2016). ‘Creative Zombie Apocalypse: A Critique of Computer Creativity Evaluation’. In: ***International Symposium of Creative Computing***. Oxford, UK (cit. on p. 181).

Raczinski, Fania, Hongji Yang and Andrew Hugill (2013). ‘Creative Search Using Pataphysics’. In: ***Proceedings of the 9th International Conference on Creativity and Cognition***. Sydney, Australia, pp. 274–280 (cit. on pp. 7, 174, 181).

Ramesh, V., Robert L. Glass and Iris Vessey (2004). ‘Research in computer science: an empirical study’. In: ***Systems and Software*** 70.1-2, pp. 165–176 (cit. on p. 23).

Ray Kurzweil’s Cybernetic Poet (2001). Kurzweil CyberArt Technologies. URL: http://www.kurzweilcyberart.com/poetry/rkcp_overview.php (visited on 22/11/2016) (cit. on p. 63).

Raymond, Eric (2004). ***The Jargon File, version 4.4.8***. URL: http://www.catb.org/~esr/jargon/ (visited on 20/11/2016) (cit. on p. 19).

Reitz, Kenneth (n.d.). ***Requests***. HTTP for Humans. URL: http://docs.python-requests.org/en/master/ (visited on 24/11/2016) (cit. on p. 162).

Rhodes, Mel (1961). ‘An analysis of creativity’. In: ***The Phi Delta Kappan*** 42.7, pp. 305–310 (cit. on pp. 54, 112).

Ricciardi, Giovanni (2014). ***Pataphysical Search Tool***. Patakosmos.com. URL: http://www.patakosmos.com/tool%5C_pataphysical%5C_search/ (visited on 03/11/2016) (cit. on p. 182).

Rimbaud, Arthur (2009). ***Poésies complètes***. Project Gutenberg (cit. on p. 147).

Ritchie, Graeme (2001). ‘Assessing creativity’. In: ***Proceedings of Symposium on Artificial Intelligence and Creativity in Arts and Science***, pp. 3–11 (cit. on p. 98).

– (2007). ‘Some Empirical Criteria for Attributing Creativity to a Computer Program’. In: ***Minds and Machines*** 17.1, pp. 67–99 (cit. on p. 98).

– (2012). ‘A closer look at creativity as search’. In: ***Proceedings of the International Conference on Computational Creativity***, pp. 41–48 (cit. on pp. 103, 230).

Ronacher, Armin (2008). ***Welcome to Jinja2***. pocoo.org. URL: http://jinja.pocoo.org/docs/dev/ (visited on 01/11/2016) (cit. on pp. 144, 163).

– (n.d.). ***Flask: web development, one drop at a time***. URL: http://flask.pocoo.org/ (visited on 31/10/2016) (cit. on pp. 144, 146).

Saint Luke (2005). ***The Gospel according to St. Luke***. Ebible.org. URL: http://ebible.org/asv/Luke.htm (visited on 15/10/2016) (cit. on p. 73).

Sawle, James, Fania Raczinski and Hongji Yang (2011). ‘A Framework for Creativity in Search Results’. In: ***Proceedings of the 3rd International Conference on Creative Content Technologies***. Rome, pp. 54–57 (cit. on pp. 96, 181, 222).

Schmidhuber, Jürgen (2006a). 'Developmental robotics, optimal artificial curiosity, creativity, music, and the fine arts'. In: ***Connection Science*** 18.2 (2006), pp. 173–187 (cit. on p. 132).

– (2006b). 'New millennium AI and the Convergence of history'. In: ***Singularity Hypotheses***. Ed. by Amnon Eden et al. 2006, pp. 60–82 (cit. on p. 96).

Schulman, Ari (2009). 'Why Minds Are Not Like Computers'. In: ***The New Atlantis*** 23, pp. 46–68 (cit. on p. 216).

Schütze, Hinrich (1998). 'Automatic Word Sense Discrimination'. In: ***Computational Linguistics*** 24 (1 1998) (cit. on p. 79).

Schütze, Hinrich and Jan Pedersen (1995). 'Information Retrieval Based on Word Senses'. In: ***Proceedings of the 4th Annual Symposium on Document Analysis and Information Retrieval***. Las Vegas, USA (cit. on p. 79).

Schwob, Marcel (2012). ***Der Kinderkreuzzug***. Trans. by Arthur Seiffhart. Project Gutenberg (cit. on p. 147).

Scott, Lee (2014). private communication. 26th May 2014 (cit. on p. 180).

GettyAPI (n.d.). ***Search For Creative Images***. Getty Images API. Getty. URL: http://developers.gettyimages.com/api/docs/v3/search/images/creative/get/ (visited on 07/08/2016) (cit. on pp. 144, 160, 209).

YouTube (n.d.). ***Search: list***. YouTube Data API. Google. URL: https://developers.google.com/youtube/v3/docs/search/list (visited on 07/08/2016) (cit. on pp. 144, 160, 162, 210, 211).

Searle, John (1980). 'Minds, Brains, and Programs'. In: ***Behavioral and Brain Sciences*** 3.3, pp. 417–457 (cit. on p. 130).

– (1990). ***Is the Brain a Digital Computer?*** American Philosophical Association. URL: http://users.ecs.soton.ac.uk/harnad/Papers/Py104/searle.comp.html (cit. on pp. 212, 213, 217).

– (1998). ***Brains and Machines: Correcting Some -Famous Mistakes-***. The DANA Foundation. URL: http://www.dana.org/Cerebrum/Default.aspx?id=39214 (visited on 15/11/2016) (cit. on pp. 214, 216).

– (2011). ***Watson Doesn't Know It Won on 'Jeopardy!'*** The Wall Street Journal. URL: http://www.wsj.com/articles/SB10001424052748703407304576154313126987674 (visited on 05/11/2016) (cit. on p. 218).

– (2015). ***Consciousness in Artificial Intelligence***. Talks at Google. URL: https://youtu.be/rHKwIYsPXLg (visited on 16/08/2016) (cit. on pp. 31, 107, 212, 213).

Shakespeare, William (2011). ***The Complete Works***. URL: http://www.gutenberg.org/cache/epub/100/pg100-images.html (visited on 30/10/2016) (cit. on p. 148).

Shattuck, Roger (1959). ***The Banquet Years***. London: Faber (cit. on p. 37).

Shu, Bo and Subhash Kak (1999). 'A neural network-based intelligent meta-search engine'. In: ***Information Sciences: Informatics and Computer Science*** 120 (1-4 1999), pp. 1–11 (cit. on p. 83).

Singh, Push (2005). 'EM-ONE: An Architecture for Reflective Commonsense Thinking'. PhD thesis. Massachusetts Institute of Technology (cit. on p. 12).

Sophia (2016). ***Sophia AI***. Hanson Robotics. URL: http://sophiabot.com/ (visited on 05/11/2016) (cit. on p. 219).

Srinivasan, P (2001). 'Vocabulary mining for information retrieval: rough sets and fuzzy sets'. In: ***Information Processing and Management*** 37.1, pp. 15–38 (cit. on p. 79).

Stahl, Bernd Carsten, Marina Jirotka and Grace Eden (2013). 'Responsible Research and Innovation in Information and Communication Technology: Identifying and Engaging with the Ethical Implications of ICTs'. In: ***Responsible Innovation***. Ed. by Richard Owen. John Wiley and Sons. Chap. 11, pp. 199–218 (cit. on pp. 116, 133).

Sternberg, Robert J (1999). ***Handbook of creativity***. Cambridge University Press (cit. on pp. 54, 56, 112, 118).

Still, Arthur and Mark d'Inverno (2016). 'A History of Creativity for Future AI Research'. In: ***Proceedings of the Seventh International Conference on Computational Creativity***, pp. 147–154 (cit. on p. 125).

Stribling, Jeremy, Max Krohn and Dan Aguayo (2016). ***SCIgen - An Automatic CS Paper Generator***. URL: https://pdos.csail.mit.edu/archive/scigen/ (visited on 05/11/2016) (cit. on p. 230).

Sutcliffe, Alistrair and Mark Ennis (1998). 'Towards a cognitive theory of information retrieval'. In: ***Interacting with Computers*** 10, pp. 321–351 (cit. on p. 79).

@tayandyou (2016). ***Tay.AI***. @tayandyou. Twitter. URL: https://twitter.com/tayandyou (visited on 05/11/2016) (cit. on p. 218).

Taye, Mohammad Mustafa (2009). 'Ontology Alignment Mechanisms for Improving Web-based Searching'. PhD thesis. De Montort University (cit. on p. 83).

TREC (2016). ***Text REtrieval Conference (TREC)***. National Institute of Standards and Technology. URL: http://trec.nist.gov/ (visited on 20/10/2016) (cit. on p. 95).

Dada Engine (2016). ***The Dada Engine***. dev.null.org. URL: http://dev.null.org/dadaengine/ (visited on 05/11/2016) (cit. on p. 230).

Stanford Parser (2016). ***The Stanford Parser: A statistical parser***. The Stanford Natural Language Processing Group. Standford University. URL: http://nlp.stanford.edu:8080/parser/index.jsp (visited on 22/11/2016) (cit. on p. 92).

Thomas, Sue et al. (2007). 'Transliteracy: Crossing divides'. In: ***First Monday*** 12.12 (cit. on pp. 26, 69).

Thrid (2002). ***An Explanation of l33t Speak***. The Hitchhiker's Guide to the Galaxy: Earth Edition. URL: http://h2g2.com/edited%5C_entry/A787917 (visited on 20/11/2016) (cit. on p. 19).

Toivanen, Jukka, Matti Järvisalo and Hannu Toivonen (2013). 'Harnessing Constraint Programming for Poetry Composition'. In: ***Proceedings of the Fourth International Conference on Computational Creativity***, pp. 160–167 (cit. on p. 47).

Top 500 (2016). ***TOP 10 Sites for June 2016***. Top 500. URL: https://www.top500.org/lists/2016/06/ (visited on 04/11/2016) (cit. on p. 214).

Transdisciplinary DMU (2013). ***The Pataphysics of the Future***. YouTube. URL: https://www.youtube.com/watch?v=UxYUZMyPE0o (visited on 03/11/2016) (cit. on p. 182).

TREC (2011). ***TREC Web, Terabyte & Blog Tracks***. Web Research Collections. University of Glasgow. URL: http://ir.dcs.gla.ac.uk/test%5C_collections/ (visited on 20/10/2016) (cit. on p. 95).

Turing, Alan (1950). 'Computing Machinery and Intelligence'. In: ***Mind*** 59, pp. 433–460 (cit. on p. 130).

– (1951). ***Can digital computers think?*** BBC Third Programme. Radio Broadcast (cit. on pp. 212, 216, 217).

– (2009). 'Computing Machinery and Intelligence'. In: ***Parsing the Turing Test***. Ed. by Robert Epstein, Gary Roberts and Grace Beber. Springer. Chap. 3, pp. 23–66 (cit. on pp. 107, 185, 212).

Copyright (2015). ***UK Copyright Law***. Factsheet No. P-01. UK Copyright Service. URL: https://www.copyrightservice.co.uk/ukcs/docs/edupack.pdf (visited on 01/11/2016) (cit. on p. 147).

van der Velde, Frank et al. (2015). 'A Semantic Map for Evaluating Creativity'. In: ***Proceedings of the Sixth International Conference on Computational Creativity***, pp. 94–101 (cit. on p. 106).

Veale, Tony (2013). 'Once More, With Feeling! Using Creative Affective Metaphors to Express Information Needs'. In: ***Proceedings of the Fourth International Conference on Computational Creativity***, pp. 16–23 (cit. on p. 16).

Ventura, Dan (2008). 'A Reductio Ad Absurdum Experiment in Sufficiency for Evaluating (Computational) Creative Systems'. In: ***Proceedings of the 5th International Joint Workshop on Computational Creativty***. Madrid, Spain (cit. on p. 98).

Verhaeren, Emile (2010). ***Poems***. Trans. by Alma Strettell. Project Gutenberg (cit. on p. 147).

Verlaine, Paul (2009). ***Poems***. Trans. by Gertrude Hall. Project Gutenberg (cit. on p. 147).

Verne, Jules (2010). ***A Journey to the Interior of the Earth***. Project Gutenberg. URL: http://www.gutenberg.org/cache/epub/3748/pg3748-images.html (visited on 15/10/2016) (cit. on pp. 73, 147).

Vries, Erica de (1993). ***Stretching the initial problem space for design problem solving: Browsing versus searching in network and hierarchy structures***. OCTO report 93/02 (cit. on p. 80).

Vykhodtsev, Alexander (2015). ***Chrome-Load-Timer***. Google Chrome extension to measure page load time and display it in the toolbar. GitHub. URL: https://github.com/alex-vv/chrome-load-timer (visited on 25/11/2016) (cit. on p. 200).

W3.CSS (n.d.). ***W3.CSS Tutorial***. Faster and Better Responsive Websites. W3 Schools. URL: http://www.w3schools.com/w3css/ (visited on 23/11/2016) (cit. on p. 143).

Walber (2014). ***File:Precisionrecall.svg***. Wikimedia Commons. URL: https://commons.wikimedia.org/wiki/File:Precisionrecall.svg (visited on 20/10/2016) (cit. on p. 95).

Walker, Richard (2012). ***The Human Brain Project***. Tech. rep. HBP-PS Consortium (cit. on pp. 129, 215).

Wallas, Graham (1926). ***The Art of Thought***. Jonathan Cape (cit. on pp. 53, 60, 66, 114).

Walsh, Dave (2001). ***Alfred Jarry: Absinthe, Bicycles and Merdre***. Blather.net. URL: http://www.blather.net/theblather/2001/05/alfred%5C_jarry%5C_absinthe%5C_bicycles/ (visited on 15/11/2016) (cit. on p. 39).

Hanson (2016). ***We bring robots to life***. Hanson Robotics. URL: http://www.hansonrobotics.com/ (visited on 05/11/2016) (cit. on p. 219).

WordNet (n.d.). ***What is WordNet?*** WordNet: A lexical database for English. Princeton University. URL: http://wordnet.princeton.edu (visited on 20/10/2016) (cit. on pp. 84, 85, 144, 229).

White, Ryen William (2013). 'Beliefs and biases in web search'. In: ***Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval***. Dublin, Ireland (cit. on p. 220).

White, Ryen William and Gary Marchionini (2004). 'Evaluating Exploratory Search Systems'. In: ***Information Processing & Management*** 44 (433-436 2004) (cit. on p. 222).

Python (2016). ***Why is it called Python?*** General Python FAQ. URL: https://docs.python.org/3/faq/general.html#why-is-it-called-python (visited on 20/11/2016) (cit. on p. 18).

Wickson, F., A.L. Carew and A.W. Russell (2006). 'Transdisciplinary research: characteristics, quandaries and quality'. In: ***Futures*** 38.9, pp. 1046–1059 (cit. on p. 185).

Widyantoro, D.H. and J. Yen (2001). 'A fuzzy ontology-based abstract search engine and its user studies'. In: ***Proceedings of the 10th IEEE International Conference on Fuzzy Systems***, pp. 1291–1294 (cit. on pp. 79, 83).

Wiggins, Geraint (2006). 'A preliminary framework for description, analysis and comparison of creative systems'. In: ***Knowledge Based Systems*** 19.7, pp. 449–458 (cit. on pp. 102, 230).

Winter, Joke de (2016). ***ArtyBollocks Generator***. URL: https://artybollocks.com/ (visited on 05/11/2016) (cit. on p. 230).

Wolf, David and Anette Wolf (n.d.). ***The Easter Egg Archive***. URL: http://www.eeggs.com/faq.html (visited on 20/11/2016) (cit. on p. 20).

Wong, Grace (2013). ***It's not a bug, it's a feature: the rise of glitch art***. The Guardian. URL: https://www.theguardian.com/artanddesign/2013/oct/25/rise-of-glitch-art (visited on 20/11/2016) (cit. on p. 20).

Yang, Hongji (2013). 'Editorial'. In: ***International journal of Creative Computing*** 1.1, pp. 1–3 (cit. on pp. 29, 65, 115).

Yossarian (2015). URL: https://yossarian.co/ (visited on 02/12/2015) (cit. on pp. 15, 16).

KTHXBYE