

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!!! 9
- relate all of these things back to my topic of AMC 9
- discuss fig 6.2 (in relation to DH methodologies) 9
- expand 6.1 (abusing stuff, creating own rules, oulipo) 9
- explain why not 10
- remove yossarian criticism 11
- remember to replace some of the chapter poems with shakespeare . . . 13
- check analindex refs and index refs 19
- semantic hierarchy visualised? 24
- i tried modeeling the three alignments but how? 24
- 24
- table showing hypo, hyper and holonyms numbers side by side for a query 25
- find links for motion sickness 30
- find links for epilepsy 30
- find links for oculus rift and pokemon go etc 30
- find pokemon links 30

| | | |
|---|-------|----|
| <input type="checkbox"/> check | | 30 |
| <input type="checkbox"/> find example | | 31 |
| <input type="checkbox"/> check | | 31 |
| <input type="checkbox"/> check | | 31 |
| <input type="checkbox"/> whats the point im making? how does this relate to my work? | | 31 |
| <input type="checkbox"/> p and H creativity for computers? | | 31 |
| <input type="checkbox"/> find k-computer reference | | 32 |
| <input type="checkbox"/> find references | | 34 |
| <input type="checkbox"/> neural networks and other models based on the brain | | 34 |
| <input type="checkbox"/> common sense research | | 34 |
| <input type="checkbox"/> again talk about how this is relevant for my project | | 34 |
| <input type="checkbox"/> add file for appendix with full git history | | 34 |
| <input type="checkbox"/> add calendar screenshot of github contributions | | 35 |
| <input type="checkbox"/> links to git and github | | 35 |
| <input type="checkbox"/> say more, check keywords, potentially generate new poems | | 37 |
| <input type="checkbox"/> software refactoring | | 41 |
| <input type="checkbox"/> write these out all in one list and then group them as fit | | 41 |
| <input type="checkbox"/> check | | 44 |
| <input type="checkbox"/> add ipa data or whatever is best for the rhyming stuff | | 44 |
| <input type="checkbox"/> work the maths out here for this example of MOE | | 45 |
| <input type="checkbox"/> look into rhyming tags in nlp | | 46 |
| <input type="checkbox"/> https://wordnet.princeton.edu/wordnet/man/wngloss.7WN.html for glossary | | 46 |
| <input type="checkbox"/> fix all chapter XYZ mentions | | 46 |
| <input type="checkbox"/> group these into better sub groups and make them proper sections rather than paragraphs | | 46 |

| | | |
|--|-------|----|
| <input type="checkbox"/> pageinate results for speed? | | 46 |
| <input type="checkbox"/> summarise thesis, contributions etc. conclude by comparing against introduction | | 50 |
| <input type="checkbox"/> refer back to these in conclusion | | 50 |
| <input type="checkbox"/> check if i need to submit a CD? | | 50 |
| <input type="checkbox"/> add chapter references | | 51 |
| <input type="checkbox"/> these are closed questions, not research questions | | 51 |
| <input type="checkbox"/> write more | | 52 |

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: 6th August 2016
Wordcount: 11568

[Go to TOC](#)

PRE

And dire au, in dire deux hommes passer
car as, deux autres, shod and the other bare.
The hamlets bare, The hamlets over prime
sugeries, will not you be content to pay a puncheon of Breton wine, the
cri en courant dans la rue, having one foot
bare White, une salle pleine le port de
courant, ne put qu'artificuler au, in dire pif paf pan, ne put
and fro in artificuler au, in dire pif paf pan, ne put

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. It is the use of digital tools (which may not be creative themselves) and the way they are used forms the creative process or product. Creativity in humans needs to be interpreted differently to machines. Humans and machines differ in many ways, we have different ‘brains/memory’, ‘thinking processes/software’ and ‘bodies/hardware’. Too often creative output by machines is judged as we would a humans. Computers which are truly artificially intelligent might be capable of true artificial creativity. Until then they are (philosophical) zombie robots: machines that behave like humans but aren’t conscious. The only alternative is to see any computer creativity as a direct or indirect expression of human creativity using digital means and evaluate it as such. [AMC](#) is neither machine creativity nor human creativity—it is both. By acknowledging the undeniable link between computer creativity and its human influence (the machine is just a tool for the human) we enter a new realm of thought. How is [AMC](#) defined and evaluated? This thesis address this issue. First a practical demonstration of [AMC](#) is presented ([pata.physics.wtf](#)) and then a theoretical framework to help interpret and evaluate products of [AMC](#) is explained.

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

¹“Too long; didn’t read”

PUBLICATIONS

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

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

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

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



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

CONTENTS

| | |
|------------------|----------|
| 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 | 8 |
| 1.1 | Influences | 10 |
| 1.2 | Pataphysicalisation | 11 |
| 1.3 | Design | 27 |
| 1.4 | Science Fiction | 28 |
| 1.5 | Meta | 34 |
| 2 | Aspirations | 40 |
| 2.1 | Technical | 41 |
| 2.2 | Creative NLP | 45 |
| 2.3 | Theoretical | 46 |
| HAPPILY EVER AFTER <hr/> | | |
| 3 | Observations | 49 |
| 3.1 | Outroduction | 50 |
| 3.2 | Issues | 51 |
| 3.3 | Answers | 51 |
| 3.4 | Contributions | 52 |
| 3.5 | And Finally | 52 |
| POSTFACE <hr/> | | |
| References | | 55 |

FIGURES

| | | |
|-----|---|----|
| 1.1 | Faustroll vs. Shakespeare poetry | 13 |
| 1.2 | Image spiral ‘blue kitten’—Flickr | 26 |
| 1.3 | Image spiral ‘blue kitten’—Getty | 27 |
| 1.4 | Results as poem | 28 |
| 1.5 | Results as list by sources | 29 |
| 1.6 | Results as list by algorithm | 29 |

TABLES

| | | |
|-----|---|----|
| 1.1 | Comparison of patalgorithms | 12 |
| 1.2 | Faustroll vs Shakespeare in numbers | 14 |
| 1.3 | Numbers per algorithm | 16 |
| 1.4 | Count and time of results | 19 |
| 1.5 | Changing base in Clinamen - 'fania' | 23 |
| 1.6 | Changing base in Clinamen - 'clear' | 23 |
| 1.7 | Changing base in Clinamen - 'moss' | 23 |
| 1.8 | Changing number of errors in Clinamen | 24 |

CODE

ACRONYMS

| | |
|---------------|--|
| AMC | Algorithmic Meta-Creativity |
| IR | Information Retrieval |
| NLP | Natural Language Processing |
| AI | Artificial Intelligence |
| API | Application Program Interface |
| RDF | Resource Description Framework |
| HTML | Hypertext Markup Language |
| OULIPO | Ouvroir de Littérature Potentielle FIX THIS SHIT |
| MLE | Maximum Likelihood Estimation |
| POS | Parts-of-Speech |
| 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 Spadefoots made their body. Who longs to plunge two fellow creatures into the water, who bends his head at the four corners. She fell on to a hillock of sand, aux montagnes d'origine, . . . Ludgate hill, till the Spadefoots made their body.

Part II

TOOLS OF THE TRADE

Made up your habill'd minds to brave me, ce train re
comme'nait quand' que'z week's silenty, a diff'rence
t'ree and' que'z week's silenty, a diff'rence
sown with the train is due, mad' voyage, against the tide, aucun employe
long' gourmerre ne l'ignorant plus, sell that
sir Excellency stooped to take it up, or in the portugall
language, 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 CORE: ΤΣΕΧΝΟ- ΛΟΓΙΚ

Do not cry and bleed to will, cloth to be sure, your blows it cringe
and definitley. A royal robe none can miss, how cold she must be, sa belle robe rose en desordre.
Come like un fillet sur le centre de la France et qui s'appela, mes bagages et regler ma note, if
pure bijoussin. Ils peuvent aller a tout le moins ay sens, there is none of the jaded
with graceful pride, death only is the lot which none can miss, how cold she must be, sa belle robe rose en desordre.

Part IV

THE CORE: TECHNO- PRACTICE

INTERLUDE II

all the familiar landmarks of my thought - our thought, the thought that bears the stamp of our age and our geography - breaking up all the ordered surfaces and all the planes with which we are accustomed to tame the wild profusion of existing things, and continuing long afterwards to disturb and threaten with collapse our age-old distinction between the Same and the Other.

(Foucault 1966)—taking about Borges

Only those who attempt the absurd achieve the impossible.

(attributed to M.C. Escher)

A great truth is a truth whose opposite is also a great truth. Thomas Mann
(as cited in Wickson, Carew and Russell 2006)

Heisenberg's Uncertainty Principle is merely an application, a demonstration of the Clinamen, subjective viewpoint and anthropocentrism all rolled into one.

(Jarry 2006)

Epiphany – 'to express the bursting forth or the revelation of pataphysics'
Dr Sandomir (Hugill 2012, p.174)

Part V

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

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

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

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

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

| | | |
|-------|-------------------------------|----|
| 1.1 | Influences | 10 |
| 1.2 | Pataphysicalisation | 11 |
| 1.2.1 | Numbers | 14 |
| 1.2.2 | Sentences | 17 |
| 1.2.3 | Index | 19 |
| 1.2.4 | Clinamen | 21 |
| 1.2.5 | Syzygy | 24 |
| 1.2.6 | Images | 25 |
| 1.3 | Design | 27 |
| 1.4 | Science Fiction | 28 |
| 1.4.1 | AI | 28 |
| 1.4.2 | Brains | 32 |
| 1.5 | Meta | 34 |
| 1.5.1 | Management | 34 |
| 1.5.2 | Thesis | 37 |



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

discuss fig 6.2 (in relation to DH methodologies)

expand 6.1 (abusing stuff, creating own rules, oulipo)



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 reason 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. It should not be forgotten that 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 even though this was originally suggested. Also, the concept of patadata was never really conceptualised properly.

explain why not

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

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

Library of Babel

§ ?? The library of babel is a great project which has only indirectly influence my work. The pataphysical elements in it are obvious even though perhaps unconscious. The seriousness with which the library is presented, the pseudo-scientific approach, the vagueness of what's actually behind it. Is it random? Or is it indeed the most gigantic digital library of any book every written or even to be written? The sheer perceived scale of the library was part motivation for calculating the numbers of the generatable poems.

Oulipo

§ ?? Given that the Ouvroir de Littérature Potentielle FIX THIS SHIT ([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.

remove yossarian criticism

1.2 PATAPHYSICALISATION

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

1. User enters single query term,
2. system transforms query term into list of pataphysicalised terms,

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

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.

| 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, domi- cile, taste, see, be | recorded, dead |

Table 1.1: Comparison of patalgorithms showing a selection of results for each.

- 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 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.



| | |
|--|---|
| 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 shoulds 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 cor-

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

pus 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.

| | Query | Corpus | Results | Reverbs | Origins | Poems |
|--------|--------------|---------------|----------------|----------------|----------------------|--------------|
| flower | Faustroll | 89 | 24 | 18 | 7.8×10^{10} | |
| | Shakespeare | 157 | 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 | |

Table 1.2: Faustroll versus Shakespeare in numbers

■ 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

self...

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 generated 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 **§ 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, rever-

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

berations 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').

| | | Clinamen | | | Syzygy | | | Antinomy | | | |
|--------------------|----------|----------|---------|---------|---------|---------|---------|----------|---------|---------|------------|
| | | Results | Reverbs | Origins | Results | Reverbs | Origins | Results | Reverbs | Origins | Total |
| | | Query | | | | | | | | | |
| Faustroll | clear | 158 | 20 | 13 | 368 | 90 | 23 | 16 | 8 | 8 | 542—79—23 |
| | shine | 228 | 29 | 19 | 154 | 61 | 16 | 0 | 0 | 0 | 382—61—20 |
| | disorder | 0 | 0 | 0 | 159 | 127 | 23 | 10 | 2 | 10 | 169—40—23 |
| | stuck | 59 | 14 | 13 | 181 | 43 | 22 | 11 | 3 | 9 | 251—47—22 |
| Shakespeare | clear | 435 | 20 | 38 | 997 | 90 | 38 | 13 | 8 | 12 | 1445—72—38 |
| | shine | 575 | 29 | 38 | 333 | 61 | 38 | 0 | 0 | 0 | 908—53—38 |
| | disorder | 0 | 0 | 0 | 326 | 127 | 38 | 29 | 2 | 29 | 355—26—38 |
| | stuck | 152 | 14 | 37 | 479 | 43 | 38 | 34 | 3 | 34 | 665—41—38 |

Table 1.3: Results-Reverberations-Origin numbers per algorithm

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 § 2 was left for future work mentioned in chapter 2. 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 cor-

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.

pus mostly produces results from each of its 38 texts. The Faustroll corpus varies a lot more. This may be due to the different languages and varying word counts of the files in the corpus.

Faustroll

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

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

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).

```
{
    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

⁶Francois Rabelais: Gargantua and Pantagruel

- 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%.

| Query | Corpus | Return all | | First only | | | Percent |
|----------|-------------|------------|-------|------------|-------|---------|---------|
| | | Count | Time | Count | Time | Percent | |
| clear | Faustroll | 8751 | 59s | 542 | 1.83s | 6.19% | |
| | Shakespeare | 11304 | 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 | 19222 | 120s | 1064 | 2.81s | 5.54% | |
| | Shakespeare | 13635 | 90s | 2125 | 4.63s | 15.58% | |
| record | Faustroll | 5199 | 38s | 275 | 1.72s | 5.29% | |
| | Shakespeare | 7631 | 49.2s | 794 | 2.09s | 10.40% | |

Table 1.4: Count, time and percentage of results retrieved

Google recommends having a “response time under 200ms”⁷. The numbers in ■ 1.4 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

check analindex refs and index refs

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

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

§ ?? 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 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 but also French and German terms since the Faustroll corpus is multi-lingual. The combined index is therefore 35.2MB large.



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

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

each word entry. This would of course be beneficial for the implementation of a § 2 rhyming scheme for the poetry generation. See also chapter 2.

```
{  
    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 this thesis, 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”.

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. (Google 2016)

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 pataphysicalised words. It also uses the Faustroll text. The way this works is as

follows. If the query term is a spelling error of size 1 or 2 of a term in the vocabulary within the faustroll text then it is included in the list of resulting terms. The logic behind this is due to the damerau levenshtein algorithm needing two words to compare with each other. It also ensures we get real words as results and not some random gibberish.

Currently the algorithm is set to accept terms that have a difference of 1 or 2 to the original query. We can lower this to 1 to allow fewer results or increase it to make it broader. I felt 1 or 2 was a good compromise. Only allowing 1 error would mean terms are too similar. Allowing 3 might mean they are drastically different.

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)

Tables 1.5, 1.6 and 1.7 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 books' 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).

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

Table 1.5: Changing base in Clinamen - query 'fania'

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

Table 1.6: Changing base in Clinamen - query 'clear'

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

Table 1.7: Changing base in Clinamen - query 'moss'

§ 1.8 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.

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

Table 1.8: Changing number of errors in Clinamen

1.2.5 SYZYGY

semantic hierarchy visualised?

i tried modeeling the three alignments but how?

The syzygy function goes through the following process.

1. A set of synonyms (a list of “synsets”) is generated.
2. For each of these, hyponyms, hypernyms and holonyms 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 holonyms. 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
]

```

| |
|---|
| table showing hypo, hyper and holonyms numbers side by side for a query |
|---|

1.2.6 IMAGES

The image search can produce quite interesting results as well. A search for “blue kitten” on Flickr produces the following resulting pataphysicalised query terms: “[artistrocratrical, depressed, blueing, drab, puritanic, wild blue yonder, kitty, dingy, blueness, blue air]” which are then passed into ten separate Application

Program Interface §[API](#) calls to retrieve one image each (see figure 1.2). The results show a variety of images seemingly unrelated to each other.



Figure 1.2: Image spiral for query ‘blue kitten’—Flickr

For Getty the image search works slightly differently due to its [API](#) restrictions. The query “blue kitten” gets turned into the word “racy” which then calls the § 1.3 [API](#) to retrieve ten results (see figure 1.3). 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).



The [API](#) functions 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 § 1.3 be seen in figure 1.3. 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



Figure 1.3: Image spiral for query ‘blue kitten’—Getty

of the APIs.



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.

1.3 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 focus groups using questionnaires and interviews or eye tracking tools to find out what users prefer or perceive as more creative for example.

Figures 1.4, 1.5 and 1.6 show the three different text result styles. The poetry  1.5 & 1.6 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.

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

Figure 1.4: Results in poem form for query ‘tree’—Shakespeare

1.4 SCIENCE FICTION

Where does this project stand in the wider world and the progress of computing, Artificial Intelligence ([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 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 perfectly human-like (and yet cannot walk up stairs). There are so many examples of advances that are hailed as the next big thing which aren’t all that great.

1.4.1 AI

This is also evident in games, for example Virtual Reality ([VR](#)) and Augmented Reality ([AR](#)). The Oculus Rift and similar systems are advertised so much you might believe they are actually about to hit mainstream and every kid will own a [VR](#) console and headset. Yet they are still way too expensive to be mainstream and motion sickness is also still an issue (and probably always will). These industries are so “hip” any publication is seen as the new cool thing without taking into account the history and work that has been done previously in per-

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 1.5: 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 1.6: Results as list by algorithm for query 'tree'—Shakespeare

haps slightly different disciplines. This is the case for example with a recent article on VR sickness and how to compat it. This is a well known problem already—motion sickness already exists in normal games. Similar to epilepsy problems.

find links for motion sickness

find links for epilepsy

find links for oculus rift and 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 Watson¹⁰, Microsoft's Twitter AI Chatbot¹¹, Google's AlphaGo¹² and Hanson Robotics Sophia robot¹³? 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.

IBM Watson

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

Microsoft Chatbot

Google AlphaGo

AlphaGo is a system for playing the game Go. It won against a top human professional player in 2015.

Hanson 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).

¹⁰See <http://www.ibm.com/watson/>

¹¹See <http://www.ibm.com/watson/>

¹²See <https://deepmind.com/alpha-go>

¹³See <http://www.hansonrobotics.com/>

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. Microsoft's 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 appear which isn't as big as a house, can play Go, Chess and hide-and-seek, genuinely 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.

what's the point I'm making? how does this relate to my work?

Perhaps 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 mentioned in chapter ?? live in this uncanny valley?

p and H creativity for computers?

1.4.2 BRAINS

I'm not talking about the beer or the zombie food but rather research into the human brain (or animal brains) and attempts to model it on a computer.

The motivation here is that once we understand how the brain works, perhaps we can understand how certain cognitive processes really work and this of course include creativity.

This is no easy task of course. Chris Chatham talks about ten “important Differences Between Brains and Computers”¹⁴ which give a good overview of some of the difficulties of trying to model a brain as is. We can't just do a 1-1 copy.

1. Brains are analogue; computers are digital
2. The brain uses content-addressable memory
3. The brain is a massively parallel machine computers are modular and serial
4. Processing speed is not fixed in the brain; there is no system clock
5. Short-term memory is not like RAM
6. No hardware/software distinction can be made with respect to the brain or mind
7. Synapses are far more complex than electrical logic gates
8. Unlike computers, processing and memory are performed by the same components in the brain
9. The brain is a self-organising system
10. Brains have bodies
11. The brain is much, much bigger than any (current) computer

Chris Chatham

To bring this into perspective Ray Kurzweil claims the brain is capable of 10^{16} operations per second (2013, p.194). Japan's K-computer (the worlds largest super computer as of 2016) currently has that power—10 petaflops. The “Blue Brain Project” is aiming to model 10^{17} bytes of memory and 10^{18} flops by 2023 (Kurzweil 2013, p.125).

find k-computer reference

There are currently some major research projects going on. One of them is the “Human Brain Project” (Walker 2012).

¹⁴<http://scienceblogs.com/developingintelligence/2007/03/27/why-the-brain-is-not-like-a-computer/>

quotes:

Our brain consumes about 30W, the same as an electric light bulb, thousands of times less than a small supercomputer. (Walker 2012, p.17)

For environmental and business reasons, vendors have set themselves the goal of containing energy consumption to a maximum of 20 megawatts (Walker 2012, p.41)

the 1 PFlop machine at the Jülich Supercomputing Centre could simulate up to 100 million neurons – roughly the number found in the mouse brain. (Walker 2012, p.41)

Cellular-level simulation of the 100 billion neurons of the human brain will require compute power at the exascale (10¹⁸ flops). (Walker 2012, p.41-42)

2017 petascale 50petabytes memory + 50 petaflops + <=4MW power

2021 exascale 200petabyte memory + 1exaflop

A second, equally important goal will be to prepare the procurement of the HBP Pre-exascale-supercomputer. By 2017/18, Jülich plans to procure a Big Data-centred system with at least 50 PBytes of hierarchical storage-class memory, a peak capability of at least 50 PFlop/s and a power consumption <= 4 MW. The memory and computational speed of the machine will be sufficient to simulate a realistic mouse brain and to develop first-draft models of the human brain. (The rest of the hardware roadmap targets an exascale machine in 2021/2022 with a capability of 1 EFlop/s and a hierarchical storage-class memory of 200 PB).¹⁵

Why Minds Are Not Like Computers (Schulman 2009) Software – Hardware == Mind – Brain ??? analogy

"The power of the computer derives not from its ability to perform complex operations, but from its ability to perform many simple operations very quickly."

Layers of abstraction in computers:

1. user interface
2. high level programming language
3. machine language
4. processor microarchitecture
5. Boolean logic gates
6. transistors

¹⁵<https://www.humanbrainproject.eu/high-performance-computing-platform>

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 concept 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.

1.5 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-tr
| \
* | 71437f6  Tue 18 Aug 2015  Flickr and Bing work, radio buttons work
* | 6c552aa  Wed 12 Aug 2015  Fixed image problem but not video.
| | * 1cbb63d  Tue 11 Aug 2015  (origin/thesis) Update textsurfer.py
| |
| |
* | 0ebff0d  Tue 11 Aug 2015  Analytics enabled again
* | 703f977  Tue 11 Aug 2015  Problems solved.
* | 74a1fae  Tue 11 Aug 2015  About to change l\dict to dict of dict
* | 0935b23  Mon 10 Aug 2015  BUG FUCKER
* | 4f7d91e  Mon 10 Aug 2015  Turn debug off
* | 58f0c2b  Mon 10 Aug 2015  Button styling done
* | 59add58  Mon 10 Aug 2015  Email problem solved
* |   f1b2d40  Sun 09 Aug 2015  Merge branch 'Deploy' into thesis
| \
| *
| * | 435cb2d  Sun 09 Aug 2015  Deployment works, added analytics
| * | 8a63dc7  Sat 08 Aug 2015  gunicorn runs locally fine.
| * | 2861407  Sat 08 Aug 2015  Revert 5f2c957..4026965
| * | 4026965  Sat 08 Aug 2015  Tests
* | | 8f2eeab  Sat 08 Aug 2015  Merge branch 'w3' into thesis
| \
| \ \
```

```

| | / /
| * | 5f2c957 Sat 08 Aug 2015 Stuff
| * | 873153c Fri 07 Aug 2015 Tiny cleanup
| * | 05d5760 Thu 06 Aug 2015 Random Poems and Emailing works
| * | 657126c Wed 05 Aug 2015 Random poems work - without links though
| * | 3d31ea9 Wed 05 Aug 2015 Randomise still only works once, count c
| * | 5f1d45b Wed 05 Aug 2015 Randomise poem works ONCE
| * | c583341 Wed 05 Aug 2015 Poem subtabs, email poems done
| * | f1b3878 Wed 05 Aug 2015 Hiding divs
| * | a6939c4 Tue 04 Aug 2015 huh?
| * | e6b411d Tue 04 Aug 2015 Poem emails WORK Fuck YEAH!
| * | 4b6b170 Tue 04 Aug 2015 Test email
| * | 24e356c Tue 04 Aug 2015 Better load icon
| * | e6ae736 Tue 04 Aug 2015 loading icon version 1
| * | 51b43e2 Tue 04 Aug 2015 Added 4th pictures
| * | f2d8a83 Mon 03 Aug 2015 Minor fixes
* | | 1ddb03d Mon 03 Aug 2015 Merge branch 'w3' into thesis
| \ \ \
| | / /
| * | ca4eab3 Mon 03 Aug 2015 Pretty good state.
| * | 9370334 Mon 03 Aug 2015 working on list display of images [REDACTED]
| * | e1flead 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
- * bala9c2 Second commit
- * 244c4b3 First commit

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.

ASPIRATIONS

2

Mid the silence that pants for breath,
when I thought myself at my last gasp,
haine ou de l'ambition et qui se,
the pale motor vessel withdrew its blue breath toward the island's horizon.

As pure and simple as a powder puff,
such also was the ambition of others upon the like occasion,
there was hardly a breath of air stirring,
mon ancien cœur en une aspiration vers la vertu.

After drawing a long breath,
the silver ring she pull'd,
the suitor cried, or force shall drag thee hence.

For wild ambition wings their bold desire,
and with thine agony sobbed out my breath,
I will pull down my barns.

| | | |
|-----|------------------------|----|
| 2.1 | Technical | 41 |
| 2.2 | Creative NLP | 45 |
| 2.3 | Theoretical | 46 |



Developing a software product never finishes. Especially with creative products, where the functional requirements are more fluid perhaps, it is always tempting to add, improve, replace bits.

software refactoring

For the purpose of this doctoral project, the artefact ([pata.physics.wtf](#)) is a snapshot of a product in constant motion. The state of the code at the time of § ?? submission of this thesis is described in chapter ?? and further elaborated on in the [Patanalysis](#) chapter.

Here, in this chapter I will lay out some of the potential/likely further work for this project. This may continue on a private basis or in a more academic environment. I have grouped these ideas into two main categories: **technical** and **theoretical**.

2.1 TECHNICAL

write these out all in one list and then group them as fit

Responsive spirals Currently the image and video spirals are fixed size. This means that when the webpage is resized the spiral stays the same size and is left aligned on the page. Ideally it would be better to scale the spiral with the width of the browser page. Percentages

Scalable image sizes At the moment images are retrieved at a given size through the various [API](#) calls. Because images in the spiral have different sizes according to where in the spiral they are located, they are scaled up or down directly in the Hypertext Markup Language ([HTML](#)) code. This means that some of them look squished and pixelated. This limits the available choice of results through the API.

Square aspect ratio Another issue is the aspect ratio of images and videos. For the spiral they need to be square. I currently achieve this by squishing them

as opposed to cropping them or specifying an option in the [API](#) calls to only retrieve square images.

Responsive poems A similar problem to the responsive spirals exists with the display of the Queneau poems. The random poems are centered on the page but the Queneau poems require a lot more formatting and styling to render them on the page and currently this is achieved by left aligning them and having a fixed ‘absolute’ position on the page. Ideally this would also be centered as in the random poems.

Startup performance The website can be slow to load. Currently speed performance was not a priority during development. In fact it is not built for speed from the ground up. Each time the server restarts, the indexing process takes place from scratch. This takes time. Google and other big web search engines do this continuously in the background to keep data up to date. The index is currently cached after startup but perhaps preprocessing it and storing it more permanently in a database would help speed up the start. However this may not be necessary, as it only affects the server startup.

Query speed The time it takes from the user entering a query term and the system displaying the results page varies between unnoticeable short and impatiently long. This is due to the pataphysicalisation process. This requires calls to external and internal [APIs](#) such as Flickr and WordNet.

Preprocessing corpora At this point the texts in the corpora consist of almost unedited plaintext (‘.txt’) files¹. Newlines and whitespace formatting varies, as does language and quality of spelling. OCR SOURCES Generally, chapter headings, chapter numberings, etc are left untouched. The Shakespeare corpus contains poetry and plays for example. STAGE DIRECTIONS With the plays, scene information is kept, voice details are kept. This means sentences that appear in the results of the search tool can contain peripheral words such as in this example: “...Athens and a wood near it ACT I...” from *A Midsummer Night’s Dream* or this example: “...Exit SHERIFF Our abbeys and our priories shall pay This expedition’s charge...” from *King John*. This could be addressed by preprocessing the individual texts in advance.

Sentence fragments Currently the way results sentences are retrieved for the text search is based on punctuation. This means once a pataphysicalised keyword has been found, the system retrieves up to 10 words prior until

¹For text files downloaded from Project Gutenberg, the Gutenberg specific copyright notices have been removed to only contain the relevant body of text

it reaches a punctuation mark and the same for after. The idea here was to get suitable sentence fragments.

More APIs Currently X APIs are used². This could be increased to include more varied sources of data. Sites like Flickr are heavily based on user tags ('folksonomies') which can be unreliable and a bit random at times.

Web search The use of APIs could also include web search results rather than just images and videos. This would need its own interface section and a suitable display style for the results. The biggest problem for this is API restrictions. Alternatively a ready-made index or crawl could be used but these are typically many terabytes in size and have a cost attached. Crawling the Web myself is not an option due to the computational power, time and space required to do so.

Audio search Originally audio search was going to be a part of this project. This has been abandoned due to time constraints. However it could be added using an API such as SoundClouds. Technically the pataphysicalisation could work similar to the image and video searches, meaning it would be based on user tags. One idea would be to search in audio waves.

More algorithms It would be nice to implement some more algorithms for the search tool. This could include the two additional algorithms suggested by § ?? drew Dennis (see chapter ??) or developing more of my own. This could involve implementing some of the other pataphysical principles, such as equivalence or anomaly. Or it could consist of implementing some of the more famous OULIPO techniques. The repertoire of them is huge (see appendix XYZ).

Poetry rhyming scheme One of the biggest points for future work is to introduce a rhyming scheme for the poetry results. This would involve some more Natural Language Processing (NLP) during the creation of the index. It would make the poems much more readable. See more in chapter XYZ.

Random sentences Adding to the source of random sentences used in the top and bottom banner on the website should be an ongoing endeavour.

Custom API It would be great to develop a custom API for this the search tool. This would allow other people to use the search remotely without going through the interface and to use the results as they want. This would have been beneficial for the Digital Opera project and certainly for other researchers/developers like Adnrew Dennis.

²Flickr, Getty, Bing, MicrosoftTranslator and YouTube

WordNet vocabulary The vocabulary in WordNet is limited. According to its website (<https://wordnet.princeton.edu/>) it contains 117000 ‘synsets’³. This affects two of my algorithms. Because of the way the process works, the link between Wordnet and source texts, results may be limited.

check

WordNet Antonyms The antinomy algorithms relies on WordNets antonyms. A lot of words simply do not have an opposite and no fallback is currently defined. This means a lot of the time the antinomy function will not produce any results.

1.2.1

Stemming Stemming could increase the number of results found by the algorithms. (See chapter XYZ). A danger of increasing the output of the pataphysicalisation is always that results become more boring. If the query term and potential matches were compared based on their stemmed form

Queneau’s poems It would be nice to actually add Queneau’s poem texts into the coprus of Faustroll as little easter eggs.

Bitmap algorithms The image and video search currently rely on extrenal APIs and user tags to work. One option to approach this in a totally differnet way would be to write algorithms that analyse and pataphysicalise the bitmaps themselves. So this could mean we could have a reverse image search that finds images related the original bitmap in pataphysical way or other.

Index One idea for the pataphysicalisation process was to add ‘patadata’ to the index. This could include pronounciation tags for example to make an implementation of a rhyming scheme for the poetry easier. So each word in teh index dictionary would contain the following items.

```
(``tree'': [``l_00'': [24, 566, 4990], ``s_14'': [234, 5943]], IPA data)
```

add ipa data or whatever is best for the rhyming stuff

storing rhyming data in index or other additional things like ranking

³Synonyms—“words that denote the same concept and are interchangeable in many contexts”—are grouped into unordered sets called synsets.

Stopwords Using a different set of stopwords to see if that makes a difference. For example we could use a spanish set of stopwords on an english text. OR the other way around.

2.2 CREATIVE NLP

Section ?? N-grams are a [NLP](#) technique introduced in chapter ???. The idea is that it allows for prediction of likely word pairs, meaning if the word ‘sunny’ often occurs just before the word ‘day’ in a given training text or corpus then the probability for this particular n-gram is higher than say for ‘sunny dog’. This can be increased to predict the probability of longer chains of words. One can immediately see the attraction of abusing this to generate pseudo sentences or even of creating a formula similar in nature but for example ranking obscure combinations of words higher than common ones. So for example instead of having a Maximum Likelihood Estimation ([MLE](#)) (see chapter XYZ and formular 6.12) we could have a ‘Maximum Obscurity Estimation’ defined as:

$$P(w_n | w_{n-N+1}^{n-1}) = \frac{C(w_{n-N+1}^{n-1} w_n)}{C(w_{n-N+1}^{n-1})} \quad (2.1)$$

work the maths out here for this example of MOE

Similarly, we could play with maximum entropy models as shown on page 112 (see chapter XYZ) together with Parts-of-Speech ([POS](#)) tagging. What if we rigged the probability such that instead of ‘in Quebec’ ranking high for a ‘location’ [POS](#) tag, it now ranks high as a ‘drug’?

Again there are endless possibilities of abusing these kinds of systems to create [AMC](#). This is also very reminiscent of [OULIPO](#) techniques. We could create a whole new language grammar based on pataphysical principles.

Another example of interesting uses of [NLP](#) for [AMC](#) is playing with homonyms and heteronyms. Homonyms are pronounced the same but mean something else (e.g. ‘write’ and ‘right’). Heteronyms are words that are spelled the same but have a different meaning (e.g. ‘close to the edge’ and ‘to close the door’). There are similar techniques in the [OULIPO](#). Homophones are often used to create puns (and remember—puns are syzygy’s of words), for example “past your eyes” and “pasteurize”.

You can tune a guitar, but you can’t tuna fish. Unless of course, you

play bass.

attributed to Douglas Adams

look into rhyming tags in nlp

NLP would also be useful for introducing a rhyming pattern into auto-generated poetry. BY doing POS tagging with pronunciation data, we could retrieve sentences that match the sound of the last word of the previous line, etc.

<https://wordnet.princeton.edu/wordnet/man/wngloss.7WN.html> for glossary

fix all chapter XYZ mentions

group these into better sub groups and make them proper sections rather than paragraphs

pageinate results for speed?

2.3 THEORETICAL

Focus group It might be interesting to look at opinions of various people (general public and experts) about the interpretation/evaluation framework. This could be done by asking them to provide their own definition of computer creativity and then to analyse and evaluate a product (such as [pata.physics.wtf](#)) according to their own criteria. Then follow this up by getting the same people to use my proposed framework to compare the results. This would include asking them about whether or not they thought that using the framework was beneficial to them or confusing.

Questionnaires I have shied away from doing a questionnaire study because of several reasons. One is that due to the creative and subjective nature of the artefact, opinions on it may vary wildly and I don't see how I could derive useful unbiased data from that. Yes, it depends what questions you ask. But even if I managed to get some half-decent data, what would that tell me? Half of the people like my site, the other half don't?

Eye-tracking To study the effects of using different styles of presenting the same results an eye-tracking experiment could be done. This would involve setting up participants with the necessary equipment and then introduce them to the website and monitor their eye movements as they navigate the site. This could also provide details about how long users spend on each results page, what kind of style of results they prefer, etc. Some may prefer image or video

search over the text search while others may not be interested in that at all. Generally of course one has to take into account that this is a creative piece of work and not everybody will like it. It has no clear immediate purpose and that may put users off.

Performance Benchmarks?

Part VI

HAPPILY EVER AFTER

Matter in our assistance in our quest, his journey in quest, matter of his undertaking. It was later before I felt the force of its Centre, I found out that these latter materials is a gas. Knew as much about the matter as I did when we met him, if here I enter, the gas to be formed out of the other due to, in spite of ate and her horn, Ulysses, the latter Ulysses, the latter I felt granting us intense vibration of his father

OBSERVATIONS

3

Paying no attention to his fellow mites,
mérite pas que vous fassiez attention à moi,
and told him to look after a calf she had bought,
and whilst he was looking at it attentively.

Phedon the fact affirm'd,
comment peux,
ne faites aucune attention à mon air,
in fact.

For sure Ulysses in your look appears,
was nearly out of her mind,
I omitted none of the common forms attending a royal audience.

And the consequences attending thereupon,
impotent of mind,
shape at the moment of looking at the time.

| | | | |
|-----|-------------------------------|-----------|----|
| 3.1 | Outroduction | | 50 |
| 3.2 | Issues | | 51 |
| 3.3 | Answers | | 51 |
| 3.4 | Contributions | | 52 |
| 3.5 | And Finally | | 52 |



summarise thesis, contributions etc. conclude by comparing against introduction

a wide range of subject areas such as computer science, psychology, linguistics, literature, art and poetry, languages and mathematics.

refer back to these in conclusion

3.1 OUTRODUCTION

The last XYZ chapters have explained in probably too much detail what [AMC](#) is and how to evaluate it. Given that this spans so many different disciplines the contextual background information necessary to understand the research was presented in a broad literature survey in chap XYZ. This also posed a problem for choosing the right methodology for the project. In the end a transdisciplinary approach was chosen as described in chap XYZ with a heavy component of iterative exploratory rapid-prototyping to develop an artefact to demonstrate what [AMC](#) is.

This artefact is presented on [pata.physics.wtf](#). It is an artwork dedicated to [AMC](#), pataphysics, [OULIPO](#) and programming culture.

A critique of computer creativity and its current evaluation formed the starting point for a new framework which was introduced in chap XYZ. The general conclusion of the thesis was made up of the critical analysis and further work chapters as well as this final concluding chapter right at the end.

The appendix contains various code snippets and peripheral pieces vaguely related or relevant for parts of this thesis. The code of the website is included on a CD CHECK attached to the back of the front cover. Of course the website is also available online at [pata.physics.wtf](#).

check if i need to submit a CD?

3.2 ISSUES

- Summarise issues in analysis Section
- summarize future work

3.3 ANSWERS

In the introduction I asked several questions that I attempted to answer with my research. This section contains brief answers from 50.000 feet¹, meaning they provide a top-down view of the answer and pointers to where in the thesis readers can find more elaborations.

add chapter references

these are closed questions, not research questions

Can computers or algorithms be considered creative?

§ ?? In short: no. In chapters ?? and ?? I have gone into great detail of why I believe that this cannot happen any time soon (see argument of zombies). They can be ‘creative’ (adj/adv CHECK) but the source of the creativity is the programmer of the machine not the machine itself.

Can pataphysics facilitate creativity?

Yes. Pataphysics provides many principles which can be turned into techniques and constraints which is well known to be able to support creativity

§ ?? ity (see chapter ??). This is also evident in the OULIPO and their use of § ?? constraints (see chapter ??).

Can a creative process be automated or emulated by a computer?

Yes, in theory. It mainly depends how you define the creative process and

§ ?? that is fairly subjective. See more in chapter ?? and ??.

Can human and computer creativity be objectively measured?

§ ?? No. As discussed in chapter ?? since the perception of creativity is subjective it cannot be quantified in objective terms. By providing a frame-

§ ?? work that takes into account all possible contextually relevant contributors though we can approximate an objective evaluation.

Can information retrieval be creative?

§ 1 Yes. There are many ways this can be achieved too as mentioned in chapter 1.

Can search results be creative rather than relevant?

Yes, although this is also subjective. What is creative to some might not

¹Inspired by Time Berners-Lee’s articles on the Web in 1998—
[/urlhttp://www.w3.org/DesignIssues/Architecture.html](http://www.w3.org/DesignIssues/Architecture.html)

be creative to everybody. The artefact also nicely showed the difference in
§ 1 perception of results simply based on design of the content (see chapter 1).

3.4 CONTRIBUTIONS

mention to whom these could be useful

write more

This doctoral project can be broken down into four main contributions.

- Three pataphysical search algorithms (clinamen, syzygy and antinomy).
- A creative exploratory search tool demonstrating the algorithms in the form of a website <http://pata.physics.wtf>.
- A set of subjective parameters for defining creativity.
- An objective framework for evaluating creativity.

In a more practical sense this project has spawned several publications, talks
sec ?? and exhibitions (a full list is in preface ??). Further talks were given by Andrew
Hugill at various conferences and events throughout the world where he men-
tioned my work. My publications were cited in other academic publications and
my website was mentioned on Reddit². My job here is done.

3.5 AND FINALLY

Pataphysics is the science...

²Although absolutely nobody seemed interested in it. No idea who posted it or how he found it.

INTERLUDE III

Part VII

POST ☹

Allows air to water, now and 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 coast of Tyre and Sidon, were made out of the list of Mankind, to move from thy regulation Policy.

REFERENCES

- Burnham, Douglas (2015). 'Immanuel Kant: Aesthetics'. In: **Internet Encyclopedia of Philosophy** (cit. on p. 3).
- Efron, Bradley and Ronald Thisted (1976). 'Estimating the number of unseen species: How many words did Shakespeare know?' In: **Biometrika** 63.3, pp. 435–447 (cit. on p. 17).
- 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).
- Google (2016). **Crawling & Indexing**. URL: <https://www.google.com/insidesearch/howsearchworks/crawling-indexing.html> (visited on 04/08/2016) (cit. on p. 21).
- Hendler, Jim and Andrew Hugill (2011). 'The Syzygy Surfer : Creative Technology for the World Wide Web'. In: **ACM WebSci 11** (cit. on p. 3).
- Hugill, Andrew (2012). **'Pataphysics: A Useless Guide**. Cambridge, Massachusetts: MIT Press (cit. on p. 6).
- Jarry, Alfred (2006). **Collected Works II - Three Early Novels**. Ed. by Alastair Brotchie and Paul Edwards. London: Atlas Press (cit. on pp. 3, 6).
- Kurzweil, Ray (2013). **How to Create a Mind**. London: Duckworth Overlook (cit. on p. 32).
- 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. 34).
- Schulman, Ari (2009). 'Why Minds Are Not Like Computers'. In: **The New Atlantis** 23, pp. 46–68 (cit. on p. 33).
- Walker, Richard (2012). **The Human Brain Project**. Tech. rep. HBP-PS Consortium (cit. on pp. 32, 33).
- 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).

KTHXBYE

[Go to TOC](#)