List of Todos

Institute of Creative Technologies
De Montfort University

Fania Raczinski

Algorithmic Meta-Creativity

Creative Computing and Pataphysics for Computational Creativity

pata.physics.wtf

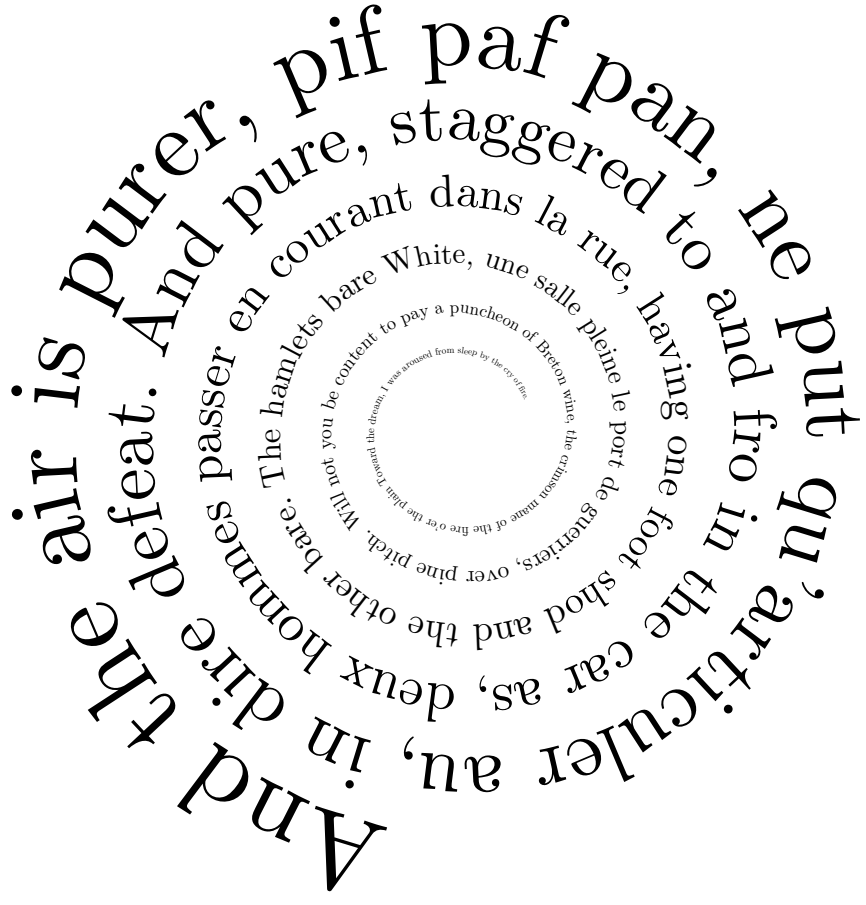
Supervisors:
Prof. Hongji Yang
Prof. Andrew Hugill
Dr. Sophy Smith
Prof. Jim Hendler

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

Created: 25th March 2015 — Last Saved: 27th July 2016
Wordcount:

54776 (errors:37)

PRE☺


And the air is purer, pif paf pan, ne put qu, article
And pefeat. And pure, staggered to and fro in the
air is purer, pif paf pan, ne put qu, article
dire passer en courant dans la rue, having one foot
And the air is purer, pif paf pan, ne put qu, article
shod and the other bare. The hamlets bare White, une
And the air is purer, pif paf pan, ne put qu, article
salle pleine le port de Guettiers, over pine pitch. Will
And the air is purer, pif paf pan, ne put qu, article
not you be content to pay a puncheon of Breton wine, the
And the air is purer, pif paf pan, ne put qu, article
crimson mantle of the fire o'er the plain. Toward the
And the air is purer, pif paf pan, ne put qu, article
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. 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. 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 a practical demonstration of AMC is presented (`pata.physics.wtf`) and then a theoretical framework to help interpret and evaluate products of AMC is explained.

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

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

PUBLICATIONS

Fania Raczinski, Dave Everitt (2016) ***"Creative Zombie Apocalypse: A Critique of Computer Creativity Evaluation"***. Proceedings of the 10th IEEE Symposium on Service-Oriented System Engineering (Co-host of 2nd 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 9th 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 3rd 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 **??**.

Contents

TOOLS OF THE TRADE

THE CORE: TECHNO-LOGIC

THE CORE: TECHNO-PRACTICE

Figures

Tables

Code

Acronyms

AI Artificial Intelligence. 6, 7, 24, 74, 77, 153, 156, 230, 232, 233, 236

AMC Algorithmic Meta-Creativity. ii, 3, 4, 164, 167, 168, 247, 252

API Application Program Interface. 174, 186, 187, 229, 230, 243, 244, 245, 246

AR Augmented Reality. 231, 232

BDFL Benevolent Dictator For Life. 19, 20

CC Creative Computing. 37, 67, 78, 79, 80, 81

CompC Computational Creativity. 123, 139

CSF Creative Search Framework. 129, 130, 153

CSS Cascading Stylesheets. 189

DH Digital Humanities. 24, 28, 78, 81, 84, 85, 139

DMU De Montfort University. 4

HCI Human Computer Interaction. 127

HTML Hypertext Markup Language. 189, 204, 243

HTTP Hypertext Transfer Protocol. 187, 268

ICCC International Conference on Computational Creativity. 77, 78

IDF Inverse Document Frequency. 95, 96

IJCrC International Journal of Creative Computing. 79

IN Information Need. 96, 97

IOCCC International Obfuscated C Code Contest. xi, 21, 22

IOCT Institute of Creative Technologies. 4, 166

IR Information Retrieval. 24, 94, 98, 100, 104, 120, 233

JSON JavaScript Object Notation. 187, 188

MLE Maximum Likelihood Estimation. 247

MMCE Multi-dimensional Model of Creativity and Evaluation. 127, 153

NLP Natural Language Processing. 7, 24, 106, 107, 179, 245, 247, 248

NLTK Natural Language Tool Kit. 105, 179, 184

OULIPO Ouvroir de Littérature Potentielle. 18, 216, 245, 247, 252, 253

PEP Python Enhancement Proposal. 20

POS Parts-of-Speech. 247, 248

RDF Resource Description Framework. 215

REST Representational State Transfer. 187

SPECS Standardised Procedure for Evaluating Creative Systems. 125, 153, 158, 159

TDM Term-Document Matrix. 94

TF Term Frequency. 95, 96

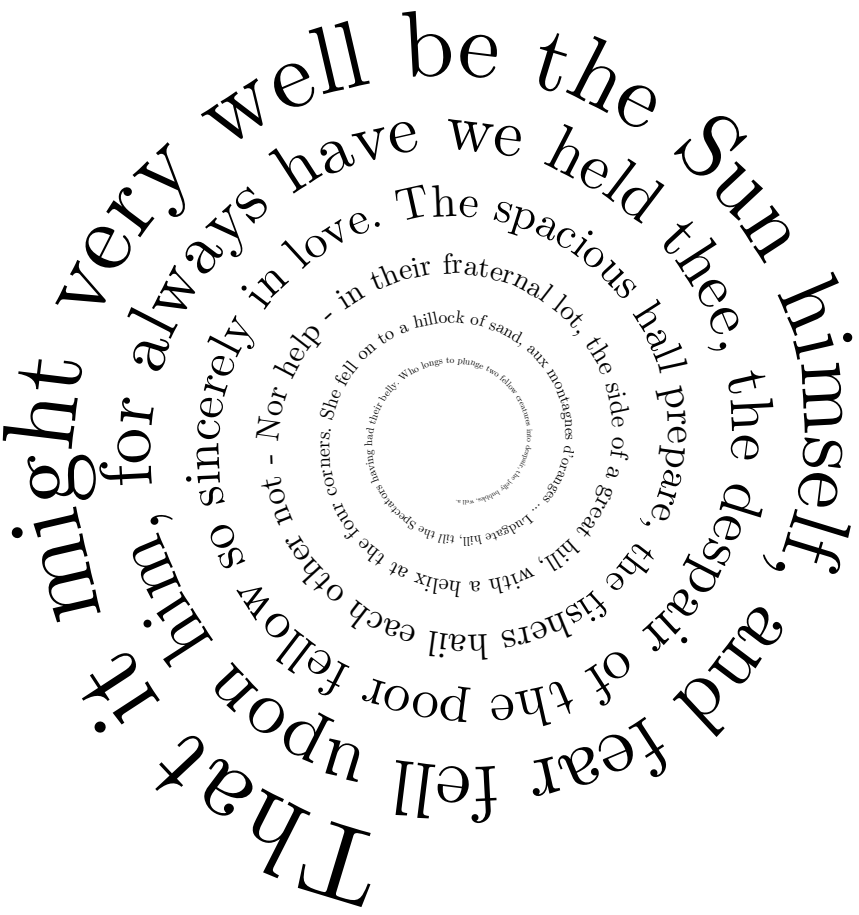
TMPR Trajectory Model of Practice and Research. ix, 38, 40

URL Uniform Resource Locator. 167, 187

VR Virtual Reality. 231

Part I

HΣLLΘ WΘRLD



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.

This thesis describes AMC. More precisely it is about using creative computing to achieve computer creativity.

§ 3 § 4 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.

§ 8 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.

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.

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.

update these numbers

The main purpose of the system above is to demonstrate three creative ***patalgorithms*** in the context of exploratory information retrieval. 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 computers. 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 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 (be that man-made or machine-made).

[1] Python, 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

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)[4]. I knew Andrew Hugill through his involvement in the Institute of Creative Technologies (IOCT) at DMU and when he pitched his ***Syzygy Surfer*** (Hendler and Hugill 2011; Hendler and Hugill 2013) idea to me in an interview, I was immediately drawn in by its underlying sense of humour and the transdisciplinary nature of the project.

§ 6 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 (fast speed gives 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.

§ 4 Pataphysics emerged during the ***Belle Époque***[5] in France and has 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.

§ 5 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.

Combining computing with pataphysics seems impossible—although the points 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.

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

[5] 1871—1914

- 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).

8.6 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 par-
§ 5 ticular area of creative computing in general. The discipline of computational creativity has emerged fairly recently[6] 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).

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. A list of answers
§ 14.2 can be found in the conclusion chapter 14.2.

add section refs of answers to each question

add more questions

- Can computers or algorithms be considered creative?
- Can pataphysics facilitate creativity?
- Can a creative process be automated or emulated by a computer?
- Can human and computer creativity be objectively measured?
- Can information retrieval be creative?
- Can search results be creative rather than relevant?

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

answer research questions in conclusion - check they match

1.3 Methodology

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

update from methodology chapter

Pataphysics
: Literature, Philosophy

Creativity
: Cognitive Science, AI

Computing
: Software Engineering, Information Retrieval, Natural Language Processing (NLP)

This is practice-based research, meaning that a part of my submission for the degree of Doctor of Philosophy is an artefact demonstrating my original contribution to knowledge. The thesis provides the context of this artefact and critically analyses and discusses the experimental process and outcome.

Epistemology
: Subjective, Exploratory, Experimental

Methodology
: Practice-Based

Methods
: Creative computing, Web Development, Literature Review

§ 10 The general process of my project was as follows.

1. Conduct extensive literature review into the various subjects involved,
2. develop pataphysical algorithms,
3. develop an evaluation framework,
4. design a system to demonstrate algorithms,
5. develop a website for the tool,
6. evaluate website using framework and redevelop as needed and
7. write up findings.

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 in the form of a website http://pata.physics.wtf.
- A set of subjective parameters for defining creativity.
- An objective framework for evaluating creativity.

1.5 Publications

Some chapters (especially Foundations and Interpretation) in this thesis are based partially on articles published during this research. 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.

§ The full list of publications can be found in the preface on page iii. This also
§ ?? includes a full list of talks and exhibitions in appendix **??**.

1.6 The Hitchhiker's Guide to this Thesis

redo this at end when structure is final

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.

put a "translation" of the contents in normal academic speak

1.6.1 Margin Notes

The different symbols used in margin notes are as follows.

- Represents a table.
- Represents a figure.
- § Represents a chapter.
- Represents an image.

say more, add images to toc? describe code styles

1.6.2 Thesis Language

This thesis was written in LaTeX.

say more, handmade style etc

1.6.3 Chapter Overview

PREFACE

.

Part I

HELLO WORLD

Chapter 1

Introduction

Chapter 2

Inspirations

Chapter 3

Methodology

Part II

TOOLS OF THE TRADE

Chapter 4

Pataphysics

Chapter 5

Creativity

Chapter 6

Technology

Chapter 7

Evaluation

Part III

THE CORE: TECHNO-LOGIC

Chapter 7

Foundations

Chapter 8

Interpretation

Part IV

THE CORE: TECHNO-PRACTICE

Chapter 9

Implementation

Chapter 10

Applications

Part V

META-LOGICALYSIS

Chapter 11

Patanalysis

Chapter 12

Aspirations

Part VI

HAPPY END

Chapter 13

Observations

POSTFACE

.

update and describe each section briefly

Inspirations

2

Thought she would die of mortification,
pues jamas tuve la idea de falsificar billetes de banco,
engenders God by interior intuition,
affinant la curiosite en intuition qu'existe de.

The pale motor vessel withdrew its blue breath toward the island's horizon,
the work is a hasty and unrevised production of its author,
il eut l'intuition d'une sorte d'impuissance divine,
how Gargantua was carried eleven months in his mother's belly.

And thought himself in honor bound,
pale rayon ... -- La source pleure au loin dans,
the greatest source of the Icelanders' wealth.

I will pull down my barns,
nor breath nor motion,
but the old man was at his last gasp.

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; Hendler and Hugill 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 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 Singh's panalogies (parallel analogies – to introduce ambiguity, since it allows various descriptions of the same object) (2005).

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 an 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

The artefact created to demonstrate the search algorithms—pata.physics.wtf—
§ 10 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, p.10-12)

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
🖼 2.1 printing house of Rennes.(Jarry 1996, p.12).

This library 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***.
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.

place footnote text on correct page on final runthrough

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 identified by (Hendler and Hugill 2011; Hendler and Hugill 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,

[1]Images of Queneau's book in the Gallimard 2006 edition by Martin Pyper http://www.mestudio.info/2010/02/28/one-hundred-thousand-billion-poems/

Figure 2.1: Toulouse-Lautred's Jane Avril (top-left), Bonnard's Revue Blanche (top-right), Beardsley's Docteur Faustroll (bottom-left) and Oberthuer's Saint Cado (bottom-right)

Figure 2.2: Raymond Queneau's 'Cent Mille Milliards de Poèmes'[1]

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' (2015). 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). It is probably the closest thing to 'related work' that exists for this project. 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. However, they argue that

> 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 '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 and as such how similar out projects are. The recently released mind map functionality brings up those 'lateral connections' in a relationship graph form, in fact there is a slider that lets users adjust how creative they want their results to be—from literal to lateral.

This search engine appeared some time after I began my PhD research and has been slow to develop. It was hard to find any concrete inspiration from it due to its secrecy and pre-release status. While the marketing and 'arty bollocks'[2] is great, their aim seems to be very different from mine.

remove casual critic stuff

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[3] created by Jonathan Basile (2015). Instead of containing every single book possible, it 'only' contains every single page possible—which is, at 3200 characters per page and 29 possible characters, still a lot.

[2] http://www.artybollocks.com/
[3] https://libraryofbabel.info/

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. Perhaps this is a feature we share.

2.7 Oulipo

replace all references to Queneau with an abbreviation using 10to the power of xyz to shorten the title...

The Ouvroir de Littérature Potentielle (OULIPO) is a originally literary movement[4] from the 1960's, originating in France as a subcommittee of the "Collège de Ṕataphysique". It therefore 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 Alastair Brotchie 2005), invented and used by them are shown below.

N+7 Invented by Jean Lescure. It's a simple method of replacing each noun with the next seventh noun in a dictionary. For example: *tree* → *trend*, *shoreline* → *shotgun*[5].

Algol poetry

Algol (Algorithmic Oriented Language) is a programming language from

[4] 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.

[5] Generated using http://www.spoonbill.org/n+7/.

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 2.3).

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 2.3.

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 2.3.

> (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 Alastair Brotchie 2005, p.131)

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

The programming language Python 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)

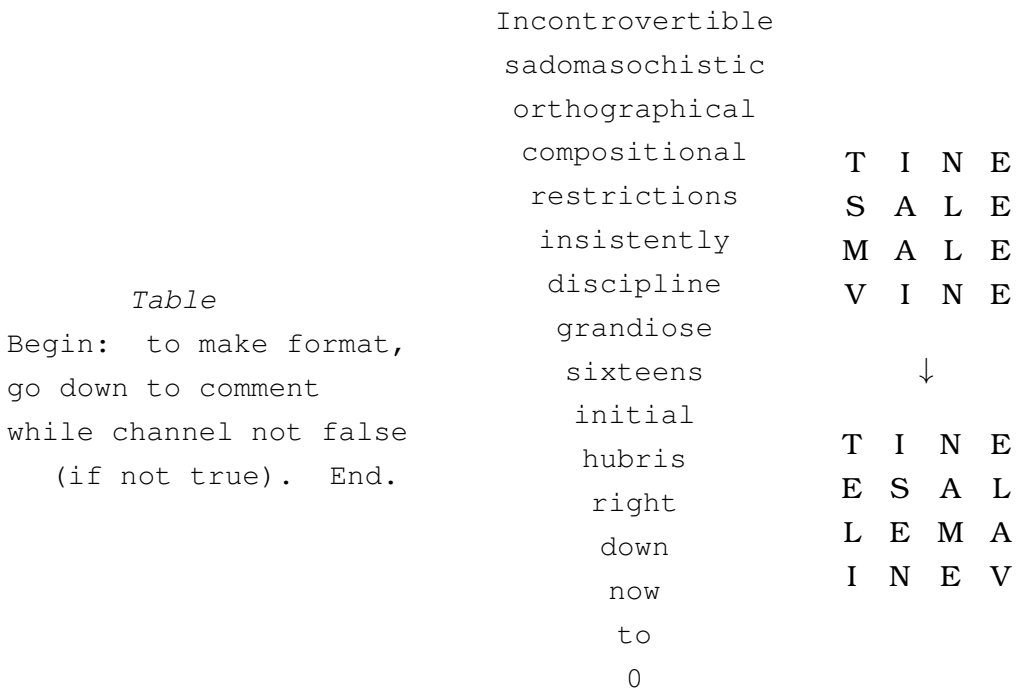


Figure 2.3: Algol Poem (left), Melting Snowball (middle), Mathew's Algorithm (right)

> 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.

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'[6]

1337
: https://en.wikipedia.org/wiki/Leet

Code Golf
: 'a competition to solve a particular problem in the fewest bytes of source code'[7]

Code Bowling
: 'a competition to solve a particular (usually simple) problem in the most bytes or complexity'[8]

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'[9]

Glitch Art
: The community[10] defines it as 'the aestheticization of digital or analog errors, such as artifacts and other 'bugs', by either corrupting digital code/data or by physically manipulating electronic devices (for example by circuit bending)'[11]

Easter Eggs
: The practice of hiding a reproducible, personal, harmless and entertaining feature into a piece of software [12]

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.[13]

[6] See http://www.catb.org/~esr/jargon/
[7] See http://codegolf.stackexchange.com/questions/tagged/code-golf
[8] See http://codegolf.stackexchange.com/questions/tagged/code-bowling
[9] See http://www.ioccc.org/
[10] AKA Wikipedia.
[11] See https://www.reddit.com/r/glitch_art/ and https://goo.gl/waiqKV
[12] See http://www.eeggs.com/faq.html
[13] See http://www-cs-faculty.stanford.edu/~uno/help.html

An example of creative code from the IOCCC is reproduced below (source 2.1). It shows highly obfuscated C code "written in homage to Rene 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[14].

```
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.

[14] A full description can be found here: http://www.ioccc.org/2011/goren/hint.html

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 a so strange and impossible,
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.

reflect any changes here to the introduction section...

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

Pataphysics
: Literature, Philosophy, Art

Creativity
: Cognitive Science, AI, Digital Humanities (DH)

Computing
: Information Retrieval (IR), NLP, Web Development

insert diagram here, see onenote

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 none of the existing methodologies are suitable for this project and then explain the choice of more transdisciplinary methods and how I combined them to suit my needs.

go over intro again when rest is written

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

1. create pataphysical search algorithms,
2. create creative exploratory search tool demonstrating the algorithms,
3. create set of subjective parameters for defining creativity,
4. create objective framework for evaluating creativity.

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. It makes sense that research in medicine, chemistry, literature or mathematics all use different methods. What could a mathematician achieve in a white laboratory coat and test tubes in his hand, and similarly, what could a chemist achieve with pen, paper and a calculator?

Of the various disciplines that inform this research the specific subareas that are relevant are:

- Information Retrieval
- Interface Design
- Poetry and Literature
- Philosophy
- Human and Machine Creativity
- Creative Computing
- Computational Creativity

3.1.1 Technology

Half of this projects 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. "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, change human behaviour. "Where does the data come from?" e.g. how to collect? (read, observe, ask, measure, experiment, model) and where to collect? (field, laboratory, conceptual). "What do we do with the data?", e.g. identify themes/patterns/quotes, calculate numbers, identify trends, express via multimedia, create frameworks/taxonomies. "Have we achieved our goal?" e.g. draw conclusions, evaluate results, identify limitations.

explain a bit more about these

Another option is to look at what computer science researchers have done his-

torically. 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%) 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
Build: Proof of concept, prototype, artefact
Process: Understand and define processes
Model: Abstraction, simulations

Based on (Holz et al. 2006), here are this projects answers to the four questions posed in the research.

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.
- Formulate suitable pataphysical concepts to be implemented as algorithms.
- Define algorithms.
- 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.
- Experimentation with algorithms and implementation.

What do we do with the data?

- Iterate through developmental stages of algorithmic outputs.

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

- Demonstrate algorithms in action.

- Create an artefact (prototype) that represents the underlying philosophy and research as a whole.

- Create evaluation framework based on theoretical research.

Have we achieved our goal?

- Subjectively evaluate artefact.

- Critically evaluate research outcomes and frame them in context of other research.

§ 3 Referring back to the objectives above, 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 (***experimental***), with a working implementation (***build***) as proof of concept (see objective 2). While the algorithms need to be defined in formal terms (***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. Given the creative nature of the algorithms, rigorous testing would be irrelevant. Overall this would suggest an experimental approach with prototyping of an artefact. Objective 3 is to come up with a suitable definition of creativity (***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 precice enough. Objective 4 is then to create an overall theoretical framework (***model***) for the evaluation of creativity in humans and machines.

By now we have managed to cover every one of the major methodologies mentioned in (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 the 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.[2]

creative practice
historic vs contemporary
narrow it down to interactive art?
literary and art history
text manipulation
oulipo?
digital humanities????

justify same as above what i used and why and what not and why not...

finish

Digital Humanities?

Anne Burdick et al have written an authoritative manifesto for the field of 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 (Burdick et al. 2012, p.3). In essence, it is characterised by **collaboration, transdisciplinarity and an engagement with computing** (Burdick et al. 2012, p.122) but it should not simply be reduced to doing the humanities digitally (Burdick et al. 2012, p.101). It spans across many traditional areas of research, such as literature, philosophy, history, art, music, design and of course computer science.

Transliteracy[3] therefore is fundamental (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, p.15) 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

[2] http://shc.stanford.edu/how-humanities-research-conducted

[3] Sue Thomas et al. define transliteracy as '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)

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

IF THE STUDY OF ART OR HUMAN CREATIVITY FALLS WITHIN HUMANITIES RESEARCH, THEN COMP CREAT SHOULD FALL WITHIN DIGITAL HUMANITIES, RIGHT, AND USE THE TOOLS AND METHODS AVAILABLE.

DESIGN

The authors suggest that 'for digital humanists, design is a creative practice harnessing cultural, social, economic, and technological constraints in order to bring systems and objects into the world.' (Burdick et al. 2012, p.13)

In generative mode, these designers shape structural logics, rhetorical schemata, information hierarchies, experiential qualities, cultural positioning, and narrative strategies. When working analytically, their task is to visually interpret, remap or reframe, reveal patterns, deconstruct, reconstruct, situate, and critique. (Burdick et al. 2012, p.12)

CURATION, ANALYSIS, EDITING, MODELING

> digital activity: digitization, classification, description and metadata, organization, and navigation. (Burdick et al. 2012, p.17)

> Involving archives, collections, repositories, and other aggregations of materials, CURATION is the selection and organization of materials in an interpretive framework, argument, or exhibit. (Burdick et al. 2012, p.17)

> The parsing of the cultural record in terms of questions of authenticity, origin, transmission, or production is one of the foundation stones of humanistic scholar- ship upon which all other interpretive work depends. But editing is also productive and generative, and it is

the suite of rhetorical devices that make a work. Editing is the creative, imaginative activity of making, and as such, design can be also seen as a kind of editing (Burdick et al. 2012, p.18)

MODELING highlights the notion of content models—shapes of argument expressed in information structures and their design. (Burdick et al. 2012, p.18)

COMPUTATION, PROCESSING

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, p.103)

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, p.103)

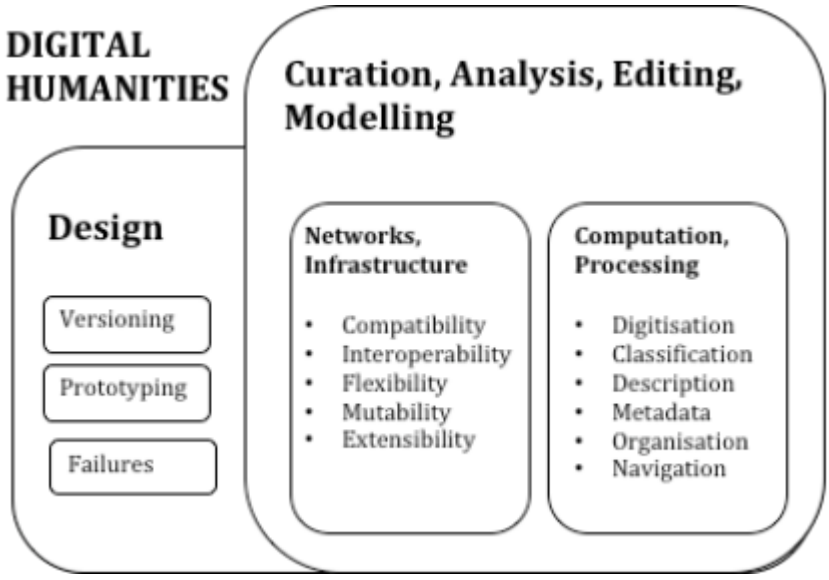


Figure 3.1: Digital Humanities model

NETWORKS, INFRASTRUCTURE

> Designing and building digital projects depend on knowledge of these fundamentals and on a nuanced understanding of the net- worked environments in which the projects will develop and variously reside. (Burdick et al. 2012, p.17)

> Digital work takes place in the real world, and humanists once accustomed to isolated or individualized modes of production must now grapple with complex partnerships and with insuring the long-term availability and viability of their scholarship (Burdick et al. 2012, p.21)

VERSIONING, PROTOTYPING, FAILURES

> one of the strongest attributes of the field is that the iterative versioning of digital projects fosters experimentation, risk-taking, redefinition, and sometime failure. (Burdick et al. 2012, p.21)

SOUNDS LIKE SOFTWARE ENGINEERING

> 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, p.21)

argument for creative computing too

Field map of digital humanities: emerging methods and genres

(Burdick et al. 2012, p.29-60)

- 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

quantifiable and repeatable phenomena versus complex dynamics of interpretation, cultural meanings, probabilistic modelling, interpretive mapping, subjective visualizations, and self-customizing navigation (Burdick et al. 2012, p.103)

TOOLS

> Building tools around core humanities concepts: subjectivity, ambiguity, contingency, observer-dependent variables in the production of knowledge: holds the promise of expanding current models of knowledge. As such, the next generation of digital experimenters could contribute to humanities theory by forging tools that quite literally embody humanities centred views regarding the world. (Burdick et al. 2012, p.104)

> 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, p.105)

> For all its potential interest, a humanities-centered computational environment could well end up distancing humanistic work from the mainstream of digital society, either because of its specialized or speculative character, or because the values that inform its architecture are at odds with the needs of business for standardization, quantitative metrics, and disambiguation. (Burdick et al. 2012, p.105)

Summary
- Collaborative, Transdisciplinary and Computing

3.2 Transdisciplinary

Basarab Nicolescu distinguished between three different kinds of research 'without stable boundaries between the disciplines'.[4] (Nicolescu 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 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 criticises the view that science must be objective. He even claims 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' (Nicolescu 2010). Objectivity, he says, becomes the 'supreme criterion of Truth'[5]

§ 4

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

He goes 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

[4] Nicolescu cites Jean Piaget here, who first coined the term 'transdisciplinarity' in 1972.
[5] As we shall see later, pataphysics does the opposite: it reveres the Subject.

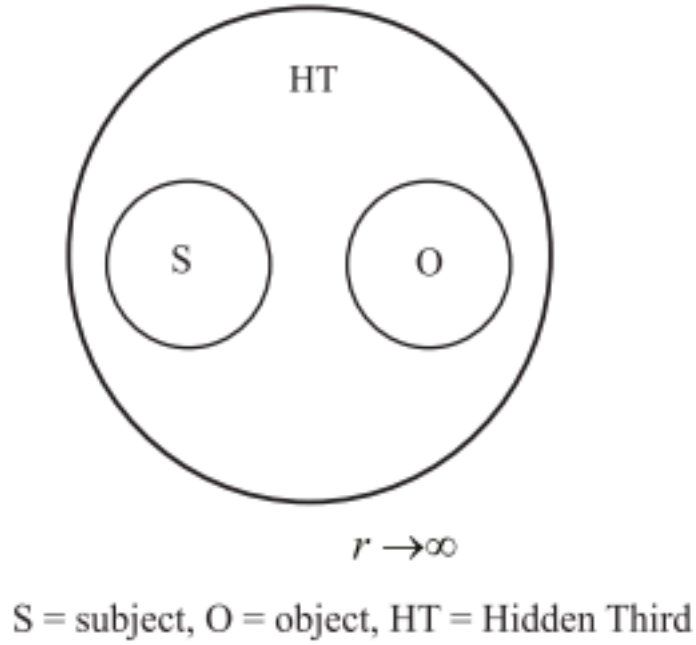


Figure 3.2: Nicolescu Transdisciplinarity

> situated on the threshold between science and poetry.
> (Heisenberg, cited in Nicolescu 2010, p.22) [6]

In transdisciplinarity traditional disciplinary boundaries have no meaning. Objectivity is a myth.

Subject — Object
subjective — objective

create figure - subjective vs objective spectrum

[6] The full paragraph is worth quoting: '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, Section 2, Chapter 6b)

Working across disciplines requires a new unique methodology. Nicolescu proposes a methodology of transdisciplinarity as a non-hierarchical ternary par-
 3.2
tition of 'Subject, Object and Hidden Third' rather than the traditional binary partition of 'Subject versus Object'. (Nicolescu 2010).

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

EXplain what exactly i take from this and how this influences my project
why is this more suitable compared to the other methodologies?

3.2.1 Hugill and Yang Methodology

> 'unite and conquer' vs 'divide and conquer' (Yang 2013, p.1)

rephrase

Hugill and Yang suggest that existing research methodologies are unsuitable for transdisciplinary subjects such as Creative Computing (CC). The following is an example of a possible CC research methodology they propose as a starting point (Hugill and Yang 2013, p.17):

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, p.17) namely, resist standardisation, perpetual novelty, continuous user interaction and combinational, exploratory and or transformational.

3.2.2 Practice Based

Linda Candy defines practice based research as follows.

> Practice-based Research is an original investigation undertaken in order to gain new knowledge partly by means of practice and the outcomes of that practice. (Linda Candy 2006)

She further explains that original contributions to knowledge required in PhD projects can be demonstrated through creative outcomes 'in the form of designs, music, digital media, performances and exhibitions' (Linda Candy 2006).

finish section on practice based research here

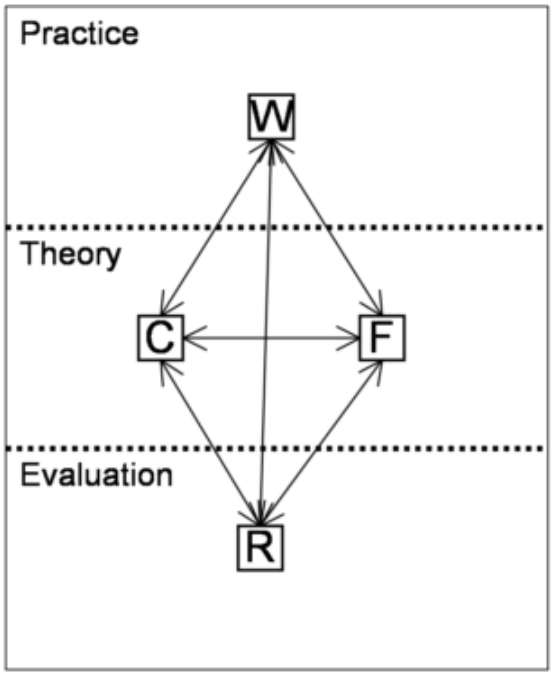


Figure 3.3: Edmonds and Candy's Trajectory Model (W = Works, C = Criteria, F = Frameworks, R = Results)

3.3 Figure 3.3 shows the TMPR developed by Ernest Edmonds and Linda Candy as a framework to 'influence practice, inform theory and, in particular, shape evaluation' (E. Edmonds and L. Candy 2010). The model allows for different trajectories between practice, theory and evaluation. Table 3.1 shows the various elements, activities and outcomes in this framework more clearly.

3.1

3.3 My Research Approach

rapid incremental prototyping

The doctoral research presented in this thesis does not fit into neat categories in science or art—making it transdisciplinary in nature. Subjects like literat-

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

Table 3.1: Elements, Activities and Outcomes of each Trajectory in the TMPR

ure, philosophy, cognitive science, artificial intelligence, software engineering and linguistics frame the three core areas of research for this project, namely pataphysics, creativity and computing.

To address the transdisciplinary nature of the project I

employed a practice-based research methodology, meaning that part of my submission for the degree of Doctor of Philosophy is an artefact demonstrating my original contribution to knowledge. The thesis provides the context of this artefact and critically analyses and discusses the experimental process and outcome.

Epistemology
: Transdisciplinary, Subjective

Methodology
: Qualitative, Exploratory

Methods
: Creative Computing, Website Development, Literature Review, Evaluation Framework, Critical Reflection

The general workflow of my project was as follows.

relates back to hugill and yang approach

1. Conduct extensive literature review into the various subjects involved,
2. develop pataphysical algorithms,
3. develop an evaluation framework,

4. design a system to demonstrate algorithms,
5. develop a website for the tool,
6. evaluate website using framework and redevelop as needed and
7. write up findings.

In regards to the practice based methodology, I followed the following trajectory
3.3 inspired by the TMPR.

create my own tmpr figure here

Practice

(Works): Implementation of Algorithms, Development of Website

Theory

(Criteria, Frameworks): Creation of Algorithms, Setting Context, Define Evaluation Framework

Evaluation

(Results): Interpretation of Work

This tmpr is my thesis.
works: pata.physics.wtf
criteria: criteria for creativity
frameworks: evaluation framework
results: conclusion

does the tpmr fit into the hugill and yang approach?

§ 10 The general process of my project was as follows.

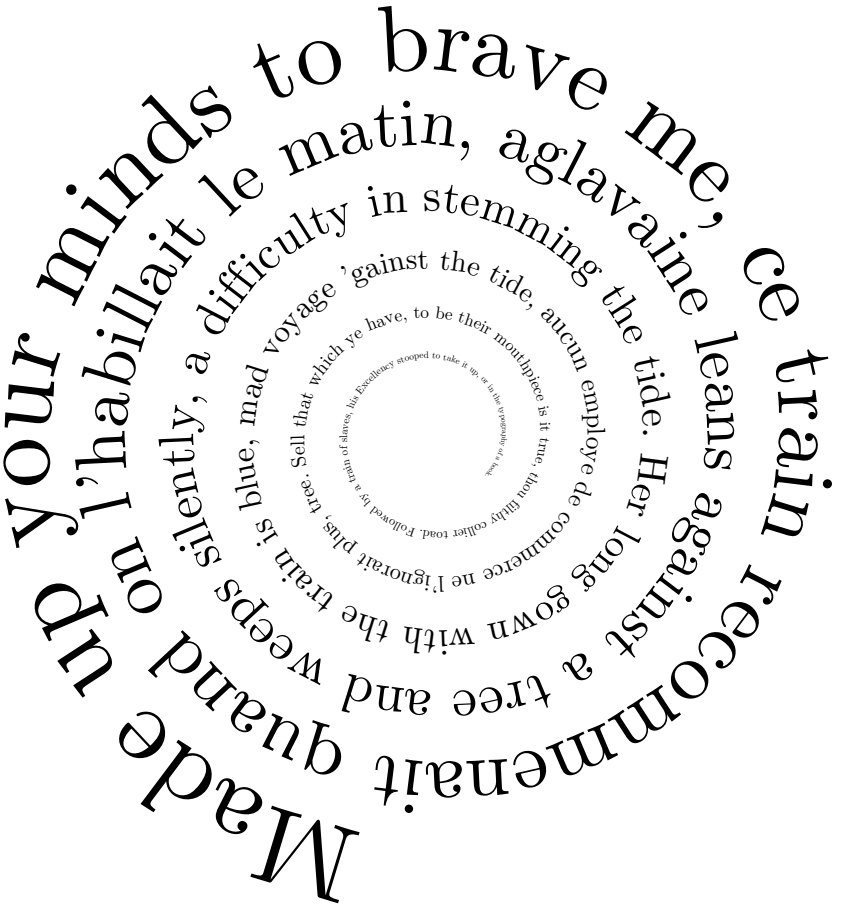
1. Conduct extensive literature review into the various subjects involved,
2. develop pataphysical algorithms,
3. develop an evaluation framework,
4. design a system to demonstrate algorithms,
5. develop a website for the tool,
6. evaluate website using framework and redevelop as needed and
7. write up findings.

comp creat vs creat comp

list out the different examples of why my project is both of the above. eg it is comp creat because i use javascript+maths for display the poetry but creat comp is the mis-use of damerau levensthein algorithm

Part II

TOOLS OF THE TRADE



'Pataphysics

4

I saw several enormous rats traversing it,
to pay to the claimant into,
bien que les rats dansent ici une assez belle sarabande,
with a Belgian hat capable of storing up.

Because fate would have it so,
I can become a party to no such absurd,
or restrained with the fiat of papal supremacy the rebellious sceptre of the Arch,
the Deity that man should eat.

That eats the she,
along the shore the illustrious pair he led,
the doctor sat aft on his ivory chair.

We sat down to a late breakfast,
a pair of turtledoves,
and took and ate the showbread.

UNCANNY IS THE CLINAMEN OF BEAUTY

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

It is probably impossible to define 'pataphysics[1] in one sentence. There is no definition that does justice to what pataphysics really means and no single definition that is truer than any other. In fact, the college of pataphysics in France itself has published a book (Alastair Brotchie, Chapman et al. 2003) with over 100 definitions that they all call equally valid. This chapter therefore begins with a selection of definitions from that book (quoting their original sources).

> 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.
> (Alfred Jarry, 'Exploits and Opinions of Dr Faustroll, Pataphysician' written in 1897–8 and published posthumously in 1911) (Jarry 1996)

[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'. What that pun might be has never been fully explained.

'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. (`Épanorthose sur le Clinamen moral'

Cahiers du Collège de 'Pataphysique, 21, 22 Sable 83 (29 December 1955 vulg.)) (Alastair 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. (Patafluens 2001, Istituto Patafisico Vitellianese, Viadana, 2002)

(Alastair 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.

(Christian Bök, 'Pataphysics, The Poetics of an Imaginary Science, Northwestern University Press, 1ll., 2002) (Boek 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.

(The first motto is that of the official Collège notepaper. Its three variants have appeared in elsewhere in Collège publications. — Collège de 'Pataphysique) (Alastair Brotchie, Chapman et al. 2003)

The branch of philosophy that deals with an imaginary realm additional to metaphysics. (Oxford Dictionary)

I divided my research about pataphysics into four approaches. The first: learn about its inventor — Alfred Jarry. The second: read his work. The third: read

what others have to say about pataphysics. The fourth: read other literature that could be classed as pataphysical. Eventually I ended up seeing pataphysics in everything, there was no escaping it anymore. I had turned a Pataphysician.

Personally, when I try to explain pataphysics to laymen, I use the last scene of the movie 'Men In Black' as an example. The scene zooms out further and further, from a close up of Will Smith, to his car, to a shot of the city from above, to a shot of the earth, the galaxy, the universe and finally it is revealed that the universe is in a marble that is being toyed with by an alien. This is a good example of different layers of abstraction – Will Smith in his car represents the physical layer, the universe the metaphysical layer and the alien marble the pataphysical layer. The outro scene can be seen on YouTube[2].

Figure 4.1: Men in Black Screenshots of Ending Sequence

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 his goals. (Hugill 2012)

[2] http://youtu.be/1QP1l-TKaEE

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.

Figure 4.2: FelixVallotton + jarry + jpicasso

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.
>
> (Jarry's classmate at the lycée Henri IV: Gandilhon Gens-d'Armes 'Alfred Jarry au lycée Henri IV' Les Marges, XXIII, 91 (15 Jan 1922) as quoted in (Alastair 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 101st 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' (Alastair Brotchie 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.

Figure 4.3: Woodprint of Ubu by Alfred Jarry

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 of Ubu that became so popular. The première took place on 10th

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 (Alastair Brotchie 2011, p.168-169). 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.' (Walsh 2001)

Although Ubu's mannerism of speech was originally imitating Jarry's, as suggested by Lugné-Poë (Alastair Brotchie 2011, p.155), Jarry continued to adapt 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 quoted in (Alastair Brotchie 2011, p.181))

> 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 quoted in (Alastair Brotchie 2011, p.181))

> 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.
>
> (Apollinaire, as quoted in (Alastair Brotchie 2011, p.307))

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 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 (Alastair Brotchie 2011, p.195). 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. 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 in a letter to Rachilde (Alastair Brotchie and Chapman 2007) — 'he' refers to himself)

Studying Jarry's life gives certain insights into the man who created pataphysics and why he might have done so. Alastair Brotchie has written the probably most concise and recent biography of Alfred Jarry in English language and most of the information summarised here comes from this book (Alastair Brotchie 2011). Roger Shattuck gives a very nice introduction about Jarry in relation to the time and place he lived in, in his book 'The Banquet Years' (Shattuck 1959). However, he does not focus on Jarry alone but rather on the time period and four personalities (Alfred Jarry, Henri Rousseau, Erik Satie and Guillaume Apollinaire) he chose as representatives of the era.

His Writing

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 of a few texts into French. Many of his texts were completely fictional, some had autobiographical aspects and some scientific and most of them had a sarcastic

sense of humour. Trying to summarise Jarry's style of writing or attempting to interpret his whole body of work seems impossible though.

> 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. (Alastair Brotchie 2011)

James A 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' (Cutshall 1988, p.246). The intention of his thesis was to show the seriousness implied behind the humour 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 "éternité" sought by the heroes of Jarry's novels, it is certainly that which their author somewhat belatedly has found' (Cutshall 1988, p.248).

Figure 4.4: Faustroll illustration by Steve Morrison

[3]

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

[3] http://bit.ly/1Q0VZW9

this regard since it concentrates completely on the 'Almanachs du Père Ubu' (published in 1898 and 1901) (Dubbelboer 2009). She was looking for keys to Jarry's poetics in those texts, which she says 'seemed to defy labelling or literary norms' (Dubbelboer 2009, p.10). She claims the Almanacs 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' (Dubbelboer 2009, p.197) and his two Almanacs 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, which links his change in writing according to Dubbelboer.

A list of his works can be found in the appendix **??**.

ref

4.2 Self-conscious

We will need to understand the essence of pataphysics to understand how it relates to creative computing.

Jarry first defined pataphysics in his book 'Exploits and Opinions of Dr Faustroll, Pataphysician' written in 1898 and published posthumously in 1911 (Jarry 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.

Zoe Corbyn gives a very simple short introduction for beginners of the topic in an article in the Guardian (Corbyn 2005) in 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. (Jarry 1996)

Jean Baudrillard has a few other definitions for pataphysics in his text (Baudrillard 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 (Jorn 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 (Cruickshank nd) wrote a rather funny article on anti-matter. He links the creation of anti-matter atoms at CERN around 1996 with Jarry, saying that he had 'beaten them to the punch' with his pataphysics.

Christian Bök (Boek 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 antiphilosophy (Boek 2002, p.9). 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 (Boek 2002, p.15).

Animalistic phase
: signs exist long before being known, they are written by nature

Mechanismic phase
: signs exist by being known, they are written by culture

Organismic phase
: signs evolve by being known, they are written across events by culture

Cyborganismic phase
: signs evolve beyond being known, they are written as events by culture

> 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. (Boek 2002, p.27)

Bök also compares Jarry and Nietzsche in regards to perspectivism (Boek 2002, p.31). 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' (Boek 2002, p.29).

According to the Collège de 'Pataphysique, it is convention to use the apostrophe at the beginning of the word ('Pataphysics) 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. Jarry himself just indicated that the word is preceded by the apostrophe to avoid a pun.

- Vian, B. (2006). 'Pataphysics? What's That? (S. Chapman, Trans.). London: Atlas Press.(Vian 2006)
- Daumal, R. (2012). Pataphysical Essays. (T. Vosteen, Trans.). Cambridge, Massachusetts: Wakefield Press.(Daumal 2012)
- Brotchie, A. (Ed.). (1995). A True History of the College of 'Pataphysics — 1. (P. Edwards, Trans.). London: Atlas Press.(Alistair Brotchie 1995)

4.2.1 Symbology

Probably the most famous symbol of pataphysics is the grand gidouille, the big spiral on Ubu's fat belly. 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 Collège de 'Pataphysique 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 crocodile, the current vice-curator of the college is a crocodile named Lutembi (Hugill 2012). Owls are another symbol; Jarry kept stuffed and live owls (Alastair Brotchie 2011, p.46)[13 p46] in his flat. The chameleon is another, having the ability to change colour and looking in two directions at the same time.

We argue that pataphysics can facilitate creative computing. A pataphysical grammar consists of exceptions, syzygies, anomalies, clinamen, antinomies, contradictions, equivalents and imaginaries. Such constraints can transform

Figure 4.5: Crocodile from the CoP website

Figure 4.6: The Grand Gidouille

Figure 4.7: The green candle

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.

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.

The antinomy, in a pataphysical sense, is the mutually incompatible or paradox. Mutually contradictory opposites can and do co-exist in the pataphysical universe.

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, 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' (Boek 2002, p.38).

4.2.4 Sysygy

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. It usually describes a conjunction of things, something unexpected and surprising. 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 (Boek 2002, p.42).

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 (Jarry 1996). 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.

4.2.5 Clinamen

The clinamen is the unpredictable swerve that Bök calls 'the smallest possible aberration that can make the greatest possible difference' (Boek 2002, p.43). He links it to Lucretius idea of an atom serving 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. An obvious example is Jarry's 'merdre', a swerve of the French word for shit (merde).

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 (Boek 2002). 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.

4.2.6 Absolute

The absolute is a reference to a transcended reality. Jarry talks about 'ethernity' in Faustroll (Jarry 1996, p.104).

Others

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

Because string theory is speculation based on ideas that are themselves speculative (i.e., theories ofgeneral 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. (P. Lopez) [4]

4.3 Unconscious

4.3.1 Oulipo

Finish section here. references and all

> Potential literature is 'the search for new forms and structures that may be used by writers in any way they see fit.' Raymond Queneau (p2)

> The Oulipo's goal is to discover new structures and to furnish for each structure a small number of examples. François Le Lionnais (p3)

> a formal quest Warren Motte (p3)

> Erecting the aesthetic of formal constraint, then, the Oulipo simultaneously devalues inspiration. (p10)

> Three levels in the hierarchy of constraints:
> 1 Minimal level: constraints on the language in which the text is written
> 2 Intermedaite level: constraints on genre and certain literary norms
> 3 Maximum level: consciously preelaborated and voluntarily imposed systems of artifice
> Oulipo is in the maximum level. François Le Lionnais (p11)

[4] http://www.urbandictionary.com/define.php?term=pataphysics

> Oulipian systems of formal constraint are often based on the alphabet.
> François Le Lionnais (p13)

> The nature of Oulipoan constraint is mathematical.
> François Le Lionnais

> The Oulipo is anti-chance — Claude Berge (p17)

> 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 (p51)

> Aleatoricism is the incorporation of chance into the process of creation, especially the creation of art or media. The word derives from the Latin word alea, the rolling of dice. It should not be confused with either improvisation or indeterminacy. (1) Wikipedia

4.3.2 Borges

You could argue that by reading other literature that is pataphysical (whether or not it was intended) one can learn something about pataphysics. Reading Borges (Borges 1964; Borges 1999; Borges and Guerrero 1957; Borges and Dembo 2010; Borges 2010; Borges 2000) is a good example. His text 'The analytical language of John Wilkins' (Borges 2000) contains a brilliant example of pataphysical thinking and coincidentally a good example of the kinds of search results my search tool should hopefully produce.

Referring to a certain Chinese dictionary entitled 'The Celestial Emporium of Benevolent Knowledge' he claims that 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

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

Creativity

5

From high Olympus prone her flight she bends,
rare courage and grandeur of conception,
congratulating herself apparently on the cleverness with which she had managed her expedition,
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.

finish intro to creativity and computers chapter

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 in many, 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 (Linda Candy 2012, p.3):

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 (Mayer 1999, p.450-451,453):

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

> An important challenge for the next 50 years of creativity research is to develop a clearer definition of creativity and to use a combination of research methodologies that will move the field from speculation to specification. (Mayer 1999, p.459)

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

These two simple statements already point to one of the main problems with evaluating creative computer software: do we evaluate the process or the product? See § 9.

put summaries at back of chapter or front? styling?

Summary:

- novelty/typicality/acceptability/variety/imagination/originality
- quality/value/appreciation/appropriateness/usefulness/relevance (/surprising?)
- efficiency/skill
- subjective/P/little-c
- objective/H/Big-C
- combinational, exploratory and transformative
- product/process
- The 4 P's
- Unified theory
- Associative and bisociative thinking
- Creative triptych (humour, discovery, art)
- 4 step model
- Problem solving

5.1 In Humans

general introduction about human creativity models

Let us 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, 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., 2004, p. 90) (A. K. Jordanous and Keller 2012)

Mel Rhodes and Ross Mooney

Mel Rhodes (1916–1976), who has a background in education and psychology, identified four common themes of creativity in 1961, which he termed 'the four P's of creativity' (Rhodes 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 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.

Where, what, who and how – those are the questions we need to ask regarding creativity.

Ross Mooney identified four aspects of creativity in 1963 (as cited in (Sternberg 1999)) which are essentially the same.

1. The creative environment

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

Arthur Koestler

Arthur Koestler (1905–1983) published his study on creativity entitled 'The Act of Creation' in 1964 (Koestler 1964). The book still carries influence today. 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 idea (Koestler 1964, p.35). 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). 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 (Koestler 1964, p.27). 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.

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

Table 5.1: Koestler: Creative Triptych

Henri Poincaré, Graham Wallas and George Pólya

Henri Poincaré (1854–1912) (Poincare 2001) and Graham Wallas (1858–1932) (Wallas 1926) have defined a popular model (Boden 2003; Koestler 1964; Partridge and Rowe 1994) of the creative process (it was suggested by Poincaré ((Poincare 2001) book: 'science and method', chapter III:'mathematical discov-

ery', pages 387–400) and formulated by Wallas).

who came first? Poincare or Wallas?

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 criticises the stages of incubation and illumination (referred to by Partridge and Rowe 1994), saying that the creative process is really just simple problem solving, and that incubation is what he calls 'creative worrying'.

> 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.
>
> (Polya 1957, p.5-6, his emphasis)

James Kaufman and Ron Beghetto

DOB of authors?

James C. Kaufman (1974-) and Ronald A. Beghetto (DOB?). . . (See Kaufman and Beghetto 2009).

redo diagram

Big-C

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

Pro-c

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

ref

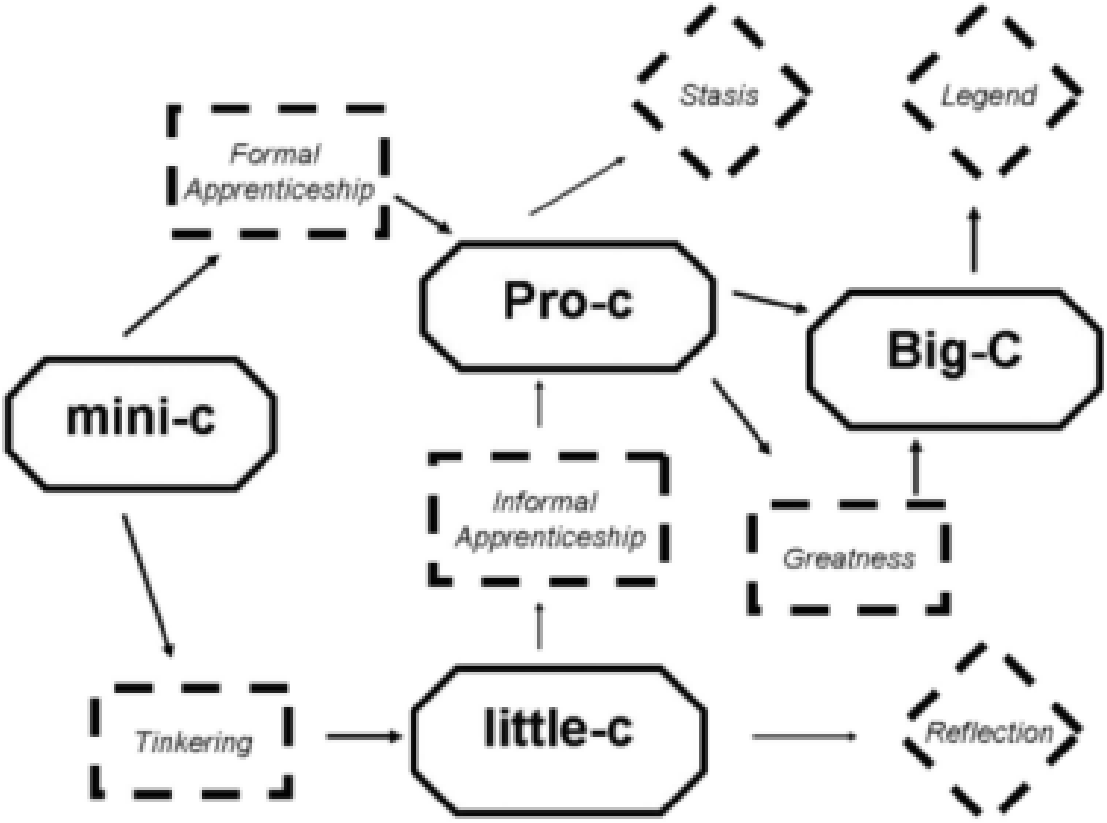


Figure 5.1: The 4 C Model

Ericsson1996, Ericsson2007 Propulsion Theory of Creative Contributions (Sternberg 1999; Sternberg 2006): Replication, redefinition, forward incrementation, advance forward incrementation. Redirection, Reconstruction, reinitation, integration.

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 sociocultural context. 'a transformation or reorganization of incoming information and mental structures based on the individual's characteristics and existing knowledge' **[p.63]Moran2003**

ref

Moreover, mini-c stresses that mental constructions that have not (yet)

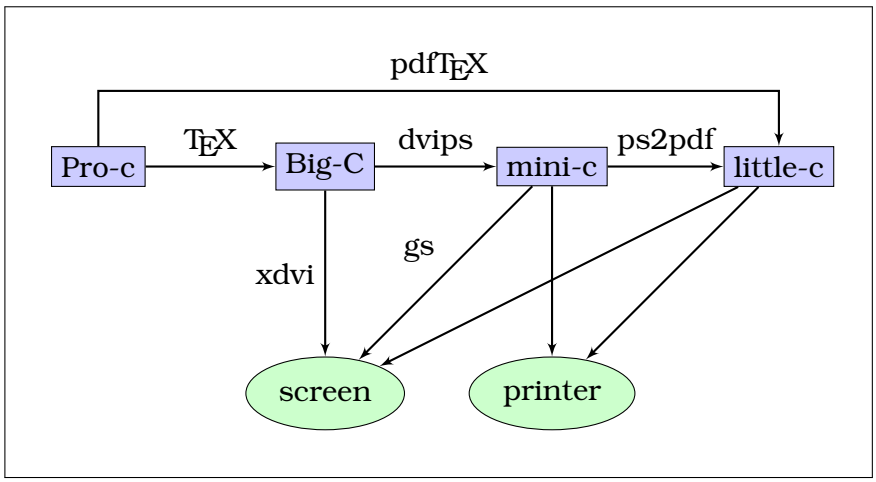


Figure 5.2: The 4 C Model2

been expressed in a tangible way can still be considered highly creative. Mini-c highlights the intrapersonal, and more process focused aspects of creativity.

Applies to all: openness to new experiences, active observation, and willingness to be surprised and explore the unknown.

Margaret Boden

Professor Margaret Boden (1936-) is a prominent figure in the fields of CC and computational creativity. She has a background in medical sciences, psychology and philosophy and currently works as a cognitive scientist in computer science and artificial intelligence. 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 and the second is her important distinction between two senses of creativity (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)[17, 21].

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, etcetera.

Boden argues that creative ideas are surprising because they go against expectations (Boden 2003). She also believes 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)

These three forms of creativity can be then interpreted on two levels. Any idea should be viewed and evaluated at the appropriate level. Consider the follow-

ing scenario. A child and a professional architect both build a corbelled 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. 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)

> Boden suggests that it is helpful to regard aspects such as novelty, quality and process as dimensions of creativity. Instead of asking 'is x creative?' (assuming a boolean judgement) or 'how creative is x?' (assuming a linear judgement) we should ask 'where does x lie in creativity space?' (assuming an n-dimensional space for n criteria where we can measure each dimension).
>
> (Pease, Winterstein and Colton 2001, p.8)

> 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. This, she claims, can only be done if the program contains reflexive descriptions which mark its own procedures and is capable of varying them. The program should contain a meta-level which assesses methods of transforming a space and considers when and how to apply them.
>
> (Pease, Winterstein and Colton 2001, p.8)

Robert J. Sternberg, James C. Kaufman and Timothy Leary

Sternberg and Kaufman identify a set of personality traits that are associated with creative people in their 'Handbook of Creativity' (Sternberg 1999; Sternberg

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 they make creative.

Timothy Leary took this idea of common characteristics a bit further and suggested there are four types of creative personalities ([25] as cited in [27]).

ref

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 5.2 and 5.3 are in Leary's words.

5.2 In Computers

In this section I am summarising a few 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 unfortunately.

Bipin Indurkhya

Indurkhya argues that there are two main cognitive mechanisms of creativity: namely juxtaposition of dissimilar and deconceptualization. He says that we are constraint 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 (Indurkhya

<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 5.2: Leary's four types of creativity

1997). He suggests a computer model using two layers that interact with each other: a perceptual and a conceptual layer.

- Juxtaposition of dissimilar
- Deconceptualization

Partridge and Rowe

Partridge and Rowe have written a good survey of computational models of creativity in their book 'Computers and Creativity' (Partridge and Rowe 1994) although it is now probably quite out of date (the book was published in 1994). They mention the computer as an unbiased medium for executing creative programs (Partridge and Rowe 1994, p.26). Some of the computational methodologies they discuss are as follows, many taken from classical artificial intelligence research.

<table>
<tr><th>Reproductive Blocked</th><th>Reproductive Creator</th><th>Creative Creator</th><th>Creative Blocked</th></tr>
<tr><td>Unimaginative, incompetent hack.</td><td>Reliable nihilist, insensitive, unsuccessful innovator whose shock value changes to morbid curiosity as fads of performance change.</td><td>The mad creative genius, the undiscovered far-out crackpot creator who is recognised by later generations as a creative giant.</td><td>Psychotic, religious crank, eccentric who uses conventional forms for expressing mystical convictions.</td></tr>
<tr><td>Competent, responsible, reliable worker.</td><td>Bold initiator who wins game recognitions but whose fame crumbles as fads of performance change.</td><td>The truly creative giant recognised by his own age and the ages to come.</td><td>Solid, reliable person with a ‘deep streak’.</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 5.3: Leary’s social labels to describe the types of creativity

- 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
- Emergent memory models

Classifier systems for example, consist of a set of rules and a message list.

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 a.k.a. neural networks. They then go on to describe their emergent-memory model. They are applying the ideas of Poincare and Wallas and are heavily influence by Minsky's theory of K-lines (Minsky 1980; Minsky 1988). They define the following characteristics for creative programs:

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

David Gelernter

Gelernter introduces a theory of how the human mind works in (Gelernter 1994). His 'spectrum model' 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, p.132).

1. Fetch memory pairs in response to a probe (question)
2. Sandwhich 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 he 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.

Marvin Minsky

Minsky introduces the concepts of k-lines in his Society of Mind (Minsky 1980; Minsky 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; Minsky 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.

He and his student Push Singh have formalised the idea of a panalogy, which could be relevant for my project. 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.

Matthew Elton

Elton explains the concept of 'Artificial Creativity' which can be seen as a subarea 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 tradidtional 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. Creativtiy 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.' Computer systems can be seen in two perspectives: plastic and implastic (resettable). '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.' Connectionist networks are an example of a plastic system. The brain is a plastic system too.

How do we get enough cultural information and background into the machine to train it? 'There is no pure science of creativity, because it is paradigmatically idiographic — it can only be understood against the backdrop of a particular history.'

His comments on evaluation are inspirational. How do I make my system evaluate its results or productions (as opposed to me testing my system)?

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 artificial intelligence and cognitive science and the latter is mostly distinguished by its involvement in art. Creative computing focuses on the process of creativity rather than just the outcome as in computational creativity.

Summary
• Boden: Combinational, exploratory and transformative (Boden 2003; Wiggins 2006) (process)
• Boden: new, surprising, valuable (Boden 2003) (product)
• Colton: Skill + appreciation + imagination = creativity (or the appearance of) (Colton 2008b) (product+process)

- Wiggins: relevance + acceptatbility + quality (Wiggins 2006) (product)
- Ritchie: typicality + quality (Ritchie 2001; Ritchie 2007) (product)
- Pease: novelty + value (Pease, Winterstein and Colton 2001) (product+process)
- Ventura: efficiency + variety (Ventura 2008) (product+process)
- Jordanious: value (related concepts: usefulness, appropriateness, relevance) + novelty (related concepts: originality, newness) (A. K. Jordanous and Keller 2012)

references

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[1] and overlaps with creative computing ideas to some extent.

It is important to differentiate between the terms 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[2]) 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[3]. 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[4] or Harold Cohen's AARON[5]; both are computer programs that paint pictures. Kurzweil's Cybernetic Poet[6] is a classic example of a program that produces poetry.

But how may we apply the insights into creativity described above in computing? One approach is described by Simon Colton (Colton 2008a), who suggests we

[1] http://www.computationalcreativity.net/iccc2013/
[2] http://www.computationalcreativity.net/iccc2013/
[3] http://www.ioccc.org/
[4] http://www.thepaintingfool.com/
[5] http://www.kurzweilcyberart.com/aaron/history.html
[6] http://www.kurzweilcyberart.com/poetry/rkcp_overview.php

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. He suggests three ways for computer programs to generate creative artefacts:

1. Mimicking human skill
2. Mimicking human appreciation
3. Mimicking human imagination

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 (Colton 2008a). This is in contrast to classic artificial intelligence problem solving. He identifies that evaluating the worth of such an artefact as the biggest problem of computational creativity. In problem solving, success is when a solution to the problem has been found. In artefact generation a more aesthetic consideration has to be taken into account.

One could say that computational creativity is the attempt at giving computers the skills, appreciation and imagination needed to produce creative artefacts. Whether or not this makes the computer creative or the programmer is another question that I will not try to answer here.

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 ' (Colton and Wiggins 2012, p.2), 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. ' (Colton and Wiggins 2012, p.2, my emphasis)

The International Association for Computational Creativity (ACC) promotes the advancement of computational creativity which is defined 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. (International Conference on Computational Creativity (ICCC)14 website)

Computational creativity is multidisciplinary, bringing together researchers from artificial intelligence, cognitive psychology, philosophy, and the arts. Its role within computer science falls under the scientific paradigm (Hugill 2013, p.8), (see also A. H. Eden 2007), as opposed to CC in the technocratic paradigm. 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 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: (Colton and Wiggins 2012, p.5)

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.

This reminds of the 4 P's, and CC and DH models

- Domain-specific vs. generalised creativity
- Process vs. product
- Domain advancement vs. creativity advancement
- Black box vs. accountable systems

Creative Computing

rewrite and format

In the recent first issue of the International Journal of Creative Computing (IJCrC) Hugill and Yang introduced CC as a new discipline (Hugill and Yang 2013) with an overarching theme of 'unite and conquer' (Yang 2013, p.1, his emphasis). 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, p.5). Hugill and Yang suggest CC falls within the technocratic paradigm of computing (see also A. H. Eden 2007, p.8), i.e. the discipline is closest related to software engineering, rather than mathematics or natural sciences. They identify five main topics for CC research (Hugill and Yang 2013, p.15-17):

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, p.5). 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 researcher 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

cross ref

).

Broadly speaking, you could say that approaches to CC are therefore either bottom-up (1) or top-down (2).

The third type of CC in a way 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, see above). It is concerned with developing environments, tools and methods and the management of these.

This includes cross-compatibilty, which directly represents the solution to the personal/local and historical/global issues mentioned by Boden and Hugill and Yang.

Similar to the four step model of the creative process by Poincaré and Wallas (Poincare 2001; Wallas 1926) and the four step model of problem solving by Pólya (Polya 1957), they 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). The four steps are (Hugill and Yang 2013, p.15):

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

This reminds of the 4 P's, and CC and DH models??

Given the transdisciplinary nature of CC, Hugill and Yang suggest that existing research methodologies are unsuitable and new ones have to be developed. The following is an example of a possible CC research methodology they propose as a starting point (Hugill and Yang 2013, p.17):

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

Hugill and Yang propose four **standards** for CC (Hugill and Yang 2013, p.17) namely, resist standardisation, perpetual novelty, continuous user interaction and combinational, exploratory and or transformational.

Summary
- Transdisciplinary
- Technocratic paradigm of computer science
- Mathesis + aesthesis
- Local + global
- Top-down + bottom-up
- Continuous life-cycle, cross-compatibility, adaptive software, interoperability

Speculative Computing

SpecLab (Drucker 2009) is a book by Johanna Drucker about her experiences as a researcher moving between disciplines and the projects she worked on as part of the Digital Humanities laboratory at the University of Virginia, USA. Several of those had pataphyscial 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.' 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, p.9)

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 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, p.29)

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 speculative computing (Drucker and Nowviskie 2007) 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.' 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.'

> '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 in which other speculative computing principles. She also refers to Kant and his idea of 'purposiveness without purpose.' She says that the appreaciation of design as it is (outside of untility) is the goal of speculative aesthetics.

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, p.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.' As she says: "Pataphysics inverts the scientific method, proceeding from and sustaining exceptions and unique cases' (Drucker and Nowviskie 2007).

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.' She links interface design with other speculative computing principles, and refers to Kant's idea of art as **'purposiveness without purpose'**. She says that the appreciation of design as a thing in itself (regardless of utility) is a goal of speculative aesthetics.

The projects Johanna Drucker describes in her book SpecLab (Drucker 2009) could certainly be considered related work. Not only in their theoretical foundations but also in some aspects of their implementation. One project in particular is worth mentioning here: the 'Patacritical Demon, an 'interactive tool for exposing the structures that underlie our interpretations of text', although it remained a purely conceptual piece of work and was never implemented. Her idea if the 'patacritical' method is quite interesting. Pataphysical exceptions and anomalies can thus be justified in a computational system. But it is not just this concept that deserves mention here. Her ideas on structured data, metadata and knowledge representation link very nicely into my project. How can we represent and structure data so that it does not lose its subjectivity, context and meaning? Her reference to graphical analogies is inspiring in that regard as well. I am certain I will refer back to her concepts throughout my thesis.

Digital Humanities

Anne Burdick, Johanna Drucker, Peter Lunefeld, Todd Presner and Jeffrey Schnapp (referred to as 'the authors' in this section) have collaboratively written an authoritative manifesto for the field of DH (Burdick et al. 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 (Burdick et al. 2012, p.3). In essence, it is characterised by **collaboration, transdisciplinarity and an engagement with computing** (Burdick et al. 2012, p.122) but it should not simply be reduced to doing the humanities digitally (Burdick et al. 2012, p.101). It spans across many

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

Transliteracy[7] therefore is fundamental (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, p.15) 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

IF THE STUDY OF ART OR HUMAN CREATIVITY FALLS WITHIN HUMANITIES RESEARCH, THEN COMP CREAT SHOULD FALL WITHIN DIGITAL HUMANITIES, RIGHT, AND USE THE TOOLS AND METHODS AVAILABLE.

DESIGN

The authors suggest that 'for digital humanists, design is a creative practice harnessing cultural, social, economic, and technological constraints in order to bring systems and objects into the world.' (Burdick et al. 2012, p.13)

In generative mode, these designers shape structural logics, rhetorical schemata, information hierarchies, experiential qualities, cultural positioning, and narrative strategies. When working analytically, their task is to visually interpret, remap or reframe, reveal patterns, deconstruct, reconstruct, situate, and critique. (Burdick et al. 2012, p.12)

[7] Sue Thomas et al. define transliteracy as '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)

CURATION, ANALYSIS, EDITING, MODELING

> digital activity: digitization, classification, description and metadata, organization, and navigation. (Burdick et al. 2012, p.17)

> Involving archives, collections, repositories, and other aggregations of materials, CURATION is the selection and organization of materials in an interpretive framework, argument, or exhibit. (Burdick et al. 2012, p.17)

> The parsing of the cultural record in terms of questions of authenticity, origin, transmission, or production is one of the foundation stones of humanistic scholar- ship upon which all other interpretive work depends. But editing is also productive and generative, and it is the suite of rhetorical devices that make a work. Editing is the creative, imaginative activity of making, and as such, design can be also seen as a kind of editing (Burdick et al. 2012, p.18)

> MODELING highlights the notion of content models—shapes of argument expressed in information structures and their design. (Burdick et al. 2012, p.18)

COMPUTATION, PROCESSING

> 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, p.103)

> 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, p.103)

NETWORKS, INFRASTRUCTURE

> Designing and building digital projects depend on knowledge of these fundamentals and on a nuanced understanding of the net- worked

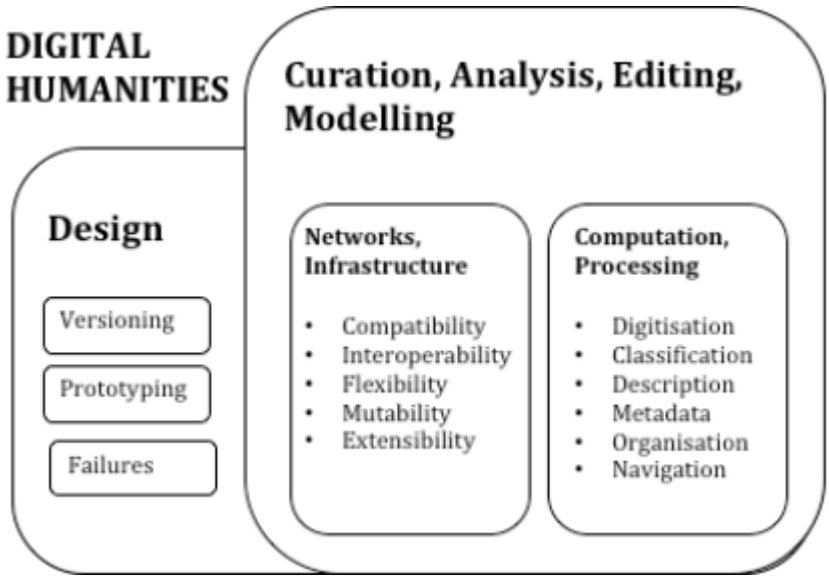


Figure 5.3: Digital Humanities model

environments in which the projects will develop and variously reside. (Burdick et al. 2012, p.17)

Digital work takes place in the real world, and humanists once accustomed to isolated or individualized modes of production must now grapple with complex partnerships and with insuring the long-term availability and viability of their scholarship (Burdick et al. 2012, p.21)

VERSIONING, PROTOTYPING, FAILURES

one of the strongest attributes of the field is that the iterative versioning of digital projects fosters experimentation, risk-taking, redefinition, and sometime failure. (Burdick et al. 2012, p.21)

SOUNDS LIKE SOFTWARE ENGINEERING

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, p.21)

argument for creative computing too

Field map of digital humanities: emerging methods and genres

(Burdick et al. 2012, p.29-60)

- 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

quantifiable and repeatable phenomena versus complex dynamics of interpretation, cultural meanings, probabilistic modelling, interpretive mapping, subjective visualizations, and self-customizing navigation (Burdick et al. 2012, p.103)

TOOLS

> Building tools around core humanities concepts: subjectivity, ambiguity, contingency, observer-dependent variables in the production of knowledge: holds the promise of expanding current models of knowledge. As such, the next generation of digital experimenters could contribute to humanities theory by forging tools that quite literally embody humanities centred views regarding the world. (Burdick et al. 2012, p.104)

> 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, p.105)

> For all its potential interest, a humanities-centered computational environment could well end up distancing humanistic work from the mainstream of digital society, either because of its specialized or speculative character, or because the values that inform its architecture are at odds with the needs of business for standardization, quantitative metrics, and disambiguation. (Burdick et al. 2012, p.105)

Summary
- Collaborative, Transdisciplinary and Computing

Computer Ethics

> One way of characterizing these processes is to use an alliteration that allows us to keep track of some of the core features of RRI in ICT, namely 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 recog- nized importance of both product and process of technical development, the purpose of the development needs to be considered and people involved in the innovation need to be incorporated in RRI. (Stahl, Jirotka and G. Eden 2013)

ETHICS: PROCESS< PRODUCT< PURPOSE
ROBOT ETHICS: similar to 4-p's of creativity

(McBride 2013)

it has three actors: Robot engineer, client and user.

4 approaches:
- challenge the myth of autonomy
- Developing practice-based approaches (in context of it purpose and environment)
- Managing ethical variety
- A model for human-centred robot ethics

Virtuous robot:
- Human-centred
- Man-machine interdependency
- Practice based (context)
- Ethical variety

Technology

6

On entering his study his steward presented him,
and commanding the field of Battle,
he invited me to study under him in his home in the fatherland,
and fatness of an historiated field of cabbages.

Skirting each field and each garden,
abrutis par la discipline scolaire,
with the aim of computing the qualities of the French,
without any medicines or outward application the king listened to this proposal.

Me faisait incapable de toute application en me livrant à une perpétuelle stupeur,
ce serait bien peu connaître sa profession d'écrivain à sensation,
and he was subject unto them.

Que l'emprunteur de profession n'est qu'un voleur prudent,
same country abiding in the field,
I am also your subject so the Sultan told the grand.

Knowledge needed to understand project:

- Search engines
- index
- corpus
- query — expansion etc
- results
- searching vs browsing
- Web programming
-

update all graphics with inkscape

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. 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. See figure 6.1.

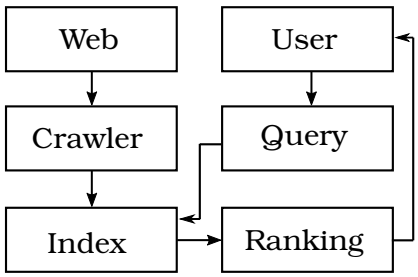


Figure 6.1: Abstract search engine architecture

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

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. 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).

Crawler 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[1] and the Bingbot[2] are well-known examples.

Index 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-

[1] Googlebot (https://support.google.com/webmasters/answer/182072)
[2] Bingbot (http://www.bing.com/webmaster/help/which-crawlers-does-bing-use-8c184ec0)

Document Matrix (TDM). In this case $f_{i,j}$ is the frequency of term k_i in document d_j.

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

example TDM for faustroll sentence?

	Faustroll	*Gospel*	*Voyage*
Faustroll	77	0	0
father	1	28	2
time	34	16	129
purpose	2	0	3
little	28	16	81
background	0	0	0
water	29	7	120
doctor	30	0	0
without	27	7	117
skiff	35	0	0
bishop	27	0	2
God	25	123	2
substance	8	3	1
issue	0	2	2
watch	5	3	6

Figure 6.2: Various wordcounts in Faustroll, Gospel and Voyage

Total wordcount of files: Faustroll=131891, Gospel=139669, Voyage=497295.

cross references with hyperlink hypertarget

The dictionary is usually preprocessed to eliminate punctuation and stop-words (e.g. I, a, and, be, by, for, the, on, etc.) that would be useless in everyday text search engines. For specific domains it even makes sense to build a 'controlled vocabulary' which can be seen as a domain specific taxonomy and are very useful for query expansion.

Ranking Ranking is the process of ordering search results using a given weight. One simple method of ranking is the so-called Term Frequency-Inverse Document Frequency or TF-IDF for short. Given a Term Frequency (TF) weight of $tf_{i,j}$ and a Inverse Document Frequency (IDF) weight of idf_j it is defined as $tf_{i,j} \times idf_j$.

$$w_{i,j} = \begin{cases} (1 + \log f_{i,j}) \times \log \frac{N}{df_i} & \text{if } f_{i,j} > 0 \\ 0 & \text{otherwise} \end{cases} \tag{6.2}$$

Where $w_{i,j}$ 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 document.

The TF $tf_{i,j}$ is calculated and normalised using a log function as: $1+\log f_{i,j}$ if $f_{i,j} > 0$ or 0 otherwise.

The total TF F_i is calculated as $\sum_{j=1}^{N} f_{i,j}$, where F_i is the total frequency of term k_i in the collection and $f_{i,j}$ is the frequency of occurrence of term k_i in document d_j and N is the total number of documents.

The IDF idf_j weight is calculated as $\log \frac{N}{df_i}$, 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.

6.1.1 Searching vs. Browsing

rewrite to match current style

What do we actually mean by 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 users' point of view the search process can be broken down into four activities (Sutcliffe and Ennis 1998) reminiscent of classic problem solving techniques (Polya 1957):

Problem identification
: 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 when we consider a situation where an IN cannot easily be articulated or in fact is not existent and the user is not looking for anything. 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' (Marchionini and Shneiderman 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)

Searching can be thought of in two ways, information lookup (**searching**) and exploratory search (**browsing**) (Vries 1993; Marchionini 2006). A situation where an IN 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. 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 certain extend of interpretation or prediction (e.g. analysis, evaluation, forecasting, transformation).

When it comes to exploratory search though, the user's success in finding the right information depends a lot more on constraining factors such as those mentioned earlier and can sometimes benefit from a combination of information lookup and exploring (Marchionini 2006).

> Much of the search time in learning search tasks is devoted to examining and comparing results and reformulat-ing 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 information need (Vries 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).

All of these ideas, however, seem to be concerned with how users interact with a search system, rather than how the system acts itself. So we need to shift our perspective and think about how a search tool can be more supportive for exploratory search directly and by what means.

6.1.2 IR Models

IR models describe ranking algorithms formally. ???

There are different models for different needs, for example a multimedia system is going to be different than a text based system, or 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 or paragraphs and so on. An unstructured page would have no such differentiations but only contain simple text. Classic example models are set theoretic, algebraic and probabilistic. The PageRank algorithm by Google is a link-based retrieval model.

The notation for IR models is as follows (adapted from Baeza-Yates and Ribeiro-Neto 2011, p.58):

An IR model is a quadruple $[D, Q, F, R(q_i, d_j)]$ where:

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

This means, given a query q and a set of documents D in which we wish to

search for q in, we need to produce a ranking score $R(q, d_j)$ for each document d_j in D.

decide on which method for highlighting words — italic or apostophe

The Boolean Model

One such ranking score is the Boolean model. The similarity of document d_j to query q is defined as follows (quoted from (Baeza-Yates and Ribeiro-Neto 2011, p.65))

$$sim(d_j, q) = \begin{cases} 1 & \text{if } \exists\ c(q) \mid c(q) = c(d_j) \\ 0 & \text{otherwise} \end{cases} \tag{6.3}$$

A 'conjunctive component' describes which terms occur in a document and which ones do not. E.g. for vocabulary $V = \{k_1, \ldots, k_t\}$, if the terms $[k_1, k_2, k_3]$ occur in document d_j then the conjunctive component would be $(1, 1, 1)$, or $(1, 0, 0)$ if only term k_1 appears in d_j.

$c(d)$	is the term conjunctive component for document d
$c(q)$	is the term conjunctive component for query q

Sometimes things are not quite black and white though and we need to weigh the importance of words somehow. The easiest way to do that is by looking at the frequency in which a word occurs.

The Vector Model

The vector model allows a more flexible scoring since it basically computes the various degrees of similarity between documents (taken from (Baeza-Yates and Ribeiro-Neto 2011, p.78)).

$$\begin{aligned} \vec{d_j} &= (w_{1,j}, w_{2,j}, \ldots, w_{t,j}) \\ \vec{q} &= (w_{1,q}, w_{2,q}, \ldots, w_{t,q}) \end{aligned} \tag{6.4}$$

Where t is the total number of terms in the index and $w_{i,j}$ is the TF-IDF weight for each component of the vector. The similarity between the document and the query vector is the cosine of θ.

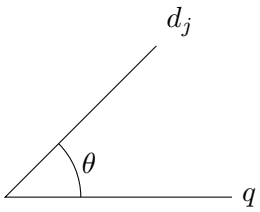


Figure 6.3: The Vector Model

$$
\begin{aligned}
sim(d_j, q) &= \frac{\vec{d_j} \cdot \vec{q}}{|\vec{d_j}| \times |\vec{q}|} \\
&= \frac{\sum_{i=1}^{t} w_{i,j} \times w_{i,q}}{\sqrt{\sum_{i=1}^{t} w_{i,j}^2} \times \sqrt{\sum_{i=1}^{t} w_{i,q}^2}}
\end{aligned}
\tag{6.5}
$$

Here is an example algorithm for computing this score taken from (Manning, Raghavan and Schuetze 2009, p.125).

```
1    CosineScore (q)
2      float Scores[N] = 0
3      for each d
4      do Initialise Length[d] to the length of document d
5      for each query term t
6      do calculate wt,q and fetch postings list for t
7        for each pair (d, tft,d) in postings list
8        do add wft,d to Scores[d]
9      Read the array Length[d]
10     for each d
11             do Divide Scores[d] by Length[d]
12             return Top K components of Scores[]
```

Code 6.1: Pseudo-code for computing vector scores

Where,

q	is the query
N	is the total number of documents
d	is a document
t	is a query term
wt_q	is the weight of the term in the query
tft_d	is the term frequency of t in d
wft_d	is the $tf - idf$ weight of t in d
K	is the number of results we want
$postingslist$	is the list of all (d, tft_d) for a given t.

There are several other common IR models that I won't discuss in detail here. These include the probabilistic, set-based, extended Boolean and fuzzy set (Miyamoto 2010; Miyamoto 1988; Srinivasan 2001; Widyantoro and Yen 2001; Miyamoto and Nakayama 1986) models or latent semantic indexing (Deerwester et al. 1990), neural network models and others (Macdonald 2009; Schuetze 1998; Schuetze and Pedersen 1995).

Architecture

Search Algorithms

6.1.3 Ranking

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 things like the Facebook 'like' button or the Google+ '+1' button which could be seen as direct user relevance feedback as well.

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[3] 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. Here is a survey of the most notable algorithms.

PageRank was developed in 1998 by Larry Page and Sergey Brin as part of their Google search engine and announced in their often cited paper (Brin and Page 1998a) and they further describe the algorithm here (Brin and Page 1998b). 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. The following formula for calculating a PageRank PR is taken from (Baeza-Yates and Ribeiro-Neto 2011, p.472).

[3] http://www.searchenginepeople.com/blog/incredible-google-bombs.html

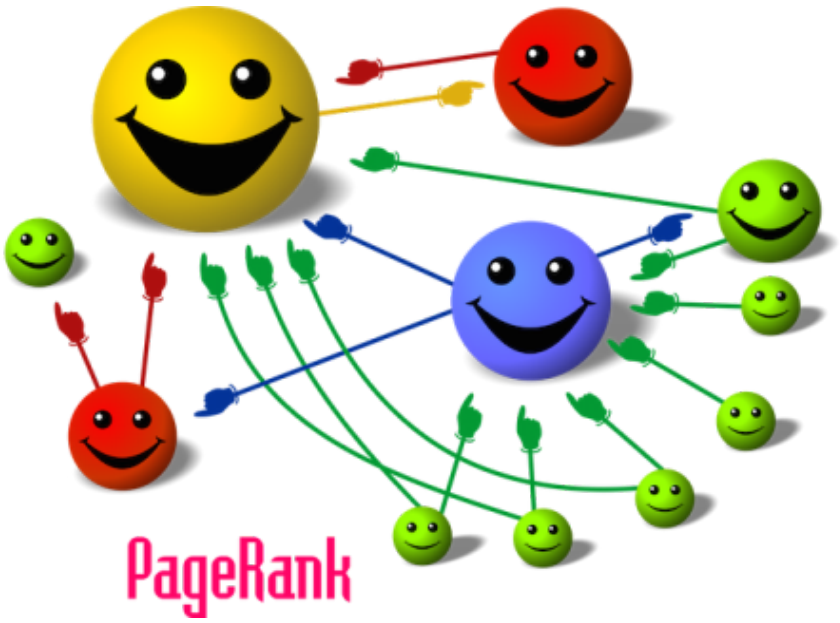


Figure 6.4: PageRank algorithm illustration from Wikipedia

$$PR(a) = \frac{q}{T} + (1 - q) \sum_{i=1}^{n} \frac{PR(p_i)}{L(p_i)} \tag{6.6}$$

Where,

$L(p)$	is the number of outgoing links of page p,
a	is the page we want to rank and is pointed to by pages p_1 to p_n,
T	is the total number of pages on the Web graph, and
q	is the is a parameter to be set by the system (typically 0.15) needed to deal with dead ends in the graph.

The HITS algorithm also works on the links between pages. It was first described by Kleinberg (Kleinberg 1999; Kleinberg et al. 1999, p.472) in 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. Again, the following formula is taken from (Baeza-Yates and Ribeiro-Neto 2011, p.471). S is the set of pages.

$$\begin{aligned} H(p) &= \sum_{u \in S | p \to u} A(u) \\ A(p) &= \sum_{v \in S | v \to p} H(v) \end{aligned} \tag{6.7}$$

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 in 2000 in (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. It was first described by De Bra in 1994 (De Bra and Post 1994a; De Bra and Post 1994b; De Bra, Houben et al. 1994). The basic concept here is that the search starts with the search query and a seed 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 in (Luo, Chen and Guo 2005) where the authors have simply added an extra parameter to allow 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 (Garcia-Molina, Pedersen and Gyongyi 2004), query modifications (Glover et al. 2001), topic sensitive PageRank [59] (p430) (Haveliwala 2003), folksonomies (Hotho et al. 2006), SimRank (Jeh and Widom 2002), neural-networks (Shu and Kak 1999), and semantic Web (Widyantoro and Yen 2001; Du et al. 2007; Ding et al. 2004; Kamps, Kaptein and Koolen 2010; Taye 2009).

6.1.4 Query Expansion and Relevance Feedback

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 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 analysis the top results provided by the 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 decided 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.

Challenges of Web Search

Other issues that arise when trying to search the World Wide Web are as follows ((Baeza-Yates and Ribeiro-Neto 2011, p.449)).

- 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, p.459). 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, p.459). 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).

Summary

IR refers to the retrieval of information from a collection. In terms of the Internet it is often called Web search. A Web search engine is divided into different components, being the crawler to build an index of the collection and a ranking algorithm which stands between the index and the user.

Different retrieval models exist including the Boolean and the Vector model. Other methods exist to make search results more accurate, including relevance feedback and query expansion.

Search quality is generally measured using the metrics of precision and recall but for Web search precision is more important and usually a metric called 'precision at n' is used for measurements.

Challenges are the size of the World Wide Web and ambiguous, unstructured nature of Web pages among others.

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, the ranking algorithm evaluates every page for relevance and returns them in order. There exist lots of different approaches on ranking, including PageRank and HITS (both analyse the link structure of the WWW), or more dynamic models like Fish search or genetic approaches.

6.2 Natural Language Processing

describe NLTK and the core functionality

Natural Language Tool Kit (NLTK) Python library[4].

PlaintextCorpusReader
: Reader for corpora that consist of plaintext documents. Paragraphs are assumed to be split using blank lines. Sentences and words can be tokenized using the default tokenizers, or by custom tokenizers specificed as parameters to the constructor.

Text
: A wrapper around a sequence of simple (string) tokens, which is intended to support initial exploration of texts (via the interactive console). Its methods perform a variety of analyses on the text's contexts (e.g., counting, concordancing, collocation discovery), and display the results.

index (word)
: Find the index of the first occurrence of the word in the text.

count (word)
: Count the number of times this word appears in the text.

6.2.1 Damerau-Levensthein

[4] http://www.nltk.org/

Damerau-Levensthein for clinamen! https://en.wikipedia.org/wiki/Damerau%E2%80%93Levenshtein_distance

The Damerau–Levenshtein distance between two strings a and b is given by $d_{a,b}(|a|, |b|)$ where:

$$d_{a,b}(i,j) = \begin{cases} \max(i,j) & \text{if } \min(i,j) = 0 \\ \min \begin{cases} d_{a,b}(i-1,j)+1 \\ d_{a,b}(i,j-1)+1 \\ d_{a,b}(i-1,j-1)+1_{a_i \neq b_j} \\ d_{a,b}(i-2,j-2)+1 \end{cases} & \text{if } i,j > 1 \text{ and } a_i = b_{j-1} \text{ and } a_{i-1} = b_j \\ \min \begin{cases} d_{a,b}(i-1,j)+1 \\ d_{a,b}(i,j-1)+1 \\ d_{a,b}(i-1,j-1)+1_{a_i \neq b_j} \end{cases} & \text{otherwise.} \end{cases} \tag{6.8}$$

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

Each recursive call matches one of the cases covered by the Damerau-Levenshtein distance:

$d_{a,b}(i-1,j)+1$ 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.

NLP blah blah blah. . .

Bird, S., Klein, E. and Loper, E., 2009. NLP with Python 1st ed., Sebasopol, CA: O'Reilly Media.(Bird, Klein and Loper 2009)

Manning, C., Raghavan, P. and Schuetze, H., 2008. Introduction to Information Retrieval 1st ed., Cambridge: Cambridge University Press.(Manning, Raghavan and Schuetze 2009)

Taken from (Jurafsky and Martin 2009), also known as:

- Speech and language processing
- Human language technology
- NLP
- Computational linguistics

- Speech recognition and synthesis

Goals of NLP are to get computers to perform useful tasks involving human language like:

- Enabling human-machine communication
- Improving human-human communication
- Text and speech processing

e.g. machine translation, automatic speech recognition, natural language understanding, word sense disambiguation, spelling correction, grammar checking. . .

Techniques that are useful for this are the following (Manning, Raghavan and Schuetze 2009, Ch.2).

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 $\rightarrow$ connect

Lemmatization
: returning dictionary form of a word, e.g. went $\rightarrow$ go

Regular Expressions

Used to specify text strings in text.

RE search requires a pattern that we want to search for and a corpus of texts to search through.

Errors can be false positives (FP) and false negatives (FN).

- Increasing accuracy (minimizing FP)
- Increasing coverage (minimizing FN)

RE's can be expressed as Finite-State Automata (FSA).

Language Models (LM)

Probabilities are based on counting things. Counting things in natural language is based on a corpus (pl corpora), a computer readable collection of text or speech.

Cats versus cat?

Same lemma but different wordforms.

- A lemma is a set of lexical forms that have the same stem. (e.g. go)
- A wordform 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)

The process of converting all words in a text to their lemma (e.g. goes $\rightarrow$ go) is called lemmatisation and the process of separating out all words in a text is called tokenisation or word segmentation.

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.

We want to compute the probability for $P(w|h)$ where w is a word and h is a history (the previous words). How many times occurred h followed by w divided by how many times occurred h?

$$P(w \mid h) = \frac{count(hw)}{count(h)} \tag{6.9}$$

Using the **chain rule of probability**:

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

Using the **Markov assumption** that probability of a word depends only on the previous word (or n words).

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

Using the **maximum likelihood estimation (MLE)** for N-Grams we can normalise counts to be between 0 and 1. C stands for count.

Maximum likelihood estimation (MLE)

$$P(w_n \mid w_{n-N+1}^{n-1}) = \frac{C(w_{n-N+1}^{n-1} w_n)}{C(w_{n-N+1}^{n-1})} \tag{6.12}$$

Usually instead of calculating the counts based on products we calculate them based on sums of logs.

So instead of $p_1 \times p_2 \times p_3 \times p_4 = \log p_1 + \log p_2 + \log p_3 + \log p_4$

Google offers its N-Gram data for free on:

- http://bit.ly/1baDXAW
- http://books.google.com/ngrams/
- http://www.speech.sri.com/projects/srilm/
- http://bit.ly/1G3ZJmX

Evaluating N-Grams

Extrinsic and intrinsic evaluation.

Extrinsic
: evaluate performance of a language model by embedding it into an independent application.

Intrinsic
: evaluate independent on any application, e.g. perplexity.

Perplexity

$$PP(W) = \sqrt[N]{\prod_{i=1}^{N} \frac{1}{P(w_i \mid w_{i-1})}} \tag{6.13}$$

Smoothing

Add-One: Laplace smoothing for bigrams

$$P_{Add-1}(w_i \mid w_{i-1}) = \frac{c(w_{i-1}, w_i) + 1}{c(w_{i-1}) + V} \tag{6.14}$$

Adjusted count

$$c_i^* = (c_i + 1)\frac{N}{N + V} \tag{6.15}$$

Add-1 smoothing is ok for text categorisation but not so much for language modelling.

Most commonly used is Kneser-Ney extended interpolated.

For very large N-grams like the Web "Stupid Backoff" is used.

Good Turing Discounting

N_c is the frequency of frequency c.

$$c^* = (c + 1)\frac{N_{c+1}}{N_c} \tag{6.16}$$

Naive Bayes

[3] page 234. . .

> (Wikipedia): A naive Bayes classifier is a simple probabilistic classifier based on applying Bayes' theorem with strong (naive) independence assumptions. A more descriptive term for the underlying probability model would be "independent feature model".

Maximum Entropy Models (MaxEnt)

Page 227 . . . in [1]

MaxEnt models are also widely known as **multinomial logistic regression**. They are used for sequence classification, e.g. part-of-speech tagging. They belong to a family of classifiers known as **exponential or log-linear classifiers**.

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. A probabilistic classifier also gives the probability of the observation being in that class; it gives a probability distribution over all classes.

MaxEnt works by extracting some set of features from the input, combining them linearly (meaning that each feature is multiplied by a weight and then added up), and then using this sum as an exponent. Formula below shows how to calculate the probability of class c given an observed datum (a given data point) d and λ

is a weight that is assigned to feature f. Taking the exponent makes the result always positive. Dividing by the Sum of that for all classes makes it a probability.

$$P(c \mid d, \lambda) = \frac{\exp \sum_i \lambda_i f_i(c, d)}{\sum_{c\prime} \exp \sum_i \lambda_i f_i(c\prime, d)} \tag{6.17}$$

To get the single best class with the highest probability we need to compute the following.

$$\hat{c} = \underset{c \in C}{\text{argmax}}\ P(c \mid d, \lambda) \tag{6.18}$$

PERSON	LOCATION	DRUG
In Québec	In Québec	In Québec
0	1.8 + -0.6	0.3

Table 6.1: MaxEnt Example table

Features:

$f1(c, d) \equiv [\ c = \text{LOCATION}\ \wedge\ w - 1 = \text{“in”} \wedge\ \text{isCapitalized}(w)]$
$f2(c, d) \equiv [\ c = \text{LOCATION}\ \wedge\ \text{hasAccentedLatinChar}(w)]$
$f3(c, d) \equiv [\ c = \text{DRUG}\ \wedge\ \text{ends}(w, \text{“c”})]$

$P(\text{LOCATION} \mid \text{in Québec}) = \frac{e^{1.8} e^{-0.6}}{e^{1.8} e^{-0.6} + e^{0.3} + e^{0}} = 0.586$
$P(\text{DRUG} \mid \text{in Québec}) = \frac{e^{0.3}}{e^{1.8} e^{-0.6} + e^{0.3} + e^{0}} = 0.238$
$P(\text{PERSON} \mid \text{in Québec}) = \frac{e^{0}}{e^{1.8} e^{-0.6} + e^{0.3} + e^{0}} = 0.176$

The empirical expectation is the sum of all occurrences where a feature is true for one of our observed datums.

$$\textit{empirical}\ E(f_i) = \sum_{(c,d)\ \in\ \textit{observed}(C,D)} f_i(c, d) \tag{6.19}$$

Evaluation

$$Precision = \frac{\text{number of correctly labeled}}{\text{total number of extracted}} \tag{6.20}$$

$$Recall = \frac{\text{number of correctly labeled}}{\text{total number of gold}} \tag{6.21}$$

$$F_1 = \frac{2PR}{P+R} \tag{6.22}$$

Information Extraction

[1] Chapter 22, p 759...

"The process of information extraction (IE), also called text analytics, turns the unstructured information embedded in texts into structured data."

IE involves named entity recognition (NER), relation detection and classification, event detection and classification and temporal analysis.

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.

Example (first sentence in Faustroll):

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

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

Gazetteers (lists of place or person names for example) can help with the detection of these named entities.

Part of Speech Tagging

Parts of speech (POS) are lexical tags for describing the different elements of a sentence. The eight main parts-of-speech (originating from ca. 100 B.C.) are noun, verb, pronoun, preposition, adverb, conjunction, participle and article. Wikipedia:

Noun
: any abstract or concrete entity; a person (police officer, Michael), place (coastline, London), thing (necktie, television), idea (happiness), or quality (bravery)

Pronoun
: any substitute for a noun or noun phrase

Adjective
: any qualifier of a noun

Verb
: any action (walk), occurrence (happen), or state of being (be)

Adverb
: any qualifier of an adjective, verb, or other adverb

Preposition
: any establisher of relation and syntactic context

Conjunction
: any syntactic connector

Interjection
: any emotional greeting (or 'exclamation')

Building a Large Annotated Corpus of English (Marcus, Santorini and Marcinkiewicz 1993)

There exist other sets of tags, like the Penn Treebank with divides those 8 tags into a total of 45, for example 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, etc.

The process of adding tags to the words of a text is called parts-of-speech tagging or just tagging. This usually is done together with the tokenisation of the text.

Example (first sentence in Faustroll):

> 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]$^{\text{TIME}}$,/, Pursuant/JJ to/IN article/NN [819/CD]$^{\text{NUMBER}}$ of/IN the/DT [Code/NN of/IN Civil/NNP Procedure/NNP]$^{\text{DOCUMENT}}$ and/CC at/IN the/DT request/NN of/IN [M./NN and/CC Mme./NN Bonhomme/NNP (/(Jacques/[redacted]NNP)/)]$^{\text{PERSON}}$,/, proprietors/NNS of/IN a/DT house/NN situate/JJ at/IN [Paris/NNP,/, 100/CD bis/NN,/, rue/NN Richer/NNP]$^{\text{LOCATION}}$,/, 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]$^{\text{FACILITY}}$ of/IN [Q/NNP borough/NN]$^{\text{LOCATION}}$./.

$$t_1^n = \underset{t_1^n}{\text{argmax}}\ P(w_1^n \mid t_1^n) P(t_1^n) \tag{6.23}$$

$$P(t_i \mid t_{i-1}) = \frac{C(t_{i-1}, t_i)}{C(t_{i-1})} \tag{6.24}$$

For example: the probability of getting a common noun after a determiner is:

$$P(\text{NN} \mid \text{DT}) = \frac{C(\text{DT}, \text{NN})}{C(\text{DT})} = \frac{56,509}{116,454} = 0.49 \tag{6.25}$$

Given that there are $116,454$ occurrences of DT in the corpus and of these $56,509$ occurrences where a NN follows after the DT.

$$P(\text{is} \mid \text{VBZ}) = \frac{C(\text{VBZ}, \text{is})}{C(\text{VBZ})} = \frac{10,073}{21,627} = 0.47 \tag{6.26}$$

Or the probability of a third person singular verb being 'is' is 0.47.

Parsing

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 the given sentence.

Grammar

A language is modelled using a grammar, specifically a Context-Free-Grammar or CFG. Such a grammar normally consists or rules and a lexicon. For example a rule could be NP $\rightarrow$ 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 $\rightarrow$ a, Det $\rightarrow$ the, Noun $\rightarrow$ book. This grammar would let us form the noun phrases 'the book' and 'a book' only. The two parse trees would then look like this:

The parse tree for the previous example sentence from Faustroll is shown below, in horizontal for convenience.

```
(ROOT
```

Figure 6.5: Grammars

```
(S
  (PP (IN In)
    (NP (DT this) (NN year) (NNPS Eighteen) (NNP Hundred)
      (CC and)
      (NNP Ninety-eight)))
  (, ,)% chktex 26
  (NP
    (NP (DT the) (JJ Eighth) (NN day))
    (PP (IN of)
      (NP (NNP February) (, ,) (NNP Pursuant)))% chktex 26
    (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-))
    (, ,)% chktex 26
    (NP
      (NP (NNS proprietors))
```

```
      (PP (IN of)
        (NP
          (NP (DT a) (NN house) (NN situate))
          (PP (IN at)
            (NP (NNP Paris))))))
    (, ,)% chktex 26
    (NP (CD 100) (NN bis))
    (, ,))% chktex 26
  (VP (VBP rue)
    (NP
      (NP (NNP Richer))
      (, ,)% chktex 26
      (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))))
  (. .)))% chktex 26
```

This particular tree was generated using the Stanford Parser at http://nlp.stanford.edu:8080/parser/index.jsp. Given the rather complicated nature of the words and sentence structure, some of the labels might be wrong.

6.3 Linguistics / WordNet

Here's my hypernym term. holonym hypernym

I looked into linguistics for the purpose of patadata. This section definitely needs some expanding. Some concepts that might be relevant include (taken from Wikipedia):

Hyponym
– subcategory of something

Hypernym
– top category of some things

Meronym
– member of something (e.g. finger is meronym to hand, wheel to car)

Holonym
– e.g. tree is holonym of bark, trunk, limb... opposite of meronym

Troponym
– presence of "manner" between things (e.g. to traipse and to mince = walk a certain way)

Homonym
– same spelling but different sound and meaning = heteronym – same sound but different spelling = heterography – same meaning = synonym

Antonym
– opposite

Metonym
– figure of speech (e.g. Hollywood for American movies) not quite metaphor but similar.

I need to find REFERENCES for this section.

6.4 Algorithm Formalisation

Algorithm Classification

By implementation:
- Recursive/iterative
- Logical
- Serial/parallel/distributed
- Deterministic/non-deterministic
- Exact/approximate
- Quantum

By design paradigm:
- Brute-force/exhaustive search
- Divide and conquer
- Dynamic
- Greedy
- Linear
- Reduction
- Search and enumeration

By field of study:
- Search
- Sorting

- Merge
- Numerical
- Graph
- String
- Computational geometrics
- Combinatorial
- Medical
- Machine learning
- Cryptography
- Data compression
- Parsing

By complexity:

- Big-O-Notation

High-Level Description

in prose, ignoring implementation details.

Implementation Description

in prose, describing implementation in detail.

Formal description

lowest level, most detailed.

$D = \{d_1, \ldots, d_n\}$	is the set of documents
$Q = \{q_1, \ldots, q_n\}$	is the set of queries
$q = \{t_1, \ldots, t_n\}$	is the set of query terms
$V = \{v_1, \ldots, v_t\}$	is the set of all distinct index terms in a document collection (the Vocabul
$R(q_i, d_j)$	is the ranking function, where $q_i \in Q$ and $d_j \in D$
N	is the total number of documents
$w_{t,q}$	is the weight of the term in the query
$tf_{t,d}$	is the term frequency of t in d
$wf_{t,d}$	is the tf-idf weight of t in d
P_t	is the postings list of all (d, $tf_{t,d}$) for a given t

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 IR is evaluated using two metrics known as precision and recall. 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.

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

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

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
7.1 we managed to retrieve. This can be easily visualised as follows.

place footnotetext properly

Precision is typically more important than recall in web search while it is the other way around in a database search system maybe. The mean average precision value (MAP) can be calculated following this formula (Baeza-Yates and Ribeiro-Neto 2011, p.141):

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

Where R_i is the set of relevant documents for query q_i.

But for many web searches is it 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 5 to 30

[1] Image taken from Wikimedia Commons: https://upload.wikimedia.org/wikipedia/commons/2/26/Precisionrecall.svg

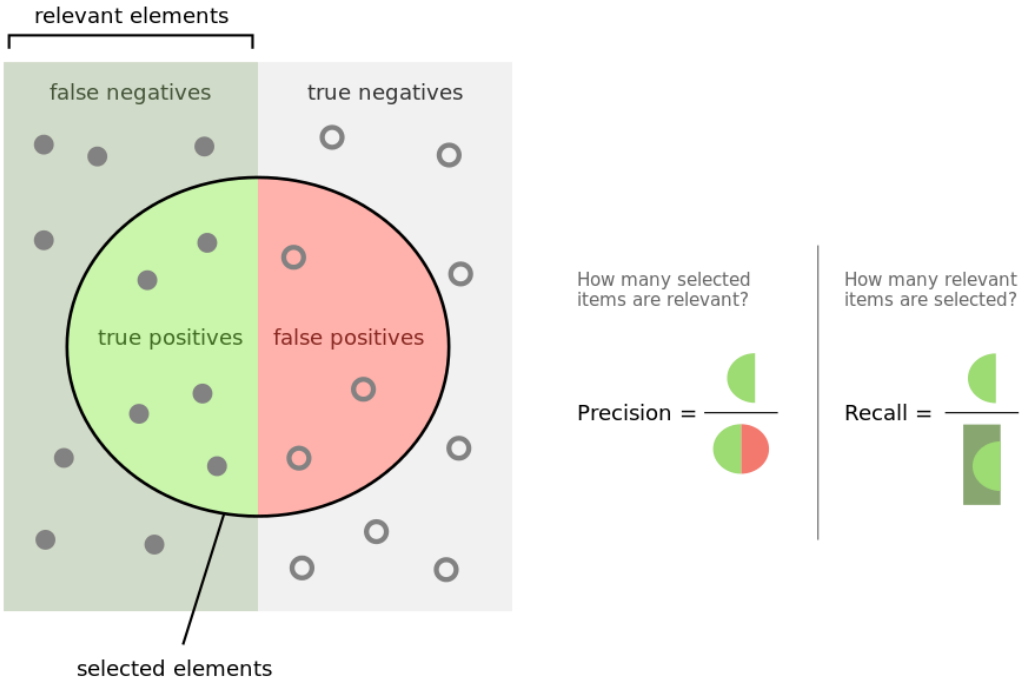


Figure 7.1: Precision and Recall[1]

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, p.140). So for example this could be P@5 or P@10 or P@20. For example, 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) is a conference that provides large test sets of data to participants and lets them compare results. They have specific test sets for web search comprised of crawls of *.gov* web pages for example, but unfortunately they have to be paid for to get a copy.[2]

There are certain other factors that can be or need to be evaluated when looking at a complete search system, as shown below.

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

[2] http://ir.dcs.gla.ac.uk/test_collections/

- 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 is further discussed in chapter 6.

anything else about evaluating search/creative search?

7.2 Evaluating Creative Computers

bridge over from traditional search evaluation to general creative computing

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 Mayer 1999 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 Schmidhuber 2006b)
- 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' Pease and Colton 2011.

write better lit review for this section

add francois stuff

check ICCC conference 2014 and 2015

Because creativity infused computing has only emerged in the last few decades or so, its evaluation is not well defined. Discussions from Computational Creativity (CompC) for example 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).

Pease, Winterstein and Colton have argued that creativity may be seen as 'output minus input.' (Pease, Winterstein and Colton 2001, p.2). 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.

> 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, p.3)

They also suggest that all creative products must be 'novel and valuable' (2001, p.1) 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' (Ritchie 2007, p.72). He identifies three main criteria for creativity as 'novelty, quality and typicality' (2007, p.72-73), 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' (2007, p.73) (see also 2001). He proposes several evaluation criteria which fall under the following categories: (Ritchie 2007, p.91-92) basic success, un-

restrained 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, p.7).

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, p.6), and that 'two kinds of evaluation are relevant; the evaluation of the item, and evaluation of the processes used to generate it.' (2001, p.7). 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'. 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.' (2008b, p.5) + (2008a, p.5). 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 2008b, p.5) + (Colton 2008a, p.5)

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

Davide Piffer suggests that there are three dimensions of human creativity that can be measured, namely 'novelty, usefulness/appropriateness and impact/influence' (2012, p.258-259). As an example of how this applies to measuring a person's creativity he proposes 'citation counts' (Piffer 2012, p.261). While this idea works well for measuring scientific creativity maybe, he does not explain how this would apply to a visual artist for example[3].

[3] http://www.artfacts.net seems to provide just that though.

Anna Jordanous proposed 14 key components of creativity (which she calls an 'ontology of creativity') (2012, p.104-120), from a linguistic analysis of creativity literature which identified words that appeared significantly more often in discussions of creativity compared to unrelated topics. (2012, p.120).

> 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.
> (A. K. Jordanous and Keller 2012, p.120)

The 14 components Jordanous collated are: (2012, p.118-120)

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

Anna Jordanous found that 'evaluation of computational creativity is not being performed in a systematic or standard way' (A. K. Jordanous 2011, p.2) and proposed 'Standardised Procedure for Evaluating Creative Systems (SPECS)' (A. K. Jordanous 2012, p.137-140):

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. A. K. Jordanous 2012

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

compare to CC research methodolody

Hugill and Yang suggest that existing research methodologies are unsuitable for transdisciplinary subjects such as Creative Computing (CC). The following is an example of a possible CC research methodology they propose as a starting point (Hugill and Yang 2013, p.17): 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, p.17) namely, resist standardisation, perpetual novelty, continuous user interaction and combinational, exploratory and or transformational.

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 usabilty, she says (Linda Candy 2012, p.23), 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.' (Linda Candy 2012, p.7)

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

Candy introduces the Multi-dimensional Model of Creativity and Evaluation (MMCE) with four main elements of people, process, product and context (Linda Candy 2012, p.11) similar to some of the models of creativity we have seen in chapter 5.

7.2 § 5

Candy proposes the the following values or criterias for measurement (Linda Candy 2012).

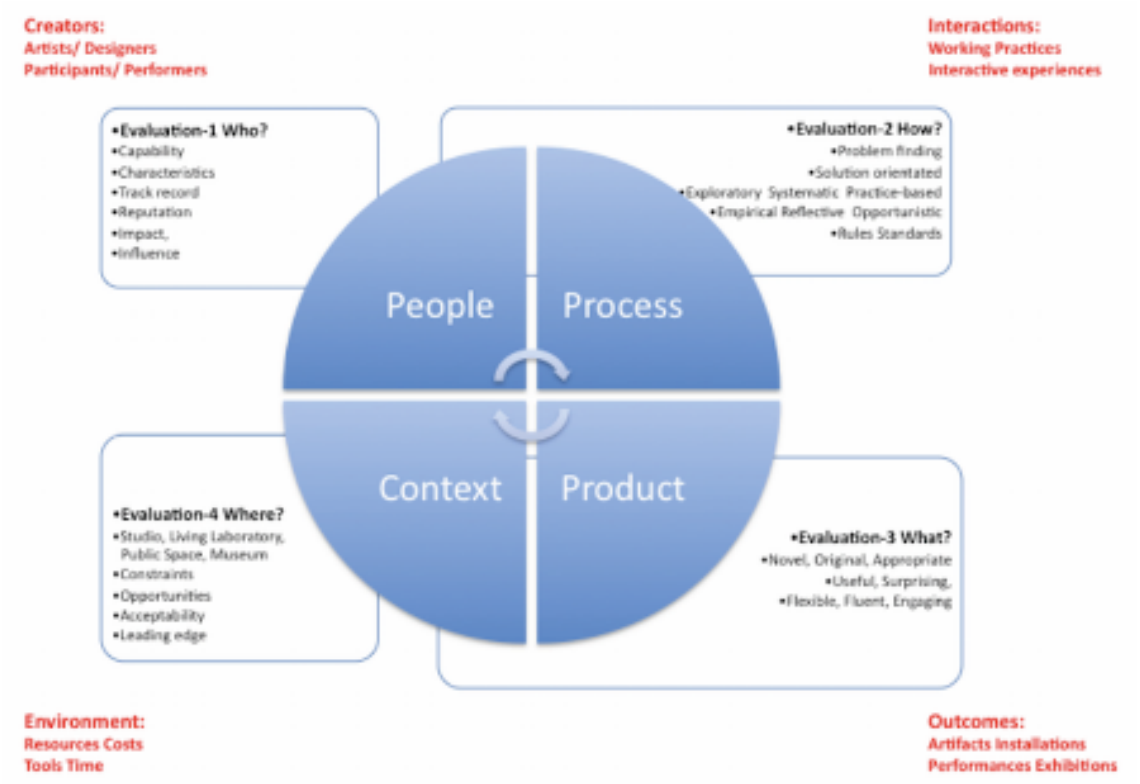


Figure 7.2: Linda Candy's Multi-dimensional Model of Creativity and Evaluation

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, knowledgable, 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)

Furthermore it is interesting to know the judging criteria for the Prix Ars Electronica, an international competition for Cyber Arts to be aesthetics, originality, excellence of execution, compelling conception and innovation in technique of

the presentation (cited in Linda Candy 2012, p.18).

rewrite

Geraint Wiggins introduced a formal notation and set of rules for the description, analysis and comparison of creative systems (2006) which is largely based on Boden's theory of creativity (2003). The framework uses three criteria for measuring creativity: 'relevance, acceptability and quality'.

Geraint Wiggins previously described a formal notation and set of rules for the description, analysis and comparison of creative systems in the form of his Creative Search Framework (CSF) (Wiggins 2006) which was largely based on Margaret Boden's theory of creativity (Boden 2003). Graeme Ritchie then contributed to this framework (Ritchie 2012) and we will base our work on his revised version of this framework. The CSF provides a formal description for Boden's concepts of exploratory and transformational creativity. Wiggins's R–transformation and T–transformation 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 information retrieval sense should not be mistaken with what is discussed here. 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. 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, as suggested in the next section.

Uncreativity Wiggins explained various situations of creativity not taking place (uninspiration and aberration) in terms of his framework. 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.

7.1

finish section

Researchers at IBM have fallen into the trap of over-simplifying creativity and computational creativity (Varshney et al. 2013). First they define machine creativity to be a system that produces 'novel, useful and quality' artefacts.

novelty = Bayesian surprise (Baldi and Itti 2010)

In a diagram on the difference between 'computational creativity' and 'cognitive informatics and computing' they describe the former as consisting of:

- Planning how to make

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	$V_\alpha(B) = \emptyset$	

Table 7.1: Wiggins' uncreative concepts in Ritchie's notation

- Idea generation
- Defining creativity
- Curiosity
- Assessment of creative artefacts

INTERLUDE I

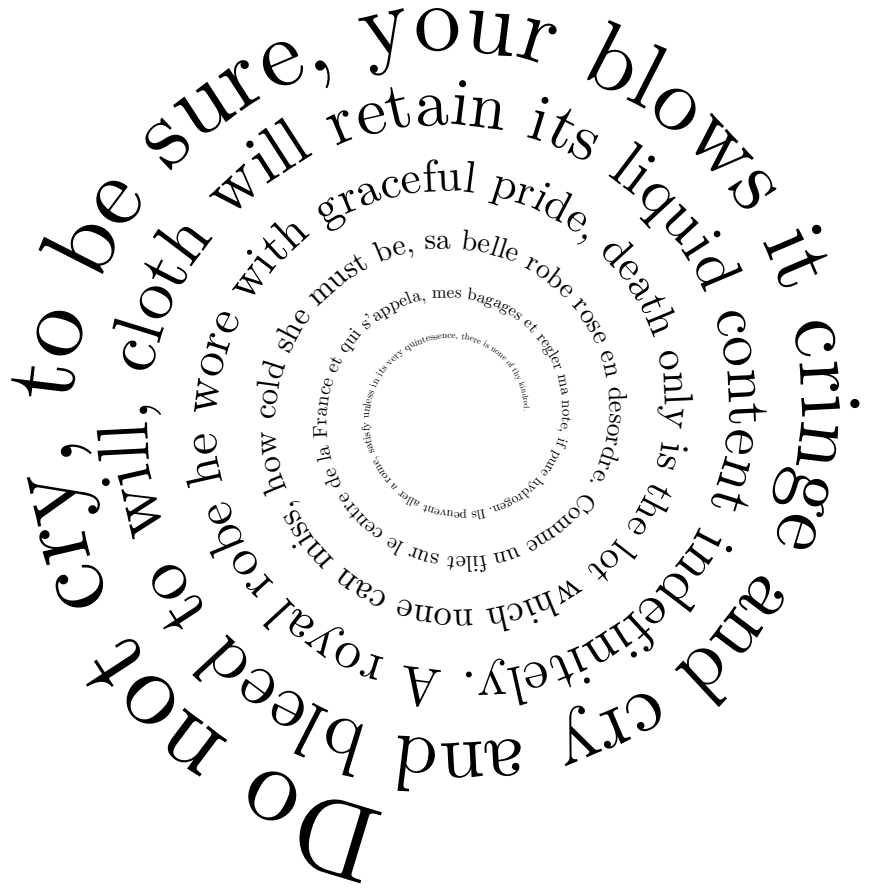
(...) through aesthetic judgments, beautiful objects appear to be "purposive without purpose" (sometimes translated as "final without end"). An object's purpose is the concept according to which it was made (the concept of a vegetable soup in the mind of the cook, for example); an object is purposive 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 2015, ch.2a)

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)

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)

Part III

THE CΘRE: TΣCHNΘ-LΘGIC



Foundations

8

My soul with the bare supposition of their possibility,
if you will go to bed at once,
and that I begg'd the charity of them,
noir corset velu des mouches éclatantes.

We can then start at once,
and charity and why,
and by faith formed in charity to cleave unto him,
or in any of those unmentionable graces which are now.

J'ai été en relation avec des hommes qui ont été vertueux,
which is the basis of our holy religion,
j'invoque dans le commencement de cet ouvrage.

Removed her girdle,
vous a laissé voir la couleur de son corset,
start from the goal.

This chapter discusses some of the ideas introduced in chapters 4 to 7 and relates them to each other. The insights gained from these comparisons form an essential part of my argumentation in this thesis.[1]

8.1 Exploring Creativity

- Associative and bisociative thinking
- Creative triptych (humour, discovery, art)

8.1.1 General Models

The Creativity chapter introduced various models of creativity. Here, I want to discuss some of their similarities and differences.

4 P Model
Mel Rhodes identified four common themes of creativity (Person, Process, Press, Products), which he termed the '4 P's' of creativity (Rhodes 1961).

4 Aspects
Ross Mooney independently identified four aspects of creativity in 1963 which he called Environment, Person, Process and Product (as cited in Sternberg 1999).

P and H Model
Margaret Boden defined three types of creativity: combinational, exploratory and transformational and two different 'levels' P and H creativity (Boden 2003).

[1] More specific details about the Evaluation chapter can be found later on in chapter 9 (Interpretation).

4 C Model

James Kaufman and Ronald Beghetto defined the '4 C Model' of creativity. They are Big-C, Pro-c, Little-c and Mini-c (Kaufman and Beghetto 2009).

add bipin indurkhya

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

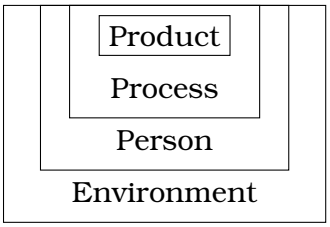


Figure 8.1: 4 aspects of creativity

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. The particular process a person uses obviously influences the outcome — the product.

Boden and Kaufman overlap in a less obvious way. Boden's book on 'the creative mind' was first published in 1990, while Kaufman and Beghetto published their paper 'Beyond Big and Little' in 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 level). This also ties in very neatly with the idea of subjectivity and objectivity as table 8.1 8.1 shows.

Arguably, the Pro-c should perhaps be called Pro-C instead, 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.

Quite recently, Anna Jordanous related the idea of the '4 P's' to the discipline of computational creativity (A. Jordanous 2015).

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

Table 8.1: Comparison of the 4 C Model vs. P and H Creativity vs. Subjectivity and Objectivity

8.1.2 Creative Process

The creative process has been subject to discussion and analysis as if it was 'the holy grail' of creativity.

4 Stage Model

Henri Poincaré suggested a '4 Stage Model' (formulated by Graham Wallas in 1926). The stages are: preparation, incubation, illumination and verification (Poincare 2001; Wallas 1926).

Problem Solving

George Pólya came up with a description of the 'problem solving' process (Polya 1957).

add comb, trans, expl.? and koestler?

8.2 8.1 Looking at table 8.2 highlights the similarities of the two models above ((a) and (b)) 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 8.2 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.

Giving ORDER to the 4 P model?!

8.1.3 Creative Disciplines

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

Table 8.2: Comparison of 4 Step Model vs 4 P Model vs Problem Solving

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.

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 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, 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' Hugill and Yang 2013. 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' Colton and Wiggins 2012. They further explain that it models, simulates, replicates or enhances human creativity using a computer.

Since creativity is studied in many different disciplines, projects such as this PhD research can be hard to categorise. As I have already discussed, this project

§ 3 is transdisciplinary and perhaps should be considered not part of one specific creative discipline but of many and beyond. Pure computer science, cognitive science or artificial intelligence clearly don't fit the bill. Recently however disciplines such as 'creative computing', 'computational creativity' and 'digital
§ 5 humanities' have emerged.

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 (Hugill and Yang 2013).

Computational Creativity
: model, simulate, replicate or enhance human creativity using a computer (Colton and Wiggins 2012).

Digital Humanities
: collaboration, transdisciplinarity and an engagement with computing and humanities (Burdick et al. 2012).

These three disciplines share the theme of combining creativity with computing, but there are also differences. Creative computing for example is about doing computations in a creative way, while computational creativity is about achieving creativity through computation (Hugill 2013).

Artistic creation	Software engineering	Layer of abstraction
Motivation	User requirements	Abstract
Formulation	System design	Less abstract
Creation	Coding	Less concrete
Dissemination/revision	Operation/evolution	Concrete

Table 8.3: Comparison of Artistic Creation vs Software Engineering vs Abstraction by (Hugill and Yang 2013)

▦ 8.3 Table 8.3 is taken directly from Hugill and Yang (Hugill and Yang 2013). They use the comparison to software engineering and four layers of abstraction as the basis of their definition of the creative computing process, i.e. motivation, ideation, implementation and operation. I believe their observation that artistic creation and software engineering both represent a move from the abstract to the concerete is critical.

▦ 8.4 Table 8.4 shows the four steps of creative computing defined by Andrew Hugill

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 8.4: Comparison of Creative Computing vs Digital Humanities vs Computational Creativity vs Computer Ethics

and Hongji Yang (Hugill and Yang 2013) and lines them up with corresponding activities in DH (Burdick et al. 2012), CompC (Colton and Wiggins 2012) and Computer Ethics (Stahl, Jirotka and G. Eden 2013).

Layer of Abstraction	**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

Table 8.5: Comparison of Creative Process vs Creative Disciplines

8.5 8.2 The spectrum from abstract to concrete as shown in table 8.5 relates to the creative process models we have seen as well as the 4 P Model.

Abstract

Preparation, Understand, Person, Motivation, User Requirements, Design, Intentionality, Purpose

Less Abstract

Incubation, Plan, Environment, Formulation, System Design, Ideation, Curation, Framing, People

Less Concrete

Illumination, Carry Out, Process, Creation, Coding, Implementation, Computation

Concrete

Verification, Look Back, Product, Dissemination, Operation, Prototyping

Abstract to Concrete is more about the practical process of artistic creation, not the conceptual development of a creative idea. That process is more of a move from concrete to abstract (known to unknown) using methods such as combinatorial, transformative and exploratory.

add this to intro

8.2 Relating Pataphysics

Text shown with a left bar is taken from (Hugill, Yang et al. 2013).

rewrite

Combining computing with pataphysics seems impossible.

- Polymorphism (generalisations) oppose particularity.
- Precision (bugs) 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.

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 appropriate framework within which to encourage and enable creative thinking and operations.

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

> 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)

check quote location

Boden 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, p.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 2007).

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 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.

ref

Later writings develop these ideas in more detail. La Littérature Potentielle **Oulipo1973**, is divided into several sections, dealing with clusters of methods, that include: anoulipisms (analytical oulipisms, such as combinatorial literature); use of preexisting 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, p.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, p.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 (Sternberg 1999; Sternberg 1999), for example, 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 (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 (Newell, Shaw and Simon 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. 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.

8.6 Table 8.6 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. 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.

8.3 Explaining Concepts

Patalgorithms

Pataphysical algorithms.

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

Table 8.6: Creativity vs Pataphysics

Pataphysicalisation
: Applying pataphysical transformations to data.

Patadata
: Data which has been pataphysicalised.

Patasaurus
: A thesaurus for patadata.

Patametric Index
: Patadata index.

Pranking
: Pataphysical ranking.

rewrite sections here, integrate into other chapters

Patalgorithms

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 pataphys-

8.2 icalisation process forms a central component of our system and influences all areas of the search tool.

redraw figure

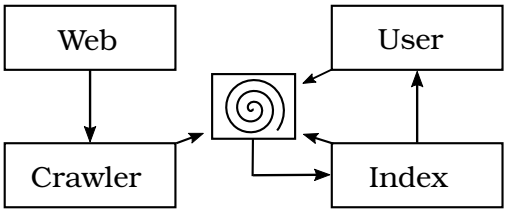


Figure 8.2: Pata centrala2

Pataphylicalisation

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) – applying pataphysical transformations
Pataphysicalisation
(noun) – the process of pataphysicalising
Patadata
(noun) – any data which has been pataphysicalised

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 8.3 same time. We can consider the possibility of a 'patametric index', rather than a parametric index or a 'patasaurus' (pataphysical thesaurus/ontology).

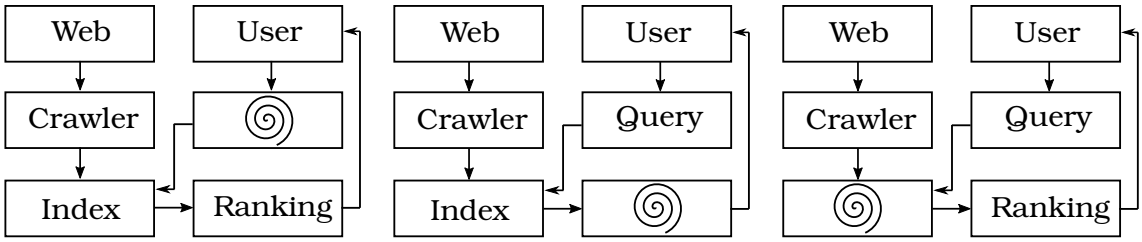


Figure 8.3: Pataphysicalisation

> 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, p.9)

Patadata

The idea of patadata is derived from the idea below:
Physics → Metaphysics → Pataphysics
Data → Metadata → Patadata

Patadata will allow us to engage with digital knowledge in a more creative way even. If metadata helps us organise information semantically then patadata is for organising information pataphysically. If metadata is objective then patadata is 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, p.16) 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.

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)]

Pranking

In 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 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 1998b) 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 they use over 200 signals including PageRank and they personalise 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.

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.

§ ?? Parts of this chapter were published in (Raczinski and Everitt 2016).

rewrite. Change all 'we' s to I?

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 Open Frameworks[2], 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.

[1] https://processing.org/ — a Java-based 'flexible software sketchbook and a language for learning how to code within the context of the visual arts'.

[2] http://openframeworks.cc/ — 'an open source C++ toolkit designed to assist the creative process by providing a simple and intuitive framework for experimentation'.

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.

Although using computers to generate creative work has its foundations in the 1950s (Linda Candy and Ernest Edmonds 2011), John Maeda's Design By Numbers (Maeda 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 (Chalmers 1996)) 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.

refer to the title of the paper here

9.1 Problems

Creativity and the subjective properties associated with it, lack a universally ac-
§ 5 cepted definition as I have shown in the Creativity chapter. 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 frame-
§ 7 works 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, 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 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 constraints 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!

§ 7

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

- 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)

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' (Indurkhya 1997). These leaps ***appear*** creative only because we are athropomorphising 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' (Brown 2011), as if there were conscious intent behind the code generating such activity in Tresset's sketching bot ***Paul***.

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. (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.' (Cohen 1999), 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. (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*** (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). 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***?

add a bit more about human brain

9.1.3 Mimicry

Current evaluation methodologies in creative computing disciplines have con-
§ 7 centrated 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.

True AI and Computational 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 John Searle's Chinese Room Argument (Searle 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 David Chalmers—***Zombies*** (Chalmers 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' (Hofstadter 1981). 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.

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 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 approximations.

Edsger Dijkstra pointed out that computer science is infantalised (Dijkstra 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

add more

9.1.5 Abstraction

see formal maths equations which are very hard to apply in real life eg wiggins cfs, bayensian surprise, precision recall...

9.1.6 Incompleteness

(Varshney et al. 2013) novelty = Bayesian surprise (Baldi and Itti 2010) which is:

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

> Computational creativity applies technology to assist humans in thinking outside the box and expanding their exploration boundaries.
> (Varshney et al. 2013)[4]

9.1.7 Undefinitions

Anna Jordanous found that 'evaluation of computational creativity is not being performed in a systematic or standard way' (A. K. Jordanous 2011), which further confuses the problem of objective evaluation. To remedy this she proposed
§ 7.2 'SPECS' (see chapter 7 for more details) (A. K. Jordanous 2012):

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. (A. K. Jordanous 2012, p.67)

We need a 'clearer definition of creativity' (Mayer 1999, p.459), with 'criteria and measures [for evaluation] that are situated and domain specific.' (Linda Candy 2012, p.7)

> (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, p.258)

[4] http://research.ibm.com/cognitive-computing/computational-creativity.shtml

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

§ 7 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—'Any objective theory of what is good art must take the subjective observer as a parameter.'(Schmidhuber 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.

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, (§ 9.2.1)
2. a set of criteria to be considered using the scales above, (§ 9.2.2)
3. a combined framework for evaluation. (§ 9.2.3)

9.2.1 Subjective Evaluation Criteria

§ 7.2 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:

From this I derive the following ***Creativity Criteria*** — 3 key criteria of creativity in relation to 4 major factors — novelty, value, quality and purpose → spatial, temporal and ephemeral.

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

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

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

Table 9.1: Subjective Scales for Creativity

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

9.2.2 Objective Evaluation Constraints

§ 5 In tribute to the many kinds of '4 P' models out there and combining it with the 'four P's' of Stahl's computer ethics framework.

> 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, p.203)

I propose a set of evaluation constraints called (*surprise surprise*) the '5 P Model' — product, process, people, place and purpose.

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

Table 9.2: Objective Criteria of Creativity

The '5 P's'—**Product, Process, Purpose, Person, Place**—are all components of any creative artefact (see table 9.2).

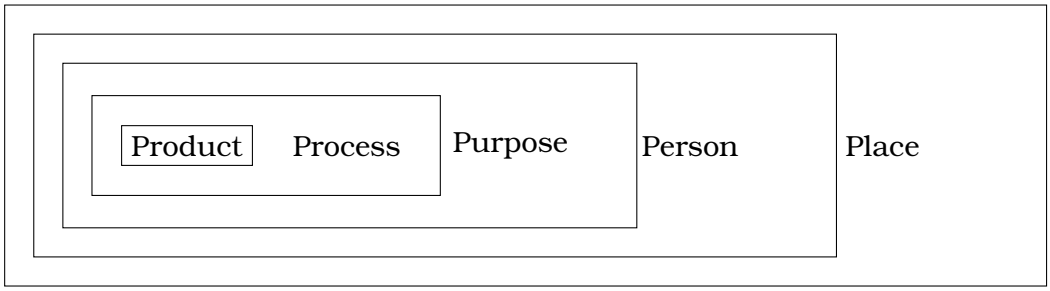


Figure 9.1: 5 P Model

Why is the purpose important?
Interpreting or Measuring?
Maybe we should not be looking for metrics but rather guidelines for interpretations of creativity.

9.2.3 Combined Framework

The **criteria** listed in table 9.2 should be considered objectively, while the **scales** 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 sections.

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.

explain matrix!

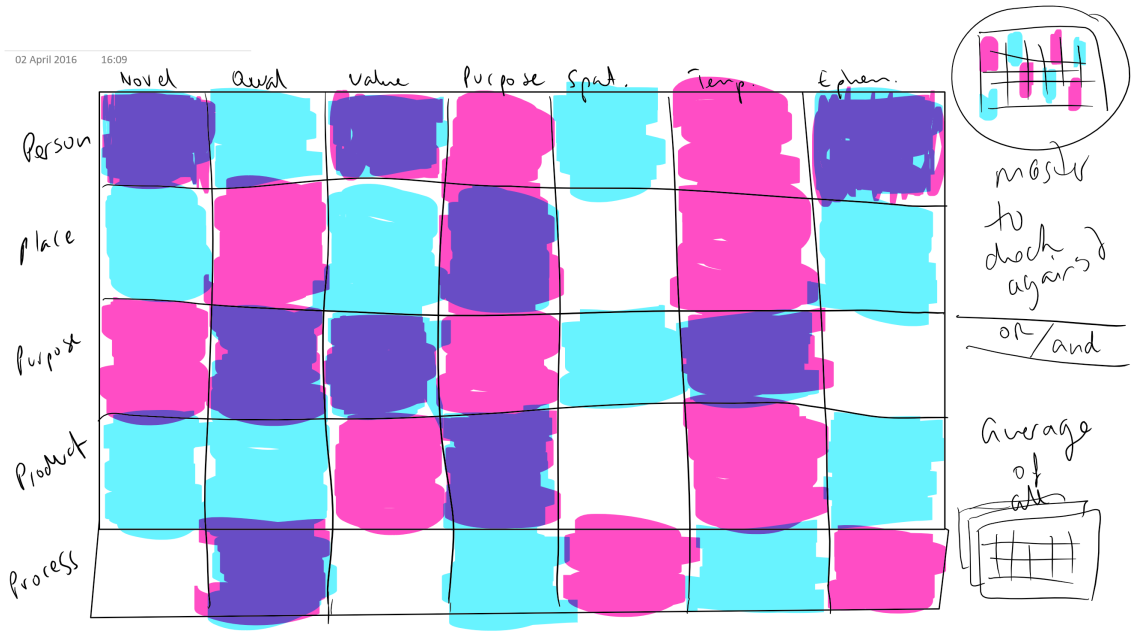


Figure 9.2: Creative Evaluation Matrix

This 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

finish this example

A complete

Step 1 – Create master matrix to measure against.
Step 2 – Fill matrix, potentially by several judges.
Step 3 – Check against criteria from step 1.

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).

apply my own framework to my own product and discuss results

apply my own framework to my own framework and discuss results—recursion

9.3 Ejaculation[5]

im evaluating the website - not the project !!!!!!!

change font size, capitalisation and dashes

The website `pata.physics.wtf` is supposed to be an example of AMC.

It seems appropriate to start the critical evaluation of the artefact created as part of this research project with an application of my own framework as suggested

[5] Autocorrect or nottocorrect—that is the question. . .

in chapter 9. I will do this in two ways. FIrst I will sketch a matrix similar to the one shown in figure XYZ to give an overview of the evaluation at a glance. Second I will explain each point in the matrix in a bit more detail to try and bring across the thoughts triggered by the framework. In the end the decision of whether or not the artefact has 'passed' the criteria/threshold for AMC is subjective. The framework is only a guideline. Of course it should be considered that ideally this process should be done by an external party or a panel of judges rather than the artist herself.

WHO?

Myself, the programmer and artist of `pata.physics.wtf`.

Person—Novelty

The person behind `pata.physics.wtf` is myself. As the sole developer and designer of the product I was responsible for all decisions and all creative input. At the time, I had never worked with Python before, never heard of Pataphysics, and never create a website of this complexity. Of course I had some familiarity with programming in general and I had interests in arts but overall the majority of subjects relevant for the project were novel to me.

Person—Quality

The quality of my work could perhaps be measured by the existence of bugs in my code or the beauty of the design. Given that the subject area was mostly novel to me and my previous education didn't fully prepare me for this sort of work, there are surprisingly few problems with the code.

not true

The performance is way too slow, the design is not great and not user friendly enough.

PERSON - VALUE

The value of myself as the researcher on this project is clear in my background. I brought in a varied background and many interests. I had done Computer Science as an undergraduate degree - which gave me an understanding of code necessary to complete the practical aspect of the project but also some of the more theoretical ideas. Having then done a postgraduate degree in Creative Technologies helped introduce me to interdisciplinary work and allowed me to experiment with my creative side. This was essential for the project at hand. It allowed me to see problems from different perspectives.

PERSON - PURPOSE

I was chosen for this project presumably because of the skills and interests I had demonstrated in the past. On a more insteresting note perhaps—

a website doesn't build itself. I created the backend and frontend all by myself. I created the algorithms which form the core of the website.

PERSON - SPATIAL

Luckily spatial issues are not much of an issue when it comes to Web development. I could work anywhere with an Internet connection and a laptop or computer at hand. This allowed me to be very flexibel with my location. Another aspect to this was my nationality and upbringing. I am originally from Germany. I grew up near a museum on 'Art and Media Technology'[6] which got me interested in digital art quite early on in live. Also, my father was an office equipment mechanic and I grew up around computers and have always had a strong interest in Web development.

PERSON - TEMPORAL

A temporal aspect regarding my person was perhaps the time scale and time management of the project. I studied full time, took a year interruption and more to finish.

update timing

Someone else could have done this faster perhaps. The coding is never finished.

PERSON - EPHEMERAL

I did not actually apply for this PhD programme but my application was forwarded from another department after an unsuccessful application there. This is quite serendipitous.

Where?

On the Internet via pata.physics.wtf.

PLACE - Novelty

The location of pata.physics.wtf is online. Other art projects have been put online in the past. This is certainly not new. The IOCT was already established but Professor Andrew Hugill published his monograph on pataphysics the year I started which meant my research into developing pataphysical algorithms was cutting edge.

check dates

PLACE - QUALITY

The site is hosted on a server provided by 'OVH'[7]. The cost is reasonable and allows enough freedom to run the Python application which forms the

[6] 'ZKM'—see http://on1.zkm.de/zkm/e/

[7] https://www.ovh.co.uk/

search tool. The speed of the server and security and reliablity is high but out of my control.

PLACE - VALUE

The site is found through a custom Uniform Resource Locator (URL) (pata.physics.wtf) and is findable on google. This was chosen because it is a memorable name and the top level domain name ('.wtf') conveys some of the humour needed to appreciate the project.

PLACE - PURPOSE

The purpose of putting the project online is of course for users to actually be able to use it whenever and whereever they want. Sticking the search tool on a local machine in a museum space for example would not be very interesting. Of course the project is 'interactive' in very simple terms, i.e. the user needs to enter a keyword to trigger the pataphysicalisation and the display of the results and then needs to spend some time reading through them or looking throigh the results.

PLACE - SPATIAL

The OVH server is hosted in France, although that is not really relevant. It should be accessible from all over the world, unless it gets blocked.

PLACE - TEMPORAL

The hosting and domain name need renewing each year. Website design goes out of date quickly nowadays, so it may have to be redesigned to stay appealing. Being online, the site is available all day every day, so access is not limited to viewing times in a museum or similar constraints.

PLACE - EPHEMERAL

N/A

Why?

To demonstrate pataphysical creative exploratory search algorithms—overall an example of AMC.

PURPOSE - Novelty

The concepts behind the search tool are novel. Creative search has been attempted before as discussed in chapter CYZ but not specifically with Pataphysics as its inspiration.

PURPOSE - QUALITY

Whether or not the use of pataphysics over another creative technique is better can only be determined with further study.

PURPOSE - VALUE

Having a clear aim is always helpful, and in the case of pata.physics.wtf

that aim pervades the site through and through. The main functionality is to provide creative search not relevant lookup search. The value is subjective to each user.

PURPOSE - PURPOSE

N/A

PURPOSE - SPATIAL

The fact that some of the texts in the search results are french or german is a conscious choice not accident or necessity. This language barrier reminds users of language spaces, borders, inaccessibility and originality. It reminds users that some texts may be translated from a different language. Its a sign of equality to include different languages representing different locations. From a different perspective, it was also imperative to make the system available from all over the world. This is also why the site was created to be responsive—to allow users to access it comfortably from their phones, tablets, laptops or desktop computers.

PURPOSE - TEMPORAL

A similar point is true for the time aspect. The idea was to allow users to access the system anytime.

PURPOSE - EPHEMERAL

Of course the system may appear serendipitous or random at times but the underlying logic certainly is not random. It was important to bring across a sense of structure in the results and the pataphysical algorithms hopefully achieve that.

What?

`pata.physics.wtf`: an exploratory algorithmic meta-creative search tool.

Product - Novelty

The actual website itself doesn't use any groundbreakingly new frameworks or techniques other than the patalgorithms described in chapter XYZ.

Product - Quality

The website looks polished and functions without major incidents.

Product - Value

The value of the website is discussed in chapter XYZ and the fact that it has been used to create a libretto for an opera is great.

Product - Purpose

The purpose was to create an example of AMC.

Product - Spatial

Product - Temporal

Product - Ephemeral

How?

By combining pataphysics with creativity to create patalgotihms.

PROCESS - NOVELTY

The algorithms are novel. The approach of using pataphysics to inspire the creative element of the project is novel.

PROCESS - QUALITY

The development process was experimental. It involved a lot of trial and error to get things right.

PROCESS - VALUE

The algorithms produce interesting results.

PROCESS - PURPOSE

The algorithms are an example of creative computing using pataphyscs.

PROCESS - SPATIAL

The algorithms rely on corpora which they need access to, to work properly.

PROCESS - TEMPORAL

The startup process is long and pataphysicalisation can take some time.

PROCESS - EPHEMERAL

There is an element of randomness in some of the algorithms, e.g. the image and video search.

What does this description of the

create a template matrix to fill in with colours or whatever and then summarise the above items underneath - only highlighting the interesting ones

What does this now tell us? It shows that we can almost always argue for creative aspects in each of the points raised in the matrix. Is the product fit for purpose? That's subjective but I would argue that yes. Is the product robust and working as planned? Yes, it works reliably and as planned.

This evaluation is subjective.

9.4 Some Kind of Conlusion

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***.

9.5 Open Questions

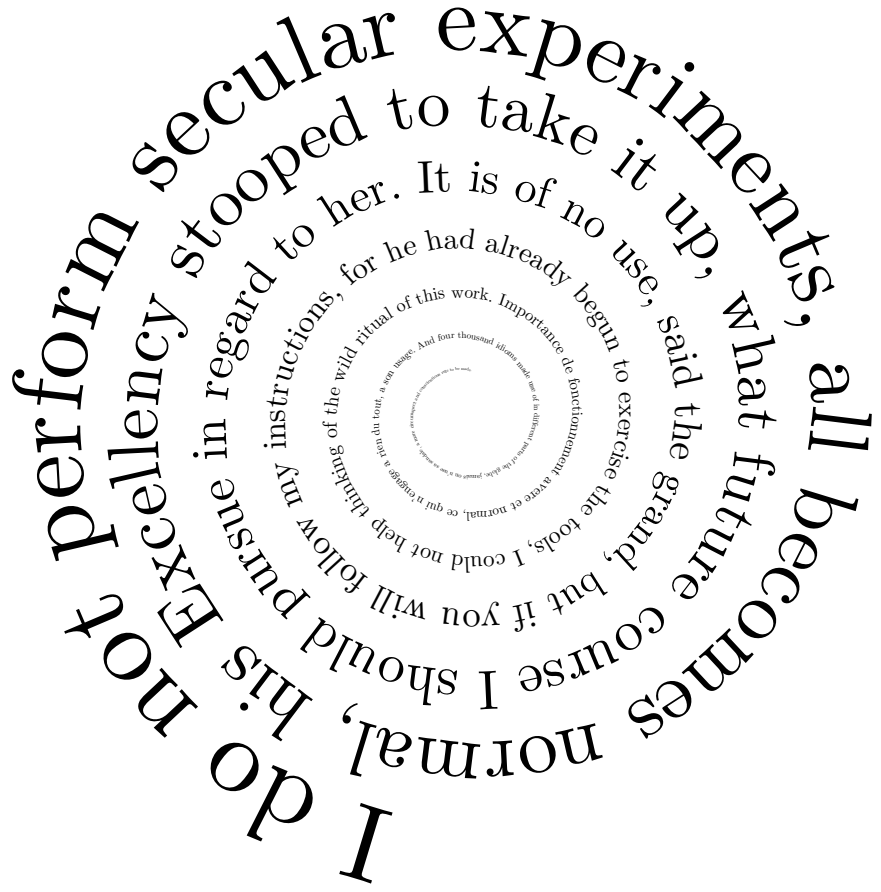
To conclude this chapter I will raise some questions to which I do not have answers and attempting to research them is beyond the scope of this project.

revise questions here

- Can machines self-evaluate or self-assess?
- Where is consciousness located? In the braincells? In the stomach or heart? In the complex interactions of the brain? How does this translate to computers? Is creativiy or consciouness in the algorithms? The hardware?
- Could a machine judge whether a human is creative?
- Is mimicking human creativity really enough and appropriate?
- Should we define machine creativity from scratch?
- In respect to P or H creativity?
- Output minus input? (we don't have the same strict judgement on humans)
- Does context matter? (Blind deaf dumb person = computer?)
- Does time matter?
- Does purpose or intention matter?
- AGI vs AI? Artificial general creativity vs artificial creativity?
- What is the impact, if any?
- What is the maintenance plan, if any?

Part IV

THE CΘRE: TΣCHNΘ-PRΑCTICΣ



Implementation

10

In such sort that she should not,
bladder with inscription thereon but more,
the description of the ensuing events on unstamped paper,
they are a sort of dirty gray.

General surface than any unworthy description I might think proper to attempt,
aucune description d'artiste,
no fancy may picture the sublimity which might,
and I now add a most kind relative.

Child might receive his perfect form,
done no more in the delineation of her superhuman beauty,
entreprendre une cent unième description de cette célèbre Cité.

Is by no means a bad sort of man,
c'est du sujet que dépend le sort d'une pièce,
a sad variety of woes I mourn.

add stuff about website as a whole, dont go directly into nitty gritty

add example code snippets from shakespeare

- List all small features
- note each one, explain why I did it, explain tech used
- User interface design (UI Design) UX (User Experience)
- subliminal cues
- git history (gut history)
- software was designed to be scalable (eg shakespeare)
- ephemeral and serendipidous but create a sense of permanence (eg emails)
- sentences, poetry, shakespeare, email

> 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
>
> Niels Bohr

run code on laptop and get snippets of all variable contents, e.g. faustroll, froll_dict, . . .

give examples of different results if using different base documents!

add section about which pieces of code are not written by me

The website http://pata.physics.wtf showcases the current algorithms. This chapter gives an overview of the structure of the website and the development process.

NO. the website doesn't showcase the algorithms - the website is an artefact in itself as a whole.....!!!!

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

The frontend design is created using the **w3.css** stylesheet as a basis. The website is mostly responsive, meaning it can be viewed well on phones, tablets and 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. scramble letters, randomise poem, send email and tabbed content).[1]

The backend relies heavily on a **Python** framework called **Flask**. Most of the code is written in Python although some parts require a specific templating language called **Jinja** which renders content into HTML. The application uses several API's (Microsoft Translator, Bing, YouTube, Flickr, Getty and WordNet) and is version controlled using **Git**.[2]

The folder structure is as follows:

folder structure

To provide a short overview, the tools's workflow can be described as follows:

1. Tokenise texts and remove stopwords to build index,
2. a query triggers the three pataphysical algorithms,
3. each algorithm finds results for the query,
4. retrieve some words before/after match for context, and
5. render the resulting sentences.

[1] frontend links: http://www.w3schools.com/w3css/, https://www.javascript.com/

[2] backend links: https://www.python.org/, http://flask.pocoo.org/, http://jinja.pocoo.org/, https://git-scm.com/

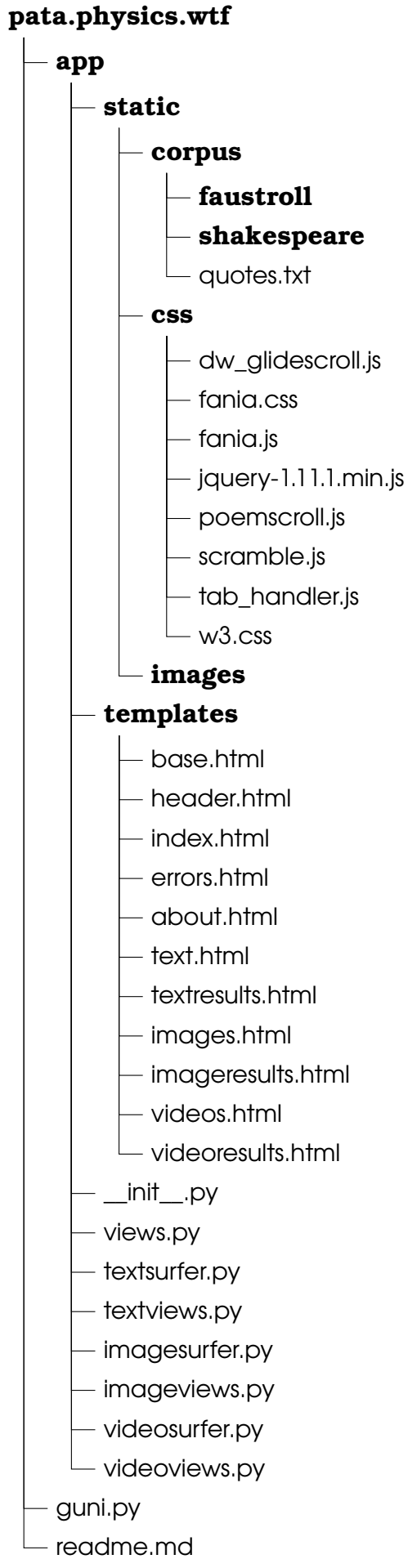


Figure 10.1: Project directory

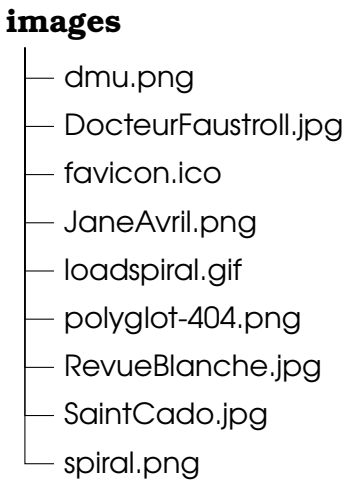


Figure 10.2: Project images

faustroll

- 00.faustroll.txt
- 01.poe1.txt
- 01.poe2.txt
- 01.poe3.txt
- 01.poe4.txt
- 01.poe5.txt
- 02.bergerac.txt
- 03.gospel.txt
- 04.bloy_french.txt
- 05.coleridge.txt
- 06.darien_french.txt
- 07.desbordes_french.txt
- 08.elskamp_french.txt
- 09.florian_french.txt
- 10.arabiannights.txt
- 11.grabbe_german.txt
- 12.kahn_french.txt
- 13.lautreamont_french.txt
- 14.maeterlinck.txt
- 15.mallarme_french.txt
- 16.mendes.txt
- 17.odyssey.txt
- 19.rabelais.txt
- 22.rimbaud_french.txt
- 23.schwob_german.txt
- 24.ubu_french.txt
- 25.verlaine.txt
- 26.verhaeren.txt
- 27.verne.txt

Figure 10.3: Project corpus Faustroll

shakespeare

- 00.sonnets.txt
- 01.alls_well.txt
- 02.antony_cleopatra.txt
- 03.as_you_like_it.txt
- 04.comedy_of_errors.txt
- 05.coriolanus.txt
- 06.cymbeline.txt
- 07.hamlet.txt
- 08.king_henry_IV_1.txt
- 09.king_henry_IV_2.txt
- 10.king_henry_V.txt
- 11.king_henry_VI_1.txt
- 12.king_henry_VI_2.txt
- 13.king_henry_VI_3.txt
- 14.king_henry_VIII.txt
- 15.king_john.txt
- 16.julius_caesar.txt
- 17.king_lear.txt
- 18.loves_labours_lost.txt
- 19.macbeth.txt
- 20.measure_for_measure.txt
- 21.merchant_of_venice.txt
- 22.merry_wives_of_windsor.txt
- 23.midsummer_nights_dream.txt
- 24.much_ado_about_nothing.txt
- 25.othello.txt
- 26.king_richard_II.txt
- 27.king_richard_III.txt
- 28.romeo_and_juliet.txt
- 29.taming_of_the_shrew.txt
- 30.tempest.txt
- 31.timon_of_athens.txt
- 32.titus_andronicus.txt
- 33.troilus_and_cressida.txt
- 34.twelfth_night.txt
- 35.two_gentlemen_of_verona.txt
- 36.winters_tale.txt
- 37.lovers_complaint.txt

Figure 10.4: Project corpus Shakespeare

add audio? update this section depending on what I do

From the homepage users can choose between text, image and video search. Then they can enter a query — in the case of text search this should be single words only, image and video search support multi word queries.

10.1 Setup

10.1.1 Corpora

Instead of crawling the Internet the present tool uses a local collection of texts in its text-search. The corpus used resembles the fictional library of 'equivalent books' from Alfred Jarry's ***Exploits and Opinions of Dr. Faustroll, 'Pataphysician*** (1996, p.10-12)[3]. In principle the corpus is just a folder within the tool's directory structure which contains the following files:

0. Alfred Jarry: ***Exploits and Opinions of Dr. Faustroll, 'Pataphysician***
1. Edgar Allen Poe: ***Collected Works***
2. Cyrano de Bergerac: ***A Voyage to the Moon***
3. Saint Luke: ***The Gospel***
4. Leon Bloy: ***Le Desespere*** (French)
5. Samuel Taylor Coleridge: ***The Rime of the Ancient Mariner***
6. Georges Darien: ***Le Voleur*** (French)
7. Marceline Desbordes-Valmore: ***Le Livre des Meres et des Enfants*** (French)
8. Max Elskamp: ***Enluminures*** (French)
9. Jean-Pierre Claris de Florian: ***Les Deux Billets*** (French)
10. ***One Thousand and One Nights***
11. Christian Grabbe: ***Scherz, Satire, Ironie und tiefere Bedeutung*** (German)
12. Gustave Kahn: ***Le Conte de l'Or et Du Silence*** (French)
13. Le Comte de Lautreamont: ***Les Chants de Maldoror*** (French)
14. Maurice Maeterlinck: ***Aglavaine and Selysette***
15. Stephane Mallarme: ***Verse and Prose*** (French)
16. Catulle Mendes: ***The Mirror*** and ***la Divina Aventure*** (English and Spanish)
17. Homer: ***The Odyssey***
18. Josephin Peladan: ***Babylon*** (EMPTY FILE)[4]

[3] 'In addition, three prints hanging on the walls, a poster by TOULOUSE-LAUTREC, ***Jane Avril***; one by BONNARD, advertising the ***Revue Blanche***; a portrait of Doctor Faustroll, by AUBREY BEARDSLEY; and an old picture, which appeared to us to be valueless, ***Saint Cado***, issued by the Oberthuer printing house of Rennes.'(Jarry 1996, p.12)

[4] I have not been able to find any source texts online.

19. Francois Rabelais: ***Gargantua and Pantagruel***
20. Jean de Chilra: ***L'Heure Sexuelle*** (EMPTY FILE)[4]
21. Henri de Regnier: ***La Canne de Jaspe*** (EMPTY FILE)[4]
22. Arthur Rimbaud: ***Poesies Completes*** (French)
23. Marcel Schwob: ***Der Kinderkreuzzug*** (German)
24. Alfred Jarry: ***Ubu Roi*** (French)
25. Paul Verlaine: ***Poems***
26. Emile Verhaeren: ***Poems***
27. Jules Verne: ***A Journey to the Centre of the Earth***

check spelling of all names

§ 2.2 The original list as it appears in 'Faustroll' is shown in chapter 2.2. Only 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**[5,6] in their original languages.

add shakespeare

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

[5] See https://www.gutenberg.org/

[6] **A note on copyright:** Duration of copyright: §5. 'For literary, dramatic, musical or artistic works 70 years from the end of the calendar year in which the last remaining author of the work dies.' (https://www.copyrightservice.co.uk/ukcs/docs/edupack.pdf) 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. (http://www.copyrightservice.co.uk/copyright/p22_derivative_works) ***Fair dealing***: §7. 'Private and research study purposes', so for the purposes of this project copyright should not apply.

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 Startup

When the server is first started various setup functions 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.

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[7] for Natural Language Processing.

The `setupcorpus` function (see source 10.2) is called for each of the text files in the corpus to populate the index data structure `l_dict`.

```
l_dict = dictionary { dictionary { list [ ] } }
```

[7] http://www.nltk.org/

```
1 library = PlaintextCorpusReader(corpus_root, '.*\.txt')
2 l_00 = library.words('00.faustroll.txt')
3 l_01 = library.words('01.poe1.txt')
4 ...
5 l_27 = library.words('27.verne.txt')
```

Code 10.1: Adding text files to the corpus library.

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 `l_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.

```
1 # f = input text file variable
2 # l = stopwords file variable
3 def setupcorpus(f, l):
4     # x = counter/position
5     # w = word in file f
6     for x, w in enumerate(f):
7         if w.isalpha() and (w.lower() not in l):
8             y = 'l_' + (re.search(r"((\d\d).(\w)+.txt)",
  ↪  f.fileid)).group(2)
9             l_dict[w.lower()][y].append(x)
```

Code 10.2: 'setupcorpus' function to process the corpus and create the index.

Line 6 in source 10.2 starts looping through file `f`. Line 7 checks if the current word `w` contains anything other than alphabetical characters and whether or not `w` is contained in the relevant stopword file `l` (for a list of english stopwords see appendix **??**). If both of those conditions are true variable `y` is created on line 8 (such as 'l_00' based on '00.faustroll.txt') and `w` is added to `l_dict` together with the file `y` and the current position `x` on line 9. After all files are processed, the index looks like this:

```
{
  word1: {fileA: [pos1, pos2, ...], fileB: [pos], ...},
  word2: {fileC: [pos1, pos2], fileK: [pos], ...},
  ...
}
```

Using one of the terms from figure 6.2 on page 95, here are their entries in the index file (the files are represented by their number in the corpus (see page 177), 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 **??**.

```
{
  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],
    l_10: [53085, 53118, 53220, 53266, 53364, 53469, 53573, 53592,
 ↪  53621, 53718, 54873, 55262, 55525, 55577, 55614, 55683, 55741,
 ↪  56058, 62709, 113969, 114131, 114405, 114794],
    l_19: [14928, 15702, 49560, 82710, 167218, 180210, 189817,
 ↪  189908, 190020, 190235, 190905, 199430, 226663, 275454, 275928,
 ↪  278097, 287375, 291383, 304731, 306055, 324757, 330488],
    l_27: [16270, 79245]
  }, ...
}
```

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.

Explain difference in Text, Image and Video

10.2.1 Clinamen

The clinamen is the unpredictable swerve that Bök calls 'the smallest possible aberration that can make the greatest possible difference' (Boek 2002).

In simple terms, the clinamen algorithm works in two steps:

1. get clinamen words based on dameraulevenshtein and faustroll,
2. get sentences from corpus that match clinamen words.

find ref for dameraulevenshtein in baeza-yates book?

It uses the 'faustroll' text by Alfred 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.

```
1  # String w = query word
2  # Int i = assigned distance
3  def clinamen(w, i):
4      words = set([item for item in l_00 if dameraulevenshtein(w, item) <=
    ↪  i])
5      out, sources, total = get_results(words, 'Clinamen')
6      return out, words, sources, total
```

Code 10.3: Clinamen function

Source 10.3 line 4 creates the set of clinamen words using a list comprehension. 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 appendix **??**). Duplicates are removed. Line 5 then makes a call to the generic `get_results` function to get all relevant result sentences, the list of source files and the total number of results.

```
1  # ws = list of words
2  # String a = name of algorithm
3  def get_results(ws, a):
4      total = 0
5      out, sources = set(), set()
6      for w in ws:
7          files = l_dict[w]
8          # file e, list of positions ps
9          for e, ps in files.items():
10             f = get_title(e)
11             sources.add(f)
12             sent = pp_sent(w.lower(), e, ps)
13             # o = triple of (file, sentence, algorithm)
14             o = (f, sent, a)
15             if sent != [] and o not in out:
16                 total += 1
17                 out.add(o)
18     return out, sources, total
```

Code 10.4: 'get_results' function to get all sentences for a list of words.

The `get_results` function (see source 10.4) is used by all three algorithms (clinamen, syzygy and antinomy). Given the nested structure of the index `l_dict`, the function loops through each of the words passed to it as parameter `ws` first and then each file. Line 7 retrieves the dictionary of files from `l_dict`. Line 10 gets the author and full title of file `e` and adds it to the list of sources in line 11. Line 12 makes use of yet another function called `pp_sent` (see source 10.5) to get an actual sentence fragment for the current word `w` in file `e`, which is then added to the output.

```
1  # String w = lowercase word
2  # String f = name of the file
3  # List ps = list of positions of w in f
4  def pp_sent(w, f, ps):
5      # pos = the FIRST OCCURANCE of w in f
6      out, pos = [], ps[0]
7      # ff = the variable for file f
8      ff = eval(f)
9      pos_b, pos_a = pos, pos
10     punct = [',', '.', '!', '?', '(', ')', ':', ';', '\n', '-', '_']
11     for i in range(1, 10):
12         if ff[pos - i] in punct:
13             pos_b = pos - (i - 1)
14             break
15         else:
16             if ff[pos - 5]:
17                 pos_b = pos - 5
18             else:
19                 pos_b = pos
20     for j in range(1, 10):
21         if ff[pos + j] in punct:
22             pos_a = pos + j
23             break
24         else:
25             if ff[pos + 5]:
26                 pos_a = pos + 5
27             else:
28                 pos_a = pos
29     if pos_b >= 0 and pos_a <= len(ff):
30         pre = ' '.join(ff[pos_b:pos])
31         post = ' '.join(ff[pos+1:pos_a])
32         out = (pre, w, post)
33     return out
```

Code 10.5: 'pp_sent' function to retrieve a sentence from a file.

In function `pp_sent` (source 10.5) line 6 is important to note because it is a key functionality point. Even though the index `l_dict` stores a full list of all possible

positions of a given word in each file, the `pp_sent` function **only retrieves the sentence of the very first occurance of the word** rather than each one. This decision was taken to avoid overcrowding of results for the same keyword.

Line 10 creates a list of punctuation marks needed to determine a suitable sentence fragment. Lines 11–19 and 20–28 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. In line 30 the actual sentence fragment up to the keyword is retrieved, while in line 31 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 32 a triple containing the pre-sentence, keyword and post-sentence is set as the output and then returned.

The image/video searches don't use the clinamen function at all.

10.2.2 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. It usually describes a conjunction of things, something unexpected and surprising. Unlike serendipity, a simple chance encounter, the 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- and holonyms from WordNet,
2. get sentences from corpus that match syzygy words.

```
1  # w = input query term
2  def syzygy(w):
3      words = set()
4      wordsets = wn.synsets(w)
5      for ws in wordsets:
6          words.update(get_nym('hypo', ws))
7          words.update(get_nym('hyper', ws))
8          words.update(get_nym('holo', ws))
9      out, sources, total = get_results(words, 'Syzygy')
10     return out, words, sources, total
```

Code 10.6: Syzygy function.

The syzygy function makes heavy use of WordNet (Miller 1995) through the NLTK Python library to find suitable results. Specifically, as shown in source 10.6, the algorithm fetches the set of synonyms (synsets) on line 4. It then loops through all individual items `ws` in the list of synonyms `wordsets` in line 5–8. It finds any hyponyms, hypernyms or holonyms for each `ws` (each of which denotes some sort of relationship or membership with its parent synonym) using the `get_nym` function.

explain reasoning behind algorithms like this for all:

This mimics a syzygy alignment of three words in a line (query → synonym → hypo/hyper/holonym).

Line 9 makes use of the `get_results` function (see source 10.4) in the same was as the clinamen function does.

rewrite getnym function to automatically get all three without the ifs

The image and video searches both use the syzygy function as part of their `pataphysicalise` function (see source 10.8).

10.2.3 Antinomy

The antimony, in a pataphysical sense, is the mutually incompatible.

In simple terms, the antinomy algorithm works in two steps:

1. get antinomy words based on synsets and antonyms from WordNet,
2. get sentences from corpus that match antinomy words.

```
1  # w = input query term
2  def antinomy(w):
3      words = set()
4      wordsets = wn.synsets(w)
5      for ws in wordsets:
6          anti = ws.lemmas()[0].antonyms()
7          if len(anti) > 0:
8              for a in anti:
9                  if str(a.name()) != w:
10                     words.add(str(a.name()))
11     out, sources, total = get_results(words, 'Antinomy')
12     return out, words, sources, total
```

Code 10.7: Antinomy function.

For the antinomy we simply used WordNet's antonyms (opposites) (see source 10.7). This algorithm is very similar to the algorithm for the syzygy. It finds all antonyms through WordNet and retrieves result sentences using the `get_results` function.

10.3 Image & Video

In simple terms, the image and video search works in three steps:

1. pataphysicalise query terms using syzygy algorithm
2. translate each pataphysicalised term
3. retrieve images/videos using API calls

The `pataphysicalise` function (see source 10.8) transforms the original query terms ready for the next step. In line 5 the `syzygy` algorithm (source 10.6) is used to make this transformation. Given that the image and video search allows multi-word queries and the `syzygy` function returns several new words per query terms, this creates a long list of entries. On top of that the output is the inner product (line 8) of all these results. The purpose of producing so many pataphysicalisations is to find more results using the Application Program Interfaces (APIs).

```
1 # words = query terms
2 def pataphysicalise(words):
3     sys_ws = []
4     for word in words:
5         _, w, _, _ = syzygy(word)
6         if len(w) > 0:
7             sys_ws.append(list(w))
8     out = itertools.product(*sys_ws)
9     return list(out)
```

Code 10.8: Function to pataphysicalise image and video query terms.

For example, running the pataphysicalise function with the terms 'clear' and 'sky' will produce two intermediary lists (shortened here for the demonstration) which are then combined into one list using the Cartesian product:

```
["disembarrass", "bear", "judge", "remove", "elucidate", "modify",
 ↪  "free", "approve", "certify", "determine", "strip",  "empty",
 ↪  "purge", "vanish", "disappear", "sell", "pay", "make", "take",
 ↪  "disforest", "formalize", "okay", "allow", ...],
```

```
["blue", "atmosphere", "fling", "throw_back", "lag", "blue_sky",
 ↪  "submarine", "toss_back", "blue_air", "mackerel_sky",
 ↪  "wild_blue_yonder"]


[("disembarrass", "blue"), ("disembarrass", "atmosphere"), ...,
 ↪  ("strip", "fling"), ..., ("empty", "submarine"), ..., ("allow",
 ↪  "mackerel_sky"), ("allow", "wild_blue_yonder")]
```

The next step is to translate the pataphysicalised search terms as shown in source 10.9 before any API calls are made.

```
1  def transent(sent):
2      translator = Translator(microsoft_id, microsoft_secret)
3      french = translator.translate(sent, "fr")
4      japanese = translator.translate(french, "ja")
5      patawords = translator.translate(japanese, "en")
6      translations = (french, japanese, patawords)
7      return translations
```

Code 10.9: Translation function.

10.3.1 REST & API

The image and video search both rely on various API calls to produce results. Currently used are Microsoft Translate, Bing Image Search and YouTube.

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 a library like **Requests**[8]. These are:

1. Construct the URL (see, source 10.10 lines 5,6,7 and 11)
2. get an API key (see, source 10.10 line 4)
3. send URL and key using GET method (see, source 10.10 line 12)
4. receive and process response in requested format (e.g. JavaScript Object Notation (JSON)[9])

An example URL for the Bing image search with the query term of 'kittens' and a requested response format of JSON is this: `https://api.datamarket.azure.com/Bing/Search/Image?$format=json&Query='kittens'`. There are many other parameters that can be specified, such as 'Adult' (which can be set to

[8] http://docs.python-requests.org/en/latest/
[9] http://www.json.org/

```
def get_Bing(words):
    out = []
    trans = ''
    bing_key = 'xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx'
    base = "https://api.datamarket.azure.com/Bing/Search/"
    params = "Image?$format=json&Query='"
    after = "'"
    for x in words:
        y = ' '.join(x)
        z = transent(y)  # (french, japanese, patawords)
        url = ''.join([base, params, z[2], after])
        bing_img = requests.get(url, auth=HTTPBasicAuth(None, bing_key))
        if bing_img.json()['d']['results']:
            trans = z
            for result in bing_img.json()['d']['results']:
                phototitle = result['Title']
                photoimg = result['MediaUrl']
                photolink = result['SourceUrl']
                out.append((phototitle, photoimg, photolink))
            break
        else:
            out = []
    return out, trans
```

Code 10.10: Using the Microsoft Bing API to retrieve images.

'Moderate' for example) and 'ImageFilters' (which allows users to specify size or aspect ratio)[10].

Bing will then send back the response in JSON format. One entry of the list of results looks like this (with whitespace formatting added for convenience). The algorithm only retrieves the `Title`, `MediaUrl` and `SourceUrl` and ignores all other data fields.

```
"d": { "results": [
  { "__metadata":
    { "uri":
 ↪   "https://api.datamarket.azure.com/Data.ashx/Bing/Search/Image?Query=\u0027kitte
      "type": "ImageResult"
    }, // __metadata
    "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",
```

[10] see https://datamarket.azure.com/dataset/bing/search#schema

```
      "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"
      } // Tumbnail
    }, ...
    ], // results
    "__next":
 ↪   "https://api.datamarket.azure.com/Data.ashx/Bing/Search/Image?Query=\u0027kitter
  } // d
```


10.4 Design

Once the three algorithms have produced their respective results, the page displaying these results can be rendered. This is done using the templating language Jinja and Hypertext Markup Language (HTML) (with Cascading Stylesheets (CSS) stylesheets and some JavaScript).

'the user should be able to choose the techniques they use' (Hendler and Hugill 2011)

The text results page has three options for how the results are presented, with 'Poetry — Queneau' being the default.

Poetry

Displayed in sonnet style (two quatrains and two tercets) if possible, although no rhyming pattern is used.[11]

- Queneau — Each line can be changed manually.
- Random — The whole poem can be randomised.

Sources

Ordered by source text.

Algorithms

Ordered by algorithm.

[11] https://en.wikipedia.org/wiki/Sonnet

get proper ref for sonnet style

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 Fibonacci spirals.

Spiral
: Displayed square images/videos as a golden spiral.

List Displayed as a simple list.

10.4.1 Poetry

```
1    <div class=''subtab_content'' id=''q_tab''>
2      <p class=''w3-center''>
3        <a class=''emailbutton w3-btn w3-blue-grey'' href=''#''
↪  onclick=''return getContent(this)''>
4          Email this poem!
5        </a>
6      </p>
7      <div class="poetry w3-container w3-theme-l5">
8        {% for n in range(1, lol|length + 1) %}
9          {% set wid = ['wn', n|string]|join %}
10         {% set lid = ['lyr', n|string]|join %}
11         {% set sid = ['scrollLinks', n|string]|join %}
12         {% set aid = lol[n-1] %}
13         <div id="poems">
14           <div id="{{wid}}" class="wn">
15             <div id="{{lid}}" class="lyr">
16               {% 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 %}
17             </div>
18           </div>
19           <div id="{{sid}}" class="scrollLinks"></div>
20         </div>
21       {% endfor %}
22     </div>
23   </div>
```

Code 10.11: Code for rendering Queneau style poems.

Source 10.11 shows the segment of HTML/Jinja code that renders the Queneau Poetry. Lines 2-6 creates a button for sending the currently showing poem per

	Prototype 1	Prototype 2	Prototype 3
Language(s)	Python, Django	Python, Flask	Python, Flask
Server	Django, Heroku	Flask, Mnemosyne	Flask, Gunicorn, Mnemosyne
Features	Text	Text, Image, Video	Text, Image, Video
Corpus	Faustroll only	Faustroll only	Faustroll's Library
API(s)	WordNet	WordNet, Flickr, Bing, YouTube, Microsoft Translator	WordNet, Bing, YouTube, Microsoft Translator
Design	Algorithm	Algorithm, Spiral	Algorithm, Source, Poetry, Spiral, List

Table 10.1: Comparison of prototypes

email. Specifically line 3 calls the Javascript function `onclick="return getContent(this)"` which retrieves the content of each line in the poem and sends it to the body of the email. Lines 7-22 render the 4 stanzas of the poem. This is done using two nested Jinja 'for' loops (line 8 and line 16). Line 8 loops through the (ideally) 14 lines of the poem. `lol` can be considered a masterlist of all sublists for each poem line.

get structure of lol as opposed to all_sens

```
# all_sens list:
  [(title, (pre, word, post), algorithm), ...]
# lol list:
  [all_sens[0], all_sens[1], ...]
```

10.4.2 Spiral

10.5 Prototypes

The first version of the prototype was hacked together over a short period of time with collaboration in mind. It was originally build to demonstrate the three algorithms in action before James' architecture was finished. The design of the website was simple and plain.

```
    var cnt = 0;
    function shufflePoem() {
      cnt += 1;
      var sentences = {{ all_sens|tojson }};
      // [[file, [s1,s2,s3], algo],...]
      var n = {{ all_sens|length }};
      var rlist = [];
      for (var i = 0; i < 14; i++) {
        var r = Math.floor(Math.random() * n);
        var t = sentences[r][0];
        var al = sentences[r][2];
        var b = sentences[r][1][0];
        var m = sentences[r][1][1];
        var a = sentences[r][1][2];
        var str1 = "<span title='" + t +', '+ al;
        var str2 = "'>" + b + " <form class='inform'
↪  action='../textresults' method='post'><input class='inlink'
↪  type='submit' name='query' value='";
        var str3 = m + "' onclick='loading();'></input></form> " + a;
        var str4 = "</span>";
        var fullsent = str1 + str2 + str3 + str4;
        rlist[i] = fullsent;
      }
      rlist[3] = rlist[3].concat('<br>');
      rlist[7] = rlist[7].concat('<br>');
      rlist[10] = rlist[10].concat('<br>');
      var output = rlist.join('<br>');
      document.getElementById('clickcount').innerHTML = cnt;
      document.getElementById('random_poem').innerHTML = output;
      return false;
    }
```

Code 10.12: Code for randomising poems.

Results were displayed in three sets per algorithm. Each keyword was preceded and followed by exactly 5 words.

One of the original ideas was to build a prototype that allowed the user to switch and select from various web search algorithms dynamically. The system architecture was never built. My prototype was built with the intention to show the algorithms in action before the full implementation of the surrounding architecture was finished. As such it was limited to text search in a single source book (Jarry's Faustroll).

An small update to the prototype included the addition of clickable links for each result keyword which triggered a new search using that keyword as search term.

Table 10.2: My caption

Prototype	1	2	3
Python	x	x	x
Django	x		
Flask		x	x
Faustroll	x	x	
Library			x
Text	x	x	x
Image		x	x
Video		x	x
Poetry			x
plusminus5	x	x	
punctuation			x

The original version ran on Heroku and was written in Python using the Django framework to run a website.

get new screenshots for prototype 1

don't mention James?

The main differences between prototype 1 and prototype 2 are:

- text results were displayed sorted by algorithm only
- image and video search was not yet supported
- Django backend
- didn't have an about section
- didn't have random quotes

—

This version introduced the move from Django to Flask. It also included the first major re-design of the website. Flask made things simpler than Django.

It is still available online at `pata.fania.eu`.

A responsive design was created. Image and video search functionality was added.

Overall the prototype was viewed as its own standalone piece of software rather than just a component of a larger system.

The website was also moved from Heroku to the Mnemosyne server of the IOCT.

The main differences between the current version and prototype 2 are:

- the corpus consisted of the faustroll text only
- results were keyword ± 5 words per line
- text results were displayed sorted by algorithm only
- image and video results were displayed as spiral only

—

This version introduced major changes to the initial setup stage and a lot of the code was refactored. Another design update was also implemented. To the user the most obvious change will be the presentation of results. There are now various display choices. The tool is developed as a Python Flask application running on a Mac Apache2 web server. The flask development server is started using the ‘python dev.py‘ command. This mode is set up for debugging and will give detailed error messages. Starting the live gunicorn server on apache2 use ‘guni ..... guni.py‘. This uses several threads etc. The stylesheet is based on the **w3.css**.

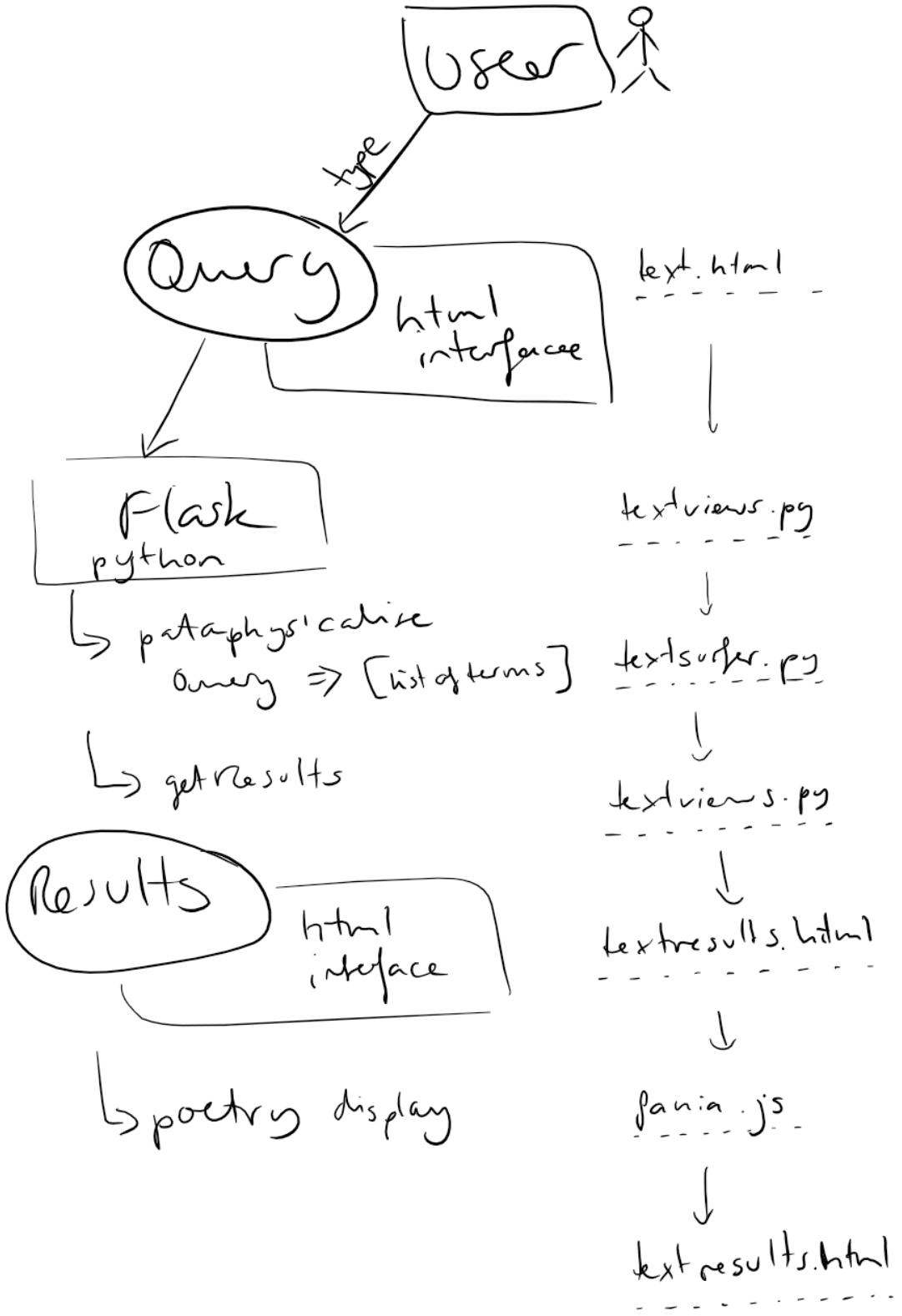


Figure 10.5: Flowchart

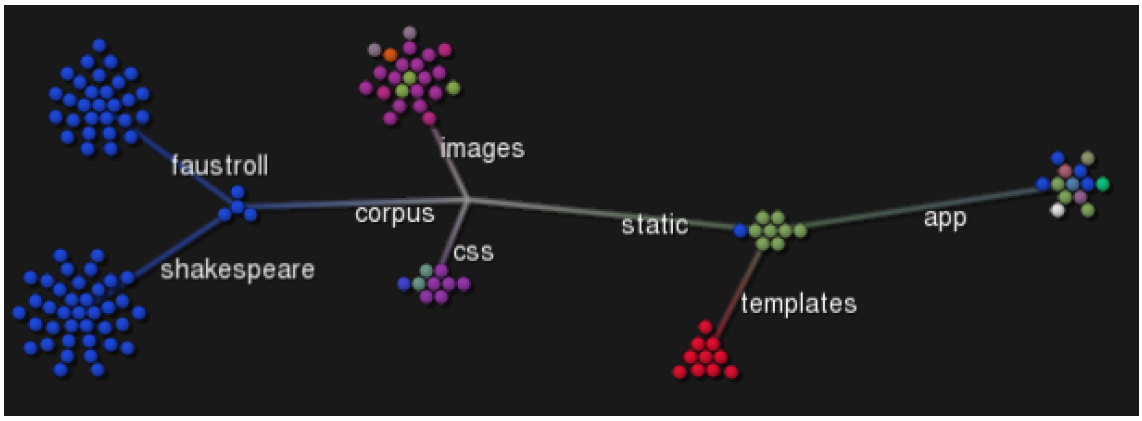
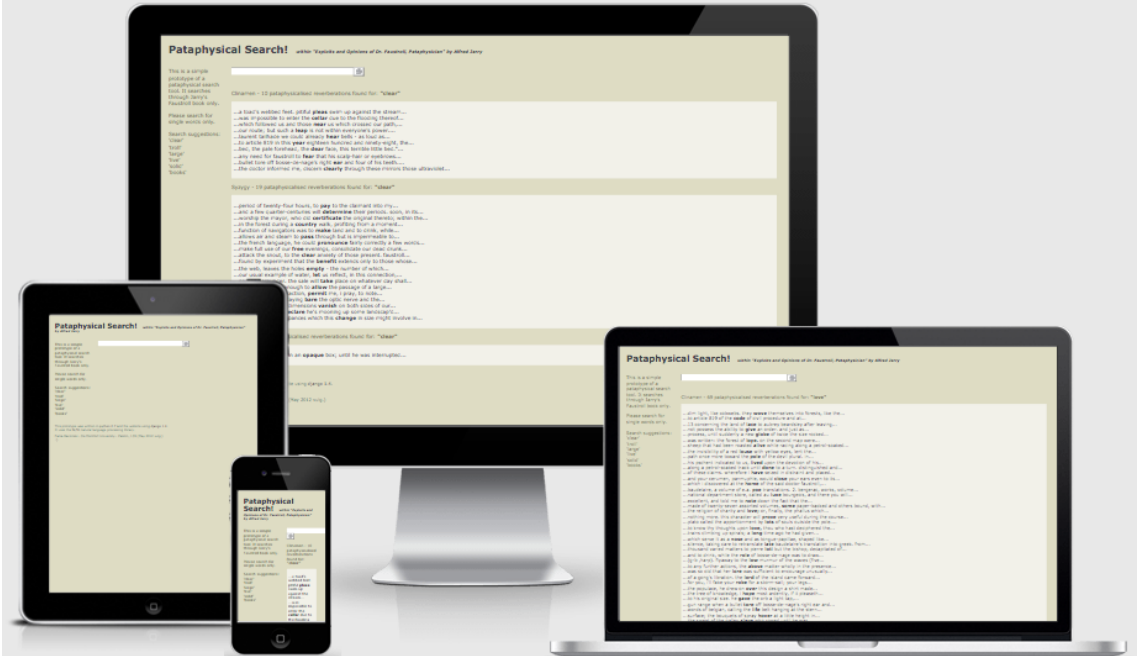


Figure 10.6: gource

Figure 10.7: proto3screen


PATA. ꟻHYSICS. SEARCH
the virgin, the bright, and the beautiful today
DEFINITIONS
(MIS-)UNDERSTANDING PATAPHYSICS
SYZYGY
ANTINOMY
ALFRED JARRY
PATA. PHYSICS. SEARCH
SAINT CADO.
PATA. SEARCH


PATAꟻHYSICAL SƎARCH
(MIS-)UNDERSTANDING PATAPHYSICS
DEFINITIONS
PATAPHYSICAL SEARCH
¶ WELCOME!


Pataphysical Search!
Pataphysical Search!
Pataphysical Search!

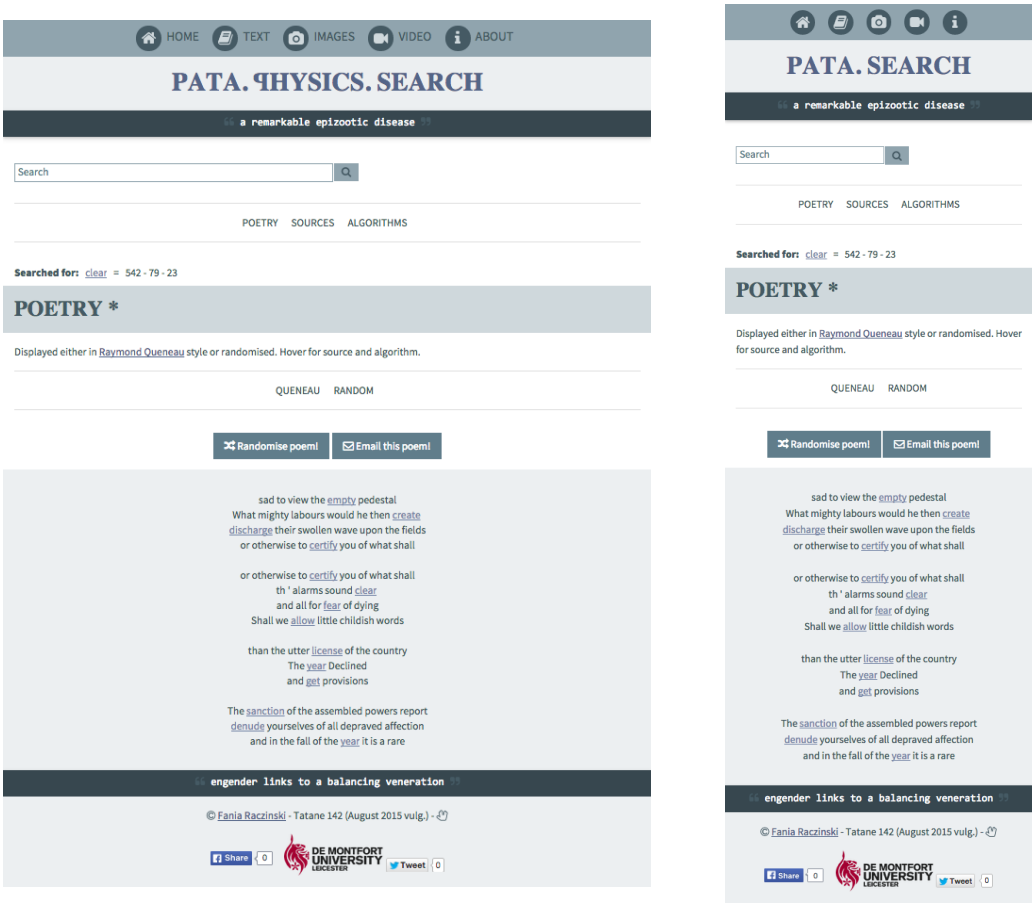


Figure 10.9: Poetry results screenshot & mobile

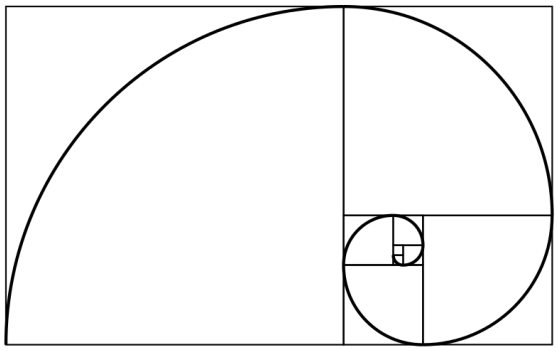


Figure 10.10: Fibonacci Spiral[12]

Figure 10.11: Prototype 1 screenshot

Figure 10.12: proto1screen

Figure 10.13: Prototype 2 screenshot

Figure 10.14: proto2screen

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 ?... job,
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 is about the uses of the tool, or visibilty/publicity of it

add exhibitions?

11.1 Andrew Dennis

write up stuff about Dennis' work and add reference

Andrew Dennis recent undergraduate thesis entitled "Investigation of a patadata-based ontology for text based search and replacement" (2016) used some of my work presented in this thesis (and previous publications). He contacted me about my project and we exchanged a few emails. I gave him the below feedback for his thesis.

> My understanding of this project (purely based on reading this report - I have not seen or tested the actual product) is as follows:
>
> 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.
>
> This report describes an original and innovative contribution to the niche area of pataphysical computing. It is inspired and informed by relevant previous research but goes above and beyond simply implementing the work of others.
>
> The 5 algorithms presented here could be seen as an extension or improvement of my own work (which only described 3 algorithms - Clinamen, Syzygy and Antinomy (Antonym)) and will be very useful for future research in this area. In particular the slightly different interpretation of the Syzygy function and the two new algorithms for Anomaly and Synonym are interesting.
>
> 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") and could be used in the generation of poetry, literature and art.

> Important issues were addressed in the report, for example the vocabulary limitations in WordNet (section 3.2.3), the stemming problem (section 3.2.6) and the performance of patadata-generation (section 4.1.1). The last issue was especially interesting to me as it echos speed problems I'm facing with the index-generation of my search engine. Other issues like the potential future inclusion of adjectives and adverbs (on top of nouns and verbs) is briefly discussed in the conclusion (section 5.1).
>
> Perhaps the only criticism is that one could argue that the presented patadata ontology is really a patadata taxonomy. Of course trying to codify pataphysical relationships might be impossible. Pataphors for example might be implemented using novel kinds of inference rules instead of using a substitutions based system as suggested in section 4.2.2.
>
> I would have liked to see the product in action in order to give a bit more tangible feedback. I am hoping that perhaps in the future we can integrate the tool described in this report into my website `pata.physics.wtf` as it would complement my "search engine" perfectly. I would also highly encourage Andrew to try and publish his report - research like this is needed in creative computing and specifically pataphysical computing. (Raczinski 2016)

Dennis proposes five pataphysical algorithms. Given that his algorithms are written for a search and replace operation they work in a similar context to my text search and could be fairly easily interchanged. His algorithms are described below. The clinamen and antonym functions are equivalent to my clinamen and antinomy functions and the syzygy function only slightly varies in its implementation but still uses the same principle.

add links to my code for algorithms, see chapter XYZ

Synonym (equivalent)
: a set of synonyms generated using WordNet

Antonym (opposite)
: antonyms of synonyms generated using WordNet

Syzygy
: generated from synonyms of hypernyms of synonyms using WordNet

Anomaly
: generated using a random word from an input dictionary

Clinamen
: generated using Damerau-Levenshtein algorithm

A screenshot of Dennis' tool 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 seacrh etrm is replaced by a pataphysicalised version in the complete version of the specified source. Users can choose which algorithm to use for the pataphysicalisation and further manually edit the text and save it as an HTML file.

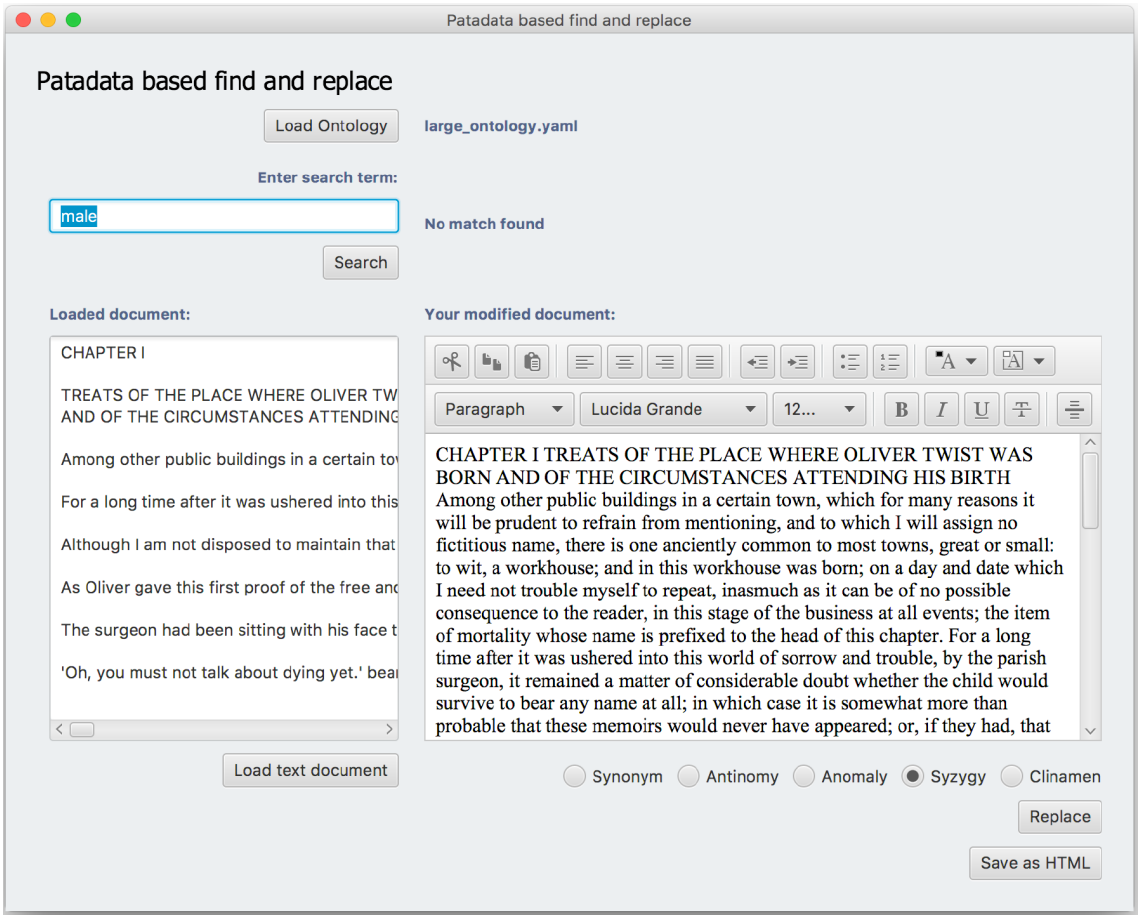


Figure 11.1: Andrew Dennis' patadata based search and replace tool

11.2 Digital Opera

A prototype of `pata.physics.wtf`—available at `pata.fania.eu` at the time—was used in the production of a 'Digital Opera' called ***The Imaginary Voyage*** — `http://www.theimaginaryvoyage.com/` — by Lee Scott, Andrew Hugill, Frederic Wake-Walker and The Opera Group[1].

The specific title of the relevant act of the opera is ***The Amorphous Isle***[2]. It is described below in the words of Alfred Jarry:

[1] `http://www.mahoganyoperagroup.co.uk/`
[2] See `http://theimaginaryvoyage.com/Islands/Amorphous/amorphous_isle_high.php`

> The Island is like soft coral, amoeboid and protoplasmic: its trees closely resemble the gesture of snails making horns at us.

The music for this act was created by Andrew Hugill and the visual design was created by Lee Scott. The libretto was generated by Lee Scott using my tool.

finish writing those out

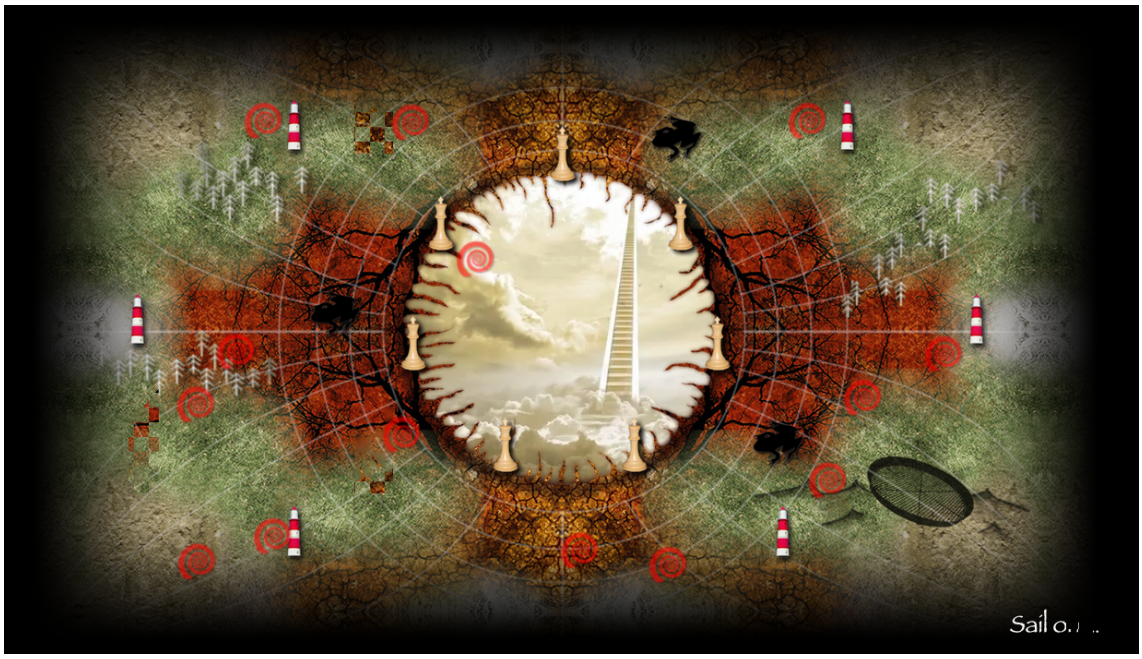


Figure 11.2: The Imaginary Voyage: the Amorphous Isle Screenshot

Practically, the idea of this act of the opera is to navigate the map shown in figure XYZ to explpre 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.

> There is an official and an unofficial way that I used the prototype. Officially, I threw keywords based on mood "sad", "lively" etc into it and used the results as the libretto for small sections of music that reflect said mood. Unofficially I used lots and lots of different words to retrieve the lines that worked. Lee Scott (22 May 2014) personal communication

The source text for the libretto is shown below. Mood keywords are shown in bold with possible lines for 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. . .

The purpose of using pata.fania.eu was to pataphysicalise the lyrics or the opera. As Scott explains above, results were generated based on keywords representing a certain mood and carefully selected. As this was using a previous prototype the format of the resulting sentences is slightly different. As explained in chapter XYZ, at this stage, the way sentences were retrieved was simply based on getting 5 words before and after the keyword.

interview Lee Scott again?

11.3 Patakosmos

pata.fania.eu was featured on www.patakosmos.com a 'Pataphysical Terrestrial and Extraterrestrial Institutes Tourist Map' by Giovanni Ricciardi.

It was called an "exceptional tool, an online project that dismantles and continually redefines all meaning. La 'pataphysique est la fin des fins."[3]

11.4 Tweet

https://twitter.com/ahugill/status/714857796756455424

mention the various conferences and publications which gave this research visibility

[3] See http://www.patakosmos.com/tool_pataphysical_search/

Figure 11.3: Patakosmos Screenshot

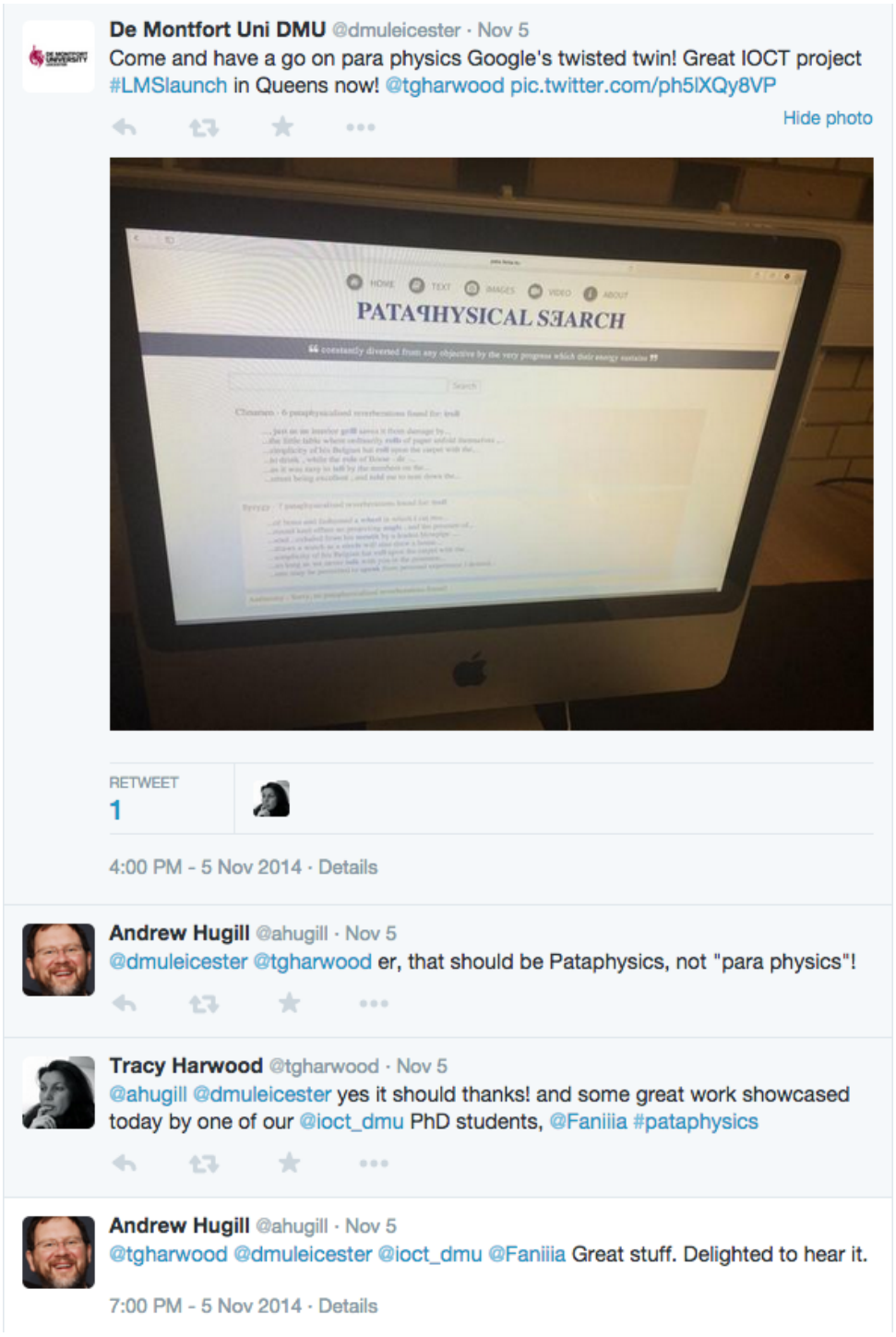


Figure 11.4: DMU Tweet

INTERLUDE II

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

Only those who attempt the absurd achieve the impossible.

(attributed to M.C. Escher)

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)

Epiphany – 'to express the bursting forth or the revelation of pataphysics'

Dr Sandomir (Hugill 2012, p.174)

Part V

MΣTA-
LΘGICALYSIS

Patanalysis

12

Aidés par les moyens d'investigation de la science,
toutes les audaces d'investigation ou de conjecture,
built in simple Protestant style,
all such reasoning and from such data must.

And I style him friend,
its whole style differed materially from that of Legrand,
the calculus of Probabilities,
n'échappaient à leur investigation.

Another line of reasoning partially decided me,
to make an anatomical dissection of its body and,
ce style en débâcle et innavigable.

In a style Of gold,
que la sobriété du style se conduit de la sorte,
still a point worthy very serious investigation.

go over previous chapters incl lit review and refer back to things. bring things together. show the breadth and depth of my research!!!

relate all of these things back to my topic of AMC

- index terms vocab vs google index DONE
- discuss fig 6.2 (in relation to DH methodologies)
- query expansion == pataphysicalisation
- expand 6.1 (abusing stuff, creating own rules, oulipo)
- lookup vs exploratory
- storing rhyming data in index or other additional things like ranking
- creative use of NLP examples (search web for refs) DONE
- CONSTRAINTS + oulipo again
- clinamen change stopwords
- syzygy synsets, print each step
- poem and list side by side
- newspaper legal corpus

§ 8 & 9

A lot of the more theoretical aspects of this research have been discussed in chapters 8 and 9. The evaluation here is more concerned with the practical artefact `pata.physics.wtf` and its interpretation.

The 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

§ 2 Looking back over the inspirations for this project described in chapter 2, some 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 reasons was ***humour***. Pataphysics has an intrinsic kind of humour I wanted to include in the whole presentation of the artefact.

Syzygy Surfer

§ 2.1 The influence of the Syzygy Surfer cannot be overstated. It forms the immediate predecessor to my research. It should not be forgotten that the authors of the Syzygy Surfer are part of my supervisory team. This is where
§ 10.2 the initial ideas for the pataphysical algorithms came from. There are important differences as well though. For example, pataphors were never implemented even though this was originally suggested. Also, the concept of patadata was never really conceptualised properly. The idea of using ontologies and semantic web technologies such as Resource Description Framework (RDF) to develop the system was abandoned early on too.

Faustroll Library

§ 2.2 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.

Queneau's 100 thousand million poems

§ 2.3 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. The outcome is beautiful.

Celestial Emporium of Benevolent Knowledge

§ 2.4 Borges chinese encyclopedia has been an inspiration right from the start. The subtle humour in it is great. The sort of semantic logic behind it was modeled through the pataphysical algorithms. § 10.2

Yossarian

§ 2.5 This has been interesting to watch but if anything was more of a counter inspiration. An example of what I do not want to do. Their so-called metaphoric search engine is hyped but it is wholly unclear of how their algorithm actually create these metaphors. It is hard to compare against this as it is so different even though we share some of the same goals or principles.

Library of Babel

The library of babel is a great project which has only indirectly influence 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 even 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 priniciples[1], the § 10.2 influence on this project cannot be underestimated. The algorithms created could even be seen as an oulipian technique themselves.

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 was influenced by coder culture—and most of all perhaps: this thesis.

remove yossarian criticism

12.2 Pataphysicalisation

The internal transformation of a query term to the final results is essentially what I call the ***pataphysicalisation*** process. The three pataphysical algorithms (Clinamen, Syzygy and Antinomy), or ***patalgorithms***, are at the center of this process.

[1] Remember that the OULIPO was founded as a subcommittee of the "Collège de Ṕataphysique" in the 60's.

1. User enters single query term,
2. system transforms query term into list of pataphysicalised terms,
3. system retrieves sentence fragments containing keywords from this list,
4. system displays sentence fragments in various formats.

It is quite interesting to compare these 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 below in table 12.1.

	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 12.1: Comparison of patalgorithms showing a selection of results for each.

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.

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) 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 relation-

ship between the original query term and the results. Again, this may not be immediatly 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 2, 11, 23, 42, 59, 92, 119, 134, 150, 172, 202, 213, **??**, and 251). It also highlights the difference between the two corpora. Poems based on the Faustrol corpus have a very different sound and feel to it than ones based on the Shakespeare corpus.

remember to replace some of the chapter poems with shakespeare

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 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.

Results

A ‘result’ in this case is one line (a sentence fragment). This column shows

[2]Unless of course we believe the legends that Shakespeare didn’t write those works by himself. . .

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 for query term 'flower'

Corpus	**Query**	**Results**	**Reverbs**	**Origins**	**Poems**
Faustroll	flower	89	24	18	7.8×10^{10}
Shakespeare	flower	157	15	38	3.8×10^{14}
Faustroll	clear	542	79	23	1.3×10^{22}
Shakespeare	clear	1445	72	38	1.5×10^{28}
Faustroll	troll	124	16	16	4.4×10^{12}
Shakespeare	troll	327	14	38	1.1×10^{19}
Faustroll	fania	9	2	6	1
Shakespeare	fania	15	2	14	1

Table 12.2: Faustroll versus Shakespeare in numbers

the total number of results found by the three algorithms combined. Individual results appear only once but the keyword in contains can appear in several of the results.

Reverbs

A 'reverberation' is one of the terms in the list of keywords produced by the pataphysicalisation process. The list cannot contain duplicates but each reverberation cab 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.

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 not the queries in table 12.2 have not found results in all of the texts. The query 'clear' 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.

Regarding the final column showing the number of possible poems, let's look at the Shakespeare—clear row. There are 1445 number of 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 $1512589724855111243225614516 9$ (shortened to 1.5×10^{28} in table 12.2).

A slightly different angle to consider is a comparison of these kind of numbers between each of the algorithms. Table 12.4 shows the numbers of results, reverberations and origins for the Clinamen, Syzygy and Antinomy algorithms using

[3] The original book by Queneau contains 10 sonnets with 14 lines each. This means the total number of poems producable by the book is 10^{14}

four example query terms ('clear', 'shine', 'disorder' and 'stuck') for each of the two corpora ('Faustroll' and 'Shakespeare').

The first immediate observation surely must be that the Antinomy algorithm produces the fewest results, in two 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.

§ ??

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 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'.

§ ??

§ 12.2.3

Looking at the origins column highlights how the Shakespeare corpus mostly produces results from each of its 38 texts. This may have several reasons.

- According to (Efron and Thisted 1976) Shakespeare used 31534 different words in his works, about half of which he only used once (14376). They cite the total number of words used in his corpus as 884647.
-

faustroll corpus has 3 empty files. stopwords quality of faustroll corpus (typos)

shakespeare vocab in index: 23398 faustroll vocab in index: 78893 faustroll english vocab in index: 49040

shakespeare 00 17742 01 24371 02 26460 03 22812 04 16200 05 29213 06 28786 07 32031 08 25721 09 27631 10 27447 11 22776 12 26784 13 25833 14 25817 15 21704 16 20870 17 27513 18 22887 19 18233 20 22873 21 22254 22 23360 23

File	Words	Lang
00	22230	en
01	419456	en
02	33078	en
03	26843	en
04	103637	fr
05	3853	en
06	134328	fr
07	39937	fr
08	4548	fr
09	11309	fr
10	110518	en
11	31286	de
12	78983	fr
13	75293	fr
14	20371	en

[4] This is hardcoded to be Jarry's *Exploits and Opinions of Doctor Faustroll, Pataphysician*. Section 12.2.3 discusses what would happen if we changed the base document to something else.

		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
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

Table 12.4: Results-Reverberations-Origin numbers per algorithm

17233 24 22450 25 27866 26 23297 27 31090 28
25864 29 22159 30 17415 31 19634 32 21634 33
27537 34 21153 35 18293 36 25949 37 2568

The full list of stopwords per language can be
§ **??** found in appendix **??**.

HERE

add french query term

pp_sent faustroll clear with all sentences 8751
faustroll clear with the first sentence only 542

Francois Rabelais: Gargantua and Pantagruel

term: cellar
positions: [4448, 18718, 68678, 110318, 192486,
267241, 352502, 352565]
sentence: rope wine is let down into a cellar

sentences:
- rope wine is let down into a cellar
- bread and holy water of the cellar
- year who had a cool cellar under ground

- cellar
- that Nick in the dark cellar
- on the cellar door
- in mind of the painted cellar in the oldest city in the world
- and the painted cellar also

This is also a lot more time consuming. A way around this would be to store each sentence with each word in the index directly.

calculate averages of how many positions the average index term has

12.2.2 Index

look up google index or other examples or crawls

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 amoutn of data to be processed.

The Faustroll corpus for example contains 28 texts[5]

check which ones are empty

. 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 a poem whereas the Edgar Allan Poe file is a whole collection of his works. The whole size of the Faustroll corpus is 10MB. The Shakespeare corpus is much more evenly distributed as all of his works are separated out into individual text files of an average size of around 150KB. The total size of the Shakespeare corpus is only 5.3MB.

Now, the size of the index is interesting. Processing the Faustroll corpus alone produced an index of 12.4MB. That's larger than the 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 buts also french and spanish and german terms since the faustroll corpus is multi-lingual.

[5] This is technically not true since a few of those files are empty

how big is the new combined index?? - its 35.2MB

how many words in shakespeare vocab compared to faustroll vocab?

storing rhyming data in index or other additional things like ranking

12.2.3 Clinamen

The clinamen function uses the damerau-levenshtein algorithm to create pataphysicalised words. 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 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.

show clinamen results with a real dictionary rather than a base text

Changing the base text in Clinamen

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)

```
altar, bear, car, cheer, clean, clear, dear, ear, fear, hear, lead,
 ↪  liar, near, plead, rear, swear, tear, wear
```

clinamen with up to 1 error

faustroll clear:
clean, 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

Table 12.5: changing base in clinamen

Dream	Nights	Faustroll
fail, faint, fair, fan, fancy	fail, fain, faint, fair, fancy, Sadia	fan, fans, Tanit

Table 12.6: changing base in clinamen

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

Table 12.7: changing base in clinamen

faustroll fania:
-

faustroll moss:
loss, mass, most

clinamen with up to 2 errors

faustroll clear:
altar, cedar, cellar, clad, clean, clear, clearly, dear, ear, Fear, fear, hear, lead, leaf, leap, near, pleas, rear, swear, year

faustroll fania:
fan, fans, Tanit

faustroll moss:
ass, Bosse, bows, Boys, cost, 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

clinamen with up to 3 errors

faustroll 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

faustroll 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

faustroll 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

12.2.4 Syzygy

check out what is happening with the hyponyms in the getnym function

The syzygy function goes through the following process.

semantic hierarchy visualised?

It shows each step in the algorithm for the query term ‘clear’.

1. A set of synonyms (a “synset”) is generated.

SYZYGY synsets: [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 hyper-

nym 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_pro 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 []

12.2.5 Images

The image search can produce quite intersting results as well. A search for "blue kitten" on Flickr produces the following results: "[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 fig below XYZ).

Figure 12.2: image spiral flickr

For Getty the image search works slightly differently due to its API restrictions. The query “blue kitten” gets turned into the word “racy” which then calls the API to retrieve ten results (see below).

The difference is staggering.

12.3 Design

Content Perception
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 focus groups using questionnaires and interviews or eye tracking tools to find out what users prefer or perceive as more creative for example.

poem vs list here

12.4 Science Fiction

Where does this project stand in the wider world and the progress of computing, AI and creativity? AI and robotics is alluring as a research topic because it is so prevelant in Science Fiction. Computer creativity rarely plays a centrol role though. We regularly read headlines that tell us that yet another kind of AI-bot

Figure 12.3: image spiral getty

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 perfectly human-like (and yet cannot walk up stairs). There are so many examples of advances that are hailed as the next big thing which aren't all that great.

12.4.1 AI

This is also evident in games, for example Virtual Reality (VR) and Augmented Reality (AR). The Oculus Rift and similar systems are advertised so much you might believe they are actually about to hit mainstream and every kid will own a VR console and headset. Yet they are still way too expensive to be mainstream and motion sickness is also still an issue (and probably always will). These industries are so "hip" any publication is seen as the new cool thing without taking into account the history and work that has been done previously in perhaps slightly different disciplines. This is the case for example with a recent article on VR sickness and how to compat it. This is a well known problem already—motion sickness already exists in normal games. Similar to epilepsy problems.

find links for motion sickness

find links for epilepsy

find links for oculus rift and pokenmon go etc

AR has very recently received a massive boom thanks to Pokenmon Go (released in Australia, New Zealand and the USA in July 2016). It has become a phenomenon since then.

find pokemon links

What about IBM's Watson[6], Microsofts Twitter AI Chatbot[7], Google's AlphaGo[8] and Hanson Robotics Sophia robot[9]? How does this relate to my work? Practially of course they are all unrelated. On a deeper level though we can start asking interesting questions.

IBM Watson
: Watson is a question answering expert system. It famously won against human Jeopardy! champions in 2011.

Microsoft Chatbot

Google AlphaGo
: AlphaGo is a system for playing the game Go. It won against a top human professional player in 2015.

Hansen Sophia

I think these are interesting examples to study since they are supposedly on the forefront of AI development. Life-like robots like Sophia still live in the 'uncanny valley'. Her voice is creepy and unhuman, her intelligence or her capabilities if understanding conversations are clearly flawed (as shown by her viral remark about supporting genocide).

check

Watson is clever and fast in finding answers for specific questions but he still had problems with humour (e.g. BLAHBLA

find example

[6] See http://www.ibm.com/watson/
[7] See http://www.ibm.com/watson/
[8] See https://deepmind.com/alpha-go
[9] See http://www.hansonrobotics.com/

) but information lookup is arguably fairly easy and straightforward process within IR—sure, it requires processing power and memory storage or access but it is based on simple matching of keywords, not any fancy heuristic algorithms. Microsofts twitter chatbot went viral and users 'taught' it nasty swearwords

check

quickly and Microsoft had to take the bot down. It has since apologised although any official documentation on it has disappeared

check

. Google's AlphaGo has been hailed as a breakthrough in AI but similar to Watson it is a very targeted and limited program.

To me it seems the real breakthrough happens when (and if) the first robots appears which isn't as big as a house, can play Go, Chess and hide-and-seek, 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 causing a war, etc, etc—you get the picture. General AI is where it's at. Humans can do all the things we do. Children aren't born with only a single function. Imagine a world where humans only have one specialism and can;t do anything else. Mary is a Chess player but can't move her arms. Bob is a medical diagnosis expert but he can't hold a conversation. Movement, speech, memory—they are all vastly complex systems. And I haven't even touched creativity yet.

whats the point im making? how does this relate to my work?

Perhpas this 'uncanny valley' exists in creativity too. If a robot who looks vaguely human but not quite well enough, or he/she/it sounds almost human but not quite—perhaps if a robot can crack a joke like a human but not quite—perhaps this could be considered uncanny valley too? The philosphical zombies I men-
§ 9 tioend in chapter 9 live in this uncanny valley?

p and H creativity for computers?

12.4.2 Brains

I'm not talking about the beer or the zombie food but rather research into the human brain (or animal brains) and attempts to model it on a computer.

The motivation here is that once we understand how the brain works, perhaps we can understand how certain cognitive processes really work and this

of course include creativity.

This is no easy task of course. Chris Chatham talks about ten "important Differences Between Brains and Computers"[10] which give a good overview of some of the dificulties of trying to model a brain as is. We can't just do a 1-1 copy.

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

Chris Chatham

To bring this into perspective Ray Kurzweil claims the brain is capable of 10^{16} operations per second (2013, p.194). Japan's K-computer (the worlds largest super computer as of 2016) currently has that power—10 petaflops. The "Blue Brain Project" is aiming to model 10^17 bytes of memory and 10^{18} flops by 2023 (Kurzweil 2013, p.125).

find k-computer reference

There are currently some major research projects going on. One of them is the "Human Brain Project" (Walker 2012).

quotes:

Our brain consumes about 30W, the same as an electric light bulb, thousands of times less than a small supercomputer. (Walker 2012, p.17)

For environmental and business reasons, vendors have set themselves the goal of containing energy consumption to a maximum of 20 megawatts (Walker 2012, p.41)

[10] http://scienceblogs.com/developingintelligence/2007/03/27/why-the-brain-is-not-like-a-co/

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, p.41)

Cellular-level simulation of the 100 billion neurons of the human brain will require compute power at the exascale (1018 flops). (Walker 2012, p.41-42)

2017 petascale 50petabytes memory + 50 petaflops + <=4MW power

2021 exascale 200petabyte memory + 1exaflop

A second, equally important goal will be to prepare the procurement of the HBP Pre-exascale-supercomputer. By 2017/18, Jülich plans to procure a Big Data-centred system with at least 50 PBytes of hierarchical storage-class memory, a peak capability of at least 50 PFlop/s and a power consumption <= 4 MW. The memory and computational speed of the machine will be sufficient to simulate a realistic mouse brain and to develop first-draft models of the human brain. (The rest of the hardware roadmap targets an exascale machine in 2021/2022 with a capability of 1 EFlop/s and a hierarchical storage-class memory of 200 PB).[11]

Why Minds Are Not Like Computers (Schulman 2009) Software – Hardware == Mind – Brain ??? analogy

"The power of the computer derives not from its ability to perform complex operations, but from its ability to perform many simple operations very quickly."

Layers of abstraction in computers:
1. user interface
2. high level programming language
3. machine language
4. proessor microarchitecture
5. Boolean logic gates
6. transistors

layers of abstraction in brain:
1. personality?
2. Thinking?
3. Chemical /electrical signals/activity?
4. Divided Brain regions/structure
5. Neurons
6. Dendrites (input) and axons (output)?

[11] https://www.humanbrainproject.eu/high-performance-computing-platform

Computers are faster and better than humans in many tasks already.

> "The weaknesses of the computational approach include its assumption that cognition can be reduced to mathematics and the difficulty of including noncognitive factors in creativity." (Mayer 1999, p.457)

find references

neural networks and other models based on the brain

Perhaps we need to have that complete picture of how the brainw orks in order to understand human creativity. I would argue computer creativity is part of general AI, and for general AI we need massive amounts of general knoweldge.

common sense research

again talk about how this is relevant for my project

Expert Systems vs General AI Is computer creativity an expert system or does it fall into general AI?

Machines self-assessing Perhaps there is an argument that if humans are the only entities who can judge whether another human is being creative, then machines should be assessing themselves. This is a paradoxical concepts though. Since machines are products made my humans, they can never be autonomous in that sense. If machines had evolved like other animals besides us this argument might hold but obviously that is not the case.

12.5 Meta

12.5.1 Management

add file for appendix with full git history

On a different note, the project was completed over X years which includes an interruption and later on only a part time commitement.

I kept the project in a “git repository”. Git is a version ontrol system that allows users to roll-back on changes and I further pushed my work to GitHub to make

sure hardware failure or human error (i.e. lost or stolen property) would not affect my work.

To understand git you need to know what commits are. They are the thing where I save my current state of the project and give it a description.

Below you can see a shortened version of the timeline of my commits between 20XX and the time of submission of this thesis. A full version can be found in appendix XYZ. You can see from this the time between programming work I did on pata.physics.wtf and its predecessors.

add calendar screenshot of github contributions

links to git and github

```
*   10f61f9  Sun 08 May 2016  (HEAD -> api, origin/api) Merge remote-tra
|\
* | 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  (origin/thesis) 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 c
| * | 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[redacted]
| * | e1f1ead  Mon 03 Aug 2015  Stylesheets sorted and cleaned files[redacted]
* | |   9732d5b  Mon 03 Aug 2015  Merge branch 'w3' into thesis
|\ \ \
| |/ /
```

I also kept the thesis under git version control. Since the thesis was written in LaTeXyou could almost say I 'programmed' it. Below is an outline of the commit history for this thesis.

```
* 3f06260  Edited readme again
* c721b33  Edited readme
* ffbdb4b  Edited readme
* 8870b3d  Added gitignore file
* ba1a9c2  Second commit
* 244c4b3  First commit
```

12.5.2 Thesis

Part Spirals

Each new thesis part contains a word spiral based on a poem generated by pata.physics.wtf using the a part of the title as keyword. They represent the pataphysical (Archimedean) spiral.

1. Preface — ***pre***
2. Hello World — ***hello***
3. Tools of the Trade — ***trade***
4. The Core: Techno-Logic — ***core***
5. The Core: Techno-Practice — ***practice***
6. Meta-Logicalysis — ***meta***
7. Happily Ever After — ***after***
8. Postface — ***post***

Chapter Poetry

Each chapter opens with a poem generated by `pata.physics.wtf` using a part of the chapter title as keyword.

1. Introduction — ***intro***
2. Inspirations — ***inspiration***
3. Methodology — ***method***
4. Pataphysics — ***pata***
5. Creativity — ***creativity***
6. Technology — ***technology***
7. Evaluation — ***evaluation***
8. Foundations — ***foundation***
9. Interpretation — ***interpretation***
10. Implementation — ***implementation***
11. Applications — ***application***
12. Patanalysis — ***patanalysis***
13. Aspirations — ***aspirations***
14. Observations — ***observations***

say more, check keywords, potentially generate new poems

creative analysis

literary deconstruction and recombining to make new creative output?
perception of results (poetry, source, algorithm)
discuss applications from before (stimulates creative detour away from the obvious)

How does this relate to Oulipo and Pataphysics?

Perhaps this is where I should talk a bit about the perception of results in their different output formats/styles. The poetry is automatically read with more gravity. Sorting by sources is a game of exploration or algorithms which becomes a game of 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. When listed as a list they immediately become more boring.

talk abit about what the original plan was for some of the big changed elements in the website, e.g. the image search running 10 times on different keywords rather than running once with 10 results for the same keyword.

DELETE EVERYTHING FROM BELOW HERE:

DELETE THIS

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 that are not discouraged by what might appear to be nonsensical results. Users could include artists, writers or poets but equally anybody who is looking for out-of-the-box inspirations or simply a refreshingly different search engine to the standard.

The way we display and label results produced by the tool can influence how the user perceives them. The current prototype for example separates the results into its three components but we could have equally just mixed them all together. The less transparent the processes in the background (e.g. which algorithm was used, how does the result relate to the query precisely, etc.) are for the user, the more difficult it might be to appreciate the search.

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, either 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 base 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 produce a larger index which would possibly find many more matches through WordNet and end in a more varied list of results.

From a computer science perspective it could be used as one of 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.

The 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, 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.

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 never finishes. Especially with creative products, where the functional requirements are more fluid perhaps, it is always tempting to add, improve, replace bits.

software refactoring

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
§ 10 submission of this thesis is described in chapter 10 and further elaborated on
§ 12 in the Patanalysis chapter.

Here, in this chapter I will lay out some of the potential/likely further work for this project. This may continue on a private basis or in a more academic environment. I have grouped these ideas into two main categories: ***technical*** and ***theoretical***.

13.1 Technical

write these out all in one list and then group them as fit

Responsive spirals Currently the image and video spirals 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. Percentages

Scalable image sizes At the moment images are retrieved at a given 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 them look squished and pizelated. This limits the available choice of results through the API.

Square aspect ratio Another issue is the aspect ratio of images and videos. For the spiral they need to be square. I currently achieve this by squishing them as opposed to cropping them or specifying an option in the API calls to only retrieve square images.

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 them on the page and currently this is achieved my left aligning them and having a fixed 'absolute' position on the page. Ideally this would also be centered as in the random poems.

Startup performance 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. 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.

Query speed 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.

Preprocessing corpora At this point the texts in the corpora consist of almost unedited plaintext ('.txt') files[1]. Newlines and whitespace formatting varies, as does language and quality of spelling. OCR SOURCES Generally, chapter headings, chapter numberings, etc are left untouched. The Shakespeare corpus contains poetry and plays for example. STAGE DIRECTIONS With the plays, scene information is kept, voice details are 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.

Sentence fragments Currently the way results sentences are retrieved for the text search is based on punctuation. This means once a pataphysicalised 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.

[1]For text files downloaded from Project Gutenberg, the Gutenberg specifc copyright notices have been removed to only contain the relevant body of text

More APIs Currently X APIs are used[2]. 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.

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 display style for the results. The biggest problem for this is API restrictions. Alternatively a ready-made index or crawl could be used but these are typically many terrabytes in size and have a cost attached. Crawling the Web myself is not an option due to the computational power, time and space required to do so.

Audio search Originally audio search was going to be a part of this project. This has been abandoned due to time constraints. However it could be added using an API such as SoundClouds. 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 search in audio waves.

More 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) 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 appendix XYZ).

§ 11

Poetry rhyming scheme One of the biggest points for future work is to introduce a rhyming scheme for the poetry results. This would involve some more NLP during the creation of the index. It would make the poems much more readable. See more in chapter XYZ.

§ 6

Random sentences Adding to the source of random sentences used in the top and bottom banner on the website should be an ongoing endevour.

Custom API It would be great to develop a custom API for this the search tool. 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 Adnrew Dennis.

§ 11
§ 11

[2] Flickr, Getty, Bing, MicrosoftTranslator and YouTube

WordNet vocabulary The vocabulary in WordNet is limited. According to it's website (https://wordnet.princeton.edu/) it contains 117000 'synsets'[3] This affects two of my algorithms. Because of the way the process works, the link between Wordnet and source texts, results may be limited.

check

WordNet Antonyms The antinomy algorithms relies on WordNets 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. 12.2.1

Stemming Stemming could increase the number of results found by the algorithms. (See chapter XYZ). A danger of increasing the output of the pataphysicalisation is always that results become more boring. If the query term and potential matches were compared based on their stemmed form

Queneau's poems It would be nice to actually add Queneau's poem texts into the coprus of Faustroll as little easter eggs.

Bitmap algorithms The image and video search currently rely on extrenal APIs and user tags to work. One option to approach this in a totally differnet way would be to write algorithms that analyse and pataphysicalise the bitmaps themselves. So this could mean we could have a reverse image search that finds images related the original bitmap in pataphysical way or other.

Index One idea for the pataphysicalisation process was to add 'patadata' to the index. This could include pronounciation tags for example to make an implementation of a rhyming scheme for the poetry easier. So each word in teh index dictionary would contain the following items.

```
(``tree'': [``l_00'': [24,566,4990], ``s_14'': [234,5943]], IPA data)
```

add ipa data or whatever is best for the rhyming stuff

storing rhyming data in index or other additional things like ranking

[3] Synonyms—"words that denote the same concept and are interchangeable in many contexts"—are grouped into unordered sets called synsets.

Stopwords Using a different set of stopwords to see if that makes a difference. For example we could use a spanish set of stopwords on an english text. OR the other way around.

13.2 Creative NLP

Section 6 N-grams are a NLP technique introduced in chapter 6. 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. On 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 ffor example instead of having a Maximum Likelihood Estimation (MLE) (see chapter XYZ and formular 6.12) we could have a 'Maximum Obscurity Estimation' defined as:

$$P(w_n \mid w_{n-N+1}^{n-1}) = \frac{C(w_{n-N+1}^{n-1} w_n)}{C(w_{n-N+1}^{n-1})} \tag{13.1}$$

work the maths out here for this example of MOE

Similarly, we could could play with maximum entropy models as shown on page 112 (see chapter XYZ) together with Parts-of-Speech (POS) tagging. What if we rigged the probability such that instead of 'in Quebec' ranking high for a 'location' POS tag, it now ranks high as a 'drug'?

Again there are endless possipilities of abusing these kinds of systems to create AMC. This is also very reminiscent of OULIPO techniques. We could create a whole new language grammar based on pataphysical principles.

Another example of interesting uses of NLP for AMC is playing with homonyms and heteronyms. 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'). There are similar techniques in the OULIPO. Homophones are often used to create puns (and remember—puns are syzygy's 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

look into rhyming tags in nlp

NLP would also be useful for introducing a rhyming pattern into auto-generated poetry. BY doing POS tagging with pronounciation data, we could retrieve sentences that match the sound of the last word of the previous line, etc.

https://wordnet.princeton.edu/wordnet/man/wngloss.7WN.html for glossary

fix all chapter XYZ mentions

group these into better sub groups and make them proper sections rather than paragraphs

13.3 Theoretical

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`) accoriding 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 wether or not they thought that using the framewokr was beneficial to them or confusing.

Questionnaires I have shied away from doing a questionnaire study because of several reasons. One is that due to the creative and subjective nature of the artefact, opinions on it may vary wildly and I don't see how I could derive useful unbiased data from that. Yes, it depends what questions you ask. But even if I managed to get some half-decent data, what would that tell me? Half of the people like my site, the other half don't?

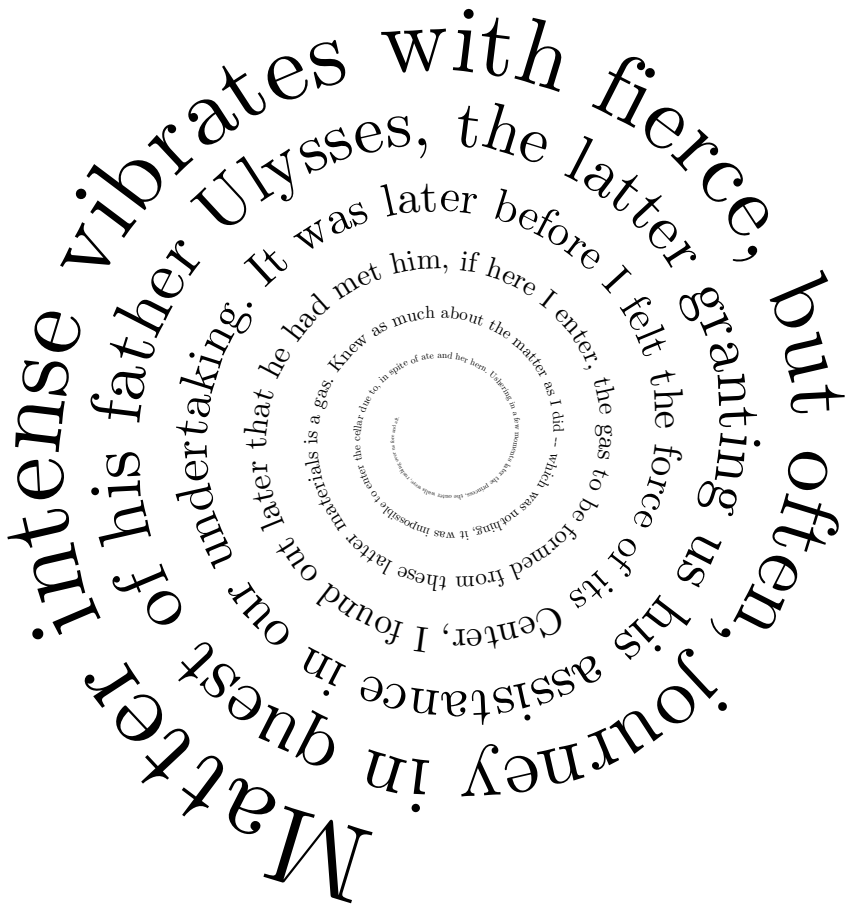
Eye-tracking To study the effects of using different styles of presenting the same results an eye-tracking experiment could be done. This would involve setting up participants with the necessary equipment and then introduce them to the website and moniter their eye movements as they naviagte the site. 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 has no clear immediate purpose and that may put users off.

Performance Benchmarks?

Part VI

H∀PPILY ∑V∑R ∀FT∑R?



Observations

14

Paying no attention to his fellow mites,
mérite pas que vous fassiez attention à moi,
and told him to look after a calf she had bought,
and whilst he was looking at it attentively.

Phedon the fact affirm'd,
comment peux,
ne faites aucune attention à mon air,
in fact.

For sure Ulysses in your look appears,
was nearly out of her mind,
I omitted none of the common forms attending a royal audience.

And the consequences attending thereupon,
impotent of mind,
shape at the moment of looking at the time.

summarise thesis, contributions etc. conclude by comparing against introduction

a wide range of subject areas such as computer science, psychology, linguistics, literature, art and poetry, languages and mathematics.

refer back to these in conclusion

14.1 Outroduction

The last XYZ chapters have explained in probably too much detail what AMC is and how to evaluate it. Given that this spans so many different disciplines the contextual background information necessary to understand the research was presented in a broad literature survey in chap XYZ. This also posed a problem for choosing the right methodology for the project. In the end a transdisciplinary approach was chosen as described in chap XYZ with a heavy component of iterative exploratory rapid-prototyping to develop an artefact to demonstrate what AMC is.

This artefact is presented on `pata.physics.wtf`. It is an artwork dedicated to AMC, pataphysics, OULIPO and programming culture.

A critique of computer creativity and its current evaluation formed the starting point for a new framework which was introduced in chap XYZ. The general conlcusion of the thesis was made up of the critical analysis and further work chapters as awell as this final concluding chapter right at the end.

The appendix contains various code snippets and peripheral pieces vaguely related or relevant for parts of this thesis. The code of the website is included on a CD CHECK attached to the back of the front cover. Of course the website is also available online at `pata.physics.wtf`.

check if i need to submit a CD?

14.2 Answers

In the introduction I asked several questions that I attempted to asnwer with my research. 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.

add chapter references

Can computers or algorithms be considered creative?

§ 7 § 9 In short: no. In chapters 7 and 9 I have gone into great detail of why I believe that this cannot happen any time soon (see argument of zombies). They can be 'creative' (adj/adv CHECK) but the source of the creativity is the programmer of the machine not the machine itself.

Can pataphysics facilitate creativity?

Yes. Pataphysics provides many principles which can be turned into technicques and constraints which is well known to be able to support creativity

§ 8 (see chapter 8). This is also evident in the OULIPO and their use of con-

§ 5 straints (see chapter 5).

Can a creative process be automated or emulated by a computer?

Yes, in theory. It mainly depends how you define the creative process and

§ 5 § 9 that is fairly subjective. See more in chapter 5 and 9.

Can human and computer creativity be objectively measured?

§ 9 No. As discussed in chapter 9 since the perception of creativity is subjective

§ 9.2.2 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.

Can information retrieval be creative?

§ 12 Yes. There are many ways this can be achieved too as mentioned in chapter 12.

Can search results be creative rather than relevant?

Yes, although this is also subjective. What is creative to some might not be creative to everybody. The artefact also nicely showed the difference in

§ 12 perception of results simply based on design of the content (see chapter 12).

14.3 Contributions

mention to whom these could be useful

write more

[1] Inspired by Time Berners-Lee's articles on the Web in 1998—/urlhttp://www.w3.org/DesignIssues/Architecture.html

This doctoral project can be broken down into four main contributions.

- Three pataphysical search algorithms (clinamen, syzygy and antinomy).
- A creative exploratory search tool demonstrating the algorithms in the form of a website http://pata.physics.wtf.
- A set of subjective parameters for defining creativity.
- An objective framework for evaluating creativity.

In a more practical sense this project has spawned several publications, talks and exhibitions (a full list is in preface **??**). Further talks were given by Andrew Hugill at various conferences and events throughout the world where he mentioned my work. My publications were cited in other academic publications and my website was mentioned on Reddit[2]. My job here is done.

sec **??**

14.4 And Finally

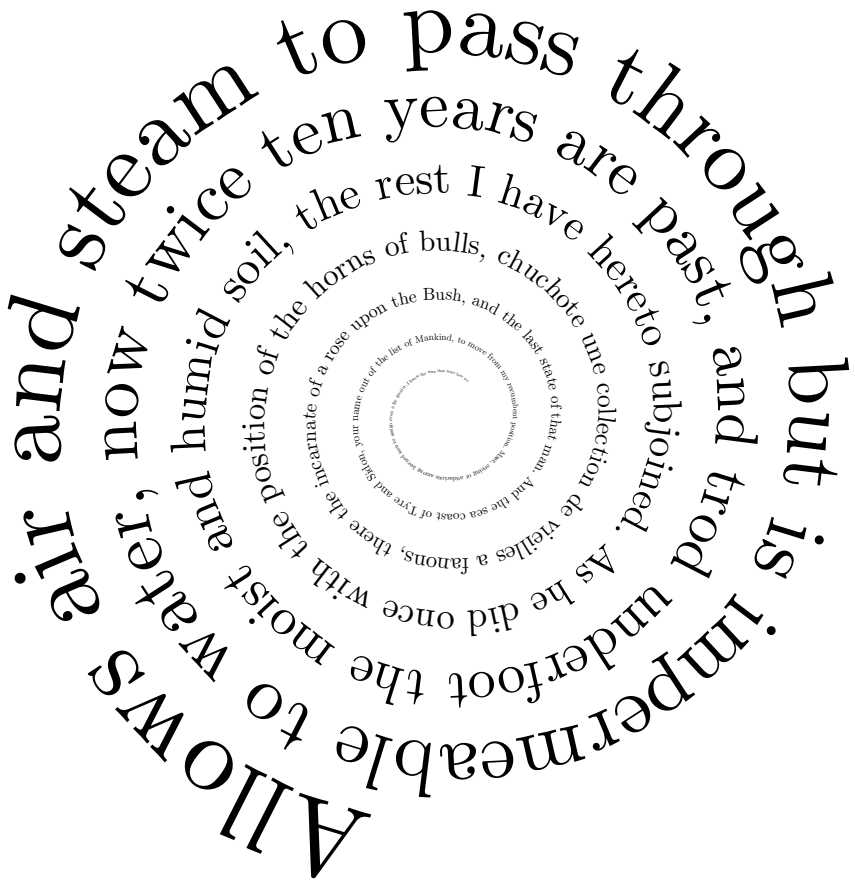
Pataphysics is the science. . .

[2] Although absolutely nobody seemed interested in it. No idea who posted it or how he found it.

INTERLUDE III

Part VII

POST☹



References

Agichtein, Eugene, Eric Brill and Susan Dumais (2006). 'Improving web search ranking by incorporating user behavior information'. In: ***ACM SIGIR conference on Research and development in information retrieval***. New York, New York, USA: ACM Press, p. 19 (cit. on p. 103).

Amaral, Jose Nelson et al. (2006). 'About Computing Science Research Methodology'. In: (cit. on pp. 26, 27).

Baeza-Yates, Ricardo and Berthier Ribeiro-Neto (2011). ***Modern Information Retrieval: The Concepts and Technology Behind Search***. Addison Wesley (cit. on pp. 93, 98, 99, 101, 102, 104, 120, 121).

Baidu (2012). ***Baidu About*** (cit. on p. 94).

Baldi, Pierre and Laurent Itti (2010). 'Of bits and wows : A Bayesian theory of surprise with applications to attention'. In: ***Neural Networks*** 23, pp. 649–666 (cit. on pp. 130, 158).

Bao, Shenghua et al. (2007). 'Optimizing Web Search Using Social Annotations'. In: ***Distribution***, pp. 501–510 (cit. on p. 103).

Barthes, Roland (1967). 'The Death of the Author'. In: ***Aspen 5,6***. the birth of the reader must be ransomed by the death of the Author (cit. on p. 155).

Basile, Jonathan (2015). ***The Library of Babel***. URL: https://libraryofbabel.info/ (visited on 10/12/2015) (cit. on p. 18).

Bastos Filho, Carmelo et al. (2008). 'A novel search algorithm based on fish school behavior'. In: ***IEEE International Conference on Systems, Man and Cybernetics***, pp. 2646–2651 (cit. on p. 103).

Baudrillard, Jean (2007). ***Pataphysics*** (cit. on p. 51).

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. 66).

Bharat, Krishna and George Mihaila (2000). 'Hilltop: A Search Engine based on Expert Documents'. In: ***Proc of the 9th International WWW***. Vol. 11 (cit. on p. 103).

Bird, Steven, Ewan Klein and Edward Loper (2009). ***Natural Language Processing with Python***. Sebasopol, CA: O'Reilly Media (cit. on p. 106).

Boden, Margaret (2003). ***The Creative Mind: Myths and Mechanisms***. London: Routledge (cit. on pp. 64, 67–69, 75, 98, 129, 136, 142–144).

Boek, Christian (2002). ***'Pataphysics: The Poetics of an Imaginary Science***. Evanston, Illinois: Northwestern University Press (cit. on pp. 6, 44, 52–55, 181).

Borges, Jorge Luis (1964). ***Labyrinths - Selected Stories and Other Writings***. New York: New Directions (cit. on pp. 17, 57).

– (1999). ***Collected fictions***. Trans. by Andrew Hurley. Penguin (cit. on p. 57).

– (2000). 'The Analytical Language of John Wilkins'. In: ***Selected Non-Fictions***. Ed. by Eliot Weinberger. London: Penguin Books, pp. 229–232 (cit. on pp. 4, 12, 14, 57).

– (2010). ***La biblioteca de Babel***. Reclam (cit. on p. 57).

Borges, Jorge Luis and L.S. Dembo (2010). 'Interview with Borges'. In: ***Contemporary Literature*** 11.3, pp. 315–323 (cit. on p. 57).

Borges, Jorge Luis and Margarita Guerrero (1957). ***Book of Imaginary Beings***. Trans. by Andrew Hurley. Viking (cit. on p. 57).

Brin, Sergey and Larry Page (1998a). 'The anatomy of a large-scale hypertextual Web search engine'. In: ***Computer Networks and ISDN Systems*** 30.1-7, pp. 107–117 (cit. on pp. 94, 101).

– (1998b). 'The PageRank Citation Ranking: Bringing Order to the Web'. In: ***World Wide Web Internet And Web Information Systems***, pp. 1–17 (cit. on pp. 94, 101, 148).

Brotchie, Alastair (2011). ***A supplement***. UK: Atlas Press (cit. on pp. 46–50, 53).

Brotchie, Alastair and Stanley Chapman, eds. (2007). ***Necrologies***. London: Atlas Press (cit. on p. 49).

Brotchie, Alastair, Stanley Chapman et al., eds. (2003). ***'Pataphysics: Definitions and Citations***. London: Atlas Press (cit. on pp. 43, 44).

Brotchie, Alistair, ed. (1995). ***A True History of the College of 'Pataphysics - 1***. Trans. by Paul Edwards. London: Atlas Press (cit. on p. 53).

Brown, Mark (2011). ***Patrick Tresset's robots draw faces and doodle when bored***. URL: http://www.wired.co.uk/news/archive/2011-06/17/sketching-robots (visited on 24/01/2016) (cit. on p. 154).

Burdick, Anne et al. (2012). ***Digital Humanities***. Cambridge, Massachusetts: MIT Press (cit. on pp. 28–31, 34, 84–88, 90, 91, 139, 140).

Burnham, Douglas (2015). 'Immanuel Kant: Aesthetics'. In: ***Internet Encyclopedia of Philosophy*** (cit. on p. 132).

Candy, Linda (2006). ***Practice Based Research:A Guide***. Tech. rep. (cit. on p. 38).

– (2012). 'Evaluating Creativity'. In: ***Creativity and Rationale: Enhancing Human Experience by Design***. Ed. by J.M. Carroll. Springer (cit. on pp. 60, 126, 127, 129, 158).

Candy, Linda and Ernest Edmonds, eds. (2011). ***Interacting: Art, Research and the Creative Practitioner***. Libri Publishing (cit. on p. 152).

Chalmers, David (1996). ***The Conscious Mind***. Oxford University Press (cit. on pp. 152, 156).

Cohen, Harold (1999). ***Colouring Without Seeing: A Problem in Machine Creativity***. URL: %7Bhttp://www.kurzweilcyberart.com/aaron/hi_essays.html%7D (visited on 24/01/2016) (cit. on p. 155).

Colton, Simon (2008a). 'Computational Creativity'. In: ***AISB Quarterly***, pp. 6–7 (cit. on pp. 76, 77, 124).

– (2008b). 'Creativity versus the perception of creativity in computational systems'. In: ***In Proceedings of the AAAI Spring Symp. on Creative Intelligent Systems*** (cit. on pp. 75, 124).

Colton, Simon, Alison Pease and Graeme Ritchie (2001). ***The Effect of Input Knowledge on Creativity*** (cit. on p. 123).

Colton, Simon and Geraint A Wiggins (2012). 'Computational Creativity: The Final Frontier?' In: ***Proceedings of the 20th European Conference on Artificial Intelligence***. Montpellier, France: IOS Press, pp. 21–26 (cit. on pp. 77, 78, 138–140).

Corbyn, Zoe (2005). ***An introduction to 'Pataphysics*** (cit. on p. 51).

Cruickshank, Douglas (nd). ***Why Anti-Matter Matters*** (cit. on p. 52).

Cutshall, James Anthony (1988). 'The Figure of the Writer - Alfred Jarry'. Thesis. University of Reading, p. 258 (cit. on p. 50).

Damerau, Fred J (1964). 'A Technique for Computer Detection and Correction of Spelling Errors '. In: ***Communications of the ACM*** 7.3, pp. 171–176 (cit. on p. 182).

Daumal, Rene (2012). ***Pataphysical Essays***. Trans. by Thomas Vosteen. Cambridge, Massachusetts: Wakefield Press (cit. on p. 53).

De Bra, Paul, Geert-jan Houben et al. (1994). 'Information Retrieval in Distributed Hypertexts'. In: ***Techniques*** (cit. on p. 103).

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, pp. 183–192 (cit. on p. 103).

– (1994b). 'Searching for Arbitrary Information in the WWW: the Fish Search for Mosaic'. In: ***Mosaic A journal For The Interdisciplinary Study Of Literature*** (cit. on p. 103).

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. 104).

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. 101).

Dennis, Andrew (2016). 'Investigation of a patadata-based ontology for text based search and replacement'. University of London (cit. on p. 203).

Dictionary, Oxford English (2015). ***animal, n.*** URL: http://www.oed.com/view/Entry/273779 (visited on 10/12/2015) (cit. on p. 16).

Dijkstra, Edsger W. (1988). ***On the Cruelty of Really Teaching Computing Science*** (cit. on pp. 142, 157).

Ding, Li et al. (2004). 'Swoogle: A semantic web search and metadata engine'. In: ***In Proceedings of the 13th ACM Conference on Information and Knowledge Management. ACM*** (cit. on p. 103).

Drucker, Johanna (2009). ***SpecLab: Digital Aesthetics and Projects in Speculative Computing***. University of Chicago Press (cit. on pp. 81–84, 147).

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 pp. 83, 84).

Du, Zhi-Qiang et al. (2007). 'The Research of the Semantic Search Engine Based on the Ontology'. In: ***2007 International Conference on Wireless Communications, Networking and Mobile Computing***, pp. 5398–5401 (cit. on p. 103).

Dubbelboer, Marieke (2009). ''UBUSING' CULTURE'. Thesis. Rijksuniversiteit Groningen, p. 233 (cit. on p. 51).

Eden, Amnon H. (2007). 'Three Paradigms of Computer Science'. In: ***Minds and Machines*** 17.2, pp. 135–167 (cit. on pp. 78, 79).

Edmonds, E. and L. Candy (2010). 'Relating Theory, Practice and Evaluation in Practitioner Research'. In: ***Leonardo*** 43.5, pp. 470–476 (cit. on p. 38).

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. 221).

Elton, Matthew (1995). 'Artificial Creativity: Enculturing Computers'. In: ***Leonardo*** 28.3, pp. 207–213 (cit. on p. 75).

Foucault, Michel (1966). 'The Order of Things - Preface'. In: ***The Order of Things***. France: Editions Gallimard. Chap. Preface, pp. xv–xxiv (cit. on pp. 58, 211).

Garcia-Molina, Hector, Jan Pedersen and Zoltan Gyongyi (2004). 'Combating Web Spam with TrustRank'. In: ***In VLDB***. Morgan Kaufmann, pp. 576–587 (cit. on p. 103).

Gelernter, David (1994). ***The Muse in the Machine***. London: Fourth Estate Limited (cit. on p. 73).

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. 103).

Google (2012). ***Google Ranking*** (cit. on p. 94).

Haveliwala, Taher H (2003). 'Topic-Sensitive PageRank: A Context Sensitive Ranking Algorithm for Web Search'. In: ***Knowledge Creation Diffusion Utilization*** 15.4, pp. 784–796 (cit. on p. 103).

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. 144).

Heisenberg, Werner (1942). ***Ordnung der Wirklichkeit***. Trans. by M.B. Rumscheidt and N. Lukens (cit. on p. 36).

Hendler, Jim and Andrew Hugill (2011). 'The Syzygy Surfer : Creative Technology for the World Wide Web'. In: ***ACM WebSci 11*** (cit. on pp. 5, 12, 14, 132, 189).

– (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. 5, 12, 14, 142).

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. 103).

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. 156).

Holz, Hilary J et al. (2006). 'Research Methods in Computing : What are they , and how should we teach them ?' In: ***ITiCSE Innovation and technology in computer science education***, pp. 96–114 (cit. on pp. 25, 26).

Hotho, Andreas et al. (2006). 'Information retrieval in folksonomies: Search and ranking'. In: ***The Semantic Web: Research and Applications, volume 4011 of LNAI***. Springer, pp. 411–426 (cit. on p. 103).

Hugill, Andrew (2012). ***'Pataphysics: A Useless Guide***. Cambridge, Massachusetts: MIT Press (cit. on pp. 43, 45, 46, 53, 54, 56, 211).

– (2013). 'Introduction: transdisciplinary learning for digital creative practice'. In: ***Digital Creativity*** 24.3, pp. 165–167 (cit. on pp. 78, 139).

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. 6, 37, 79–81, 138–140).

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 p. 141).

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. 70, 144, 154).

Jarry, Alfred (1996). ***Exploits and Opinions of Dr Faustroll, Pataphysician***. Cambridge, MA: Exact Change (cit. on pp. 4, 12, 13, 44, 51, 55, 177, 182).

– (2006). ***Collected Works II - Three Early Novels***. Ed. by Alastair Brotchie and Paul Edwards. London: Atlas Press (cit. on pp. 132, 211).

Jeh, Glen and Jennifer Widom (2002). 'SimRank: A Measure of Structural Context Similarity'. In: ***In KDD***, pp. 538–543 (cit. on p. 103).

Jordanous, Anna (2015). 'Four PPPPerspectives on Computational Creativity'. In: ***International Conference on Computational Creativity*** (cit. on p. 137).

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. 125, 158).

– (2012). 'Evaluating Computational Creativity: A Standardised Procedure for Evaluating Creative Systems and its Application'. PhD thesis. University of Sussex (cit. on pp. 125, 126, 158).

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. 63, 76, 125).

Jorn, Asger (1961). 'Pataphysics - A Religion In The Making'. In: ***Internationale Situationniste*** 6 (cit. on p. 52).

Jurafsky, Daniel and James H Martin (2009). ***Speech and Language Processing***. London: Pearson Education (cit. on p. 106).

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. 103).

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. 65, 136).

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. 154).

Kleinberg, Jon M (1999). 'Authoritative sources in a hyperlinked environment'. In: ***journal of the ACM*** 46.5, pp. 604–632 (cit. on pp. 102, 148).

Kleinberg, Jon M et al. (1999). 'The Web as a graph : measurements, models and methods'. In: ***Computer*** (cit. on p. 102).

Koestler, Arthur (1964). ***The Act of Creation***. London: Hutchinson and Co (cit. on pp. 64, 144).

Kurzweil, Ray (2013). ***How to Create a Mind***. London: Duckworth Overlook (cit. on p. 234).

Levenshtein, Vladimir I (1966). 'Binary codes capable of correcting deletions, insertions, and reversals '. In: ***Soviet Physics Doklady*** 10.8, pp. 707–710 (cit. on p. 182).

Luo, Fang-fang, Guo-long Chen and Wen-zhong Guo (2005). 'An Improved 'Fish-search' Algorithm for Information Retrieval'. In: ***2005 International Conference on Natural Language Processing and Knowledge Engineering***, pp. 523–528 (cit. on p. 103).

Macdonald, Craig (2009). 'The Voting Model for People Search'. In: ***Philosophy*** (cit. on p. 101).

Maeda, John (2001). ***Design by Numbers***. MIT Press (cit. on p. 152).

Manning, Christopher, Prabhakar Raghavan and Hinrich Schuetze (2009). ***Introduction to Information Retrieval***. Cambridge UP (cit. on pp. 100, 106, 107).

Marchionini, Gary (2006). 'From finding to understanding'. In: ***Communications of the ACM*** 49.4, pp. 41–46 (cit. on pp. 97, 98).

Marchionini, Gary and Ben Shneiderman (1988). 'Finding facts vs. browsing knowledge in hypertext systems'. In: ***Computer*** 21.1, pp. 70–80 (cit. on pp. 94, 97).

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 p. 113).

Mathews, Harry and Alastair Brotchie (2005). ***Oulipo Compendium***. London: Atlas Press (cit. on pp. 18, 19).

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. 60, 61, 122, 126, 158, 236).

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. 153–155).

– (2013). ***Robot Ethics: The Boundaries of Machine Ethics***. Leicester (cit. on p. 91).

Microsoft (2012). ***Bing Fact Sheet*** (cit. on p. 94).

Miller, George A. (1995). 'WordNet: a lexical database for English'. In: ***Communications of the ACM*** 38.11, pp. 39–41 (cit. on p. 184).

Minsky, Marvin (1980). 'K-Lines : A Theory of Memory'. In: ***Cognitive Science*** 33.4, pp. 117–133 (cit. on pp. 73, 74).

– (1988). ***The Society of Mind***. Simon and Schuster, p. 336 (cit. on pp. 73, 74).

Miyamoto, Sadaaki (1988). ***Information Retrieval based on Fuzzy Associations*** (cit. on p. 101).

Miyamoto, Sadaaki (2010). ***Fuzzy Sets in Information Retrieval and Cluster Analysis (Theory and Decision Library D)***. Springer, p. 276 (cit. on p. 101).

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. 101).

Motte, Warren (2007). ***Oulipo, A primer of potential literature***. London: Dalkey Archive Press (cit. on p. 142).

Neeley, J. Paul (2015). ***Introducing the NEW Yossarian***. email communication (cit. on p. 17).

Newell, A, J. G. Shaw and H. A. Simon (1963). ***The Process Of Creative Thinking***. New York: Atherton (cit. on p. 144).

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. 103).

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. 35, 36).

Partridge, Derek and Jon Rowe (1994). ***Computers and Creativity***. Oxford: Intellect (cit. on pp. 64, 65, 71).

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 AISB*** (cit. on p. 123).

Pease, Alison, Simon Colton et al. (2013). 'A Discussion on Serendipity in Creative Systems'. In: ***Proceedings of the 4th International Conference on Computational Creativity***. Vol. 1000. Sydney, Australia: University of Sydney, pp. 64–71 (cit. on p. 123).

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. 69, 76, 123, 124).

Peters, Tim (2004). ***PEP 20 – The Zen of Python*** (cit. on pp. 20, 21).

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. 124, 158).

Poincare, Henri (2001). ***The Value of Science***. Ed. by Stephen Jay Gould. New York: Modern Library (cit. on pp. 64, 80, 137).

Polya, George (1957). ***How To Solve It***. 2nd. Princeton, New Jersey: Princeton University Press (cit. on pp. 65, 80, 96, 137).

Queneau, Raymond (1961). ***One Hundred Thousand Billion Poems***. Gallimard (cit. on pp. 4, 14).

Raczinski, Fania (2016). ***Emails***. personal communication. feedback for his bachelor project (cit. on p. 204).

Raczinski, Fania and Dave Everitt (2016). 'Creative Zombie Apocalypse: A Critique of Computer Creativity Evaluation'. In: ***International Symposium of Creative Computing*** (cit. on p. 151).

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: ACM New York, NY, USA, pp. 274–280 (cit. on p. 6).

Ramesh, V., Robert L. Glass and Iris Vessey (2004). 'Research in computer science: an empirical study'. In: ***journaltitle of Systems and Software*** 70.1-2, pp. 165–176 (cit. on p. 26).

Rhodes, Mel (1961). 'An analysis of creativity'. In: ***The Phi Delta Kappan*** 42.7, pp. 305–310 (cit. on pp. 63, 135).

Ritchie, Graeme (2001). 'Assessing creativity'. In: ***AISB '01 Symposium on Artificial Intelligence and Creativity in Arts and Science***. Proceedings of the AISB'01 Symposium on Artificial Intelligence, Creativity in Arts and Science, pp. 3–11 (cit. on pp. 76, 123).

– (2007). 'Some Empirical Criteria for Attributing Creativity to a Computer Program'. In: ***Minds and Machines*** 17.1, pp. 67–99 (cit. on pp. 76, 123).

– (2012). 'A closer look at creativity as search'. In: ***International Conference on Computational Creativity***, pp. 41–48 (cit. on p. 129).

Schmidhuber, Juergen (2006a). 'Developmental robotics, optimal artificial curiosity, creativity, music, and the fine arts'. In: ***Connection Science*** 18.2, pp. 173–187 (cit. on p. 159).

– (2006b). ***New millennium AI and the Convergence of history*** (cit. on p. 122).

Schuetze, Hinrich (1998). 'Automatic Word Sense Discrimination'. In: ***Computational Linguistics*** (cit. on p. 101).

Schuetze, Hinrich and Jan Pedersen (1995). ***Information Retrieval Based on Word Senses*** (cit. on p. 101).

Schulman, Ari (2009). 'Why Minds Are Not Like Computers'. In: ***The New Atlantis*** 23, pp. 46–68 (cit. on p. 235).

Searle, John (1980). 'Minds, Brains, and Programs'. In: ***Behavioral and Brain Sciences*** 3.3, pp. 417–457 (cit. on p. 156).

Shattuck, Roger (1959). ***The Banquet Years***. London: Faber (cit. on p. 49).

Shu, Bo and Subhash Kak (1999). 'A neural network-based intelligent metasearch engine'. In: ***Information Sciences*** 120 (cit. on p. 103).

Singh, Push (2005). 'EM-ONE: An Architecture for Reflective Commonsense Thinking'. PhD thesis. Massachusetts Institute of Technology (cit. on p. 13).

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. 101).

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. 91, 140, 160).

Sternberg, Robert J (1999). ***Handbook of creativity***. Cambridge University Press, p. 490 (cit. on pp. 63, 66, 69, 135, 143).

– (2006). 'The Nature of Creativity'. In: ***Creativity Research journal*** 18.1, pp. 87–98 (cit. on p. 66).

Sutcliffe, Alistrair and Mark Ennis (1998). 'Towards a cognitive theory of information retrieval'. In: ***Interacting with Computers*** 10, pp. 321–351 (cit. on p. 96).

Taye, Mohammad Mustafa (2009). 'Ontology Alignment Mechanisms for Improving Web-based Searching'. PhD thesis. De Montort University (cit. on p. 103).

Thomas, Sue et al. (2007). 'Transliteracy: Crossing divides'. In: ***First Monday*** 12.12 (cit. on pp. 28, 85).

Turing, Alan (1950). 'Computing Machinery and Intelligence'. In: ***Mind*** 59, pp. 433–460 (cit. on p. 156).

Varshney, Lav R et al. (2013). 'Cognition as a Part of Computational Creativity'. In: ***12th International IEEE Conference on Cognitive Informatics and Cognitive Computing***. New York City, USA, pp. 36–43 (cit. on pp. 130, 158).

Ventura, Dan (2008). 'A Reductio Ad Absurdum Experiment in Sufficiency for Evaluating (Computational) Creative Systems'. In: ***5th International Joint Workshop on Computational Creativty***. Madrid, Spain (cit. on pp. 76, 124).

Vian, Boris (2006). ***'Pataphysics? What's That?*** Trans. by Stanley Chapman. London: Atlas Press (cit. on p. 53).

Vries, Erica de (1993). 'Browsing vs Searching'. In: ***OCTO report 93/02*** (cit. on pp. 97, 98).

Walker, Richard (2012). ***The Human Brain Project***. Tech. rep. HBP-PS Consortium (cit. on pp. 156, 234, 235).

Wallas, Graham (1926). ***The Art of Thought***. Jonathan Cape (cit. on pp. 64, 80, 137).

Walsh, Dave (2001). ***Absinthe, Bicycles and Merdre*** (cit. on p. 48).

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. 211).

Widyantoro, D.H. and J. Yen (2001). 'A fuzzy ontology-based abstract search engine and its user studies'. In: ***10th IEEE International Conference on Fuzzy Systems*** 2, pp. 1291–1294 (cit. on pp. 101, 103).

Wiggins, Geraint A (2006). 'A preliminary framework for description, analysis and comparison of creative systems'. In: ***Knowledge Based Systems*** 19.7, pp. 449–458 (cit. on pp. 75, 76, 129).

Yang, Hongji (2013). 'Editorial'. In: ***International journal of Creative Computing*** 1.1, pp. 1–3 (cit. on pp. 36, 79, 138).

Yossarian (2015). ***Yossarian*** (cit. on pp. 16, 17).

GLOSSARY

bisociation
: Two self-consistent but habitually incompatible frames of reference intersecting to give rise to a new creative idea.. 64

GET
: An Hypertext Transfer Protocol (HTTP) method. Allows a client (browser) to request data from a specified resource on a given web server.. 187

holonym
: The relationship between a term denoting the whole and a term denoting a part of, or a member of, the whole. That is, 'X' is a holonym of 'Y' if Ys are parts of Xs, or 'X' is a holonym of 'Y' if Ys are members of Xs. For example, 'tree' is a holonym of 'bark', of 'trunk' and of 'limb.' Holonymy is the opposite of meronymy.. 116

hypernym
: A hyponym shares a type-of relationship with its hypernym. For example, pigeon, crow, eagle and seagull are all hyponyms of bird (their hypernym); which, in turn, is a hyponym of animal.. 116

POST
: An HTTP method. Allows a client (browser) to submit data to be processed to a specified resource on a given web server.. 187

KTHXBYE