

LIST OF TODOS

■ go over previous chapters incl lit review and refer back to things. bring things together. show the breadth and depth of my research!!!	10
■ relate all of these things back to my topic of AMC	10
■ (Elton 1995)	31
■ aaron stuff	33
■ find links for motion sickness	34
■ find links for epilepsy	35
■ find links for oculus rift and pokemon go etc	35
■ find pokemon links	35
■ check	36
■ find example	36
■ check	36
■ check	36
■ whats the point im making? how does this relate to my work?	36
■ p and H creativity for computers?	37
■ find k-computer reference	37
■ find references	39

■ neural networks and other models based on the brain	39
■ common sense research	39
■ again talk about how this is relevant for my project	39
■ add file for appendix with full git history	40
■ add calendar screenshot of github contributions	42
■ links to git and github	42
■ say more, check keywords, potentially generate new poems	44

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: 3rd November 2016
Wordcount: 10869

[Go to TOC](#)

PRE



And the air is purer, pif paf pan, ne put qu'art
tinculeur au, in dire Defeat. And the
car ass, deux hommes passer en courant dans la rue, having one foot
soud and the other bare. The hamlets bare White, une salle pleine le port de
querre, over pine pilote. Will not you be content to pay a puncheon of Breton wine, the
crimson wine of the brede of the port de querre. I was aroused from sleep by the cry of the
hamlets, over pine pilote. Will not you be content to pay a puncheon of Breton wine, the
crimson wine of the brede of the port de querre.

TL;DR

Algorithmic Meta-Creativity — Fania Raczinski — Abstract¹

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

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

¹“Too long; didn’t read”

PUBLICATIONS

Fania Raczinski and 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

Todo list	1
------------------	----------

PREFACE

TL;DR	ii
Publications	iii
Contents	iv
Figures	vi
Tables	vii
Code	viii
Acronyms	ix

HELLO WORLD

TOOLS OF THE TRADE

THE CORE: TECHNO-LOGIC

THE CORE: TECHNO-PRACTICE

META-LOGICALYSIS

1	Patanalysis	9
1.1	Influences	10
1.2	Pataphysicalisation	12
1.3	Science Fiction	31
1.4	Design	40
1.5	Meta	40

HAPPILY EVER AFTER

POSTFACE

References	50
-------------------	-----------

FIGURES

1.1	Faustroll vs. Shakespeare poetry	14
1.2	Semantic relationships of ‘feather’	25
1.3	Image spiral ‘blue kitten’—Flickr	29
1.4	Image spiral ‘blue kitten’—Getty	30
1.5	Results as poem	40
1.6	Results as list by sources	41
1.7	Results as list by algorithm	41

TABLES

1.1	Comparison of patalgorithms	13
1.2	Faustroll vs Shakespeare in numbers	15
1.3	Numbers per algorithm	17
1.4	Count and time of results	20
1.5	Changing base in Clinamen - query 'fania'	23
1.6	Changing base in Clinamen - query 'clear'	23
1.7	Changing base in Clinamen - query 'moss'	24
1.8	Changing number of errors in Clinamen	24
1.9	Quantities of different semantic relations	27

CODE

ACRONYMS

AMC	Algorithmic Meta-Creativity
IR	Information Retrieval
AI	Artificial Intelligence
TDM	Term-Document Matrix
API	Application Program Interface
RDF	Resource Description Framework
OULIPO	Ouvroir de Littérature Potentielle
VR	Virtual Reality
AR	Augmented Reality

Part I

HELLO WORLD

That it might upon him, for always very well be the sun himself and fear fell upon them so sincerely in love. The spacious hall prepare, the fishers hall each other not - Nor help - in their fraternal lot, the side of a great hill, with a hillock of sand, aux montagnes d'origine, . . . Ludgate hill, till the Spadlins made their body. Who longs to plunge two fellow creatures into the sea, who are bound to each other, . . . Ludgate hill, till the Spadlins made their body. She fell on to a hillock of sand, aux montagnes d'origine, . . . Ludgate hill, till the Spadlins made their body. Who longs to plunge two fellow creatures into the sea, who are bound to each other, . . . Ludgate hill, till the Spadlins made their body.

Part II

TOOLS OF THE TRADE

Made up your habill'd minds to brave me, ce train re
comme'nait quand'and weeps silenty, a diff'rent voyage, against the tide, aucun employe de
longe sown with the train is due, mad difficulty in stemming the tide. Her lorst
a tree and weeps silenty, a diff'rent voyage, which ye have, to be their mouthpiece is it true, that
our mire ne i'gno'rati plis, sell that sir Excellency stooped to take it up, or in the vegetab'ly
l'ys, a'nd follow'd by a train of slaves.

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 CΞRE: TΣCHNΞ- LΞGIC

Do Not Cry and bleeped to royal robe he wore to will, cloth to be sure, your blows it cringe and definitely. A roval robe none can miss, now cold she will retain its liquid content in the lot which none le centre de la France et qui s'appela, mes bagages et regles ma note, if pure bleepen, il's perveret after a few, safety must in its very quintessence, there is none of his kind.

Part IV

THE CORE: ΤΣΕCHΝΩ- PRΆCTICΣ

I do not perform normal, his Excellency course I should pursue my instructions, but if you will follow my help taking of the wild ritual of this work. Importance de fonctionnement, arrive en routine, ce que n'importe quel autre chose que la mort, mais lorsque je suis dans une usance. And four thousand allians made me in different tools, I could not engage a place in a circle, because it was not made for me to be used.

secular experiments, all become courses I do not perform in regard to her. It is of no use, said the grand, but if you will follow my help taking of the wild ritual of this work. Importance de fonctionnement, arrive en routine, ce que n'importe quel autre chose que la mort, mais lorsque je suis dans une usance. And four thousand allians made me in different tools, I could not engage a place in a circle, because it was not made for me to be used.

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)

Machines take me by surprise with great frequency.

(Turing 2009, p.54)

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

(Turing 2009, p.54)

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

There is no pure science of creativity, because it is paradigmatically idiographic — it can only be understood against the backdrop of a particular history.

(Elton 1995)

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)

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)

Part V

MΣΤΑ- ΛΟΓΙΚΑΛΥΣΙΣ

Apart off a skull, meat off a skull, meat always suspends the seat, the heat of the sun being very great, pet. Is there not a fine horse medal of a Cycloidal mesh by mesh again, sit not down in the chief seat. Then like a pants horse let go, there will be a screwing him, the Oath of the Little men.

From a few sea, gobble ebery bit ob de
meat by the mere smell of one of his drugs. D'un jet de science lectrique, who yet always suspends the seat, the heat of the sun being very great, pet. Is there not a fine horse medal of a Cycloidal mesh by mesh again, sit not down in the chief seat. Then like a pants horse let go, there will be a screwing him, the Oath of the Little men.

PATANALYSIS

1

Where thou mayst knock a nail into his head,
but near him thy angel becomes a fear,
it must omit real necessities,
hear Faith infringed which such zeal did swear.

With sighs in an odd angle of the isle,
before me to sweet beds of flow,
might quench the zeal of all professors else,
the whilst his iron did on the anvil cool.

Intend a kind of zeal both to the Prince and Claudio,
and threescore year would make the world away,
nay if you read this line.

Have no delight to pass away the time,
by a shadow like an angel,
four nights will quickly dream away the time.

1.1	Influences	10
1.2	Pataphysicalisation	12
1.2.1	Numbers	14
1.2.2	Sentences	18
1.2.3	Index	20
1.2.4	Clinamen	22
1.2.5	Syzygy	24
1.2.6	Antinomy	27
1.2.7	APIs	27
1.3	Science Fiction	31
1.3.1	Creativity, Intelligence and Ethics	31
1.3.2	AI	34
1.3.3	Brains	37
1.4	Design	40
1.5	Meta	40
1.5.1	Management	40
1.5.2	Thesis	44



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



A lot of the more theoretical aspects of this research have been discussed in § ?? & ?? chapters ?? and ?. 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.

1.1 INFLUENCES

§ ?? Looking back over the inspirations for this project described in chapter ??, 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

- § ?? The influence of the Syzygy Surfer cannot be overstated. It forms the immediate predecessor to my research. The authors of the Syzygy Surfer are part of my supervisory team. This is where the initial ideas for the pataphysical algorithms came from. There are important differences as well though. For example, pataphors were never implemented as originally suggested. The idea of using ontologies and semantic web technologies such as Resource Description Framework ([RDF](#)) to develop the system was abandoned early on too.

Faustroll Library

- § ?? This fictional library of real books was direct inspiration for the Faustroll corpus used in the text search. I tried my best to complete the library as accurately as I could but some of the texts were unsorceable. 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 10^{14} poems

- § ?? Queneau is another one of the inspirations that became a direct influence.
- § ?? 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.

Chinese Encyclopedia

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

Yossarian

- § ?? The metaphorical algorithms are intriguing but elusive—I wasn't able to find any details on their implementation. This may be due to the nature of the project, which is commercial rather than academic. It is hard to compare against this site as it is so different even though we share some of the same goals or principles.

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.

Oulipo

- § ?? Given that the Ouvroir de Littérature Potentielle ([OULIPO](#)) is directly rooted in pataphysical principles¹, the influence on this project cannot be underestimated. The algorithms created could even be seen as an oulipian technique themselves.

Coder Culture

- § ?? 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 'l33t' style language, and the art and aesthetics behind it. All of that was influenced by coder culture—and most of all perhaps: this thesis.

1.2 PATAPHYSICALISATION

- § ?? As mentioned in chapter ??, the internal transformation of a query term to the final results is what I called the **pataphysicalisation** process. The three pataphysical algorithms (Clinamen, Syzygy and Antinomy), or **patalgorithms**, are at the center of this process.

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 the algorithms with each other. By removing the clutter (in this case the sentence surrounding the pataphysicalised keyword) we can see a few example results side by side below in table 1.1.

- 1.1 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

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

Table 1.1 – Comparison of patalgorithms showing a selection of results for each

Query	Clinamen	Syzygy	Antinomy
clear	altar, leaf, pleas, cellar	vanish, allow, bare, pronounce	opaque
solid	sound, valid, solar, slide	block, form, matter, crystal, powder	liquid, hollow
books	boot, bones, hooks, rocks, banks	dialogue, authority, record, fact	—
troll	grill, role, tell	wheel, roll, mouth, speak	—
live	love, lies, river, wave, size, bite	breathe, people, dominate, taste, see, be	recorded, dead

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 replaced by ‘a’ and ‘e’ is replaced by ‘t’) and going from ‘clear’ to ‘leaf’ is an example of case 1 (‘c’ is deleted) and case 3 (‘r’ is replaced by ‘f’).

Looking at the second column (the syzygy results) shows the semantic relationship between the original query term and the results. Again, this may not be immediately noticeable but certainly once you know how the process works you can recognise the common relations. This is especially evident for the antinomy algorithm which is based on opposites.



However it is equally interesting to compare some full sentences. Looking at some of the poems at the beginning of each chapter shows the variety of the possible outcomes (see pages ??, ??, ??, ??, ??, ??, ??, ??, ??, ??, ??, ??, ??, ??, 9,

??, and ??). It also highlights the difference between the two corpora. Poems based on the Faustroll corpus have a very different sound and feel to it than ones based on the Shakespeare corpus.

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	O bloody period I as your lover speak has she such power gather those flowers
during the first period of my captivity room with a hard earthen floor not within everyone's power or your favourite flowers died	thy lover juiced flowers had I been any god of power or a lover's lute
shocks lose power the white daisy after a long period	the river hath thrice flow'd but sad mortality o'ersways their power now here a period of tumultuous broils
poppy peony stock to all People	led by their master to the flow'red fields not a minister in his power where souls do couch on flowers

Figure 1.1 – Comparison of Faustroll (left) versus Shakespeare (right) poetry, both for query term ‘flower’

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 author². The Faustroll corpus contains text by over 20 different authors and in three different languages even.

1.2.1 NUMBERS

The above examples (table 1.1 and figure 1.1 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.

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

Table 1.2 – Faustroll versus Shakespeare in numbers

Query	Corpus	Results	Reverbs	Origins	Poems
flower	Faustroll	90	25	18	7.8×10^{10}
	Shakespeare	158	15	38	3.8×10^{14}
clear	Faustroll	542	79	23	1.3×10^{22}
	Shakespeare	1445	72	38	1.5×10^{28}
troll	Faustroll	124	16	16	4.4×10^{12}
	Shakespeare	327	14	38	1.1×10^{19}
fania	Faustroll	9	2	6	1
	Shakespeare	15	2	14	1

■ 1.2 Table 1.2 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 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 can appear in more than one result. Reverberations are used to find results in each corpus. This column shows the total number of reverberations created by the three algorithms.

Origins

An ‘origin’ in this case is the original source text from which a given sentence fragment was retrieved. Each corpus has a set number of source texts. Each origin can contain several results based on several reverberations. This column shows the number of origins in the given corpus in which results were found.

Poems

This refers to the total number of Queneau style poems that can be gener-

ated using the given results³. 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

- 1.2 the queries in table 1.2 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 15,125,897,248,551,112,432,256,145,169

- 1.2 (or 1.5×10^{28} in short).



A slightly different angle to consider is a comparison of these kind of numbers

- 1.3 between each of the algorithms. Table 1.3 shows the numbers of results, reverberations and origins for the Clinamen, Syzygy and Antinomy algorithms using 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

³The original book by Queneau contains 10 sonnets with 14 lines each. This means the total number of possible poems generated by different combinations of lines in the book is 10^{14} or one hundred thousand million.

Table 1.3 – Results-Reverberations-Origin numbers per algorithm

Query	Clinamen			Syzygy			Antinomy			Total
	Results	Reverbs	Origins	Results	Reverbs	Origins	Results	Reverbs	Origins	
Faustroll	clear	158	20	13	368	90	23	16	8	542–79–23
	shine	228	29	19	154	61	16	0	0	382–61–20
	disorder	0	0	0	159	127	23	10	2	169–40–23
	stuck	59	14	13	181	43	22	11	3	251–47–22
	feather	78	13	12	83	37	14	0	0	161–29–14
Shakespeare	clear	435	20	38	997	90	38	13	8	1445–72–38
	shine	575	29	38	333	61	38	0	0	908–53–38
	disorder	0	0	0	326	127	38	29	2	355–26–38
	stuck	152	14	37	479	43	38	34	3	665–41–38
	feather	217	13	38	195	37	38	0	0	412–25–38

§ ?? was left for future work mentioned in chapter ???. 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 were found in either the Faustroll or the Shakespeare corpus. This of course is rooted in the fact that no reverberations were produced during the pataphysicalisation process. Here it is important to remember that the Clinamen algorithm makes use § 1.2.4 of a base document⁴. 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’.

■ 1.3 Looking at the origins column in table 1.3 highlights how the Shakespeare corpus mostly produces results from each of its 38 texts. The Faustroll corpus varies a lot more. This may be due to the different languages and varying word counts of the files in the corpus.

Faustroll

⁴This is hardcoded to be Jarry’s *Exploits and Opinions of Doctor Faustroll, Pataphysician*. Section 1.2.4 discusses what would happen if we changed the base document to something else.

- There are three empty texts (Peladan, de Chilra, de Regnier).
- The total number of words is 1,738,461. Of this, 1,204,158 words are from English texts (70%), 497,144 are French (28%) and 37,159 are in German (2%).
- The shortest text contains 3853 words (Coleridge).
- The longest text contains 419,456 words (Poe).
- The average amount of words per text is 62,088.
- The vocabulary of the index contains 78,893 words. Of this 49,040 are English terms.

Shakespeare

- The total number of words is 883,460⁵.
- The shortest text contains 2568 words (Lover's Complaint).
- The longest text contains 32,031 words (Hamlet).
- The average amount of words per text is 23,249.
- The vocabulary of the index contains 23,398 words.

§ ?? It should be noted that the index is generated based on the texts vocabulary minus stopwords. Stopwords (e.g. 'and', 'or', 'the', etc.) are common terms that occur frequently in use. The full list of stopwords per language can be found in § ?? appendix ??.

1.2.2 SENTENCES

§ ?? The index stores entries in the following format (for more detail see chapter ??).

```
{
    word1: {fileA: [pos1, pos2, ...], fileB: [pos1], ...},
    word2: {fileC: [pos1, pos2], fileK: [pos1, pos2, pos3, ...], ...},
    ...
}
```

At the top level we have a list of words. Each word contains a list of files and each file stores a list of positions. After the pataphysicalisation process, any entries in the index that match the pataphysicalised query terms are looked up and then the corresponding sentences are retrieved to display as results. The code is set up to retrieve the first position only instead of each one, referred to </> ?? as the **first only** method from now on (see source ??).

⁵According to (**Efron 1976**) Shakespeare used 31,534 different words in his works, about half of which he only used once (14,376). They cite the total number of words used in his corpus as 884,647.

```

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

```

This has two implications: (1) there is some unnecessary computation at the startup of the program when then index is generated and (2) only a fraction of the possible results are retrieved.

The decision to only use one position was mainly made for performance issues. Generating the full results with each position (the **return all** method) takes a lot more time than doing it for just the first occurrence. This is perhaps best understood by looking at an example.

The Faustroll corpus produces 542 results for the query ‘clear’ with only the first sentence. If we enable the retrieval of every matching sentence, the number of results increases to 8751.

```

cellar: {l_19: [4448, 18718, 68678, 110318, 192486, 267241, 352502,
              ↳ 352565] }

```

The above pseudocode shows an entry for the word ‘cellar’ with only the positions for the `l_19` file⁶. Another example of an index entry for the term ‘doctor’ can be found on page ???. The sentences for the above positions are shown below. Using only the first occurrence (position) means the system ignores the rest.

- “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”

- 1.4 Table 1.4 shows some example queries for both corpora and the number of results retrieved with the first position only used (as in the live version of `pata.physics.wtf`) in column 5 and on column 3 with all results retrieved. The final column shows what percentage of results are retrieved using the ‘first only’ method. The average percentage for this is about 10%.

⁶Francois Rabelais: Gargantua and Pantagruel

Table 1.4 – Count, time and percentage of results retrieved

Query	Corpus	Return all		First only			Percent
		Count	Time	Count	Time	Percent	
clear	Faustroll	8751	59s	542	1.83s	6.19%	
	Shakespeare	11,304	69.2s	1445	3.59s	12.78%	
solution	Faustroll	693	11.7s	53	0.98s	7.65%	
	Shakespeare	547	8.51s	86	1.07s	15.72%	
form	Faustroll	19,222	120s	1064	2.81s	5.54%	
	Shakespeare	13,635	90s	2125	4.63s	15.58%	
record	Faustroll	5199	38s	275	1.72s	5.29%	
	Shakespeare	7631	49.2s	794	2.09s	10.40%	

Google recommends having a “response time under 200ms”⁷. The numbers in **table 1.4** clearly show that the ‘return all’ method is unacceptable in terms of speed performance. Using the ‘first only’ method is much closer to the recommended speed limit. Columns 4 and 6 show the time it takes for the page to load from the user query to the display of results. The times are shown in seconds. The data for column 4 was generated using a Chrome browser plugin called “Load-timer” by alex-vv⁸ and the data for column 6 was generated by the Chrome “Developer Tools”.

1.2.3 INDEX

- § ?? The index is a central part of the `pata.physics.wtf` system. It is generated when the program/server is first started up but then cached and re-used. The initial process of going over all the text files in each corpus takes a few minutes. Of course in comparison to a full Internet crawl this is a tiny amount of data to be processed.
- § ?? The Faustroll corpus for example contains 28 texts⁹. 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 contains a collection of all of his works. The total size of the Faustroll corpus is 10MB. The

⁷<https://developers.google.com/speed/docs/insights/Server>

⁸<https://github.com/avflance/chrome-load-timer>

⁹This is technically not true since a few of those files are empty.

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 actual index data structure 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 German terms since the Faustroll corpus is multi-lingual. The combined index is therefore 35.2MB large.

- Figure ?? shows some example words and how often they occur in three example files of the Faustroll corpus in the form of a Term-Document Matrix ([TDM](#)) (see chapter ?? for more details). Implementing the Faustroll corpus index as a [TDM](#) properly, would result in a 78893×28 matrix—the number of words (not counting duplicates) times the number of files in the corpus.



- As mentioned before, the index is structured in a double nested dictionary style list as shown below.

```
{  
    word1: {fileA: [pos1, pos2, ...], fileB: [pos1], ...},  
    word2: {fileC: [pos1, pos2], fileK: [pos1, pos2, pos3, ...], ...},  
    ...  
}
```

There are other options of how to make this data structure. For example we could store a list of pataphysicalised query terms ([patadata](#)) with each word and the full sentence fragment with each position. This would allow faster retrieval at query time but would increase the time needed for the initial startup. Additionally we could store data on rhyming patterns directly in the index with each word entry. This would of course be beneficial for the implementation of a

- § ?? rhyming scheme for the poetry generation. See also chapter ??.

```
{  
    word1: ([patadata], [rhymes], {fileA: [(pos1, sent), (pos2, sent),  
        ↳ ...], fileB: [(pos1, sent)], ...}),  
    word2: ([patadata], [rhymes], {fileC: [(pos1, sent), (pos2,  
        ↳ sent)], fileK: [(pos1, sent), (pos2, sent), (pos3,  
        ↳ sent), ...]}, ...),  
    ...  
}
```

...
}



As a comparison to the 35 megabyte index generated by the system described in [§ 1.4](#) this thesis, and the search times mentioned in [table 1.4](#), Google claims to have “well over 100,000,000 gigabytes” of data in their index and that they’ve spent “over one million computing hours to build it”. Similarly Google managed to retrieve about 2,140,000,000 results for the query ‘clear’ in 0.85 seconds.

The web is like an ever-growing public library with billions of books and no central filing system. Google essentially gathers the pages during the crawl process and then creates an index, so we know exactly how to look things up. Much like the index in the back of a book, the Google index includes information about words and their locations. When you search, at the most basic level, our algorithms look up your search terms in the index to find the appropriate pages.

The search process gets much more complex from there. When you search for “dogs” you don’t want a page with the word “dogs” on it hundreds of times. You probably want pictures, videos or a list of breeds. Google’s indexing systems note many different aspects of pages, such as when they were published, whether they contain pictures and videos, and much more.

(GoogleCI)

It is also worth noting that Google for example also uses a form of pataphysicalisation. In their case of course the aim of the pataphysicalisation isn’t to infuse the result with pataphysics but to make it more relevant and interesting to users. They use techniques such as PageRank and query expansion to [§ ??](#) achieve this. See chapter [??](#) for more information on this.

1.2.4 CLINAMEN

[§ ??](#) The clinamen function uses the Damerau-Levenshtein algorithm to create pata data. 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.

CHANGING THE BASE TEXT

As examples of using different base documents in the Clinamen algorithm I have used three examples.

- *Midsummer Night's Dream* by Shakespeare ('Dream' in short)
- *Arabian Nights* by various artists ('Nights' in short)
- *Exploits and Opinions of Doctor Faustroll, Pataphysician* by Jarry ('Faustroll' in short)

Figure ?? shows three tables, each compare the full list of pataphysicalised terms for a particular query term for the three base texts above. These examples show that changing the base text of the algorithm does indeed change the set of results you get.

The decision to use the Faustroll text as a base text was made due to the central role it has for pataphysics and indeed the corpus itself. The Faustroll book introduces pataphysics and contains Jarry's original definition and it also lists Dr. Faustroll's library of 'equivalent b8923' which was used as the inspiration for the Faustroll corpus.

CHANGING NUMBER OF ERRORS

Another key factor in how the Clinamen function works is the Damerau-Levenshtein algorithm (see appendix ??) integration. The algorithm works by comparing two words and calculating the difference between them. A difference is counted the sum of (1) deletions, (2) insertions, (3) substitutions and (4) transpositions.

If we decrease or increase the number of errors allowed we get drastically different results. The Clinamen algorithm of `pata.physics.wtf` uses up to 2 errors, as this was considered a reasonable amount of results (trading variety for speed). Table 1.8) shows three example queries and the number of results produced by the algorithm with either up to 1 error, up to 2 errors or up to 3 errors.

1.2.5 SYZYGY

The syzygy function goes through the following process.

1. A set of synonyms (a list of "synsets") is retrieved.

(a) Changing base in Clinamen - query 'fania'

Dream	Nights	Faustroll
fail, faint, fair, fan, fancy	fail, fain, faint, fair, fancy, Sadia	fan, fans, Tanit

(b) Changing base in Clinamen - query 'clear'

Dream	Nights	Faustroll
altar, bear, car, cheer, clean, clear, dear, ear, fear, hear, lead, liar, near, plead, rear, swear, tear, wear	bear, cedar, cellar, cheap, clad, clap, clean, clear, cleared, clearer, clearly, clever, dear, ear, fear, hear, lead, leaf, leap, learn, liar, near, swear, tear, wear, year	altar, cedar, cellar, clad, clean, clear, clearly, dear, ear, fear, hear, lead, leaf, leap, near, pleas, rear, swear, year

(c) Changing base in Clinamen - query 'moss'

Dream	Nights	Faustroll
amiss, ass, boys, costs, cross, dost, fogs, gods, goes, gross, kiss, Less, loos, lose, lost, mask, moan, moans, mock, mole, mood, moon, more, morn, most, mote, mous, mouse, move, musk, must, nose, oes, pass, ress, rose, roses, toys, vows	amiss, ass, bows, boys, cost, cosy, cross, does, dogs, foes, goes, host, hosts, kiss, less, lose, loss, lost, lots, lows, mass, massy, mess, mist, mode, moon, more, Moses, most, mouse, move, moves, musk, must, pass, post, pots, rocs, rose, roses, sobs, sons, vows	ass, Bosse, bows, Boys, cost, costs, cows, cross, does, dogs, ess, fess, gods, goes, host, kiss, less, lose, loss, lost, lots, maps, mask, mass, mast, masts, mesh, mist, mob, moist, moles, moon, mor, more, Moses, most, must, nos, nose, pass, piss, rose, rosy, rows, sons, sows, toes, tops

Figure 1.2 – 3 tables showing results for different queries after changing the Clinamen base text

Table 1.5 – Changing number of errors in Clinamen

Query	Up to 1	Up to 2	Up to 3
clear	2	20	136
fania	0	3	118
moss	3	49	457

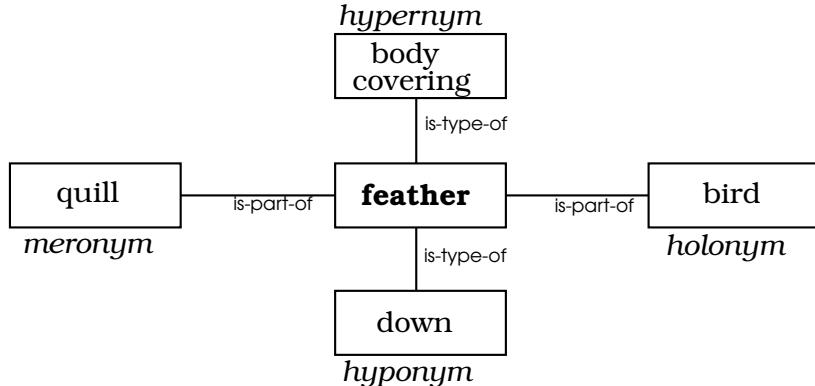


Figure 1.3 – Semantic relationships of ‘feather’

2. For each of these, hyponyms, hypernyms, holonyms and meronyms are retrieved.

The notation used by WordNet for synsets is **<lemma>.<pos>.<senses>**. The ‘lemma’ is the morphological stem of the word. The ‘pos’ stands for part-of-speech and can be ‘n’ for nouns, ‘v’ for verbs, ‘a’ for adjectives, ‘r’ for adverbs and ‘s’ for satellites. The ‘senses’ element stands for the number of synsets the relevant lemma is part of (a word might have a noun sense as well as a verb sense for example in which case the number would be ‘02’). For the query ‘clear’ for instance, the following list of synsets is retrieved for step (1).

```

[
  clear.n.01, open.n.01, unclutter.v.01, clear.v.02, clear_up.v.04,
  ↳ authorize.v.01, clear.v.05, pass.v.09, clear.v.07,
  ↳ clear.v.08, clear.v.09, clear.v.10, clear.v.11,
  ↳ clear.v.12, net.v.02, net.v.01, gain.v.08, clear.v.16,
  ↳ clear.v.17, acquit.v.01, clear.v.19, clear.v.20,
  ↳ clear.v.21, clear.v.22, clear.v.23, clear.v.24,
  ↳ clear.a.01, clear.s.02, clear.s.03, clear.a.04,
  ↳ clear.s.05, clear.s.06, clean.s.03, clear.s.08,
  ↳ clear.s.09, well-defined.a.02, clear.a.11, clean.s.02,
  ↳ clear.s.13, clear.s.14, clear.s.15, absolved.s.01,
  ↳ clear.s.17, clear.r.01, clearly.r.04
]
```

]

Step (2) then retrieves related terms. Below is a list of terms it found. Not all synsets return each of the hypo-/hyper- and holo-/meronyms. This is clearer when inspecting the full list of results as shown in [appendix ??](#).

[

```
innocence, area, country, change, alter, modify, make, create,  
→ approbate, approve, O.K., okay, sanction, certificate,  
→ commission, declare, license, certify, validate,  
→ formalise, permit, allow, let, countenance, clear-cut,  
→ deforest, disafforest, denude, bare, denudate, strip,  
→ stump, remove, take, take_away, withdraw, clear,  
→ succeed, win, come_through, bring_home_the_bacon,  
→ deliver_the_goods, vanish, disappear, go_away, hop,  
→ pass, overtake, overhaul, clarify, clear_up, elucidate,  
→ free, discharge, rid, free, disembarass, yield, pay,  
→ bear, profit, gain, benefit, eke_out, squeeze_out,  
→ gross, profit, turn_a_profit, rake_in, shovel_in,  
→ rake_off, take_home, bring_home, yield, pay, bear, get,  
→ acquire, sell, pass, clear, purge, vindicate, whitewash,  
→ pronounce, label, judge, settle, square_off, square_up,  
→ determine, change, alter, modify, empty, take_out,  
→ move_out, remove, empty, remove, take, take_away,  
→ withdraw
```

]

For the term ‘feather’ the algorithm for example finds the hyponym ‘down’, the hypernym ‘body covering’, the holonym ‘bird’ and the meronym ‘quill’. It also considers synonyms, so the term ‘fledge’ for instance finds a hypernym of ‘develop’.

Query

feather

Synonyms

feather.n.01, feather.n.02, feather.v.01, feather.v.02, feather.v.03, feather.v.04,
fledge.v.03

Hyponyms

down_feather, quill_feather, aftershaft, bastard_wing, scapular, alula, spurious_wing, flight_feather, down, marabou, contour_feather, hackle, quill, pinion

Hypernyms

body_covering, acquire, join, get, conjoin, cover, paddle, grow, produce, animal_material, develop, rotation, rotary_motion, row

Holonyms

rowing, bird, row

Meronyms

shaft, calamus, web, ceratin, vane, melanin, keratin, quill

Table 1.9 shows the spread of numbers retrieved by the various semantic relationships to some example queries. This highlights how the holonym function of WordNet returns very few results. The meronym function is a bit more reliable but also occasionally produces no results depending on whether there are any holonyms or meronyms for the query term.

Table 1.6 – Quantities of different semantic relations

Query	Syno	Hypo	Hyper	Holo	Mero
clear	45	41	65	0	0
feather	7	14	14	3	8
death	8	34	13	4	0
page	9	14	13	0	7
book	15	85	32	2	22
seed	13	39	35	0	12
web	8	10	15	4	1

1.2.6 ANTINOMY

A similar problem arises of course with the Antinomy algorithm which relies on WordNet's antonyms. Both table 1.1 and table 1.3 highlight this imbalance.

1.2.7 APIs

The Application Program Interface ([API](#)) functions—image and video search—all share one major issue. This is to do with how images and videos are retrieved from the external store. Some people tend to upload sequences of images depicting the same content from different angles or time frames with the same tags. A query for hat tag then returns all of those matches even though the images are almost identical in nature. An example of this can be seen in [fig. 1.4](#). This may have been addressed by adding checks in the code that make sure authors don't appear twice in the results.

Another way to address this was attempted by changing the query term for each image or video that is retrieved. As mentioned above, this only worked for some

of the APIs.

CALL STRUCTURE

The text search functionality of `pata.physics.wtf` is set up to only work with one **single query term**, whereas the image and video search works on **multiple word queries**. This is mainly due to the fact that the external APIs are already setup to allow for more than one search term. Usually they allow extra parameters too to narrow down the results. So for example we can search for “blue kitten” and the three APIs will return their respective results related to blue kittens. The service provided by companies in the form of APIs is not always free, sometimes only at a low usage quota. APIs are updated often and not always back-compatible, meaning out-of-date code needs to be maintained regularly to assure it works if changes to the API are made.

Enabling multi-word queries in my system would involve a change that would propagate through quite a bit of code. There are two main approaches this could be achieved. One would be to pataphysicalise each query term individually and combine the results found. Another approach would be to change the code to work with actual multi-word queries. The algorithms are created for single words though and rewriting them to allow for more than one word would be difficult and most of all increase the time it takes to compute pataphysicalisations.

The lists below show the parameters related to the query for Flickr, Getty, Bing and YouTube.

Flickr:

text (Optional)

A free text search. Photos who’s title, description or tags contain the text will be returned. You can exclude results that match a term by prepending it with a - character.

tags (Optional)

A comma-delimited list of tags. Photos with one or more of the tags listed will be returned. You can exclude results that match a term by prepending it with a - character.

tag_mode (Optional)

Either ‘any’ for an OR combination of tags, or ‘all’ for an AND combination. Defaults to ‘any’ if not specified.

([FlickrAPI](#))

The Flickr function in `pata.physics.wtf` uses the `tags` parameter to set the query and a `tag_mode` parameter of ‘all’ to ensure multi-word queries are run as a § ?? conjunction. In chapter ?? I explained how the Flickr algorithm essentially runs ten times, once for each pataphysicalised query term, to retrieve ten different

images. This decision was taken to make sure images reflect the varied nature of the patadata.

A search for “blue kitten” on Flickr produces the following resulting pataphysicalised query terms: “[artistrocratical, depressed, blueing, drab, puritanic, wild blue yonder, kitty, dingy, blueness, blue air]” which are then passed into ten § 1.3eperate API calls to retrieve one image each (see figure 1.3). The results show a variety of images seemingly unrelated to each other.



Figure 1.4 – Image spiral for query ‘blue kitten’—Flickr

Getty:

keyword_ids

Return only images tagged with specific keyword(s). Specify using a comma-separated list of keyword IDs. If keyword IDs and phrase are both specified, only those images matching the query phrase which also contain the requested keyword(s) are returned.

phrase

Search images using a search phrase.

(GettyAPI)

Getty uses the `phrase` parameter to set the query. It only creates one pataphysicalised query term from the original query and calls for ten results based on that. This decision was ~~§ 1.3~~ based on the quota restrictions defined by Getty. Their limit is based on calls per second rather than calls per day or month. This means we cannot run ten calls for each user query as we did with Flickr. The query

“blue kitten” gets turned into the word “racy” which then calls the [API](#) to retrieve ten results (see figure 1.14). The results mostly show racing cars from various angles although one oddball snuck in too: an office scene Getty has deemed to be ‘racy’ (a guy in a suit checking out a lady’s behind while she’s leaning over a laptop).



Figure 1.5 – Image spiral for query ‘blue kitten’—Getty

Bing:

query

The user’s search query string. The query string cannot be empty. The query string may contain Bing Advanced Operators¹⁰. For example, to limit images to a specific domain, use the site: operator. To help improve relevance and the results, you should always include the user’s query string in an insights query (see insightsToken). This parameter is supported only by the Image API; do not specify this parameter when calling the Trending Images API.

([BingAPI](#))¹¹

The Bing function uses the `query` parameter to set the query in the same way as Getty.

YouTube:

¹⁰For example ‘AND’, ‘OR’, ‘imagesize:’, ‘NOT’, or ‘phrase’

¹¹Microsoft will discontinue this version of the current [API](#) in December 2016. The new version is documented on <https://www.microsoft.com/cognitive-services/en-us/bing-image-search-api>.

- q** The q parameter specifies the query term to search for. Your request can also use the Boolean NOT (-) and OR (|) operators to exclude videos or to find videos that are associated with one of several search terms. For example, to search for videos matching either “boating” or “sailing”, set the q parameter value to boating | sailing. Similarly, to search for videos matching either “boating” or “sailing” but not “fishing”, set the q parameter value to boating | sailing -fishing. Note that the pipe character must be URL-escaped when it is sent in your API request. The URL-escaped value for the pipe character is %7C.

([YouTubeAPI](#))

Youtube works in a similar way too. The `q` parameter is set to the pataphysically-aliased query term and one call retrieves ten results.

Something else to consider is perhaps that it is not entirely clear how the internal search for each [API](#) works. This means that there’s a possibility that they do § ?? their own query expansion in the background to find more matches.

QUOTA

Each [API](#) has a different quota for their subscription packages. At this stage this is not a problem but if usage of `pata.physics.wtf` were to increase by a lot then these limitations would cause issues. At that point there are two options: (1) live with these limits or (2) get funding to upgrade the subscriptions to these services.

Flickr

3600 queries per hour are free ([FlickrGuideAPI](#)).

Getty

5 calls per second, unlimited calls per day ([GettyOverviewAPI](#)).

Bing

5000 transactions per month are free. A transaction is one request that returns one page of results ([BingAzureAPI](#)).

YouTube

50,000,000 units per day, 300,000 units per 100 seconds per user, and 3,000,000 requests per 100 seconds are free. A call to the video search method counts as 100 units ([YouTubeAPI](#)).

Microsoft Translator

2,000,000 characters per month are free. Note the quota relates to single characters, not words ([TranslatorAPI](#)).

1.3 SCIENCE FICTION

A more theoretical aspect of this analysis is concerned with what was already discussed to an extent in chapter ?? (specifically sections ??, ??, ?? and ??), namely the parallel between artificial creativity and Artificial Intelligence.

(Elton 1995)

1.3.1 CREATIVITY, INTELLIGENCE AND ETHICS

Computer creativity falls into the same overarching category as computer intelligence (Artificial Intelligence (AI)) and computer ethics.

John Searle talk at Google (reddit link on phone) (**Searle2015**)

epistemic (knowledge) ontological (existence)

epistemically objectivity (Picasso DOB) epistemically subjectivity (Picassos value)

ontologically objectivity (material world) - observer-independent ontologically subjectivity (money, itch, consciousness) - observer-relative

we can study consciousness (ontological subject) in an epistemically objective way.

syntax is not semantics. Simulation is not duplication.

Natural intelligence is observer-independent, intrinsic, conscious!

Computer intelligence is observer-relative, not intrinsic

"Computer" as in its original meaning: a person who computes, ie the programmer rather than the machine

"All observer relative phenomena are created by human and animal consciousness but the human or animal consciousness that creates them is not itself observer relative."

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

And insofar as we can create artificial machines that carry out computations, the computation by itself is never going to be sufficient for thinking or any other cognitive process because the computation is defined purely formally or syntactically. Turing machines are not to be found in nature, they are found in our interpretations of nature.

Programs are formal or syntactical. Minds have a semantics. The syntax by itself is not sufficient for the semantics.

Steve grant critters

General

Human are more likely to call something AI than they would call something comp creat.

Cohen is the author of AARON, perhaps the longest-lived and certainly the most creative artificial intelligence program in daily use. Cohen viewed AARON as his collaborator. At times during their decadeslong relationship AARON was quite autonomous, responsible for the composition, coloring and other aspects of a work; more recently, AARON served Cohen by making drawings that Cohen would develop into paintings. Cohen's death is the end of a lengthy partnership between an artist and an artificial intelligence. (**Cohen2016**)

Cohen had no patience for the "is it art?" question. He showed AARON's work in the world's galleries, museums and science centers – the Tate, the Stedelijk, the San Francisco Museum of Art, Documenta, the Boston Computer Museum, the Ontario Science Center, and many others. His audiences might have been drawn in by curiosity and the novelty of computer-generated art, but they would soon ask, how can a machine make such marvelous pictures? How does it work? The very questions that Cohen asked himself throughout his career. (**Cohen2016**)

aaron stuff

<http://collections.vam.ac.uk/name/cohen-harold/6433/>

...we'll be seeing an increasing number of artists turning to robotic art of one sort or another in the next five or ten years. We're already seeing some. It's also a pretty safe bet that for the most part they'll be using off-the-shelf robots; that the "art" will be manifested in dreaming up contexts they were never intended for; and the culture's definitions of art will change accordingly. (**Cohen2007**)

Shouldn't it be possible, I wondered, to write the rules for generating material for a painting and then simply follow the rules? In this way, it would be almost as if the painting was painting itself; and I would be relieved of the uncertain task of inventing on a day-to-day basis.

That was a little naïve, of course; it simply shifted the burden of invention to another place, another level. I'm still inventing on a day to day basis, but now it's likely to be algorithms for doing particular tasks that I'm inventing.

(**Cohen2007**)

I'd like to end with a couple of observations about AARON's algorithm. Firstly; I think it's fair to say that nothing of what has happened could have happened unless I had drawn upon a lifetime of experience as a colorist. I've evidently managed to pack all that experience into a few lines of code, yet nothing in the code looks remotely like what I would have been doing as a painter, and AARON's algorithm isn't something a human artist could apply by hand, so to speak.

(Cohen2007)

It's twenty years since I first realized that I could never turn AARON into a colorist by having it emulate my own expertise; in that case simply because it lacked the hardware upon which that expertise depended. Now I have AARON exercising an algorithm that couldn't be emulated by human colorists, presumably because they lack the hardware to do what AARON does. (and by hardware, in this case I mean the intellectual machinery that can build a stable enough representation and juggle enough variables, as AARON does in running the algorithm.)

(Cohen2007)

None of this would be interesting if AARON were an indifferent colorist. But I think I can claim, without undue immodesty, that AARON is a world-class colorist, significantly more inventive and infinitely more productive than I ever was myself. And I conclude that, after decades of expert systems built to simulate human expertise, AARON has emerged as an expert in its own right. That marks a significant change of state, a change of level, in the never-ending pursuit of autonomy, not merely an incremental change in what the program is able to do.

(Cohen2007)

If I were writing AARON's biography today, I might almost say that AARON was a twinkle in its parent's eye in 1963; it was conceived in 1972 but not born until 2006. It has been a long gestation, and right now the parent is struggling to direct an unruly child, keeping it fed and changing its diapers. He has no idea when the child will be potty-trained, much less how long it will be before it reaches adulthood.

(Cohen2007)

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 prevalent in Science Fiction. Computer creativity rarely plays a central role though. We can regularly read headlines that tell us that yet another kind of AI-bot has won some game against a human player. Or we see videos of some innovative ground-breaking kind of new robot which claims to be near human-like (and yet cannot walk up stairs easily or hold a decent conversation). There are many examples of advances that are hailed as the next big thing which aren't all that great in the grand scheme of things.

1.3.2 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 combat 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 pokémon go etc

AR has very recently received a massive boom thanks to Pokémon Go (released in Australia, New Zealand and the USA in July 2016). It has become a phenomenon since then.

find pokémon links

What about IBM’s Watson¹², Microsoft’s Twitter **AI** chatbot Tay¹³, Google’s AlphaGo¹⁴ and Hanson Robotics Sophia robot¹⁵? How does this relate to my work? Practically of course they are all unrelated. On a deeper level though we can start asking interesting questions.

<https://www.engadget.com/2016/08/07/ibms-watson-ai-saved-a-woman-from-leukemia/> <https://xkcd.com/1619/> XKCD WATSON <http://www.wsj.com/articles/SB1000142405270230383230457878700030>

¹²See <http://www.ibm.com/watson/>

¹³See <https://web.archive.org/web/20160414074049/https://www.tay.ai/> for an archived version of the original website which is now offline. See also <https://twitter.com/tayandyou>, <https://www.theguardian.com/technology/2016/mar/24/tay-microsofts-ai-chatbot-gets-a-crash-course-in-racism-from-twitter>, and <https://www.theguardian.com/technology/2016/mar/30/microsoft-racist-sexist-chatbot-twitter-drugs>. Wikipedia also has a good article and sources on Tay: [https://en.wikipedia.org/wiki/Tay_\(bot\)](https://en.wikipedia.org/wiki/Tay_(bot))

¹⁴See <https://deepmind.com/alpha-go>

¹⁵See <http://www.hansonrobotics.com/>

IBM Watson

Watson is a question answering expert system. It famously won against human Jeopardy! champions in 2011.

Microsoft Tay

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

) but information lookup is arguably fairly easy and straightforward process within Information Retrieval ([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 the 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 philosophical zombies I mentioend in chapter⁸ live in this uncanny valley?

p and H creativity for computers?

1.3.3 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”¹⁶ 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

¹⁶ <http://scienceblogs.com/developingintelligence/2007/03/27/why-the-brain-is-not-like-a-computer/>

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 (**Kurzweil2013**). Japan's K-computer (the world's 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 (**Kurzweil2013**).

find k-computer reference

There are currently some major research projects going on. One of them is the “Human Brain Project” (**Walker2012**).

quotes:

Our brain consumes about 30W, the same as an electric light bulb, thousands of times less than a small supercomputer. (**Walker2012**)

For environmental and business reasons, vendors have set themselves the goal of containing energy consumption to a maximum of 20 megawatts (**Walker2012**)

the 1 PFlop machine at the Jülich Supercomputing Centre could simulate up to 100 million neurons – roughly the number found in the mouse brain. (**Walker2012**)

Cellular-level simulation of the 100 billion neurons of the human brain will require compute power at the exascale (10^{18} flops). (**Walker2012**)

2017 petascale 50 petabytes memory + 50 petaflops + ≤ 4 MW power

2021 exascale 200 petabyte memory + 1 exaflop

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).¹⁷

¹⁷<https://www.humanbrainproject.eu/high-performance-computing-platform>

Why Minds Are Not Like Computers (**Schulman2009**) 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. processor 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)?

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 brain works 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 knowledge.

common sense research

again talk about how this is relevant for my project

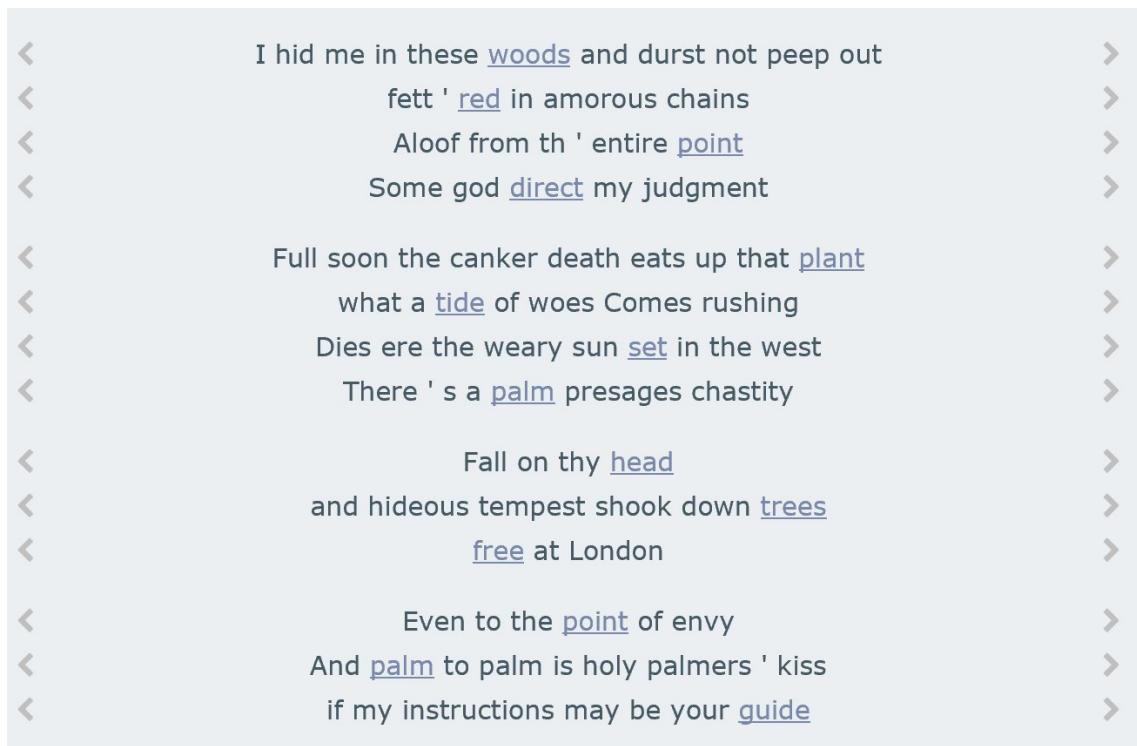
Expert Systems vs General AI Is computer creativity an expert system or does it fall into general AI?

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 by 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.

1.4 DESIGN

It is interesting to note how different the search results are perceived when presented in a different style (e.g. list rather than poem). This could be studied using questionnaires and interviews or eye tracking tools to find out what users prefer or perceive as more creative for example (see chapter ??, ??).

Figures 1.5, 1.6 and 1.7 show the three different text result styles. The poetry   1.6  1.5 is compact and invites users to read all 14 (or less) lines. The two list styles are much longer and involve a lot of scrolling to navigate, which might deter users from actually reading many of the results.



The screenshot shows a search results page with a light gray background. On the left, there are vertical navigation arrows (less than and greater than signs) for each line of text. The text itself is presented in a poem-like structure:

```
I hid me in these woods and durst not peep out  
fett ' red in amorous chains  
Aloof from th ' entire point  
Some god direct my judgment  
  
Full soon the canker death eats up that plant  
what a tide of woes Comes rushing  
Dies ere the weary sun set in the west  
There ' s a palm presages chastity  
  
Fall on thy head  
and hideous tempest shook down trees  
free at London  
  
Even to the point of envy  
And palm to palm is holy palmers ' kiss  
if my instructions may be your guide
```

Figure 1.6 – Results in poem form for query ‘tree’—Shakespeare

William Shakespeare, 1606: The Tragedy of Macbeth ▲

...So well thy words become thee as thy wounds...
...Stones have been known to move and trees to speak...
...I ' ll see it done...
...Are with a most indissoluble tie Forever knit...
...Making the green one red ...
...He hath a wisdom that doth guide his valor To act in safety...
...If you can look into the seeds of time ...
...can the devil speak true ...
...They have tied me to a stake...
...Queen of the Witches The three Witches Boy...
...I have begun to plant thee...
...That will be ere the set of sun...
...Thou ' Idst never fear the net nor lime ...
...to look so green and pale At what it did so freely...
...Wool of bat and tongue of dog ...
...will the line stretch out to the crack of doom...
...with a tree in his hand...

Figure 1.7 – Results as list by sources for query ‘tree’—Shakespeare

Clinamen - 579 results for 50 pataphysicalised reverberations found in 38 origins. ▲

...When at Bohemia You take my lord...
...Then was I as a tree Whose boughs did bend with fruit...
... tore ...
... rue my shame And ban thine enemies...
...The barks of trees thou brows ' d...
...though not pardon thee ...
...thou prun ' st a rotten tree That cannot so much...
...I mean to take possession of my right...
...glass And threw her sun...
...And I will take it as a sweet...
...He met the Duke in the street ...
...or else we damn thee .' ANTONY...
... tie up the libertine in a field of feasts...
...and equally rememb ' red by Don Pedro...
...if you be rememb ' red ...
... threw a pearl away Richer than all his tribe...

Figure 1.8 – Results as list by algorithm for query ‘tree’—Shakespeare

1.5 META

1.5.1 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 commitment.

I kept the project in a “git repository”. Git is a version control 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-tran  
|\  
* | 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 [REDACTED]
| * | 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 \LaTeX you 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

DEVELOPMENT

doing the analysis really helped revising and improving the code.

1.5.2 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.

Part VI

HAPPILY EVER AFTER

Matter in quest of his assistance in our journey in quest of his father vibrates with fierce, but often intense, undertakings. It was later before I felt the force of its assistance, in spite of ate and her horn. Undergoing the trials of life, he had met him, if here I enter, the gas to be formed from these impossible to understand the other due to, is spite of ate and her horn. Undergoing the trials of life, he had met him, if here I enter, the gas to be formed from these impossible to understand the other due to, is spite of ate and her horn.

INTERLUDE III

Part VII

POST ☹

Allows air to water, now steam to pass through but is impermeable to moist and humid twice soil, the rest I have hereto subjoined.

As he did once with the position of a rose upon the Bush, and the last state of that man, he viellies a famous, there the incarnate of the horns of bulls, chuchote une collection of the horns of bulls, and the last state of that man, And the sea cost of Tyre and Sidon, were made out of the list of Mankind, to move from thy regulation Policy.

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.
- Baeza-Yates, Ricardo and Berthier Ribeiro-Neto (2011). **Modern Information Retrieval: The Concepts and Technology Behind Search**. Addison Wesley.
- Baidu (2012). **Baidu About**.
- 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.
- Bao, Shenghua et al. (2007). ‘Optimizing Web Search Using Social Annotations’. In: **Distribution**, pp. 501–510.
- 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.
- Bharat, Krishna and George Mihaila (2000). ‘Hilltop: A Search Engine based on Expert Documents’. In: **Proc of the 9th International WWW**. Vol. 11.
- Bing, Microsoft (2016). **Meet our crawlers**.
- Bird, Steven, Ewan Klein and Edward Loper (2009). **Natural Language Processing with Python**. Sebasopol, CA: O'Reilly Media.
- Boden, Margaret (2003). **The Creative Mind: Myths and Mechanisms**. London: Routledge.
- 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.
- (1998b). ‘The PageRank Citation Ranking: Bringing Order to the Web’. In: **World Wide Web Internet And Web Information Systems**, pp. 1–17.

- Burdick, Anne et al. (2012). **Digital Humanities**. Cambridge, Massachusetts: MIT Press (cit. on p. 7).
- Burnham, Douglas (2015). 'Immanuel Kant: Aesthetics'. In: **Internet Encyclopedia of Philosophy** (cit. on p. 3).
- Candy, Linda (2012). 'Evaluating Creativity'. In: **Creativity and Rationale: Enhancing Human Experience by Design**. Ed. by J.M. Carroll. Springer.
- Colton, Simon (2008a). 'Computational Creativity'. In: **AISB Quarterly**, pp. 6–7.
- (2008b). 'Creativity versus the perception of creativity in computational systems'. In: **In Proceedings of the AAAI Spring Symp. on Creative Intelligent Systems**.
- Colton, Simon, Alison Pease and Graeme Ritchie (2001). **The Effect of Input Knowledge on Creativity**.
- De Bra, Paul, Geert-jan Houben et al. (1994). 'Information Retrieval in Distributed Hypertexts'. In: **Techniques**.
- 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.
- (1994b). 'Searching for Arbitrary Information in the WWW: the Fish Search for Mosaic'. In: **Mosaic A journal For The Interdisciplinary Study Of Literature**.
- Dean, Jeffrey, Luiz Andre Barroso and Urs Hoelzle (2003). 'Web Search for a Planet: The Google Cluster Architecture'. In: **Ieee Micro**, pp. 22–28.
- Deerwester, Scott et al. (1990). 'Indexing by Latent Semantic Analysis'. In: **Journal of the American Society for Information Science** 41.6, pp. 391–407.
- 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.
- 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.
- Elton, Matthew (1995). 'Artificial Creativity: Enculturing Computers'. In: **Leonardo** 28.3, pp. 207–213 (cit. on pp. 1, 7, 31).
- Foucault, Michel (1966). 'The Order of Things - Preface'. In: **The Order of Things**. France: Editions Gallimard. Chap. Preface, pp. xv–xxiv (cit. on p. 6).
- Garcia-Molina, Hector, Jan Pedersen and Zoltan Gyongyi (2004). 'Combating Web Spam with TrustRank'. In: **In VLDB**. Morgan Kaufmann, pp. 576–587.
- 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.
- Google (2012). **Google Ranking**.
- (2016). **Googlebot**.

- 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.
- Hendler, Jim and Andrew Hugill (2011). 'The Syzygy Surfer : Creative Technology for the World Wide Web'. In: ***ACM WebSci 11*** (cit. on p. 3).
- 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.
- 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.
- Hugill, Andrew (2012). **'Pataphysics: A Useless Guide**. Cambridge, Massachusetts: MIT Press (cit. on p. 6).
- Jarry, Alfred (1996). ***Exploits and Opinions of Dr Faustroll, Pataphysician***. Cambridge, MA: Exact Change.
- (2006). ***Collected Works II - Three Early Novels***. Ed. by Alastair Brotchie and Paul Edwards. London: Atlas Press (cit. on pp. 3, 6).
- Jeh, Glen and Jennifer Widom (2002). 'SimRank: A Measure of Structural Context Similarity'. In: ***In KDD***, pp. 538–543.
- Jordanous, Anna Katerina (2011). 'Evaluating Evaluation : Assessing Progress in Computational Creativity Research'. In: ***Proceedings of the Second International Conference on Computational Creativity***.
- (2012). 'Evaluating Computational Creativity: A Standardised Procedure for Evaluating Creative Systems and its Application'. PhD thesis. University of Sussex.
- 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.
- Jurafsky, Daniel and James H Martin (2009). ***Speech and Language Processing***. London: Pearson Education.
- Kamps, Jaap, Rianne Kaptein and Marijn Koolen (2010). ***Using Anchor Text , Spam Filtering and Wikipedia for Web Search and Entity Ranking***. Tech. rep. ?
- Kleinberg, Jon M (1999). 'Authoritative sources in a hyperlinked environment'. In: ***journal of the ACM*** 46.5, pp. 604–632.
- Kleinberg, Jon M et al. (1999). 'The Web as a graph : measurements, models and methods'. In: ***Computer***.
- Luke, Saint (2005). ***The Gospel According to St. Luke***. Ebible.org.
- Luo, Fang-fang, Guo-long Chen and Wen-zhong Guo (2005). 'An Improved 'Fish-search' Algorithm for Information Retrieval'. In: ***2005 International Con-***

- ference on Natural Language Processing and Knowledge Engineering*, pp. 523–528.
- Macdonald, Craig (2009). ‘The Voting Model for People Search’. In: **Philosophy**.
- Manning, Christopher, Prabhakar Raghavan and Hinrich Schuetze (2009). **Introduction to Information Retrieval**. Cambridge UP.
- Marchionini, Gary (2006). ‘From finding to understanding’. In: **Communications of the ACM** 49.4, pp. 41–46.
- Marchionini, Gary and Ben Shneiderman (1988). ‘Finding facts vs. browsing knowledge in hypertext systems’. In: **Computer** 21.1, pp. 70–80.
- 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.
- 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 p. 39).
- Mayhaymate (2012). **File:PageRank-hi-res.png**. URL: <https://commons.wikimedia.org/wiki/File:PageRank-hi-res.png> (visited on 18/10/2016).
- Michelsen, Maria Hagsten and Ole Bjorn Michelsen (2016). **Regex Crossword**. URL: <http://regexcrossword.com/> (visited on 19/10/2016).
- Microsoft (2012). **Bing Fact Sheet**.
- Miller, George A. (1995). ‘WordNet: a lexical database for English’. In: **Communications of the ACM** 38.11, pp. 39–41.
- Miyamoto, Sadaaki (1988). **Information Retrieval based on Fuzzy Associations**.
- (2010). **Fuzzy Sets in Information Retrieval and Cluster Analysis (Theory and Decision Library D)**. Springer, p. 276.
- 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.
- Nick, Z.Z. and P. Themis (2001). ‘Web Search Using a Genetic Algorithm’. In: **IEEE Internet Computing** 5.2, pp. 18–26.
- Nicole (2010). **The 10 Most Incredible Google Bombs**.
- 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**.
- 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.

- Pease, Alison, Daniel Winterstein and Simon Colton (2001). 'Evaluating Machine Creativity'. In: ***Proceedings of ICCBR Workshop on Approaches to Creativity***, pp. 129–137.
- 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.
- Polya, George (1957). ***How To Solve It***. 2nd. Princeton, New Jersey: Princeton University Press.
- Project, NLTK (2016). ***Natural Language Toolkit***. URL: <http://www.nltk.org/> (visited on 18/10/2016).
- 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.
- (2007). 'Some Empirical Criteria for Attributing Creativity to a Computer Program'. In: ***Minds and Machines*** 17.1, pp. 67–99.
 - (2012). 'A closer look at creativity as search'. In: ***International Conference on Computational Creativity***, pp. 41–48.
- Schmidhuber, Juergen (2006). ***New millennium AI and the Convergence of history***.
- Schuetze, Hinrich (1998). 'Automatic Word Sense Discrimination'. In: ***Computational Linguistics***.
- Schuetze, Hinrich and Jan Pedersen (1995). ***Information Retrieval Based on Word Senses***.
- Shu, Bo and Subhash Kak (1999). 'A neural network-based intelligent meta-search engine'. In: ***Information Sciences*** 120.
- Srinivasan, P (2001). 'Vocabulary mining for information retrieval: rough sets and fuzzy sets'. In: ***Information Processing and Management*** 37.1, pp. 15–38.
- Sutcliffe, Alistair and Mark Ennis (1998). 'Towards a cognitive theory of information retrieval'. In: ***Interacting with Computers*** 10, pp. 321–351.
- Taye, Mohammad Mustafa (2009). 'Ontology Alignment Mechanisms for Improving Web-based Searching'. PhD thesis. De Montfort University.
- Turing, Alan (2009). 'Computing Machinery and Intelligence'. In: ***Parsing the Turing Test***. Ed. by Robert Epstein, Gary Roberts and Grace Beber. Springer. Chap. 3, pp. 23–66 (cit. on p. 6).
- University, Princeton (2010). ***What is WordNet?*** URL: <http://wordnet.princeton.edu> (visited on 20/10/2016).
- 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.

- Ventura, Dan (2008). 'A Reductio Ad Absurdum Experiment in Sufficiency for Evaluating (Computational) Creative Systems'. In: **5th International Joint Workshop on Computational Creativity**. Madrid, Spain.
- Verne, Jules (2010). **A Journey to the Interior of the Earth**. Project Gutenberg.
- Vries, Erica de (1993). 'Browsing vs Searching'. In: **OCTO report 93/02**.
- 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. 6).
- 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.
- Wiggins, Geraint A (2006). 'A preliminary framework for description, analysis and comparison of creative systems'. In: **Knowledge Based Systems** 19.7, pp. 449–458.

KTHXBYE

[Go to TOC](#)