Case Board, Traces, & Chicanes

Diagrams for an archaeology of algorithmic prediction through critical design practice

David Benqué

January 2020

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

Royal College of Art
School of Communication

Word count: 39,889

Abstract

This PhD thesis utilises diagrams as a language for research and design practice to critically investigate algorithmic prediction. As a tool for practice-based research, the language of diagrams is presented as a way to *read* algorithmic prediction as a set of intricate computational geometries, and to *write* it through critical practice immersed in the very materials in question: data and code. From a position rooted in graphic and interaction design, the research uses diagrams to gain purchase on algorithmic prediction, making it available for examination, experimentation, and critique. The project is framed by media archaeology, used here as a methodology through which both the technical and historical "depths" of algorithmic systems are excavated.

My main research question asks:

> How can diagrams be used as a language to critically investigate algorithmic prediction through design practice?

This thesis presents two secondary questions for critical examination, asking:

> Through which mechanisms does thinking/writing/designing in diagrammatic terms inform research and practice focused on algorithmic prediction?
>
> As algorithmic systems claim to produce objective knowledge, how can diagrams be used as instruments for speculative and/or conjectural knowledge production?

I contextualise my research by establishing three registers of relations between diagrams and algorithmic prediction. These are identified as: *Data Diagrams* to describe the algorithmic forms and processes through which data are turned into predictions; *Control Diagrams* to afford critical perspectives on algorithmic prediction, framing the latter as an apparatus of prescription and control; and *Speculative Diagrams* to open up opportunities for reclaiming the generative potential of computation. These categories form the scaffolding for the three practice-oriented chapters where I evidence a range of meaningful ways to investigate algorithmic prediction through diagrams.

This includes, the 'case board' where I unpack some of the historical genealogies of algorithmic prediction. A purpose-built graph application materialises broader reflections about how such genealogies might be

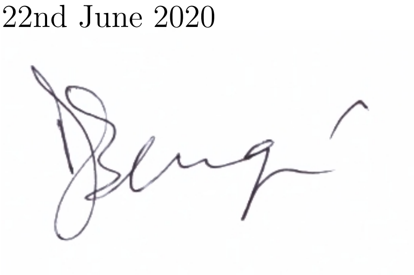
conceptualised, and facilitates a visual and subjective mode of knowledge production. I then move to producing 'traces', namely probing the output of an algorithmic prediction system—in this case YouTube recommendations. Traces, and the purpose-built instruments used to visualise them, interrogate both the mechanisms of algorithmic capture and claims to make these mechanisms transparent through data visualisations. Finally, I produce algorithmic predictions and examine the diagrammatic "tricks," or 'chicanes', that this involves. I revisit a historical prototype for algorithmic prediction, the almanac publication, and use it to question the boundaries between data-science and divination. This is materialised through a new version of the almanac—an automated publication where algorithmic processes are used to produce divinatory predictions.

My original contribution to knowledge is an approach to practice-based research which draws from media archaeology and focuses on diagrams to investigate algorithmic prediction through design practice. I demonstrate to researchers and practitioners with interests in algorithmic systems, prediction, and/or speculation, that diagrams can be used as a language to engage critically with these themes.

Declarations

Author's Declaration
This thesis represents partial submission for the degree of Doctor of Philosophy at the Royal College of Art. I confirm that the work presented here is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis. During the period of registered study in which this thesis was prepared the author has not been registered for any other academic award or qualification. The material included in this thesis has not been submitted wholly or in part for any academic award or qualification other than that for which it is now submitted.

David Benqué
22nd June 2020



Acknowledgements

This research is supported by Microsoft Research Cambridge (UK) through its PhD Scholarship Programme.

With special thanks to my supervisors for their support and dedication:
Prof. Teal Triggs (RCA) and Richard Banks (Microsoft Research Cambridge) and to examiners Prof. Jussi Parikka and Prof. Anthony Masure for engaging with my research.

Previous supervisors:
Dr. Kevin Walker (RCA, June 2016 to Sept. 2019)
Dr. Alex Taylor (Microsoft Research, Sept. 2015 to Sept. 2017)
Dr. Jonathan Edelman (RCA, Sept. 2015 to June 2016)

I would like to thank:

Prof. Anthony Dunne and Dr. James Auger for their feedback and help during the application process.

Dr. Betti Marenko for collaborating with me on the work discussed in Chapter 4, and for the conversations that helped me shape this thesis.

Tom Simmons (RCA) for his feedback in meetings and examinations.

For their time, conversation, and questions: Dr. Wesley Goatley, RYBN, Dr. James Auger, Dr. Alex Taylor, Dr. Joshua Ramey, Matt Ward, Paul Bailey. For asking short questions that stayed with me for a long time: Dr. Nick Seaver, David Chatting.

Editors and reviewers of the chapters and papers written during this research: Dr. Georgina Voss, Prof. Sandra Kemp, Dr. Benjamin Aldes Wurgaft, and the anonymous reviewers of the *Research Through Design* conference 2019.

Podcast hosts and guest: Prof. Roberto Trotta for being interviewed (*VVFA*), Jasmin Blasco (*VVFA*) and Elise Rigot (*Bio is the New Black*) for interviewing me.

Fellow RCA students: Ben Stopher, Claire Van Rhyn, Caroline Yan Zheng, Chenyi Liao, Welmoet Wartena, Larissa Nowicki, Graham Hutchinson, Kirsty Smith, and Carmen Hannibal for demonstrating LaTeX.

Family for their support: Eileen and Larry O'Brien, Maryvonne Wetsch, Charles Benqué.

My partner Claire-Anne O'Brien, and children Aoife & Tadhg Benqué both born during this PhD research.

In memory of Sascha Pohflepp with whom I was looking forward to sharing this research.

Contents

Abstract . . . iii
Declarations . . . v
Acknowledgements . . . vi
Contents . . . ix
List of Figures . . . xi

1 Introduction **1**
1.1 Research Questions . . . 5
1.2 Methodology: an archaeology of/with diagrams . . . 6
1.3 Methods . . . 9
1.4 Contributions to knowledge . . . 15
1.5 Structure of the thesis . . . 16

2 Context Review **19**
2.1 Introduction . . . 19
2.2 Data Diagrams . . . 21
2.2.1 Vectorisation and operations . . . 22
2.2.2 Visualisation . . . 30
2.2.3 Predictive positivism . . . 36
2.3 Critical Diagrams . . . 39
2.3.1 Framed and framing . . . 39
2.3.2 Prescription machines . . . 43
2.3.3 Diagrammatic excavations . . . 49
2.4 Speculative Diagrams . . . 54
2.4.1 All possible combinations . . . 55
2.4.2 Designing in/with possibility space . . . 60
2.4.3 Computational speculation . . . 66
2.5 Conclusion . . . 69

Preamble to the practice chapters **73**

3 Case Board 77
3.1 Introduction . . . 77
3.2 The history of prediction: a diagram of diagrams . . . 78
3.2.1 Big data's historical burden . . . 80
3.2.2 Taming the future into a space . . . 81
3.2.3 Genealogies of prediction . . . 85
3.3 Practice . . . 88
3.3.1 Case boards and yarnwork . . . 88
3.3.2 Diagrams of the Future . . . 91
3.3.3 A patacritical graph . . . 97
3.4 Discussion . . . 100
3.4.1 Starting in the middle . . . 100
3.4.2 The x-axis as an abductive arc . . . 104
3.4.3 Publishing the *DOTF* archive . . . 107
3.5 Conclusion . . . 109

4 Traces 111
4.1 Introduction . . . 111
4.2 YouTube recommendations: the funnel . . . 113
4.2.1 YouTube as capture apparatus . . . 113
4.2.2 Map the trap . . . 118
4.2.3 Seeing vs. knowing . . . 121
4.3 Practice . . . 122
4.3.1 Data Collection . . . 122
4.3.2 Visualisation . . . 124
4.3.3 Traces . . . 130
4.4 Discussion . . . 134
4.4.1 Enclosed and hosted . . . 134
4.4.2 Traces as conjectural knowledge . . . 137
4.4.3 Publishing and broadcasting traces . . . 139
4.5 Conclusion . . . 140

5 Chicanes 143
5.1 Introduction . . . 143
5.2 The almanac as a prototype for divinatory data analytics . . . 146
5.2.1 A cosmic imaginary . . . 146
5.2.2 Data for divination . . . 149
5.2.3 Almanacs in the petabyte age . . . 153
5.3 Practice . . . 157
5.3.1 Computation and divination in critical practice. . . . 157
5.3.2 Almanac tooling . . . 160

5.3.3 Cast of widgets 162
5.4 Discussion 169
5.4.1 Sincere chicanery 169
5.4.2 Stochastic Arrows 171
5.4.3 Automated publication of predictions 173
5.5 Conclusion 174

6 Conclusion and Future Work **177**

A Note on repositories **183**

B Thesis repository **185**

C Diagrams of the Future **191**

D Architectures of Choice Vol.1: YouTube **199**

E The Monistic Almanac **203**

F Possibility Space **227**

G The Presage Range **239**

H List of conferences, exhibitions, and events attended **245**

Bibliography **249**

List of Figures

1.1 Experimental design (Box, 1976) 12

2.1 The data science process (O'Neil and Schutt, 2013) 22
2.2 Iris Dataset (excerpt) (Fisher, 1988) 23
2.3 Vector, matrix, and tensor (Goodfellow et al., 2018) 23
2.4 Nearest Neighbours (Bernhardsson, 2015) 24
2.5 Iris dataset in Python 25
2.6 Iris Nearest Neighbours (scikit-learn) 25
2.7 Scikit-learn classifier comparison (Pedregosa et al., 2011) . . . 27
2.8 Over and Under Fitting (Lavrenko, 2015) 28
2.9 *Neural Network Zoo* (van Veen, 2016) 29
2.10 *Sémiologie Graphique* (Bertin, 1967) 31
2.11 Python visualisation landscape (VanderPlas, 2017) 32
2.12 Iris scatter-plot matrix (Bostock, 2018) 33
2.13 t-SNE (van der Maaten and Hinton, 2008) 35
2.14 The Vectoralist Class (Wark, 2015; Anthony, 2017) 41
2.15 *White Collar Crime Risk Zones* (Clifton et al., 2017) 42
2.16 COMPAS matrix (NYS Division of Criminal Justice Services, 2012) 44
2.17 *Face Cage 2* (Blas, 2013) 47
2.18 Facebook Patent US9740752B (Pereira, 2019) 48
2.19 Zā'irja (Link, 2010) 50
2.20 Police Drawing Workshops (Prévieux, 2011) 54
2.21 Ramon Llull's algorithms (1305) 56
2.22 *Predictive Art Bot* Diagrams (Roszkowska and Maigret, 2015) 57
2.23 *Predictive Art Bot* (Roszkowska and Maigret, 2015) 58
2.24 Semantic Space (Parrish, 2015) 59
2.25 Futures Cone (Candy, 2010) 61
2.26 Alternative presents and speculative futures (Auger, 2012) . . 62
2.27 *Real Prediction Machines* Bayesian network (Auger, 2014) . . 62

2.28 *Extrapolation Factory* method diagrams (Montgomery and Woebken, 2016) 63
2.29 Impossible Object (Schmitt, Dunne and Raby, 2018) 65
2.30 Project timeline 75

3.1 *Diagrams of the Future* Case board 79
3.2 Correlation diagram (Galton, 1886) 82
3.3 Eugenist view of British social structure (Galton) 83
3.4 *The Evolution of Machinic Intelligence* (Pasquinelli, 2015a) . . 86
3.5 *World Government* (Bureau d'Etudes, 2004). (redacted) . . . 89
3.6 *Diagrams of the Future* homepage 92
3.7 *Diagrams of the Future* topos page 93
3.8 *Diagrams of the Future* editing interface 94
3.9 Nodes and relationships in Neo4j graph database. 95
3.10 Example Cypher query 96
3.11 Digital Humanities vs. Speculative Computing (Drucker, 2009) 98
3.12 *Temporal Modelling* sketch (Drucker, 2009) 99
3.13 *Diagrams of the Future* sketch of the editing interface 99
3.14 *Diagrams of the Future* graph structure 102
3.15 *Diagrams of the Future* graph 103
3.16 *Diagrams of the Future* iterations 107

4.1 *Architectures of Choice* poster 114
4.2 YouTube's funnel architecture (Covington et al., 2016) 115
4.3 Funnel as distiller vs. funnel as trap 118
4.4 *YouTube Map* (Chaslot et al., 2016a) 120
4.5 Three probes 125
4.6 Circular visualisation 126
4.7 Mermaid.js graph 127
4.8 Gephi graph 128
4.9 Neo4j Browser visualisation 129
4.10 Trace print 131
4.11 Trace animation 132
4.12 Trace sequence 133
4.13 The paper as primary mechanism for research (Peyton Jones, 2016) 138

5.1 *The Monistic Almanac* instalation detail 145
5.2 *Annuaire pour l'an 1875* (Bureau des longitudes, 1875) 148
5.3 Old Farmer's Almanac calendar (Thomas, 1976) 150
5.4 Euro Millions Astro Indicator (Old Moore's Almanack, 2017) . 151

5.5 Old Moore's Almanac homepage (Old Moore's Almanac) . . . 154
5.6 Daily calendar web page (Farmers' Almanac) 156
5.7 *Co—Star* astrology iPhone app (Co – Star Astrology) 157
5.8 Amazon Prime horoscope (Katz, 2019) 158
5.9 Cosmic Commodity Chart 163
5.10 Crisis Proximity Index 164
5.11 Carbon Prophet 165
5.12 Electional Astrology Calendar 166
5.13 Power Grid Harmonics 168

C.1 Timeline Prototype 192
C.2 MSRC Summer School Poster 193
C.3 *Counting the Future* MVP 194
C.4 *Counting the Future* version 2 195
C.5 *Counting the Future* graph redesign 196
C.6 *Diagrams of the Future* 197

D.1 Trace animation 200
D.2 Traces streaming intructions 201

E.1 *Supra Systems Office Rites* 204

F.1 *Possibility Space* - Title Card 234
F.2 Futures Cone 234
F.3 Clusters 235
F.4 Decision Surface 235
F.5 Haruspex 236
F.6 Probability Space 236
F.7 Tea Cup 237
F.8 Probability Distribution 237

G.1 Boltzmann Scanner 241
G.2 Waddington Surface 242
G.3 Richardson Observatory 243
G.4 Accessories 244

Chapter 1

Introduction

Your data already knows the future.

The website of a data-analytics start-up (Verteego, 2017) makes a big promise: data have predictive powers. No isolated spreadsheet is likely to start dispensing predictions on its own, however. Data are operationalised as part of algorithmic systems—entanglements of data, people, and computation—in order to produce predictions. These elements are impossible to separate from each other (Geuter, 2017), and from culture itself (Seaver, 2017). Processes of *algorithmic prediction* move data from collection/production, through mediation, to prediction, in flows that are currently re-shaping the world in powerful ways.

Predictions—defined for the purposes of this research as knowledge inferred through algorithmic processes on the basis of past data—are currently proliferating on an unprecedented scale. They are not limited to the future. Many predictive systems are focused on very short term results—for example, predicting what advertisements are likely to be clicked on a web page (McMahan et al., 2013)—while others explicitly aim to 'predict the present' (Choi and Varian, 2012). Part of my argument in this thesis is that neither data nor algorithmic prediction are new, however they are being mobilised now as part of a paradigm shift that is challenging the very nature of knowledge. This shift, once centred around "big data" before being re-branded as "artificial intelligence," is well documented (Crawford et al., 2014; Kitchin, 2014; Rieder, 2016). It can be characterised as the redefinition of all problems as prediction problems, aided by vast amounts of data and 'cheap' computation (Agrawal et al., 2018, 38).

Knowledge produced through algorithmic prediction is made valid and legitimate through a digital form of positivism: 'the understanding that the social world can be known and explained from a value-neutral, transcendent view from nowhere in particular' Jurgenson (2014).[1] Corollary to this position is a view of data as a 'raw' material (Gitelman, 2013), a direct and un-mediated measurement of reality (Mosco, 2014; Fuchs, 2017). With reality considered stable—and observable from an objective distance through data—it follows that the future too can be known, through the inference of causes and effects. In this epistemological position, the only challenge to total knowledge of the present and future is to gather *enough* data to capture reality in its entirety. 'With enough data, the numbers speak for themselves' Anderson (2008) famously proclaimed in the pages of *Wired Magazine*.

Promises made about and with data are part of what Beer (2019) calls the 'data imaginary.' As illustrated by the quote at the beginning of this chapter, this imaginary positions data as 'panoramic' (29) and 'prophetic,' (32) among other characteristics. An industry of intermediaries, arguably a whole new form of capitalism (Srnicek, 2017), mobilises this imaginary as they transform data into "insights," peeks into an unknown future. They advertise a form of 'prosthetic vision' (Beer, 2019, 7), a *gaze* with 'both sight *and* foresight' (32). To produce knowledge about the future, this gaze peers *deep*, below the surface level of perception (28) and into high-dimensional mathematical spaces constructed from data through algorithmic processes such as machine learning. This knowledge is legitimised by the ideological claim that the depths and volumes of data contain knowledge that is unavailable, and *superior*, to that available to our senses (McQuillan, 2017). Once predictions are surfaced through algorithmic processes, they are folded back onto the present to inform decision-making, re-shaping the world to bring computed futures into being (Adams et al., 2009; Beer, 2019, 29).

This thesis is about finding ways to *gaze* back at algorithmic prediction, to grapple with how "deep volumes" are created from data, and with the operations that "extract" predictions from them. My entry point was to consider the ways in which predictions are materialised, for example through the design of visualisations and interfaces. As I approached prediction with a design sensibility, I realised that these processes are *diagrammatic* all the way down—from the "normal distributions" of

[1]Positivism dates back to Auguste Comte and his *Course in Positive Philosophy* 1830-1842 (Moatti, 2017).

probability theory to "big data" operationalised as high-dimensional vector spaces in machine learning (Mackenzie, 2017). Algorithmic prediction activates data by spatialising them, and by constructing statistical *relations* between them, in other words by turning them into diagrams. From the standpoint of design research and practice, thinking with diagrams provides some purchase on algorithmic prediction. Through them, my research position is engaged on one side with the 'operational formation' (Mackenzie, 2017, 43) of data and code, and on the other with critical perspectives on their social, cultural, and political ramifications. As a designer practitioner and researcher I am at once in active contact with materials through practice, and attentive to the ideas, concepts and theories that these materials reify. Design's unique capacity for 'moving ideas through practice'[2] is, in my view, the main vector for its potential to contribute to broader debates in the (digital) humanities (Masure, 2017). In this context, diagrams provide a vocabulary for research and practice, a way to make prediction available for examination, experimentation, and critique. I use this language to "read" algorithmic prediction as a sophisticated set of computational geometries, and to "write" it as I conduct my research through critical practice immersed in the very diagrams in question: data and code.

Through diagrams this research aims to appropriate some of the materials and tools of algorithmic prediction, and to mobilise them towards a different kind of knowledge production; one that is rooted in creative practice, situated, incomplete, and speculative. I take the diagram as a device that connects practice to theoretical considerations. Specifically, diagrams are in constant 'oscillation between systematisation and openness' (Leeb, 2017, 31). On the one hand diagrams are instruments of control, instructions for 'the process whereby power relations are produced through relationships of strength.' (32) On the other, they are speculative instruments, generating "lines of flight" in a 'kind of cognitive sweep that extends the possibilities of thought' (Knoespel, 2001, cited in Leeb, 2017, 33).[3] It is important to note that these are not separate types of diagrams, but qualities present in each and every one. In this research, I bring these oscillations to bear on algorithmic prediction as I read them as control diagrams, and as I attempt to construct generative diagrams.

[2] Prof. Teal Triggs (supervisor) in personal conversation (April 2020).

[3] These theoretical views of the diagram famously come from the work of Deleuze (1988, and other parts of his work). In this thesis, my contribution is not on this level of theory, rather I draw from it to inform my research and practice. I mainly engage with discussions of these works by others (Knoespel, 2001; Leeb, 2017; Marenko and

Taken together, these considerations about algorithmic prediction, its proliferation, its diagrammatic nature, and the potential of diagrams for research through creative practice, lead me to articulate the research questions outlined in the next section.

Brassett, 2015). The idea of oscillations between systematising/classification/ordering and opening/speculation/generation is the central notion I draw from.

1.1 Research Questions

My initial questions, as posed in the proposal for this research, were centred around finding new ways to visualise predictive models, and on the influence of visual languages on the perception of these models by end users. Through this research, the diagram has emerged as a way of seeing algorithmic prediction beyond purely issues of representation. Hence the questions below are both the *result* of the research process throughout this Ph.D. project, and the *starting point* for presenting the research in its final form in this thesis.

My primary research question is as follows:

RQ How can diagrams be used as a language to critically investigate algorithmic prediction through design practice?

As I began to unpack the terms as indicated in this main question, the process led to the following two sub-questions :

RQ1 Research question one positions diagrams as bridge between research and practice: Through which mechanisms does thinking/writing/designing in diagrammatic terms inform research and practice focused on algorithmic prediction?

RQ2 Leading to research question two which posits that to "critically investigate" algorithmic prediction means to interrogate the ways in which it produces knowledge: As algorithmic systems claim to produce objective knowledge, how can diagrams be used as instruments for speculative and/or conjectural knowledge production?

1.2 Methodology: an archaeology of/with diagrams

Diagrams are the focus of my research because they afford a practice-based mode of engaging with algorithmic prediction. They open up a space for critical practice *with* rather than simply *against* materials such as data and code. Much of the critical scholarship on algorithmic prediction operates from a 'high' theoretical position (Fuller, 2018, 251), with little contact to the "stuff" itself. In contrast, diagrams offer ways to reconcile theory and critique with materials and practice. In other words they create opportunities for design to operate critically while 'moving at the same level of abstraction as the algorithm' (Pasquinelli, 2015b). Through the process of this research, I came to define this methodological position in line with media archaeology—a field of media-studies and artistic practice focused on "digging up" the forgotten artefacts and overlooked genealogies of media history, drawing in part from Foucault's (1972) *Archaeology of Knowledge*. In this section I discuss the specific terms on which I engage with, and borrow from, media archaeology.

While this was not declared from the outset of my research, the adoption of this methodology was less a "turn" than a realisation that the research I had already begun to undertake was of an archaeological nature.

Like the data-gaze, media archaeology aims to explore 'depths' but it does not claim to surface total or superior knowledge. Instead it stays down, in the 'underground material strata of media' (Bardini et al., 2016, my translation) to explore, decipher, unravel strategies[4], find poetry and aesthetics. Its claim to knowledge production are much more humble. It does not propose an objective gaze but instead offers glimpses distorted by the very machines they attempt to observe.

The original proposal for this research set out a project around the aesthetic materialisation of predictive models. I moved through a series of stages to refine this as I focused the research. First I was concerned with the "aesthetics of accuracy," then with "possibility space," and finally found the right combination of practical and discursive possibilities in "diagrams". This was informed by the work of Mackenzie (2017), about halfway through the research, which resonated with my own trajectory of 'learning to machine learn' (18). Theirs is an immersive practice that takes

[4]Industrial, economic, logistical, geopolitical, and ideological strategies (Bardini et al., 2016, my translation)

the materials of algorithmic prediction (diagrams, data, code, formulas, books, tutorials, etc) as the focus for a critical investigation informed by theory.

> My somewhat risky, naive immersion in technical practice seeks to support an alternative account of machine learning, an account in which some schema, analogy, imagining, and sense of agency can take root. Mundane technical practices, sometimes at a quite low level (e.g., vectorization) and other times at a high level of formalization (e.g., in discussing mathematical functions), are elements to be drawn—sometimes literally, sometimes operationally—on a diagram. (Mackenzie, 2017, 19)

Mackenzie combines in-depth practical understanding of machine learning with 'the theoretical resources of a media-focused archaeology of knowledge and a science studies-informed ethnographic sensibility' (18). He uses the diagram throughout as an elastic and morphing concept, stretching from the intricate geometries of machine learning all the way to critical theory. Mackenzie's work was instrumental in focusing my research on the deeply diagrammatic nature of algorithmic prediction, and the ways diagrams can be used to draw relations to critical literature from the 'depths' of algorithmic machines. As a design researcher my approach differs significantly from this, though, as it is not aimed at theory but informs a creative research practice. While my focus on diagrams was largely informed by Mackenzie's work on machine learning, it is also broader. It considers machine learning as only one instance of diagrammatic algorithmic prediction, although a very prominent one.

Mackenzie's 'archaeology of knowledge' approach is focused on power and knowledge, close to the original formulation by Foucault. This was my entry point to the broader field of media archaeology, which includes artists and creative practitioners as well as theorists (Bardini et al., 2016; Huhtamo, 2011b; Parikka, 2012), some of whom are specifically focused on studying algorithmic systems from an archaeological standpoint (Link, 2016; Pasquinelli, 2017, 2019; RYBN). They share a critical stance, centred around unpacking power relations and knowledge production as they manifest in technological artefacts. The practice-oriented side of media archaeology takes creative license with these theoretical questions however, it puts them through the test of practice and through the filter of intuition, aesthetics, and interpretation.

Media archeology informs my creative research practice. I focus specifically on the core notion of *excavation* as I unpack some of the material/diagrammatic qualities of algorithmic prediction, both historically and in the current 'regime of anticipation' (Adams et al., 2009; Mackenzie, 2013). This follows:

> ... a two-fold understanding of media archaeology: as excavating longer time-spans in order to understand the conditions for the contemporary scientific media culture, and as excavating the technicalities of current technologies in order to understand how they frame our living world. (Parikka, 2012, 151)

Parikka (2012) points to the 'privileged position' of designers to conduct these excavations, as they operate within the 'actual practices that constitute media culture' they can mobilise this towards critical ends (156). As it gets mobilised in practice, the notion of excavation gets stretched to include the examination and creation of new and/or imaginary media, with this comes further stretching of the original, Foucauldian definition of an archaeology of knowledge. In this context, I understand excavation to be a disposition towards critically unpacking technological systems through practice, whether these are historical, current, or imaginary. In this thesis I move through these three stages following the distinction between 'media remembered, media observed, and media imagined' (Blegvad, cited in Huhtamo and Parikka, 2011, 55). This is a porous distinction however, as my remembering and observing are also acts of imagining. Parikka (2019), emphasises the inherently speculative nature of media-archaeological research and practice where he proposes:

> ... a methodological way of approaching reality not as ready-made and finished, but as produced and open to further variations, potential, and a temporality that includes the possibility of something else. (Parikka, 2019, 205)

This resonates with my approach through creative practice, and with my interests in prediction and speculation. While the idea of *excavating* algorithmic prediction may imply that I could somehow get "to the bottom" of it, this approach points to a more speculative activity where the very 'coordinates of space and time' (206) are interrogated and re-configured.

In summary, my use of media archaeology as a methodology comes from a shared positioning in the "depths" of technical systems, a disposition towards making and 'thinkering' (Huhtamo cited in Parikka, 2012, 157) as a way of knowing with and about digital media, and an attention to material forms—such as diagrams—as manifestations of cultural, economic, and political forces. I also subscribe to the interdisciplinary nature of the field, merging inputs from the (digital) humanities, social sciences, computer science, art, and design. However, this thesis does not contain an in-depth theoretical review of archaeologies of knowledge,[5]. Instead, literature and key ideas such as *excavation* are presented and mobilised throughout the text, according to the specific contexts of my practice-based projects and their discussion.

1.3 Methods

The research presented in this thesis is conducted *through* design practice (Frayling, 1993), meaning it presents and discusses knowledge produced by the process of designing a series of artefacts. In my case, this practice is primarily rooted in graphic and interactive design, and produces digital media artefacts such as web applications and publications. To borrow from industry parlance I take a "full stack" approach where I consider everything from how my projects look to how they are hosted and "served" on the internet as part of the design process.

Research *through* design also means, in my case, that the projects presented here are not discussed as finished artefacts that encapsulate knowledge but as processes leading to the substantiation, and/or re-definition, of my initial hunches. My focus is not on outcomes—technically advanced or aesthetically polished—but on trajectories where my initial position shifts through making. I consider practice as 'a conversation with the materials of a situation' (Schön, 1983, 78). In this case the 'situation' is algorithmic prediction. Practice, in this sense, means a space and time of making reflexively to engage with the 'conceptual matter' (Sayers, 2017) of algorithmic prediction. In this section I discuss the key methods that have informed my practice in this research.

Instrumented research

Archaeology is conducted through instruments, from shovels and brushes, to scanning probes, to specimen databases. Similarly, I consider diagrams as instruments to excavate algorithmic prediction. However, these are not scientific instruments aimed to produce objective and definite knowledge.

[5]For example, I do not detail the differences between Foucauldian and Kittlerian approaches to media archaeology.

As I have mentioned above in Methodology: an archaeology of/with diagrams, they instead aim to unpack and create sites for interpretation and critical reflection. The tensions between computation and knowledge production are the subject of ongoing concern for researchers[6] that take social and cultural dimensions into account. While my instruments use data and computation, they are like the 'mechanical aids to humanities interpretation' described by Nowviskie (2004, 90), making their indeterminacies and imperfections inseparable from the knowledge they produce.

While computational instruments promise new, otherwise unattainable knowledge, they often "smuggle in" positivist assumptions of observer independence and mechanical objectivity; especially into humanities inquiries that aim to preserve interpretation (Drucker, 2011). These tensions are well known to the digital humanities and are summarised by Masure (2017) as a stratified landscape, where three combinations of computation and knowledge production have been evolving over time, not replacing each other but sedimenting above one another. The first stratum establishes the *application* of computation to humanities research through the "mining" of digitised texts, the automation of linguistical analysis, and so on. This is a quantitative effort where data storage and processing power dictate the 'modalities of access' to knowledge (Masure, 2017, 29). The second stratum is a more qualitative turn (see Schnapp and Presner, 2009) where criticality and interpretation are foregrounded. The role of design in shaping arguments made through digital tools is recognised (Burdick et al., 2012, 118) and other overlaps emerge, such as using the "project" as a unit of work.

I position my instrument-making in Masure's third stratum, that pushes further against an instrumentalist view of technology while also moving closer to examine its intricacies, 'sufficiently immersed in the "making" with digital matter to detect singular aesthetics within it' (Masure, 2017, 37 my translation). Crucially, in this combination of computation and knowledge production, the tools themselves come under scrutiny. Meanwhile the final goal of the research is not known in advance, but allowed to be defined and redefined in conversation with the instruments themselves. In the words of Fuller (2018) computational instruments are not means to get definitive answers, but 'one of the fields in which the crystal grows' (252)—before being ground up, mixed, ingested or otherwise

[6] For example, those at the intersection of media studies and digital humanities (Sayers, 2018) or in social sciences using digital methods (Hargittai and Sandvig, 2015).

observed (255). The identification, or growth conditions, of these "crystals" is not determined in advance but emerge through experimentation and a speculative approach, to which I turn next.

Speculative/abductive practice

I have described an immersive practice of excavation with materials such as data and code. While I seek to get as close as possible to the materials and practice of algorithmic prediction, I draw a key distinction with it in terms of knowledge production by refusing *inference* through either induction or deduction from data. Instead, to address RQ2, my practice-based method is *abductive*. As I have discussed earlier in Methodology: an archaeology of/with diagrams, the *immersions* and *excavations* of my archaeological approach are inherently speculative. This translates into a 'method for the unnatainable' as set out by Luciana Parisi (2012):

> The true method of speculation is like the flight of an airplane. It starts from the ground of particular observation; it makes a flight in the thin air of imaginative generalization; and it again lands for renewed observation rendered acute by rational interpretation. (Whitehead 1929 quoted in Parisi, 2012)

There are multiple understandings of abduction and what it does,[7] notably the level of rationality to which the result of the "flight" should be returned,[8]or the use of abduction as part of the political regime of anticipation, using the future to demand action in the present (Adams et al., 2009, 255). In the context of this research I draw from Parisi's work, and make a practice-oriented reading of it. In this light I present research projects that begin with interrogations, hunches or '*vibes*' (Parisi, 2012, 235), from there a conversation between practice and research begins, launching a *probehead* that morphs and changes through the work before returning to be examined and discussed. This echoes definitions of design as a 'methodology of doubt' (Sottsass quoted in Masure, 2017, 12). The key notion for me is to allow for the intricacies of practice and the critical considerations of research to inform and re-shape each other through the abductive process, which I refer to as "abductive arc" in the discussion

[7] Adams et al. (2009) for example consider abduction as part of the 'regime of anticipation' that includes algorithmic prediction. 'Abduction moves reasoning temporally from data gathered about the past to simulations or probabilistic anticipations of the future that in turn demand action in the present.' (255). Others, for example Drucker (2009, 25, see fig. 3.11) position abduction in opposition to computational logics of induction and deduction. I use the latter sense.

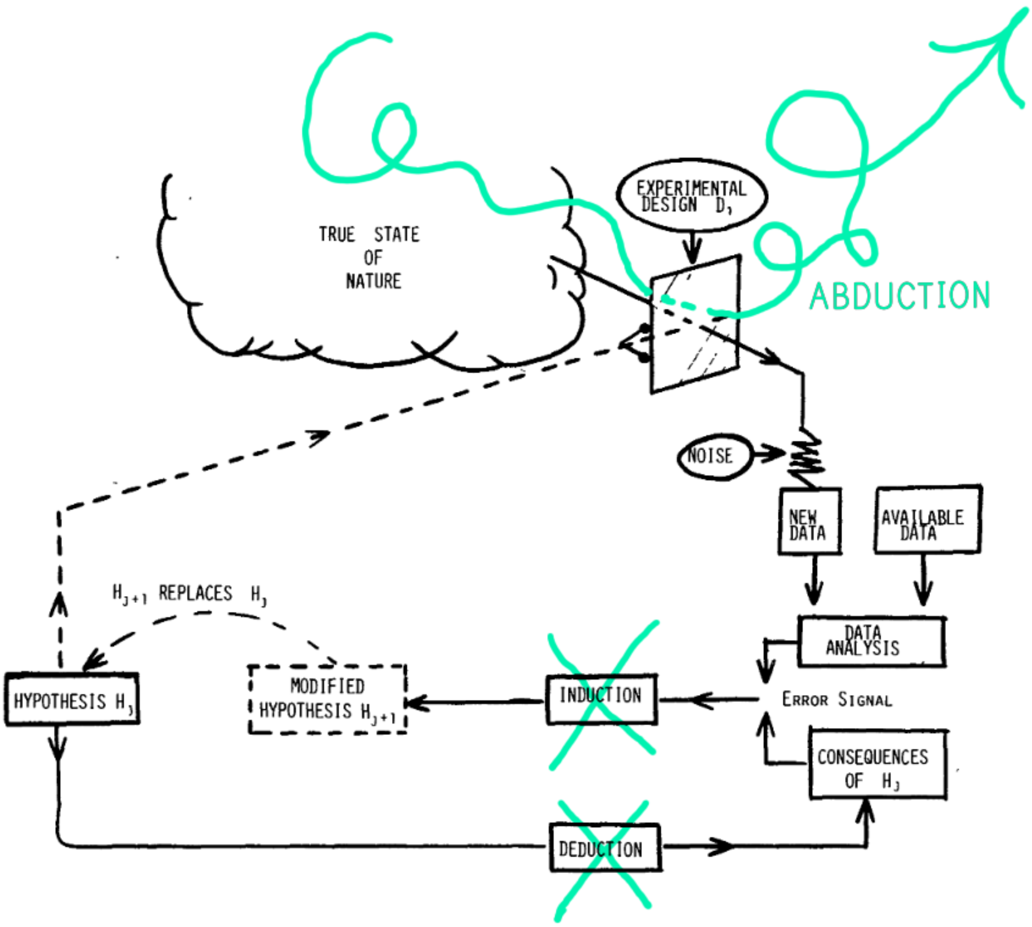


Figure 1.1: Experimental Design. Diagram in black from Box (1976, 796) 'The experimental design is here shown as a movable window looking onto the true state of nature.' My annotations added in green.

section of the practice chapters (3.4, 4.4, 5.4). At the core of this speculative approach is a challenge to the 'positive assumption that *knowing* should order *making*' (Monjou, 2014, cited in Masure, 2017, 54).

Publication practices

Finally, my research and practice are geared towards the production of publications. By publication, I specifically mean media artefacts with the

[8]My understanding of the initial notion of abduction by Charles Sanders Pierce is that the brief "flight" of the probehead was quickly folded back into a pragmatic, scientific mode of knowledge production.

explicit purpose to be accessed and disseminated. Additionally, these artefacts are considered as the result of a publishing process, encompassing everything from editorial intent to technical constraints. In this research, examples of publications include an archive of diagrams, an academic paper, animations (considering the broadcast as a form of publishing), and a web-page that produces booklets. These are published in different ways, but all make use of the internet and related technologies for their dissemination.

Following Monjou's (2014) insightful discussion into the publicity of design research, the three projects presented in this dissertation all involve one or more forms of publishing. This ranges from an archive of predictive diagrams, to a peer reviewed conference paper, to streamable web-based animations, to an automated almanac, to book chapters. This leads me to question what it means to publish, and the forms through which design practice, and design research, gets propagated and validated. As design seeks to establish itself as a valid scientific discipline it conforms to the formats and processes of scientific publishing. With this move comes a 'dessiccation' (Blanc and Haute, 2018, my translation) of the specific types of knowledge that design produces; rooted in experimentation, practice, and interpretation. Instead, Blanc and Haute (2018) argue, design should celebrate a multiplicity of forms of publishing. Seeking alternative modes of publication is, in my view, an integral part of design inquiry.

Publishing as a mode of creative practice creates opportunities for 'site specific gestures and critical interventions' (Gilbert, 2016, 20) through processes of 'filtering and amplification' (11). Publishing also raises questions around distribution, as well as commitments to temporality and maintenance.[9] For each of the projects presented here, I consider their 'publicity': the reasons for, and specificity of, their mode of publishing (Monjou, 2014).

Note on speculative and critical design

The words "design," "speculative," and "critical" are used in proximity to each other throughout this thesis. Given my background in close association with the small, research and/or education oriented sub-field

[9]Considerations around what kind of temporality one commits to when putting a design project out in the world, publication or otherwise, were raised by Carl Di Salvo when he visited the Royal College of Art's School of Communication Post Graduate Research Summer School in July 2017. I believe they are particularly relevant when publication is involved.

known as *Speculative and Critical Design* (SCD) this warrants some clarifications on the positioning of this research.

SCD attempts to unpack the socio-cultural implications of technology by exploring "possible futures" through design (Auger, 2014; Dunne and Raby, 2013). This involves fictional scenarios that are made tangible through props, objects, products, videos, and so on. In this research, I take prediction and speculation themselves as research topics. I make no use of scenarios, nor do I extrapolate on what current modes of prediction might look like in the future. Effectively, I am doing the opposite, turning to the past and history to unpack our current moment. As such this research is not SCD.

Diagrams have played a key role in establishing SCD, in particular the *Futures Cone* (fig.2.25) as a way to map and generate futures. However SCD rarely examines the "coordinate space" in which its extrapolations take place. In this thesis I turn my attention to the transformation of data into coordinate, multi-dimensional spaces that are operationalised as part of predictive systems. The primary focus of my research is to investigate how futures are produced through diagrams, rather than using diagrams to generate new ideas as is the case in SCD.

However, I do not see this as a hard rupture, if anything I am a "designer interested in speculation" rather than a "speculative designer." In my view there are significant overlaps between speculative design and the media archaeology informed approach I am taking here; as Parikka (2019) argues, both fields might in fact benefit from 'cross-fertilising' (205) their approaches. In practice my SCD work has always involved digging up histories of science and technology and taking their overlooked oddities as departure points for new narratives. Gradually this took over as the focus of my practice. Overall, as Parikka (2019) points out, the two fields share an interest in 'rethink[ing] the usual coordinates of time and space.'

There are also continuities, for example I review the use of diagrams as speculative devices in SCD in section 2.4.2. More broadly, in my view, SCD was always about media. It took the "language" of industrial design and subverted it for discursive purposes. *The Monistic Almanac* project I discuss in Chapter 5 can be viewed through this angle, applying a kind of 'industrial realism' (Dunne, 2005, 90) to a graphic design oddity.

Overall, my focus on media archaeology, on excavations rather than extrapolations, and on producing publications rather than exhibitions, are enough of a rupture to keep mentions of SCD as a method confined to this note, and to refer instead to the methods and methodology outlined in this chapter throughout this thesis.

Discussion criteria

In each of the practice chapters, I discuss the research through a set of criteria informed by the methodology and methods outlined above, in summary these are:

- Diagrammatic oscillations and movements, relating to RQ1 and Instrumented research: I discuss how aspects of algorithmic prediction can be read as diagrams, and how my practice writes diagrams in relation to these. Oscillations between control/systematising and openness/speculation are a recurring focus throughout the practice chapters. I am also attentive to the specific forms and movements that each of my proposed diagrams affords.
- The "arc" of the research project, relating to RQ2 and Speculative/abductive practice: I evaluate and discuss my practice-based research as trajectories rather than only the final outcome. In each case, I locate knowledge production in the way that my initial position has been refined or changed through the research.
- Modalities of Publication practices: Finally I discuss how my focus on producing digital publications manifests in each of the practice chapters. This includes the choice of format(s) and its implications in relation to critical practice and technical implementation.

1.4 Contributions to knowledge

The original contribution to knowledge with this research is the development of an approach to practice-based research, drawing from media archaeology and focused on diagrams, to investigate algorithmic prediction through design practice. This work is rooted in design practice, but looks to other fields such a media studies and the digital humanities for methodological framing. My intent is to demonstrate to researchers and practitioners with interests in algorithmic systems, prediction, and/or speculation, that diagrams can be used as a language to engage critically with these themes.

This contribution is not a generalisable template or framework; it emerges out of three distinct diagram types—Case Board, Traces, and Chicanes—that are each manifested by a chapter in this thesis and a code repository as a record of original creative work. These demonstrate my approach and produce smaller, practice-oriented contributions. While they are tailored to the specific conditions of each research project, these proposed diagram types define the characteristics of a broader diagrammatic language—namely its purchase on algorithmic systems, making it suitable for an archaeology of algorithmic prediction; its range, including both design practice and research writing; and its disposition towards speculative/abductive modes of knowledge production. While this language is not readily applied to other contexts, it serves as a reference to be extended or adapted.

Finally, by being itself a code repository and a media object—reflecting my archaeological approach by using the LaTeX language, widely used in scientific publishing—this thesis contributes to recent efforts to make research more open and accessible. Researchers in design (Maudet, 2017; Masure, 2014) and elsewhere (Guy, 2017) have recently published their doctoral work as websites and/or repositories. Each of these examples have merits and trade-offs. I am contributing another attempt which is undoubtedly imperfect as well, but nonetheless offers an alternative by exploring the use of LaTeX in a design context.

1.5 Structure of the thesis

This thesis begins by establishing the diagrammatic nature of algorithmic prediction. Chapter 2 Context Review details the diagrammatic forms and processes that mediate the transformations of data into predictions. These are described as the technical foundations of digital positivism. The chapter continues in its examination of the oscillations of algorithmic prediction between control diagrams: conservative prescription machines, and speculative diagrams: generative probe-heads.

My main research question - How can diagrams be used as a language to critically investigate algorithmic prediction through design practice? - is addressed through three practice chapters, each proposing one type of diagram as an example of archaeological practice. Each of these chapters is based on one practice-led project, and mobilises slightly different sets of literature. Together they make a broader case for a diagrammatic approach to the archaeology of algorithmic prediction.

In Chapter 3 Case Board I present the case board as a diagrammatic instrument to *remember* the historical genealogies of algorithmic prediction as a diagram of diagrams. While the history of prediction is made of attempts to spatialise the future through statistical techniques, these also form threads or *topoï*, which can be followed and woven together on a case board.

Chapter 4 Traces examines one specific predictive algorithm, the recommendation system on YouTube, that is *observed* through traces. While traces are the fragmented data points used to entrap users through recommendations, they can also be seen as sources of conjectural knowledge. I use this to question modes of mapping algorithmic systems and the claims to transparency associated with them.

Finally, in Chapter 5 Chicanes I move inside the operations of algorithmic predictions as I re-*imagine* a historical genre of publication, the almanac. I use the almanac as a prototype for a cosmic imaginary of data, that survives both in contemporary data analytics and in divinatory practices such as astrology. I use the chicane to characterise the "tricks" performed by the operations of algorithmic prediction, and to question their "sincerity".

Each of these chapters follow a similar internal structure to their presentation. They start by 1) setting up background and situation for the research, and establishing the diagrammatic form under investigation. This is followed by 2) outlining the practice, its tools, and any additional literature that has informed my making, and finally 3) discussing the practice and research, referring back to my research questions via the criteria detailed above: abductive arc, diagrammatic movements, and the publicity of the research.

Finally, I conclude by summarising the qualities of the three diagrams and how they address my research question. I outline the contributions of the research and its limitations before suggesting possibilities for future work.

In the next chapter, I introduce the elements that come together in my diagrammatic language to investigate algorithmic prediction: the computation of predictions as/with spaces made up of data, criticisms of these spaces as prescriptions forever reproducing past behaviours and power relations, and the possibility that computation might be used to open up new, less stable spaces of creativity and speculation.

Chapter 2

Context Review

2.1 Introduction

I have found diagrams to be a generative focus for research and practice because of their highly flexible and "stretchy" nature. In particular they connect to the operational and historical core of computation. Diagrams extend from today's processors all the way to the diagrammatic devices of Ramon Llull (Fidora and Sierra, 2011) in the 14th century, through to Gottfried Wilhelm Leibniz (Gray, 2016) and his mythical links to modern computing (see section 2.4.1). Diagrams were key to the formalisation of logic and semiotics, notably through the work of Charles Peirce in the late 19th century (Gardner, 1958). After World War II, diagrams were instrumental in underpinning the ideology of cybernetics in its various flavours (Veloso, 2014). As the world came to be conceptualised as a computer (Van de Velde, 2003), diagrams were also devices for critical theory pushing back against instrumentalisation and control, for example in the work of Deleuze and Guattari (Knoespel, 2001, as analysed by). It is this 'schizophrenic identity of diagram' (148) that has made it a generative instrument for my work in this thesis. It is at once a device of systematic containment and computation, and a way to generate 'lines of flight' (Deleuze and Guattari, 1987, 135) of speculative escape.

However malleable and elastic, the diagram can still be differentiated from other forms, such as the schema or data-visualisation, along clear lines of demarcation. While these words may be colloquially interchangeable, the key difference is that both schema and visualisation have (or claim) a descriptive relationship to the object they represent. The schema describes through visual synthesis while data-visualisation relies on an 'indexical relationship' (Cubitt, 2015). Diagrams, on the other hand, do not describe objects but relations. They are 'a kind of icon that resembles not the object itself but the relations necessary for generating an object' (Munster, 2013, 24, drawing from C.S. Peirce).

In this chapter, I review diagrams and algorithmic prediction from three perspectives—data and computation, control, and speculation. In diagrammatic fashion, the relations between these registers form the necessary connections extending through the body of research and practice in this thesis.

This chapter is not an exhaustive literature review on diagrams, algorithmic prediction, or their overlap. It does not assemble a limited set of texts to review them in detail and reference them throughout the thesis. Rather, it is a broad scoping exercise that aims to set the backdrop for my research, in response to research questions of an interdisciplinary nature. This has led me to draw from a wide range of texts, from science-studies canon (e.g. Hacking, 1990), to graphic design (e.g. Drucker, 2014b; Bertin, 1967), media studies (e.g. Gitelman, 2013; Mackenzie, 2017), and the practical literature on machine learning (e.g. Grus, 2015). This chapter collects these texts, and connects them to three main polarities, opening a space between them for research and practice.

I start with diagrams in data science and machine learning as I detail some of the ways through which data are turned into predictions via processes of spatialisation. I describe the epistemology afforded by these diagrams as a positivist view of data and algorithms as neutral, objective processes that observe and predict from an outside position. I examine how design incorporates, implements, and propagates the positivist imaginary of algorithmic prediction.

In the remaining two sections, I examine how diagrams can be seen as 'oscillating' (Leeb, 2017) between apparatus of control that enclose and pin down, and instruments for speculation that unsettle and open up.

I first turn to describing data and predictive algorithms as diagrams of control. Here data are not considered as a neutral "raw" material but as a frame that, like a photograph, puts some things into view and crops out others. Furthermore, by using past data as a coordinate system for predictions, the operations of algorithmic prediction are inherently conservative; they are *prescription* rather than *prediction* machines. I conclude this section by aligning my position as a researcher with the figure of the archaeologist of algorithms that spans both academic research and creative practice. Here a practice-led *excavation* of the diagrams of algorithmic prediction is used to engage critically with technological artefacts, to relate them to their social, political, and cultural contexts.

I then turn to uses of diagrams as generative/speculative devices. First, combinatorial diagrams go way back in the history of algorithmic processes. While they still subscribe to the positivist claims to universal computation, their use by artists highlight their generative potential. I then turn to uses of diagrams in speculative design, starting with what has become the canonical *Futures Cone* (Candy, 2010, see fig.2.25). Here diagrams are meant to support the imagination in radically shifting or re-configuring the coordinates of what is considered possible, however these practices do not acknowledge or position themselves with regard to the diagrammatic workings of algorithmic prediction. Finally, I review practices that aim to reclaim computation as a tool for speculation.

2.2 Data Diagrams

In this section I summarise the algorithmic processes that "extract" predictions out of data [fig. 2.1]. This is far from an exhaustive review, the complexity of the mathematical techniques discussed here are beyond the scope of this thesis,[1]and are under intensive and constant development. I simply aim to establish the basic diagrammatic principles at play in algorithmic prediction. I draw the technical aspects from the practice-oriented literature on data-science and machine learning (O'Neil and Schutt, 2013; Grus, 2015; VanderPlas, 2016; Chollet, 2018), and from my own experience following an *Introduction to Machine Learning* course as part of this research (by Malone and Thrun, 2015).

[1]The extent of my immersion in the complexity of algorithmic prediction is centred on coding practice. This means that I limit my research and practice to code, not mathematical formulas.

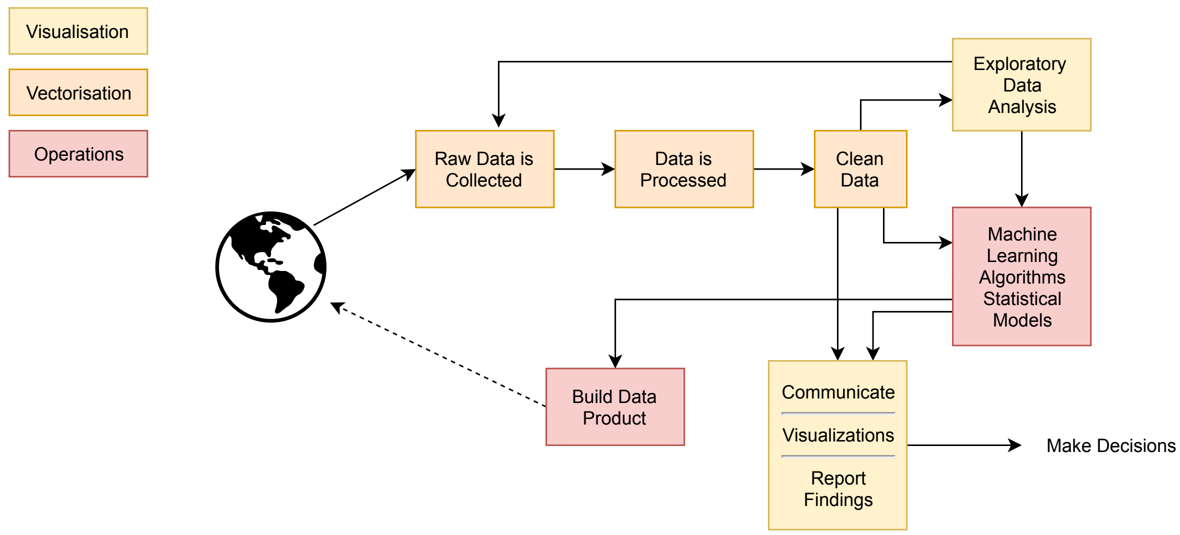


Figure 2.1: *The data science process* redrawn from O'Neil and Schutt (2013, 41) with added colour-coding relating to the parts in this section 2.2.

2.2.1 Vectorisation and operations

The most basic form of data storage and display is the table [fig. 2.2]. While this is 'often overlooked' (Drucker, 2014b, 70), tables are diagrams, they are the 'rationalization of a surface' (71) through alignment and ordering. From the very start, then, data exist as and through diagrams, turning inert white space into an 'active element supporting crucial tasks of differentiation' (84). The use of tables is ubiquitous in algorithmic prediction. Data are stored, combined, and/or transformed through them at various stages of the prediction process. If tables are omnipresent—the one common format in the wide range of datasets and techniques—they are only a basic form of data spatialisation that gets 'kaleidoscopically transmuted' in practices such as machine learning (Mackenzie, 2017, 58). These mutations are the pre-requisite for data to be operationalised in predictive operations. They can be characterised as turning data into a 'feature space where a notion of "distance" make sense.' (O'Neil and Schutt, 2013, 81).

This spatialisation of data as mathematical shapes is called *vectorisation*. The key shift from the rows and columns of the table is that data are turned into spatial mathematical objects with *dimensions*—vector: 1 dimension, matrix: 2 dimensions, tensor: n dimensions [fig. 2.3]. As data

sepal length (cm)	sepal width (cm)	petal length (cm)	petal width (cm)	species
5.0	3.0	1.6	0.2	Iris-setosa
6.3	2.5	5.0	1.9	Iris-virginica
4.6	3.6	1.0	0.2	Iris-setosa
5.6	2.7	4.2	1.3	Iris-versicolor
7.7	2.8	6.7	2.0	Iris-virginica
6.2	3.4	5.4	2.3	Iris-virginica
7.1	3.0	5.9	2.1	Iris-virginica
5.2	3.4	1.4	0.2	Iris-setosa
5.5	4.2	1.4	0.2	Iris-setosa
...	...	...	...	...

Figure 2.2: Excerpt from the Iris dataset, showing measurements of three species of iris flowers (Fisher, 1988). This dataset is commonly used for demonstration purposes in data-science and machine learning, for more on the historical context of its creation and use see section 3.2.2.

$$\begin{bmatrix}1\\2\end{bmatrix} \qquad \begin{bmatrix}1 & 2\\3 & 4\end{bmatrix} \qquad \begin{bmatrix}\begin{bmatrix}1 & 2\end{bmatrix} & \begin{bmatrix}3 & 2\end{bmatrix}\\ \begin{bmatrix}1 & 7\end{bmatrix} & \begin{bmatrix}5 & 4\end{bmatrix}\end{bmatrix}$$

(a) (b) (c)

Figure 2.3: a) vector, b) matrix, and c) tensor redrawn from Goodfellow et al. (2018).

are 'mapped onto a coordinate volume' (Mackenzie, 2017, 51) their original provenance, differences, and meanings are all collapsed in the formation of a "purely" mathematical construct. Measures of different things, expressed in different units (e.g. centimetres and flower species), and coming from different places, all merge through vectorisation to produce a space where new kinds of differentiations can take place: measures of distance between vectors. One simple example is the Nearest Neighbours algorithm. As its name indicates, it derives meaning from measures of proximity in vector space. The example shown in figure 2.4 comes from an implementation of this for music streaming company Spotify (Bernhardsson, 2013), suggesting that they use this technique as part of their predictions of what music a user "may like," based on proximity with other users in a multi-dimensional space defined by the data points held on each user. We continue to follow the Iris dataset as an example through the data → vectorisation → operations process in figure 2.6 that shows a Nearest Neighbours classifier separating the flower species.

In my own encounters with vector space, the mathematics of vectors and their associated operations (i.e. linear algebra) are abstracted behind code

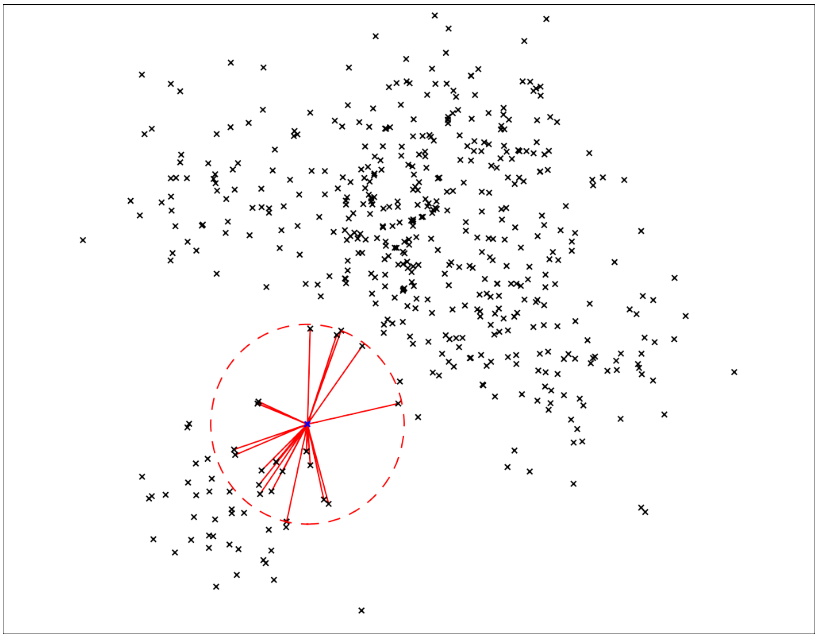


Figure 2.4: Visualisation of ‘approximate nearest neighbours’ by Bernhardsson (2015) author of the *Annoy* package for Spotify (Bernhardsson, 2013). (used with permission)

written in the Python language. The vectorisation of data is explained briefly in some of the practice-oriented textbooks (e.g. O’Neil and Schutt, 2013), but is largely taken for granted as a prerequisite for any data-science or machine learning to take place. Figure 2.5 shows the data being loaded into Python[2] as an array of vectors.[3] The last line of the figure demonstrates the spatial characteristic of vectorised data, as the `.shape` of the array is displayed.

Once data are vectorised, the *operations* of prediction can begin. These can be summarised as drawing lines, planes, or hyper-planes[4] through vector space. If the variable x being predicted is “continuous” such as a price, the surface represents the *regression*, or the statistical relationship of all other variables with x. For any new data point where x is not known, it can be predicted using the regression function. If the variable is “discrete”, for example a species of flower, the surface represents the boundaries for

[2]Iris is part of the test data included with Scikit-Learn, making this example even easier. The data could have easily been read from file with a similar amount of code (1 line).

[3]In Python, N-dimensional arrays are handled by the *Numpy* package (van der Walt et al., 2011).

[4]Hyper-planes are multidimensional surfaces.

```
>>> from sklearn import datasets
>>> iris = datasets.load_iris()
>>> iris.data[:10]
array([[5.1, 3.5, 1.4, 0.2],
       [4.9, 3. , 1.4, 0.2],
       [4.7, 3.2, 1.3, 0.2],
       [4.6, 3.1, 1.5, 0.2],
       [5. , 3.6, 1.4, 0.2],
       [5.4, 3.9, 1.7, 0.4],
       [4.6, 3.4, 1.4, 0.3],
       [5. , 3.4, 1.5, 0.2],
       [4.4, 2.9, 1.4, 0.2],
       [4.9, 3.1, 1.5, 0.1]])
>>> iris.data.shape
(150, 4)
```

Figure 2.5: The Iris dataset in Python (excerpt of the 10 first data points). Data are loaded from Scikit Learn and printed as an array of vectors (Numpy array). The *shape* of the array is (150,4).

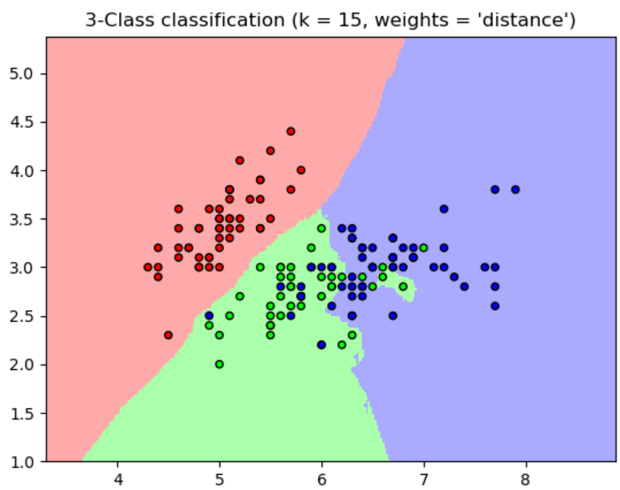


Figure 2.6: Nearest Neighbour classification on the Iris dataset (scikit-learn).

classification. For any new data point, the class it belongs to can be predicted from its position in the vector space. The field of machine learning is dedicated to the development of methods to summarise, separate, and transform vector spaces in ways that are too sophisticated and numerous to cover here. Figure 2.7 shows the range of classification techniques available in Scikit-Learn (Pedregosa et al., 2011). They are all diagrammatic operations that delineate vector space.

The goal of these operations is finding an optimal *fit* of a model through vector space. This is a negotiation between the specificity of the "training" data and producing a model that is generalisable to other cases. This process was described by Hayashi (1996)—in his definition of the field of data science—as a 'dynamic movement of both simplification or conceptualisation and diversification', an iterative synthesis of individual observation (data) into general insights (models). *Fitting* models is a geometrical exercise, as demonstrated by figure 2.8 it is literally about drawing lines through and around data points.

These considerations are meant to summarise the diagrammatic and spatial underpinnings of algorithmic prediction. They are not exhaustive by far, diagrams proliferate in a variety of ways in this domain. For example, Decision Trees algorithms are among the more simple ways to delineate vector space through a series of straight lines. Taken together as *forests* however these trees become themselves part of a vector space where they are ranked for their fit to the data.

Another way in which diagrammatic complexity proliferates is the field of neural networks, or connectionist models, that are the focus of much of the current research, discourse, and claims made around algorithmic prediction. The diagrammatic characteristics I described are also found in neural networks, but dramatically amplified as they are able to deal with vector spaces of very high dimensions. This has made them popular in uses such as image recognition, where each pixel is treated as a dimension. Like other techniques, neural networks produce 'statistico-topological construct[s]' (Pasquinelli, 2017). They add a diagrammatic element with the fact that their *architecture* is made up of layers that fragment the data, these can be combined in a wide range of different configurations [see fig. 2.9]. Chollet (2018, 44-45) gives a 'geometric interpretation' of neural networks:

> ...[which] consist entirely of chains of tensor operations and that all of these tensor operations are just geometric transformations of the input data. It follows that you can

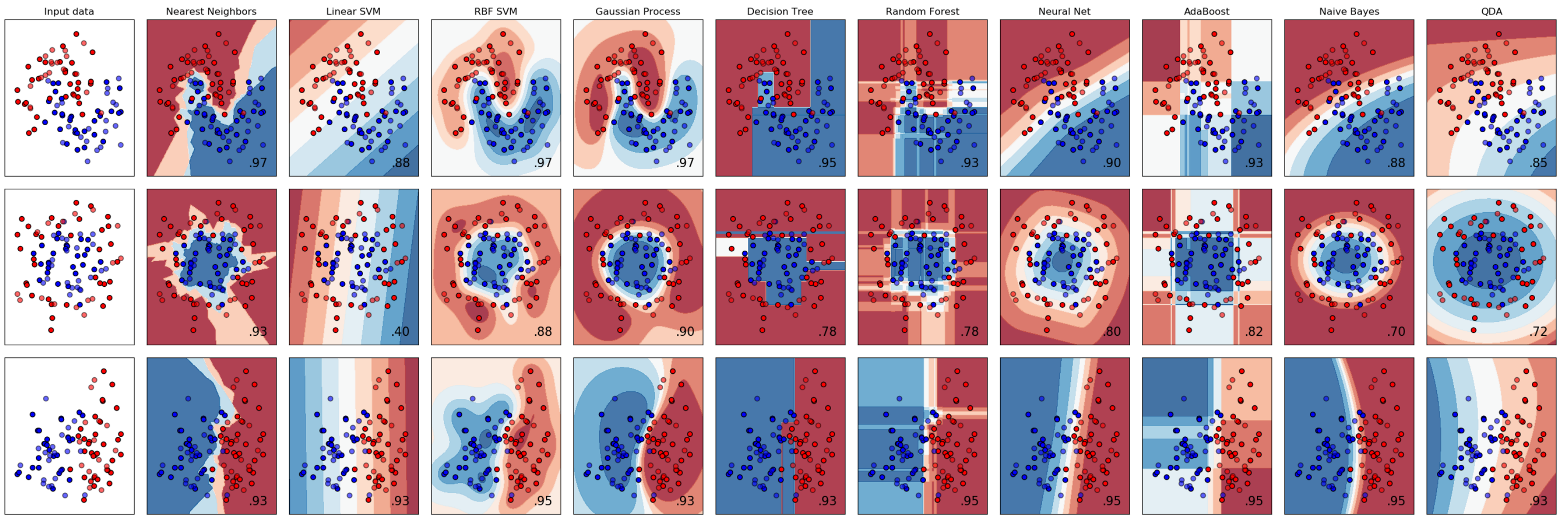


Figure 2.7: scikit-learn classifier comparison (Pedregosa et al., 2011).

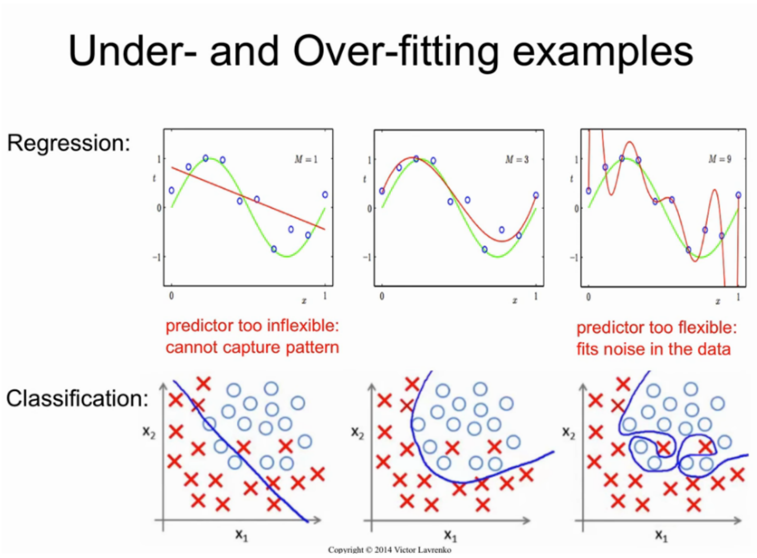


Figure 2.8: Examples of under (left) and over (right) fitting in regression and classification (Lavrenko, 2015).

> interpret a neural network as a very complex geometric transformation in a high-dimensional space, implemented via a long series of simple steps. [...] Uncrumpling paper balls is what machine learning is about: finding neat representations for complex, highly folded data manifolds.

With the diagrammatic foundations of algorithmic prediction established, I now turn to reviewing ways of *seeing* data through visualisations.

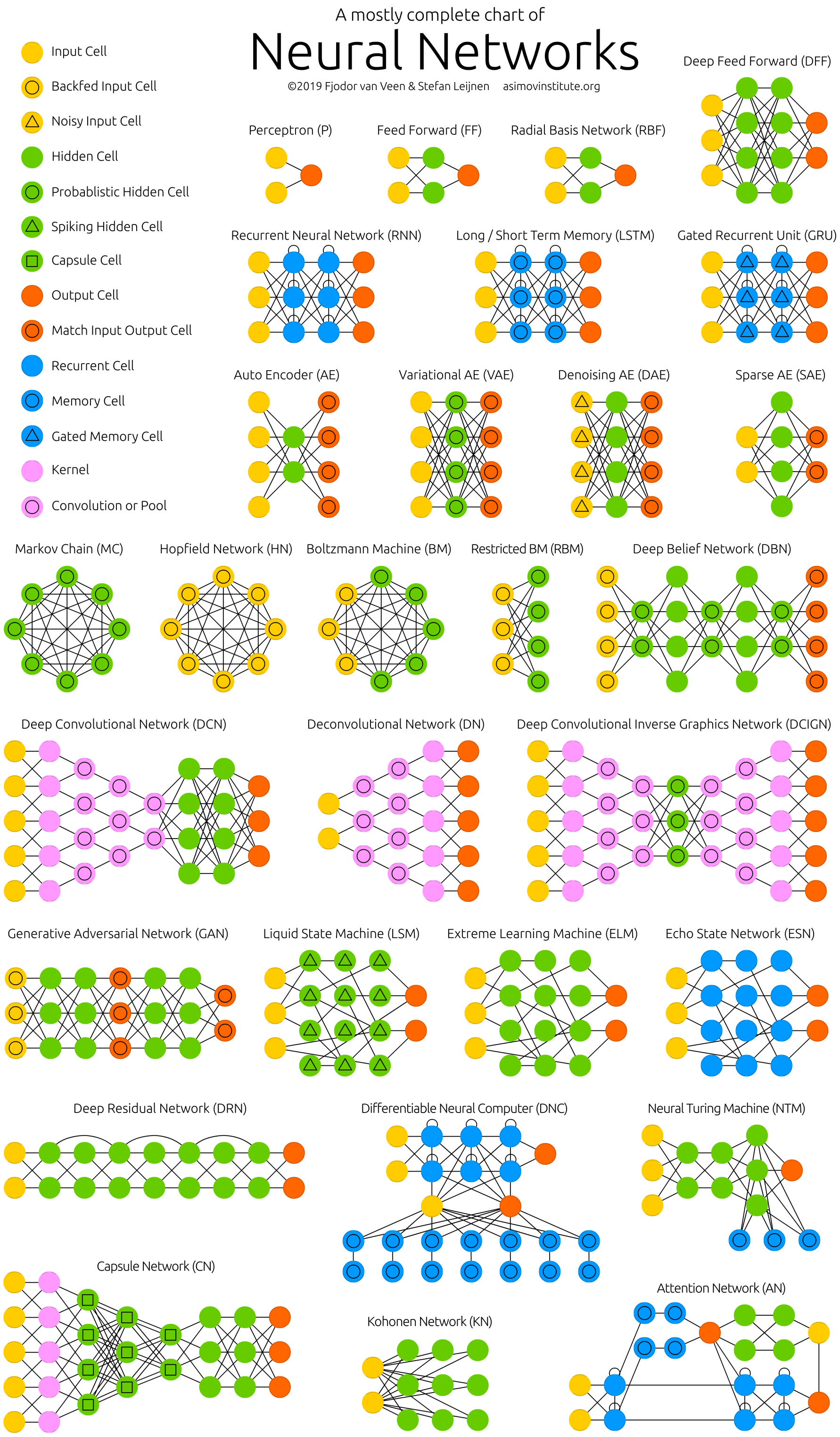


Figure 2.9: *Neural Network Zoo* (van Veen, 2016). (used with permission)

2.2.2 Visualisation

Data visualisation is a slight detour from the purely diagrammatic paths between the table and vector space I described in the previous section. The diagrams I described above in section 2.2.1 (e.g. fig.2.3) are mathematical constructs that are beyond representation. Multi-dimensional vector spaces are, by definition, impossible to represent as forms that human cognition can recognise (i.e. in two or three dimensions). *Seeing* is, however, a crucial part of data practices, from the 'graphical detective work' (Tukey, 1977, 1) of exploratory data analysis, to the development of models, and the communication of predictions and "insights." Data visualisations are not technically diagrams, they follow a logic of representation as they map data visually to a two-dimensional surface. While visualisations claim a 1:1 relationship between data and the objects they describe (see positivism in the next section 2.2.3), diagrams are generative structures based on relations (Drucker, 2014a). However, they offer glimpses into vector space—a view that is always incomplete through flattening, projecting, or intersecting (Schmitt, 2020)—and as such they play a key role in this research.

Tables are efficient forms for storage and computation but they soon boggle human cognitive capabilities (especially when data are "big"). Graphical forms such as plots, graphs, and curves are used to *see* data, to render it intelligible. The display of data, their mapping onto image surfaces, was established and developed through their joint history with media production (e.g. the printing press, computer displays, see Friendly and Denis, 2005; Friendly, 2006, 2008). Conventions about how numerical values should relate to visual characteristics such as position, size, or colour, are the "language" of data, formulated as a visual semiotics by Bertin (1967), a computational grammar by Wilkinson (1999), or a more general set of guidelines by Tufte (2001).

These visual languages of data are routinely used today in graphs and charts (e.g. bar chart, line chart, scatter-plot) that permeate literature from scientific research to mainstream newspapers. A wide range of software is available to produce visualisations, from business software (e.g. Microsoft Excel, Tableau) to programming languages. In the Python language alone, figure 2.11 shows the proliferation of code libraries each catering to specific use cases and offering variations on the visual "grammar" of data. I make extensive use of visualisation tools in this research, for example *Matplotlib* (Hunter, 2007), *Networkx* (Hagberg et al., 2008), and *D3* in JavaScript (Bostock et al., 2011). This latter example is

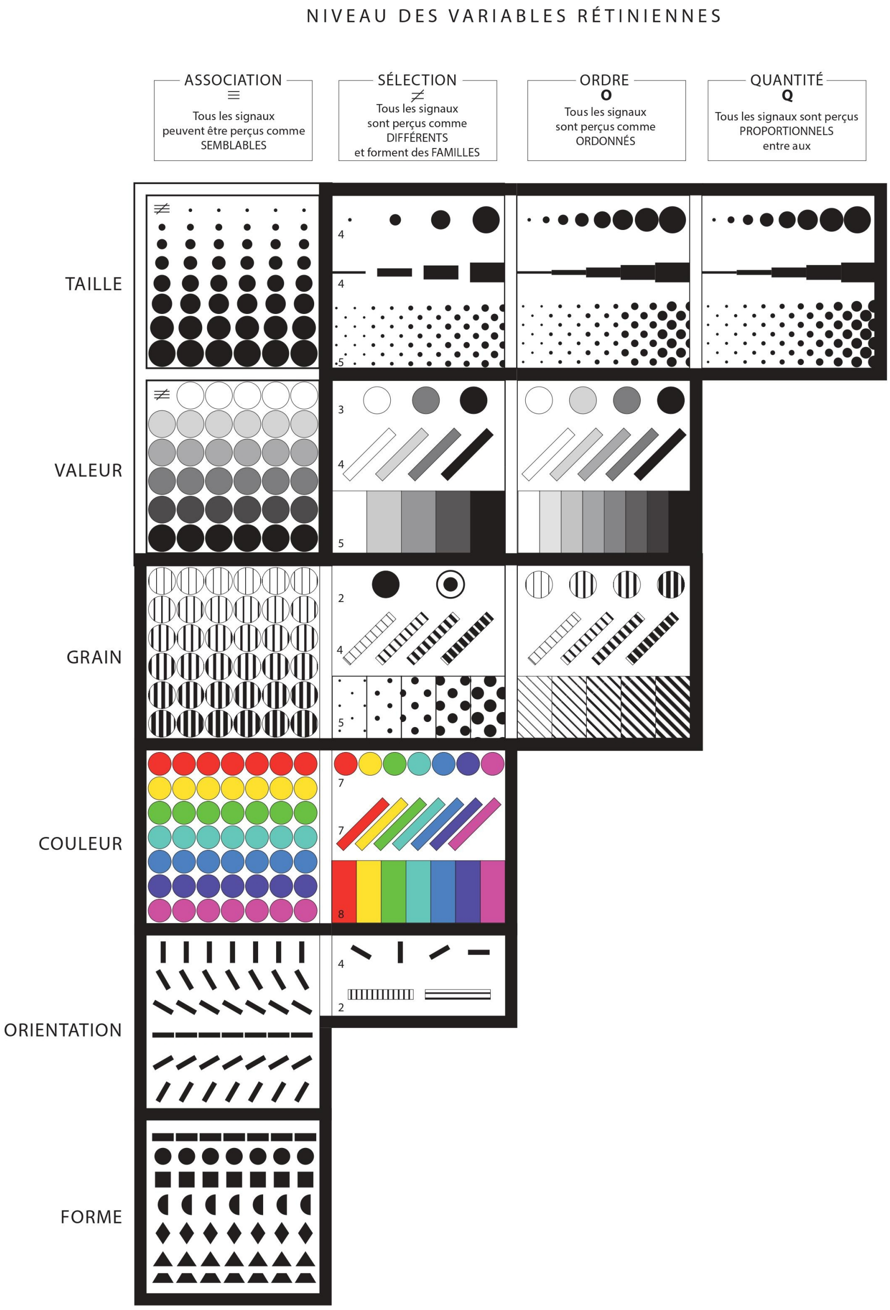


Figure 2.10: Visual variables in Bertin's *Sémiologie Graphique* (Bertin, 1967, 96). (copyright Éditions EHESS, used with permission)

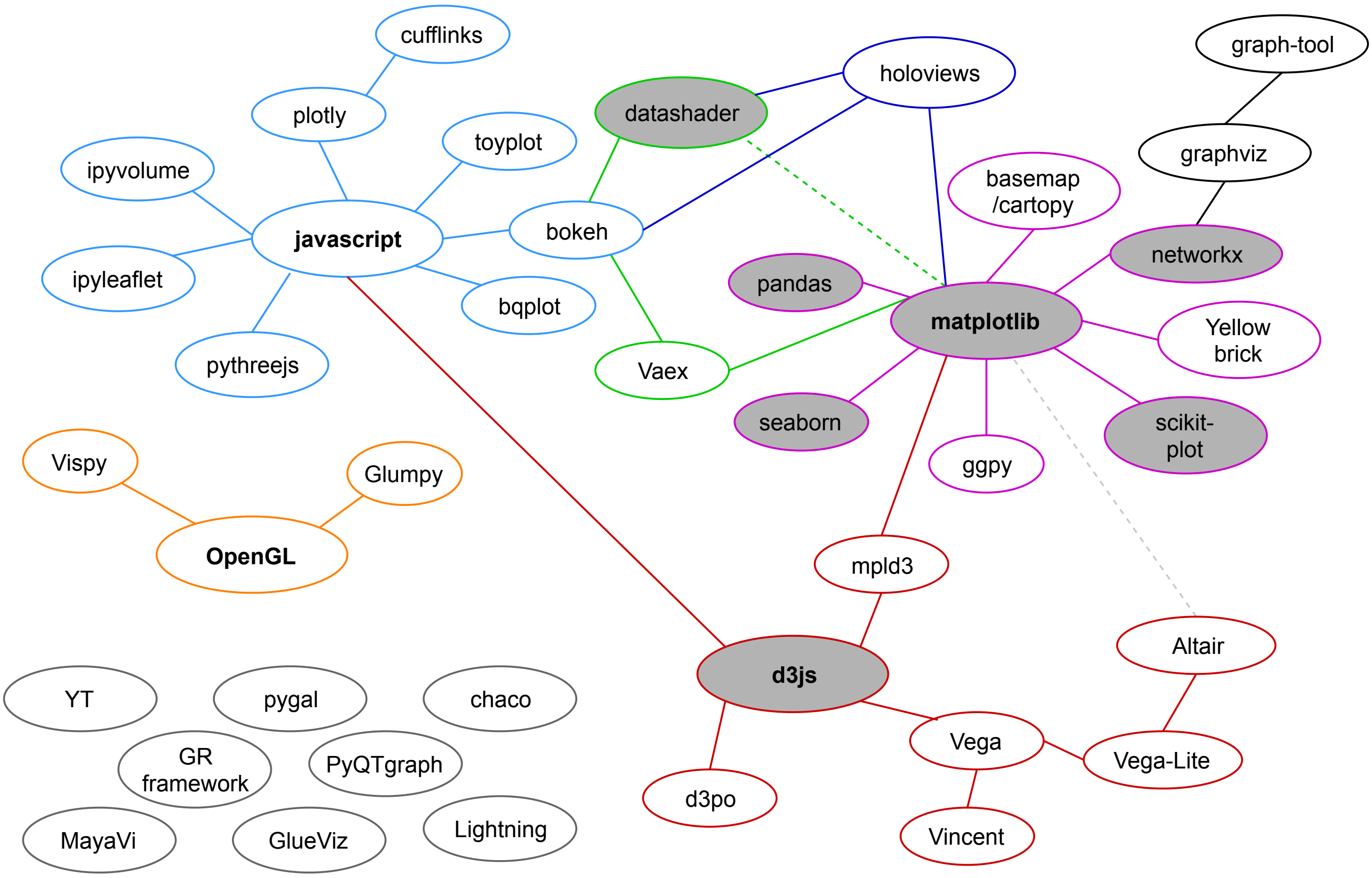


Figure 2.11: The Python visualisation landscape. Redrawn from VanderPlas (2017) with software used in this research shaded in grey.

the most widely used as it makes use of web technologies to produce visualisations, it 'binds' data to HTML and SVG elements effectively constructing and manipulating web pages as data visualisations.

Coming back to the dimensionality of vectorised data discussed in the previous section, data visualisation faces the challenge of how to represent multi-dimensional space. This has been addressed in a variety of ways ranging from parallel lines (Wegman, 1990), to interactive rotating displays (Fisherkeller et al., 1988), to cartoonish faces (Chernoff, 1973). Figure 2.12 shows the example of the Iris dataset visualised as a matrix of scatterplots displaying pairs of variables. This method, suitable only for a small number of dimensions, is one of many ways to "intersect" N-dimensional space to make it visible through permutations. For higher dimensional spaces, much more sophisticated methods are employed. Figure 2.13 shows

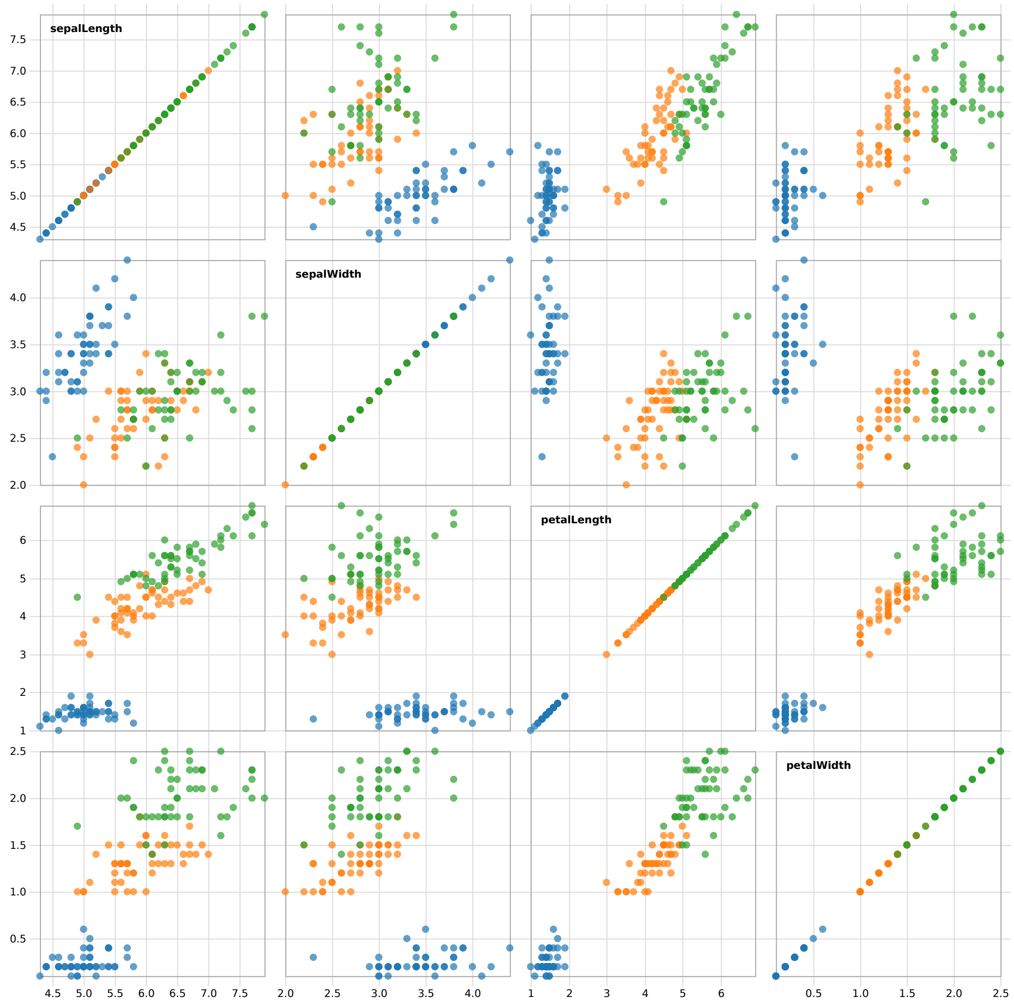


Figure 2.12: The Iris dataset visualised as a scatter-plot matrix using D3 (Bostock, 2018). Note how some flower species are linearly separable in some of the variable pairs.

one of these methods, t-SNE (van der Maaten and Hinton, 2008) displaying the high-dimensional space of the *MNIST database of handwritten digits*[5] (LeCun et al., 1998). In the high dimensional space of pixel data, the algorithm shows the images representing the same digit (and labelled as such) as clusters of proximity in the vector space.

My aim here is not to exhaustively review the vast and developing field of data visualisation. Rather, I have established a few key references and highlighted some lines of demarcation between the high-dimensional spaces in which the diagrammatic operations of algorithmic prediction take place and attempts to represent these data through visual languages. In the next section, I turn to another aspect of the language of data diagrams, the ontological claims that are made about and with them.

[5]The MNIST dataset contains 70,000 handwritten digits, stored as images of 28x28 pixels. If each pixel is considered a dimension, the vector space produced has 784 dimensions.

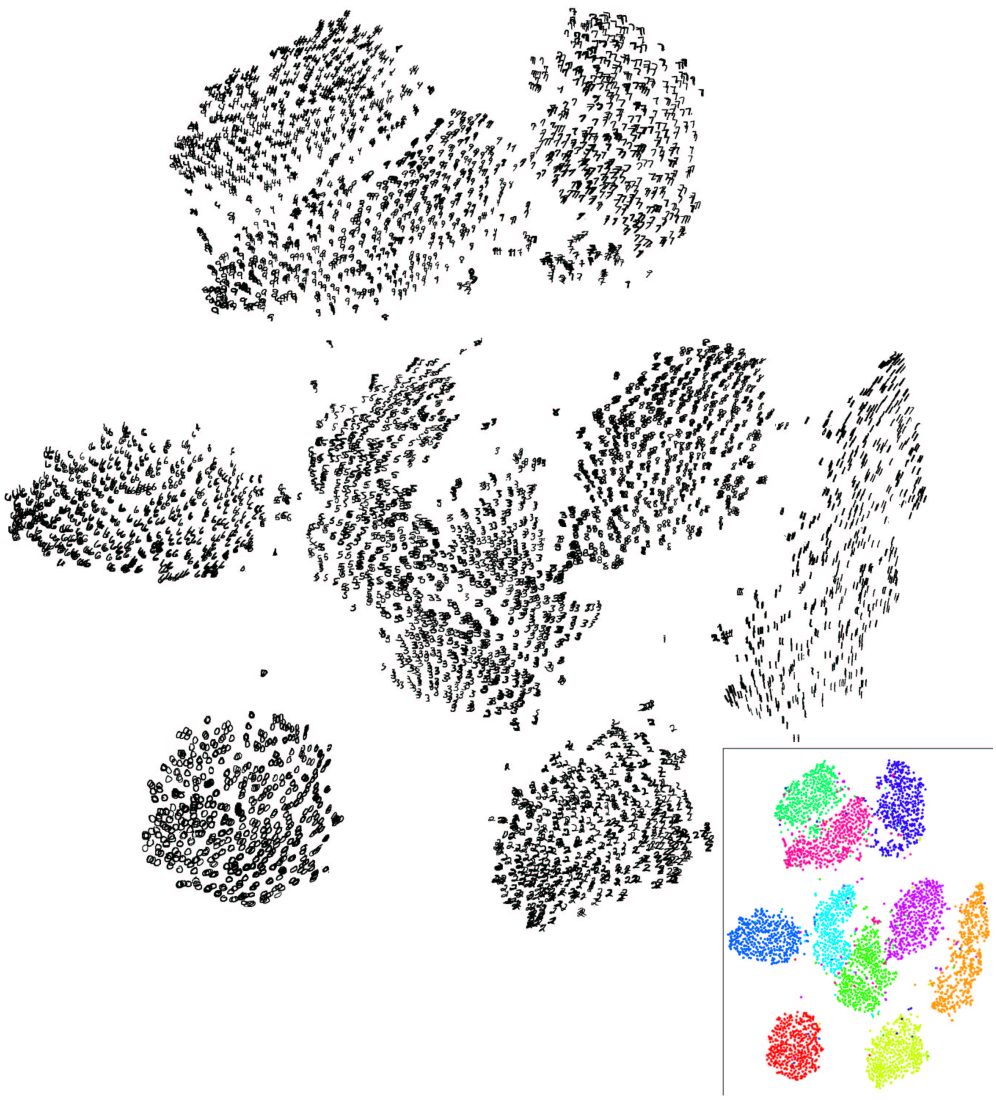


Figure 2.13: 'Visualization of 6,000 digits from the MNIST data set produced by the random walk version of t-SNE (employing all 60,000 digit images).' (van der Maaten and Hinton, 2008, 18).

2.2.3 Predictive positivism

An interesting feature of the data science process diagram at the top of this section [fig. 2.1] is that the "pipeline" from data to prediction is presented as external to the world. This suggests that the sophisticated methods of vectorisation, prediction and visualisation described above (sections 2.2.1 and 4.3.2) are neutrally and objectively observing a reality that exists independently. Central to this view—and to the data imaginary as analysed by Beer (2019)—is the notion that data contain truth. Intricate processes of mediation (gathering, vectorising, modelling, visualising data) are only intermediaries that 'reveal' (Beer, 2019, 28) this knowledge that is otherwise buried and in-accessible. This position is described by Mosco (2014, 206) as *digital positivism*, the 'specific belief that the data, suitably circumscribed by quantity, correlation, and algorithm, will, in fact, speak to us.' I would argue for my case in this research that this 'speaking' in fact happens visually.

This positivist imaginary is relayed by metaphors[6] of "mining" knowledge out of data that 'already knows the future' (Verteego, 2017). The layered complexity of vector spaces and diagrammatic operations are seen as ways of accessing 'a hidden mathematical order that is ontologically superior to the one available to our everyday senses' (McQuillan, 2017). This is found in the discourse propping up the data imaginary with claims to reveal 'who we are when we think no one is looking' by OkCupid founder Rudder (2015), or the trope that 'Facebook knows you better than you know yourself' (Evans, 2015; Lapowsky, 2015). In this view, the elaborate mediation of data via vector space and algorithmic operations is seen as a way to get closer to the "ground truth" of a pre-existing geometrical order. According to McQuillan this is 'an echo of the neo-platonism that informed early modern science in the work of Copernicus and Galileo' that is generalised through computation.

Digital positivism resonates with historical views of data and knowledge, promising that with more data comes more truth. This promise was set out by mathematician Pierre Simon Laplace in the 19th century.

> An intellect which at a certain moment would know all forces that set nature in motion, and all positions of all items of which nature is composed, if this intellect were also vast enough to submit these data to analysis, it would embrace in a single

[6] For the role of metaphors in shaping imaginaries of data see Stark and Hoffmann (2019)

> formula the movements of the greatest bodies of the universe and those of the tiniest atom; for such an intellect nothing would be uncertain and the future just like the past would be present before its eyes. (Laplace, 1814)

It takes on a new life in the data imaginary which affirms that this superior 'intellect' may, finally, be realised. Domingos (2015) exemplifies this in his quest for a *Master Algorithm* that summarises the positivist claim to universal knowledge through data and algorithmic processes.

> All knowledge—past, present, and future—can be derived from data by a single, universal learning algorithm. (Domingos, 2015, 25)

Two centuries apart, the similarities between these two claims are striking. The digital positivist stance is pushed to its extreme in the famous piece by Anderson (2008), arguing that the 'scientific method is obsolete' and the hypotheses themselves will soon be formulated by data that 'speak for themselves.'

This imaginary of data as a serum of truth is reflected in design practices such as data visualisation. These are seen as a crucial layer in the mediation of a 'tsunami of data' that remains unintelligible unless visualisations are 'properly designed' (Few, 2006, 6). Ideals of 'truth and beauty' (Stefaner) bring an aesthetic dimension to the digital positivist ideal, while also positioning the "proper" designer of data as yet another intermediary. This figure is caught in a double imperative, needing both aesthetic sense to communicate effectively and neutrality to avoid corrupting the truths they are revealing out of the data. They have to negotiate the 'awkward relationship' between scientific claims and 'style and beauty' that can either distract from the data, or give it an 'impression of truth' (McInerny et al., 2014, 149). As major newspapers now include data visualisation departments (such as *Guardian Graphics*, *The New York Times*' *The Upshot*), Cairo (2016) describes *The Truthful Art* of visualisations that should strive to be at once 'truthful, functional, beautiful, insightful, and enlightening' (60). All of these qualities are in line with the positivist imaginary of data as a substrate containing truth. They further it by adding an aesthetic dimension; data are seen as beautiful, like a wonder of the natural world (David McCandless, 2012).

Beyond visualisation, design also plays a role in the implementation of algorithmic prediction in digital products. While "user experiences" were traditionally centred on presenting choices to users, they increasingly leverage prediction in the background to provide 'seamless' (Chalmers and Maccoll, 2003) interactions. Once designers have been "caught up" on the new paradigm of machine learning (Hebron, 2016), they can materialise it in digital products and other artefacts. For example the move towards "anticipatory design" on e-commerce platforms relieves users reportedly suffering from 'decision fatigue' and delegates choice to automated systems (Shapiro, 2015). The logical conclusion for "anticipatory design" is 'negative latency' where questions are answered before they have even been asked (Nguyen and Andreessen Horowitz, 2015). Design is tasked with presenting the output of algorithmic systems, sometimes removing its own role in hierarchising information. In other cases—such as the red "risk" flags on the payment platform PayPal discussed by Leese (2016)—the design of "user experiences" introduces the 'affective triggers' that activate predictions, 'fold[ing them] back onto the present' (Leese, 2016, 144).

This begins to suggest the performative nature of algorithmic prediction, in which user-interface and user-experience design partake I return to this in the next section 2.3. As design implements "anticipatory" non-interfaces, it effectively materialises the imaginary of algorithmic prediction and facilitates the enforcement of algorithmic prescriptions. The range of digital personal assistants in various operating systems (e.g. Apple's *Siri*, Microsoft's *Cortana*) exemplifies the former. These software agents are as seamless as possible, offering only a voice interface, and conjure science fictional imaginaries of artificial intelligence through personified characters.[7]

In this section 2.2 I have described the diagrams of algorithmic prediction, as well as the positivist epistemology that surrounds them. In this aspect of the research I take data diagrams as a material for critical study, but position myself outside of the visualisation, implementation, or promotion assigned to designers. In the next sections (2.3 and 2.4), I cover how diagrams may be used towards this end, first by relating data diagrams to critiques of data and algorithms, and second by reviewing how data diagrams are being reclaimed as instruments for speculation, foregrounding imagination and interpretation.

[7] In the case of *Cortnana*, "she" is named and designed after a character in the video game *Halo*. For more on digital assistants, especially their gendering, see Søndergaard and Hansen (2018); Phan (2017).

2.3 Critical Diagrams

I now turn to critical views of the data diagrams and digital positivism discussed in the previous section. I draw from critical studies of data and algorithms, media archaeology, and art and design practices to discuss data diagrams as 'centers of power knowledge' (Mackenzie, 2017, 17), sites where 'power relations are produced through relationships of strength' (Leeb, 2017, 32).

The critical literature on the objectivity of data debunks the notion that they provide a mode of seeing the world through a neutral scientific gaze. In addition, the operations of algorithmic systems might not be characterised as predictions, but as prescriptions that produce futures. Finally, I position this research in line with the figure of the archaeologist of algorithms—bringing critical views and speculative re-configurations in conversation with the materials of algorithmic systems.

2.3.1 Framed and framing

The positivist claims of the data imaginary rest on the assumption that data are a "raw" material. Gitelman (2013) famously counters this view that 'data are transparent, that information is self-evident, the fundamental stuff of truth itself' (2) and points instead to the processes of data collection, transformation, storage, processing, and computation, as a sequence of interpretive steps that produce knowledge and authority out of data rather than "reveal" what was already there. The data imaginary, in this view, starts with the fact that 'data need to be imagined *as* data to exist and function as such' (3).

This critical view of the data imaginary establishes a parallel between modes of seeing through data, and the 'mechanical objectivity' (Daston and Galison, 2007, 115) afforded to scientific imaging techniques. According to Lohr (2015, 18) 'Big-data technology is the digital-age equivalent of the telescope or the microscope.' This echoes claims to mechanical objectivity made around photography as a replacement to drawing as part of scientific enquiry. As Daston and Galison (2007) show, the machines were supposed to protect scientists from their own interpretive biases and produce more truthful images. However this is contested ground and objectivity—as its definition itself shifts through time and places—was never fully reached. In the more contemporary setting of the "big data" era, mechanical objectivity translates as an 'ideal of framelessness' (Andrejevic, 2018, 256) that continuously fails to deliver on its promise of un-mediated access to "true" knowledge. Like a photograph, Gitelman (2013, 5) argues, data should always be seen as 'framed and framing.' The data pipeline diagram in the previous section [fig. 2.1] is not a neutral camera providing a 'view from nowhere' (Jurgenson, 2014) but very much part of the mess on the ground, embedded in systems of knowledge and power, and informed by interpretation at every step.[8]

If seeing through data is framed and framing, what might this frame be focused on, or cropping out? This might also be asked as: what gets counted, and gets to count, in data? This question is at the core of the field of critical data studies, in particular its origins in geography and questions around what gets included or excluded in geographical data/maps: 'What is quantified, stored, and sorted? What is discarded?' (Dalton and Thatcher, 2014).[9] According to Drucker (2014b, 87) graphical forms such as the table, however generalisable, always carry the traces of the original purpose for which they were created. With much of these forms originating in commerce and administrative tasks, one way of seeing the "framing" of data is that it "sees" and re-makes everything as either business or bureaucracy.

The vectors discussed in the previous section do not just exist, they have to be produced and have economic value. Their ownership, according to Wark (2004, 2015), re-structures society as the *vectoralists* succeed landlords and industry capitalists as the dominant class [see fig. 2.14]. This is easy to see at play in the current landscape where the biggest datasets, and much of the research on algorithmic prediction, are collected and produced by big platform companies such as Google, Amazon, Facebook, and Microsoft. In Wark's theory of vectoralism the mathematical shapes I discussed in the previous section are not just abstract constructs but form the very structure of power relations in a data-centric digital age (Wark, 2004).

A field where data vectors frame power relations is predictive policing. Here the vectoralist class is tied up with government through procurement contracts of local law enforcement agencies in the US (Harris, 2017).

[8]More recently, local (Loukissas, 2019) and feminist (D'Ignazio and Klein, 2016, *Data Feminism* book forthcoming) approaches to data are countering narratives of mechanical objectivity and outlining strategies for counter-practices.

[9]See also *Counting by Other Means* a session at the 2016 4S/EASST Conference that I took part in during this research, Taylor and Kember (2016) introduce the theme with a review of critical positions on counting and computational counts.

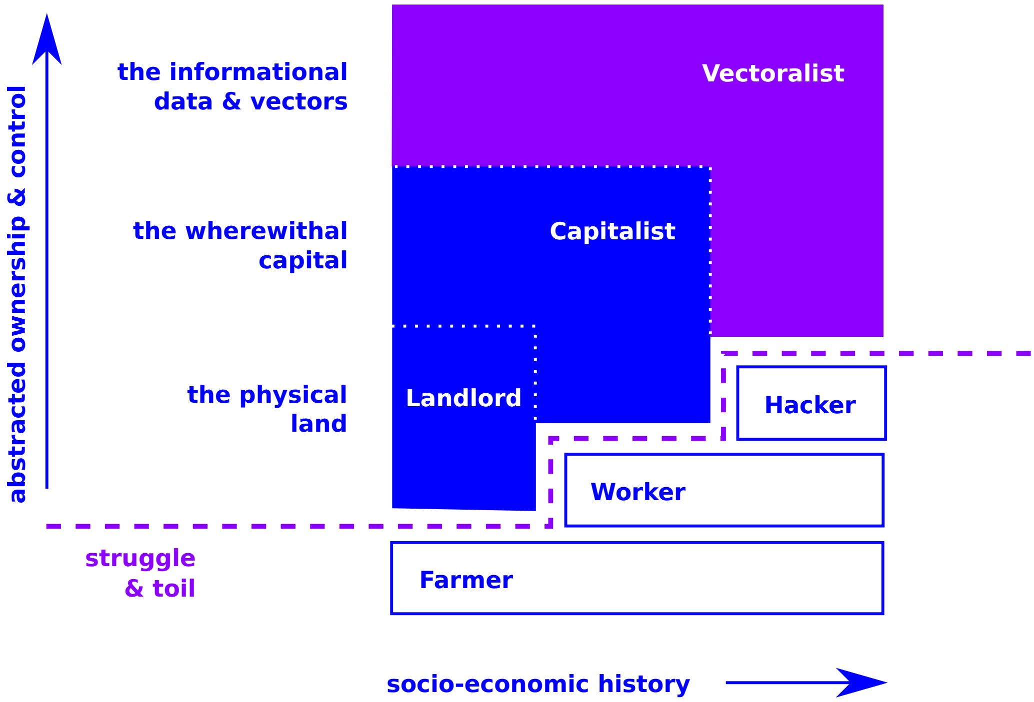


Figure 2.14: Diagram of the 'vectoralist class' as theorised by Wark (2004, 2015) redrawn and coloured from Anthony (2017) CC-BY.

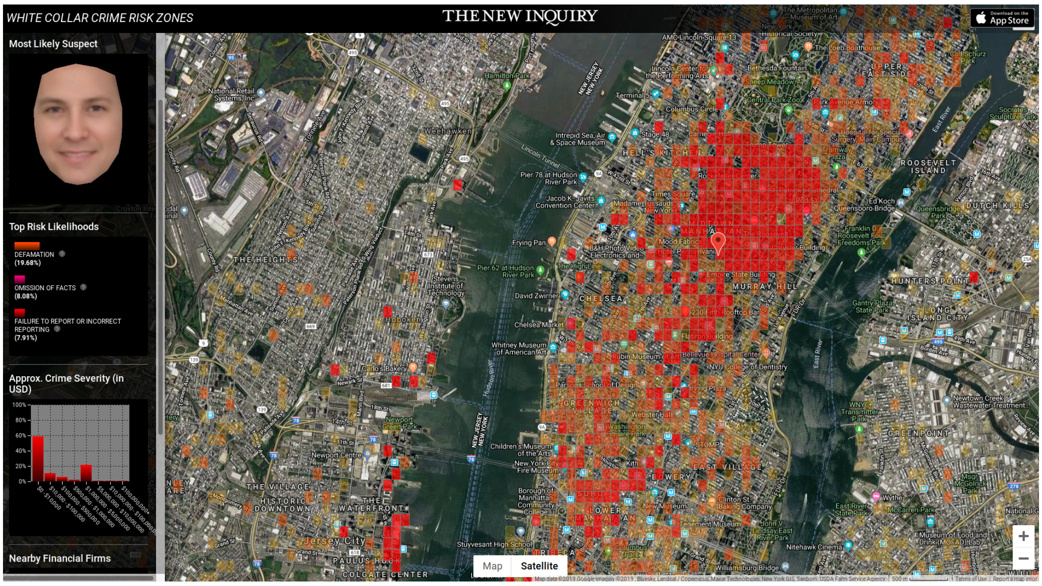


Figure 2.15: *White Collar Crime Risk Zones* web interface showing the heat map of the 'risk surface' where white collar crime is most likely to happen (Clifton et al., 2017). (used with permission)

Predictive policing systems have been widely criticised for reproducing the biases of the 'bad' data they are based on (Richardson et al., 2019). Artists and technologists Clifton et al. (2017) draw attention to the ways predictive policing *frame* low income populations by reversing the target and building a predictor for *White Collar Crime Risk Zones* [fig. 2.15]. Here the language of predictive policing, for example 'risk surfaces' and heat-map visualisations, is preserved but simply aimed at a segment of population that is never targetted by it.

Data can be seen as a frame rather than the 'panoramic' neutral view promoted by the data imaginary. That frame serves to reinforce power relations, and draws new class lines—demarcating who stands to profit, or not, from the ownership of data vectors. As I have shown in the previous section 2.2, vector space only defines the space in which algorithmic prediction operates. In the next section I turn to these operations and how they might be re-characterised as *prescriptions* rather than predictions.

> These diagrammatic schemes are performative. They make the world by structuring our experience of it. (Drucker, 2014b, 74)

2.3.2 Prescription machines

The engineers of one of the largest neural networks in operation (the YouTube recommendation system covered in Chapter 4), describe in their own words the penchant of algorithmic prediction towards reproducing the past.

> Machine learning systems often exhibit an implicit bias towards the past because they are trained to predict future behavior from historical examples. (Covington et al., 2016, 193)

As I have shown in the first section, the vectorisation process constructs spaces out of data which are then transformed through geometrical operations to make predictions about unknown data. It is not hard to see how inherently *conservative* these operations are, past data define the coordinate system in which the future is allowed to happen.

A poster example of this is the investigation of the COMPAS score by investigative journalists Angwin et al. (2016). Here the fate of convicts is based on their COMPAS score that identifies how likely they are to re-offend. It has been shown to be racially biased. Figure 2.16 shows how a matrix of vectors (answers to the COMPAS questionnaire covering many aspects of the suspect's life, including family and friends) turns into a control diagram that determines an individual's future based on how "similar" people have been handled by the court system in the past. In the case of COMPAS this has been shown to be biased against people of colour (Angwin et al., 2016).

Staying with the justic system, Aradau and Blanke (2017) articulate the link between vector space and control. In predictive policing, 'heat maps' of areas where crimes are likely to occur are produced through the analysis of past data. These 'risk surfaces' are spatialisations of crime data in vector space[10] that are folded back onto actual neighbourhood maps and used to dispatch police forces in a self-fulfilling prophecy. This mode of governing the future, of *producing* rather than predicting, has been theorised by Rouvroy (2011) using Foucault's notion of governmentality. While Foucault's notion was developed with 'the stranger, the plague, mental illness, or leprosy' as targets of biopolitics, Rouvroy's algorithmic governmentality shifts to target uncertainty itself. As Rouvroy argues:

> Pre-emption, algorithmic governmentality's mode of operation, consists in making certain things, which are only possibilities,

[10] Aradau and Blanke (2017) use the term 'feature space.'

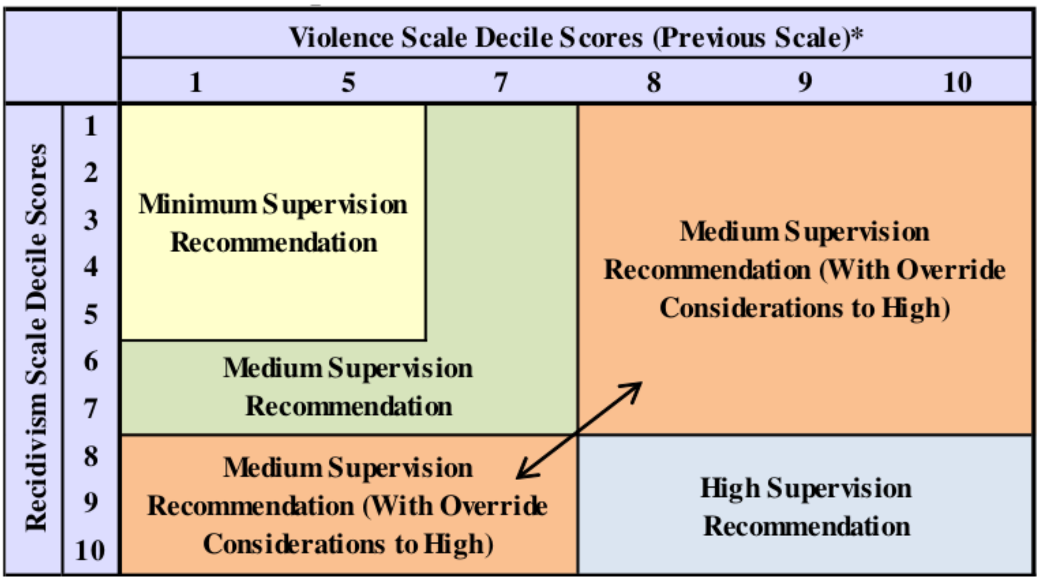


Figure 2.16: COMPAS-Probation Classification Matrix for Supervision-Level Recommendations (NYS Division of Criminal Justice Services, 2012, cited in Angwin et al., 2016).

> either happen for sure or not happen at all. (Rouvroy, 2016, my translation)

This is not, according to Rouvroy, inherent to predictive algorithms themselves, but to the 'governmental rationality' they are serving, one in which 'the focus on contingency and risk minimisation has shadowed most other political goals' (Rouvroy, 2011, 136).

Cheney-Lippold (2011) also focuses on biopower but zooms in on the inference of identity by algorithmic systems. Classification systems, he argues, do not merely sort existing subjects from an external and neutral viewpoint, but actively participate in 'the digital construction of categories of identity' such as race, gender and class. This is not enforced directly through norms, but a 'soft-biopower', a modulating 'self-deforming cast that will continuously change from one moment to the other' (Deleuze, 1992, cited in Cheney-Lippold, 2011). Identity, as a set of vectorised data points, forms a matrix of control that:

> . . . configures life by tailoring its conditions of possibility. Regulation predicts our lives as users by tethering the potential for alternative futures to our previous actions as users based on consumption and research for consumption. (Cheney-Lippold, 2011, 169)

Like the heat-maps of predictive policing this fine-grained control over, and production of, identities happens at the intersection of vector space and literal space. For example, this is seen through biometric data, as artist Blas (2013) exposes in his 'dramatization of the abstract violence of the biometric diagram' [fig. 2.17]. Here the "space" of the human face is considered as the vectors that define everything from a person's unique identity (facial recognition) to personality traits such as their 'criminality' (Wu and Zhang, 2016).[11] But vector spaces are not always folded back onto actual physical space, in fact they most often are not. Another example [fig. 2.18] shows the algorithmic diagram used to predict user personality on Facebook, as summarised visually by Pereira (2019) from a patent document. Here a variety of data points, such as the language of status updates, connections to other users, demographics, and so on are used as the predictors in a simplistic classification of human personality.[12]

The core criticism here is that 'prediction takes down potential' (Munster, 2013, 42). From individual identities to the broader regime of governmentality they are a part of, algorithmic prediction *prescribes* rather than objectively *predicts* futures. While the positivist ideal of the data imaginary affirms that algorithmic systems observe the world from a neutral and objective place, the critiques summarised in this section consider them as playing an integral part in producing it. In other words the intricate process of mediation via data, vector space, and algorithmic operations are ways of shaping and re-shaping the world with the core assumption that the past is a coordinate system in which to predict the future. This entrenches and amplifies the status quo—existing biases, power relations, and discriminations.[13]

In this research, I aim to probe the control diagrams of algorithmic prediction through design practice. What unites the examples above, apart from the recurrent theme of policing and sentencing, is an attention to the materiality and diagrammatic qualities of algorithmic prediction. The white-collar crime predictor, face cages, or the visual mapping of a patent, all dive into algorithmic systems and their diagrams to either expose them or implement alternatives. They begin to suggest ways of prying open control diagrams through creative and critical means. In the next section

[11] This paper by Wu and Zhang (2016) has been widely criticised for reviving phrenology—a long discredited view linking cranial features to personality traits—under the guise of "A.I.". See for example Biddle (2016)

[12] The O.C.E.A.N. model, or "Big Five" is part of a long history of attempts to classify personality traits.

[13] See for example Noble (2018) and Eubanks (2019)

2.3.3 I discuss what emerges in these works as the figure of the archaeologist of algorithms, the practice-based inquiry of media archaeology on the specific topic of algorithmic systems. This starts with the realisation that algorithmic prediction dramatically amplifies the scale and reach of concerns linked to governing the future through data that are far from new.

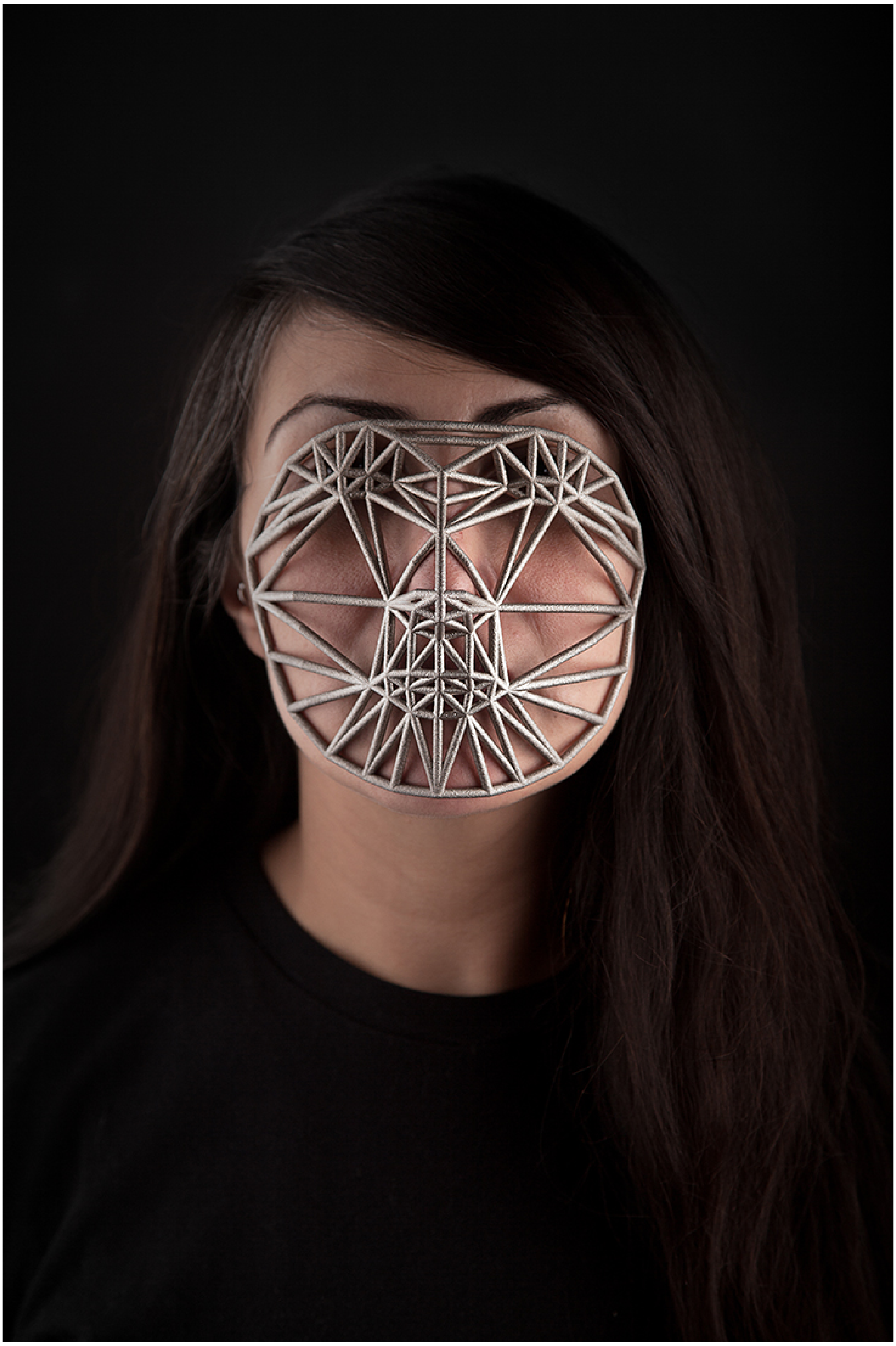


Figure 2.17: Zach Blas, *Face Cage 2*, endurance performance with Elle Mehrmand, 2013. (courtesy of the artist)

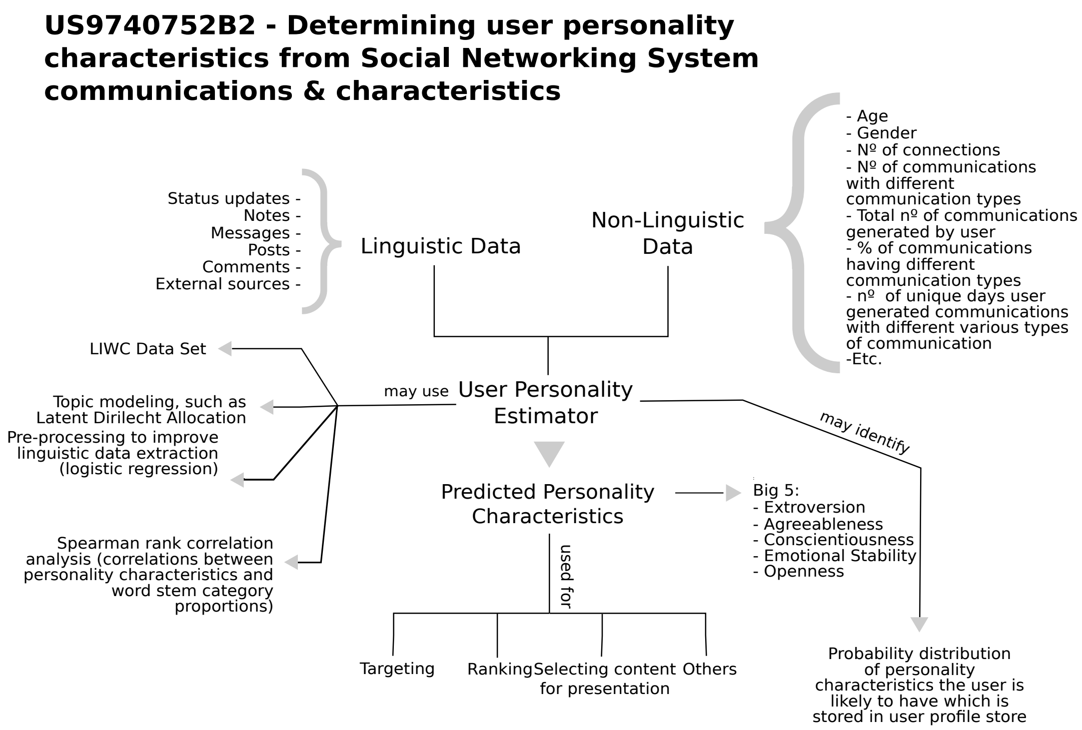


Figure 2.18: Patented system to predict personality by Facebook (Nowak and Eckles, 2017), visually summarised by Pereira (2019, 79) used with permission.

2.3.3 Diagrammatic excavations

In the dominant narrative of the data imaginary, the algorithmic prediction methods presented in the previous section are presented as new. They are part of a paradigm shift, famously framed by Anderson (2008) as the *End of Theory*, and relayed by others (Agrawal et al., 2018; Brynjolfsson and McAfee, 2014; Lohr, 2015) as a tale of unprecedented progress, profits, access to, and production of, knowledge. When history is mentioned, for example by Domingos (2015), it is quite narrowly focused around post-World War II technical advancements leading up to the different flavours of machine learning in operation today.

Media archaeology challenges these tales of new-ness and progress by digging-up, or excavating, forgotten artefacts and bringing them to bear on contemporary narratives. ‘All excavations into the past are meant to elaborate our current situation’ (Parikka, 2012, 6). The spatial logics of algorithmic prediction are often buried behind terms such as “artificial intelligence” that conjure science-fictional imaginaries, or simply inside “black boxes” that evade scrutiny through corporate secrecy or sheer complexity (Burrell, 2016). Media archaeology, especially when it focuses on algorithmic artefacts, proposes to ‘confront this tendency of burying’ (Link, 2016, 11). As I have shown in the previous section, stories of groundbreaking technologies smuggle with them age-old epistemological positions, such as neo-platonism and positivism, that date back centuries. Excavating the past is, therefore, key to gaining purchase on present systems and to counter some of the narratives told with, and about, them.

Media archaeologists, whether theorists, researchers, practitioners, artists, or often some combination of these, go into the historical and technical depths of algorithmic systems. From there, they (in this case Bardini et al., 2016) state a triple aim: to unpack the political and economical forces that shape the very foundations of technical systems; to dissect old and dead media forms and understand how they come into being; and finally to search for, or generate, the poetics of machines. Many of these excavations can be characterised as diagrammatic, from re-activating old circuits to tracing genealogies. The use of diagrams in media archaeology is contested ground (Parikka, 2011), and highlights some lines of fracture between different approaches in the field. Some see diagrams as part of the ‘cold gaze’ of the engineer (Ernst, 2011) while others use them to trace narrative lineages that relate to cultural and political contexts (Huhtamo, 2011a).

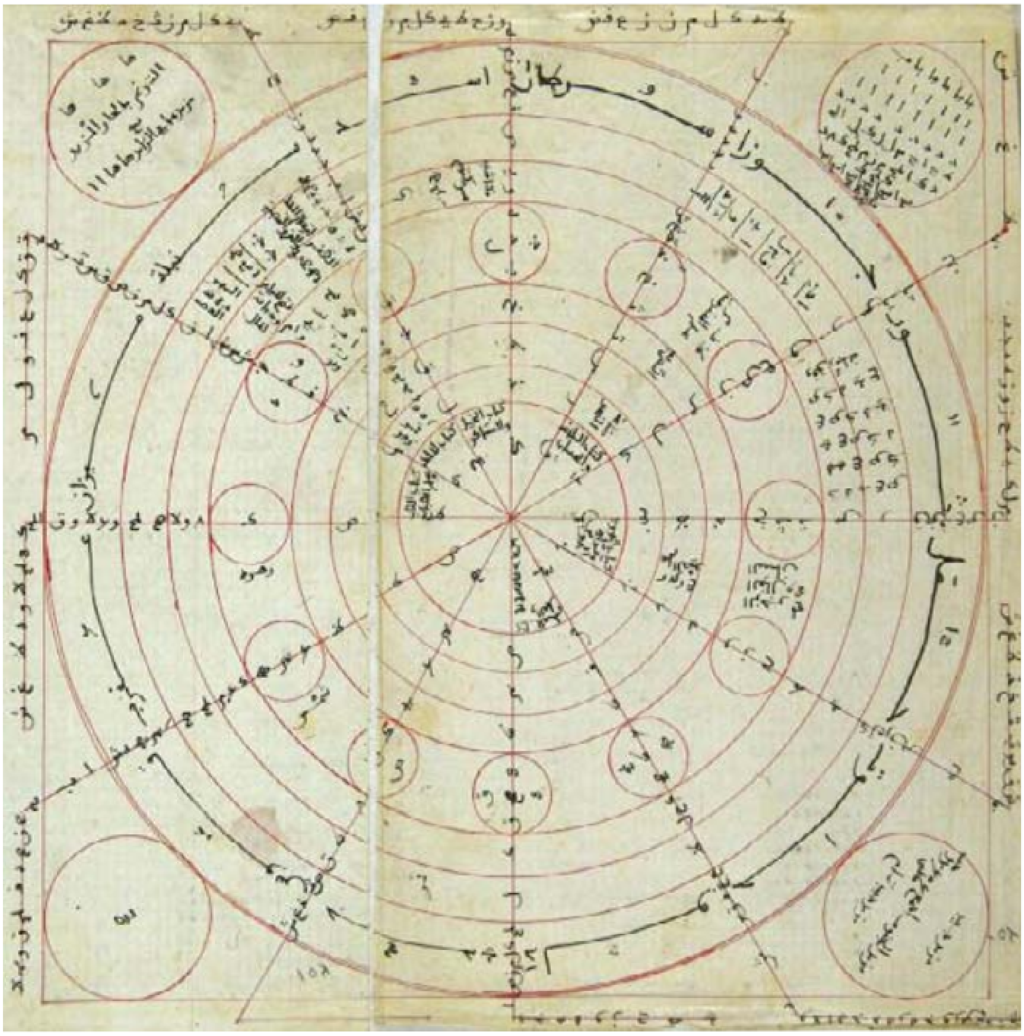


Figure 2.19: Zāʿirja front from a circa 1394 Turkish manuscript of the Muqaddima cited in Link (2010).

Diagrams are especially key to archaeologies of algorithmic artefacts, as they are inextricably linked to computation since well before the advent of the computer in the second half of the twentieth century. Link (2016) has demonstrated the practical and theoretical rigour with which these excavations can be conducted, diving back several centuries to revive artefacts such as the Zāʿirja [fig. 2.19] as fully functional diagrammatic machines. Link embodies the figure of the archaeologist of algorithms, conducting "deep time" excavations with an in-depth knowledge of history, cryptography, theory, and programming, and a particular attention to the potential of this work to change the way in which technology is perceived today. Link's work is saturated with diagrams. In many cases including his work on the Zāʿirja they are the central component of his studies in reviving and exposing the workings of algorithmic artefacts. In stark contrast with the data diagrams of the previous section, these diagrams do not exist in a vacuum but serve as vectors to investigate tortuous and overlooked histories of technology. My research is inspired by research practices such as Link's, although as a non-specialist I cannot always follow the depths of his excavations. Gitelman (2017) summarises the figure of the archaeologist in a way that is aligned to my research purposes.

> The archaeologist is more adventuresome than the historian, however, less wed to chronological narration and historiographical citation, more prone to deducto-speculative admixture and creative, illuminating connections.' (Gitelman, 2017)

Researchers and practitioners embody this figure in various ways, mixing theory, research, and practice in different proportions. Mackenzie (2017) conducts an immersion into the materials in and around machine learning. This does not cover the same time scales as Link but is more directly related to my subject here, algorithmic prediction. Mackenzie also demonstrates the use of the diagram as a language, rather than purely an object of study. This language extends from practical code experiments woven within the writing[14] and extends from practice all the way to critical theory. As I have noted my approach differs from that of Mackenzie's as it extends the diagram outside of machine learning, and is informed by creative practice rather than critical theory. Another key example is Pasquinelli (2019), researching in an art context, and providing important deep-time studies of algorithms as 'computations of space' that can be traced back to Veidic and Greek rituals. Once again by excavating deep genealogies, the archaeologist of algorithms debunks mystifying tropes around "AI" and describes algorithmic prediction more 'modestly' as the result of a 'topological turn' where data are activated through spatialisation (Pasquinelli, 2019). He remarks

> ...the genealogy of the algorithm shows that its form has emerged from material practices, from a mundane division of space, time, labor, and social relations. (Pasquinelli, 2019)

This demonstrates how a focus on space and diagrams move algorithms out of a discourse centred on technnological promise and into a much wider social, cultural, and political context. While seeing algorithms as computational geometries accurately describes contemporary computer science it also includes much longer histories that are, importantly, not centred on modern western tales of invention and innovation[15].

Scholars such as Mackenzie, Link, and Pasquinelli outline the silhouette of what an archaeologist of algorithms might look like. Their attention to material—diagrammatic—forms of computation speaks to my design background providing entry points into the broader histories of algorithmic prediction that I can begin to engage with on my own terms. I take their important contributions as the starting point for creative practice. As a material practice media archaeology provides a meeting point between academic research and creative practice, rigour and interpretation. I situate my research within this nexus, perhaps not as bent towards theory as the examples cited above but using their work as foundation to produce my own 'deducto-speculative admixture' (Gitelman, 2017).

For this I draw from art and design practices with an archaeological sensibility. Artist collective RYBN are the primary example here, and a significant influence on this research. Founded in 1999, this multi-discplinary collective conducts investigations into financial and algorithmic systems, such as offshore finance, high frequency trading, or more recently digital labour platforms. Their practice takes the production of computational artefacts—such as programs, installations, and algorithms—as a means to examine and question the histories, politics, and power relations within existing systems. For example their piece *AAI Chess* (RYBN, 2018) presents a 'pseudo AI' chess computer to draw attention to the hidden and underpaid human labour that support systems

[14] Mackenzie's *Machine Learners* is written in R-Markdown, that combines plain text with the statistical programming language R. Code can be displayed and executed within the text to produce graphs, tables, and other output.

[15] Pasquinelli points to the 'epistemic colonialism' that comes with considering algorithms only as part of Western industrialised cultures, his reading of Vedic rituals as part of the genealogy of AI attempts to remedy this.

presented as autonomous. The work includes *Human Computers*, an archive of documentation on the history of entanglements between human labour and computation[16] culminating in the notion of "faux" or "Potemkin" AI. My research approach is very much aligned with RYBN's artistic practice—I come back to their work in sections 3.4.3 and 5.3.1—from their critical outlook, to their archaeological methods, creative outputs, and interests in prediction. Points of differentiation are my focus on diagrammatic forms—while RYBN use/excavate plenty of diagrams in their work, they are not an explicit focus—and the format of their productions—installations, workshops, and online bibliographies—while I am explicitly focusing on publications in this research.

Other practitioners, while not explicitly labelled as media archaeology, still overlap with its approach. I am especially attentive to works focused on computation and its diagrammatic forms. Drulhe (2015), for example, re-spatialises the internet through speculative cartographies of its structure, using a wide range of media from computer generated graphics, to drawings, physical models, and video. This opens new ways of seeing, and interpreting, media infrastructures and their socio-cultural implications—for example the silo-ing of knowledge by various proprietary platforms. Her practice of 'spatial analysis as a tool for socio-political purposes' has distinct aims from my research, but is applicable to my purposes here. Another example tackles the algorithms of predictive policing. Artist Prévieux (2011) organised drawing workshops with Parisian policemen, re-tracing the process through which predictive policing operates [fig.2.20]. The slow process of reconstructing algorithmic geometries renders them absurd, useless for their original aim of dispatching police forces in real time, but opens up time for reflection and discussion. Prévieux remarks:

> It took us several days, even weeks at the beginning of our collaboration when our drawing methods were not really fine-tuned, to obtain what the computer traces in a fraction of a second. What we lost in efficiency, we gained in other aspects. It allowed us to reclaim agency over the algorithm that produces the diagrams, whereas this technology normally obscures the steps it goes through. (Bruno, Didier and Prévieux, 2014, 94, my translation)

These are examples of practices that reclaim the diagrams of algorithmic prediction in one way or the other. They challenge control diagrams with

[16]See Grier (2007) on the history of computation as performed by humans.

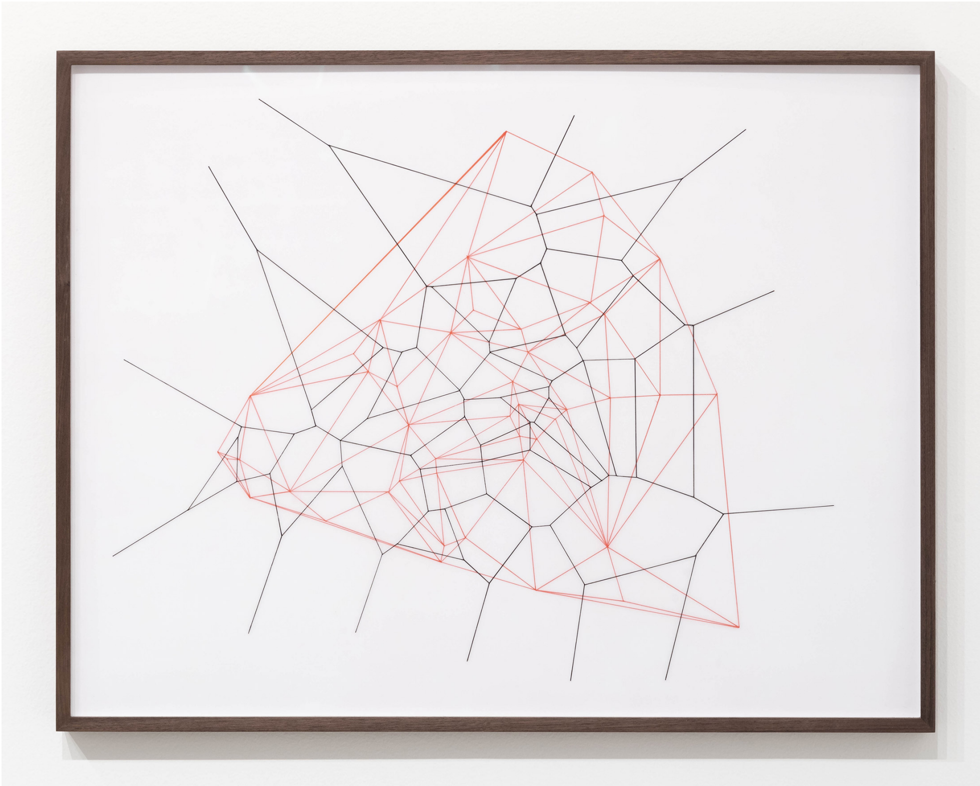


Figure 2.20: Drawing of predictive policing algorithm by Parisian policemen as part of workshops by Prévieux (2011). (used with permission)

creative interpretation, based on excavations of algorithmic material. They draw a line against the epistemic claims of the positivist data imaginary, and position themselves on the side of imagination and interpretation. In the next section I follow the oscillations of diagrams to see them not as control apparatus that pins down futures but as instruments for opening up, for extending and reclaiming.

2.4 Speculative Diagrams

While diagrams can be seen as apparatus of control, they are also generative instruments. These seemingly opposing notions are not mutually exclusive, they co-exist as diagrams constantly oscillate between the two (Leeb, 2017, 31). While the previous section framed algorithmic prediction as retrospective and systematising, I now turn to the projective qualities of diagrams, defined as 'vectors pointing in unknown directions' (ibid.). From this angle, 'diagrams work to generate a kind of cognitive sweep that extend the possibilities of thought' (Knoespel, 2001, 148). Here diagrams can be used to branch out of the data imaginary to reclaim possibilities and openings. Instead of constraining the possible to the coordinates of the past, I consider diagrams as modes of thinking (Gansterer, 2017) and knowing that relate computation and code to the imagination (Cramer, 2005).

In the next section 2.4.1, I begin by discussing combinatorial imaginaries, a core element of the early history of computing with figure such as Ramon Lull. I then move to consider the use of "possibility space" in speculative design as a way to open up spaces for the imagination through diagrams. These however do not relate or acknowledge the diagrams of algorithmic prediction. Finally I discuss the possibilty of reclaiming speculation not against but with computation, especially in a practice-based research context.

2.4.1 All possible combinations

The most obvious way in which computational diagrams can be seen as generative is through combinatorial imaginaries that date back, in part,[17] to the work of Ramon Llull in the 14th century. Llull devised diagrammatic algorithms [fig. 2.21] with the aim to generate all possible combinations of philosophical arguments. The figure of Llull is the subject of numerous studies (for example Nowviskie, 2004; Cramer, 2005; Fidora and Sierra, 2011; Gray, 2016; Vega et al., 2019). Yet he remains shrouded in mystery as an almost prophetic figure. His computational diagrams were intended to correct "false" opinions—for example to win over arguments to convert Muslims to Christianity—and to 'arrive at "true intellectual certitude removed from any doubt"' (Gray, 2016). This can be seen as the precursor for the kinds of totalising, universal imaginaries of computation that underpin modern ideals such as digital postivism; especially as Llull's ideas were later picked up by Leibniz who amplified them and contributed directly to the lineage of contemporary computing with the invention of binary numbers.

Llull also demonstrates how these computational ideals, while presented as rational and neutral, cannot be disentangled from spirituality and religious beliefs. Llull's contributions to logic and computation, via diagrams, are mirrored by his appeal to 'alchemists, cabbalists, and general mystics'

(Sales, 2011, 26). I return to the entanglements between computing and divination in chapter 5.

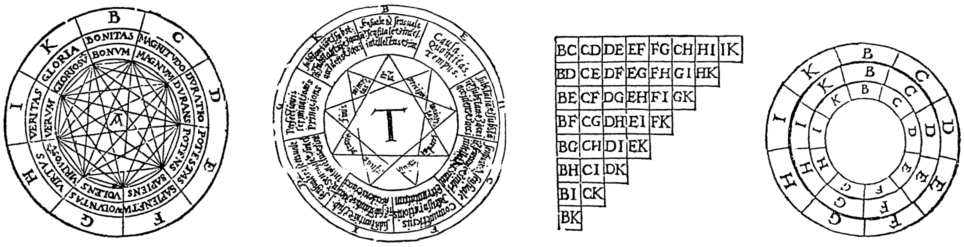


Figure 2.21: The four algorithms of Llull‘s *Ars Generalis Ultima* (1305) as cited in Cramer (2005, 36).

These combinatorial imaginaries date back centuries, but their legacy endures, embedded in the chips of contemporary computers, and in the narratives that suround them. One example is the figure of Gottfried Wilhelm Leibniz, credited for helping to establish the binary number system at the very core of computing. He was, supposedly, inspired i this by the *I Ching*—or *Changes*—an ancient Chinese combinatorial divinatory system (Gray, 2016).[18] They are being put to creative use in generative art, especially in the relatively new forms of “art bots” that have flourished on social media such as Twitter.[19] Generative systems like *Tracery* (Compton et al., 2015) work with a “grammar” of interchangeable signs selected at random to generate surprising, humorous, and/or unexpected outputs. When drawing from large corpora such as dictionaries, the permutations are seemingly endless, although they always draw from a fixed number of possible outcomes. These combinatorial artworks and bots tend to be based on very simple processes such as random number generators, enhancing their appeal as they are very easy to implement (see for example *Cheap Bots Done Quick* (Buckenham, 2015)).

One example of this generative approach that speaks to my interests here is the *Predictive Art Bot* (Roszkowska and Maigret, 2015). The bot generates briefs for new ‘potential’ artworks from headlines in the

[17] Outside of the western world, other examples include the Zā‘irja [fig. 2.19] discussed in section 2.3.3 through the work of Link (2016), the Chinese I-Ching (Smith, 2012), Hebraic Kabbalah (Cramer, 2005, 29), Algerian divination disks (Sales, 2011, 32) and many others that pre-date and inspired western combinatorial sytems such as Llull’s.

[18] For more on the *I Ching* see Smith (2012)

[19] For more on bots see Plummer-Fernandez (2019) and the work of artists such as Kazemi.

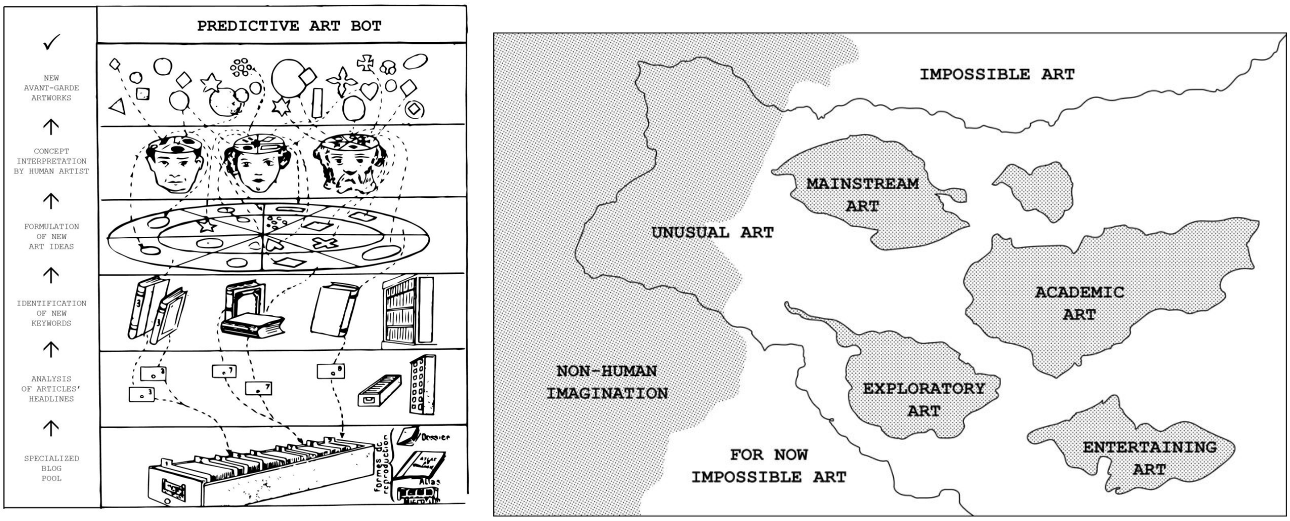


Figure 2.22: *Predictive Art Bot* Diagrams, as cited in Debatty (2016). (Roszkowska and Maigret, 2015). (used with permission)

technology press [fig. 2.23]. I see it as a commentary on the jargon used, both in art and technology, and on algorithmic prediction. It is a deceptively simple system that endlessly re-hashes existing language into new permutations that claim to predict the future. The project 'aims to stimulate unbridled, counter-intuitive and even disconcerting associations of ideas.' (Disnovation, 2018). In this sense combinatorial projects can be seen to push the boundaries of a coordinate space to expand them to unknown territories, diagrammatically relating existing elements into new configurations [see fig. 2.22].

While "generative" artworks are often based on combinatorial techniques, this is not always the case. (Palmer, 2017) for example, explores the generative potential of computational diagrams, and bridges different notions of 'scripting' for producing films and performances. Others make use of sophisticated techniques akin to the ones discussed in section 2.2. Markov-chains (for example as used by Parrish, 2018) are a probabilistic method especially popular as a way to generate text from a corpus of "training" data. Parrish (2015) explicitly uses vector space in generative and creative ways. Using N-gram data,[20] Parrish constructs a vectorised version of 'semantic space' [fig. 2.24] but she reverses the expected operations of algorithmic prediction when she uses these spaces to explore

[20] Bi-grams (pairs) or tri-grams (triples) of words used together, for example in the large corpus of books digitised by Google (Google, 2012)

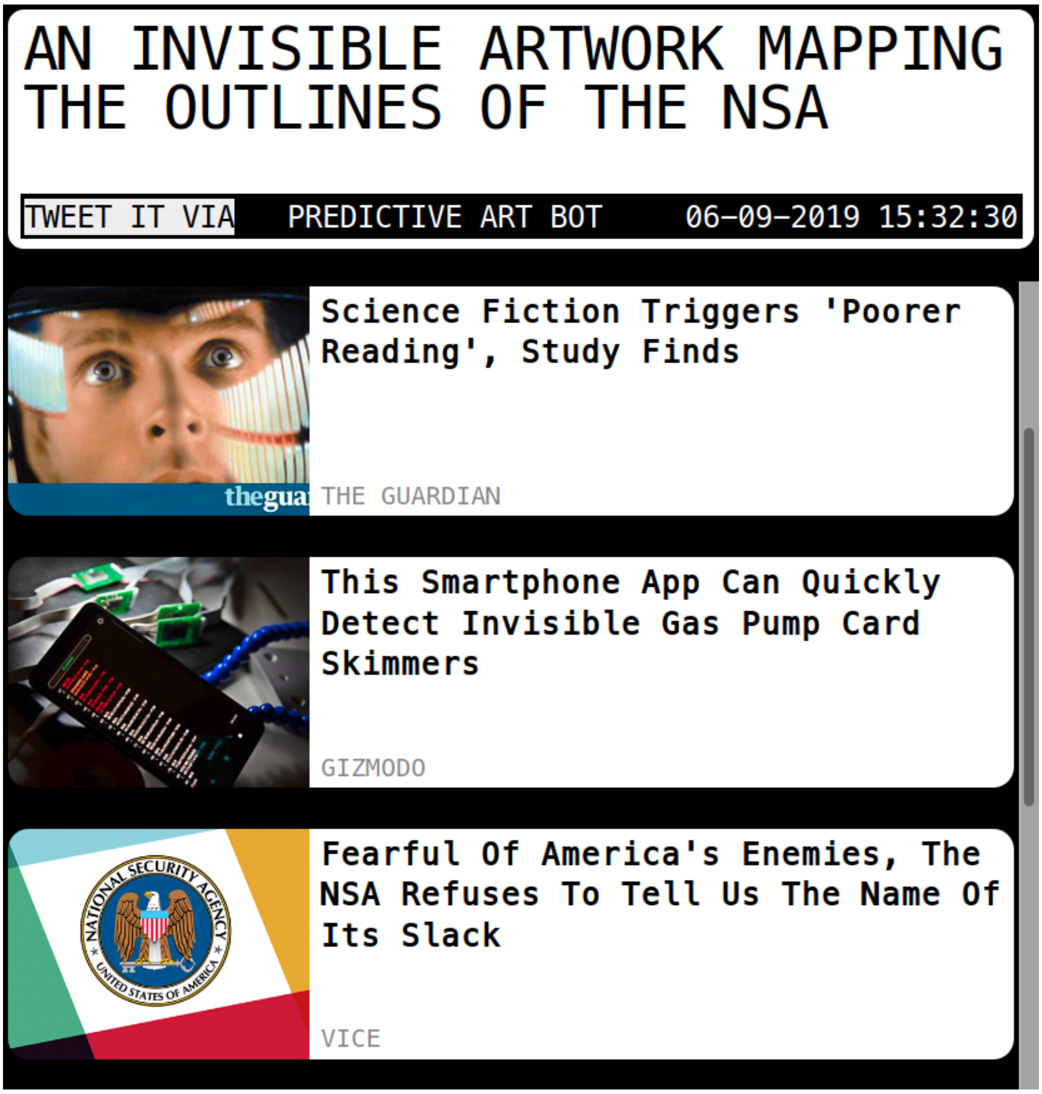


Figure 2.23: *Predictive Art Bot* example output, (accessed 6 September 2019). (Roszkowska and Maigret, 2015). (used with permission)

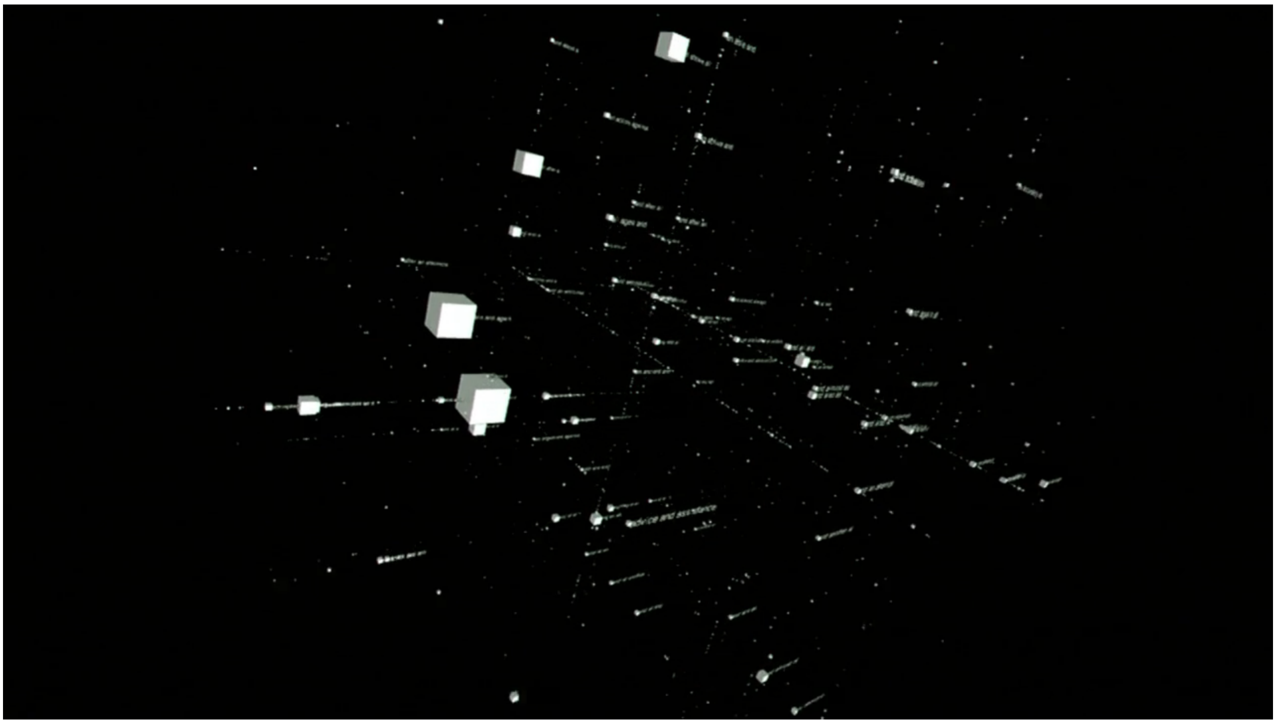


Figure 2.24: Semantic Space Parrish (2015). (used with permission)

the uncharted zones with computer programs that search for combinations of words that have never been used together.

> I like to think of these computer programs as automated probes that send back telemetry from the frontiers of "sense," exposing us to previously unforeseen possibilities for how words can behave, and allowing us to establish way stations in regions of language previously thought uninhabitable. (Parrish, 2015)

The combinatorial approaches summarised here begin to suggest that computational diagrams can be generative as much as they are controlling. Although they are still defined by the past, generating new knowledge through unexpected combinations, sometimes as in Parrish's work, through poetic means that cut right to the heart of the epistemology of vector space. Can they also expand the reach of speculations and imaginations out of pre-defined coordinate systems?

2.4.2 Designing in/with possibility space

The field of speculative and critical design aims to project and open up spaces for imagination and debate around the uses, misuses, and potential effects of technology on society and culture (Dunne and Raby, 2013; Auger, 2014). This is underpinned by diagrammatic foundations, namely the concept of *possibility space*, as represented by the future's cone diagram [fig. 2.25]. The futures cone, imported to design by Candy (2010) from future studies (Hancock and Bezold, 1994), has served as a visual defintion for the field of speculative design since its inception. The cone represents the beam of possibilities 'radiating from the present moment' (Candy, 2010, 33), shining a metaphorical flashlight into the unknown future (34). It segments these possible futures in degrees of likelihoods (probable, plausible, possible) intersected by what is politically 'preferable.'[21] The futures cone has become an icon for speculative design, and that is perhaps its biggest failing. While Candy (2010, 37-38) proposed it as one example of a 'new *way* to map, ideally one which invites and empowers more of us to make our own, rather than taking existing maps as given,' the cone has turned into an icon, the *only* map in use in this field—bar a few exceptions which I discuss below—and has been folded back into the very discourses of market-driven innovation it was initially positioned against.

Alongside the pervasive use of the futures cone, a small number of speculative designers have used diagrams as ways to map out, probe, and expand "possibility space." The *Extrapolation Factory* make extensive use of diagrams, taken from corporate and/or government forecasting methods, and democratise them as part of public, site-specific workshops [fig. 2.28]. Auger (2012) positions his research into the 'domestication' of technology, thinking in terms of parallel timelines along which technologies do, or do not, become mainstream products [fig. 2.26]. On the more specific topic of algorithmic prediction, the project *Real Prediction Machines* (Auger and Loizeau, 2015) uses a Bayesian network to predict the occurrence of events such as heart attacks, elections, or domestic arguments [fig. 2.27]. This is perhaps the only use of "data diagrams" in speculative design, with the intent to question the performative nature of algorithmic predictions. Displayed through a purpose-built connected device, the system explores if the signal of an impending argument nudges the partners into changing their behaviour. Aside from this example—a speculation that has since

[21] The way in which these sections intersect differ in the many versions of the cone diagram. In some versions (e.g. Dunne and Raby, 2013) the "preferable" is contained within the "possible" while in others (Candy, 2010, fig. 2.25) it bleeds over into the impossible, opening what is in my view the most interesting space.

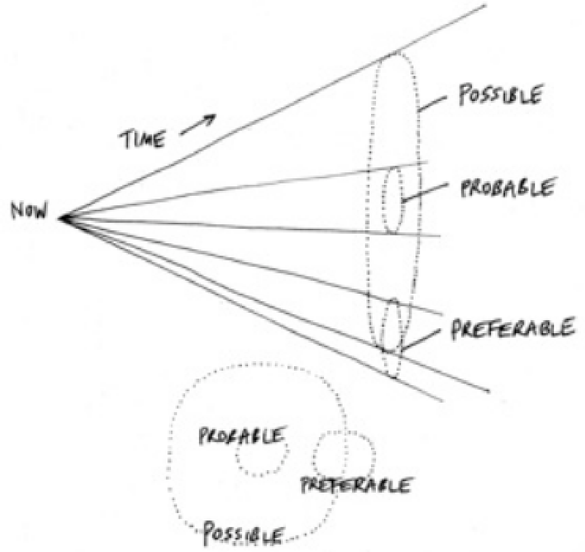


Figure 2.25: Futures Cone, as redrawn by Candy (2010, 35) from an earlier version by Hancock and Bezold (1994). (used with permission)

been realised through the rise of connected "smart" speakers in a domestic context—the diagrams of speculative design, futures cone and others, do not engage with the continuous diagramming of the future by algorithmic systems I have covered in section 2.2. They use diagrams as tools and methods for speculation but do not position themselves with regard to computational diagrams, algorithmic prediction, or algorithmic governmentality.

There are, however, some echoes between speculation as practised by designers and the notions relating to data, algorithms, and prediction I have discussed in this chapter. The first is the positioning of speculative design by Auger (2014, 43) using the concept of 'coordinates of reality' developed by Žižek in his film *The Pervert's Guide to Cinema* (Fiennes, 2006). The task of speculation, in this view, is to 'stretch'[22] these coordinates to challenge existing systems and narratives. This can be read as the beginning of a push back against the conservative nature of data diagrams enforced through the coordinate system of vector space. The second echo is the delimitation of possibility space by Bratton (2016b) between the extremes of the *Pharmakon*: remedy and poison. In his view, both are present in emerging technologies, and foreclosing either positive or negative potential is 'incomplete and/or dishonest'. The goal of

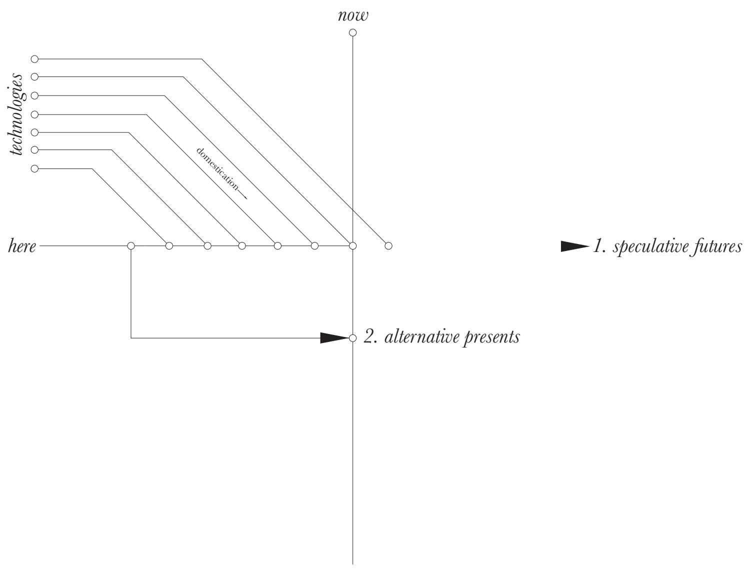


Figure 2.26: Alternative presents and speculative futures. (Auger, 2012, 135). (used with permission)

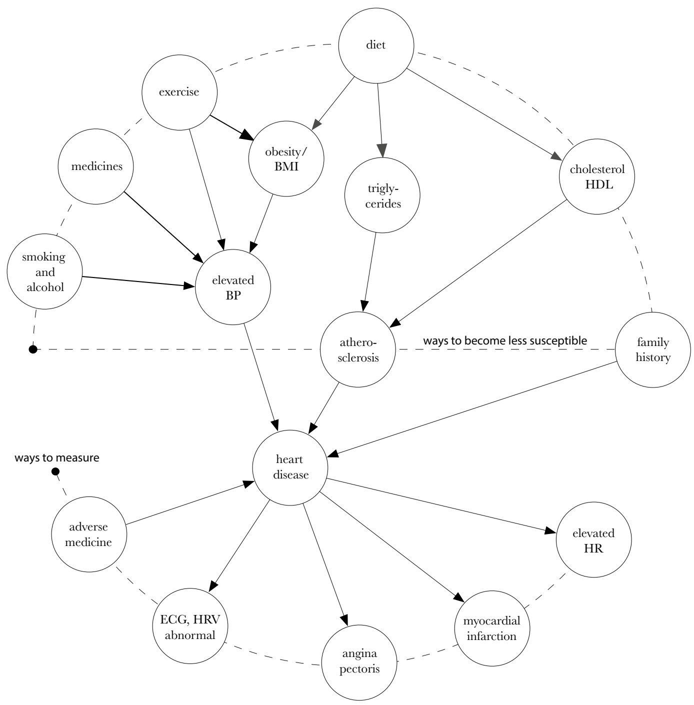
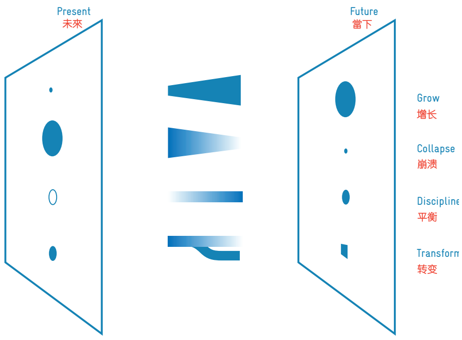
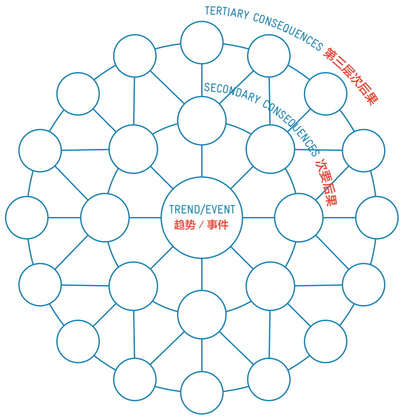


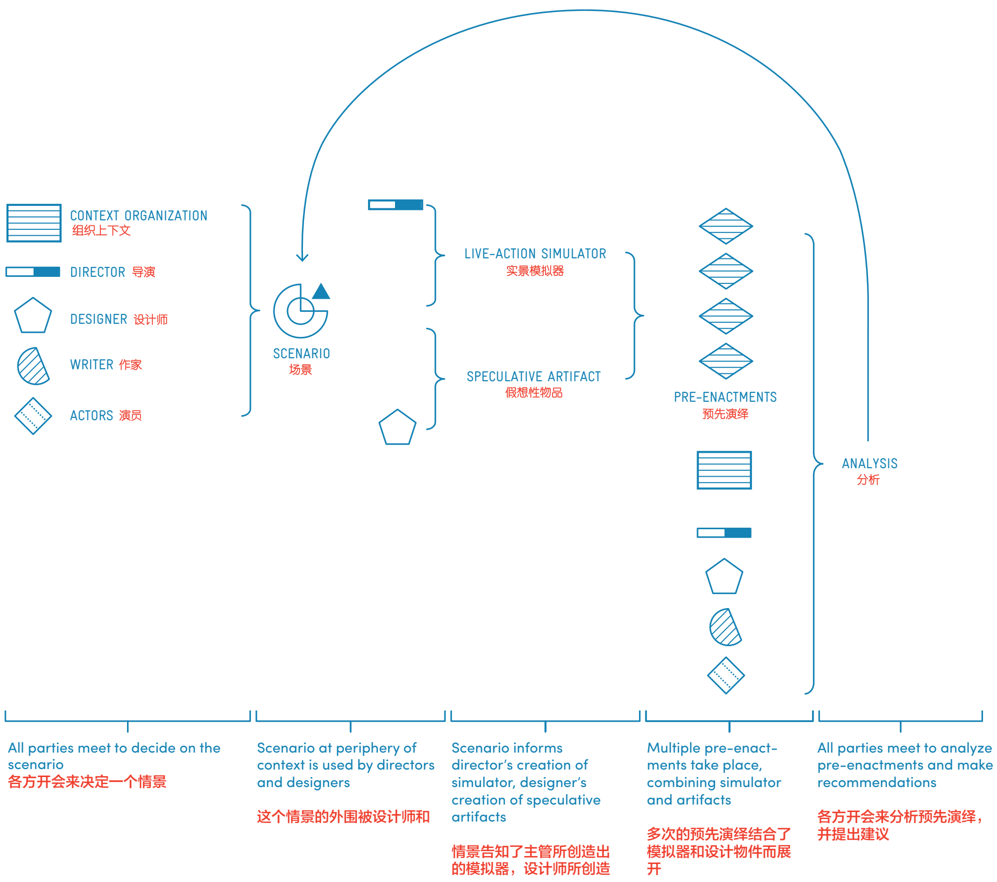
Figure 2.27: *Real Prediction Machines* Bayesian network showing the factors used to predict a heart attack. (Auger, 2014, 49). (used with permission)



(a)



(b)



(c)

Figure 2.28: *Extrapolation Factory* methods a) 4 Arcs Diagram, referencing a concept by Jim Dator and illustrated by Sungmy Kim b) Futures Wheel Diagram, referencing a concept by Jerome Glenn c) Alternative Unknowns. (Montgomery and Woebken, 2016) used with permission.

speculation, then, is to keep the 'search space' between the two as open as possible. To this end, Bratton proposes to 'rotate' the use of predictive algorithms,

> ...less to predict what is most likely to happen (deriving value from advance simulation of given outcomes) than to search the space of actual possibility (even and especially beyond what any of us would conceive otherwise.) (Bratton, 2016b).

In their more recent work Schmitt, Dunne and Raby (2018) use neural networks to search for 'impossible objects.' In a poetic exploration of vector space, they "train" a generative network on images of optical illusions and patent diagrams, and consider the multi-dimensional space of the model as a landscape of possibilities 'not constrained by the laws of physics.' While the framing of the project may obscure some of what is actually at play[23] this work suggests a break with the scenario-based premise of speculative design, and shows designers engaging with algorithmic speculation and moving towards reclaiming its materials, such as vector space, as the grounds for speculative and critical explorations. In another example, Schmitt (2020) documents the *Curse of Dimensionality*, term used to describe the challenges posed by working with data as high-dimensional spaces. Schmitt generates collages from diagrams found in the scientific literature, deriving a form of visual poetry from scientific attempts to grapple with impossible geometries.

I have focused here on examples from design that engage with *how* futures are produced through computational systems, and how agency might be reclaimed over these systems. Speculative design, while aiming to imagine provoking scenarios, rarely examines the data-centric modes of knowledge production described by Kitchin (2014), or the imaginaries analysed by Beer (2019). Apart from a few noted exceptions I have discussed here (Auger, 2014; Schmitt et al., 2018), it operates within the frame of vector space, and tends to extrapolate along the "axes" set by industry discourse—such as the data imaginary—without questioning the coordinate system itself. As an example in the field of biotechnology, the work of Ginsberg (2018) is thought provoking but its framing as 'critical' is

[23]The multidimensional objects of vector space are only "impossible" from the three-dimensional perspective of design, they get created routinely by algorithmic systems and are reified in many of their outputs.

[23]Stretching the coordinates of reality is used in contrast to 'shattering' them, which is the word Žižek uses. For Auger this means keeping a degree of 'plausibility' in speculative design scenarios in order to maximise their impact on the audience, to avoid them being too easily dismissed as fiction (Auger, 2014, 44).

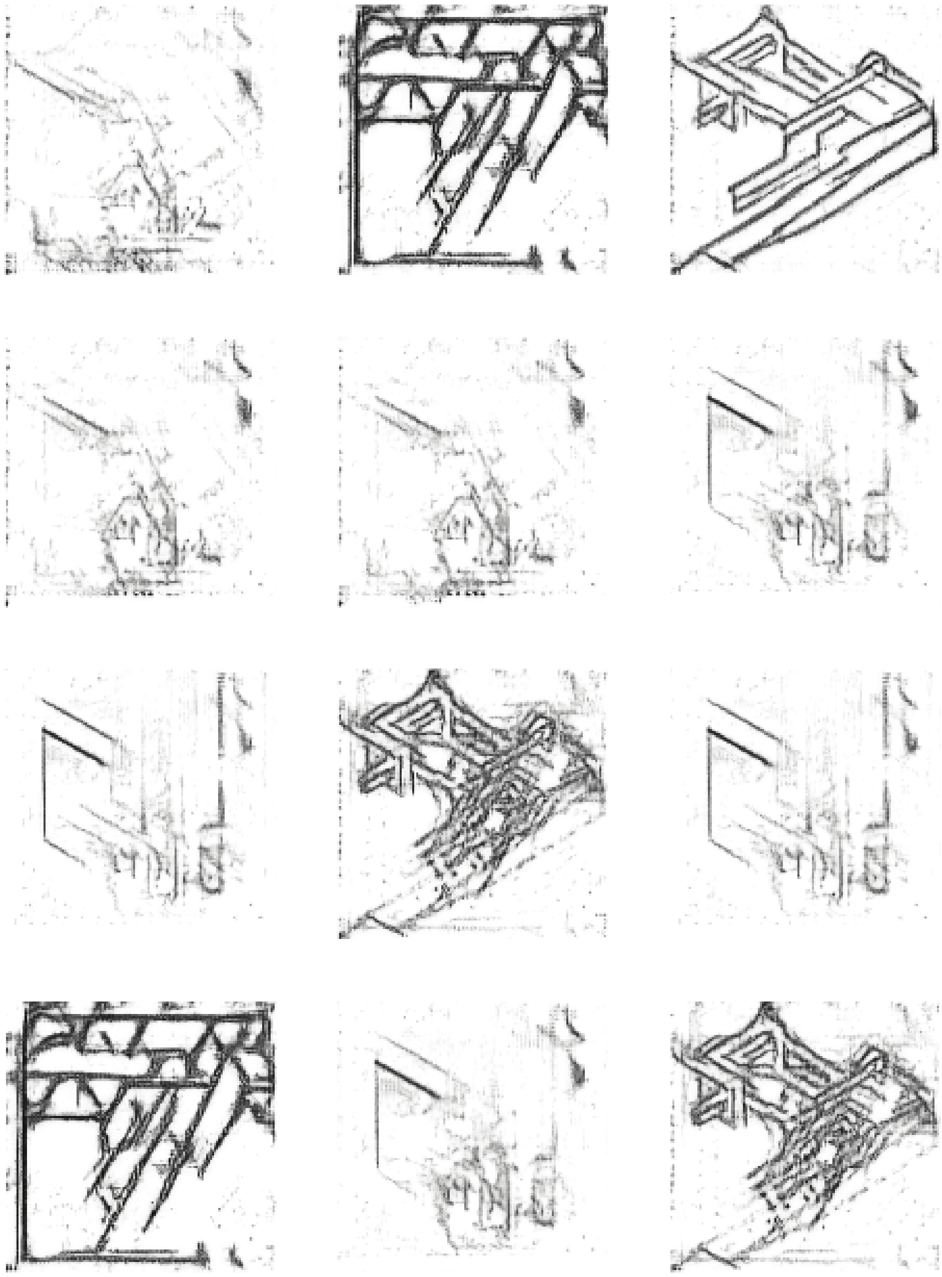


Figure 2.29: Searching for an 'impossible object' with neural networks (Schmitt, Dunne and Raby, 2018). (used with permission)

questionable when imagery is readily adopted by the biotech industry as communication material. This is in line with speculative design's broader shortcomings, as Parikka (2019) argues 'it has remained attached to a Modernist appreciation of practice; it is elitist; it is committed to a rather narrow idea of future; it fails to ascribe to a stronger sense of politics.' (25). In the last section, I turn to practices and research that aim to reclaim speculation as a mode of knowledge production, specifically those that aim to challenge digital positivism through computational means.

2.4.3 Computational speculation

As the positivist imaginaries of data and algorithmic prediction take hold, they re-shape what counts as knowledge; universal and computable ground-truths. This challenges what it means to do research, as the "big data" paradigm produces new, automated and emboldened, forms of empiricism that are purely inductive—they suppose knowledge emerges from data (Kitchin, 2014). This is a concern for many in qualitative fields such as social science, who have long considered the role of instruments in shaping the reality being observed (boyd and Crawford, 2012, 665). It is perhaps even more acutely felt in creative fields such as design that strive to establish themselves as legitimate research (Monjou, 2014) but operate through practice, creative processes, interpretation, and subjectivity, none of which are accounted for through data. Data-centric modes of knowledge production also go hand in hand with the re-shaping of research as an economic activity. As I have noted with Drucker (2014b, 87), the graphical forms of data never fully shed their initial purposes. As research is increasingly conducted with the tools of accounting, they reshape the kinds of questions being asked, and answers being sought, in these terms (boyd and Crawford, 2012, 665, citing Du Gay and Pryke, 2002).[24]

Faced with these drastic shifts in the nature of research and knowledge, speculation appears as a counter-position that embraces uncertainty and subjectivity instead of suppressing them. Like "diagram," speculation encompasses seemingly opposed meanings. One the one hand the economic register of 'firmative' speculation aims to profit from uncertainty, to solidify it, 'to pin down, delimit, constrain, and enclose—to make things definitive, firm' (Uncertain Commons, 2013, loc.127). On the other hand, the cognitive register of 'affirmative' speculation 'sabotages the exploitation of potentialities, produces the common, and opens up innumerable

[24]Here data, knowledge, and economics are entangled in bigger shifts around the nature of research and academia under neoliberalism, see Vernon (2018).

possibilities' (loc.1148). A number of scholars, researchers, and practitioners have called for a reclaiming of speculation, from calculative logics to projections and conjectures, from induction to abduction (Uncertain Commons, 2013; Debaise and Stengers, 2015; Wilkie et al., 2017; Venuturupalli Rao et al., 2015; Parisi, 2012). Speculation finds a new relevance in a moment saturated with data and computational rationality, undergoing 'a generalised crisis of modes of thinking that, one way or the other, owed their authority to notions of progress, rationality, and universality' (Debaise and Stengers, 2015, loc.3, my translation).

After the 'computational turn' (Berry, 2011) doing speculative research starts with reclaiming the possibility that 'the goal is not known in advance', which means pushing back against 'imperatives of efficiency, profitability, and objectivity' (Masure, 2017, 39, my translation). However this does not necessarily mean rejecting computation and/or data. Didier and Tasset confirm this, suggesting:

> We do not react like those who reject [statistics] wholesale and shout: "No to quantification! No to numbers! Yes to qualities!" because, in doing so, they leave a monopoly over these instruments to the powerful. There is no reason for quantification to always be on the side of the state and of capital. (Didier and Tasset, 2013, 124, my translation)

The challenge here is to speculate with, not against, computation, to find opportunities for critique and to open up creative possibilities from within (Bruno et al., 2014). Design, Masure (2017, 56) argues, should reject the (statistical) *model* as a mode of enquiry, and embrace a more risky and stimulating position that disrupts, dis-orients, and opens spaces previously thought as universal to a multiplicity of interpretations. Once again the argument here is for this to happen through design's immersion in, and unique capacity for, digital making (37). This combination of an immersion within algorithmic systems while refusing them as a mode of knowledge production is aligned with the media-archaeological stance I discussed in section 2.3.3. As I have previously discussed[25] there are significant overlaps between a practice-based archaeology of algorithms and notions of speculation (Parikka, 2019).

Calls for new forms of speculative research and practice using computation hinge on a key distinction with the positivist view of data and algorithms: the refusal of the universal in favour of the singular. This is central to

[25] See Note on speculative and critical design.

arguments by Masure (2017); Uncertain Commons (2013), and perhaps best summarised by Drucker (2009) and Nowviskie (2004) in their accounts of 'speculative computing' research and practice. Their inspiration comes from the OuLiPo movement—famous for using constraints as a generative space of creativity—and pataphysics, the science of exceptions with its 'emphasis on "the particular" over "the general"' (Drucker, 2009, 25). They build on these in their *patacritical* position that they apply to the design and implementation of computational tools. These are not general purpose but uniquely suited to their approach and research in a singular context, considering exceptions as 'valuable to speculation in a substantive, not trivial sense' (26). This challenges inductive modes of knowledge production at the core of the "paradigm shift" of big data epistemology. It prefers *abduction* (28) that considers the relations between entities rather than assume an underlying logic.[26] It also suggests that this process is more interested in activities than in final results. I come back to this in section 3.3.3 and fig. 3.11.

More extreme in breaking with computational logics is the figure of the idiot, as put forward by Stengers (2005),[27] as 'the one who always slows the others down, who resists the consensual way in which the situation is presented.' This figure has been used to theorise speculation and speculative design on a few occasions (Michael, 2012; Wilkie et al., 2017), demonstrating its appeal for design practitioners. Stengers (2005, 994) uses the idiot as part of her *Cosmopolitical Proposal* which is aimed at practitioners who have 'learned to shrug their shoulders at the claims generalizing theoreticians that define them as subordinate.' The idiot is in line with the refusal of models by Masure (2017), but also, although it may be hard to imagine them operating a computer, with the reclaiming of speculation with computation. Michael Guggenheim et al. (2017) make that link as they highlight the role of instruments and technical infrastructure in mediating certain types of speculation. While firmative, financial speculation is facilitated by the terminals and algorithms used by traders, Michael Guggenheim et al. 'suggest the building of speculative machines that encourage idiotic speculation' (146). The machine, in their case, is a sandbox. Their suggestion holds for computational/algorithmic machines however, and open the possibility that computational diagrams may be used as instruments for affirmative speculation.

[26] Drucker draws this notion from C.S. Pierce, see Speculative/abductive practice in the methods section, and figure 3.11.

[27] Building the conceptual persona by Deleuze and Guattari (1991).

In this section 2.4, I have discussed how computational diagrams are sites of speculation, contrasting their controlling nature covered in the previous section. I first discussed the combinatorial nature of early examples of computational thinking such as Ramon Llull's devices, and the use of combinatorial diagrams in art projects today. I then moved to speculative design and its uses of diagrams. The field is defined by one, the futures cone, but otherwise mostly does not relate its diagrams and speculations to data diagrams and their politics. Finally, I discussed the reclaiming of speculation as an emerging critical position for practice and research as big-data epistemology re-defines the nature of knowledge.

2.5 Conclusion

In Chapter 2, I have described three registers of diagrams as they overlap with algorithmic prediction. I drew from a wide range of material to establish broad positions that serve as polaritie. My intent is to navigate and draw relationships between them in the following practice-oriented chapters 3, 4, and 5.

In relation to RQ1, I described a range of diagrammatic forms, movements, and mechanisms involved in algorithmic prediction and in its critical investigation. With data diagrams I discussed vector space, summarised how it is produced from data, and the geometrical operations that it undergoes to produce predictions. I differentiated vector space from data visualisation, the translation of high-dimensional constructs into the two-dimensional space of graphics through visual languages, or grammars. I then focused on the oscillations of diagrams between control and openness. First I reviewed how diagrams can help to see data and algorithmic prediction critically. Like a photograph, data can be seen as framing, not an neutral objective material but actively produced. While it claims to reveal an objective ground truth, algorithmic prediction actually re-shapes the world—prescribing rather than predicting—in the image of data diagrams. I described excavation as a key movement in critically interrogating these diagrams through practice, relating them to socio-political contexts and challenging their claims of new-ness. I then moved to the possible reclaiming of data diagrams as instruments for speculation, opening up multiple possibiblities rather than pinning them down. I discussed combinatorial imaginaries and the generative potential of permutations, although they are grounded in ideals of universality. I then moved to the diagrams of speculative design, aiming to expand "possiblity space," and in some instances begging to contrast this aim with data diagrams. Finally I discussed the possibility of reclaiming computation as instruments for speculation, through a focus on singular exceptions rather than universal totalities.

With regards to RQ2, I have described how data diagrams are operationalised as part of a digital postivist modes of knowledge production. Here the geometries of vector space—although constructed through intricate and layered mediations—are considered a pre-existing mathematical order that can be objectively observed from a neutral position. I discussed how these epistemic claims are debunked, starting with data themselves and the construction of their authority. The characterisation of algorithmic prediction as an engine of prescription re-frames it as entangled in the world—not external to it—and serving specific socio-political interests. Archaeologists of algorithmic systems examine these untanglements, and through this challenge their mode of knowledge production. They acknowledge the role of interpretation and conjecture, and recognise how instruments themsleves shape the knowledge being produced. Finally speculative computing practices demonstrate that the polarity is not between a machinic positivism and a humanist imagination, but that diagrams can be reclaimed as instruments to produce singular, original research that does not know in advance where it will arrive.

Turning to manifesting this research through practice and publications, I have put this context review of algorithmic prediction in relation to practices that inform my own research outcomes. Data diagrams provide a range of materials and instruments—datasets, programming languages, visualisation grammars and tools—used throughout this research. Critical uses of data and algorithms, more or less closely related to media archaeology and its excavations, demonstrate the range of media through which creative practice unpacks algorithmic prediction and challenges its claims to knowledge production. The figure of the archaeologist of algorithms points to a variety of possible outcomes for such practice-based research, from books and academic articles, to installations, drawings, and web-based artworks. Finally speculative diagrams demonstrate how a few artists and designers have taken vector space itself as a material and reclaimed it for poetic aims. This suggests further possibilities in taking the spaces and materials of algorithmic prediction as the basis for creative practice.

In the next three chapters 3, 4, and 5, I turn to discussing the practice work undertaken for this research, weaving connections and relations between the key notions laid out in this chapter. These begin to outline the registers of diagrams to describe algorithmic prediction, articulate criticisms in close contact with its actual material operations, and to suggest alternative modes of knowledge production and creative practice.

Preamble to the practice chapters

From the proposal stage of this research, I set out to move towards or "into" algorithmic prediction in three steps. The first was "documentation" a way of scoping out the context and literature on algorithmic prediction, with the aim to produce an outcome such as an archive or publication(s). Second, an "experimentation" phase aimed to make contact with algorithmic prediction through making, to engage with its materials and produce prototypes. Finally, "fiction" was going to open a space for imagination, reconfigurations, and narrative proposals.

This initial plan morphed and shifted significantly through the research, as illustrated by figure 2.30. The practice chapters still broadly follow my proposal's phases, but they are now each centred around one diagram type, and combine the terms outlined in the previous chapter 2 in various permutations.

- The *Case-Board* retraces the genealogies of algorithmic prediction through its diagrammatic forms. I consider control diagrams as the result of long processes of spatialisation that have aimed to pin the future down. These are data diagrams enmeshed with political ideals. My intent is to map them using the speculative diagram of the case board, borrowed from detective films and TV-series, that supports a "thread" based investigation where knowledge production is performed through an endless 'yarn-work' (Mackay, 2017).
- *Traces* are the data used by recommender systems to provide endless streams of personalised content to users. These systems have been characterised as traps (Seaver, 2018), or another form of control diagram. In the process of mapping such a system, I rely on another understanding of traces as sources of conjectural knowledge (Ginzburg, 1980), or speculative diagrams.

- Finally I revisit a prototype for algorithmic prediction, the almanac publication. I use its form to examine the *chicanes* through which predictions are produced. Here the polarisation between control and speculation recedes as I examine how both rely on “sharp turns” and “tricks” to operate. I draw parallels between algorithmic prediction and divination as a way to examine the motivations behind these tricks (Ramey, 2016).

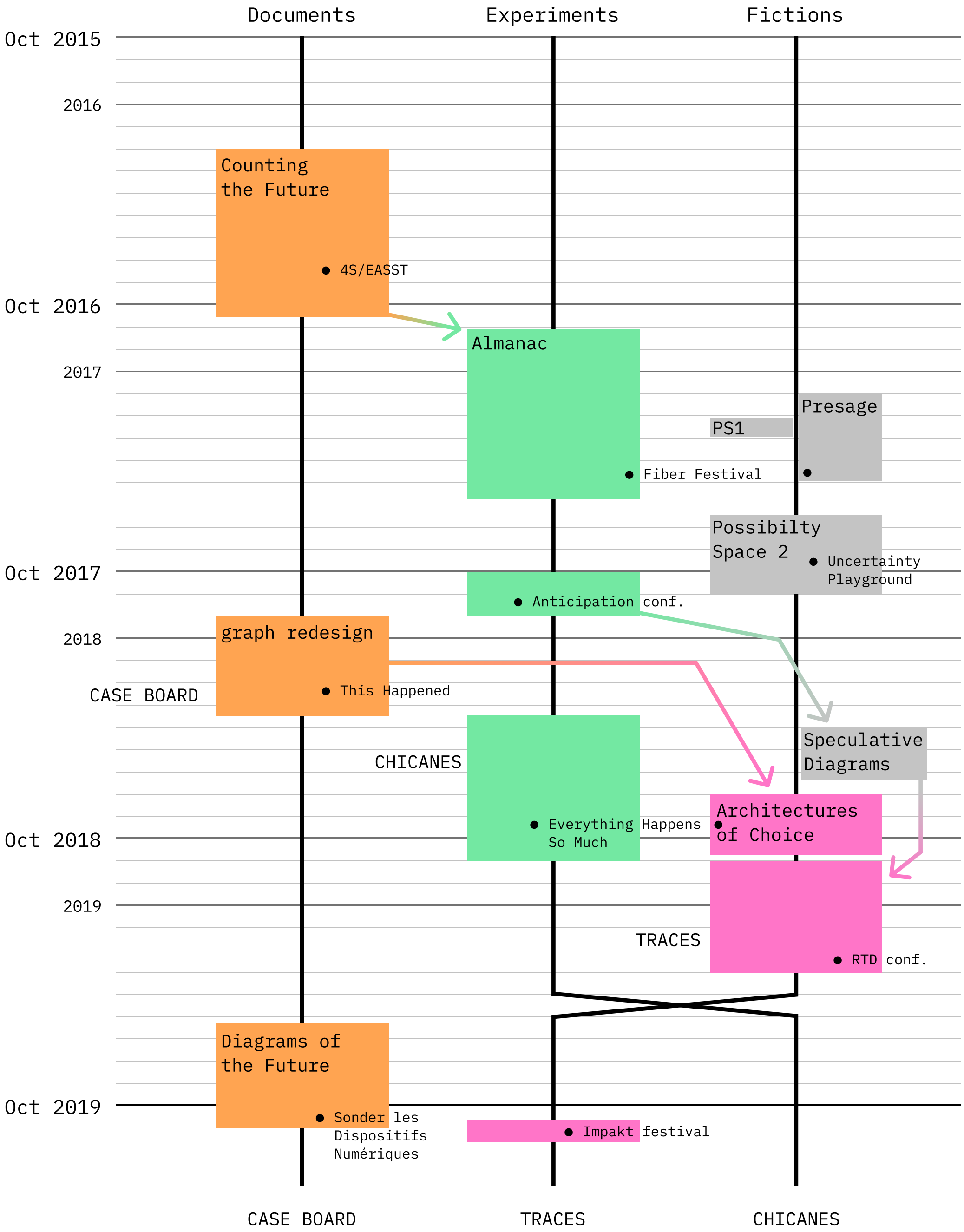


Figure 2.30: Project timeline showing the relationships between the three practice projects through the duration of the PhD, as well as key milestones. For full details of the events mentioned, see appendix H.

Please see appendix C for the practice submission related to this chapter, including: code repository, website, and other supporting material. *Diagrams of the Future* is online at: `http://dotf.xyz`

Chapter 3

Case Board

3.1 Introduction

Algorithmic prediction involves planetary-scale computational systems[1] surfaced through mundane outputs such as notifications on a smart phone or a box with recommendations on the side of a web page. This bridging of scales is further emphasised by the high-dimensionality of vector space I discussed in section 2.2. One way to describe such objects, with a dimensionality that is impossible to grasp in its entirety at any one moment, is Morton's 2013 notion of hyperobject.[2] 'Hyperobjects occupy a high-dimensional phase space that results in their being invisible to humans for stretches of time.' (1) Mackenzie (2017) makes a link between algorithmic prediction and the qualities of hyperobjects, in his study of *The Elements of Statistical Learning* (Hastie et al., 2001), a popular machine learning textbook. He writes,

> In the range of references, combinations of code, diagram, equation, scientific disciplines, and computational elements, and perhaps in the somewhat viscous, interobjectively diverse referentiality that impinges on any reading of it, *Elements of Statistical Learning* betrays some hyperobject-like positivity

[1] For more on planetary-scale computation see Bratton (2016a).

[2] Hyperobjects are a concept by Morton (2013), drawing on object-oriented ontology: 'entities of such vast temporal and spatial dimensions that they defeat traditional ideas about what a thing is in the first place.' Morton uses the term primarily to describe climate change but points to other instances such as capitalism or the Solar System.

> (Morton 2013). It is an accumulation of forms, techniques, practices, propositions, and referential relations. (Mackenzie, 2017, 30)

Described in these terms, algorithmic prediction is at the same time 'sticky' (Morton's term), that is very much *there*, present, saturating the daily experience of digital media, while also constantly evading full scrutiny, boggling cognition and comprehension. I faced a number of key challenges in dealing with such an amorphous subject. These had to with issues of representation at first, and were later addressed through using the notion of diagram to conceptualise algorithmic prediction and regain purchase on it.

In this chapter I locate cracks in the hyperobject of algorithmic prediction by mapping some of its history. I discuss how *Diagrams of the Future* (DOTF), a purpose-built mapping tool, acts as a digital *case-board* to support an excavation of the history of algorithmic prediction. Media archaeology focuses on overlooked historical artefacts and challenges narratives of technological progress. This chapter is about finding ways to excavate the diagrammatic forms that come together in the current predictive regime, to look back at them as 'media remembered' (Huhtamo and Parikka, 2011, 55). By putting these diagrams in relation to each other, I aim to challenge the new-ness of contemporary prediction techniques; to re-frame them as part of a long history of constructing the future as a diagrammatic *space* made of vectors prone to operations, and control.

I begin this chapter by proposing that the history of algorithmic prediction can be understood as a diagram of diagrams. By this I suggest that while individual techniques are diagrammatic in nature, their relations across time form genealogies. I then move to discussing the practical implementation of *DOTF*, using a graph database to reflect the diagrammatic nature of my investigation. Finally, I discuss the *case-board* as a mode of diagramming the history of algorithmic prediction, a visual, interpretive, and situated mode of knowledge production.

3.2 The history of prediction: a diagram of diagrams

In this section I establish the theoretical context for *Diagrams of the Future*, making the case for excavating the history of algorithmic prediction

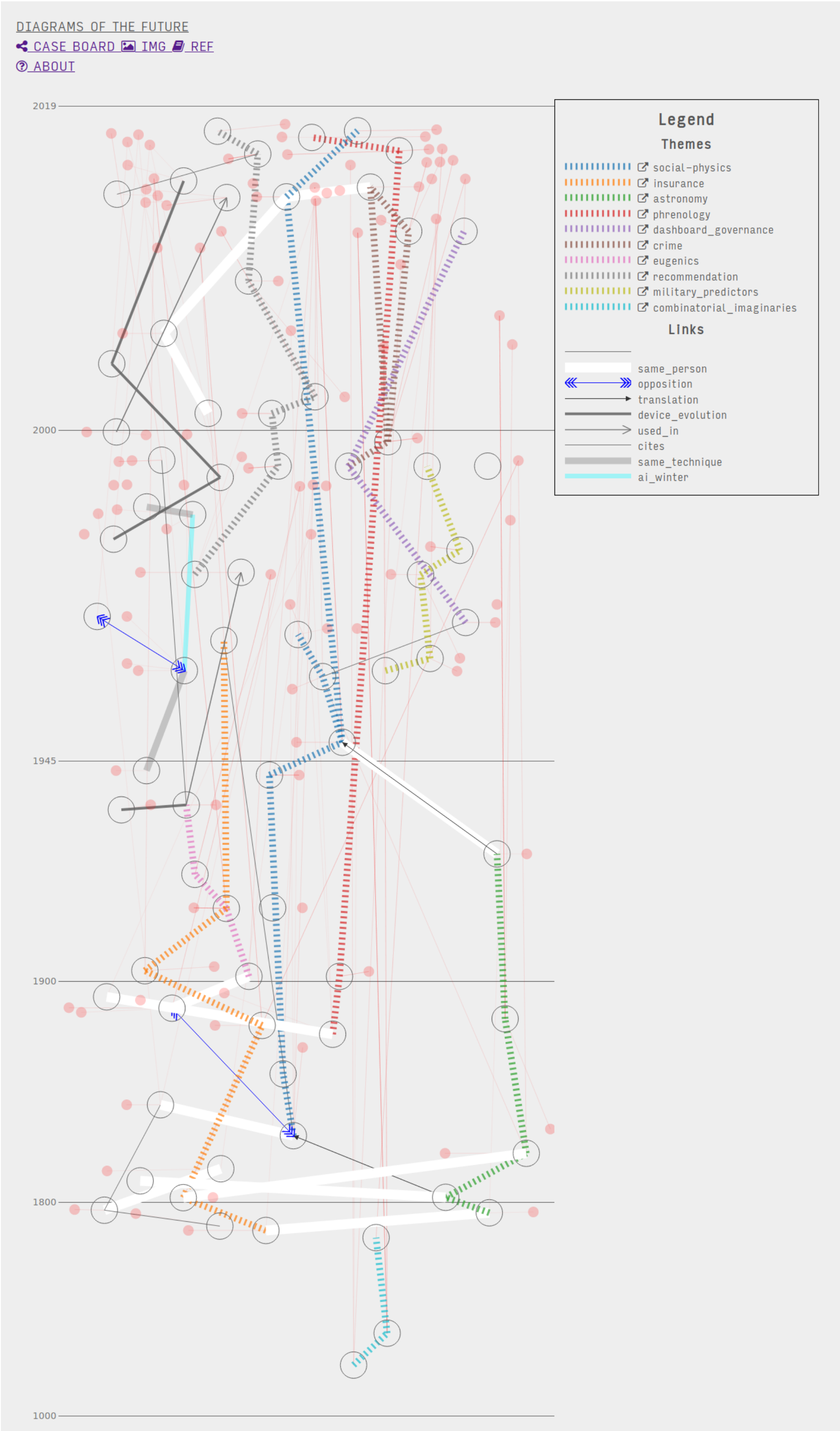


Figure 3.1: *Diagrams of the Future* Case board page - http://dotf.xyz/timeline.

as/with diagrams. This forms the basis for the practical work discussed in the next two sections.

3.2.1 Big data's historical burden

Diagrams of the Future or DOTF is the first project generated within the first few months of this PhD research. The project was the first meaningful meeting point between research and practice, as I grappled for ways to engage with algorithmic prediction from a visual standpoint. I started from the premise that critical research and/or practice that seeks purchase on algorithmic prediction should start by looking at its history. Prediction is framed as a breakthrough "innovation" that has recently, or will soon, "disrupt" just about every sector of social, cultural, and economic activities (Agrawal et al., 2018). Across economic sectors, businesses are, reportedly, either experiencing or about to experience an 'AI moment' (24) where prediction will change things forever. The focus of this discourse, as buzzwords shift from "Big Data" to "Artificial Intelligence," is centred on 'modalities of change rather than forms of continuity' (Rieder, 2016, 2). But there are indeed continuities. Making predictions from data is not new at all. It builds on at least three centuries of developments in statistics and probabilities, and their instrumentalisation 'for business profit, population control, and governance' (Elish and boyd, 2018, 58—59).

Critical geographer Barnes (2013) argues that 'with big data comes big history' (298). He turns to science studies to challenge the 'disembodied universal logic that has neither history nor geography' (ibid.) that permeates the provocative claims by Anderson (2008) that "Big Data" will soon 'speak for themselves' and replace the scientific method. If data have never been bigger, the underlying logics, politics, and statistical methods, are far from new, and have been extensively researched (e.g. Gigerenzer et al., 1990; Hacking, 1990; Porter, 1995; Daston and Galison, 2007; Desrosières, 1998). The social implications of the current "big data revolution" are, likewise, drastic amplifications of old problems that originated in the 19th century with modern data analysis (Robertson and Travaglia, 2015). They include what can, and what cannot, be *counted* in systems of data collection and analysis; the politics of ordering and scoring; and the way statistics and probability rationalise inequalities, and entrench biases. This is not to say that there is nothing new about the current version of algorithmic prediction but that, following Barnes (2013), it should be seen as 'the particular conjuncture of different elements, each with their own history, coming together at this our present moment' rather than “innovations” in a cultural and political vacuum. In more material terms, Barnes and Wilson (2014) propose that ‘unpack[ing] some of Big Data’s historical burden’ is the way to make it ‘available for verbalist discussion and contestation.’ This position is aligned with media archaeology and its consideration that ‘the facade of innovation may mask tradition, and apparent ruptures disguise hidden continuities’ (Huhtamo, 2011a, 28).

My aim with *DOTF* is to do this unpacking, or excavation, from a visual and diagrammatic perspective on two levels: first by collecting and examining the diagrams of prediction—to reframe it as a process of spatialisation—and second by relating these diagrams across time, to examine their evolutions and genealogies. I come back to the practical description of this unpacking in section 3.3.

3.2.2 Taming the future into a space

The *Taming of Chance* described by Hacking (1990) can be read as a spatialisation of the future as geometries through statistical curves and data plots. Predictive relationships between data—such as correlation and regression, developed by Francis Galton in the early 20th century—were not only illustrated graphically but *originated* in graphical methods of data analysis. They combined political ideas (e.g. about heredity) with visual representations of data (e.g. height measurements), through ‘a willingness to use both mathematical smoothing and [Galton’s] own eye-brain smoothing’ to make them fit together (Friendly and Denis, 2005, 113). In other words, Galton employed graphical methods to substantiate and make visible his ‘firm belief’ in statistical relations between parent and offspring. This was rendered as a perfectly smooth ellipse [fig. 3.2] through a negotiation between the limited number of available data points, smoothing techniques, and his sense of what the relationship “should” be like. The work of Galton demonstrates how the measurement of populations turned them into data, and made them subject to mathematical laws and geometries. The deep links between diagrams and social values are also visible by the “normal distribution” [see fig. 3.3] that underwrote Galton’s notion of “normality” as a mediocre condition which humans ‘regress’ to (Hacking, 1990, 178). Observed probabilities, and the now famous bell curve, were a foundation for eugenicist politics. Hacking characterises Galton‘s re-definition of “normal” and its embedding in social norms that survive until today as a ‘consummated’ match between word and curve (184).

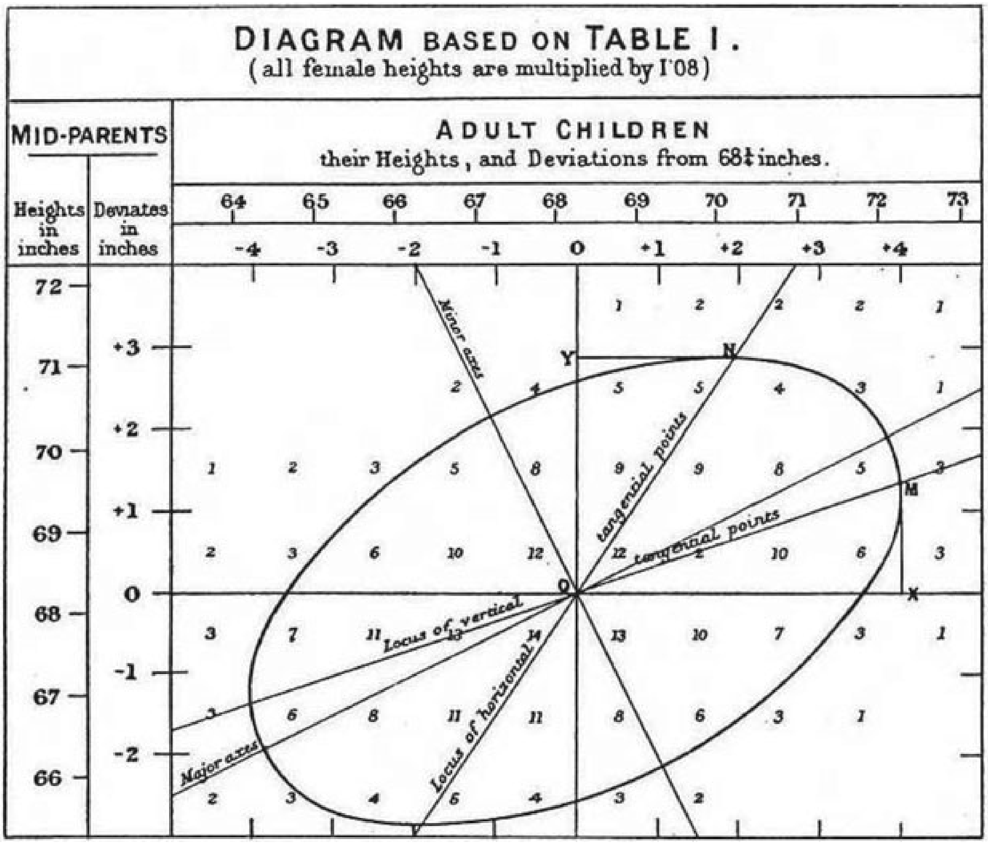


Figure 3.2: 'Galton's smoothed correlation diagram for the data on heights of parents and children, showing one ellipse of equal frequency.' (Galton 1886, cited in Friendly and Denis, 2005, 111).

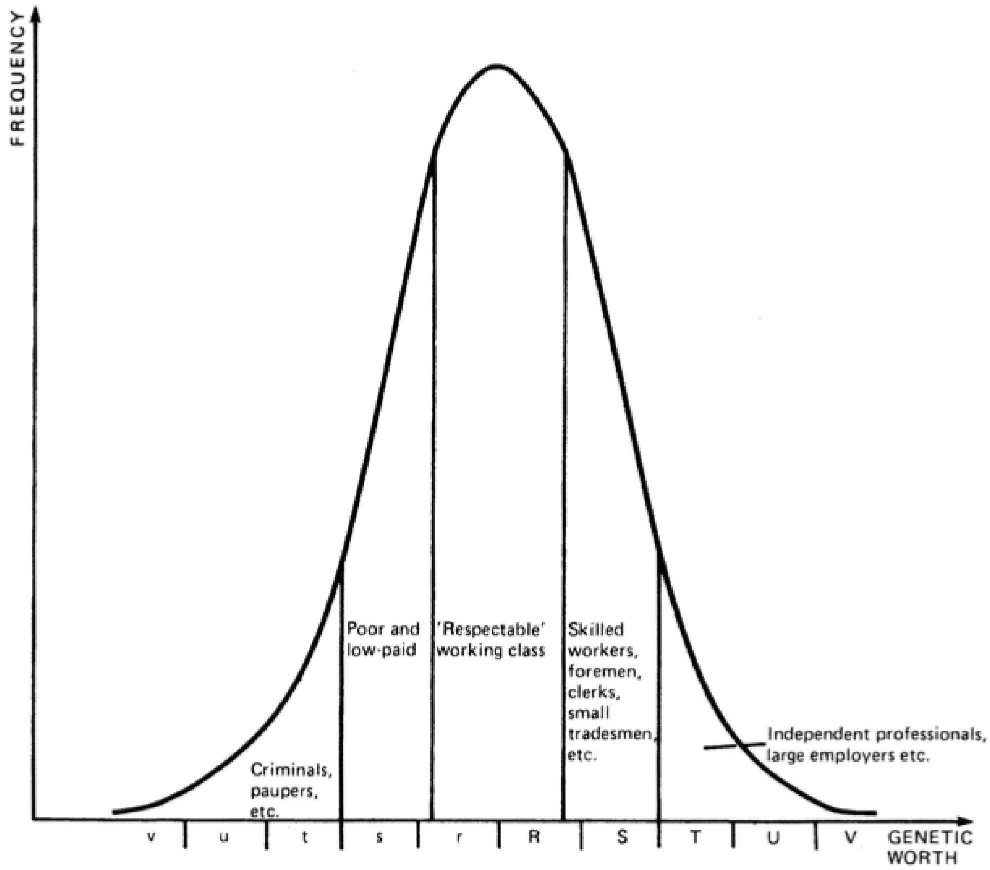


Figure 3.3: Galton's eugenist view of British social structure (MacKenzie, 1976, 514).

Another example of spatialisation is the famous[3] Iris Dataset (the example used in section 2.2) and its associated discovery by Fisher (1936) of a statistical method to separate flower species based on their measurements. The set of flowers were measured and vectorised as three dimensional models by Anderson (1936) into a dataset with four dimensions: sepal length, sepal width, petal length and petal width. Fisher used the data to derive mathematical functions that separated the vector space, effectively drawing lines between species. While more elaborate classification techniques have since been developed, this remains their core mode of operation: vectorise a set of data into a space, find the lines that separate the categories, predict the categories of new data according to their position in the space. The Iris Dataset itself is still routinely used in machine learning demonstrations. Like Galton's work, Fisher's was not politically neutral, published in the *Anals of Eugenics* his paper starts with the statement:

> When two or more populations have been measured in several characters, $x_1, \ldots, x_8$, special interest attaches to certain linear functions of the measurements by which the populations are best discriminated. (Fisher, 1936)

The political implications of "discriminating between species" in 1936 should be obvious. They are another example of the spatialisation of the future through the vectorisation of data and operations on the resulting space.

Excavating the diagrams of algorithmic prediction foregrounds the underlying geometry of the cultural and political shifts that are well covered by the science studies literature on data, probability, and statistics (Porter, 1995; Gigerenzer et al., 1990; Hacking, 1990). This geometry is not mere representation but an intrinsic part of the construction of statistical methods. My research on *DOTF* aims to re-examine the history of prediction as the conceptualising of the future as a diagrammatic space, prone to operations and control. Examples like the normal distribution or the iris dataset, both iconic and well known, show how the geometries of prediction reify political ideals; predictive diagrams are '*socio*-technical concepts' that cannot be separated from their 'contexts of development and use' (Elish and boyd, 2018, 58).

[3]At the time of writing, Google Scholar (accessed 8 December 2019) lists 16477 papers citing the original paper by Fisher (1936) on classifying iris species. This does not account for the many courses, blog-posts, demonstrations, and other formats that use the dataset as a test-case.

If looking at individual examples through a diagrammatical lens can be helpful, this approach also lends itself to considering relations and evolutions between techniques and across time. However the relationships are far from direct lineages or simple progressions of technology.

3.2.3 Genealogies of prediction

Retracing the "family tree" of algorithmic prediction quickly becomes a complicated endeavour. If the history itself can be seen as a diagram, what shape does it have? The *genealogy* of technology is a major interest of media archaeology, thinking with Foucault about 'how to think historically but avoid the idea that there are such things as simple origins.' (Parikka, 2012, 13). Pasquinelli (2015a) provides a broad overview of the *Evolution of Machinic Intelligence* [4] [fig. 3.4]. 'Technological' and 'political' axes define a space that is activated by movements connecting computing, critical theory, and politics. Pasquinelli's diagram *pins down* terminology, artefacts, theories, and scholars/thinkers in a coordinate space that *opens up* new connections and interpretations.

Pasquinelli's map shows broad "currents" to unpack 'machinic intelligence' as a set of evolutionary relations between technologies, ideas, narratives, and so on. The map conceptualises genealogy on a surface—activated by a set of axes, spatial positions, and arrows signalling movements—and offers a top-down view of it. My aim with *DOTF* differs from this in the sense that it supports building my knowledge from specific artefacts—diagrams collected in/through the literature—expanding as the research progresses. *DOTF* is a tool for 'starting in the middle' (Sayers, 2017, loc.117) with individual examples and following/connecting threads from there. One way to characterise these threads is the re-purposing by Huhtamo (2011a) of the literary concept of *topos* for media archaeology. A topos is 'a stereotypical formula evoked over and over again in different guises and for varying purposes' (28), in other words a kind of space or trope which one can track as it morphs, repeats, or ruptures. One definition of the practice of media archaeology is, according to Huhtamo, 'identifying topoï, analyzing their trajectories and transformations' (ibid.).

The first topos I encountered is the notion of *Social Physics*, which is the part of 'Big Data's historical burden' that Barnes and Wilson (2014) unpack. The basic idea of this narrative is that the "movements" of society

[4] Part of a syllabus for a seminar in 'Critical Artificial Intelligence' at HfG Karlsruhe.

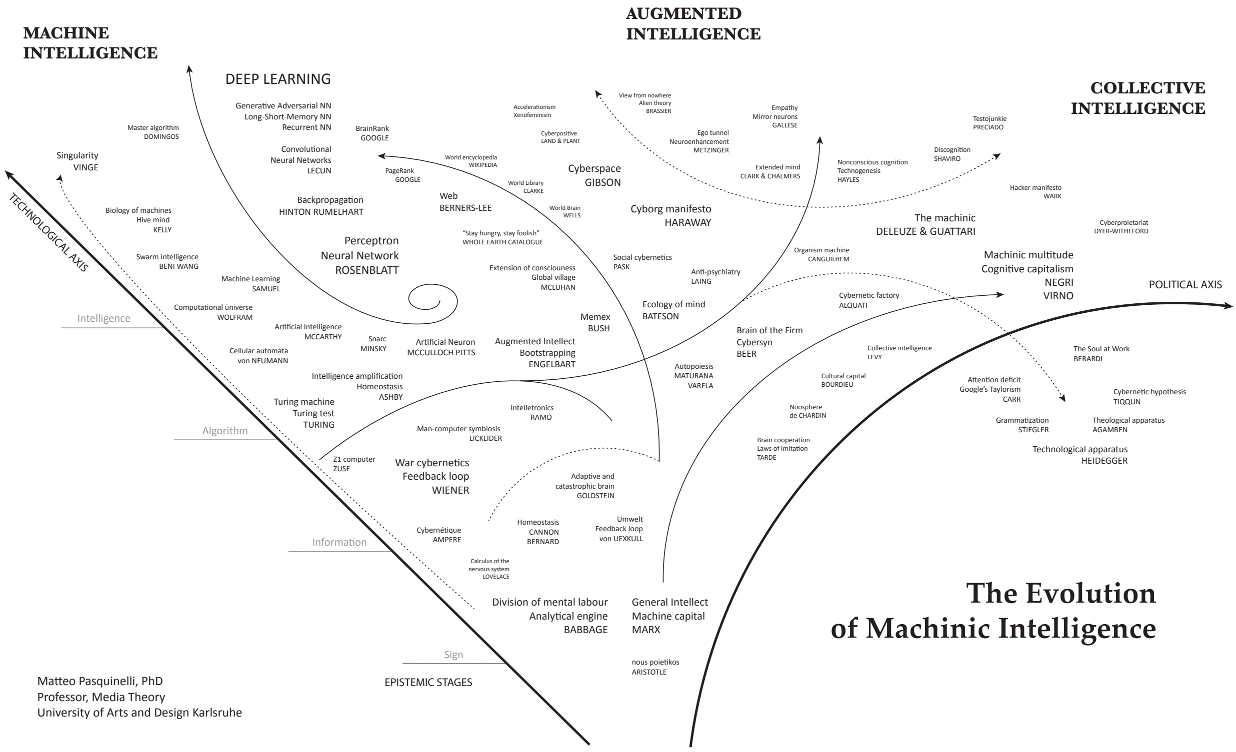


Figure 3.4: *The Evolution of Machinic Intelligence* Pasquinelli (2015a). (used with permission)

are governed by the same mathematical laws of physics as the natural world, and are therefore predictable. The phrase originates in the positivist philosophy of Auguste Compte, describing what was later called "social science." Barnes and Wilson (2014) start tracking it in 19th century Belgium, as Adolphe Quetelet used techniques to calculate smooth orbit curves in astronomy, and applied them to society in order to predict marriages and crime. Quetelet's construct of an *Average Man* was 'designed to facilitate the recognition of laws analogous to those of celestial mechanics in the domain of society' (Gigerenzer et al., 1990, 41). From there *Social Physics* ripples and morphs through the 20th century, for example with the application of the power law to analyse language by Zipf (1942, cited in Barnes and Wilson, 2014). It is a cornerstone of the field of spatial analysis, where the use of *gravitational* potential to project *population* potentials by Stewart (1948, cited in Barnes and Wilson, 2014) proposed that the distribution of human populations was subject to the laws of gravity, a kind of viscous liquid spreading through the land. Similar ideas get embedded in early forms of computer mapping, especially at the Laboratory for Computer Graphics at Harvard's Graduate School of Design (Chrisman, 2006). Extending beyond Barnes and Wilson's work, *Social Physics* continues in the big data era, this time coming from MIT's Human Dynamics laboratory. Pentland (2014) re-appropriates the term as a 'new science' premised on big data, which is the basis for Endor, a platform for business analytics predictions backed by a cryptocurrency (Endor, 2018).

Social Physics demonstrates the capacity of topoï for 'wandering across time and space' (Huhtamo, 2011a, 36) from 19th century astronomy to a 21st century crypto-currency. Their use by Huhtamo 'deviates' significantly from their origins in literary studies (34), to make them useful for the study of media and culture. For my purposes with *DOTF*, thinking about the history of prediction as a diagram of diagrams, they provide a kind of malleable scaffolding which is in turn a loose adaptation of Huhtamo's ideas as they meet the constraints of practice. In summary, topoï begin to point towards ways of addressing my research questions. While algorithmic prediction was conceived as/with data diagrams from the outset, topos study offers a way to follow the threads of its evolution diagrammatically.

3.3 Practice

After describing the history of algorithmic prediction as a diagram of diagrams, I turn in this section to engaging with this form in my practice. I discuss the instrument used for my excavations, the case board, and the set of references that helped me define it as a mode of practice. In particular the notion of 'yarnwork' (Mackay, 2017)—drawing from popular culture and the figure of the detective—frames my practice in relation to "investigating" as an activity represented by, and performed with, case boards.

3.3.1 Case boards and yarnwork

If algorithmic prediction can be considered a hyperobject, it overlaps with—or perhaps is simply within—another even larger one, namely capitalism.[5] It is with regard to the latter that Jameson (1988) proposes 'cognitive mapping' as a possible strategy to begin to reclaim a sense of situation and agency. If 'no one has ever met or seen the thing [capital] itself' (354) Jameson argues for an aesthetic that would represent capitalism spatially. Such an aesthetic could, in his view, bridge the separation between the space of lived experience and the abstracted, distributed and opaque systems that govern capital. Jameson left his proposal open-ended and un-resolved, which allows more recent scholars to pick it up and revisit this line of inquiry. Jameson's work is, for example, the starting point for research on contemporary modes of mapping capitalism by Toscano and Kinkle (2015), against a context in which:

> ... an inability to cognitively map the gears and contours of the world system is as debilitating for political action as being unable mentally to map a city would prove for a city dweller. (Toscano and Kinkle, 2015, 24)

Toscano and Kinkle provided a key connection between cognitive mapping and popular culture, as I was struggling to characterise my practice of mapping the history of prediction with the *DOTF* project. I had the sense that I was engaging in a speculative form of mapping (Corner, 1999), or a critical visualisation of sorts (Hall, 2008), but needed a more precise framing to move the research forward. Toscano and Kinkle draw from a wide range of aesthetic forms across film, art practices such as Bureau d'Études [fig. 3.5], as well as TV series such as *The Wire*. In this latter

[5] Morton (2013) includes 'the sum of all the whirring machinery of capitalism' as an example of hyperobject.

[redacted]

Figure 3.5: *World Government* (Bureau d'Etudes, 2004). (redacted)

example, one prop resonated with my work on the *DOTF* project as an investigative practice: the detective's 'case board'. The authors argue,

> Attention to visual and material mediations also shows *The Wire* to be a peculiarly reflexive study on what modalities of mapping and representation are bearers of effective knowledge. Hence **the key role of the case board as an epistemic tool** – one with interesting resonances to the artworks of the likes of Lombardi or Bureau d'Études . The case board of course cannot escape working through segments, fragments, compartments; it is never a truly 'totalising' tool, nor can it simply 'reveal' the routes of money. (Toscano and Kinkle, 2015, loc. 298.8, **bold** emphasis added)

Familiar to viewers of films or TV series where criminal investigations take place, the 'case board' is a plot device where the progress of the detective's discovery is tracked, often by pinning up pictures of suspects and marking their connections with thread.[6] In the context of this research, the case board is appealing because it does not make any claim to reflect "total" knowledge but only that of the detective—or, in this case, my role as a researcher—working through segments. Forgetting for a moment the fact that TV detectives tend to *solve* crimes, I take the case board as an end in itself, a tool for producing partial, speculative, and incomplete knowledge. It is a true *map* in the sense of what Deleuze and Guattari (1987) have observed as 'open and connectable in all of its dimensions; it is detachable, reversible, susceptible to constant modification' (12).

In a thorough analysis of '*yarnwork*'—the marking of connections with threads—that takes place on/with case boards, Mackay (2017) notes that this device is a trope.[7] Case boards exist only on camera as a materialisation of the protagonist's thought process, 'a diagram of a diagram of thought in action' (Mackay, 2017). Case boards are a mode of cognitive mapping, they allow protagonists and audience to situate themselves within the entanglements of criminal networks by materialising relations. They are key props in detective stories as an epistemological genre: 'dramatisations of the process of obtaining and configuring knowledge' (Mackay, 2017). This resonates with my research interests as it positions the case board as a site where knowledge production is "played

[6]Case boards are also known as "crazy walls" see `crazywalls.tumblr.com` (accessed 12 October 2019) or "investigation boards". I have used crazy wall in the past (Benqué, 2018b) but choose the term case board here for its less derogatory connotations and to link conceptually with cognitive mapping as covered by Toscano and Kinkle (2015).

out" through a proposed diagramming practice. I discuss this process of diagramming further in the following section 3.4.

While case boards are physical objects, Srnicek (2012) suggests that cognitive mapping should make full use of computation if it seeks purchase on the systems of contemporary capitalism. In his view, 'computer algorithms, simulation models, econometrics, and other statistical analyses' (4) should be used as 'cognitive prostheses' towards the goals of Jameson's aesthetic proposal. With this in mind, I turn to the principles that inform my design of *DOTF*, a digital case board to support the process of excavating the history of algorithmic prediction's diagrams.

3.3.2 Diagrams of the Future

In practice, *DOTF* is a case board which has been built as a web application. Instead of the detective's board, push-pins and thread, data are stored in a graph database. *DOTF* represents "artefacts," namely examples of algorithmic prediction's diagrams, as nodes that are connected by various types of relationships and displayed along a time-based axis.[8] This is built from three main components: a database to store data and their relations as a graph, a back-end in the form of a canvas on which the contents of the graph database can be visually edited, and a front-end that provides a series of pages/views into the database content.

I have designed the back-end [fig. 3.8] to be a 'visual mode of knowledge production' (Drucker, 2014b). The interface is centred around creating, editing, and connecting nodes through drag and drop operations. Artefact nodes and their connections (in orange in the figure) form the visible part of the case board that is used for the front-end display and navigation. Reference, image, and quotes nodes are used to attach content to artefacts. Considering all content as nodes and relationships is a reflection of the case board as an epistemic instrument. This goes beyond a simple visual, skeumorphic gesture. The Neo4j graph database does not require a strict schema.[9] Node and relationship types, as well as their properties, can be

[7] I am slightly stretching the meaning of 'yarnwork.' Mackay (2017) uses it to refer to the case board itself, I use it to describe the type of work or activity that the case board affords.

[8] By contrast, "regular" databases generally store data as tables. In the case of relational databases links between data points (or rows) are stored in an extra table.

[9] A schema defines the structure of a database. Many database formats require a formalised description of the type of data and their attributes before they can be stored and managed.

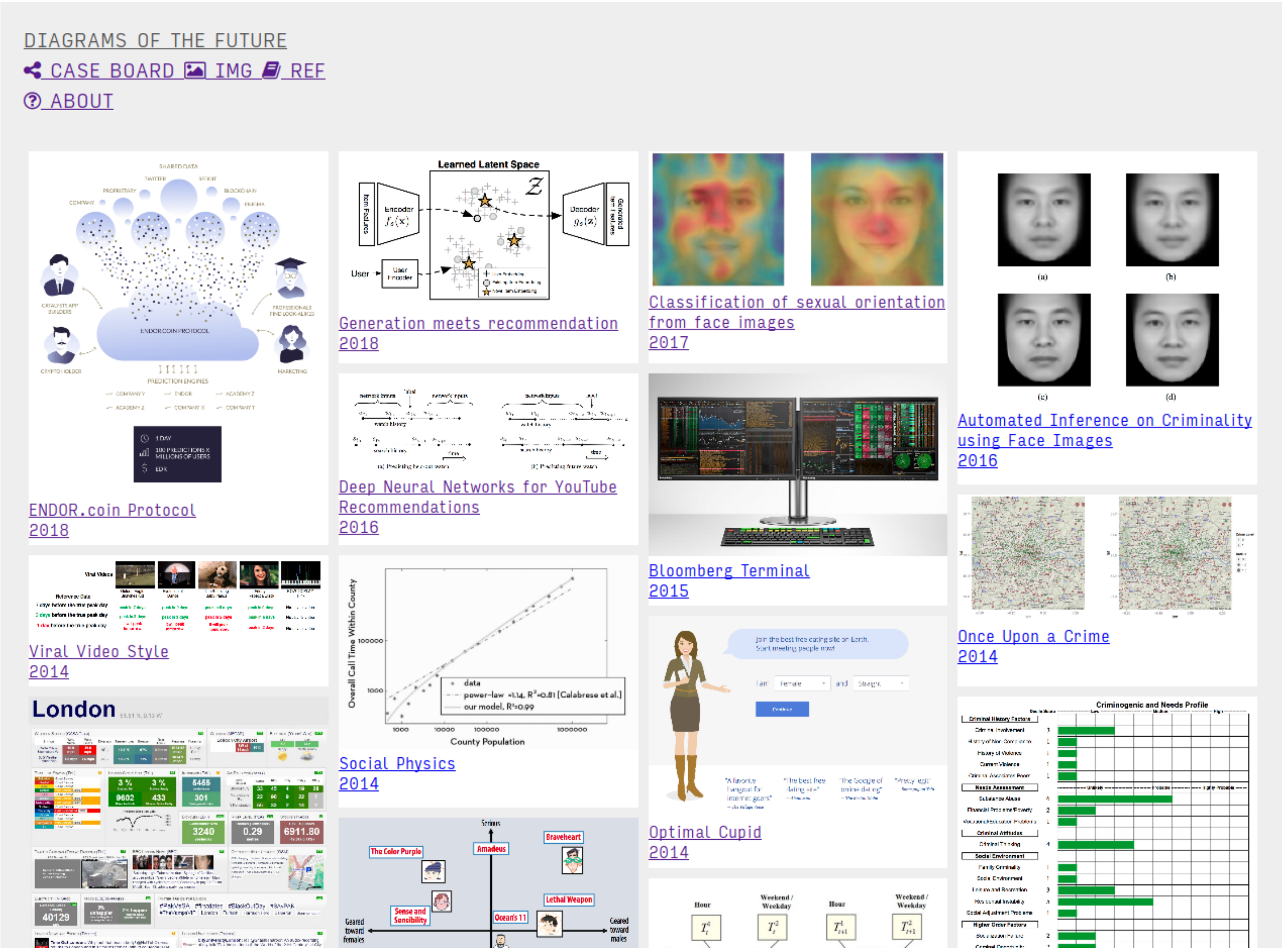


Figure 3.6: *Diagrams of the Future* homepage - http://dotf.xyz/.

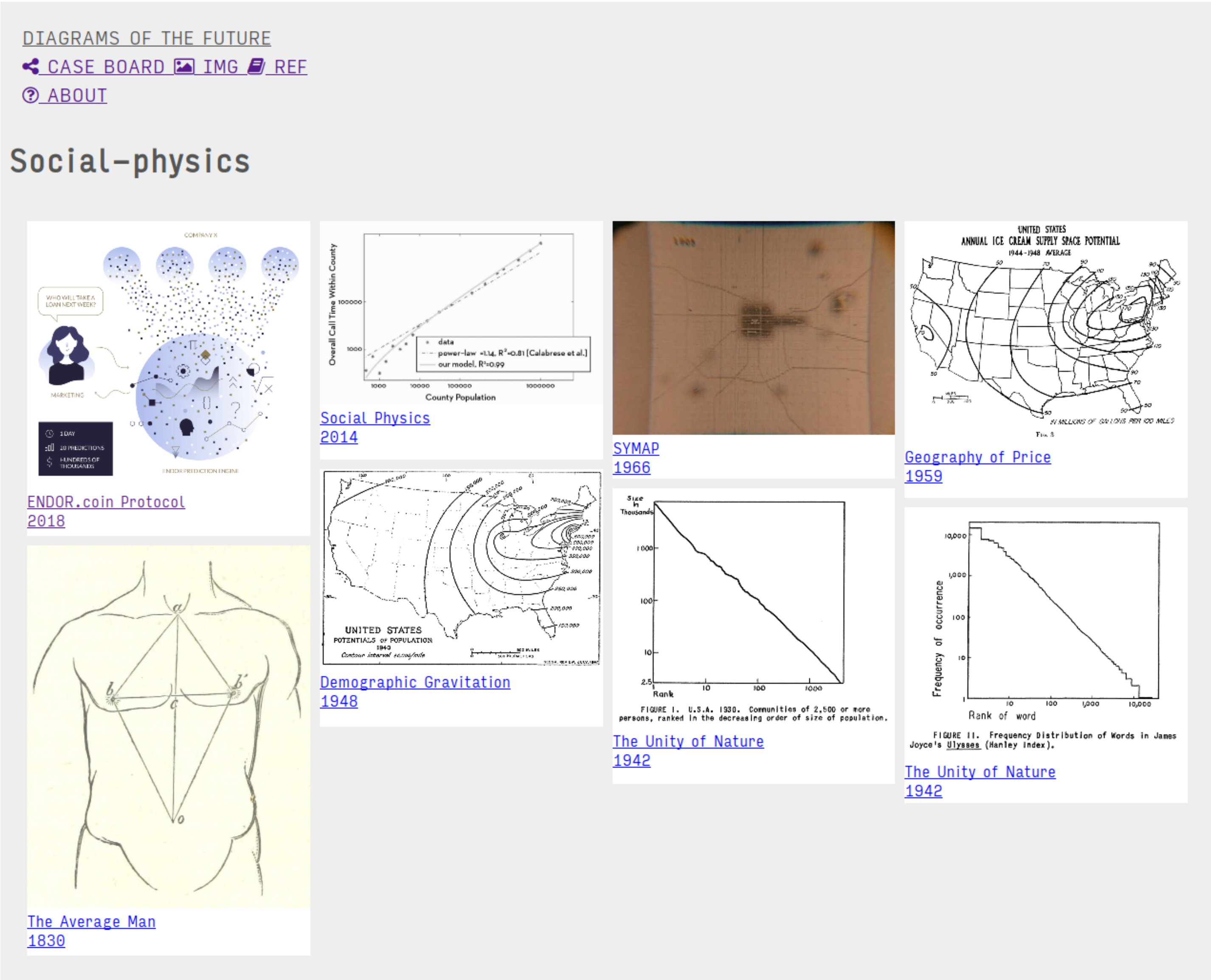


Figure 3.7: *Diagrams of the Future* Social Physics topos page - http://dotf.xyz/topos/social_physics.

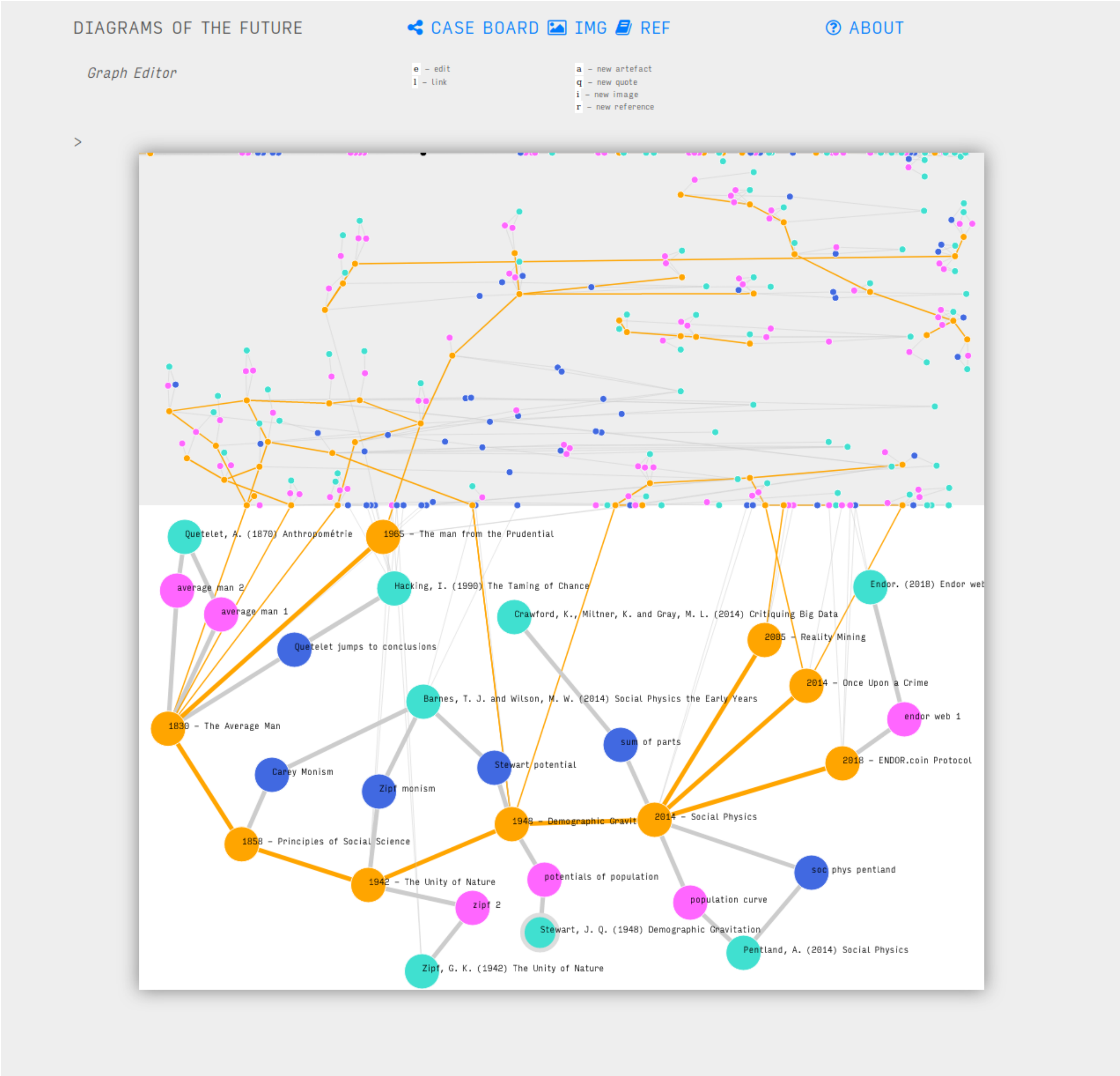


Figure 3.8: *Diagrams of the Future* editing interface showing the social physics topos staged in the editing area.

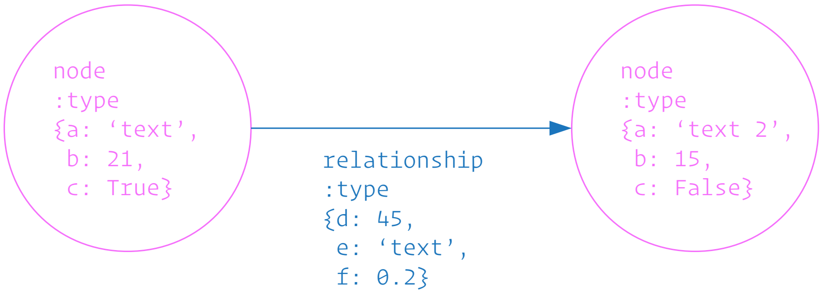


Figure 3.9: Nodes and relationships in Neo4j graph database.

created and modified as needed; in my case as my knowledge was accumulated and I started to relate to it in new ways. I elaborate on this trajectory in the next section 3.4.2. The flexibility of the graph database format also allowed me to work on building up and (loosely) modelling my knowledge independently of the ways it is displayed in the front end. I was later able to build new views into the same underlying graph, highlighting aspects of the database that I had not initially considered. These include: a home page where all diagrams are displayed on a grid layout [fig. 3.6]; individual artefact pages showing all the content attached to the artefact as well as its links; topos pages showing all artefacts bound by a specific topos or theme [fig. 3.7]; a reference page showing the full bibliography; and a case board page with a visualisation of the timeline with all artefact nodes, their relations, and a legend [fig. 3.1].

DOTF is built around a focus on producing knowledge diagrammatically, by entering data about singular artefacts, but more importantly by putting them in *relation* to each other. As different colours of thread or other markings summarise the detective's knowledge on a case board, the ongoing design challenge of *DOTF* has been to characterise and represent types of relations. Huhtamo's 'topos study' (2011a) lends itself to visualisation, if only a bit simplistically, as lines connecting artefacts. Navigating using these connections is a core activity within *DOTF*, with both viewing and "walking" the graph, by following topoï being a prominent option both on the artefact and case board pages. However, topoï are not the only type of relation I encountered. I felt other types

```
create (Neo:Crew {name:'Neo'}),
(Morpheus:Crew {name: 'Morpheus'}),
(Trinity:Crew {name: 'Trinity'}),
(Cypher:Crew:Matrix {name: 'Cypher'}),
(Smith:Matrix {name: 'Agent Smith'}),
(Architect:Matrix {name: 'The Architect'}),
(Neo)-[:KNOWS]->(Morpheus),
(Neo)-[:LOVES]->(Trinity),
(Morpheus)-[:KNOWS]->(Trinity),
(Morpheus)-[:KNOWS]->(Cypher),
(Cypher)-[:KNOWS]->(Smith),
(Smith)-[:CODED_BY]->(Architect)

match (n:Crew)-[r:KNOWS*]->(m)
where n.name = 'Neo' return n as Neo,r,m
```

Figure 3.10: Example Cypher query in Neo4j to 1) create a small graph of characters in the movie *The Matrix* and their relationships and 2) query this graph with the question "which characters know Neo?" Note that directional relationships are created and queried with an ASCII art diagram syntax `(a)-[:knows]->(b)` [Neo4j console https://console.neo4j.org/ accessed 23 Apr. 2019].

needed to be represented on the timeline, especially those established during the early exploratory phase before I encountered the very concept of 'topos study' Huhtamo (2011a).

A broader set of relationships was defined and re-defined as an iterative process as I built up my knowledge of the field, including: same person, same device, diagrammatic movements such as transposition (e.g. the transposition between astronomy and society at the root of social physics) or opposition (e.g. the conflicting visions of normality between Galton and Quetelet), and even singular relations that occur only once, such as the "A.I. Winter" between 1960 and 1986. For each of these, a special case was made in the visual language, that in turn has to be coded into the front end. These types of negotiations and adjustments highlight the tensions between the case board as tool for speculative knowledge production and the constraints imposed by web technologies and computational tools more broadly. It results in a highly idiosyncratic tool that is of little use in any other context than this research. Unlike many general purpose software solutions for tasks like mind-mapping, or image collections, *DOTF* is only suited to the specific task of excavating algorithmic prediction and reflects the speculative, incomplete and imperfect state of my own knowledge.

3.3.3 A patacritical graph

The use of a graph database seems well suited to the design of a case board, and more broadly to my diagrammatic focus in this research. *Relations* are an intrinsic feature of the data structure, rather than an addition as in traditional "relational" database formats. It is diagrammatic all the way down to the Cypher language, which expresses queries and operations as ascii-art diagrams [see fig.3.10]. However, the graph database was not readily co-opted into a time based, editable case board. I found that "pinning" nodes along a time axis runs against most of the state of the art in graph visualisation, which tends to either "animate" the whole graph, or to represent time along a single edge or node (Beck et al., 2014). Solutions to edit graph databases visually—for example through "drag and drop" with the mouse—were also, surprisingly rare. I surveyed a number of available tools[10] that were either proprietary, cumbersome to use, or were geared towards specific types of data and/or use cases.[11] The software used to make the case board would inform the knowledge produced with it. I realised after testing a number of "off the shelf" tools that they would not allow me to scrutinise this interdependence to a sufficient degree.

The idea of building *DOTF* "from scratch"[12] is not simply the result of minor inconveniences with available solutions. Rather, the decision to build my own tool was informed by a *patacritical* position which embraces computation but refuses generalisation (Drucker, 2009, 26). As part of a shift from '*digital tools* in humanities contexts' to '*humanities tools* in digital contexts' (25, emphasis in original), Drucker and Nowviskie (2004) propose that *speculative computing*, while still subjected to the constraints of computational logics, does not surrender its methodology to them. While it involves computational instruments, they are used for abduction, interpretation, and exceptions [see fig.3.11].

The principles of speculative computing are reified in the *Temporal Modelling* project, as proposed by Drucker (2009) and Nowviskie (2004)

[11] For example *nodegoat* is aimed at tracing the location of historical figures across time

[11] I tested *Arrows* (Jones), *Graffeine* (Browne, 2014), *Graphileon* (Graphileon), *neo4j-js-ng2* (Bruckert, 2018), and *nodegoat* (Bree and Kessels, 2013)

[12] DOTF still makes use of a number of pre-existing code libraries such as D3.js, the *Flask* Python web framework, and others.

Digital Humanities	Speculative Computing
Information technology/formal logic	'Pataphysics/the science of exceptions
Quantitative methods (Problem-solving approaches)(practical solutions)	Quantum interventions (Imagining what you do not know)(imaginary/imaginative solutions)
Self-identical objects/entities (Subject/object dichotomy)	Autopoiesis/constitutive or configured identity (Codependent emergence)
Induction/Deduction	**Abduction**
Discrete representations (Static artifacts)	Heteroglossic processes (Intersubjective exchange/discourse fields)
Analysis/observation (Mechanistic)	Subjective deformance/intervention (Probabilistic)

Figure 3.11: 'Attributes of digital humanities versus speculative computing.' Redrawn from Drucker (2009, 25, **emphasis** added.).

[fig.3.11]. Their reflections on this work show how abstract notions such as abduction can be implemented in practice, through a negotiation with digital tools. One key aspect is the priority given to visual and intuitive forms of knowledge production over technical implementation. They reverse the trend of considering visualisation as a cosmetic layer that comes after the ' "real work" of content modelling' has been done (Nowviskie, 2004, 100), and instead sketch out visual languages and interactions 'in advance of creating a database' (Drucker, 2009, 40) [fig.3.12]. The result is what they call a *PlaySpace* where visualisation is considered 'as an activity rather than a result' (Nowviskie, 2004, 101). Attempting to reproduce such a patacritical approach would go against its very definition as a practice focused on exceptions. Nevertheless, the positioning and description of the open ended *Temporal Modelling* project were instrumental in shaping my approach to *DOTF*, particularly around the time where I redesigned it to use a graph database [fig.3.13]. I took the patacritical work of Drucker and Nowviskie as an injunction—or a permission—to make a tool that was suited *only* to my specific investigation of algorithmic prediction, and to let the technical implementation reflect the "space" of my research, in all its incompleteness and uncertainty.

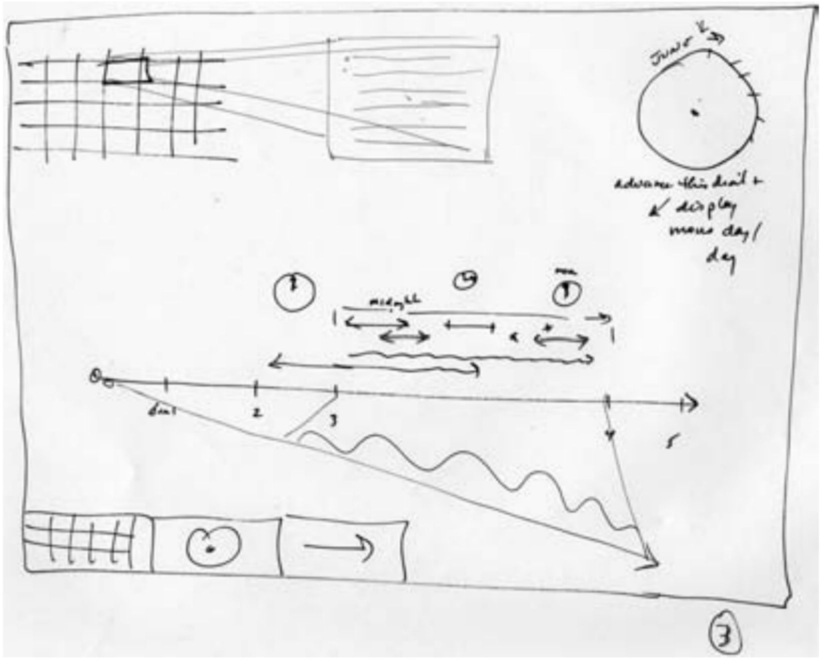


Figure 3.12: *Temporal Modelling* project sketch (Drucker, 2009, 57). (used with permission)

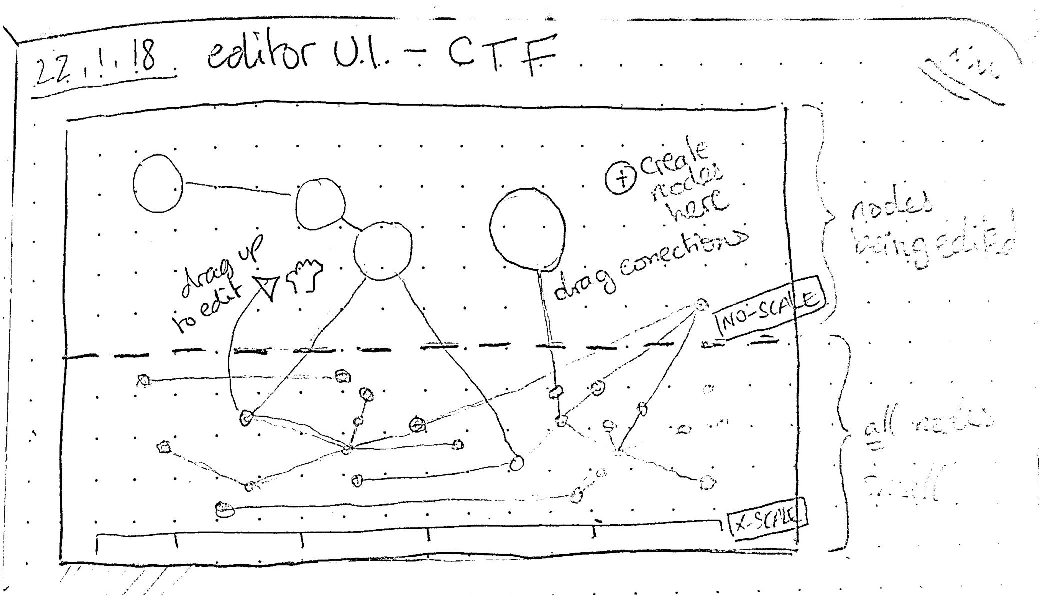


Figure 3.13: *Diagrams of the Future* sketch of the editing interface - 22 Jan. 2018.

3.4 Discussion

I now return to my research questions (RQ, RQ1, RQ2) and discussion criteria to reflect on the practice of sketching, prototyping, implementing, and iterating on *DOTF*, as well as its use as a research instrument. I discuss how *DOTF* addresses my research questions by providing a diagrammatic language to read/write the genealogy of algorithmic prediction. Working with diagrams provided purchase on my research area, affording a range of movements to guide my investigations. The successive versions of *DOTF* facilitated a shift from a systematic approach to the research towards a more abductive one where the instrument reflects the speculative and subjective nature of the investigation.

3.4.1 Starting in the middle

In line with Nowviskie's (2004) vision of digital instruments for humanities research, *DOTF* is an instrument to support excavating the history of prediction as an *activity* rather than a result. This is perhaps the key difference with case boards in detective movies, as mine never *arrives*, the investigation is never solved. Rather *DOTF* stays in a perpetual state of incompleteness and uncertainty. In relation to RQ1, this research activity is produced, or performed, through the movements and mechanisms of the case board and *yarnwork*.

The key movement afforded by the case board is *starting in the middle* (Sayers, 2017); to build up knowledge from a situated starting point while being attentive to the negotiations, indeterminacies, and surprises of making with digital media. Moving along threads, topoï, citation trails, biographies, hyperlinks, and so on, *DOTF* is an instrument for grabbing on to some point in the history of prediction and burrowing out from there. This reflects a 'tendency towards speculation or unlearning rather than proving or "wrangling" things with technologies' (Sayers, 2017, loc.244) as an 'intricate awareness of how making and scholarship start in the middle' (loc.273). This idea also resonates with a media archaeology approach, which is collapsing any linear notion of history and/or genealogy. Parikka has asked,

> Do you start with past media, like a 'proper' historian? Or from our own current world of media devices, software, platforms, networks, social media, plasma screens and such, like a 'proper' analyst of digital culture would? The proposition of this book is that **you start in the middle** – from the

> entanglement of past and present, and accept the complexity this decision brings with it to any analysis of modern media culture. (Parikka, 2012, 5, bold my emphasis added)

The design and implementation of the instrument itself can also be seen to have started in the middle. *DOTF* started as a printed poster with a linear timeline before being continuously developed, built, rebuilt, and adjusted to reflect the reframing of my research focus and questions (see appendix C for the full version history). With the refactoring of the project as a graph came a move from a rigid focus on "artefacts" to a more rhizomatic approach where all of the references, images, and quotes became explicitly part of the graph [figs.3.14 and 3.15]. The database became a more situated tool that reflects the provenance of, and gaps in, my knowledge. "Artefact" nodes are still used to order the front-end visualisation but only as focal points that point to other sources (e.g. images or text from primary sources, combined with critical analysis). My own presence as a researcher is also acknowledged, if only through a symbolic gesture, as my user account to edit the graph and administer the *DOTF* app is itself an entity in the graph-database and can be seen as a black dot in the editing interface (along the top edge in fig.3.8).

What emerges from this rhizome is a topology of the history of prediction. As I have discussed previously (section 2.2), processes of spatialisation are at the core of the operations of algorithmic prediction, through building *DOTF* I am exploring how the practice of spatialising its genealogy produces an alternative, critical account of what makes prediction today. By putting the "threads" of topoï 'into a conversation with each other' (Huhtamo, 2011a, 29) my aim has been to produce an alternative coordinate space that challenges narratives of continuity and rupture. For example by placing divination and gambling in the same space as machine learning, I suggest that they may have something in common, which is never advertised by the advocates of algorithmic prediction (Domingos, 2015; Agrawal et al., 2018).[13] Tracing topoï through time reveals continuities whereas narratives of "artifical intelligence" present these technologies as new and disruptive ruptures.

Like all diagrams, *DOTF* oscillates between systematisation and openness. The main drive in this research was to design and implement a speculative computational instrument, a diagram that remains open to interpretations, re-configurations, and reflects the situatedness of my knowledge. In section

[13]See also Cronon (1991) for an account of the early efforts to differentiate futures contracts from gambling on the corn futures exchange in 1860's Chicago.

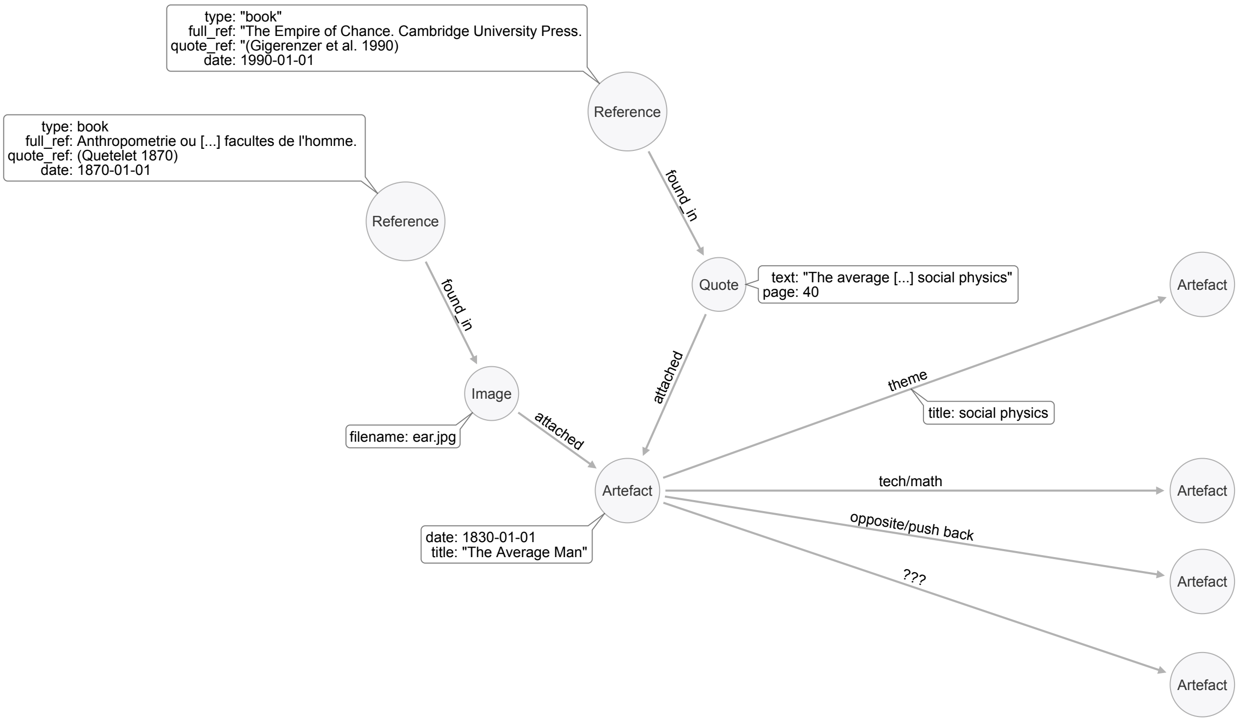


Figure 3.14: Structure of the graph database showing how reference, image, and quote nodes can be used to attach content to artefact nodes. Artefact nodes and their connections are displayed in the *Diagrams of the Future* front end and visualisation. diagram drawn with the *Arrows* tool (Jones).

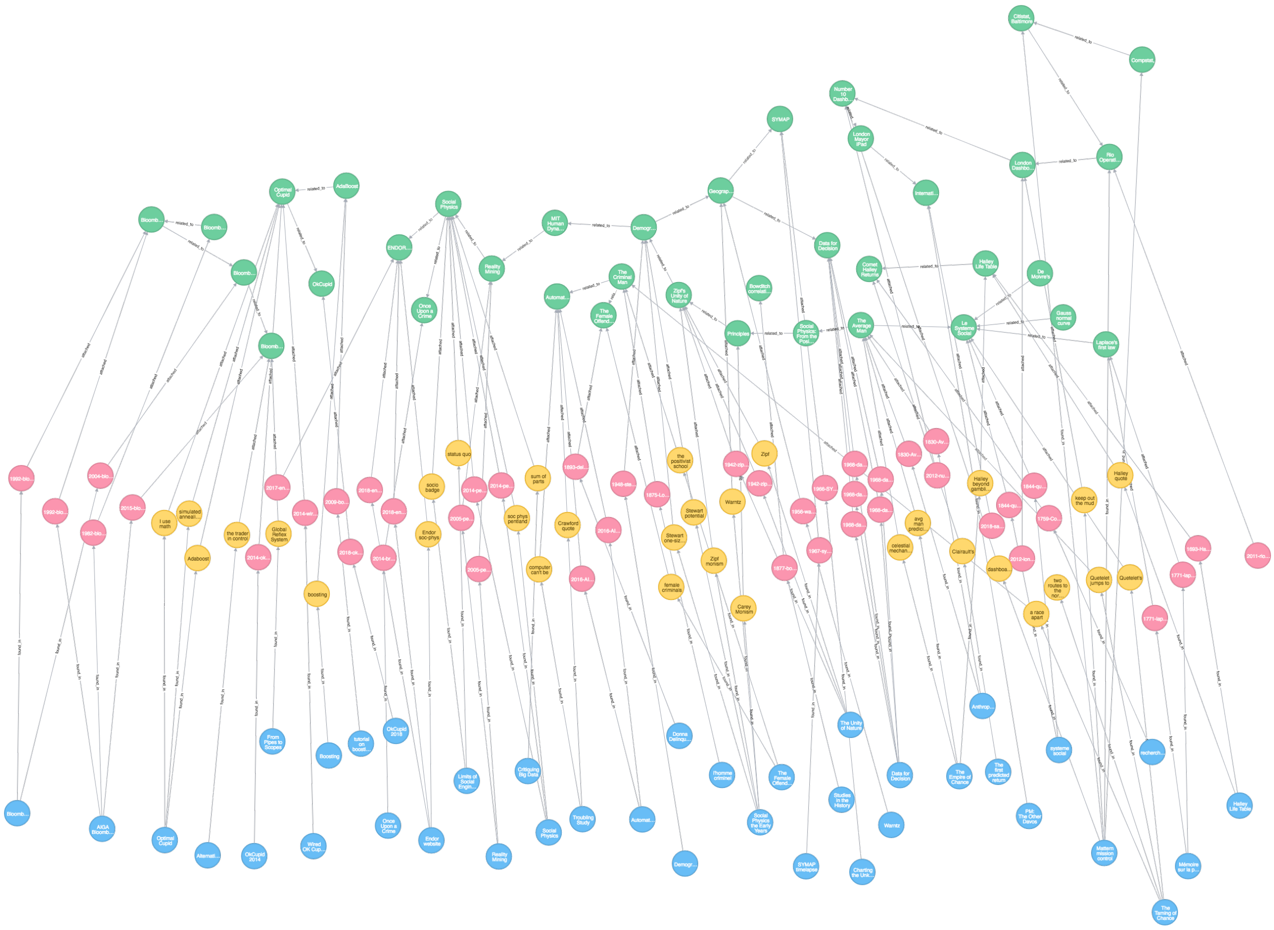


Figure 3.15: The full *DOTF* graph database as captured on 16th Feb. 2018, showing artefact nodes in green, images in pink, quotes in yellow, and references in blue.

3.4.3 I reflect on how this knowledge can also be transferable through publication. In many ways *DOTF* is a generative diagramming practice, it helped refine my research question by surfacing questions around selection criteria (what was to be included or not), and surfaced issues such as ordering through very practical means (e.g. in the x-axis of the visualisation, see section 3.4.2 below).

DOTF acts as a generative foundation for building an argument through this PhD project. The practical skills I developed with graph-databases were re-purposed to map YouTube recommendations in the next project *Architectures of Choice* (discussed in ch.4), and one of the threads followed here led me to investigate almanac publications (discussed in ch.5). An honest view of the oscillations at play in *DOTF* must of course consider the counterpart to this generative aspect; the systematising, pinning down, and even tedium involved in this project. As Toscano and Kinkle (2015) note, the case board in *The Wire* is inseparable from 'seemingly ubiquitous paperwork' (loc.298.8), likewise programming and re-programming *DOTF* involves considerable amounts of "systematising" work. I would estimate roughly the same amount of time was spent engineering and maintaining the intrument itself than conducting research *with* the intrument. Although *DOTF* is a visual instrument, making use of drag and drop, drawing connections, and so forth, data entry still requires a considerable amount of form-filling. The practice of entering new nodes and managing connections is still quite a bureaucratic process for which I had to develop strategies such as gathering all data for one session into a text document before pasting them into the interface. I have been continuously improving this process, for example by building an auto-fill functionality for references using my Zotero library. Such features crystalise the negotiation between researching with instruments and developing the instrument itself—with the added possibility that any addition to the code might break existing data or functionality and render the instrument unusable. With instrument building explicitly used as a method, it is impossible to separate it from the "actual research." Nevertheless, an ongoing trade-off has to be negotiated between the primary focus of the research, in this case algorithmic prediction, and web-app development.

3.4.2 The x-axis as an abductive arc

In relation to RQ2, *DOTF* can be seen as a stage for *yarnwork*, a performance or dramatisation of the production of knowledge (Mackay, 2017). The design and programming of this digital version of the TV detective's case board has provided opportunities to unpack, and question, the way I was conducting research—highlighting tensions and negotiations between my position as a researcher, computational tools, and the history of algorithmic prediction. Specifically it led me to articulate the diagrammatic mode of knowledge production at play in *yarnwork* as abductive, in contrast to the inference that algorithmic prediction relies on. This in turn raised questions and doubts about the validity of my research.

> Rather than accurately diagramming reality, the yarnworker is always in danger of projecting onto the blank wall his own preoccupations, vendettas, personality flaws, and, when things get really murky, sheer delirium. (Mackay, 2017)

Detective fiction, and by association the case board, is arguably a product of the modern scientific world that 'reflects the predominance of empirical evidence and rational deductive methods' (Mackay, 2017, citing Boltanski). I am, however, missing a key piece in the narrative arc it is usually employed in: the 'threshold' of resolution. Instead *DOTF* remains at an earlier stage, a tension between a systematic diagramming relating "pieces" together, and the fact that they still refuse to fully 'add up.' In my case the "arc" of the research centred around issues of ordering the case board itself, as materialised in very practical concerns such as which variable to use along the x-axis of the visualisation.[14]

The *DOTF* project extended throughout the full duration of this PhD research. It followed the definitions and re-definitions of my research focus, questions, and methodology, going through multiple iterations. It started as a linear timeline with a few artefacts [fig.C.2] and a simple website using a pre-existing timeline visualisation [fig.C.1] in the spring of 2016. I started developing my own visualisation later that summer, as a "minimum viable product" presented at the 4S/EASST conference [fig.C.3] (Benqué, 2016) and at the RCA School of Communication's research work in progress show. Based on feedback from both events I further developed the project that winter [fig.C.4], before focusing on other aspects of the research. I came back to it after nearly a year (winter 2017), after discovering graph databases and re-centering my research questions around *diagrams*. This resulted in a complete re-write of the code and the introduction of the visual editing interface [fig.3.8]. This was then further developed and

[14] The orientation of the case board visualisation changed twice during the project, as seen in the evolution in appendix C. I am referring to the x-axis in the latest version, that is to say the axis that is not time-based.

refined, especially as I re-framed my research using media archaeology and from there started to think in terms of genealogies.

From the very beginning, I ordered artefacts along a time-axis, but the question of ordering on the other axis came to crystalise the negotiations and tradeoffs between *pinning down* the genealogy of prediction and *opening up* new connections and multiple readings. The layout problem of how to order things visually became a mode of confronting much deeper questions about what kind of knowledge was being produced through *DOTF* and acknowledging my own subjectivity and interpretations. From the straight line of the beginning, the first real confrontation with this problem was building my first visualisation prototype for 4S/EASST. Here the artefacts were ordered along 4 pre-defined themes that were defined along activity domains and physical scales[15] I had already identified movements between different domains as a point of interest (for example the transposition of mathematical techniques in social physics covered in section 3.2.3) so a "knotting" mechanism was built in to allow one artefact to belong to two themes at once, forming a juncture. This approach felt like imposing too much of an arbitrary constraint on the project, especially as themes were defined at the start, the point where I knew the least about the project.

From there started a process of un-ordering [fig.3.16], a series of attempts to loosen my initial idea of "classifying the classifiers" to find a more organic way of growing the timeline and to reflect the state of my knowledge. This started by assigning an arbitrary scale score to artefacts to position them along the y-axis [fig.C.4]. The move to the graph database and complete refactoring of the project was a decisive shift, as the relations between artefacts became the main focus, and I did away with ordering altogether. However this did not solve the issue of positioning nodes on a two dimensional canvas; one dimension was now handled by a randomised force layout,[16] a common solution to the layout of network visualisations. This added distracting animations, and my negotiations with the settings did not result in an aesthetically or conceptually satisfactory result. Finally the last iteration comes back to the relations

[15] See Benqué (2016). *Heavens* was about astronomy, climate and planetary scale, *Masses* about populations, elections, and public opinion, *Gambles* about economics and finance, and *Fates* about genomes and bodies.

[16] The force layout in D3.js starts by giving all nodes a random position and subsequently adjusts their positions through a simulation of forces (e.g. how rigid or elastic are the links, is there gravity, do nodes attract or repel each other, and so on) the layout gradually settles to a stable state where the node can remain draggable.

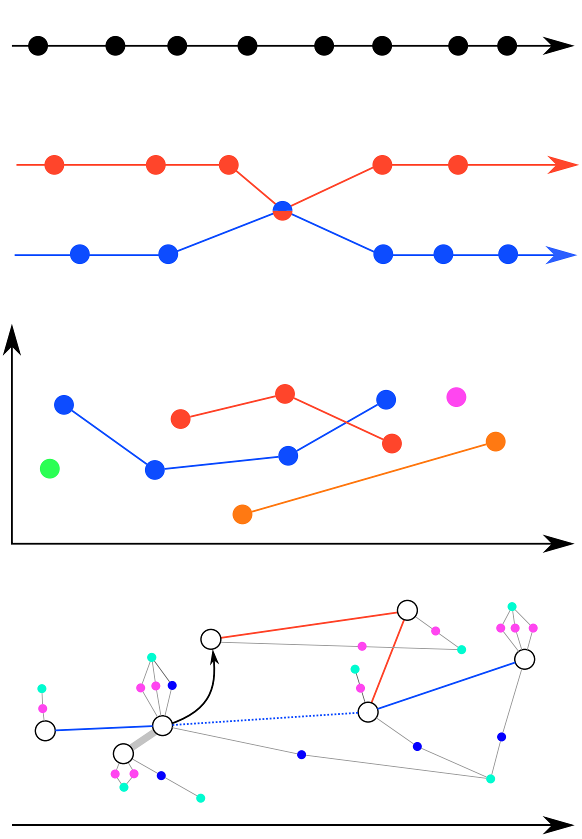


Figure 3.16: Iterations of timeline threads in *Diagrams of the Future*. See appendix C for images of each version.

between artefacts as the driver of layout. Each theme or thread is assigned a base position, and nodes are laid out from there using only a collision force to prevent form overlapping in the manner of a 'bee-swarm' plot (Eklund, 2016).

3.4.3 Publishing the *DOTF* archive

In March 2018, I presented some of my work on *DOTF* to an audience of technologists, artists, designers, and others in London as part of an event titled *Colossal Dust: practices of obsession and investigation* (Benqué, 2018b). After my presentation, a member of the audience asked about the purpose for the project as a publicly accessible website, rightfully casting doubt on the fact that it would be intelligible to anyone but myself. At the time the website homepage presented a set of dots and lines, bouncing around under the effect of a force layout. While I was still focused on the tool itself at that point, this question echoed doubts I was already having about the legibility of the project, and about the need for legibility in the first place. This stayed with me as I developed my broader research position and refined my research questions. In line with the aim of producing publications, I persisted in considering *DOTF* as a form of publishing, an open archive of sorts, however idiosyncratic. I took steps to render it more legible and potentially useful to others in a last round of revisions before submission. While the case board is a highly personal artefact—embodying the specific knowledge of an investigation team or individual detective—it is activated through its presence on screen as a plot device. Similarly, by making *DOTF* accessible online my case board becomes part of a performance of research, which activates it beyond my own personal use as a research outcome in its own right.

My point of reference here are art practitioners such as RYBN, who regularly produce bibliographies, archives, or other documentation as part of their work. Projects such as *Cybernetics Gamblers Hall of Fame* and *Human Computers* (RYBN, 2019) follow a media archaeology approach and place as much value on excavating research, scholarship, and forgotten oddities than on the final artwork. I take this as an important contribution, importing some elements of academic work to creative practice, sharing one's sources allows others to locate where the work comes from, follow them, use them in different ways, perhaps reaching other conclusions. I have myself drawn from RYBN's work and archives during this research on more than one occasion. One way to view *DOTF* as part of this thesis is, in a similar way to RYBN's archives, as an extended, visual, literature review that ends up as an outcome in its own right, an integral part of the final work. This process is not just about accumulation, but an editorial practice of 'filtering and amplification' (Gilbert, 2016, 11) that is in itself a creative, and potentially critical act.

Publishing imposes constraints and expectations that might conflict with the patacritical approach outlined earlier. If the instrument is tailored to my specific approach and knowledge, how could it be readable by anyone else? Seeing a detective's case board may not make sense to an outsider viewing it out of context. However, following the conversation with the audience member and my own ideas about the project, I decided to include some element of readability into what the *DOTF* needed to achieve. The last refactoring of the project aimed to address these questions, starting with renaming the project *Diagrams of the Future* rather than *Counting the Future* to better reflect the focus of my research. These developments include: relegating the graph view to a secondary page and preferring a more generic "tumblr" like layout for the landing page; giving nodes and topoi their own pages with links that can be bookmarked or shared. I intend to use these links to publish the archive through social media in future developments. Finally, as with any publishing project the question of temporality and maintenance should be raised. In this case this is a long-term commitment to maintain the site, as it will continue to serve as the base from which to formulate future research and projects.

3.5 Conclusion

In this chapter I have presented *Diagrams of the Future*, an archive of predictive diagrams in the form of a digital case board. I considered the genealogy of algorithmic prediction as a diagram of diagrams. This provided a common ground between research and practice (RQ1) as I mapped this genealogy using a graph database, following the threads of topoï (Huhtamo, 2011a). I characterised this process as a form of *yarnwork* (Mackay, 2017), itself a *trope* of detective/investigation movies where knowledge production is dramatised and performed. This activity can be described as diagrammatic sleuthing, with movements (RQ1) such as "starting in the middle." It oscillates between systematically ordering knowledge on the case board and subjective conjectures (RQ2) that refuse to completely 'add-up,' and contain the researcher's situated interpretations. Finally I considered the graph-based archive as a form of publication, a research outcome in its own right that provides a foundation for the other projects in this research while also offering a multiplicity of readings to other researchers or practictioners.

This chapter describes the first contribution to knowledge in this research, the *case board* as a diagrammatic mode of investigation into algorithmic prediction as 'media remembered' (Blegvad). This draws from a media-archaeological view of the genealogy of technological artefacts, and combines it with notions of 'cognitive mapping' (Jameson, 1988; Toscano and Kinkle, 2015; Srnicek, 2012) and 'yarnwork' (Mackay, 2017). In practice this translates into a novel use for graph databases, purpose-built through, and for, my practice-based research approach. This is materialised in a web application that provides a new way to edit graphs visually as a kind of diagramatic content management system. The front end of the application displays the contents of the graph through various views, including a case board visualisation that displays artefacts and their relations across time.

The case board, while computational, stays at a distance from algorithmic prediction. It is a canvas to track its developments but never touches actual operations. In the next chapter, my position shifts as I move closer to one predictive algorithm and diagram its contours through much more direct means.

Please see appendix D for the practice submission related to this chapter, including: code repository and other supporting material.

Chapter 4

Traces

4.1 Introduction

In July 2016, a headline in *Wired* magazine read 'This AI learned to predict the future by watching loads of TV' (Burgess, 2016). It described a study where researchers used neural networks to predict handshakes, hugs, kisses, and high-fives in American sitcoms with 43% accuracy (Vondrick et al., 2015). The contrast between the headline and the actual outcomes of the study highlights the gap between inflated claims around "AI" and a reality that is far less exciting. While this AI has not learned much about the future by 'watching loads of TV,' algorithmic prediction does play a key role in delivering a staggering amount of video content through a much more familiar and mundane medium: recommendations of "what to watch next." In this way prediction does shape the future but through small, mundane, and short-term nudges.

> Algorithmic recommendation has settled deep into the infrastructure of online cultural life, where it has become practically unavoidable. (Seaver, 2018, 14)

In this chapter, I move to probing algorithmic prediction as 'media observed' (Huhtamo and Parikka, 2011, 55), as it operates in a contemporary recommendation system, the YouTube platform. I discuss instruments developed as part of *Architectures of Choice Vol.1: YouTube* (*Arc-choice*), and reflect on the challenges and possibilities of *seeing* such systems through another ubiquitous medium: data visualisation.

This work brings together two strands of my research that were initially developed separately. The first is a collaboration with design theorist Betti Marenko[1] on speculative diagrams and possible ways to reclaim algorithmic prediction. We presented an initial position for this work in spring 2018 at the *Art, Materiality and Representation* conference in London (Marenko and Benque, 2018). The second is a short[2] practice pilot project presenting five initial attempts to map YouTube, this was a collaboration between *Supra Systems Studio* (SSS) at the London College of Communication and British Broadcasting Corporation's Research & Development (BBC R&D) as part of SSS's inaugural exhibition *Everything Happens so Much* (*EHSM*) for the London Design Festival 2018 (September/October). The two came together as Betti Marenko and I took the YouTube work as a first case study of our speculative diagrams approach, and submitted it to the *Research Through Design (2019)* conference in Delft, the Netherlands (RTD). This involved substantial developments on both the theoretical and practical sides of the project. Our collaboration was articulated primarily along a theory/practice axis, with Betti Marenko contributing a Deleuzian design theory perspective, and I contributed practical explorations and some technical knowledge of machine learning. We shared a critical disposition towards algorithmic prediction, and the diagram as a device, both theoretical and practical, to focus our conversation.

I start by discussing YouTube recommendations as a capture apparatus. While the funnel shape of the recommendations system is presented by its engineers as a way to distil content out of a huge corpus, it can also be characterised as a trap that relies on *traces* to keep users watching. I discuss the possibility of mapping this apparatus and the promises of transparency that they imply.

I then describe our attempts to map YouTube, confronting these issues through practice. I present the first experiments produced for the *EHSM* exhibition, and how these were developed for the *RTD* conference. This was a major shift in our approach as we started to produce out own *traces*, informed by a conversation between the theory and practice of diagrams.

Finally I discuss the project in line with my research questions, describing its abductive "arc" going through the peer review process. I describe *traces*

[1]Dr. Betti Marenko is a reader in contextual studies at Central St. Martins, University of the Arts London.

[2]The project had a time budget of about a week, in which I produced a series of experiments with a graph-database and a variety of visualisation tools, see the `02-Viz` section of the code repository.

not as part of algorithmic systems of capture, but as sources of conjectural knowledge, in other words as speculative diagrams.

4.2 YouTube recommendations: the funnel

In this section I characterise YouTube recommendations as a form of control diagram in the shape of a funnel. I discuss problematic claims that this funnel can be seen/known through data, and that this provides a form of transparency that would address the problems of recommendation systems. I engage with these claims through my practice in the next section, and propose an alternative mode of knowing through data in the following one.

4.2.1 YouTube as capture apparatus

According to YouTube engineers, YouTube's recommendations are 'one of the largest scale and most sophisticated industrial recommendation systems in existence' (Covington et al., 2016). The purpose of algorithmic recommendation, in this case, is guiding nearly 2 billion users as they navigate an enormous corpus of video content, growing at the rate of 400 hours a minute. To do this, the system acts like a 'funnel' [fig. 4.2] by combining two deep neural networks to 1) generate 'candidates' out of the large corpus of videos, and 2) rank these candidates and present most relevant to the user (Covington et al., 2016). The first step is a "rough cut," selecting candidates out of millions of videos, using collaborative filtering[3] with 'coarse features.' The second pass is much more fine-grained, and orders the candidates by taking into account a 'rich set of features' to ensure the user is presented with 'personalized and engaging' content. At both stages, a prediction of 'watch time'[4] drives the selection and ordering of video candidates (Covington et al., 2016).

For all their technical sophistication, YouTube recommendations have been criticised for their detrimental cultural and political effects (Tufekci, 2018;

[3]Collaborative filtering makes predictions based on similarities between users, using a matrix of preferences to infer what similar users "may also like." (see Seaver, 2012)

[4]"Watch time" has replaced "click through rate" as the metric of engagement YouTube seeks to optimise. The latter was said to reward "clickbait" such as suggestive titles or thumbnails. "Watch time" designates the likelihood that the user will not only watch the given video in its entirety, but that they will continue watching YouTube. 'Now when we suggest videos, we focus on those that increase the amount of time that the viewer will spend watching videos on YouTube, not only on the next view, but also successive views thereafter.' (Meyerson, 2012)

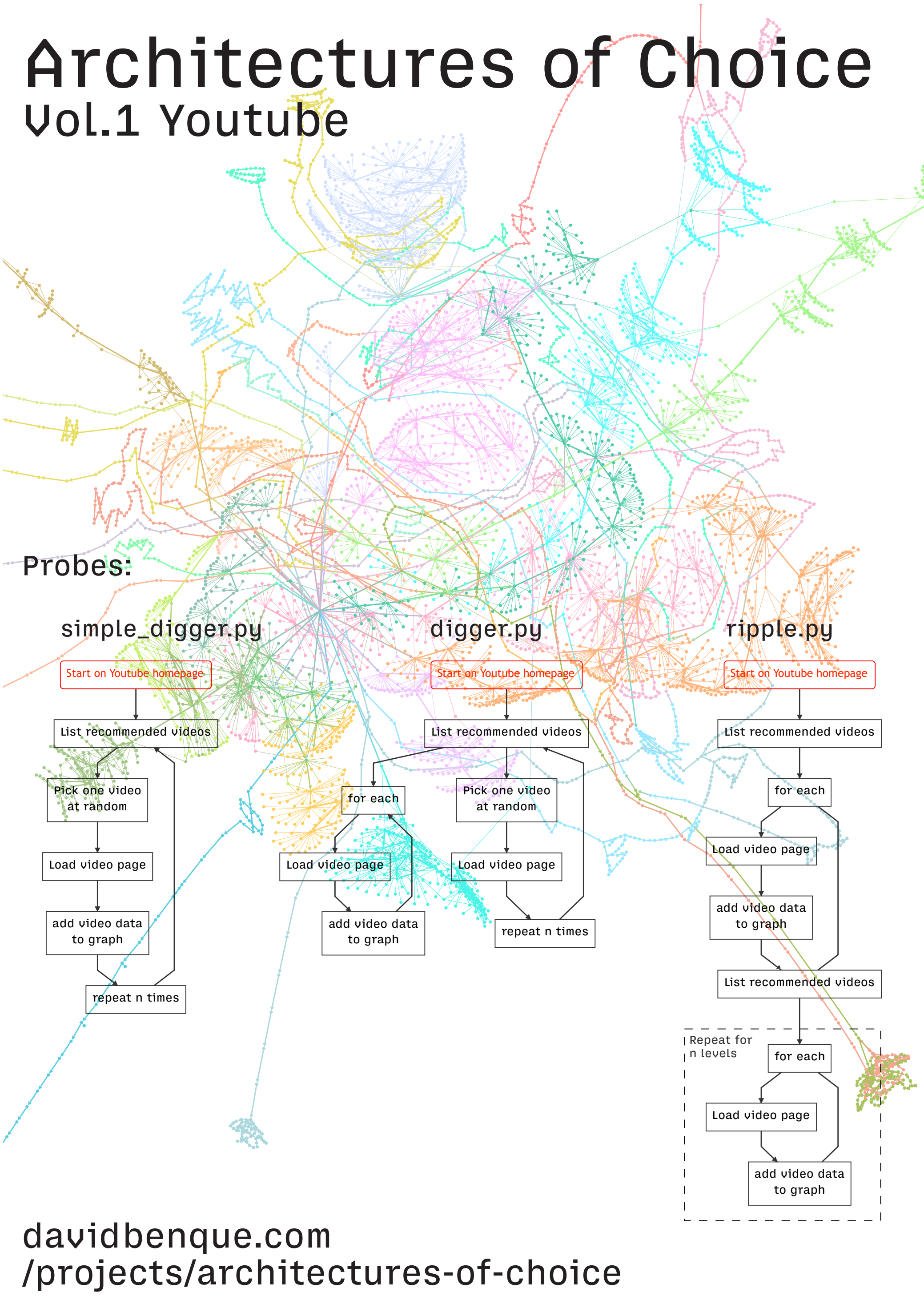


Figure 4.1: *Architectures of Choice Vol.1 YouTube* poster produced for the *Everything Happens So Much* exhibition (Sept. 2018).

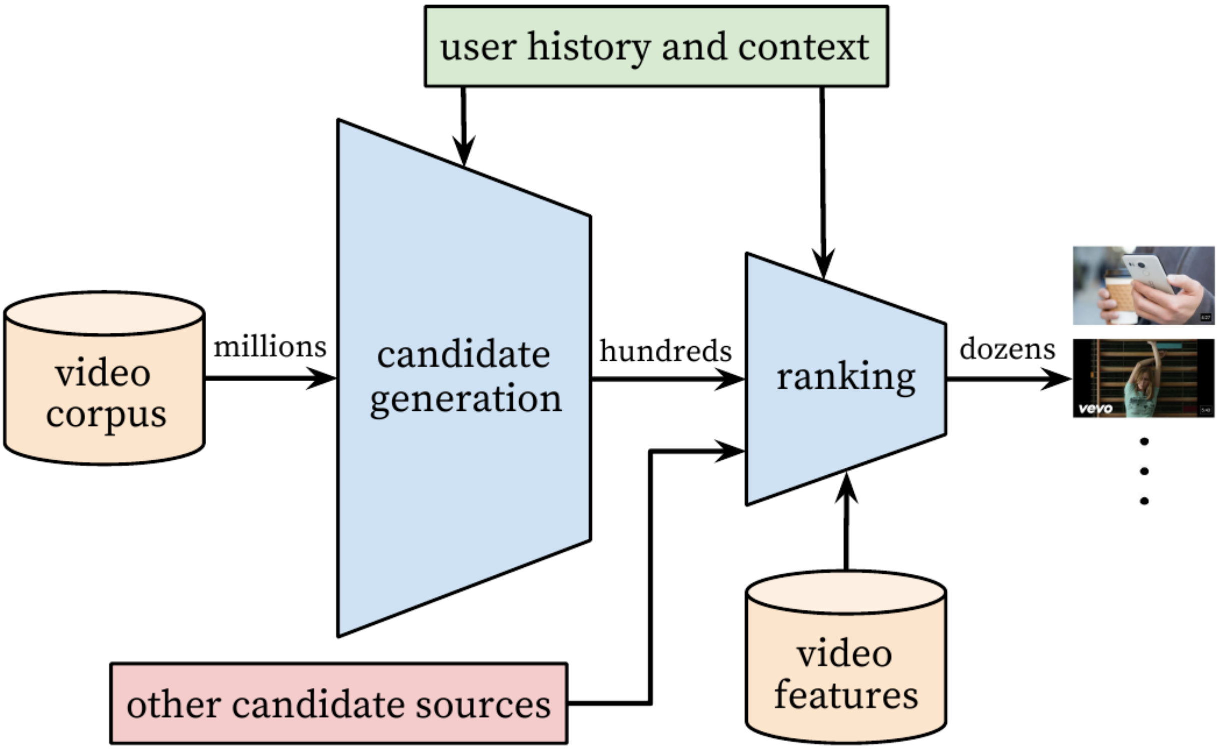


Figure 4.2: YouTube's 'Recommendation system architecture demonstrating the "funnel" where candidate videos are retrieved and ranked before presenting only a few to the user.' (Covington et al., 2016, 192).

Bridle, 2017; Lewis, 2018a). The 'funnel' pushes viewers towards increasingly extreme versions of whatever argument they might be watching, acting as a 'great radicaliser' (Tufekci, 2018). These logics have been particularly well mobilised by conservative and right-wing commentators and creators, who tailor their content to the platform's criteria and have gained massive followings in the process (Roose, 2019). Recommendations play a key role in the propagation of conspiracy theories such as "Flat Earth" with many believers pointing to YouTube as the gateway to these ideologies (Landrum, 2019; Paolillo, 2018). The system is also easily gamed by networks of automated accounts that produce computer-generated and/or "fake" news content (Albright, 2017), or take advantage of the popularity of content aimed at children to promote disturbing and violent videos (Bridle, 2017).

These issues illustrate the need for a critique of algorithmic prediction as a form of prescription—as I have discussed in Section 2.3. The promise of 'helping' users 'discover personalized content' (Covington et al., 2016) implies that the algorithm selects content that the user already likes. Many of these cases show that it produces and re-shapes user-preferences instead. YouTube recommendations aim to maximise engagement metrics such as predicted watch time. As Arnold (2016) argues in her study of the film streaming company Netflix, 'engaged' also means 'subscribed' which is the measure that the company seeks to maximise while the term "recommendation" 'maintains the perception of choice' (96).[5] Arnold frames her critique using algorithmic governmentality (Rouvroy, 2013) and the algorithmic production of identities (Cheney-Lippold, 2011) to argue that recommendations effectively *remove* choice and *produce* stereotypes out of users. These logics reward controversial content and entrench pre-existing positions. The apparatus[6] of YouTube recommendations modulates and re-configures the behaviour of its audience based on fragmented and granular behavioural data. This echoes Deleuzian notions of control that take hold by fragmenting individuals into 'dividuals' (Deleuze, 1992, cited in Cheney-Lippold, 2011, 169).

[5] Netflix and YouTube differ in their subscription models. However, many of the arguments made by Arnold about Netflix—especially ones drawing on algorithmic governmentality—are applicable to YouTube.

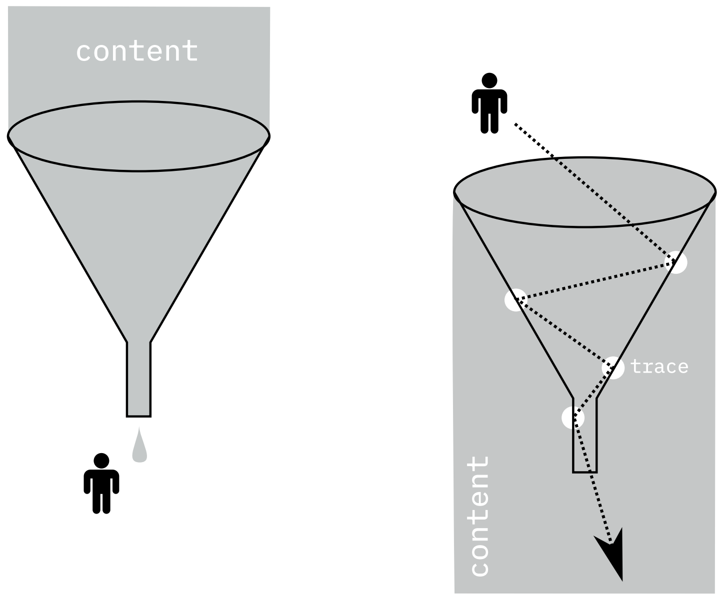
[6] Betti Marenko and I use Giorgio Agamben's notion of apparatus in our paper: 'anything that has in some way the capacity to capture, orient, determine, intercept, model, control, or secure the gestures, behaviours, opinions, or discourses of living beings' (Agamben, 2009, 14, cited in Marenko and Benqué, 2019). This notion is also used extensively by Masure (2014).

The 'funnel' of YouTube recommendations may be less of a distiller of personalised content, and more like a *trap* that the user walks/watches into [fig. 4.3]. Seaver (2018) brings the rich anthropology of traps and 'captology' to bear on recommendation systems, and moves beyond simply 'policing the boundary between freedom and coercion' (4) that is the focus of many other criticisms of algorithmic prediction.[7] Seaver highlights the central role of *traces* in the process of algorithmic entrapment. He describes the industry's move away from predicting 'explicit ratings' such as star scores and towards 'captivation metrics' based on the implicit data found in interaction logs.[8] Rather than trying to accurately predict the rating a user would give to a film or series, this approach uses signals produced by users as they navigate the site—for example whether or not the video was watched until the end, which buttons were clicked, and so on. These more fine-grained data are much more abundant. In this way, more aspects of the user experience can be measured, with applications designed to maximise the number of interactions. The more active users in the system produce more traces that in turn provide them with more recommendations. Traces are also perceived by industry as more 'truthful' (Seaver, 2018, 12), reflecting what the users "really want" rather than what they say in explicit ratings—echoing the positivist views I discussed previously in section 2.2.3.

In summary, recommender systems demonstrate an instance of algorithmic prediction as a control diagram. They produce identities and categories, trapping users in self-reinforcing loops of data production. These diagrams are formed by the relations between *traces*, implicit data-points generated by users as they interact with an application or service, and interpreted by the service for signs of "satisfaction," meaning retention and sustained engagement (Seaver, 2018, 13). The next section will look more specifically at the promise of mapping as a counter-strategy to expose the inner workings of this apparatus.

[7] With this, Seaver nuances criticisms of algorithmic prediction based on Deleuze's *control* and Foucault's *governmentality* (e.g. Rouvroy (2011); Cheney-Lippold (2011)). He focuses specifically on "redeemed" designers of addictive algorithmic systems such as Nir Eyal or Tristan Harris who have turned into vocal critics of these systems. In their counter-visions focused on ' imagining an escape' from traps, these figures do not challenge the 'behaviorist frame' of captology they helped implement.

[8] See for example Amatrian and Basilico (2012) on Netflix moving 'beyond the 5 stars.'



(a) (b)

Figure 4.3: Two views of the YouTube recommendation system as a) a funnel that distills personalised content (Covington et al., 2016), or b) a trap (Seaver, 2018).

4.2.2 Map the trap

Revealing a system's inner workings, and making it accountable for its negative effects, is an appealing counter-measure against an apparatus of capture and entrapment. This promise has been gaining traction, especially since the 2016 US election that saw the so-called weaponisation of platforms for the spread of misinformation (Subramanian, 2017; Albright, 2016), and data collection for the targeting political ads (Cadwalladr and Graham-Harrison, 2018). Strategies for increasing the accountability of algorithmic systems include auditing (Sandvig et al., 2014), and reverse engineering (Diakopoulos, 2015). However, these have to contend with various form of opacity (Burrell, 2016) and generally have to operate from outside the systems they claim to scrutinise.

I am specifically interested in attempts to *map* YouTube, as part of this broader discourse on transparency and accountability, and in the way these mapping exercises produce knowledge through data visualisations. In research/academic settings, two examples stand out. In her study of the reactionary right's use of the platform, Lewis (2018b) maps the network of co-appearances between different personalities hosting each other on their channels. This is the result of in-depth data collection done "by hand" in order to surface connections that are not visible through automated means. The map follows the codes of network visualisation with a force directed layout, and uses analysis measures such as betweenness centrality and closeness centrality to highlight influence and connections (10). A couple of 'paths' are also highlighted (11) showing how guest appearances between channels create stepping stones to more and more radical ideology.

In another setting, Rieder et al. (2018) use purpose built tools to query data from the YouTube API (Rieder, 2015). They visualise the ordering of YouTube search results and their 'morphology' as they change over time for specific topics. They characterise 'patterns of change' on a spectrum from 'stable' to 'newsy,' using the *RankFlow* visualisation tool (Rieder and Simon, 2016) as the basis for qualitative analysis of the social, cultural, and political work of algorithms. These examples are both focused on specific aspects of the YouTube platform, either a small subset of interest (the reactionary right from conservative influencers to white supremacists) or a methodological approach to observe the capture diagram from a specific angle (the ordering of search results and their patterns of change over time). They produce data visualisations to reveal a particular side of the platform's mode of operation, whether a specific network of actors or a mode of ordering. This happens in the context of in-depth qualitative research that acknowledges the methods, and limitations, of the collection and visual display of data. They contribute to a 'countervisuality' (Sandvig, 2014) that support critiques of algorithmic work.

Another map of YouTube embodies a more simplistic and problematic approach to the promise of revealing the inner workings of YouTube, one that has received attention in mainstream press. Led by a former Google employee, the *AlgoTransparency* project is said to provide 'the world's first window into YouTube's opaque recommendation engine' (Lewis, 2018a). Chaslot (2016) started this work prior to the 2016 US election, where they monitored YouTube by collecting data on 8,052 YouTube videos in the run-up to the vote and found that recommendations were 'highly skewed toward content critical of the Democratic nominee [Hilary Clinton]' (Lewis, 2018a). In the years since, they have taken an advocacy role, providing a website to conduct searches and explore YouTube recommendations. This website also displays a 'YouTube Map' [fig. 4.4] where the platform is displayed as a network visualisation clustered around broad categories.

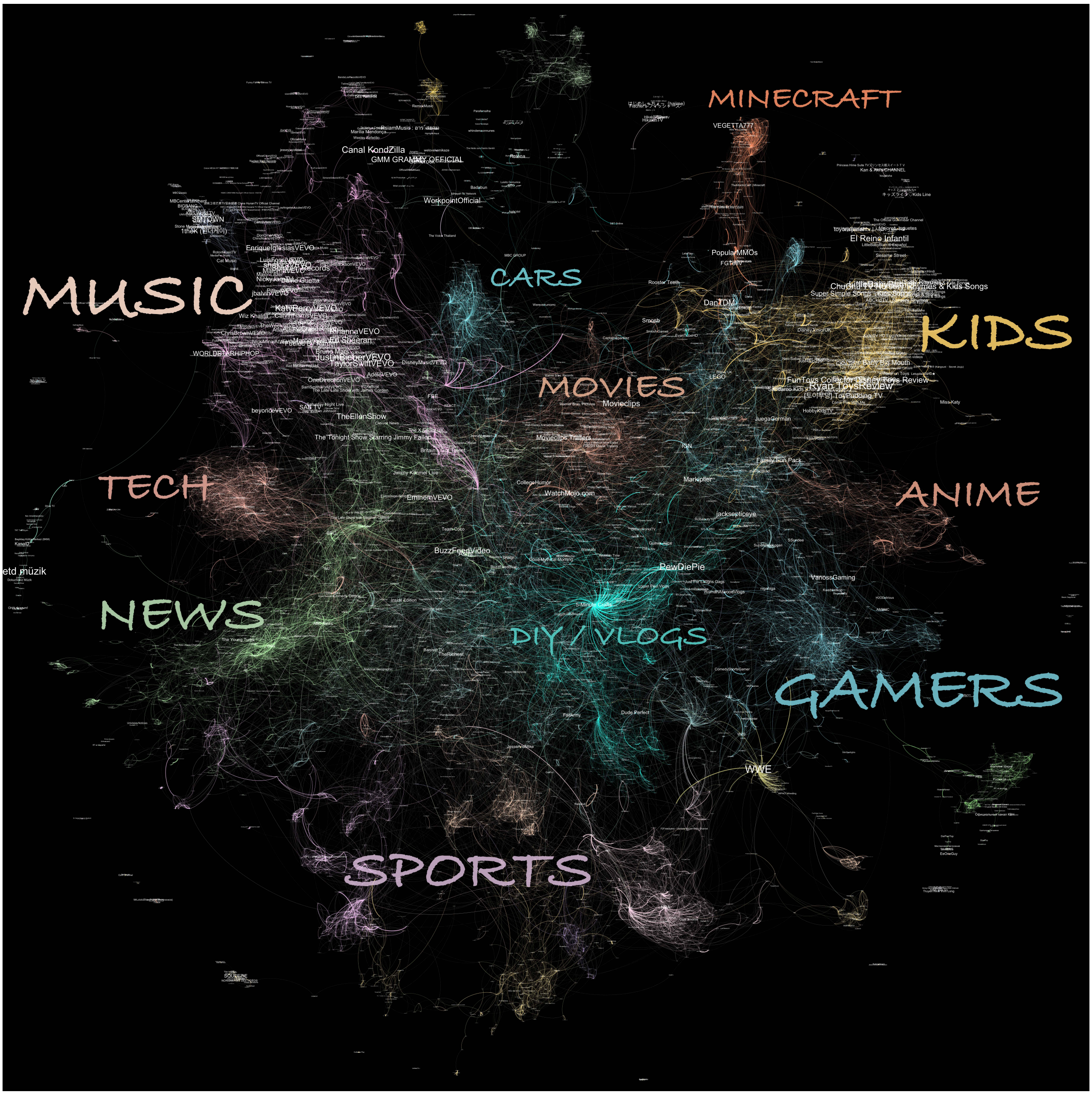


Figure 4.4: *YouTube Map* Chaslot et al. (2016a).

4.2.3 Seeing vs. knowing

The YouTube map by Chaslot et al. (2016a) materialises the paradox at the core of research that aims to critique the epistemology of data and algorithms while also relying on them as a source of knowledge (see Kitchin, 2014). It exemplifies the critique of data visualisation articulated by Boehnert (2016) around the themes of 'digital positivism', 'darkdata', and 'datawash'. By declaring that 'the YouTube algorithm distorts truth' (Lewis, 2018a) Chaslot fully subscribes to the positivist ontology that underpins YouTube's recommendations, but only shifts its target to the algorithm itself. There is still a ground, objective truth that is said to be measured and seen through data. The map qualifies as datawash, as it actually 'conceals or obscures knowledge' (Boehnert, 2016). As a 'hairball' or 'spaghetti bowl' visualisation (Bounegru et al., 2017) the *YouTube Map* is just another form of opacity, mistaking visual richness for knowledge. Finally by implying that a system that is in constant growth and shift can be viewed in its totality, the map glosses over dark data, for example content outside of the most well known '1000+ US channels' used to map recommendations.[9] While the idea of turning instruments of data collection and visualisation back against YouTube's capture apparatus is appealing, this map does in fact very little to *reveal* anything about YouTube. Seaver suggests that,

> ...these critiques take issue with the current scope and power of captology, they generally share its behaviorist common sense. They identify the problem as misaligned corporate incentives, rather than behaviorist premises. (Seaver, 2018, 16)

In diagrammatic terms, the YouTube Map claims to be a *tracing*, which implies a likeness to the object it describes.[10] This logic of reproduction fully subscribes to the positivist stance of the platform (YouTube) it seeks to criticise, if YouTube *creates* categories and identities while it claims to predict them, *AlgoTransparency creates* a version of YouTube that is fully knowable, and therefore fixable, by claiming to reveal its inner workings. It misses the diagrammatic *relations* key to gaining knowledge about the system, its actors, their motivations, and politics. This 'seeing without knowing' is described in detail by Ananny and Crawford (2016) as they characterise the ideal of rendering algorithmic systems accountable through transparency as over-simplistic, inadequate, and misleading. Meaning, they

[9] `https://algotransparency.org/methodology.html` (accessed 10 July 2019)

[10] This can be generalised to the field of data visualisation. see Drucker (2011) and Boehnert (2016).

argue is 'achieved through *relations* not *revelations*' (5). Thinking through relations means considering predictive algorithms as part of broader systems, made up of corporations, infrastructure, programmers, content producers, and users. It can also be used to scrutinise the recommendations themselves and encourage interpretations as to how they might have originated.

With the challenges inherent to *mapping* a platform such as YouTube established, I now turn to confronting these issues through my practice. In the next section I describe my own mapping of YouTube, first in a way similar to Chaslot's for the *EHSM* exhibition, and then after a shift through the collaboration with Betti Marenko and the review process for the *RTD* conference.

4.3 Practice

My own early mappings of YouTube started with a collaboration between Supra Systems Studio and BBC R&D. I was prompted to respond to previous work by BBC R&D on *Tellybox*, an exploration of physical interfaces to display, and interact with, recommendations (Miller et al., 2017). I positioned my proposal to map YouTube at the intersection of my own research interests in algorithmic prediction, diagrams, and BBC R&D's project.

4.3.1 Data Collection

I have established the challenges and epistemic issues surrounding the idea that YouTube can be mapped. I was aware of these, at least instinctively, when I started the *Arc-choice* project, but moved to confront these issues first hand, through practice. I gained technical knowledge about graph databases while working on *Diagrams of the Future*, and these skills were suited to another form of archaeological practice, moving up-close to a specific predictive system currently in operation. In fact the way I use graphs in the *Arc-choice* project is more directly related to their mainstream use for network analysis than my excavation of the history of predictive diagrams.

I designed a series of "probes"[11] to scrape data from YouTube using a headless web browser (Selenium) that simulates user interaction with web-pages such as clicking through recommendations. This approach allows for quick prototyping using the browser's developer tools to identify parts of the application where relevant data and interface elements are located. Unlike the work of Rieder et al. (2018), my data collection did not make use of the official API but simulated a user interacting with YouTube from a web-browser. While this has serious limitations—for example having to load web-pages in memory and program recursive loops instead of querying only the meta-data through the API—I made this choice in order to stay as close as possible to what a YouTube user might see and do.

The YouTube API is aimed at developers wanting to integrate with YouTube from their applications, and therefore did not seem like a relevant source of data on the recommendations provided in the YouTube web interface.[12] I took steps to make each session as "blank" as possible through the available settings, clearing cookies for example. It is difficult if not impossible to know what data persists, or how each session might still be identified. Location was shown to be factored in by the system for example. Using a VPN or a server based in the Netherlands for the *RTD* conference yielded results that were noticeably different and tied to the country of origin of the server's IP address.

Starting from the list of videos on the YouTube homepage, I follow links to video pages and recursively list recommendations [fig.4.5]. This differs from other studies such as Rieder et al. (2018) or Chaslot et al. (2016b) that focus on results from a specific search query. My interest here was on probing, or "sampling" with a degree of randomness, the YouTube catalogue from the homepage as its most obvious entry point. Three different probe designs [fig.4.5] balance the breadth and depth of the mapping in different ways. *Ripple.py* follows all recommendations for a given number of recursions. This exhaustive approach was very time

[11] My use of "probes" here has more to do with 'computer programs as automated probes that send back telemetry' Parrish (2015) than with previous uses of this term in a design or Human Computer Interaction context (Gaver et al., 1999).

[12] This echoes broader questions about digital methods to study platforms that are increasingly reluctant to publish data through their APIs, and the limitations of 'Post-API Research' (Digital Methods Initiative, 2019). These concerns are summarised by James (2019) in her review of *Spotify Teardown* (Eriksson et al., 2019): 'how does one study someone who doesn't want to be studied and (unlike Spotify's users) has the power to enforce that wish?'

consuming, and the resulting visualisations became unreadable after the first recursion [fig.4.6]. *Simple_digger.py* is a drastically simpler approach which selects only one video at random in the recommendations list and repeats the process a given number of times. Finally, *Digger.py* is a compromise which captures all of the recommendations at each step before choosing one at random. The use of random choices is another limitation to note here. My aim was simply to probe the vast architecture of recommendations on the platform. While other studies focus on specific topics related to politics or the news, my aim here—at least in the first instance—was much more tentative.

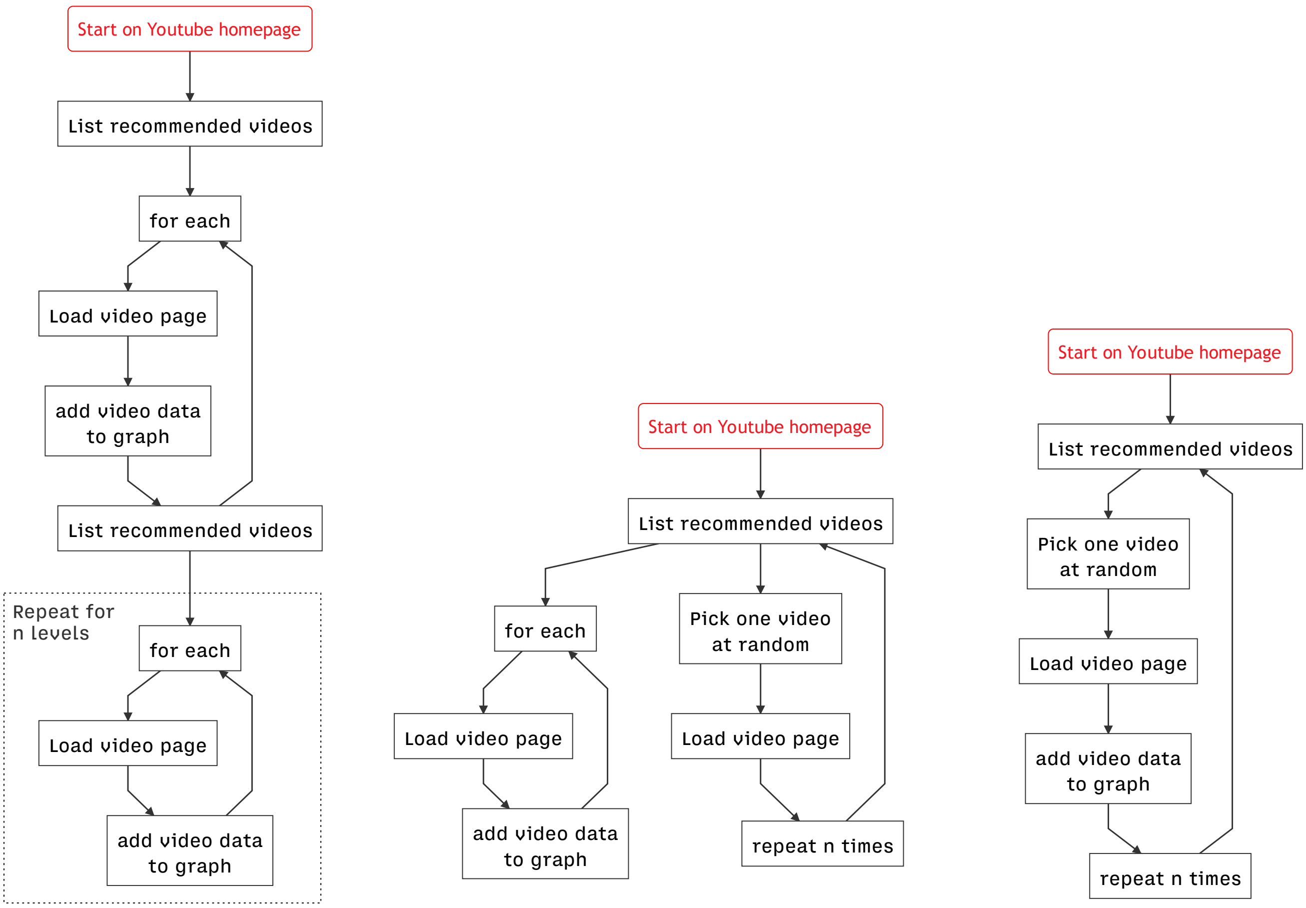
The collected data are stored in a graph database (Neo4j, similar to the structure used in *DOTF*) where videos are represented by nodes, and recommendations by edges connecting them. Each session is logged with a time stamp and indicates the type of probe used. This database can then be queried to return particular sessions or sub-graphs for visualisation.

4.3.2 Visualisation

The design and development of the probes was driven by experiments in visualising the data they scraped, as the goal of the project was to produce a visual outcome. This pushed me to confront the gap between seeing and knowing (Ananny and Crawford, 2016) in a rather direct way. I moved from my first failing attempts to map the totality of the recommendations to increasingly focused and specific paths through the "landscape" of YouTube content.

Neo4j browser was the starting point for my interactions with the collected data, as it is the tool provided with the database. It returns the result of queries as interactive graphs where nodes can be dragged and positioned by hand, and their size, colour and labels adjusted. This proved very useful for experimentation, for example to sketch sequential layouts and identify that video titles were the most suggestive data to foreground as the node labels [fig.4.9]. The main limitation however was the lack of automation, meaning that any layout realised by hand could not be standardised or reproduced with another set of data, only exported as static images.

To move beyond the manual layouts of the exploratory stage, I tested a number of off-the-shelf tools, each with their own trade-offs. *Matplotlib*, *Networkx*, and *Datashader*, industry standards for data visualisation, especially for network visualisation in the Python ecosystem, were obvious choices and purpose built to display network data. However they lost some



(a) *ripple.py* (b) *digger.py* (c) *simple_digger.py*

Figure 4.5: Three probes to collect YouTube recommendation data:
a) ripple.py b) digger.py c) simple_digger.py.

of the qualities of the Neo4j experiments. Displaying long video titles was not possible, and layouts were force-based or circular which suggested a sense of totality I was seeking to avoid [fig.4.6]. Flowchart markup language Mermaid.js produced interesting results [fig.4.7], but the amount of data and relations I was attempting to draw stretched the layout engine well beyond its capabilities, eventually breaking it completely. Finally I tested Gephi, an open-source tool well known to scholars in the digital humanities (Bastian et al., 2009). Like the Python libraries, it was very well suited for dealing with large networks, but showed limitations with a propensity for producing "hairball" visualisations [fig.4.8]. Simple things such as titles, which were impossible to "wrap" on more than one line of text, revealed underlying assumptions that were incompatible with my

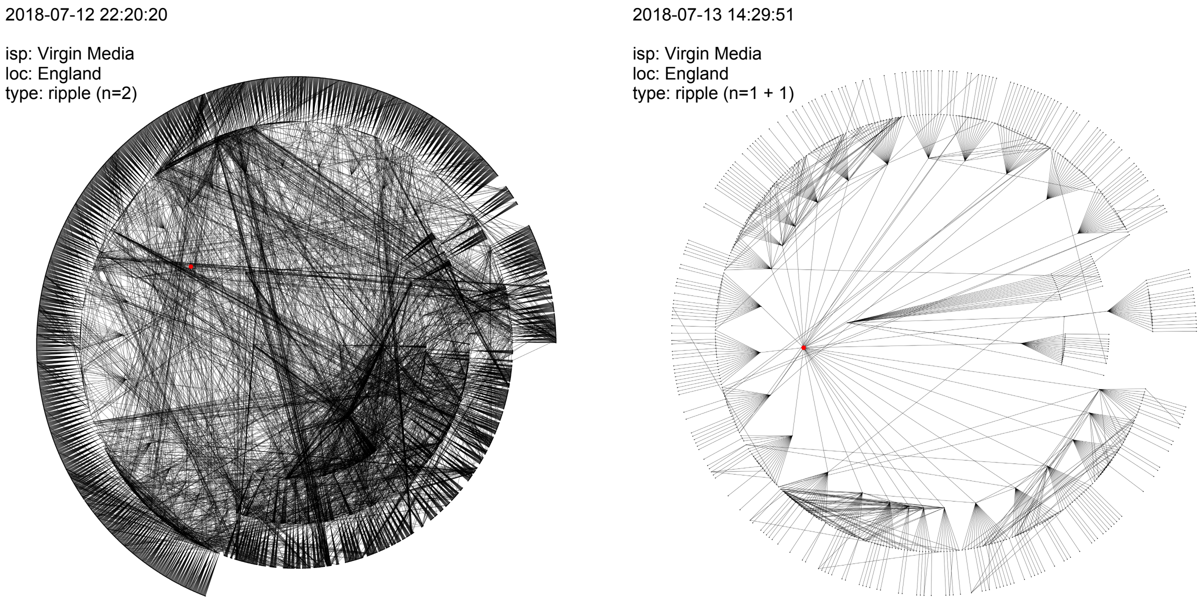


Figure 4.6: Visualisation created with the *ripple* probe with 2 levels of recommendations (left) and one with one random choice (right). Circular layout in Matplotlib with Netowrkx.

modest requirements. My take away from these tests using popular data-visualisation tools is that the issues I had with the *AlgoTransparency* map—the materialisation of a positivist premise—are partly encoded in the default settings of many of these tools.

For the *EHSM* exhibition I produced a set of prints that included most of these explorations. Effectively, they amounted to a small survey of the state of the art in network data visualisation tools, and of my technical ability to use them. Neither of the methods I tested was fully satisfactory however. Neo4j Browser was suited for visual exploration but only if manipulations and layout were done by hand. The other tools were readily automated but came with built in assumptions that reproduced the positivist stance I was aiming to avoid. The project for *EHSM* made me reach the conclusion that I would need to design my own solution if I was to develop the project further beyond the exhibition. I discuss this shift in the next section.

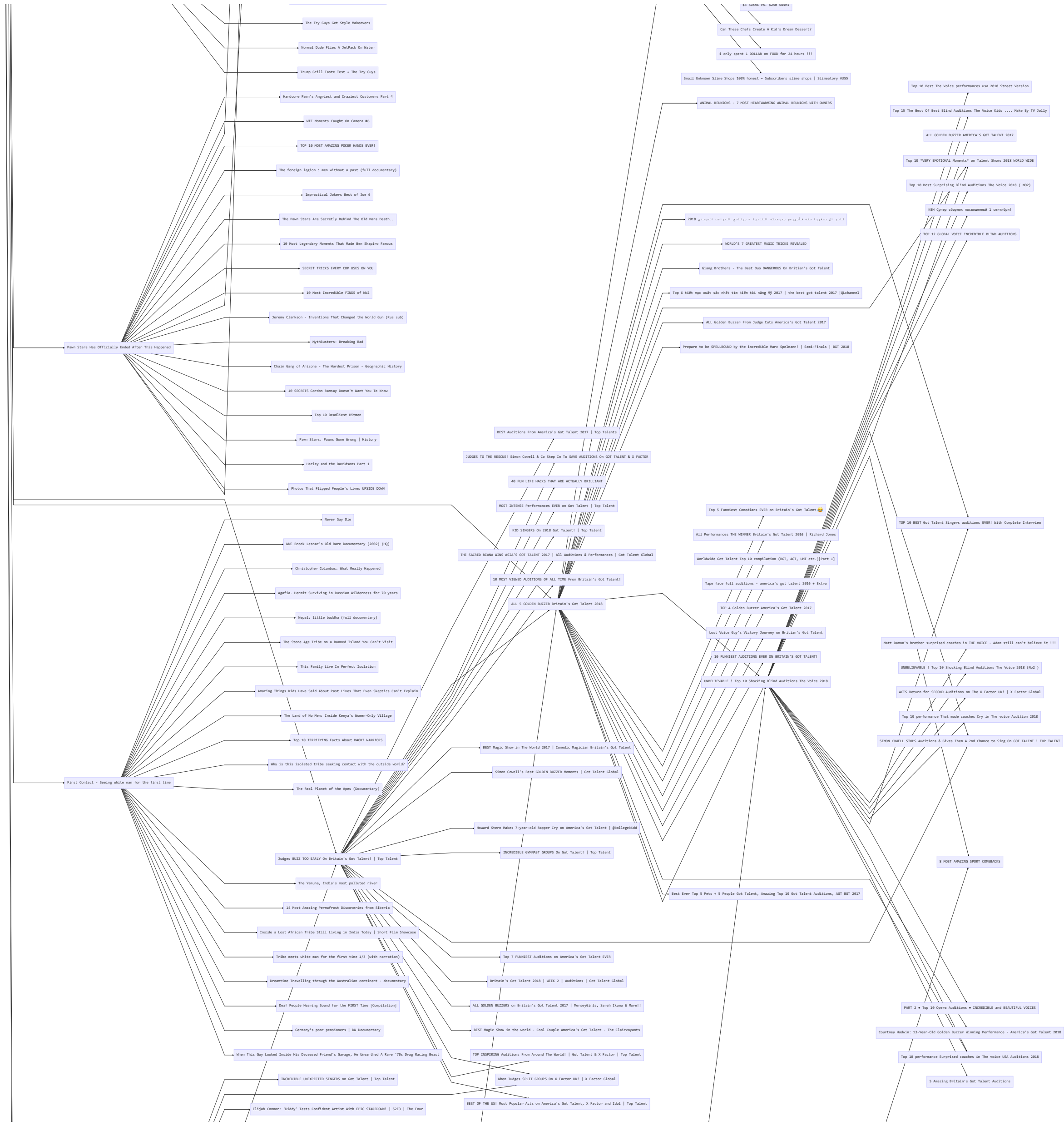


Figure 4.7: Mermaid.js visualisation (cropped).

Figure 4.8: Gephi visualisation with manual layout of main nodes (cropped).

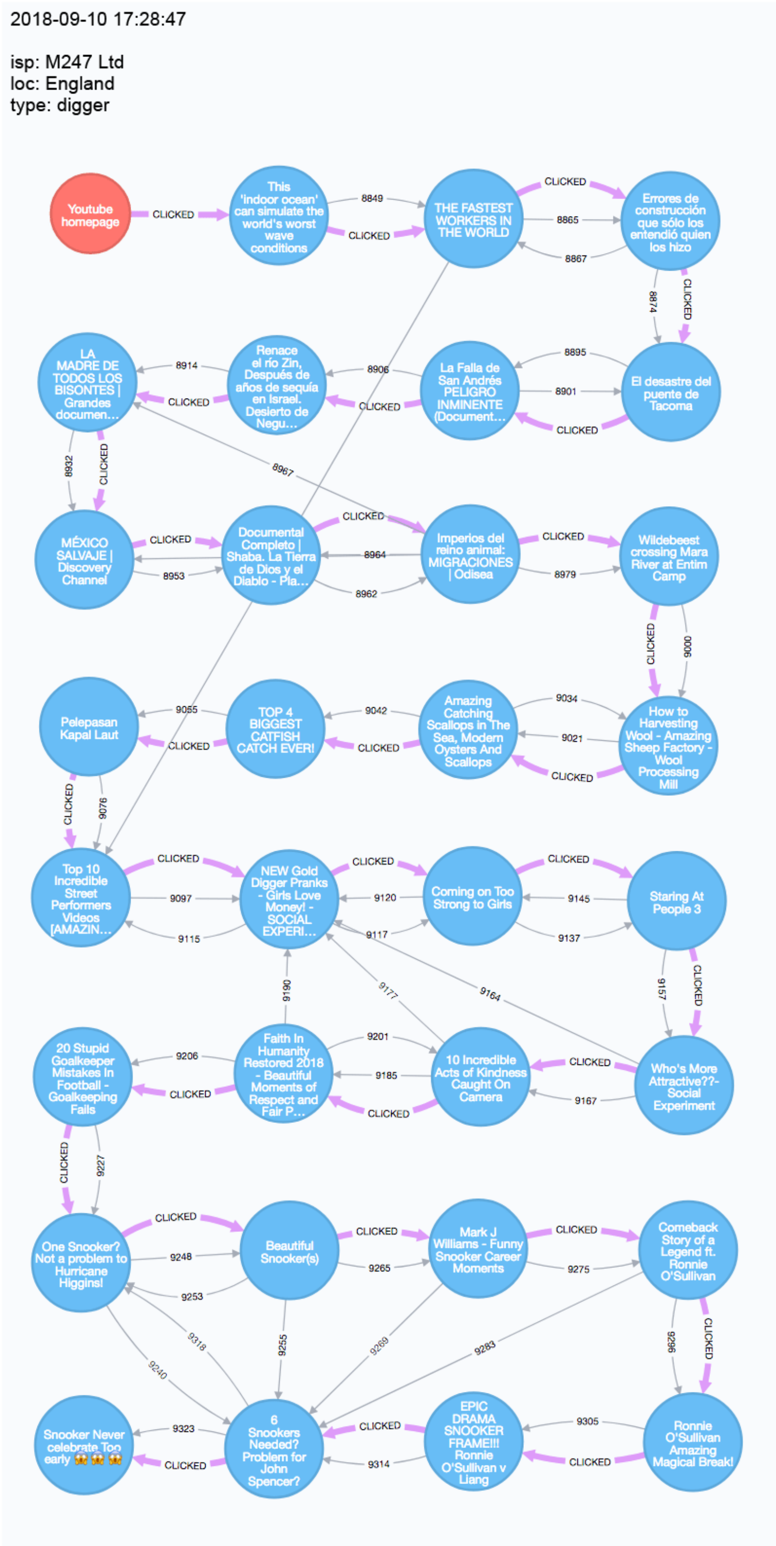


Figure 4.9: Neo4j Browser visualisation, manual layout by dragging nodes on the canvas.

4.3.3 Traces

The *Research Through Design* conference (RTD) combines research and practice in a unique format where papers are presented alongside artefacts. It encourages experimental approaches while also adhering to a rigorous academic peer-review process. The conference provided an opportunity to further develop the *Arc-choice* project, and to bring the ongoing collaboration with theorist Betti Marenko (first articulated in Marenko and Benque, 2018) to bear on my early set of experiments. Looking at the specific case of YouTube, and at my preliminary work visualising recommendations, focused our discussion and allowed us to examine and critique both algorithmic prediction and my work in diagrammatic terms. In effect, the visualisations I had produced were not the 'speculative diagrams' we had been discussing. They were embedded in a positivist ontology of representation. I developed the practice work in response to our conversations. This involved producing visualisations much more deliberately through a purpose-built program. In the next section, 4.4, I turn to discussing how this shift was informed by our collaborative research process, and in turn contributed to our distinctions between types of tracings and traces. I first conclude this section by describing these visualisations from a practical and technical standpoint.

The traces visualisations were built with D3.js to display the output of the the *simple_digger* probe, namely "strings" of videos that were recommended and clicked in a sequence. Following my experiments in Neo4j browser [fig.4.9] I used a grid based layout to display the sequence of recommendations.This was to "activate" the space of the diagram, bringing focus to the relations between recommended videos and to their meta-data; title, channel, number of views. The system was used to produce vector based images for our paper, displaying around 15 videos on an A4 page. [fig.4.10]

For the exhibition setting, I developed this system to display much larger data-sets—up to a few thousand videos—as animated sequences. This used a similar layout to display the more recent videos in the probe's trajectory while previous steps gradually faded away. This was initially programmed as a live installation, with the data collection and visualisation happening simultaneously. However due to the technical constraints of the exhibition setting (ie. running the memory intensive program from a Raspberry Pi computer), I added the ability to "play back" the data collected and saved as CSV files. The installation displayed a counter to indicate the progress and total number of videos in the data-set [fig.4.12].

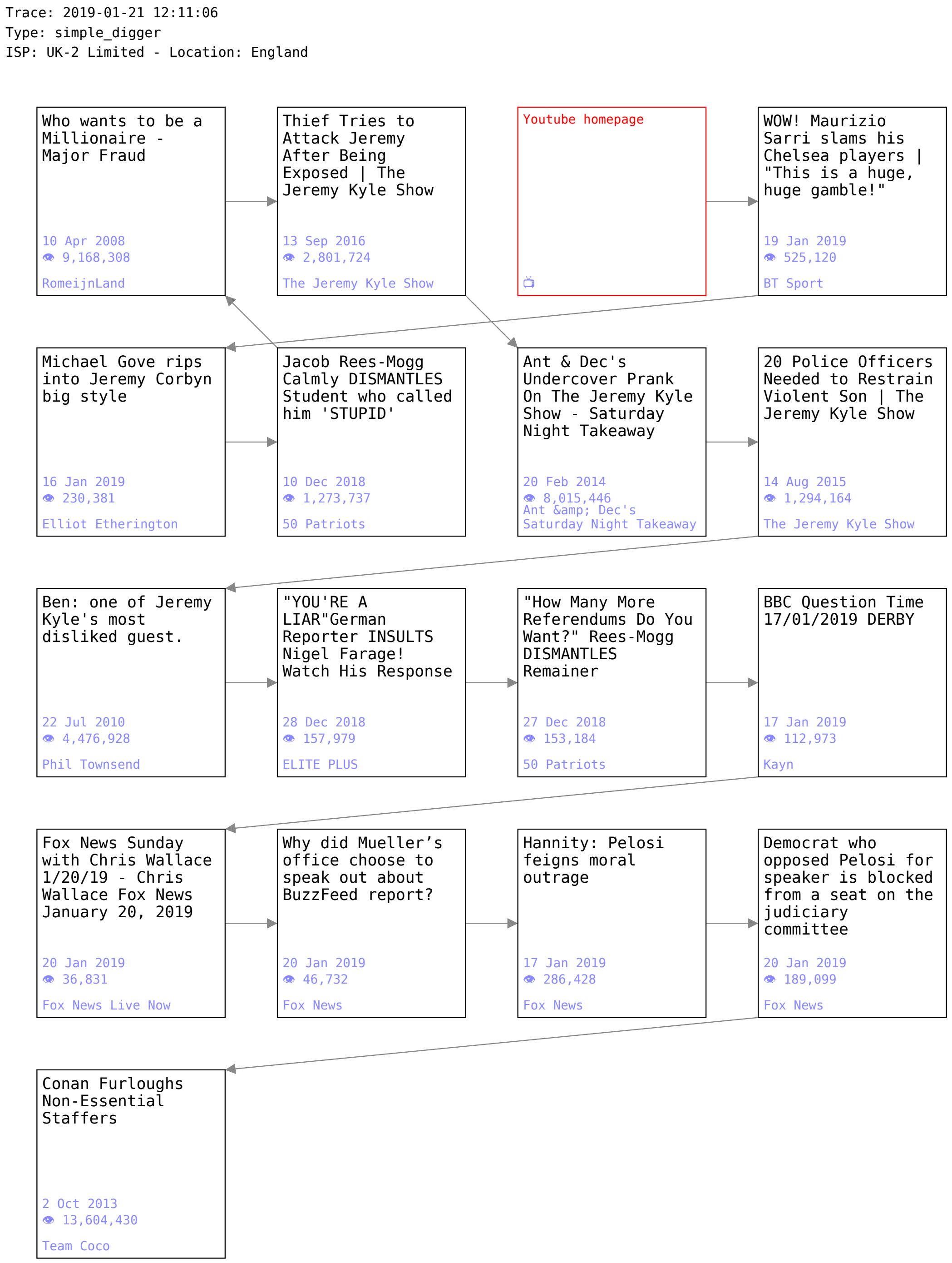


Figure 4.10: Output of the print version of the trace visualisation.

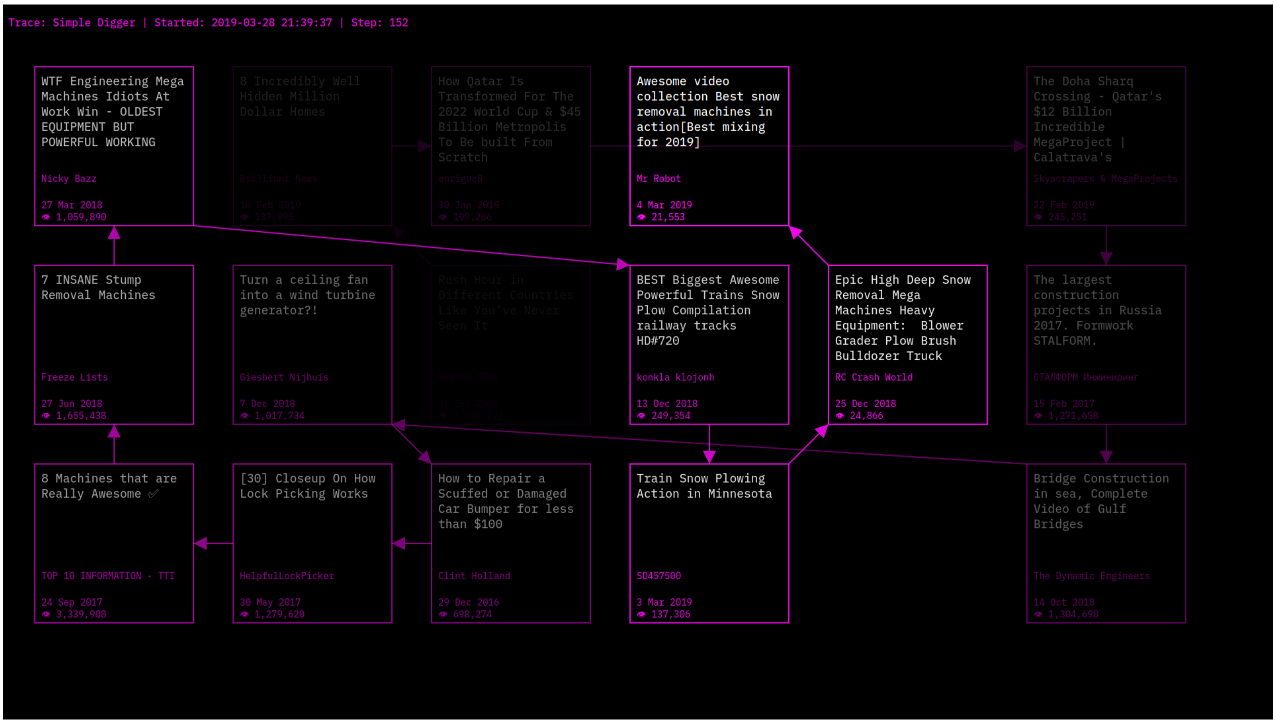


Figure 4.11: Example frame from traces animation.

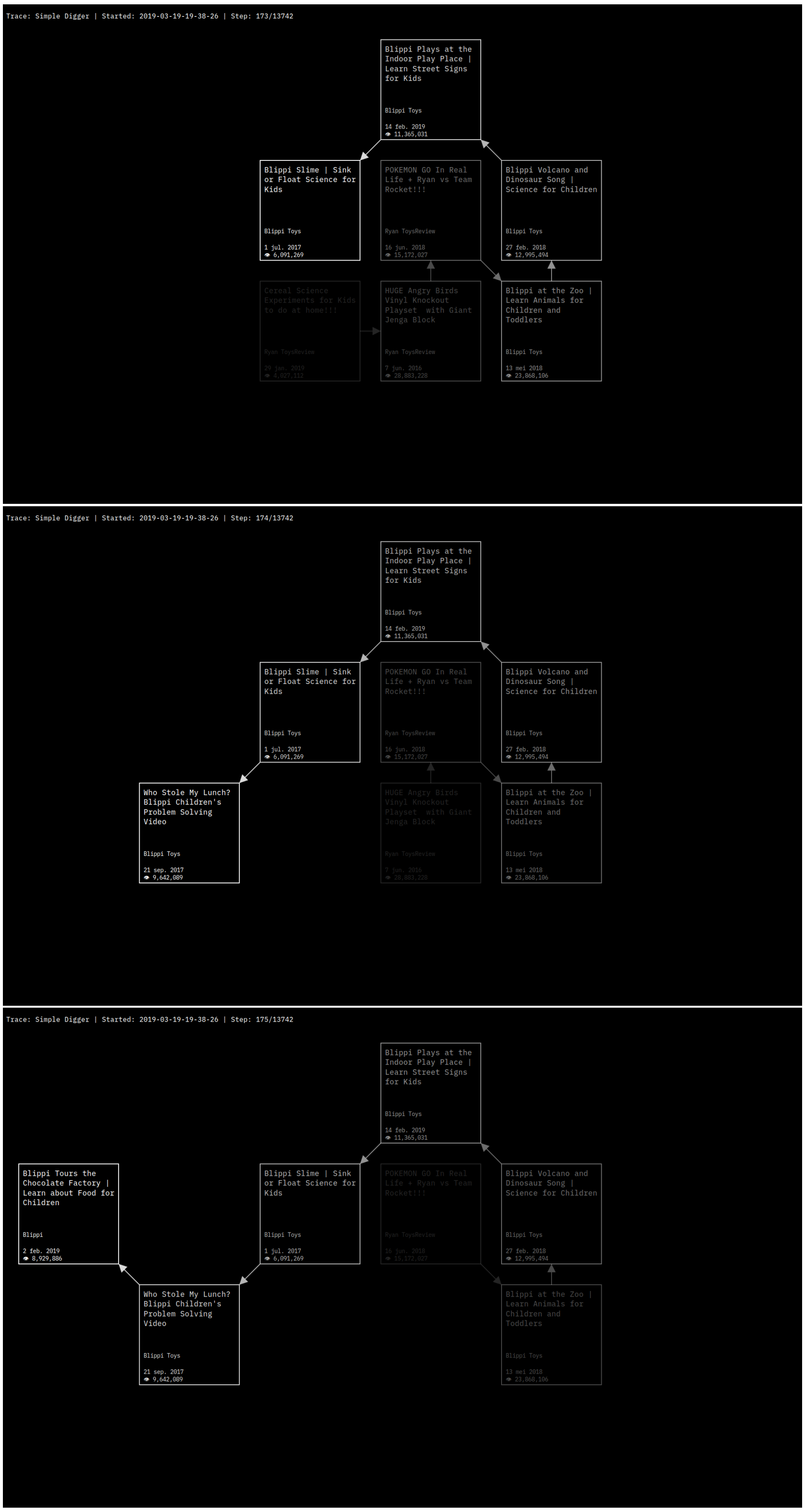


Figure 4.12: Animation frames as presented at *Research Through Design* note the counter of steps also indicates the number of videos in the pre-recorded session, in this case: 13,742.

4.4 Discussion

I now turn to discussing the *Arc-choice* research project through my research questions (RQ, RQ1, RQ2) and discussion criteria. I reflect on the language of diagrams as a mode of reading the logics of entrapment at play in YouTube recommendations. Here the movements of the research are informed by algorithmic outputs, as collected by the probing instruments I have described. The notion of trace, central to the control diagram of YouTube, affords a critical pivot that challenges the of knowledge production at play in the study of algorithmic systems.

4.4.1 Enclosed and hosted

The minute, granular nature of traces—the interaction-log data that recommendation systems rely on to make their predictions—stretches the very definition of my research focus. Are traces diagrams? Like other data, they are put in relation with each other as part of the operations of algorithmic prediction and as such can be considered, at the very least, a kind of substrate on which diagrams proliferate. In response to RQ1, my focus with *Arc-choice* was to find ways of probing these relations, to bring them into focus through a deliberate use of visualisation instruments. Moving away from off-the-shelf tools and their built-in assumptions meant adopting a much more considered approach to the aesthetic representation of data, what Goatley (2019) calls 'critical data aesthetics.' This research provides a way to think *through* practice (Frayling, 1993) about the problematic nature of algorithmic recommendations, and the claims made with data-visualisation. I arrived at some outcomes that encourage a contemplative mode of interrogating each relation as it comes into view, instead of promising to "pry open" or "reveal" the workings of algorithmic prediction.

Our focus on traces emerged out the development of the research after the *EHSM* exhibition, in part addressing my own doubts about the work, and in part through the peer-review process of the *RTD* conference. One peer reviewer in particular—among otherwise positive comments on our initial submission—summarised these concerns and urged us to move from a totalising view to a more specific one. Reviewer 1 fed back the following:

> I challenge the authors to pull out 1 single compelling path of content, so I can map this global diagrammatic contouring, with what actually matters to ME, the USER; [reviewer 1]

The trace provided a common language between research and practice, specifically in the case of my collaboration with Betti Marenko, between theorist and practitioner (RQ1). The development of the practice work was driven by, and in turn informed, our theoretical position. With each of us coming from one side of a theory/practice divide, we used diagrammatic language to grapple with crucial distinctions between traces, tracings, maps, and mapping. While at risk of being lost in subtle word-plays, the delineations between these terms focused our conversations and the practice of programming the visualisations. One the one hand data-visualisation relies on *tracing*, a logic of reproduction that claims a likeness to the objects it describes (Drucker, 2014a, see also section 2.2.2). This is exemplified in the work of Chaslot (2016) and its coverage (Lewis and McCormick, 2018), where YouTube is said to be made transparent. This logic can be extended to the recommendation system itself. Here the assumption is that *traces* accurately describe users and their preferences, while they in fact capture them into diagrammatic constructions that produce identities rather than cater to them. On the other hand, our critical position was informed by the Deleuzian[13] injunction to 'make a map, not a tracing' (Deleuze and Guattari, 1987, 12), that is to explore more speculative and contingent modes of knowing about algorithmic prediction, ones that didn't stay "trapped" in its positivist logics. This eventually resulted in a shift in our use of the word *traces* that I cover in the next section.

If recommender systems can be characterised as traps (Seaver, 2018), then perhaps this work is best seen as a way to probe the movements and mechanisms (RQ1) that this entrapment entails. I have pointed to the limitations and tentative nature of my probe design in section 4.3.1, in this light they can perhaps be compared to the practice of dye tracing, where a coloured agent is released in a river to analyse its flows. My probes, however random and limited,[14] were released in the capture apparatus of YouTube recommendations in the hope of highlighting some of their mechanisms, and to provide a way of reflecting on them through practice. They effectively produce the spectacle of watching a simple automated agent get endlessly *trapped* in the recommendation apparatus. The animated version of the piece was in this regard the most accomplished, as it displays a potentially infinite sequence of recommendations. This is in my view the most accomplished version of the work in terms of relating

[13]The Deleuzian influence on our work came from Betti Marenko's side, she has a long involvement with this line of thinking and "wrote the book" on its intersection with design (see Marenko, 2015).

research and practice, as it had more time to mature in the three months between the print deadline for our paper and the exhibition at the *RTD* conference. As I have outlined in my Methods, my criteria for "accomplishment" here is in terms of trajectory, or distance travelled, between the initial set of visualisations and the focus on individual traces. This focus was enhanced by the time-based nature of the medium, allowing for very long traces to be displayed in a way that linked back to notions of entrapment.

As my probes move within an algorithmic trap, they echo the hopeful and constructive side of the argument made by Seaver (2018). Simply rejecting or refusing these traps is actually quite limited, if possible at all. It relies on a polarity between capture and freedom that does not question the 'behiorist common sense' (14) at the core of these systems, and focuses instead on somehow correcting the business models of massive corporations, or on putting the burden back on users—for example through apps that monitor and ration "screen time." Instead, what Seaver points to is the possibility of existing—or for my purposes here of critically and creatively practising—from within traps. This is not an abdication but rather a recognition of the generative potential of traps. If algorithmic prediction can be seen, in this instance, as a diagrammatic trap made up of relations between traces, then like any other diagram it oscillates between pinning down and opening up (Leeb, 2017). In Seaver's words this translates into an oscillation between enclosing and hosting.

> To be caught at this speed is not to be dead, rather it is to be enclosed, known, and subject to manipulation. In other theoretical registers, this is akin to Deleuze's "control"(1992; Cheney-Lippold 2011) or Foucault's "governmentality" (1991); styles of enclosure that are no less sinister for being less than absolute. But to be caught at this speed is also to be hosted—to be provided with conditions for existence that facilitate activity while constraining it (Swancutt 2012; Derrida 2000). (Seaver, 2018, 15)

Relating back to RQ1, the movements of my research and practice are at once constrained by the algorithmic system and facilitated by it as the environment and material for this work to exist in the first place. While the project does not "reveal" anything about the YouTube recommendation algorithm, *Arc-choice* is about examining the possibilities

[14]Especially *simple_digger.py* that was the main source of data for the developments on traces.

for research and practice from within traps. Betti Marenko and I articulated this reclaiming around a shift in the definition of *traces*, which I explain further in the next section.

> The question to ask of traps may not be how to escape from them, but rather how to recapture them and turn them to new ends in the service of new worlds. (Seaver, 2018, 16)

4.4.2 Traces as conjectural knowledge

The key shift between the work I presented at the *EHSM* exhibition and the paper and corollary animations Betti Marenko and I presented at the *RTD* conference is contained in the reading of the word *traces*. Specifically, in relation to RQ2, we were attentive to the mode of knowledge production by which traces are mobilised. As I have discussed above in section 4.2.1, trace data form the loci of a control diagram that entraps users to maximise their "watch time." This is grounded in a positivist imaginary of data that considers traces as objective and neutral representations of users, meanwhile it produces categories and promotes extreme viewpoints.

We borrowed another meaning for *traces* from historian Carlo Ginzburg (1980). For Ginzburg, the interpretation of traces is part of a *conjectural* model of knowledge, rooted in 'the tracking skills of the first hunters, as much as in Mesopotamian divination' (13). This is the kind of knowledge production that hunters perform when they 'reconstruct the appearance of an animal they have never seen' (22) from interpreting the *traces* left behind. The interpretation of clues is central to the modern scientific paradigm, mobilised to enforce state control (recognising individuals through facial features and later fingerprints), and colonial power (27). Ginzburg points to conjecture as an overlooked and not easily formalised mode of knowing from clues, rooted in the senses and still universally used. This form of 'low intuition'—by comparison to the 'high intuition' of superior knowledge—relies on an 'elastic rigour' (28) that evade a purely rational and quantitive gaze: hunches, whiffs, and glances.

In this light, *Arc-choice* is an example of conjectural knowledge applied to digital media. It generates *traces* as clues, allowing for images of an elusive recommendation system to emerge. It focuses these conjectures on the relations between videos, where the recommendations operate. Prompting the viewer to make a speculative guess, an ellipsis, to fill in the gaps as each new step comes into view. With this our collaboration made a wider point, to encourage more qualitative and situated modes of knowing with data, in contrast with algorithmic prediction itself, but also with critical positions that make claims to total or objective transparency—which is to say, that stay within the bounds of the positivist data-imaginary. We pointed to the ways in which data visualisations often smuggle in a positivist ontology, even to research that aims to be critical of its negative side effects. To counter this we proposed that computational tools such as data, visualisations, and algorithms could be re-purposed towards more conjectural modes of knowing.

The very process of my collaboration with Betti Marenko on *Arc-choice* relates to RQ2, and to the notion of Speculative/abductive practice. While we started from tentative hunches and intentions, these were reified through the peer-review process of the *RTD* conference. This echoed the call by computer scientist Simon Peyton Jones—as he was addressing my cohort of doctoral students at the Microsoft Research PhD Summer School (2016)—to treat academic papers as a 'primary mechanism for doing research, not just reporting on it afterwards' (Peyton Jones, 2016). His injunction was to start writing first, then conduct the research [see fig 4.13].

Figure 4.13: The paper as primary mechanism for research, redrawn from Peyton Jones (2016).

Betti Marenko and I did not exactly follow this method—our process was more loosely structured and pushed forward by the successive deadlines of the peer review process—our collaborative paper can be considered a 'mechanism for doing research' (ibid.). We started with two elements 1) the broad position we had established as the aim of our collaboration between theory and practice: reclaiming algorithmic prediction through speculative diagrams (Marenko and Benque, 2018) and 2) my set of visualisation experiments produced for the *EHSM* show. We submitted these as an abstract to *RTD* in a speculative manner. We did not know how we were going to resolve them together. Through the stages of our paper's review we meshed these two strands together to produce something new. We took the work on *Arc-choice* as a an opportunity to anchor what was a very tentative set of ideas, applying them to a specific context namely the predictions on YouTube. The *RTD* reviewers pushed us to confront the gap between our inital argument—centred around diagrams as a way to reclaim the capture apparatus of algorithmic governmentality—and our practice-based artefacts that were effectively, at that point, indistinguishable from the problematic "transparency ideal" visualisations I discussed in section 4.2.3.

While our framing was critical of the positivist ontology of data visualisation, we did not effectively discuss or demonstrate an alternative mode of knowledge production. As I have discussed in section 4.3.2, I already had doubts of my own after the *EHSM* exhibition, but the review process confirmed these and compelled us to act on them. Their input informed our research, moving us closer to the diagrams of capture we were claiming to unpack. It was also reflected on the practice side, prompting me to use data visualisation in a much more deliberate way as discussed in section 4.3.3.

4.4.3 Publishing and broadcasting traces

Arc-choice has functioned as a publication in an obvious sense, it was peer-reviewed and published as part of the *RTD* proceedings. The generative review process I described in section 4.4.2 made this a positive research experience, and the open-access publication may provide avenues for future collaborations and developments. Similar to Rieder et al. (2018), our research contribution is more on a methodological level than any particular findings. We did not discover a single trace with particular value but rather presented a system to generate and display traces. We used this to make a broader point about the importance of restoring conjecture as part of critical/creative practice oriented work that aims to interrogate algorithmic systems.

As I have mentioned above in section 4.4.1, the animated, screen-based version of the work was in my view the most successful at displaying the traces we were talking about, and at encouraging the contemplative stance that would prompt conjectures to be made. This could not be included in the published paper, first because it was developed after the camera ready version was sent to print (nearly three months before the conference), and second because of the constraints of the pdf format. However, in its exhibited form the project led to some novel ways of publishing data visualisations, between a website and an animation. As the data was being collected from the probe (or replayed in the case of logged traces) I made use of the auto-refresh functionality in the browser to animate the piece by only changing a very lightweight file.[15] This allowed me to avoid any video

[15] 2.1 kilobytes.

media where the frames would have to be saved as images and relied instead on web technologies. This led to experiments in "streaming" data visualisations using the DAT protocol and *Beaker*, an experimental peer-to-peer web browser (Frazee et al.). I used these tools to stream our *RTD* installation—admittedly to a niche audience, if existent at all—with promising results from a practice standpoint. This contributes a novel form of data visualisation publishing that should be explored in further work. This use of peer-to-peer networks means the mode of publication is consistent with my research focus, effectively weaving a diagram of connections that piggybacks on internet infrastructure while avoiding the "traps" of the big platforms.[16]

4.5 Conclusion

In this chapter I have presented work on the project *Architectures of Choice Vol.1: YouTube*, a mapping of algorithmic recommendations on the YouTube platform. I used this work to reflect on *traces* as the data points between which the diagrams of algorithmic prediction are formed. I described the "funnel" of YouTube recommendations as an example of such a diagram—described by its engineers as a distiller of personalised content, or seen as a trap using the work of Seaver (2018) (RQ1). I reflected on the ways these traps exemplify control diagrams, producing categories of users and prescribing identities rather than predicting them. I discussed the delineations between traces and tracings, and how these formed a terrain for research between theory and practice through my collaboration with Betti Marenko (RQ1). This practice moves within algorithmic traps, being at once constrained and aiming to reclaim the capture apparatus, to open opportunities for contemplation and speculation. The notion of traces was key to this reclaiming, as we rejected its meaning as objective data points that implicitly reflect user's preferences, and considered its use by hunters and diviners in a conjectural mode of knowledge production (Ginzburg, 1980). I argued that this mode of knowing is more conducive to a critique of algorithmic prediction through creative practice than claims to "reveal" its working or to render it "transparent" (RQ2). Not only are these claims impossible to realise but they are based on the very same epistemological premises as the systems they aim to scrutinise. Finally I described how we published traces, both in a static form in a research paper, and as

[16] My knowledge of the intricacies of the DAT protocol, for example the exact way in which it uses network infrastructure, is limited, for more details see Dat Foundation.

animations showing our probes getting endlessly trapped in the algorithmic recommendation system.

This chapter describes my second contribution to knowledge in this research, the production of *traces* as a tentative and conjectural activity that probes algorithmic prediction systems on their terms—that is programatically—but from the outside. In other words, producing traces is a way to engage with prediction as 'media observed' (Blegvad). In practice this translates into a series of automated probes and visualisation work aimed towards the production of traces, and their display through purpose-built software tools. These also produced novel modes of animating data visualisation using very lightweight means, and of broadcasting them over a peer to peer network.

Producing traces, in my case, means moving up close towards algorithmic prediction, to examine its outputs with computational probes. In the next chapter I move one step closer yet, into the operations of algorithmic prediction as I actually produce predictions and examine the diagrammatic forms and "tricks" involved in the process. I use this to reflect further on conjectural knowledge production as I read this process as a form of divination.

Please see appendix E for the practice submission related to this chapter, including: code repository, website, and other supporting material.
The Monistic Almanac is online at: `https://almanac.computer/`

Chapter 5

Chicanes

5.1 Introduction

'To be rational is to think in ratios, like the ratios that govern the geometry of the stars' (Levinovitz, 2016). The American religion professor Alan Jay Levinovitz questions the authority of economics and argues that the discipline is a modern version of astrology. His concerns are particularly resonant in an era saturated with data, and therefore with opportunities to find and interpret ratios between them. The point is also relevant as the study of cosmic movements have informed the development of mathematical techniques such as least squares regression, now one of the basic building blocks of algorithmic rationality.[1]

In this chapter, I examine how data and cosmic ratios have been used and legitimised for algorithmic prediction. I do this by excavating a particular type of publication, the almanac, and by using it as a site for creative practice. While the previous chapters *remembered* and *observed* algorithmic prediction, I move here to *imagining* a new media form (Huhtamo and Parikka, 2011). First, I reconstruct the almanac as a cultural site spanning different types and times of publication, and characterise its cosmic imaginary of data and prediction. I then take this as the basis for creative practice as I design my own version of the

[1] Least squares regression was discovered at the turn of the 19$^{\text{th}}$ century by Carl Friedrich Gauss, as he was calculating the orbit of the Ceres asteroid (Gigerenzer et al., 1990, 80—81, see also Forbes, 1971). I encountered it in practice as part of the machine learning course by Malone and Thrun (2015).

almanac. After mapping prediction from the outside (Chapter 3), probing its output (Chapter 4), I now move to the inside and produce predictions. I discuss *The Monistic Almanac*, a project that revisits almanac publications as a prototype of contemporary data analytics. It consists of a series of predictive widgets that combine the tools of data science with divinatory rationalities. This project was self-initiated as part of this research. It is published online, presented through talks (Anticipation2017, Fiber Festival, City of London tour), exhibitions (Everything Happens So Much [fig. 5.1], Supra Systems Office Rites [fig. E.1]), and writing (Benqué, 2018a, forthcoming).

I borrow my last diagram in this thesis, the chicane, from divination studies (Cornelius, 2016), and from its use by Ramey (2016) as a way to examine the politics of our mediated relationship with chance. I start by positioning almanacs as early prototypes for algorithmic prediction as we know it today. They pre-figured not only some of the ways predictions are produced and published today but a cultural space, a topos, that complicates the dividing lines along which the legitimacy of predictions is drawn. I then move to my practice of building *The Monistic Almanac*, an automated daily publication of predictions for the year ahead. I situate this work within a broader set of art and design projects that examine the entanglements between computation and divination before detailing some of the tools and predictive widgets that comprise my almanac. I then discuss the work using the chicane to describe the sharp turns, or "tricks" involved in the production of predictions. I discuss chicanes as key to examining the politics of prediction.

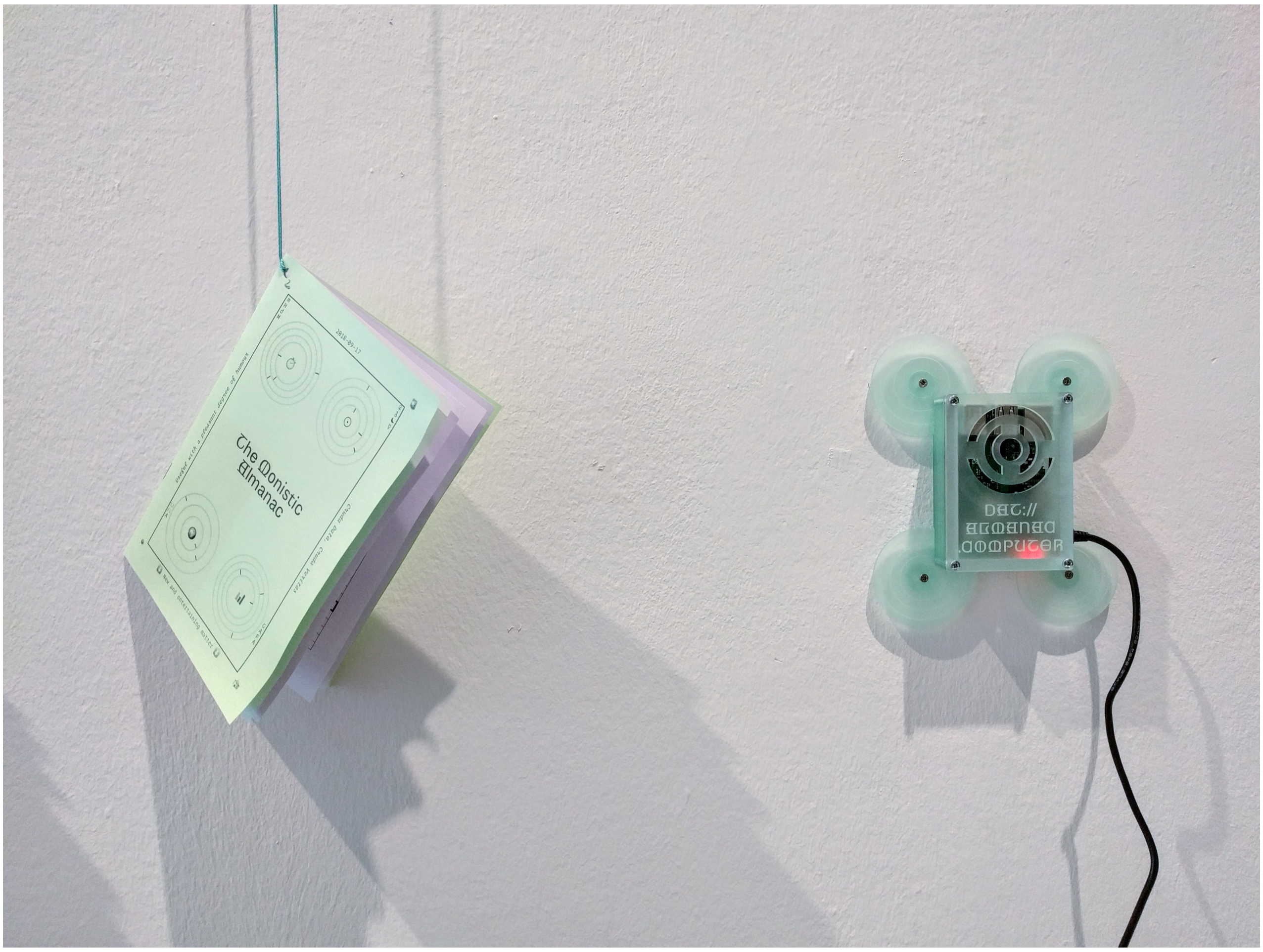


Figure 5.1: Detail of *The Monistic Almanac* as exhibted at the *Everything Happens So Much* exhibition (Sep. 2018). Showing the Raspberry Pi computer (right) and a printed publication (left).

5.2 The almanac as a prototype for divinatory data analytics

In this section I establish the background for *The Monistic Almanac*. I describe almanac publications as precursors to algorithmic prediction as it exists today. As such, the almanac affords a unique perspective on the data imaginary of prediction, one that considers data and algorithms as entangled with cosmic forces.

5.2.1 A cosmic imaginary

In Chapter 3, I introduced the notion of topos as a way of following the continuities, turns, and ruptures in the genealogy of prediction. In this chapter, I focus on one topos which is the almanac, as it materialises a broader cosmic imaginary of algorithmic prediction; observing the starts to predict events on earth. Almanacs are practical guides to the year ahead published since at least the seventeenth century[2] in areas such as farming, nautical navigation, and finance. Technology and politics writer Adrienne LaFrance (2015) frames *The Old Farmer's Almanac*, published since 1792, as a 'precursor to the information age', an early prototype for smart-phones with "apps" such as weather forecasts, calendars, navigation, and so on. This parallel is not only about the practical character of the almanac as 'a handheld, portable device' providing predictions for daily life, but about a 'cultural space' around the authority and legitimacy of data. One defining characteristic of this space is that it relates cosmic movements with daily events on earth. The almanac as a peculiar publication materialises a cosmic imaginary of data, 'prophetic' like the 'data imaginary' analysed by Beer (2019), but using data from the moon, stars and planets. This relationship is either direct, through recorded observations, or indirect through the borrowing of statistical techniques used in astronomy.

The almanac is a product of enlightenment ideals, a mechanistic view of 'the universe as machine: Once you get the operating instructions, you can tell the future.'[3] Breakthroughs such as the predicted return of comet Halley in 1759 (Broughton, 1985), an early prowess of computing (as

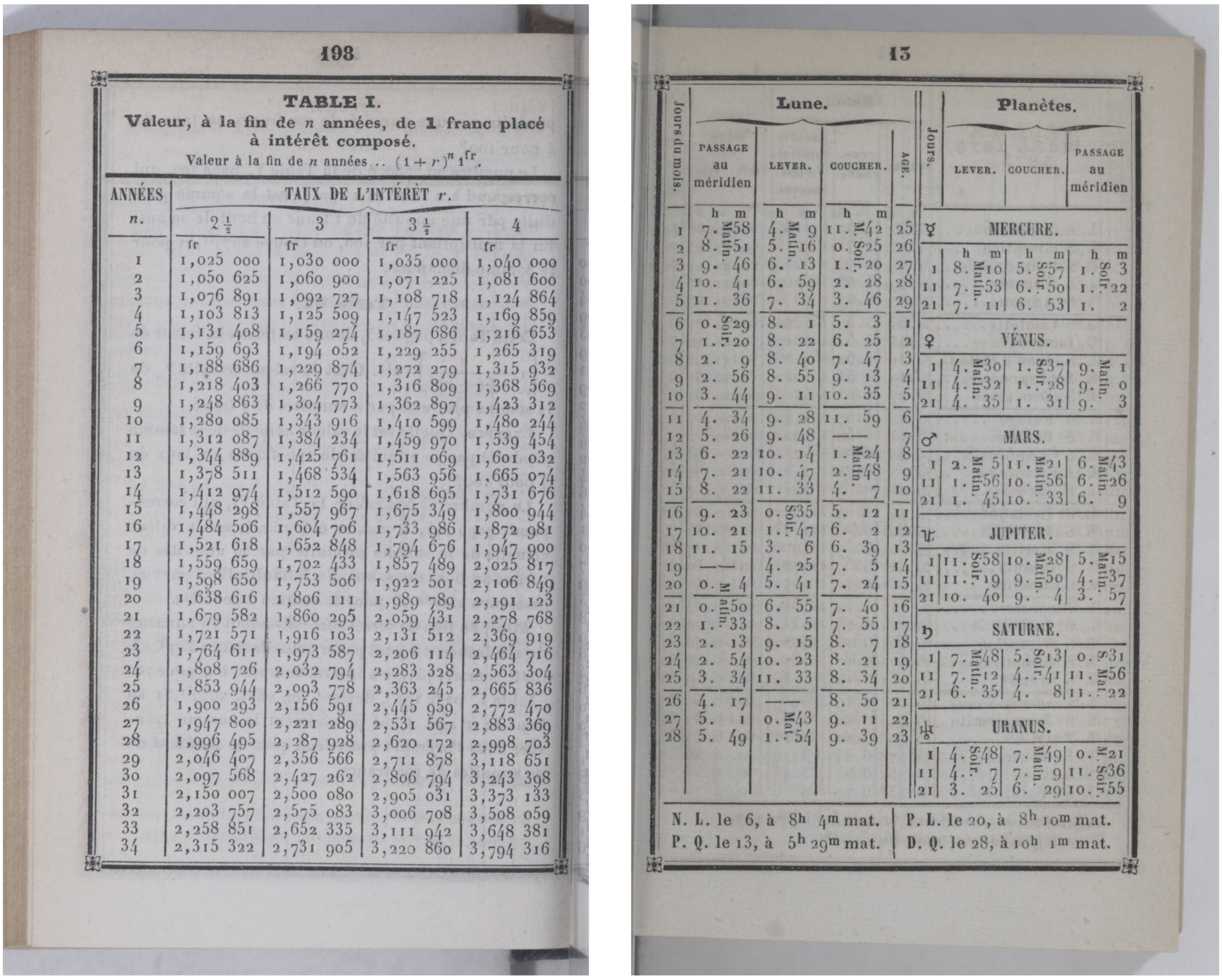
[2]Almanacs or their equivalent have been used for millennia in cultures across the world. In this chapter I focus on almanacs as mass-produced publications in the US and Europe as part of the history of statistics and prediction in the western world.

[3]Tim Clark, executive editor of The Old Farmer's Almanac, interviewed in *The New Yorker* in 1988. Cited in LaFrance (2015).

labour performed by human workers) and Newtonian physics, reinforced the sense of an underlying order to the world. Almanacs packaged this cosmic certainty and delivered it to populations such as farmers whose livelihoods were tied to the whims of the weather. This was not limited to the natural world. As I have discussed in section 3.2.3 statistical breakthroughs in astronomy were then transposed to society by astronomer Adolphe Quetelet, and served as the foundation for *social physics* that inform data imaginaries to this day. The sense that cosmic order must be reflected in human societies is grounded in a stance Barnes and Wilson (2014) call monism, 'the idea that there is only one set of principles that applies to the explanation of both natural and social worlds.' These universal laws of motions were derived diagrammatically from observed data. As orbit curves were made smooth and predictable, it was expected that the same statistical techniques would provide insights on anything, including human populations. While Hacking (1990, 106) calls this transposition by Quetelet blatant 'jumping to conclusions,' he also highlights the lasting influence of these ideas, as demonstrated for example by the lasting legacy of social physics.

Almanacs and their cosmic imaginary were also part of the legitimation of speculative finance. This time not through curves but the tables of data practices such as double-entry book- keeping, originating in astronomy and transferred to business 'it may be described as the first cosmos built up on the basis of mechanistic thought' (Sombart cited in Ashworth, 1994, 410). An example of these entanglements is the figure of Edmond Halley who gave his name to the famous comet but also authored the first mortality table for use in life insurance. In nineteenth-century Britain the 'business astronomers' used scientific techniques of data management to give finance the veneer of scientific rigour (Ashworth, 1994). Figures such as Francis Baily, John Herschel, and Charles Babbage 'set about putting speculation on the same footing as astronomy' (416) by applying the same data management rigour to their practice of both science and finance. They established their 'accountant's view of the world' (409) partly through their takeover—as the Astronomical Society—of the *Nautical Almanac* as a showcase for the importance, efficiency, and accuracy of their calculations (430). The production of astronomical tables and almanacs was also the stage for the development of computing as a form labour, so that by the time of comet Halley's return in 1835 the computing rooms resembled factories (Grier, 2007, ch.3).

The cosmic imaginary of almanacs ties prediction, data, and computing with geometries from the cosmos. Monism and the "accountant's view of



198

TABLE I.

Valeur, à la fin de n années, de 1 franc placé à intérêt composé.

Valeur à la fin de n années .. $(1+r)^n\, 1^{fr}$.

ANNÉES n.	TAUX DE L'INTÉRÊT r. 2 ½	3	3 ½	4
	fr	fr	fr	fr
1	1,025 000	1,030 000	1,035 000	1,040 000
2	1,050 625	1,060 900	1,071 225	1,081 600
3	1,076 891	1,092 727	1,108 718	1,124 864
4	1,103 813	1,125 509	1,147 523	1,169 859
5	1,131 408	1,159 274	1,187 686	1,216 653
6	1,159 693	1,194 052	1,229 255	1,265 319
7	1,188 686	1,229 874	1,272 279	1,315 932
8	1,218 403	1,266 770	1,316 809	1,368 569
9	1,248 863	1,304 773	1,362 897	1,423 312
10	1,280 085	1,343 916	1,410 599	1,480 244
11	1,312 087	1,384 234	1,459 970	1,539 454
12	1,344 889	1,425 761	1,511 069	1,601 032
13	1,378 511	1,468 534	1,563 956	1,665 074
14	1,412 974	1,512 590	1,618 695	1,731 676
15	1,448 298	1,557 967	1,675 349	1,800 944
16	1,484 506	1,604 706	1,733 986	1,872 981
17	1,521 618	1,652 848	1,794 676	1,947 900
18	1,559 659	1,702 433	1,857 489	2,025 817
19	1,598 650	1,753 506	1,922 501	2,106 849
20	1,638 616	1,806 111	1,989 789	2,191 123
21	1,679 582	1,860 295	2,059 431	2,278 768
22	1,721 571	1,916 103	2,131 512	2,369 919
23	1,764 611	1,973 587	2,206 114	2,464 716
24	1,808 726	2,032 794	2,283 328	2,563 304
25	1,853 944	2,093 778	2,363 245	2,665 836
26	1,900 293	2,156 591	2,445 959	2,772 470
27	1,947 800	2,221 289	2,531 567	2,883 369
28	1,996 495	2,287 928	2,620 172	2,998 703
29	2,046 407	2,356 566	2,711 878	3,118 651
30	2,097 568	2,427 262	2,806 794	3,243 398
31	2,150 007	2,500 080	2,905 031	3,373 133
32	2,203 757	2,575 083	3,006 708	3,508 059
33	2,258 851	2,652 335	3,111 942	3,648 381
34	2,315 322	2,731 905	3,220 860	3,794 316

13

Lune.

Jours du mois.	PASSAGE au méridien	LEVER.	COUCHER.	AGE.
	h m	h m	h m	
1	7. 58 Matin.	4. 9 Matin.	11. 42 M.	25
2	8. 51	5. 16	0. 25 Soir.	26
3	9. 46	6. 13	1. 20	27
4	10. 41	6. 59	2. 28	28
5	11. 36	7. 34	3. 46	29
6	0. 29 Soir.	8. 1	5. 3	1
7	1. 20	8. 22	6. 25	2
8	2. 9	8. 40	7. 47	3
9	2. 56	8. 55	9. 13	4
10	3. 44	9. 11	10. 35	5
11	4. 34	9. 28	11. 59	6
12	5. 26	9. 48	— —	7
13	6. 22	10. 14	1. 24 Matin.	8
14	7. 21	10. 47	2. 48	9
15	8. 22	11. 33	4. 7	10
16	9. 23	0. 35 Soir.	5. 12	11
17	10. 21	1. 47	6. 2	12
18	11. 15	3. 6	6. 39	13
19	— —	4. 25	7. 5	14
20	0. 4 Matin.	5. 41	7. 24	15
21	0. 50	6. 55	7. 40	16
22	1. 33	8. 5	7. 55	17
23	2. 13	9. 15	8. 7	18
24	2. 54	10. 23	8. 21	19
25	3. 34	11. 33	8. 34	20
26	4. 17	— —	8. 50	21
27	5. 1	0. 43 Mat.	9. 11	22
28	5. 49	1. 54	9. 39	23

Planètes.

Jours.	LEVER.	COUCHER.	PASSAGE au méridien
☿ MERCURE.			
	h m	h m	h m
1	8. 10 Matin.	5. 57 Soir.	1. 3 Soir.
11	7. 53	6. 50	1. 22
21	7. 11	6. 53	1. 2
♀ VÉNUS.			
1	4. 30 Matin.	1. 37 Soir.	9. 1 Matin.
11	4. 32	1. 28	9. 0
21	4. 35	1. 31	9. 3
♂ MARS.			
1	2. 5 Matin.	11. 21 Matin.	6. 43 Matin.
11	1. 56	10. 56	6. 26
21	1. 45	10. 33	6. 9
♃ JUPITER.			
1	11. 58 Soir.	10. 28 Matin.	5. 15 Matin.
11	11. 19	9. 50	4. 37
21	10. 40	9. 4	3. 57
♄ SATURNE.			
1	7. 48 Matin.	5. 13 Soir.	0. 31 S.
11	7. 12	4. 41	11. 56 Mat.
21	6. 35	4. 8	11. 22
♅ URANUS.			
1	4. 48 Soir.	7. 49 Matin.	0. 21 M.
11	4. 7	7. 9	11. 36 Soir.
21	3. 25	6. 29	10. 55

N. L. le 6, à $8^h\ 4^m$ mat. — P. L. le 20, à $8^h\ 10^m$ mat.

P. Q. le 13, à $5^h\ 29^m$ mat. — D. Q. le 28, à $10^h\ 1^m$ mat.

Source gallica.bnf.fr / Observatoire de Paris

Source gallica.bnf.fr / Observatoire de Paris

Figure 5.2: *Annuaire pour l'an 1875* (Bureau des longitudes, 1875) showing tables for interest on loans (left), and positions of planets (right).

the world" can be seen literally in the pages of almanacs, for example with tables for unit conversions, or with the positions of planets and interest rates for loans using similar table layouts and visual language [fig. 5.2]. The modern almanac materialises a shift where considerable resources were put towards the observation of diagrams in the sky, their recording and ordering as data, and their transposition to legitimise financial speculations and social ideals. The legacy of these efforts endures, and is taken for granted, in the current imaginaries of data. However, remembering the cosmic aspect of our relationship with data is also a form of rupture. The cosmic imaginary of almanacs bypasses the separation of predictive techniques along lines of scientific rationality and validity to place algorithmic prediction in a space shared with other predictive geometries such as astrology.

5.2.2 Data for divination

The cosmic imaginary of almanacs is not only about establishing astronomical data and associated techniques as scientific sources of predictions. This imaginary spans a multiplicity of scientific theories, traditions, and folklores. The *Old Farmer's Almanac* calendar [fig.5.3] shows the visual vernacular of the cosmic imaginary, combining data from lunar cycles, tides, weather, biblical events, astrology, with americana, proverbs, and customs. This mix, alongside the other sections of the almanac which might offer recipes, gardening tips, and other trivia, is 'downright *internetty*' according to LaFrance (2015) when seen from our current cultural moment. Like the internet, the almanac is a 'happy grab bag of marginalia' (Macgregor, 1997). Zooming out from single almanac publications we see the genre as a whole, ranging from nautical almanacs at the scientific end, to *The Old Farmer's Almanac*, and *Old Moore's Almanack* at the tabloid end of the spectrum.

Almanacs are part of the foundations of the authority of data and computing, what later became algorithmic prediction as we know it. They show these foundations as part of a bigger set of practices, multiple ways of relating cosmic movements to events on earth. Algorithmic prediction is shown in its construction, having more in common with astrology and folk theories than might be acknowledged today. Forms of prediction and speculation that have since been de-legitimised, such as astrology and gambling,[4] continue to thrive in almanacs, as illustrated by the *Old Moore's Almanack* predictions of lottery numbers by astrological sign [fig.5.4]. The tone of almanacs also suggests that, unlike the big promises

[redacted]

Figure 5.3: Old Farmer's Almanac calendar (Thomas, 1976). (redacted)

[redacted]

Figure 5.4: Euro Million Astro Indicator (Old Moore's Almanack, 2017, 78). (redacted)

of the "Big Data" imaginary, they do not take themselves too seriously. A 2017 issue of The Old Farmer's Almanac bills itself as 'useful with a pleasant degree of humour' while another cover, from 1847, advertises 'new, useful, and entertaining matter.' (Thomas, 1847). These tag-lines seem to suggest that all predictions in almanacs, however scientific or useful, should be taken with a grain of salt. My point here is not to dismiss the validity of almanacs, but to put its various modes of producing predictions, ranging from astrology to data-science, on par with each other as modes of 'ask[ing] a more-than-human intelligence for guidance' (Curry, 2016), that is, as divination.

Viewing almanacs as collections of divinatory techniques helps to see past the differentiation along the lines of scientific legitimacy—i.e. considering which methods of prediction are still considered valid today, and which are dismissed—and to think of them as modes of querying the cosmos for predictive diagrams. This transfer from the stars to earth is never a straight line, it involves a defining characteristic of divination: the chicane. Whether they are angles and aspects in astrology, smooth curves in social-science, or tables in finance, cosmic data always go through a chicane, a 'trick or subterfuge' (Ramey, 2016, 65) to be turned into predictions. Chicanes are sharp double-bends in roads or race-tracks. They are also, according to Cornelius (2016), a 'defining characteristic of divination' [114], a 'play of semblances' through which divinatory practice produces its outcome (predictions, but also healing, or other effects). This is not always a 'trick' of the diviner on the client, pretending that an effect is real or that an interpretation is true, but a mutual performance between the two[5]. While the prediction may '*feel true* or *ring true*', it is not absolute. 'The client does not have to agree, ever, that the diviner knows the unknown perfectly (or has channelled it appropriately)' (Ramey, 2016, 66). The chicane as 'double thinking' means that the symbols and metaphors used in divination have a wide range of interpretations, and that even mutually exclusive readings can be entertained at the same time (66).

[4]The study of games of chance was instrumental in the development of probability theory. However, considerable efforts were made to legitimise scientific and financial probability, and to separate them from the field of gambling, seen as immoral. see for example Cronon (1991) or Gigerenzer et al. (1990, 268).

[5]Cornelius describes the interaction between Azande witch-doctors (as notoriously studied by anthropologist Evans-Pritchard) and their patient as an example of chicane: 'The chicane is the symbolic instantiation of the cure in the physical removal of hte witchcraft object. The witch-doctor "knows" this even if he does not use our words.

One example of chicane at play is the method used by the *Old Farmer's Almanac* to produce its famously long-term weather predictions. Since its inception, the publication has linked sun-spot activity to weather patterns on earth (Boekmann, 2019). This method, based on 11 year solar cycles, is used to produce maps that indicate broad weather trends for a season or a year. It has survived through the history of climate modelling and alongside much more sophisticated forecasting techniques used by meteorological agencies.[6] Part of their success, LaFrance (2015) argues, is due to the fact that the maps only 'had to be right some of the time.' The single map for the year is open enough to interpretation to 'ring true' to the farmer, and to be read as accurate. This interpretive process allows for 'double thinking' as any weather and its opposite are easily accommodated under a single icon representing the "trend" for the year. The prediction is therefore as much *produced* by the map itself than by the farmer who verifies it through their lived experience.

Sun-spot theories have also been used to justify climate-denialism—including in the very pages of the *Farmer's Almanac* (Berman, 2019)—arguing that the current solar cycle indicates we are actually about to enter a 'new ice-age' (Landscheidt, 2003). This demonstrates that chicanes are not neutral, or always "sincere." I will come back to this in section 5.4, for now I turn to locating the almanac, and its cosmic imaginary, in the contemporary media landscape.

5.2.3 Almanacs in the petabyte age

Today the almanac and elements of the cosmic imaginary are alive and well, diluted in the current regime of algorithmic prediction. The "accountant's view of the world" has been amplified dramatically to the point of 'actuarial saturation' (Adams et al., 2009). Monism is taken for granted, as seen for example in the long relationship between physics and Wall Street (Weatherall, 2013), and in the high numbers of physics PhDs in the field of data-science. As astronomy and genomics battle for the title of "big(gest) data," (Stephens et al., 2015) the genome has replaced the horoscope: 'We used to think that our fate was in the stars. Now we know that, in large measure, our fate is in our genes.' (James Watson quoted in Davis, 1998, 157). Chicanes are still present, in the form of computer code.

But know does the patient "know" about this sleight of hand? It is not necessary for patients to be naive or credulous, even if they often are.' (Cornelius, 2016, 126).

[6] Lafrance discusses this with Paul Edwards, author of extensive work on the history of climate models and weather prediction (Edwards, 2010).

[redacted]

Figure 5.5: Old Moore's Almanac homepage (Old Moore's Almanac). (redacted)

These highly sophisticated computational intermediaries may hold some power of their own as 'society has become utterly enthralled by the idea of the code' (327), giving rise to new entanglements between computation, cybernetics, and the cosmic (Calheiros, 2012; Gauquelin, 1967; Landscheidt, 1973).

Almanacs are still being published, including *The Old Farmer's Almanac*, *Old Moore's Almanack*, and many others. Some are available online, such as the *Farmer's Almanac* and its 'daily calendar' [fig.5.6], or *Old Moore's Almanac* [fig.5.5].[7] Their endurance in a modern era testifies to the persistence of divination, suggesting that 'even a putatively modern, secular, and rationalist culture does not and cannot survive without some form of divination' (Ramey, 2016, 52). Astrology has done especially well on this front, as it successfully adapted to consumer capitalist society (Willis and Curry, 2004, 51), morphing with the times and integrating new scientific methods[8] and aesthetics [fig.5.6]. Although they use contemporary modes of distribution, and keep abreast of the 'latest

[7]This is the Irish Old Moore's Almanac, not to be confused with the British Old Moore's Almanack cited earlier.

technological trends' [farmer's almanac 2017], almanacs do seem like quaint relics in the current media landscape. They don't hold the comparison with the vast infrastructures and networks currently deployed for storing data and computing predictions. But there is more to the continuation of the cosmic imaginary. I argue that this persists in a variety of other ways, still connected to the 'cultural space' of almanacs but fragmented across a wide range of digital media and services.

The navigational aspect of the almanac survives through an obvious lineage with GPS and map apps, now enhanced with predicted routes and "Estimated Times of Arrival." It also endures in the ubiquitous form of the data dashboard, used to manage anything from businesses to cities and countries, channeling mechanical metaphors in a wide range of settings, promising a governance where one 'fl[ies] by instrument' (Mattern, 2015). Finance also has its almanacs, with feeds such as the *Bloomberg Terminal*, or multi-screen desktop rigs relaying high-frequency trading signals. The art of technical analysis (Murphy, 1999; Archer and Bickford, 2007), the reading of angles, ratios, and alignments in financial charts, is a form of financial astrology popular with traders on crypto-currency markets where high volatility produces a lot of patterns.

In a striking evolution of the almanac's cosmic imaginary, astrology merges with "A.I." in the *Co—Star Astrology* iPhone app (Co – Star Astrology). The search for meaning beyond 'technorationalism' is combined with the promises of the data imaginary to 'reveal' hidden insights in 'real time'; this astrology is 'Powered by NASA and generated with AI.' [fig.5.7]. This astrology-as-a-service approach claims to sidestep hyper-rationality and modern life anxieties, while relying on the very technologies that underpin them. The imaginary promoted by *Co—Star* is a new twist on the data imaginary, where a proprietary software black box combines personal information with cosmic data to produce an almanac of sorts, a new blend of scientific validity and mystical appeal. In another re-purposing of astrological lore at the service of platform capitalism, Amazon provides horoscopes to its Prime members with recommendations for its products [fig.5.8]. More genuine expressions of astrological practice in the connected age are found on social media platforms such as Instagram or Reddit, under tags such as `#astrologymemes`, where it is mixed with the vernacular of meme culture.

[8] Examples of systematising and modernising astrology include the Gauquelin's efforts in the 1960's (Gauquelin, 1967, discussed in Willis and Curry, 2004, 6) and others such as *Cosmic Cybernetics* by Landscheidt (1973).

[redacted]

Figure 5.6: Farmer's Almanac daily calendar web page (Farmers' Almanac). (redacted)

[redacted]

Figure 5.7: *Co—Star* astrology iPhone app (Co – Star Astrology). (redacted)

These examples show ways in which the almanac survives, and how its cosmic imaginary can be brought to bear on the current media landscape of prediction. With this backdrop established I now turn to using the almanac as a site for creative practice.

5.3 Practice

After establishing the almanac as a topos linking computation, prediction, the cosmos, and divination, I now turn to revisiting this publication through practice by making my own version of the almanac. I have shown how the almanac, or some of its sections, can be seen to endure to this day through a wide range of digital media. In this project, I bind these sections back together, and bring this peculiar predictive artefact and its history to bear on contemporary algorithmic prediction.

5.3.1 Computation and divination in critical practice.

As a popular culture artefact, the almanac presents an opportunity to consider data science and divination from an aesthetic perspective. In

[redacted]

Figure 5.8: Amazon Prime horoscope (Katz, 2019). (redacted)

recent years, a number of art and design projects have explored the interplay between data, divination, and computation. For example, American artist Ingrid Burrington creates astrological charts for the Five Eyes[9] spy agencies (Burrington, 2014). Artist collective RYBN's *The Golem* is a computer that applies ancient kabbalistic hermeneutics to its own processes and daemons (RYBN, 2017). RYBN are also the organisers and facilitators of *ADMXI* (2015), a platform for 'heretic, irrational and experimental' trading algorithms designed by artists. In a similar vein, designer Shing Tat Chung's *Superstitious Fund* trades on the stock market according to 'lunar cycles and numerology' as well as an internal logic of 'lucky and unlucky values.' (Chung, 2012). Computational poet Allison Parrish teaches a class at NYU's Interactive Telecommunications Program that interrogates forms of divination in digitally-mediated environments, 'from the casting of lots to computer-generated randomness to the contemporary revival of Tarot; from reading entrails to astrology to data science.' (Parrish, 2017). Artists Wesley Goatley and Tobias Revell's piece *Augury* references Greek and Roman divinatory practices based on patterns found in the flight of birds. They use data from aeroplane

[9]Five Eyes is an alliance between the signals intelligence agencies of Australia, Canada, New Zealand, the United Kingdom, and the United States.

positions within a 20km radius in combination with local tweets to produce predictions via a neural network, 'as a parody of the contemporary obsession with algorithmic prediction' (Goatley and Revell, 2018).

While images/stories of magic and divination have long been leveraged to promote the supernatural powers of computers (Stahl, 1995), these critical approaches in art and design have a different aim. They use divination as a reminder that despite dominant narratives of objective data-science, technological progress, and computational powers, claims of predicting the future are underpinned by social and political constructs. They develop cultural languages to examine and/or critique opaque technological systems of prediction and control. In some of the cases cited above (RYBN, Chung) this critique is directed at predictions on financial markets, where the "accountant's view of the world" uses the veneer of scientific rationality to legitimise speculative practices.

These works interrogate the relationships between data, computation, prediction, and divination through making and creative practice. They do this in different ways. Some invoke the lore of divinatory practice as a parodic device (Goatley and Revell) while others create more ambiguous, and arguably more genuine, moments of divinatory potential from computational processes (RYBN's *Golem*). Marenko and Brassett (2015) looks at digital making as a way of foregrounding the uncertainty and contingency inherent in computation. She focuses on the 'glitch event' as a moment of opening of unexpected potentials, an 'irruption of the unplanned' that sound artist Kim Cascone calls 'Errormancy' (Cascone, 2013). To achieve their effects, the artists cited above often re-purpose or reconfigure computational systems to provoke such glitches, and open them up to interpretation. Majaca and Parisi (2016) on the other hand, warn against the fetishisation of errors as just another form of mystification. They write:

> …celebrating error for its own sake is a form of mystification that can only lead to depoliticized, naive triumphalism.

In their view, celebrating the creative potential of errors is not enough. If one is serious about reclaiming contingency and uncertainty, one should ask instead whether these errors are 'fatal to the system or entirely anticipated' (Majaca and Parisi, 2016).

Whether or not we assign divinatory potential to computational errors, the selected art and design projects cited above all seek ways to engage with these questions through practice. I situate my work on *The Monistic*

Almanac in this field of digital making that critically interrogates algorithmic prediction as a diagram between data, computation, and divination. In the next section I discuss my own contribution to this field, first by presenting the tools I employed and then the widgets that make up my version of the almanac.

5.3.2 Almanac tooling

The practice of building *The Monistic Almanac* involved getting acquainted with a range of data science tools including: data-sets and APIs from various sources; software libraries for tasks such as: data handling, machine learning, astronomical and astrological computations; interactive notebooks, and code editors.

Almanac widgets produce predictions by putting data from a particular domain (finance, the power grid, daily life) in *relation* with data from cosmic bodies (planets, angles, ratios). For the former, data are drawn from publicly available sources, many of them proprietary, such as the Quandl API for financial data,[10] the Elexon portal for UK power grid data,[11] and historical UK National Lotto Winning Numbers from unofficial sources.[12] For the later, software packages such as Skyfield (Rhodes) provide an interface to the Lunar and planetary ephemerides from NASA's Jet Propulsion Laboratory (Folkner et al., 2014). These are datasets providing positions of planets from deep in the past to deep in the future; in the case of the DE431 ephemeris it covers years –13,200 to +17,191. There are multiple versions of ephemeris data, each offering a different trade-off between precision and file storage size. Flatlib for example—a software package aimed at developers of astrology software (Ventura)—uses the *Swiss Ephemeris*, a modified version of the JPL ephemeris, compressed to reduce storage (Koch and Treindl, 2014). Negotiations between storage size and precision comparisons down to the milli-arcsecond raise the question of uncertainty of these highly precise data sets. Although they are jewels of scientific achievement and sophistication, the accuracy of these cosmic data is always uncertain and contested, as Stanley (2013) discusses for the case of lunar eclipse data. This aspect is conveniently glossed over when using the software packages, as they simply make the "clean" data available to programmers and developers.

Once acquired, data need to be "cleaned" and turned into *vectors* in order for them to be used in predictive operations. A key exchange surface between data-sets[13] and vector space is the dataframe object provided by the Pandas Python package (McKinney, 2010; McKinney). Through the dataframe, data from earth (e.g. stock market prices) and the cosmos (e.g. planet positions) are 'juxtaposed' in a 'common space' of vectors (Mackenzie, 2017, 73). In other words these data are put in *relation* with each other. The history of the Pandas package also highlights the "accountant's view of the world" coming full circle. While the business astronomers transposed data practices from astronomy to finance and business, Pandas originated as a tool to handle data in a hedge fund and is now widely used in the sciences, including astronomy (Kopf, 2017).

Once data are normalised and juxtaposed into a common space, predictive *operations* can begin. In *The Monistic Almanac* this takes various forms. Some follow the expected pipeline of data science, fitting rudimentary machine learning models to do the data using scikit-learn (Pedregosa et al., 2011) (e.g. Cosmic Commodity Charts in the next section), while others interpret the vectors and notions of distance in more idiosyncratic ways (e.g. Crisis Proximity Index in the next section). They all use a form of vector space as the basis for their operations.

Once predictions are produced, their results are visualised as charts, maps and tables as part of the publication. As with the other projects discussed in this thesis, *The Monistic Almanac* involved negotiations and frictions with data visualisation software (D3.js and Python's Matplotlib). These frictions involved moving away from the implicit assumptions of default settings and creating new, more deliberately designed forms. However, the visualisations in *The Monistic Almanac* have an added dimension to this exercise as they intentionally subvert the codes of data visualisations. They are more like dramatisations of the results of predictive operations as visualities, amplifying the meaning produced by turning a graphic

[10] Provided by https://www.quandl.com (accessed 19 July 2019). For example 'Corn Futures, Continuous Contract #10' https://www.quandl.com/data/CHRIS/CME_C10-Corn-Futures-Continuous-Contract-10-C10

[11] https://www.bmreports.com/bmrs/ (accessed 19 July 2019)

[12] http://lottery.merseyworld.com/ and http://lottery.merseyworld.com/Winning_index.html (accessed 19 July 2019)

[13] e.g. "flat files" such as CSV, JSON, or as returned from an API.

semiology (Bertin, 1967) into a semiotics of fortune telling (Aphek and Tobin, 1990).

5.3.3 Cast of widgets

I connect the tools described in the previous section 5.3.2 to construct my own chicanes through a layered process of 1) mapping cosmic and other data to the same vector space 2) operating on that space to produce predictions and 3) presenting these predictions as visualisations. I now turn to describing these widgets and their operations.

Cosmic Commodity Charts [fig.5.9] predict prices on commodity futures markets using the positions of the planets of the solar system. Historical price data are used as a "target" and planet positions[14] from the JPL ephemeris as "features" to "train" a support vector regression algorithm[15]. Once the statistical relationship between positions and price established in the "fit" of the model, future planet positions from the ephermeris are used to predict future prices. The chart displays the curve for the year ahead and highlights the high and low prices.

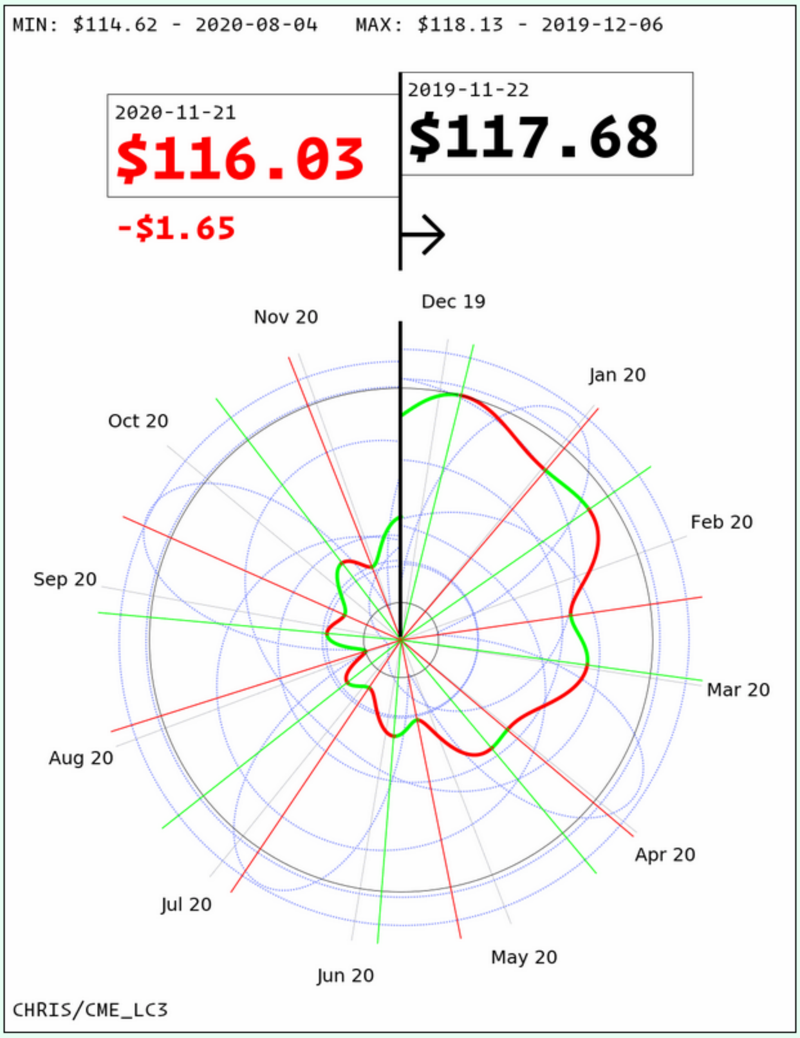
Crisis Proximity Index [fig.5.10] is an astrology based on the 2008 financial crisis. The Index is based on the reference point of August 9, 2007, when BNP Paribas froze three of its investment funds, triggering the first signs of panic among investors. Daily planet positions are compared to this base vector, using distance as an indicator of a possible new crisis. The fluctuation in distances are shown projected on the celestial sphere and as a table, both colour-coded to convey a negative (red) or positive (green) meaning respective of the distance decreasing or increasing.

Carbon Prophet [fig.5.11] uses Facebook's *Prophet* (Taylor and Letham, 2017), a tool to make probabilistic forecasts with business data (e.g. periodical sales), to predict the amount of carbon dioxide (CO_2) in the earth's atmosphere. The data are taken from *Carbondoomsday.com* a website that draws attention to the alarming shape of the *Keeling Curve* that describes the steady rise of the concentration of CO_2 (see Keeling et al., 2001). The probabilistic model produces an area of confidence around, and beyond, the training data. The graph shows this area around the historical data, and into the future.

[14]Specifically their distance from the Solar System's barycentre (centre of gravity).
[15]https://scikit-learn.org/stable/modules/generated/sklearn.svm.SVR.html

Full Live Cattle Chart

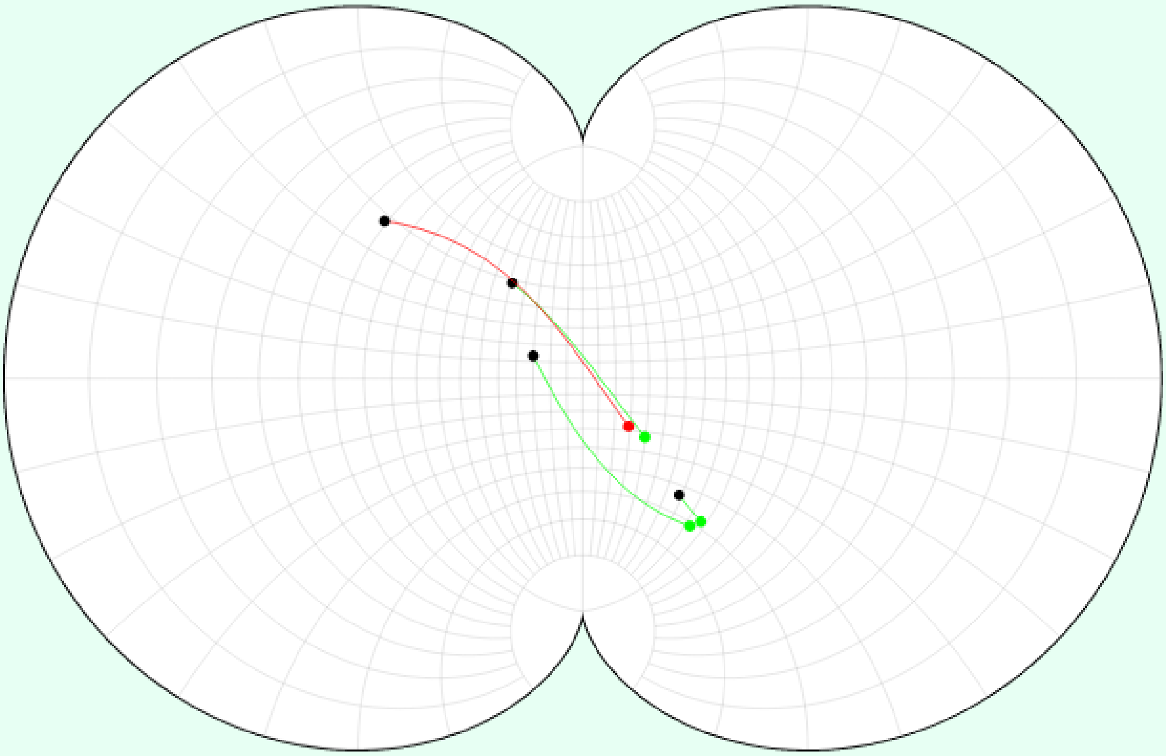
[☿♀♁♂♃♄⛢♆♇] 🐄LE 365Days



Opportunities		
◢◣	2019-12-06	$118.13
◣◢	2020-01-02	$117.66
◢◣	2020-01-17	$117.76
◣◢	2020-02-13	$116.82
◢◣	2020-02-29	$117.09
◣◢	2020-04-02	$116.28
◢◣	2020-04-12	$116.32
◣◢	2020-05-11	$115.34
◢◣	2020-05-27	$115.58
◣◢	2020-06-26	$114.94
◢◣	2020-07-14	$115.12
◣◢	2020-08-04	$114.62
◢◣	2020-08-27	$115.22
◣◢	2020-09-15	$115.02
◢◣	2020-10-14	$115.59
◣◢	2020-10-30	$115.29

Figure 5.9: Cosmic Commodity Chart: Full Live Cattle Chart.

Crisis Proximity Index



	☿	♀	♂	♃	CPI
2019-11-16	⬇ 1.53	⬆ 1.62	⬆ 3.61	⬆ 1.97	⬆ 8.72
2019-11-17	⬆ 1.53	⬆ 1.62	⬆ 3.61	⬆ 1.99	⬆ 8.74
2019-11-18	⬆ 1.53	⬆ 1.62	⬆ 3.61	⬆ 2.01	⬆ 8.77
2019-11-19	⬆ 1.54	⬆ 1.62	⬆ 3.61	⬆ 2.03	⬆ 8.8
2019-11-20	⬆ 1.55	⬆ 1.63	⬆ 3.61	⬆ 2.05	⬆ 8.83
2019-11-21	⬆ 1.56	⬆ 1.63	⬇ 3.61	⬆ 2.07	⬆ 8.86
2019-11-22	⬆ 1.57	⬆ 1.63	⬇ 3.61	⬆ 2.09	⬆ 8.9
2019-11-23	⬆ 1.59	⬆ 1.63	⬇ 3.61	⬆ 2.11	⬆ 8.94
2019-11-24	⬆ 1.61	⬆ 1.63	⬇ 3.61	⬆ 2.13	⬆ 8.98
2019-11-25	⬆ 1.64	⬆ 1.63	⬇ 3.61	⬆ 2.15	⬆ 9.03
2019-11-26	⬆ 1.66	⬆ 1.63	⬇ 3.6	⬆ 2.18	⬆ 9.07

Figure 5.10: Crisis Proximity Index.

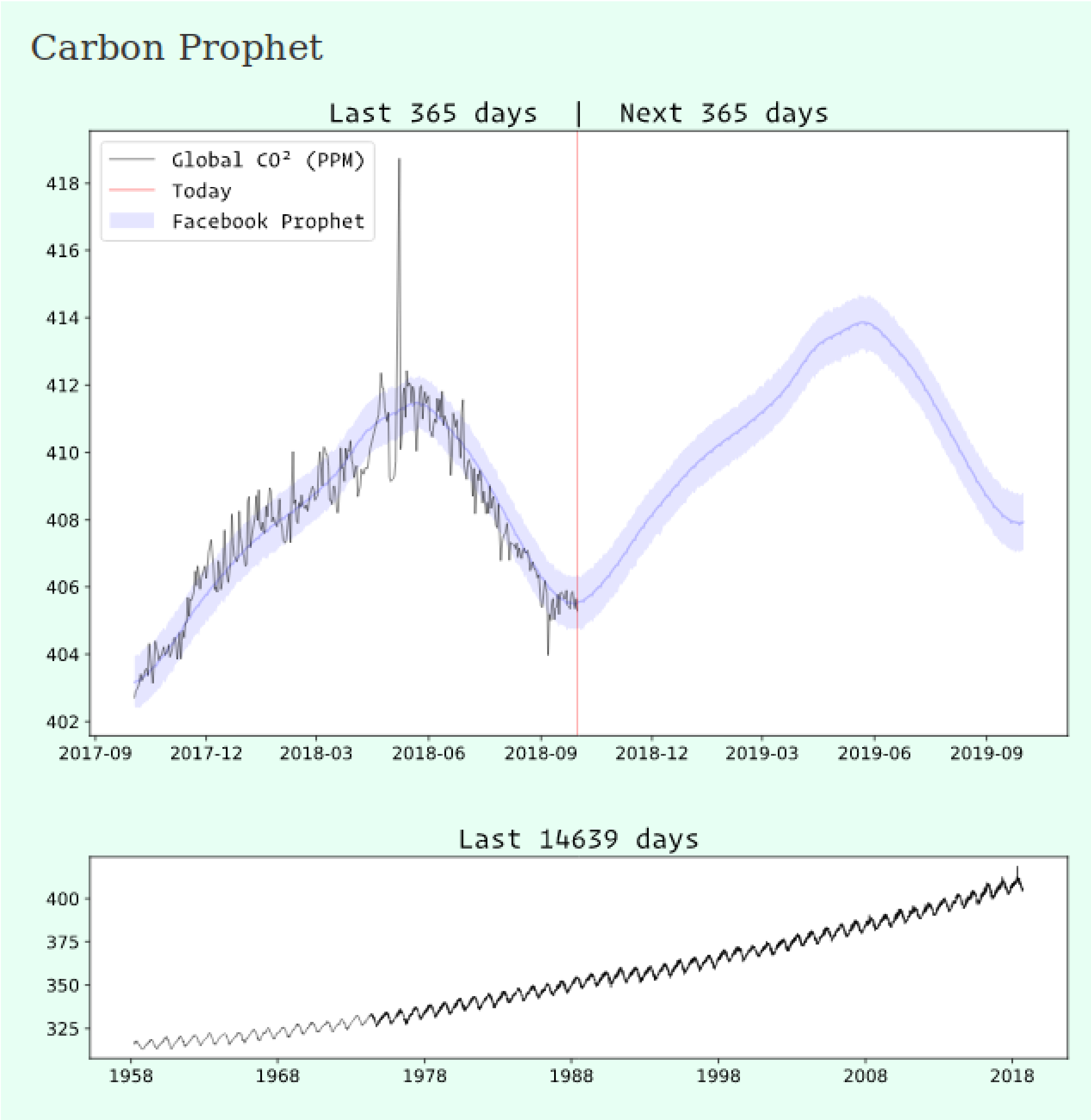
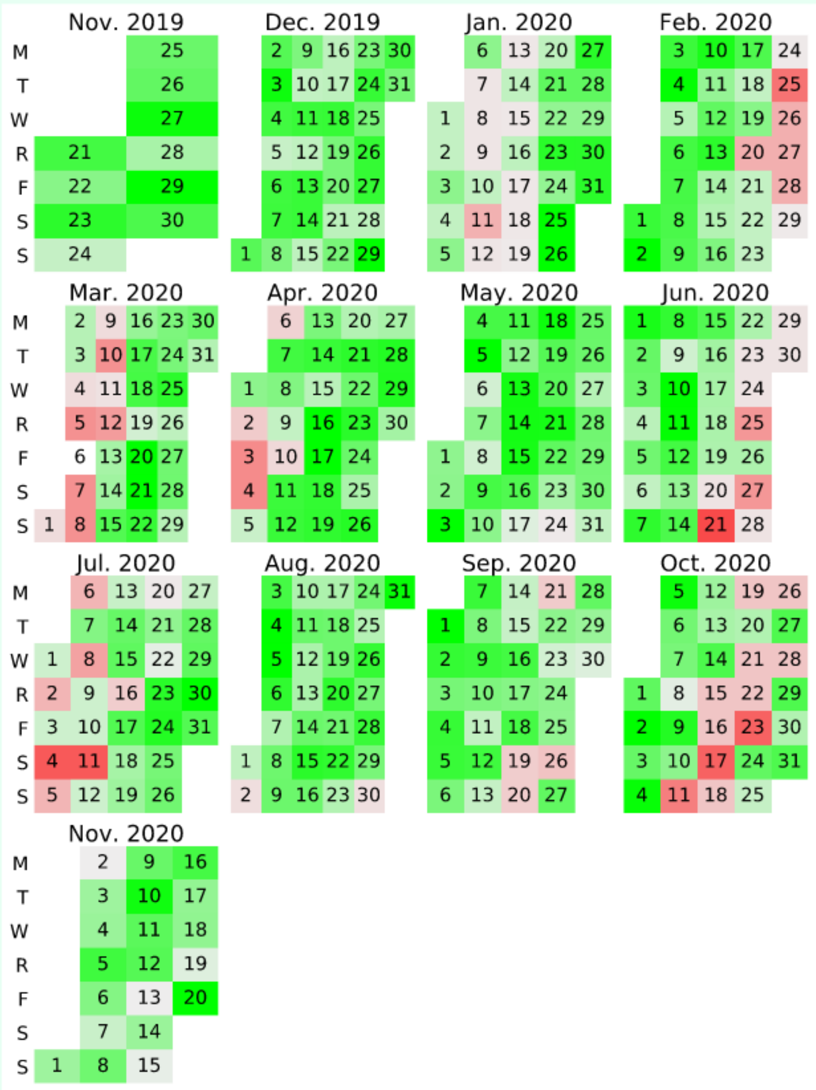


Figure 5.11: Carbon Prophet.

When to purchase a computer in MWADUI, TANZANIA



Best Days	Score
2020-01-25	0.800
2020-01-26	0.800
2020-01-27	0.642
2020-01-23	0.641
2020-01-30	0.637

Worst Days	Score
2020-03-06	-1.000
2020-10-17	-0.800
2020-07-04	-0.800
2020-06-21	-0.800
2020-07-11	-0.800

Figure 5.12: Electional Astrology Calendar: When to buy a computer in Mwadui, Tanzania.

Electional Astrology [fig.5.12] is a type of astrology that looks forward for the best days to plan particular tasks or events. In this widget I translated the set of criteria for a set of tasks—for example not to plan meetings when Mercury is retrograde (Orion, 2007)—into code that produces calendars for a specific tasks/location pair. The criteria are first computed individually before being aggregated into a score, grading each day of the year ahead for the task in question. This widget was further developed for a "live reading" setting as part of *Supra Systems, Office Rites* in October 2018. In this setting the date and place of birth of visitors was taken into account to produce personalised calendars on site at the Victoria & Albert museum [see appendix E].

Power Grid Harmonics [fig.5.13] uses elements from financial technical analysis to detect patterns in the UK power grid's fluctuations. The theory of 'Harmonic Trading' (Carney, 1999) is a form of technical analysis based on Fibonacci angles and numbers, promising profits by detecting the 'natural cycles of the market' (10). While this does not directly rely on cosmic data, the universal geometries of Fibonnacci numbers and their associated "golden" ratios are, in my view, completely aligned with the cosmic imaginary detailed in the previous section. I applied these techniques, using an implementation in Python (PythonParseltongue, 2017), to the frequency[16] of the UK power grid, suggesting some kind of harmony or "hum" to the fluctuations of electricity.

These widgets are all hosted on a Raspberry Pi computer that produces a new issue of *The Monistic Almanac* on a daily basis. Each section has a form of variation or temporality built in: the Cosmic Commodity Chart selects a different commodity future and a set of planets, the Electional Astrology varies different task and location, and the Carbon Prophet updates to the latest data.

[16]The frequency of the UK power grid reflects changes in supply and demand in the market, energy providers are legally bound to keep it close to 50Hz.

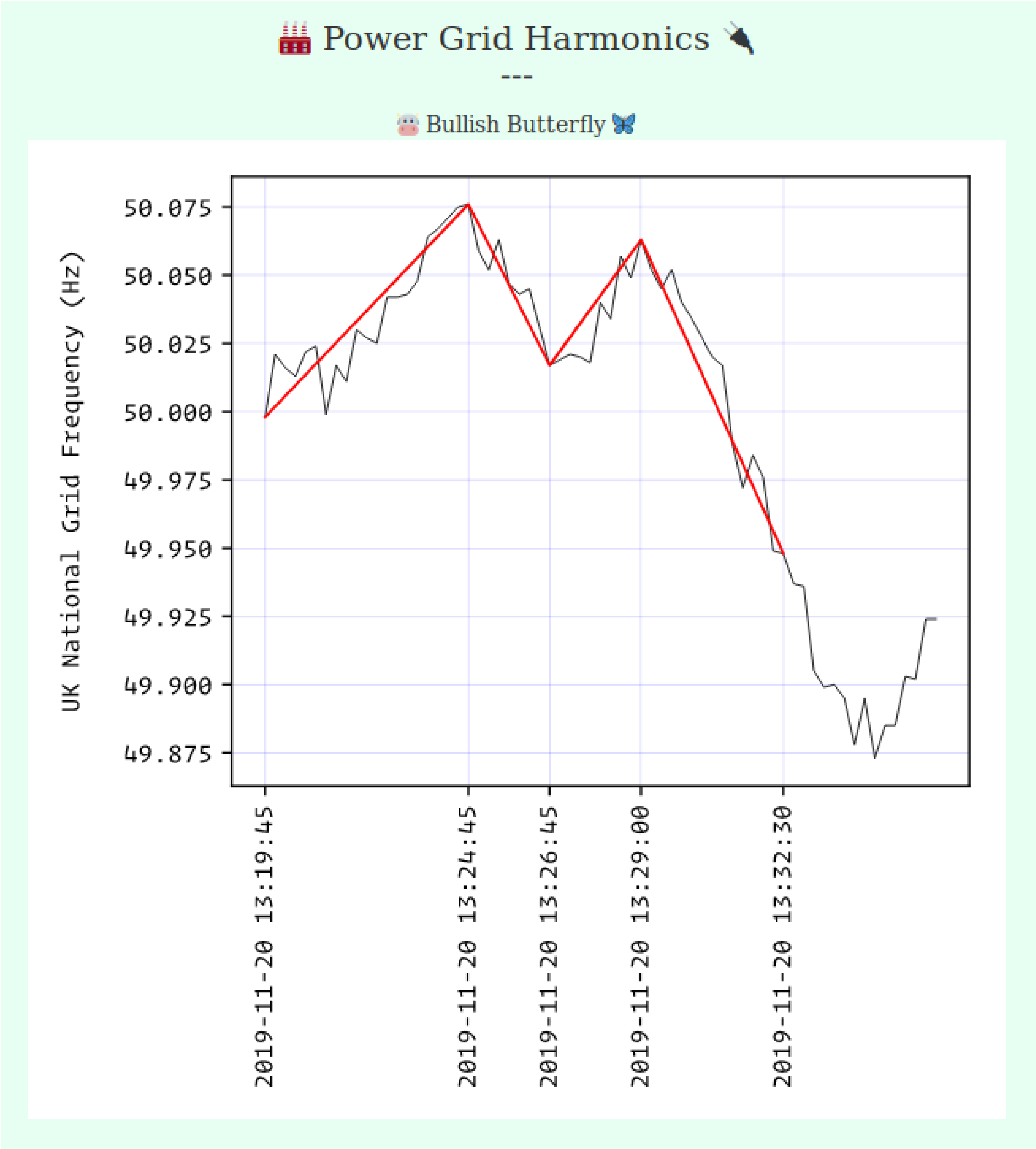


Figure 5.13: Power Grid Harmonics: Bullish Butterfly.

5.4 Discussion

In this section, I turn back to my research questions (RQ, RQ1, RQ2) and discussion criteria to reflect on *The Monistic Almanac* research project. The excavation of almanac publications brought up divinatory imaginaries of data and prediction, reified as diagrams between cosmic bodies and events on earth. I re-appropriated these forms in my own version of the almanac, playing tricks with the tools of algorithmic prediction. These tricks, or chicanes, are key to divinatory knowledge, and key to the legitimacy that is assigned, or not, to algorithmic prediction.

5.4.1 Sincere chicanery

In relation to RQ1, practical work on *The Monistic Almanac* involved producing a large number of data-diagrams. They proliferated in the exploratory process of constructing "cosmic" vector spaces and attempting to perform predictive operations. The key link between practice and research, however, was reading these experiments as *chicanes* that framed algorithmic prediction as a form of divination.

I began working on this project as a result of my skepticism towards the promises of algorithmic prediction. I took the cosmic imaginary of almanacs, rooted in monism and the "accountant's view of the world," as a creative license to put anything in relation to the cosmos through predictive models. I wanted to push the cosmic imaginary to its absurd extreme, in a similar vein as what Tyler Vigen's *Spurious Correlations* (2015) does with statistics as a whole, but in the more specific setting of the almanac. The widgets described in section 5.3.3 were chicanes from the start, in the sense that they are "tricks,"[17] experiments and mis-uses of data handling, visualisation, and machine learning tools that are neither good data-science nor genuine astrology. This initial position was one of satire, somewhat aligned with the 'parody' of divination by Goatley and Revell (2018, see 5.3.1). However, this changed through the project, as a more intricate diagram of relation emerged between computation, prediction, and divination. Drawing on Ramey's *Politics of Divination* and its coverage of chicanes, I saw that my "tricks" could be more than flippant jokes. Like *The Old Farmer's Almanac*, they became 'useful with a pleasant degree of humor.' This arc describes a change in my understanding of chicanes. I moved from questioning the validity of predictions to realising that chicanes are a central and immovable part of

[17]For more on the designer as trickster see Singleton (2014)

all predictions, and that questions are better directed at their sincerity. Ramey (2016, 66) describes this sincerity as the "honesty" of the negotiation between diviner and client as they are both involved in the chicane with the aim to produce a prediction that 'ring[s] true to the inquirer' (ibid.).

My first impulse was to use astrology to dismiss data science, to imply that they are equally pseudo-rational. Robin James argues this very well in her update of *The Stars Down to Earth*—Adorno's (1994) critique of the *LA Times* astrology column—for the big data era (James, 2015). She argues that forecasts, whether from astrology or data science aestheticise 'unfashionable superstitions' through charts and tables. Instead of predictions, they produce conservative prescriptions, 'only ever reproduc[ing] society and its most conventional norms, values, and practices.' By bringing the stars down to earth, both data science and astrology insist that society, like a planet, must be on a regular and stable orbit. In James and Adorno's arguments, the problem with predictions is that they are produced through 'pseudo-rationality,' in other words through chicanes. The very presence of this trickery invalidates both astrology and algorithmic prediction as forms of deception. James' compelling argument concludes with a call to 'shoot for the stars' instead of bringing them down to fit a conservative view of the future. However, she leaves a big question un-answered: how to deal with the uncertainty of the future? If data-science and astrology are dismissed in equal measure, is there a "true" rationality, a direct way to engage with chance that would avoid any chicanes? This position risks circling all the way back to the digital-positivism of the data imaginary.

One way around this dead-end is to follow Ramey's (2016) injuction to take divination seriously as a 'generic, even universal dimension of human culture' (49). He argues that humans have had the need to 'read chance aloud' [29, quoting Heimlich]—to relate to it in some way—since ancestral times. However, these relationships to chance are always mediated, never direct or apolitical. Chicanes are an integral part of this mediation, perhaps even a condition for divination to be successful:

> The bluff is not in the diviner who pretends to know something that the client does not know. The client agrees that the diviner knows something that she does not or cannot know, otherwise than through divination. [...] true divinations occur only *as* true performances or, as it were, well-played games. (Ramey, 2016, 66)

The focus, then, moves from questioning the mere presence of chicanes to examining their sincerity. Ramey and Cornelius argue that both diviner and client partake in the chicane to produce predictions. The "trick" here is not played on the client but a mutual, ritualised suspension of disbelief, what Ramey calls 'sincere chicanery' (66).

The sincerity, not the existence, of chicanes is what should be interrogated in algorithmic prediction. For Ramey the neoliberal market caters to the ancestral need for divination, but betrays it through an insincere chicane. By presenting itself as an objective and rational oracle, the market obscures the chicane and presents political choices as natural and inescapable. The 'double thinking' and interpretation of divination are replaced by a one-way foreclosure of sub-optimal and unprofitable futures. As many market predictions are algorithmic, and market logics are applied far beyond trading stock, I argue that Ramey's points can be extended to algorithmic prediction more broadly. Data-positivism is also a form of insincere chicanery, as it pretends to differentiate itself from divination and produce objective and rational predictions.

Through building *The Monistic Almanac*, divination emerged as a critical position from which to acknowledge the presence of chicanes in algorithmic prediction, and to pin-point insincere claims that obfuscate them. The rich, cosmic imaginary of almanacs persists but it is fragmented, and the "accountant's view of the world" has claimed most of the legitimacy. Recognising the existence of chicanes may be a way towards questioning these credentials, and restoring the multiplicity of voices we glimpse in almanacs, including those who Isabelle Stengers calls the 'story-tellers, quacks, popular customs and creeds, knowledge without credential'; (Stengers, 2011, cited in Curry, 2016, loc.252)

In summary the "abductive arc" of this project revolves around the chicane. It is a change from considering its mere presence as grounds to delegitimise both astrology and data-science to acknowledging that chicanes are always present as mediators in our relationship to chance. Reckoning with chicanes helps to reframe critiques directed at their existence towards their sincerity, to interrogate the politics and credentials of algorithmic prediction.

5.4.2 Stochastic Arrows

If chicanes are instrumental to the production of divinatory predictions, another diagrammatic form is at play when they inform actions: the arrow.

This form of vector relates to RQ2 and to conjectural knowledge. Astrology scholar Gieseler Greenbaum (2016) discusses the stochastic nature of astrology, not in the modern meaning of stochasticity associated with random processes, but in the ancient Greek meaning 'to "aim at", "conjecture" or "guess" ' (174). This echoes Ginzburg (1980), discussed in Chapter 4, and places the notion of stochasticity at the origin of both the modern concept of science and divinatory practices, in particular astrology. Stochasticity, for Gieseler Greenbaum, evokes the metaphor of the archer. She uses this to discuss the relationships between goals, targets, and aiming in astrology, a way to separate the veracity of predictions from their effect: 'Even if the target, the correct prediction, is not achieved, the goal, helping the client, can still be achieved' (181).

Taking this back to *The Monistic Almanac*, once chicanes have played their "tricks" through various diagrammatic operations, the results are "flattened" back into charts and tables. These visualities (Leese, 2016) are designed to trigger affects and reactions, for example through the use of colours such as green and red to denote positive or negative predictions. Reading these as "arrows," the fact that these predictions are likely not accurate by any data-scientific or astrological standards does not matter. Instead, their mere existence already produces an impulse, a direction, that influences the reader in however small ways. One anecdote from the *Supra Systems Office Rites* exhibition/performance (Oct. 2018) illustrates this. As I was producing personalised Electional Astrology calendars, two visitors separately asked for "when to get married" calendars. They later came back and told me they were indeed planning to get married in the year, and were super-imposing their calendars and holding them to the light to find an overlap in their best dates.

This highlights the potential for the performative nature of prediction, where regardless of the purported accuracy or objectivity of the method involved, predicting is inescapably prescribing. This is problematic when it is obscured behind claims of objectivity, when in fact it is used to foreclose futures (see discussion of Ramey in section 5.4.1), as with many cases of control diagrams. The moment I shared with the exhibition visitors points to an oscillation towards a more generative diagram. As we all acknowledged the chicane at play—i.e. we were all aware that the prediction came from a program of my making that was based on an arbitrary encoding of astrological principles—we produced a best day for them to get married in a way that "rang true" (Ramey, 2016, 66) to both parties.

5.4.3 Automated publication of predictions

The practice work discussed in this chapter addresses the question of publication in two different ways. First, and foremost, focusing on almanacs as a way to unpack entanglements between algorithmic prediction and divination aligns with the view of publication as a 'site specific gesture and critical intervention' (Gilbert, 2016, 20). Secondly, the project also aligns with media archaeology as it focuses on an overlooked published artefact to challenge dominant narratives about contemporary technology.

In practical terms, *The Monistic Almanac* explores novel ways of producing publications. The automated system produces a new issue every day which contains predictions for the next 365 days—following the yearly format of many almanacs. This process is orchestrated through a series of scripts, one for each widget, that output data files (e.g. CSV, JSON, GEOJSON). These are produced on the almanac.computer, a Raspberry Pi, and published to the internet via a version control system (GIT) and automated deployment. Each issue is a minimal and lightweight set of data files completely separate from the code used for display.[18] Issue data can be seen by exploring the repository, and past issues can be re-published by simply replacing the data folder `current_issue` with the desired archived issue.

Another publishing format was key to my work on *The Monistic Almanac*: the Jupyter notebook (Pérez and Granger, 2007; Kluyver et al., 2016). This is an interactive environment where text, images, executable code, plots, and other outputs can all be combined for exploratory, reproducible research[19]. The notebook is a staple of the data-science world and, according to journalist Somers (2018), challenges the very foundations of scientific publishing. It also, incidentally, reflects some of the qualities of almanacs in its openness to multiple uses, languages, and media assemblages. While it originated in *Mathematica*, software created by Stephen Wolfram (Somers, 2018), the notebook came into its own when it was re-appropriated as an open-source project. Somers contrasts the centralised 'cathedral' approach of Wolfram with the 'bazaar' of open-source (drawing from Eric Steven Raymond, 2000). This resonates with the multiplicity of languages and rationalities I have discussed in the

[18] 19 files and approximately 800 kilobytes per issue.

[19] I combined work in Jupyter notebooks with other notebook-like environments, such as the Hydrogen plugin for the Atom code editor. This replicates the features of the notebook (e.g. kernel and cell-based execution, inline output of data and graphs) but in a code editor.

almanac. While Jupyter notebooks are generally used for scientific demonstrations, tutorials, or research, they seamlessly allowed more experimental creative work. In *The Monistic Almanac* I use notebooks as a kind of appendix to each widget, showing the data-sources, code, and chicanes that lead to the published results [see example notebook in appendix E].

5.5 Conclusion

In this chapter, I have discussed *The Monistic Almanac*, a project that revisits almanac publications as a site to interrogate the boundaries and entanglements between algorithmic prediction and divination. The almanac occupies a singular cosmic imaginary where data from the stars and events on earth are geometrically intertwined, more (e.g. astrological ratios) or less (e.g. statistical smoothing techniques) directly. Within this "space," I used the *chicane* as a common ground between research and practice to interrogate the "tricks" and mediations between cosmic data and predictions (RQ1). I implemented such tricks in practice as I programmed my own version of the almanac, comprised of widgets that combine data diagrams with astrological rationalities. Through these experiments, the chicane emerged as a focus point to interrogate algorithmic prediction critically in new ways. With the chicane acknowledged as an inescapable mediation in the production of predictions, the question is not about whether or not trickery exists, but about how *sincere* it is, in other words to consider its politics (Ramey, 2016). This helped me move away from a satirical position that took the mere presence of chicanes as grounds to dismiss both astrology and data-science. Instead, I took chicanes seriously as a diagrammatic form that relates algorithmic prediction with divination. Another form, the arrow, connected these considerations to antique notions of stochasticity (Gieseler Greenbaum, 2016), where the 'aims' and 'goals' of vectors are part of a conjectural mode of knowledge production (RQ2). These reflections were materialised through the production of an automated almanac publication, that anchored my research in specific entanglements between cosmic data, diagrammatic predictions, and their visualisations.

In this chapter, I discussed my third contribution to knowledge in this thesis, the diagrammatic form of the chicane as a way to conceptualise the mediations performed by algorithmic prediction through the lens of divination. Considering chicanes helps to side-step overly simplistic satirical positions that aim to bring algorithmic prediction "down" by comparing it to practices such as astrology. These arguments, including my own at the start of the *Monistic Almanac* project, leave the positivist ideal of a "ground-truth" unchallenged. If chicanes—and therefore mediation—are seen as an intrinsic part of any divinatory process, then critical questions can shift to the more fruitful ground of their politics. In practice this is materialised through the production of a new almanac, a piece of 'media imagined' (Blegvad) that produces predictions through a series of astrological rationalities implemented with the tools of algorithmic prediction, and publishes them in an automated fashion.

This marks the end of the third and final practice-oriented chapter in this thesis. In the following conclusion I turn back to my research questions and summarise how they have been addressed in this research. I then open up potential routes for building on this research in future work.

Chapter 6

Conclusion and Future Work

This thesis is, by now, itself a dense diagram of citations, (cross-)references, figures, external links, code, and so on. The table of contents, thesis repository in Appendix B—that relates in turn to the repositories for practice projects—and the bibliography give some views into the diagrams of relations produced by this research. They bind the pieces of this investigation together, and evidence how I have addressed my primary and secondary research questions.

Main research question RQ: How can diagrams be used as a language to critically investigate algorithmic prediction through design practice?

From the outset, I considered diagrams as intrinsically linked to an archaeological approach. They are a spatial and material entry point for an investigation of algorithmic prediction that is fully compatible with the technical literature (e.g. mechanisms of machine learning, linear algebra) while also affording critical perspectives that unsettle the very foundations of dominant imaginaries such as "Artificial Intelligence." This relies, in a large part, on an attention to historical narratives (e.g. continuations and ruptures) and overlooked artefacts. As it excavates algorithmic prediction as diagrammatic media, this approach makes it available for interrogation, contestation, and re-appropriation through practice.

In chapter 2, I outlined three registers for this diagrammatic language: data diagrams, control diagrams, and speculative diagrams. In the subsequent three practice chapters I put this language to the test of usage, learning to read and write it, and stretching it all the way from practical to theoretical considerations. This process did not result in a finite vocabulary or grammar—that would be falling back onto claims to universality that I have positioned myself against throughout this work—but in a sensibility that morphed according to the specific materials and conditions of each research project. Out of the infinity of possible connections, I attached to particular forms, terms, and ideas that put practice and research into generative relations with each other.

The diagrammatic language I am contributing with this research can therefore be characterised in general terms as framed by an archaeological approach—that is, suited to excavating the material geometries of algorithmic prediction in order to critically examine them. As I have evidenced, this general disposition can take a wide variety of forms in practice as it stretches and morphs according to the specific materials and conditions of each project.

Turning to my secondary research questions details the constraints and possibilities that come with this diagrammatic language.

Subquestion RQ1: Through which mechanisms does thinking/writing/designing in diagrammatic terms inform research and practice focused on algorithmic prediction?

In this thesis, I have discussed my research in diagrammatic terms while diagrams were also my primary material for practice. This provided a key constraint as I drew from a wide range of literature—e.g. media archaeology, science studies, critical algorithm studies, divination studies—from an "outside" position in design. The language of diagrams provided a lens to focus my research and contribution, anchoring both in a critical practice that "speaks" in spatial and visual terms. This has particular relevance with the subject matter of algorithmic prediction that is itself saturated by diagrammatic forms and operations. The primary mechanism that informs this research is therefore one of *relating* the subject under investigation, the research literature, and my mode of operation through practice.

These relations do not, however, result in a stable form. To the contrary they have been in flux throughout this research, making diagrams an elastic rather than solid scaffold. While diagrams focused my research, and eventually stabilised enough to be written about, the key mechanism of *oscillation* was present throughout. I have referred to one type of oscillation between control and openness, but this is generalisable as a feature of this language. To be useful in this approach, diagrams have to be considered as fluctuating forms that can stretch to contain everything and its opposite. This makes them, in my view, a particularly fertile exchange surface between the intuitive nature of creative practice and the rigour of research.

These oscillations were channelled through the *movements* afforded by each of the diagrammatic forms I have discussed. These came out of relations between research and practice, and were key to turning tentative explorations into diagrammatic practices that could be articulated and developed with a sense of direction, both within each project and as part of the thesis as a whole. The *case-board* was used to approach the subject of algorithmic prediction and its history, it used the structure of a graph-database to follow the threads of topoï, "starting in the middle" and burrowing out. The production of *traces* was in more direct contact with an algorithmic system, using automated probes to move within its entrapping logics. Finally, *chicanes* moved through the tricks that mediate between data and predictions, combining ephemeris data with astrological algorithms.

While they "move" in different ways, each of these diagrams was produced with a similarly thorough practical involvement with computational tools. I took code libraries, techniques, datasets, coding environments, and so on, used in data science and algorithmic prediction and put them in conversation with critical perspectives through the language of diagrams. With this in mind, my contribution to knowledge is a diagrammatic language that informs every aspect of my research, from manipulating and visualising data with the tools of data science, to writing critically about data practice and operations. Each of my proposed diagram types is effectively a mode of knowledge production. With this I turn to summarising the way I have addressed my second sub-question.

Subquestion RQ2: As algorithmic systems claim to produce objective knowledge, how can diagrams be used as instruments for speculative/conjectural knowledge production?

With each diagrammatic form, I have demonstrated a key feature of my diagrammatic language: it provides alternatives to algorithmic inference as a mode of knowledge production. In contrast to induction or deduction, I have described an abductive process where each project departed from an hunch or intuition that was refined and transformed through the research.

I described each of my diagrammatic form as an instrument for conjectural knowledge production. The *case-board* is a site for yarnwork, the performance of knowledge production enacted by detectives in TV series.

It is a tool for piecing together an investigation, an arc that usually resolves in definitive answers. In my case, the work remains suspended in the in-between state where the pieces do not fully "add up" (Mackay, 2017). *Traces* are the fine-grained data used by recommender systems as they predict "watch time" and entrap users (Seaver, 2018). Betti Marenko and I considered them as sources for conjectural knowledge, a way to imagine algorithmic systems as hunters that 'reconstruct the appearance of an animal they have never seen' (Ginzburg, 1980, 22). Finally, *chicanes* are the tricks and turns involved in the production of divinatory knowledge (Ramey, 2016). While these are not necessarily deceitful, their "sincerity" is put in question by claims that data are an un-mediated source of knowledge about the future. Paying attention to chicanes means acknowledging mediation, and interrogating its politics.

As speculative devices, these diagrams do not "arrive" at definitive conclusions. Instead they facilitate ongoing and open ended activities, modes of questioning algorithmic prediction and of challenging the very ways in which it produces knowledge.

The diagrammatic language I am contributing with this research is fully "compatible" with the computational logics of algorithmic prediction. However it is geared towards an abductive mode of knowledge production that emphasises speculation and conjecture.

Contribution to knowledge

In this thesis I have shown how diagrams can be used as a language for critical research and practice into algorithmic prediction. This language is conducive to an archaeological approach that is immersed in the materials of algorithmic prediction, while critical of its dominant imaginaries and modes of knowledge production. With this I aim to demonstrate to researchers and practitioners with an interest in algorithmic systems, prediction and/or speculation that diagrams are a generative language to engage with these themes through practice.

Through this work I have also made a number of secondary contributions of a more practical nature, namely: a novel use for graph databases; a method for animating and broadcasting data visualisation over a peer-to-peer network; and a method for automated production and publishing of almanac publications.

Future work

The diagrammatic language outlined in this thesis is open-ended. I have located and demonstrated three specific applications of it, but it lends itself to being re-purposed and adapted. While it cannot be readily applied to other cases and settings, it provides a foundation to continue investigating algorithmic systems with a particular attention to their diagrammatic qualities.

Diagrams have proven a fertile ground for cross-disciplinary work, rooted in design but looking to media archaeology, the digital humanities, media studies, science studies, philosophy, and other fields. I plan to use the language established here to continue engaging with these fields, and to collaborate with other researchers and practitioners.

I plan to continue excavating what is obfuscated by the positivist imaginary of algorithmic prediction, namely producing or re-producing diagrams of power relations, and a race to either foreclose or capitalise on any form of uncertainty. With the approach established through this research, I plan to focus future work more specifically on the politics and economics of algorithmic diagrams, for example through more detailed studies of financial trading or political campaigning.

This research sets the stage for, but does not develop, more focused explorations of how power and violence are enacted and reified through algorithmic diagrams. As predictive operations rely on measures and/or classifications of difference in vector space, they are mobilised to entrench differences along lines of race, class, and gender—impacting very real communities and bodies in actual space. After travelling to the level of abstraction of algorithmic systems, I look forward to engaging with the ramifications of these systems in more local and tangible settings.

Towards these ends, I plan to build on my publishing practice; to broaden its reach and to open it to others, specifically in light of recent developments in open-source federated social media—a particularly promising type of diagram. I also plan to continue exploring the generative potential of algorithmic systems, for example with a project exploring the possibility for an "A.I that searches for alternatives to capitalism" that I began articulating as part of this research before realising it would completely explode the scope of this thesis.

Imaginaries of data and prediction continue to take hold, notably through the promises of "artificial intelligence" that businesses (Agrawal et al.,

2018) and governments (e.g. Villani et al., 2018) relentlessly reinforce. The "paradigm shift" of turning every problem into a prediction problem (Agrawal et al., 2018) is re-shaping the world, and challenges what counts as valid knowledge. This practice-based PhD thesis goes some way towards contributing to these debates and proposes some strategies for reclaiming the diagrams of algorithmic prediction.

Appendix A

Note on repositories

The practice work for this research is submitted in the form of three main code repositories. Two of these, *Diagrams of the Future* and *The Monistic Almanac* are also available as websites.

Submitting digital practice work raises a number of issues relating to maintenance and future availability. The repositories are hosted online, using a third party service called Gitlab[1] as it is easily accessible and free of charge. There is no guarantee that this service will continue to offer similar terms, or even to exist, in the future.

I am committing to keeping these repositories online for reference for a period of 10 years from the submission date. I will endeavour to keep them on Gitlab in order for the links provided in this thesis to keep working for this period. Should this not be possible, I will find a new location for the repositories and provide directions as to where they can be found.

I plan to continue developing and maintaining the projects presented here for years to come. In order to clearly mark the state of the repositories at the time of submission, I make use of the "release tag" functionality of the GIT version control system, using the tag `PhD Submission`. This way, the version of the work at the time of submission will be available even if changes are made to the code at a later date.

[1] https://gitlab.com/

Appendix B

Thesis repository

This repository is available at https://gitlab.com/davidbenque/thesis

```
|-- 00_Intro
|   |-- abstract.tex
|   |-- acknowledgements.tex
|   |-- declaration.tex
|   |-- introduction.assets
|   |   |-- box_abduction.kra
|   |   |-- box_abduction.png
|   |   |-- box_abduction.xcf
|   |   `-- Screen Shot 2018-05-23 at 10.38.31.png
|   `-- intro.tex
|-- 01_Context_Review
|   |-- images
|   |   |-- citekey_outline.txt
|   |   |-- citekey_parser.py
|   |   |-- data-science-process.mmd
|   |   |-- data_science_process.xml
|   |   |-- hayashi.odg
|   |   |-- iris_sample.csv
|   |   |-- iris_vectors.py
|   |   |-- iris_vectors.svg
|   |   `-- TOUSE DOC-20190422-094924 FB Patent.jpg
|   |-- main.assets
|   |   |-- 02 4 arcs.jpg
|   |   |-- 03 wheel.jpg
|   |   |-- 04 alternative unknowns.jpg
|   |   |-- 1548187496686.png
|   |   |-- 1548337505501.png
|   |   |-- 1548338484701.png
|   |   |-- 1548932994039.png
|   |   |-- 1548942335364.png
```

```
|   |   |-- 1548949404203.png
|   |   |-- Auger - 2012 - Why robot speculative design, the domestication -cro
|   |   |-- Auger_2014_Why heart attacks could be a thing of the past-cropped.p
|   |   |-- bernhardsson_KNN.png
|   |   |-- Bertin.jpg
|   |   |-- candy-futurescone.png
|   |   |-- COMPAS-matrix.png
|   |   |-- data_science_process.pdf
|   |   |-- data-science-process.svg
|   |   |-- DnR_impossible_object.png
|   |   |-- EF-4arcs.pdf
|   |   |-- EF-alts.pdf
|   |   |-- EF-wheel.pdf
|   |   |-- face-cage-2-elle-mehrmand_portrait.jpg
|   |   |-- image_9.png
|   |   |-- iris-matrix-cropped.pdf
|   |   |-- iris_vectors-1548327783767.svg
|   |   |-- iris_vectors.svg
|   |   |-- Lavrenko-over-under-fit2019-08-17 11-50-31.png
|   |   |-- Lull-cramer-cropped.pdf
|   |   |-- Maaten_Hinton_2008_Visualizing data using t-SNE-cropped.pdf
|   |   |-- neuralnetworks.png
|   |   |-- parrish-semantic_space.png
|   |   |-- previeux-atelierbac03.jpg
|   |   |-- previeux-atelierbac06.jpg
|   |   |-- previeux-recherche05.jpg
|   |   |-- Project_map.svg
|   |   |-- PythonVisLandscape-768x561.jpg
|   |   |-- PythonVis-redrawn.drawio
|   |   |-- PythonVis-redrawn.pdf
|   |   |-- schematicpatent.pdf
|   |   |-- scikit_iris_nearest_neighbour_002.png
|   |   |-- Screenshot_2019-09-06 Predictive Art Bot - Disnovation org.png
|   |   |-- sphx_glr_plot_classifier_comparison_001.png
|   |   |-- unspecifiedimpossibleart.jpeg
|   |   |-- unspredicschemaapecified.jpeg
|   |   |-- vectorialist_class.pdf
|   |   |-- vectorialist_class.svg
|   |   |-- vectorialist_diagram.jpg
|   |   |-- white-collar-crime.png
|   |   |-- zairja.png
|   |   `-- zoo.png
|   |-- main.tex
|   |-- Practice_Map.pdf
|   `-- practice_preamble.tex
|-- 02_Case-Board
|   |-- main.assets
|   |   |-- 1564423218081.png
|   |   |-- 1564423307909.png
```

```
|    |    |-- 1564423410456.png
|    |    |-- 1564424028869.png
|    |    |-- 1901-mackenzie.png
|    |    |-- 2004_World Government.pdf
|    |    |-- 20180216-neo4j structure.png
|    |    |-- CTF-editor_social-physics-topos.pdf
|    |    |-- DOTF_evolution.pdf
|    |    |-- DOTF_evolution.svg
|    |    |-- drucker_speclab.jpg
|    |    |-- galton_in_friendly_111.pdf
|    |    |-- graph_ex.pdf
|    |    |-- graph_ex.svg
|    |    |-- Pasquinelli_The Evolution of Machinic Intelligence (map).pdf
|    |    |-- scan_CTF_editor.jpg
|    |    |-- schema-arrows-cropped.pdf
|    |    |-- Screenshot from 2019-07-29 18-51-11.png
|    |    `-- TODO_CTF_evolution.png
|    |-- main_old.assets
|    |    |-- 1555943003032.png
|    |    |-- 1556022872298.png
|    |    |-- 1556115594748.png
|    |    |-- 1556280862298.png
|    |    |-- 1901-mackenzie.png
|    |    |-- 2019-04-26 14.34.33 CTF Sketch.jpg
|    |    |-- all.png
|    |    |-- CTF_editor_Galton thread_better_2.odg
|    |    |-- CTF_editor_Galton thread_better_2.pdf
|    |    |-- CTF_editor_Galton thread_better_3.pdf
|    |    |-- CTF_editor_Galton thread_better.pdf
|    |    |-- CTF_editor_Galton thread_better.svg
|    |    |-- CTF_editor_Galton thread.pdf
|    |    |-- DOTF_timeline.odg
|    |    |-- DOTF_timeline.svg
|    |    |-- drucker_speclab.jpg
|    |    |-- example_neo4j.png
|    |    |-- graph example.odg
|    |    |-- graph_ex.svg
|    |    |-- histograph.svg
|    |    |-- HOP Timeline00004.png
|    |    |-- Nowviskie_0.jpg
|    |    |-- Nowviskie.jpg
|    |    |-- Pasquinelli_evolution.svg
|    |    |-- scan_CTF_editor.jpg
|    |    |-- scan_graph_rels_sketch.jpg
|    |    |-- scan_graph_strategies.jpg
|    |    |-- Screen Shot 2018-01-20 at 12.40.27.png
|    |    |-- Screen Shot 2018-01-20 at 12.41.37.png
|    |    |-- TODO_CTF_evolution.png
|    |    `-- youtube-map.jpg
```

```
|   `-- main.tex
|-- 03_Traces
|   |-- main.assets
|   |   |-- 1562664586056.png
|   |   |-- A1-arcchoice-poster-cropped.pdf
|   |   |-- all.png
|   |   |-- Digger.mmd2.svg
|   |   |-- funnel_a_youtube.pdf
|   |   |-- funnel_a_youtube.svg
|   |   |-- funnel_b_trap.pdf
|   |   |-- funnel_b_trap.svg
|   |   |-- funnel_trap.svg
|   |   |-- Gephi 50x37cm-cropped.pdf
|   |   |-- Matplotlib 50x25cm-cropped.pdf
|   |   |-- Matplotlib A4-cropped.pdf
|   |   |-- mermaid_out04-cropped.pdf
|   |   |-- mermaid.png
|   |   |-- preview_comp.png
|   |   |-- probe_digger.pdf
|   |   |-- probe_print_4.png
|   |   |-- probe_ripple.pdf
|   |   |-- Ripple.mmd.svg
|   |   |-- Screenshot from 2019-07-30 18-26-11.png
|   |   |-- Screenshot from 2019-07-30 18-26-13.png
|   |   |-- Screenshot from 2019-07-30 18-26-15.png
|   |   |-- Simple_digger.mmd.svg
|   |   |-- simple_simpledigger.pdf
|   |   |-- trace_print_2019-01-21-12-11-06.pdf
|   |   |-- trace-print.png
|   |   |-- trace-web.png
|   |   `-- youtube-map.jpg
|   `-- main.tex
|-- 04_Chicanes
|   |-- main.assets
|   |   |-- 01 Annuaire 1875.jpg
|   |   |-- 02 Thomas_1976_The Old Farmer's Almanac.jpg
|   |   |-- 2018-09-17 15.40.14-1.jpg
|   |   |-- 500px-Mauna_Loa_CO2_monthly_mean_concentration.svg.png
|   |   |-- 5 Image10.png
|   |   |-- 6 Image11.png
|   |   |-- 7 Image12.png
|   |   |-- eadc2813f64499af66d6da0afb177176.png
|   |   |-- IMG_4212.jpg
|   |   |-- mauna-loa-cropped.pdf
|   |   |-- old-moore-astro-cropped.pdf
|   |   |-- old-moores-website-14-04-2018.png
|   |   |-- Screenshot_2019-07-17 CROP.png
|   |   |-- Screenshot from 2019-04-24 15-43-52.png
|   |   |-- TMA-CCC.png
```

```
|    |    |-- TMA-CPI.png
|    |    |-- TMA-CP.png
|    |    |-- TMA-EA.png
|    |    `-- TMA-PGH.png
|    `-- main.tex
|-- 05_Conclusion
|    `-- main.tex
|-- 06_Appendix
|    |-- Almanac
|    |    |-- 20191028195713 Electional Astrology for John Doe.pdf
|    |    |-- 20191028200018 Electional Astrology for Jane Doe.pdf
|    |    |-- 20191028200239 Electional Astrology for Jean Do.pdf
|    |    |-- 20191028200556 Electional Astrology for Jane Doe.pdf
|    |    `-- Example-notebook-CCC.pdf
|    |-- Almanac.tex
|    |-- arc-choice
|    |    |-- screencast_trace.mp4
|    |    |-- Screenshot from 2019-12-09 21-35-55.png
|    |    `-- WhatsApp Image 2019-03-08 at 11.45.16.jpeg
|    |-- Arc-choice.tex
|    |-- DOTF
|    |    |-- DB-poster.pdf
|    |    |-- graph-redesign.png
|    |    |-- HOP Timeline00004.png
|    |    |-- HOP Timeline00005.png
|    |    |-- Screen_Shot_2016-08-22_15-55-05.png
|    |    |-- Screen_Shot_2017-03-15_10-26-51.png
|    |    `-- Screenshot from 2019-11-18 21-19-39.png
|    |-- DOTF.tex
|    |-- list_of_conferences.tex
|    |-- Note-on-repos.tex
|    |-- poss-space
|    |    |-- 01-title-card.png
|    |    |-- 02-futures-cone.png
|    |    |-- 03-clusters.png
|    |    |-- 04-decision-surface.png
|    |    |-- 05-haruspex.png
|    |    |-- 06-probability.png
|    |    |-- 07-tea-cup.png
|    |    |-- 08-probability-distribution.png
|    |    |-- Possibility Space Draft1.pdf
|    |    `-- Poss_Space_LCC_FULL.mp4
|    |-- Poss-space.tex
|    |-- Presage
|    |    |-- 20170428 ACC.jpg
|    |    |-- 20170428 BOL.jpg
|    |    |-- 20170428 RIC.jpg
|    |    `-- 20170428 WAD.jpg
|    |-- presage.tex
```

```
|    |-- Thesis_repo.tex
|    `-- tree.txt
|-- bibliography.bib
|-- DB_thesis_style.bst
|-- DB_thesis_style.dbj
|-- main_abstract.pdf
|-- main_abstract.tex
|-- main.pdf
|-- main.tex
|-- notes.md
|-- Pipfile
|-- Pipfile.lock
|-- plainnat_DB.bst
|-- preamble.tex
|-- Readme.md
|-- tree.txt
|-- wordcount.sh
`-- wordcount.tex

19 directories, 225 files
```

Appendix C

Diagrams of the Future

The practice submission consists of a web application, visible at http://dotf.xyz, and a code repository https://gitlab.com/davidbenque/diagrams-of-the-future

please see supplementary material in the following pages.

From timeline to case board

Version history of *Diagrams of the Future.*

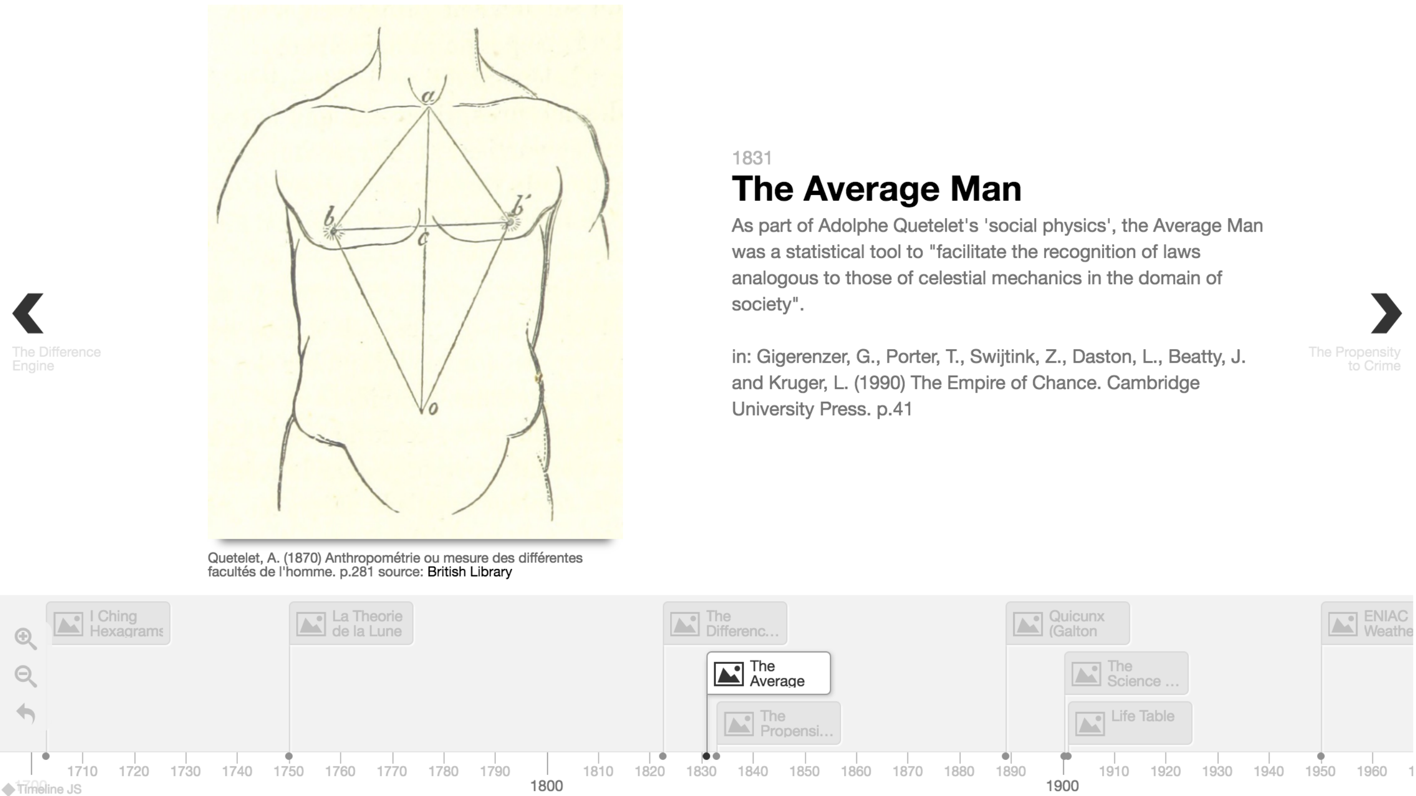


Figure C.1: Timeline Prototype, Spring 2016. First exploration using timeline.js https://timeline.knightlab.com (Accessed: 18/11/2019).

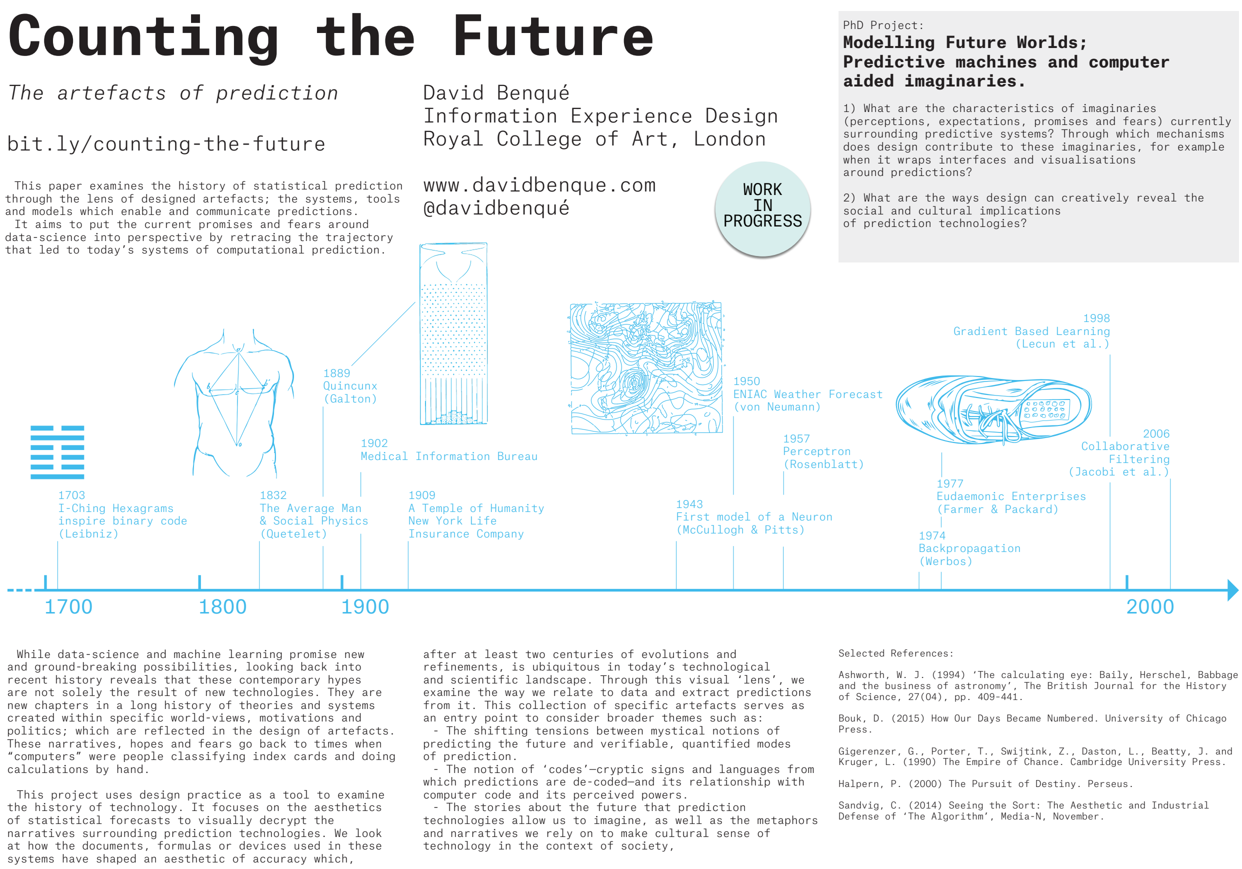


Figure C.2: Poster presented at the Microsoft Research PhD Summer School, July 2016 in Cambridge, UK.

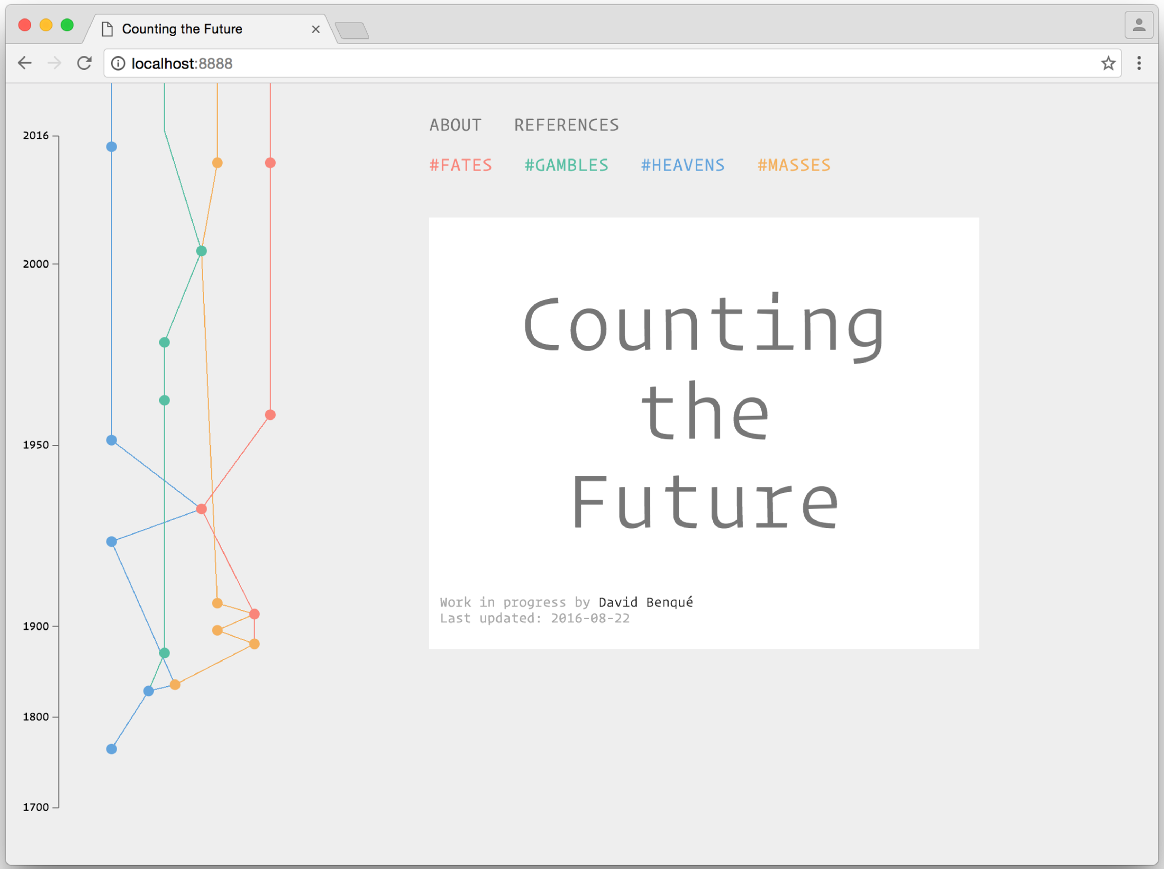


Figure C.3: *Counting the Future* Minimum Viable Product, presented at 4S/EASST in Barcelona (August, 2016) and at *Communicating the Intangible*, work-in-progress show at the Royal College of Art (September, 2016), see Benqué (2016).

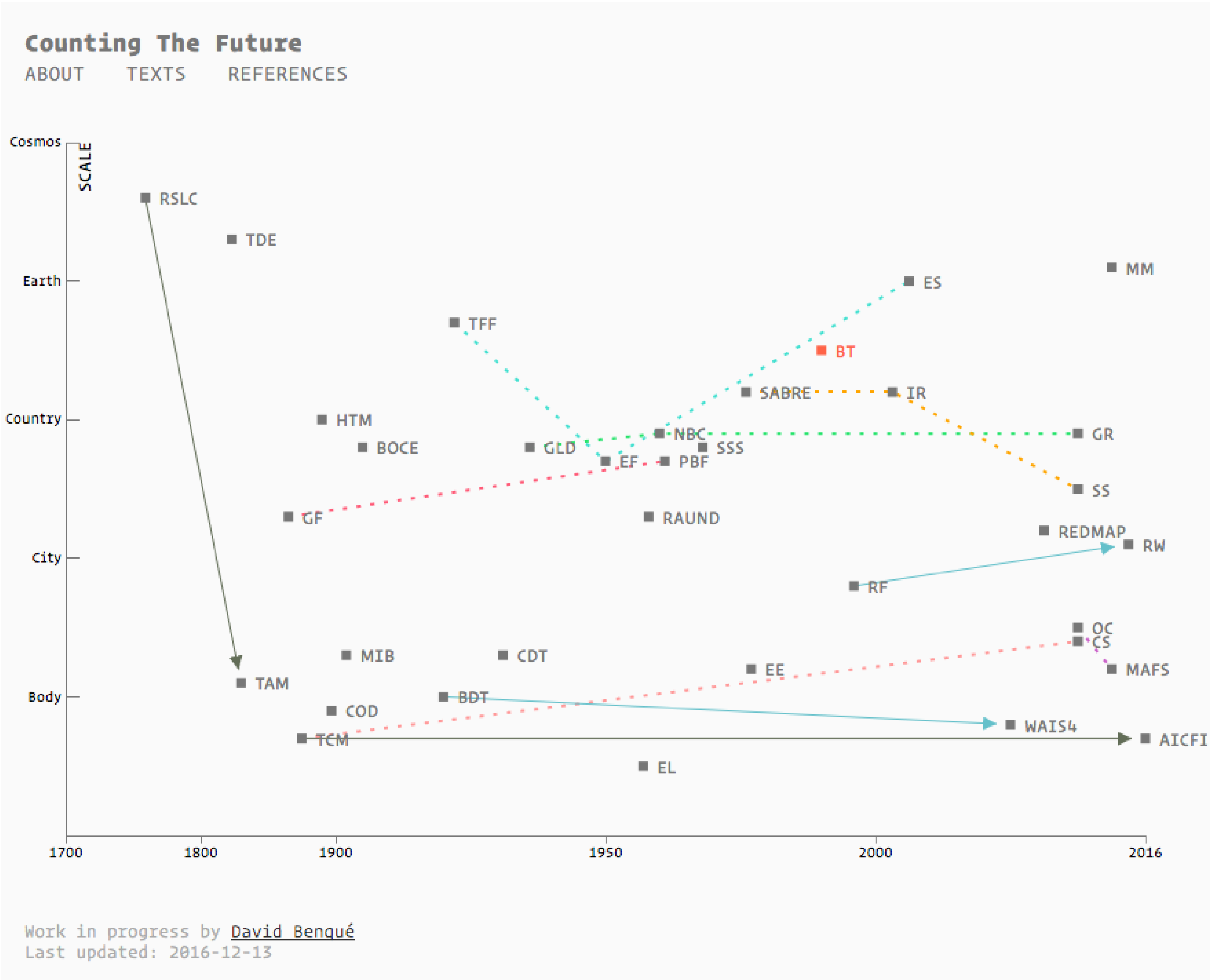


Figure C.4: *Counting the Future* redesigned after feedback from 4S/EASST and *Communicating the Intangible*, winter 2016-17.

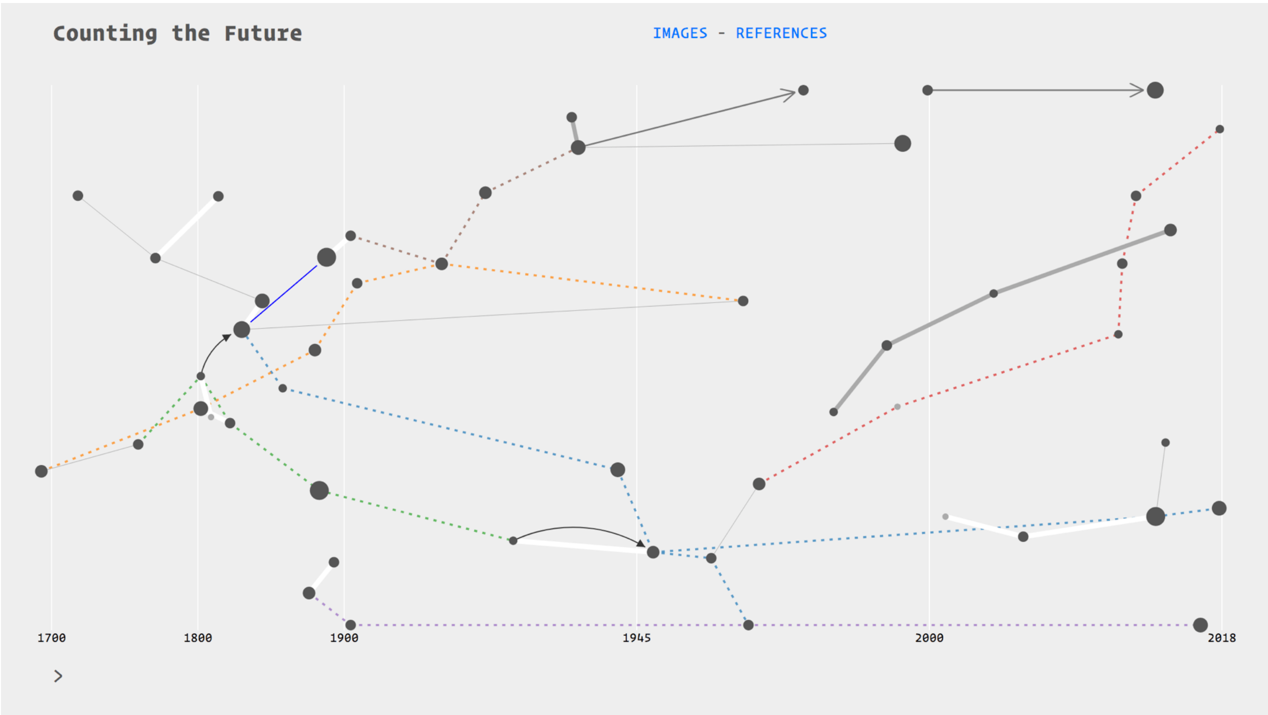


Figure C.5: *Counting the Future* re-built and redesigned with a graph database, February 2018.

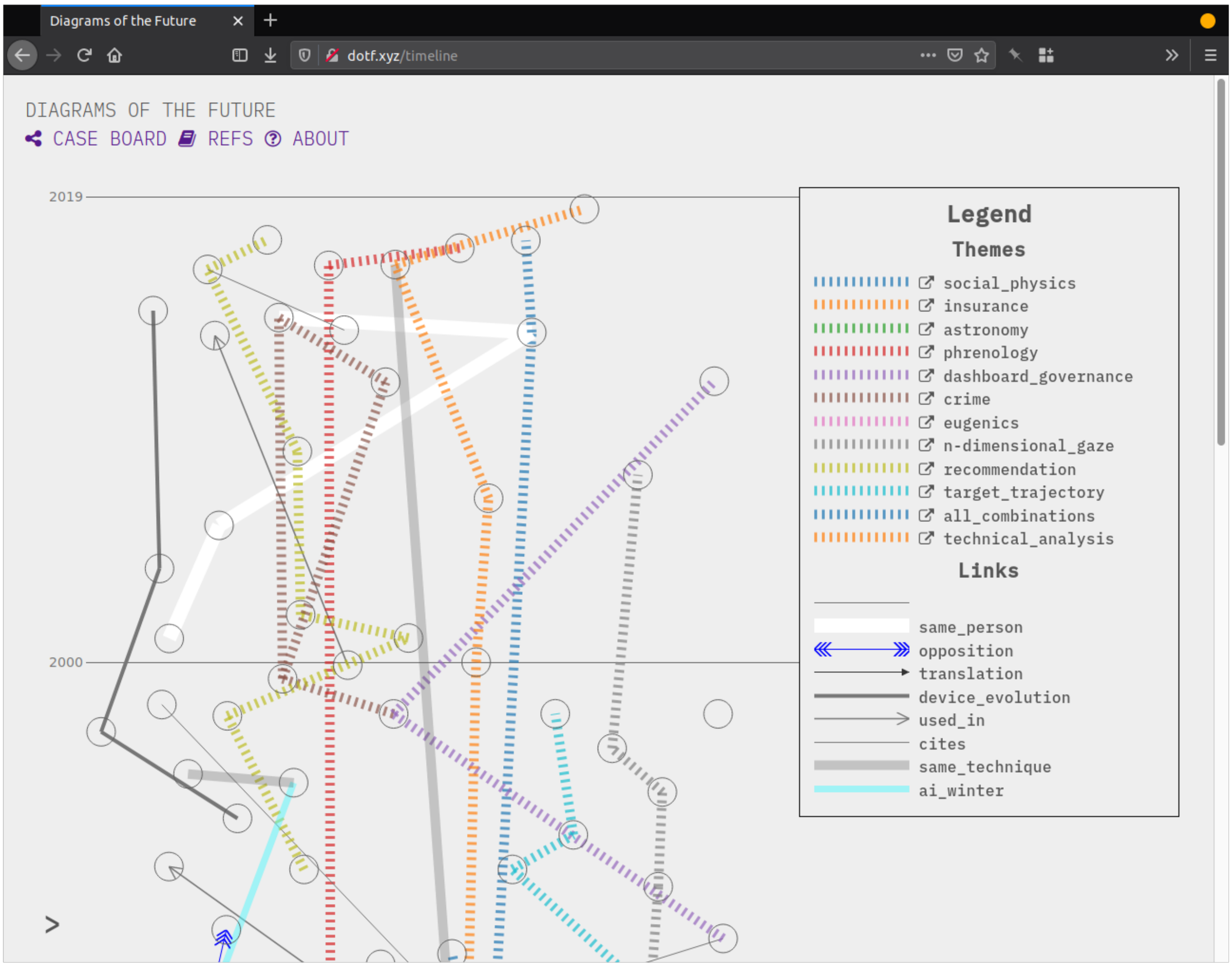


Figure C.6: *Diagrams of the Future* refactored with a vertical orientation and change of title, Summer 2019.

Appendix D

Architectures of Choice Vol.1: YouTube

The practice submission consists of a code repository
https://gitlab.com/davidbenque/arc-choice

please see supplementary material in the following pages.

Trace animation

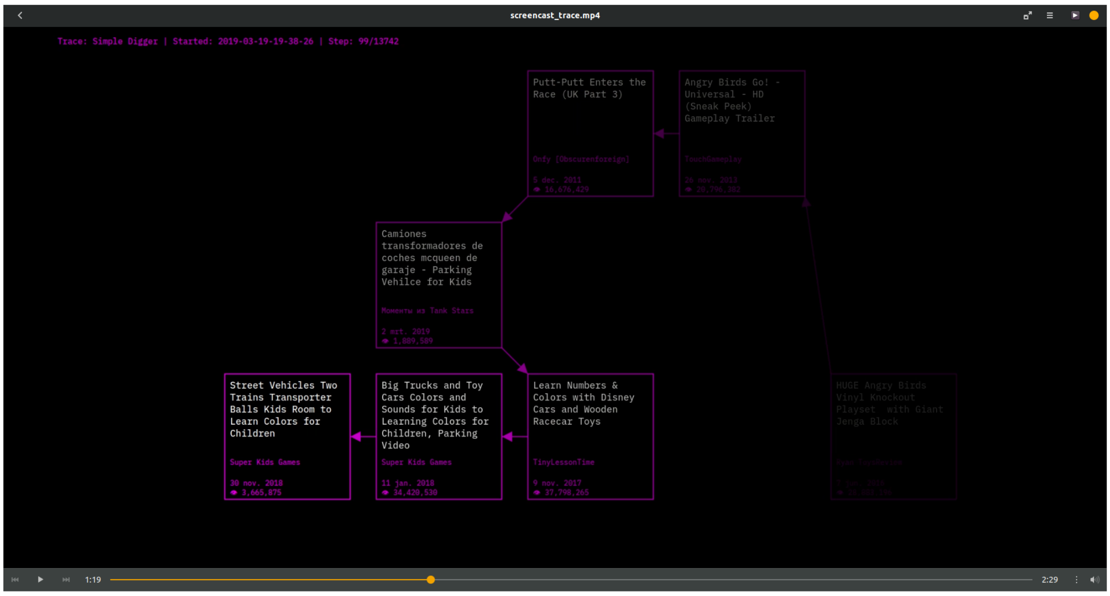


Figure D.1: Trace animation preview, as displayed in the web browser but rendered as video. Available to view online at: https://gitlab.com/davidbenque/thesis/raw/master/06_Appendix/arc-choice/screencast_trace.mp4 (best viewed in Firefox or Chromium browser).

Instructions for viewing the animated traces

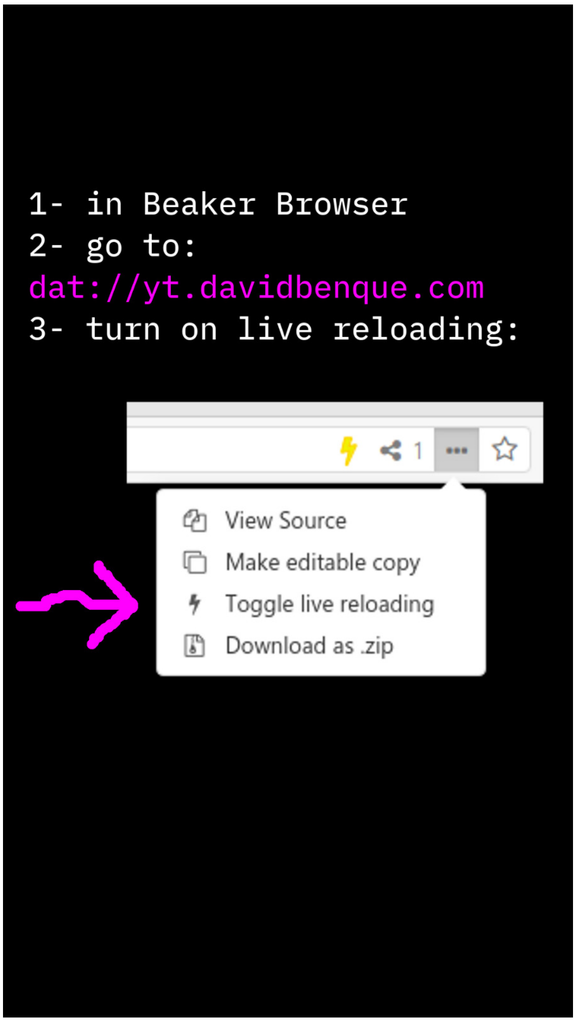


Figure D.2: Instructions for viewing the stream of traces animations in the *Beaker Browser* (Frazee et al.). Shared through *Instragram Stories* during the *Research Through Design* conference 2019.

Appendix E

The Monistic Almanac

The practice submission consists of a web publication, visible at https://almanac.computer, and a code repository https://gitlab.com/davidbenque/almanac.computer

please see supplementary material in the following pages.

Electional Astrology Calendars

The following four pages are example outputs from the personalised electional astrology calendars program used for *Supra Systems Office Rites* at the Victoria & Albert Museum Digital Design Weekend, 22-23 Oct. 2018.

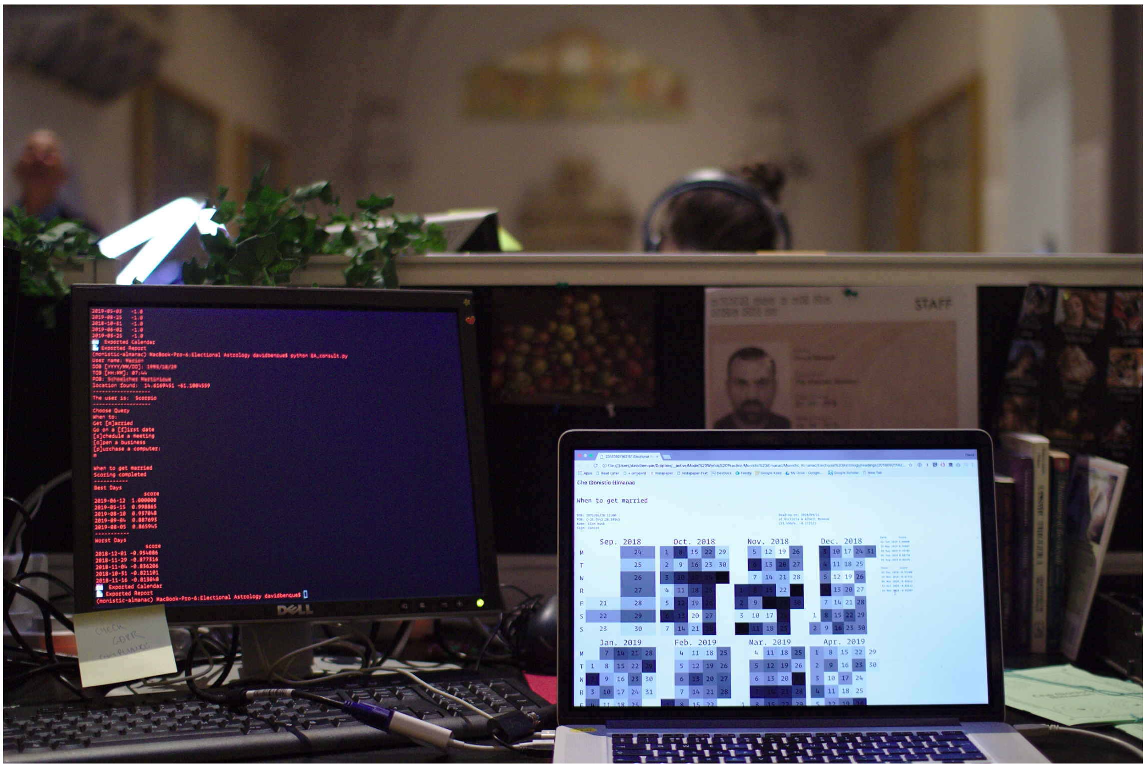


Figure E.1: Electional Astrology desk at *Supra Systems Office Rites*, Victoria & Albert Museum Digital Design Weekend, 22-23 Oct. 2018.

The Monistic Almanac

When to open a buisness

Name: John Doe
DOB: 1979/08/27 06:35
POB: Dublin (53.3497645,-6.2602732)
Sign: Virgo

Reading on: 2019/10/28
at Victoria & Albert Museum
(51.49674,-0.17252)

Oct. 2019

M	28
T	29
W	30
R	31
F	
S	
S	

Nov. 2019

M		4	11	18	25
T		5	12	19	26
W		6	13	20	27
R		7	14	21	28
F	1	8	15	22	29
S	2	9	16	23	30
S	3	10	17	24	

Dec. 2019

M		2	9	16	23	30
T		3	10	17	24	31
W		4	11	18	25	
R		5	12	19	26	
F		6	13	20	27	
S		7	14	21	28	
S	1	8	15	22	29	

Jan. 2020

M		6	13	20	27
T		7	14	21	28
W	1	8	15	22	29
R	2	9	16	23	30
F	3	10	17	24	31
S	4	11	18	25	
S	5	12	19	26	

Feb. 2020

M		3	10	17	24
T		4	11	18	25
W		5	12	19	26
R		6	13	20	27
F		7	14	21	28
S	1	8	15	22	29
S	2	9	16	23	

Mar. 2020

M		2	9	16	23	30
T		3	10	17	24	31
W		4	11	18	25	
R		5	12	19	26	
F		6	13	20	27	
S		7	14	21	28	
S	1	8	15	22	29	

Apr. 2020

M		6	13	20	27
T		7	14	21	28
W	1	8	15	22	29
R	2	9	16	23	30
F	3	10	17	24	
S	4	11	18	25	
S	5	12	19	26	

May. 2020

M		4	11	18	25
T		5	12	19	26
W		6	13	20	27
R		7	14	21	28
F	1	8	15	22	29
S	2	9	16	23	30
S	3	10	17	24	31

Jun. 2020

M	1	8	15	22	29
T	2	9	16	23	30
W	3	10	17	24	
R	4	11	18	25	
F	5	12	19	26	
S	6	13	20	27	
S	7	14	21	28	

Jul. 2020

M		6	13	20	27
T		7	14	21	28
W	1	8	15	22	29
R	2	9	16	23	30
F	3	10	17	24	31
S	4	11	18	25	
S	5	12	19	26	

Aug. 2020

M		3	10	17	24	31
T		4	11	18	25	
W		5	12	19	26	
R		6	13	20	27	
F		7	14	21	28	
S	1	8	15	22	29	
S	2	9	16	23	30	

Sep. 2020

M		7	14	21	28
T	1	8	15	22	29
W	2	9	16	23	30
R	3	10	17	24	
F	4	11	18	25	
S	5	12	19	26	
S	6	13	20	27	

Oct. 2020

M		5	12	19	26
T		6	13	20	27
W		7	14	21	
R	1	8	15	22	
F	2	9	16	23	
S	3	10	17	24	
S	4	11	18	25	

Best Days

Date	Score
23 Apr 2020	1.00000
26 Nov 2019	0.99617
24 Mar 2020	0.97114
20 Jul 2020	0.87045
24 Jan 2020	0.84274

Worst Days

Date	Score
20 Jun 2020	-1.00000
19 Jun 2020	-1.00000
22 Jun 2020	-1.00000
23 Jun 2020	-1.00000
25 Jun 2020	-0.74155

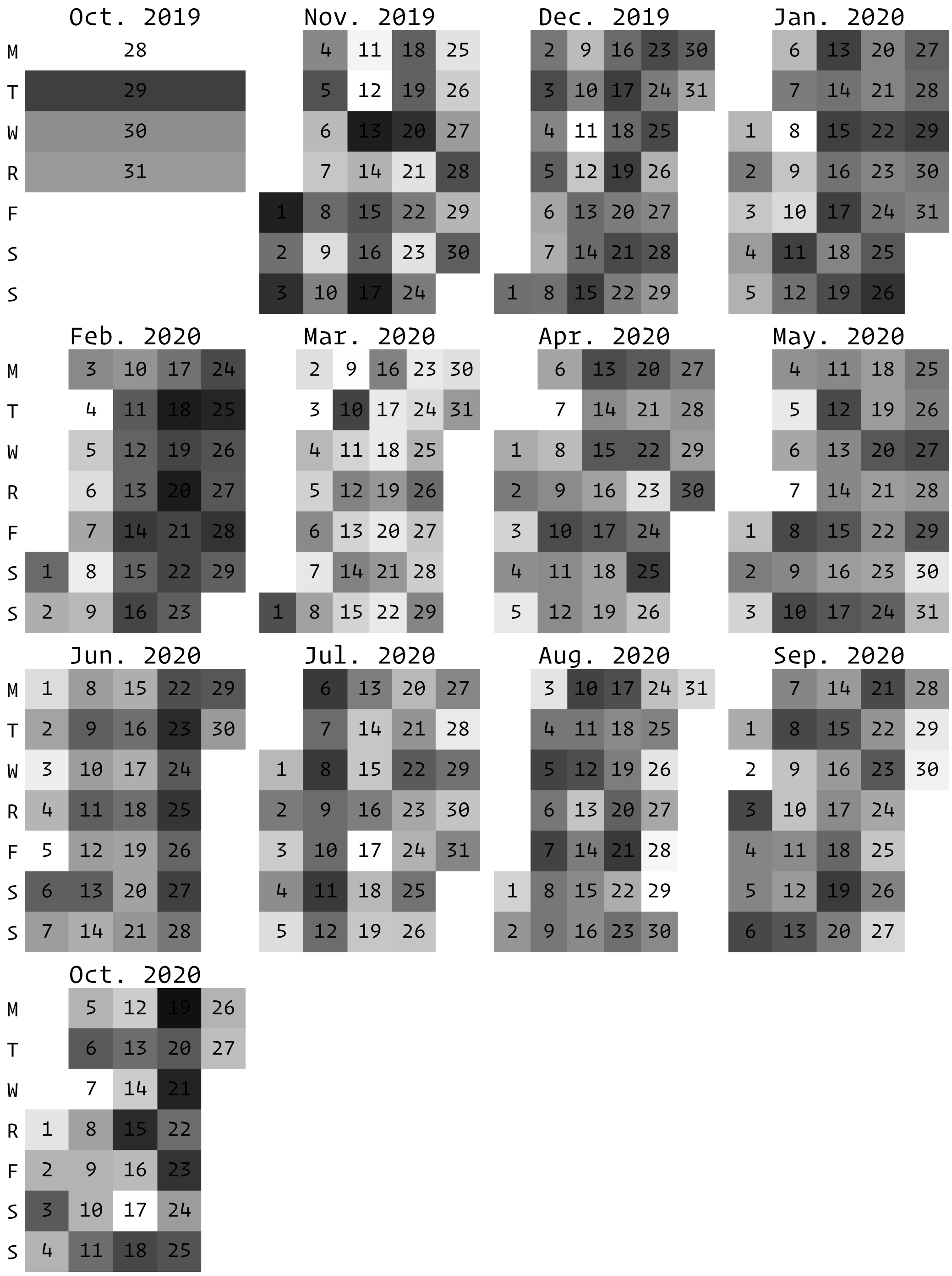
https://ALMANAC.COMPUTER

The Monistic Almanac

When to schedule a meeting

Name: Jane Doe
DOB: 1982/01/31 23:45
POB: Paris (48.8566101,2.3514992)
Sign: Aquarius

Reading on: 2019/10/28
at Victoria & Albert Museum
(51.49674,-0.17252)



Best Days

Date	Score
04 Feb 2020	0.85483
05 Jun 2020	0.77811
11 Dec 2019	0.76458
08 Feb 2020	0.75382
08 Jan 2020	0.69474

Worst Days

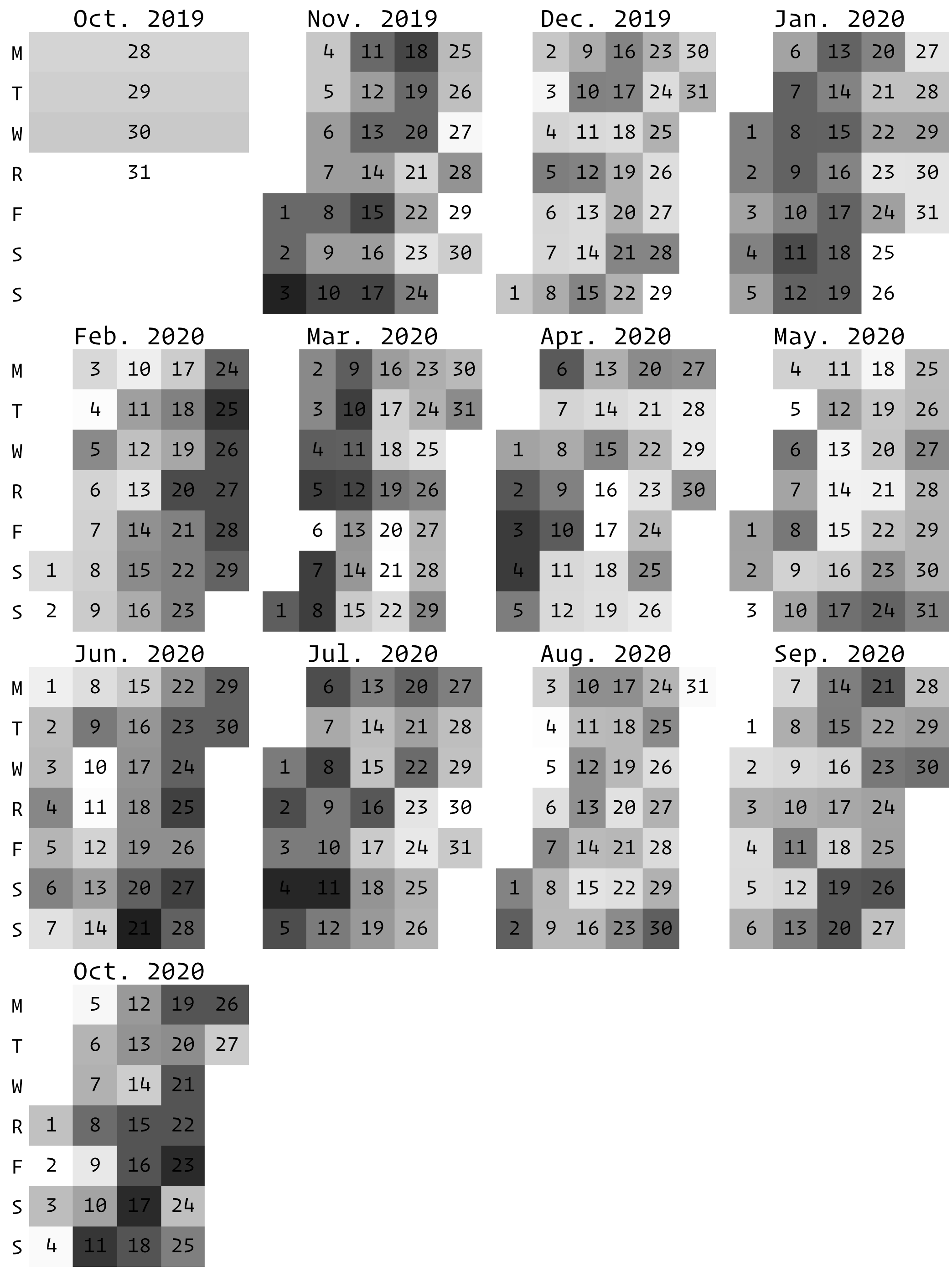
Date	Score
17 Oct 2020	-0.81546
19 Oct 2020	-0.71951
13 Nov 2019	-0.66863
17 Nov 2019	-0.66863
01 Nov 2019	-0.65427

https://almanac.computer

When to purchase a computer

Name: Jean Do
DOB: 2016/05/26 22:34
POB: London (51.5073219,-0.1276474)
Sign: Gemini

Reading on: 2019/10/28
at Victoria & Albert Museum
(51.49674,-0.17252)



Best Days

Date	Score
25 Jan 2020	0.80000
26 Jan 2020	0.79993
27 Jan 2020	0.64164
23 Jan 2020	0.64140
30 Jan 2020	0.63724

Worst Days

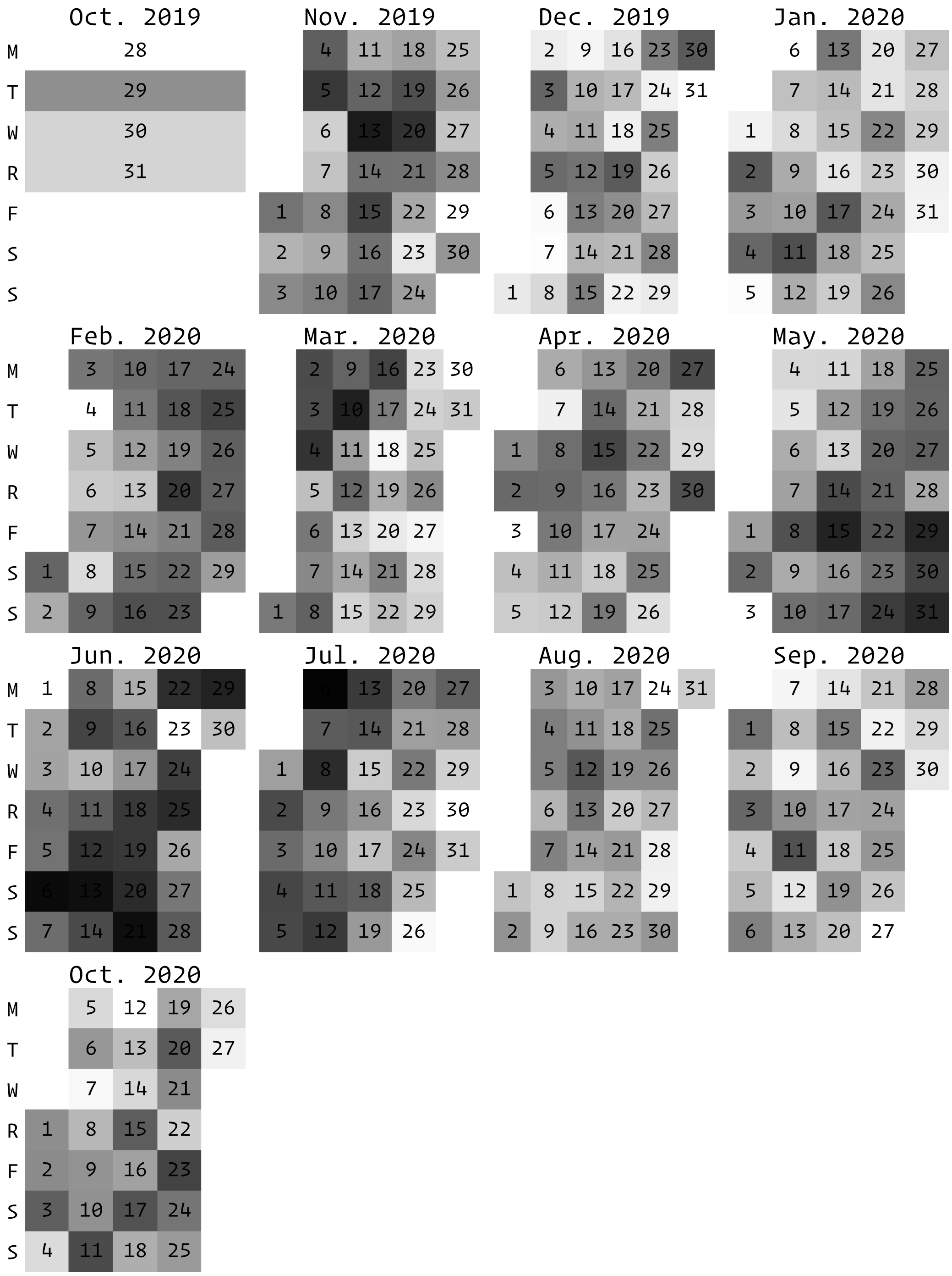
Date	Score
06 Mar 2020	-1.00000
04 Jul 2020	-0.80000
23 Oct 2020	-0.80000
03 Nov 2019	-0.80000
17 Oct 2020	-0.80000

The Monistic Almanac

When to get married

Name: Jane Doe
DOB: 1980/03/12 14:38
POB: Greenwich (52.0367323,1.168934)
Sign: Pisces

Reading on: 2019/10/28
at Victoria & Albert Museum
(51.49674,-0.17252)



Best Days

Date	Score
24 Aug 2020	0.99510
06 Jan 2020	0.99452
05 Jan 2020	0.97709
03 Apr 2020	0.96855
04 Feb 2020	0.96588

Worst Days

Date	Score
23 Jun 2020	-0.96868
06 Jul 2020	-0.92183
06 Jun 2020	-0.89269
21 Jun 2020	-0.86681
13 Jun 2020	-0.86303

https://ALMANAC.COMPUTER

Example Notebook: Cosmic Commodity Charts

The following pages display an example Jupyter notebook (Pérez and Granger, 2007) used to develop the *Cosmic Commodity Chart* section of *The Monistic Almanac*. Please note that while they can be exported to a variety of formats (e.g. PDF as seen here), notebooks are best viewed as web pages in a browser window. Too see this notebook online please visit: `https://nbviewer.jupyter.org/urls/gitlab.com/davidbenque/almanac.computer/raw/master/Cosmic%20Commodity%20Chart/Cosmic%20Commodity%20Charts.ipynb`

Cosmic Commodity Charts

Predicting prices of commodity futures according to the positions of solar system planets.

Stock market data: Quandl (https://www.quandl.com/)
Astronomical data: NASA Jet Propulsion Laboratory de430 ephemeris [PDF] (https://naif.jpl.nasa.gov/pub/naif/generic_kernels/spk/planets/de430_and_de431.pdf) accessed via Skyfield (https://rhodesmill.org/skyfield/)

In [1]:
```
import numpy as np
import pandas as pd
import datetime
import matplotlib.pyplot as plt
%matplotlib inline
```

Chart title and query

In [2]:
```
# all commodity options

commodities = {
    # Grain and foods
    "Corn": "CHRIS/CME_C1",
    "Wheat": "CHRIS/CME_W3",
    "Soybean": "CHRIS/CME_S3",
    "Rough Rice": "CHRIS/CME_RR1", # only goes back to 1986
    "Cotton No.2": "CHRIS/ICE_CT2",
    "Sugar": "CHRIS/ICE_SB2",
    "Orange Juice": "CHRIS/ICE_OJ4",
    "Coffee C": "CHRIS/ICE_KC3",
    # Livestock
    "Live Cattle": "CHRIS/CME_LC3",
    "Feeder Cattle": "CHRIS/CME_FC2",
    "Lean Hog": "CHRIS/CME_LN4",
    # energy
    "Crude Oil": "CHRIS/CME_CL38",
    # Metals
    "Copper": "CHRIS/CME_HG2",
    # Precious
    "Gold": "CHRIS/CME_GC5",
    "Silver": "CHRIS/CME_SI7",
    }
```

In [3]:
```
# planet options

planet_lists = {
    "Full": ['mercury', 'venus', 'earth', 'mars',
            'jupiter', 'saturn', 'uranus', 'neptune', 'pluto'],
    "Inner": ['mercury', 'venus', 'earth', 'mars']
}
```

In [4]:
```
# chart lengths in days

lengths = [365, 730]
```

In [5]:
```
# Build the query

from random import choice

plans = list(planet_lists.keys())
coms = list(commodities.keys())

query = [choice(plans), choice(coms), choice(lengths)]
print("{0} {1} Chart, {2} Days".format(*query))
```

```
Full Soybean Chart, 730 Days
```

Commodity Price

In [6]:
```
import quandl

with open("../_data/quandl-key.txt", "r") as text:
    for line in text:
        quandl_key = line
quandl.ApiConfig.api_key = quandl_key
```

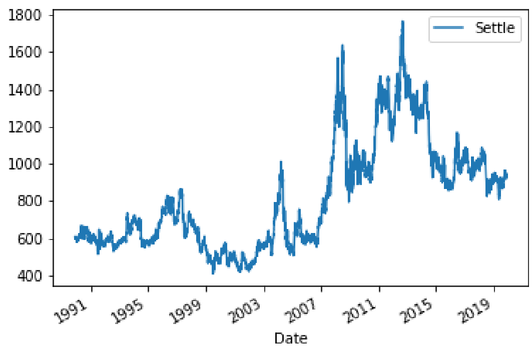
In [7]:
```
today = datetime.date.today()
start_date = today - datetime.timedelta(days = 10950) # 30 years ago

commodities = {
    "Corn": "CHRIS/CME_C1",
    "Wheat": "CHRIS/CME_W3",
    "Soybean": "CHRIS/CME_S3",
    "Rough Rice": "CHRIS/CME_RR1", # only goes back to 1986
    "Cotton No.2": "CHRIS/ICE_CT2",
    "Sugar": "CHRIS/ICE_SB2",
    "Orange Juice": "CHRIS/ICE_OJ4",
    "Coffee C": "CHRIS/ICE_KC3",
    # Livestock
    "Live Cattle": "CHRIS/CME_LC3",
    "Feeder Cattle": "CHRIS/CME_FC2",
    "Lean Hog": "CHRIS/CME_LN4",
    # energy
    "Crude Oil": "CHRIS/CME_CL38",
    # Metals
    "Copper": "CHRIS/CME_HG2",
    "Zinc": "CHRIS/SHFE_ZN3", # only goes back to 2007
    # Precious
    "Gold": "CHRIS/CME_GC5",
    "Silver": "CHRIS/CME_SI7",
    }

price_data_30y = quandl.get(commodities[query[1]], start_date=start_date, end_date=today)
price_data_30y.head()
```

Out[7]:

	Open	High	Low	Last	Change	Settle	Volume	Previous Day Open Interest
Date								
1989-11-30	609.00	610.25	606.75	607.25	NaN	607.25	7850.0	74900.0
1989-12-01	607.75	609.75	601.25	601.75	NaN	601.75	12380.0	75750.0
1989-12-04	600.00	600.00	596.50	597.00	NaN	597.00	10960.0	75310.0
1989-12-05	597.50	600.75	597.00	599.75	NaN	599.75	8640.0	76665.0
1989-12-06	600.00	604.25	597.00	603.50	NaN	603.50	12810.0	79635.0

In [8]:
```
price_data_30y.plot(y="Settle")
```

Out[8]: <matplotlib.axes._subplots.AxesSubplot at 0x7fd2941aaf28>



Planets & Orbits Training Data

In [9]:
```
# load skyfield and ephemeris
from skyfield.api import load, Loader
from astropy import units as u

# Skyfield data
load = Loader('../_data/skyfield')
planets = load('de430.bsp')
ts = load.timescale()
```

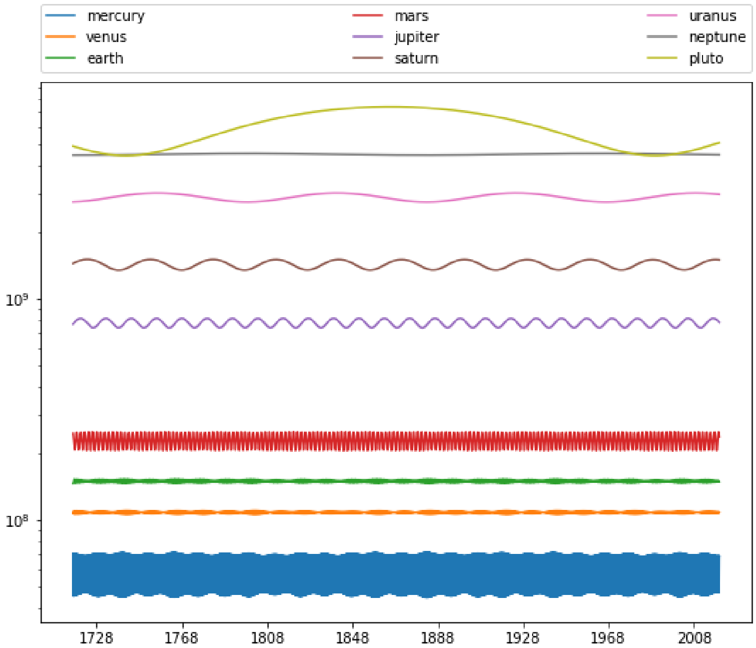
In [10]:
```
planet_lists = {
    "Full": ['mercury', 'venus', 'earth', 'mars',
             'jupiter', 'saturn', 'uranus', 'neptune', 'pluto'],
    "Inner": ['mercury', 'venus', 'earth', 'mars'],
    "Outer": ['jupiter', 'saturn', 'uranus', 'neptune', 'pluto']
}
```

In [11]:
```
planet_list = planet_lists[query[0]]
```

In [12]:
```
def add_planet_positions(df, planet_list):
    '''
    input: dataframe with datetime index, list of planets
    output: dataframe with planet distance (barycentric) columns
            added for the planets in planet_list and dates in df.index
    '''
    for planet in planet_list:
        query_name = planet + " barycenter"
        col_name = planet + "_dist"
        p = planets[query_name]
        df[col_name] = p.at(ts.utc(df.index)).distance().to(u.km)
    return df
```

Scaling

The planet positions are scaled according to their relative orbits, with 0.0 and 1.0 as the closest and furthest points from the solar system barycentre on that planet's orbit.



In [13]:
```
# Mininum and maximum for each orbit
# 300 year window to allow for a full orbit of Pluto
# exported to a CSV file to avoid computing them each time
planets_minmax = pd.read_csv("planets_minmax.csv", index_col = 0)
planets_minmax
```

Out[13]:

	min	max
mercury_dist	4.473931e+07	7.108315e+07
venus_dist	1.062104e+08	1.102090e+08
earth_dist	1.458724e+08	1.533281e+08
mars_dist	2.052792e+08	2.505484e+08
jupiter_dist	7.390777e+08	8.158063e+08
saturn_dist	1.346026e+09	1.508435e+09
uranus_dist	2.735045e+09	3.006735e+09
neptune_dist	4.459924e+09	4.536853e+09
pluto_dist	4.435661e+09	7.375996e+09

In [14]:
```
def orbit_scaler(planet_dist, x):
    min_dist, max_dist = planets_minmax.loc[planet_dist]
    return ((x - min_dist)/(max_dist-min_dist))
```

In [15]:
```
training_data = price_data_30y["Settle"].to_frame()
training_data["Date"] = training_data.index

import pytz
from pytz import timezone

utc = timezone('UTC')
def to_utc(x):
    # add utc timezone to datetime index
    a = x.replace(tzinfo=pytz.utc)
    return a

training_data["Date"] = training_data["Date"].apply(to_utc)
training_data.set_index(['Date'], inplace=True)
training_data = training_data.rename(columns={"Settle":"Price"})
training_data = add_planet_positions(training_data, planet_list)
training_data.head()
```

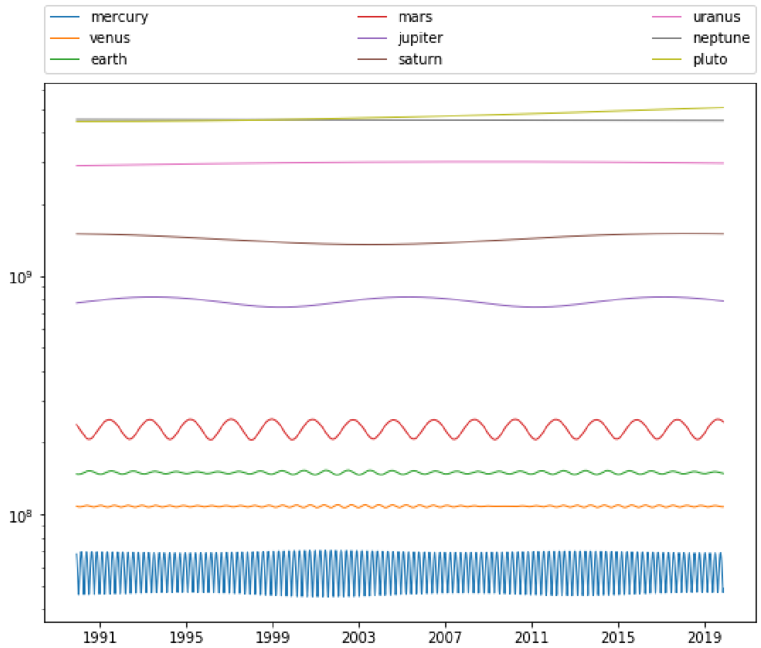
Out[15]:

	Price	mercury_dist	venus_dist	earth_dist	mars_dist	jupiter_dist	saturn_dist	uranus_dist	neptu
Date									
1989-11-30 00:00:00+00:00	607.25	6.810045e+07	1.081717e+08	1.475215e+08	2.374056e+08	7.688738e+08	1.500307e+09	2.898312e+09	4.5194
1989-12-01 00:00:00+00:00	601.75	6.770740e+07	1.081548e+08	1.474997e+08	2.372357e+08	7.689272e+08	1.500298e+09	2.898338e+09	4.5194
1989-12-04 00:00:00+00:00	597.00	6.629697e+07	1.081041e+08	1.474381e+08	2.367209e+08	7.690873e+08	1.500268e+09	2.898419e+09	4.5194
1989-12-05 00:00:00+00:00	599.75	6.575223e+07	1.080873e+08	1.474188e+08	2.365477e+08	7.691407e+08	1.500258e+09	2.898446e+09	4.5194
1989-12-06 00:00:00+00:00	603.50	6.517171e+07	1.080705e+08	1.474001e+08	2.363737e+08	7.691942e+08	1.500248e+09	2.898473e+09	4.5194

In [16]:
```
target = training_data["Price"]
features = training_data.drop(["Price"], axis=1)
```

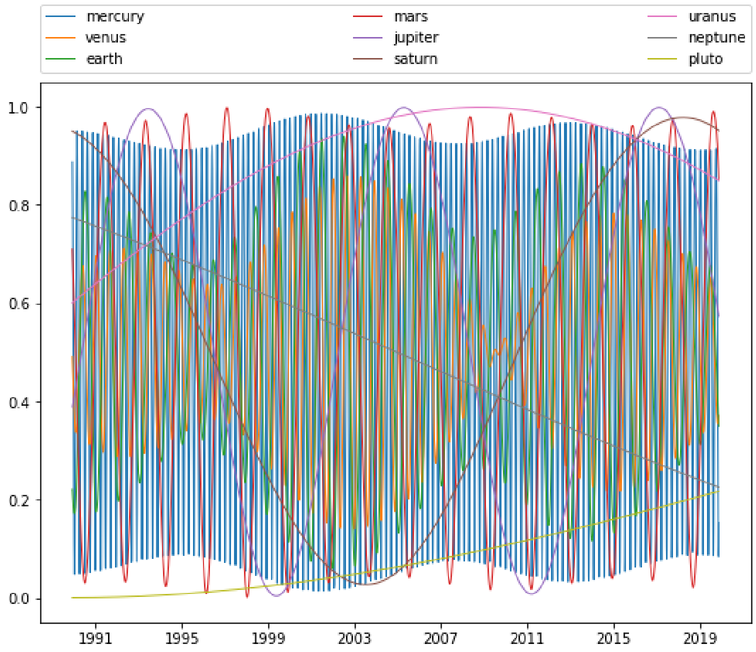
In [17]:
```
planet_distances = plt.figure(figsize=(9,7))
ax = plt.subplot()
for c in list(features):
    line = ax.plot(features.index, features[c], label=c.split("_")[0], linewidth=1)
ax.set_yscale('log')
ax.legend(bbox_to_anchor=(0., 1.02, 1., .102), loc=3,
           ncol=3, mode="expand", borderaxespad=0.)
```

Out[17]: <matplotlib.legend.Legend at 0x7fd28ff11940>



In [18]:
```
features_scaled = pd.DataFrame()
for column in features:
    features_scaled[column] = orbit_scaler(column, features[column])

planet_distances_scaled = plt.figure(figsize=(9,7))
ax = plt.subplot()
for c in list(features_scaled):
    line = ax.plot(features_scaled.index, features_scaled[c], label=c.split("_")[0], linewidth=1)
ax.legend(bbox_to_anchor=(0., 1.02, 1., .102), loc=3,
           ncol=3, mode="expand", borderaxespad=0.)
```

Out[18]: <matplotlib.legend.Legend at 0x7fd28fd5b588>



Training

In [19]:
```
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(
    features_scaled.values, target.values, test_size=0.3, random_state=0)
```

In [20]:
```
from sklearn import svm

clf = svm.SVR(kernel='rbf')
clf.fit(X_train, y_train)
clf.score(X_test, y_test)
```

Out[20]: 0.17058580436119397

Prediction

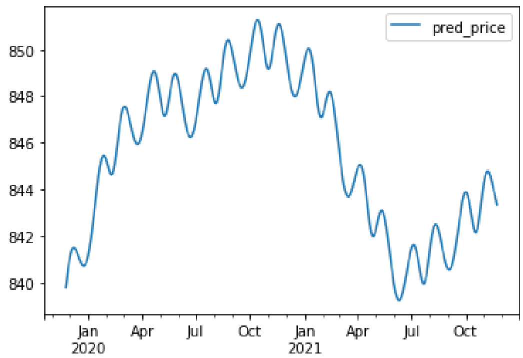
In [21]:
```
tomorrow = datetime.date.today() + datetime.timedelta(days=1)
end_date = tomorrow + datetime.timedelta(days=query[2])
dates = pd.date_range(tomorrow, end_date, freq = 'D', tz='utc')
dates = pd.to_datetime(dates)
future = pd.DataFrame(index=dates, columns=['A'])
future = add_planet_positions(future, planet_list)
future.drop(["A"], axis=1, inplace=True)

# scale
for column in future:
    future[column] = orbit_scaler(column, future[column])

# predict price
future["pred_price"] = clf.predict(future.values)

# export
future["pred_price"].to_csv("_temp/pred_price.csv")

future.head()
```

Out[21]:

	mercury_dist	venus_dist	earth_dist	mars_dist	jupiter_dist	saturn_dist	uranus_dist	neptune_dist	pluto_dist	pre
2019-11-24 00:00:00+00:00	0.195543	0.374960	0.345818	0.846282	0.571786	0.950731	0.849069	0.225019	0.217084	839
2019-11-25 00:00:00+00:00	0.219898	0.376914	0.344221	0.843307	0.571076	0.950642	0.848998	0.224975	0.217118	839
2019-11-26 00:00:00+00:00	0.245800	0.378945	0.342675	0.840306	0.570365	0.950553	0.848927	0.224930	0.217151	840
2019-11-27 00:00:00+00:00	0.272998	0.381054	0.341181	0.837278	0.569655	0.950464	0.848857	0.224885	0.217184	840
2019-11-28 00:00:00+00:00	0.301246	0.383239	0.339739	0.834223	0.568944	0.950375	0.848786	0.224841	0.217217	840

In [22]:
```
future.plot(y="pred_price")
```

Out[22]: <matplotlib.axes._subplots.AxesSubplot at 0x7fd28fd5b7b8>



In [23]:
```
future["change"] = future["pred_price"] - future["pred_price"].shift(-1)
price_preds = future[["pred_price", "change"]].copy()
price_preds = price_preds.fillna(0)
price_preds['direction'] = price_preds.change.map(lambda x: 0 if x == 0 else x/abs(x))
price_preds['dir_change'] = (price_preds.direction != price_preds.direction.shift(1)).astype(int)

# remove change flag for first and last dates
for date in [tomorrow, end_date]:
    price_preds.set_value(date, 'dir_change', 0)

price_preds["block"] = price_preds.dir_change.cumsum()
price_preds.head()
```

```
/home/david/.local/share/virtualenvs/Monistic_Almanac-me3ioVh4/lib/python3.6/site-packages/ipykernel_launc
her.py:9: FutureWarning: set_value is deprecated and will be removed in a future release. Please use .at[]
or .iat[] accessors instead
  if __name__ == '__main__':
```

Out[23]:

	pred_price	change	direction	dir_change	block
2019-11-24 00:00:00+00:00	839.774863	-0.221325	-1.0	0	0
2019-11-25 00:00:00+00:00	839.996188	-0.216123	-1.0	0	0
2019-11-26 00:00:00+00:00	840.212310	-0.206837	-1.0	0	0
2019-11-27 00:00:00+00:00	840.419148	-0.193985	-1.0	0	0
2019-11-28 00:00:00+00:00	840.613133	-0.178147	-1.0	0	0

In [24]:
```
# Format and Output peaks and valleys data

# remove weird warning for now
pd.options.mode.chained_assignment = None

#select rows with a direction change
peaks_valleys = price_preds.loc[price_preds["dir_change"] == 1]

#dump unncessary columns
peaks_valleys.drop(["change","dir_change", "block"], axis=1, inplace=True)

# make a date column, format and dump the index
peaks_valleys["date"] = peaks_valleys.index
#peaks_valleys["date"] = peaks_valleys["date"].apply(lambda x: datetime.datetime.strftime(x, '%Y-%m-%d'))
peaks_valleys = peaks_valleys.reset_index(drop=True)

# truncate price to 2 decimals
#peaks_valleys["pred_price"] = peaks_valleys["pred_price"].apply(lambda x: '%.2f'%(x))

# export csv wihtout the index column
peaks_valleys.to_csv("_temp/peaks_valleys.csv", index=False)
peaks_valleys
```

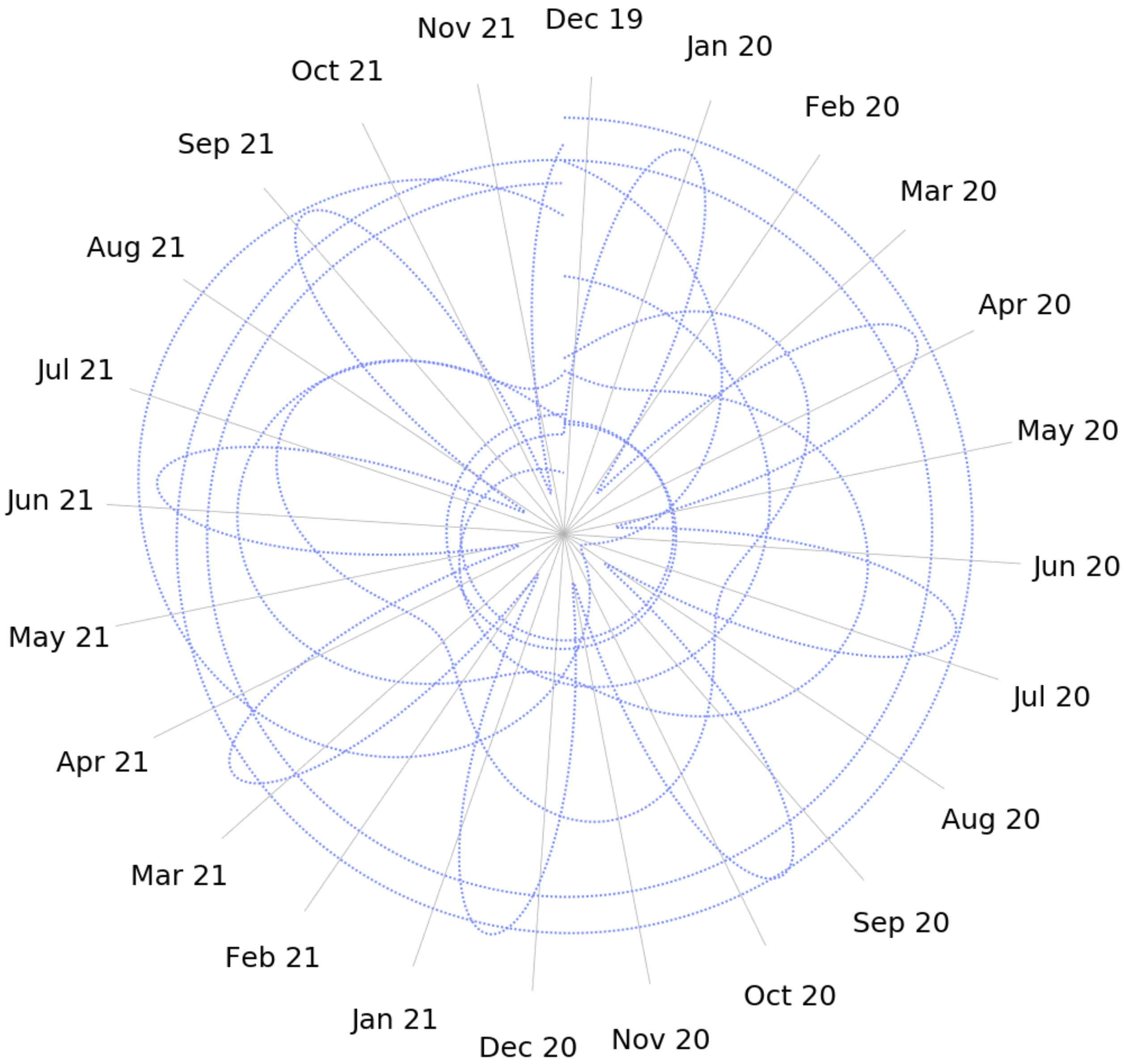
Out[24]:

	pred_price	direction	date
0	841.486027	1.0	2019-12-07 00:00:00+00:00
1	840.708833	-1.0	2019-12-24 00:00:00+00:00
2	845.453643	1.0	2020-01-27 00:00:00+00:00
3	844.620229	-1.0	2020-02-09 00:00:00+00:00
4	847.557736	1.0	2020-03-02 00:00:00+00:00
5	845.922454	-1.0	2020-03-24 00:00:00+00:00
6	849.083526	1.0	2020-04-21 00:00:00+00:00
7	847.146554	-1.0	2020-05-09 00:00:00+00:00
8	848.961004	1.0	2020-05-27 00:00:00+00:00
9	846.235501	-1.0	2020-06-22 00:00:00+00:00
10	849.182710	1.0	2020-07-19 00:00:00+00:00
11	847.688958	-1.0	2020-08-04 00:00:00+00:00
12	850.404279	1.0	2020-08-25 00:00:00+00:00
13	848.372492	-1.0	2020-09-17 00:00:00+00:00
14	851.281445	1.0	2020-10-14 00:00:00+00:00
15	849.172357	-1.0	2020-11-01 00:00:00+00:00
16	851.097627	1.0	2020-11-19 00:00:00+00:00
17	847.990035	-1.0	2020-12-16 00:00:00+00:00
18	850.044612	1.0	2021-01-08 00:00:00+00:00
19	847.080302	-1.0	2021-01-29 00:00:00+00:00
20	848.182510	1.0	2021-02-12 00:00:00+00:00
21	843.683764	-1.0	2021-03-16 00:00:00+00:00
22	845.052088	1.0	2021-04-05 00:00:00+00:00
23	841.966451	-1.0	2021-04-27 00:00:00+00:00
24	843.077891	1.0	2021-05-12 00:00:00+00:00
25	839.226117	-1.0	2021-06-10 00:00:00+00:00
26	841.600419	1.0	2021-07-05 00:00:00+00:00
27	839.930071	-1.0	2021-07-22 00:00:00+00:00
28	842.495473	1.0	2021-08-11 00:00:00+00:00
29	840.535895	-1.0	2021-09-03 00:00:00+00:00
30	843.879361	1.0	2021-10-01 00:00:00+00:00
31	842.138410	-1.0	2021-10-18 00:00:00+00:00
32	844.777843	1.0	2021-11-07 00:00:00+00:00

Charting layers

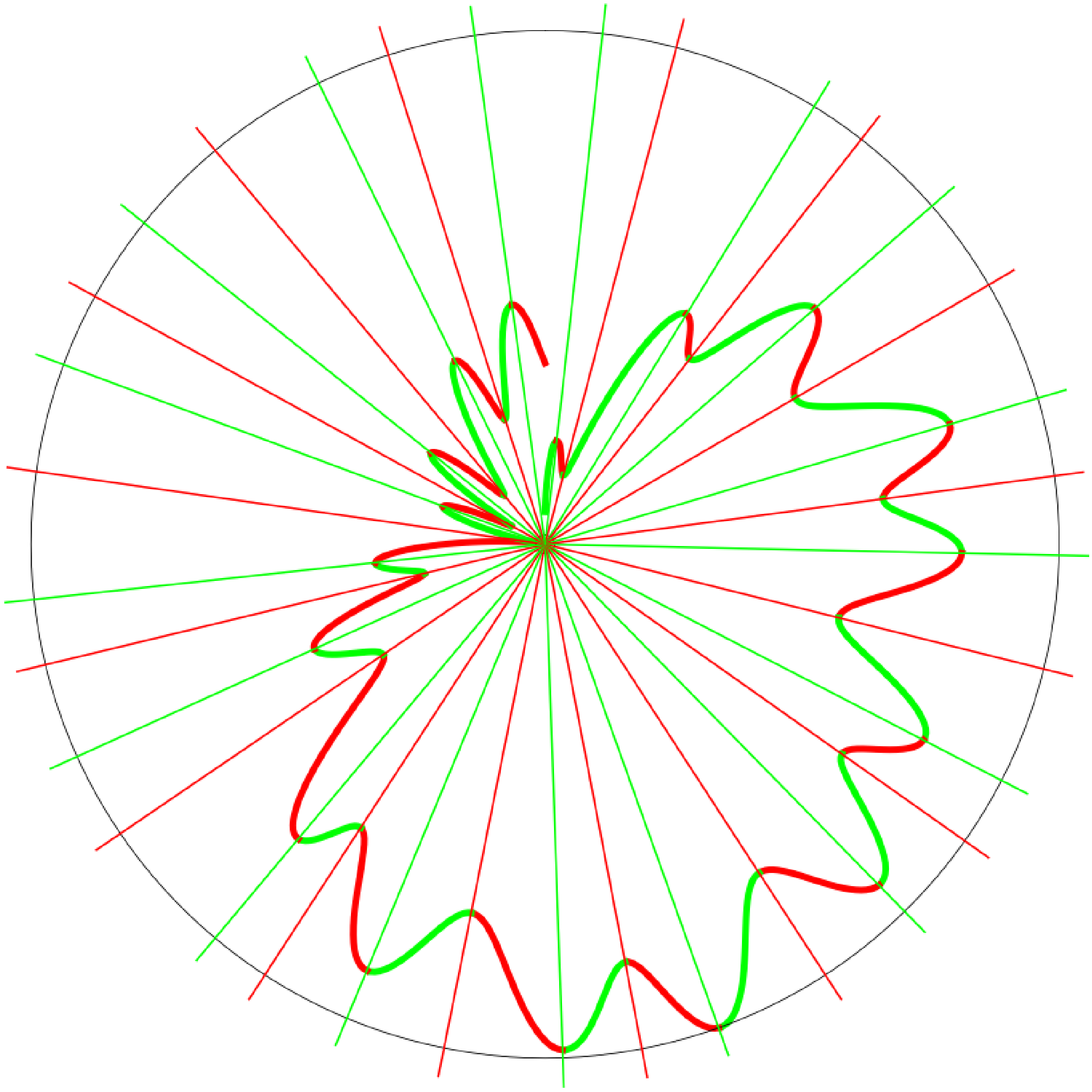
In [25]:
```
import matplotlib.dates as mdates
import matplotlib.ticker as ticker
```

In [26]:
```
# Features (planets) chart

# months for axis ticks - fill month_starts and drop columns
future["month"] = future.index.strftime('%m').astype(int)
future["month_diff"] = future.month.diff().fillna(0)
month_starts = future.loc[future['month_diff'] != 0]
future.drop(["month","month_diff"], axis=1,inplace=True)

# remove backgrounds (has to be specified again for savefig())
plt.rcParams['figure.facecolor'] = ((0,0,0,0.))
plt.rcParams['axes.facecolor'] = ((0,0,0,0.))

# create plot
fig = plt.figure(figsize=(15,15))
ax = plt.subplot(projection='polar')
ax.set_theta_direction(-1)
ax.set_theta_zero_location("N")

# normalise theta axis and ticks
t = mdates.date2num(future.index.to_pydatetime())
tnorm = (t-t.min())/(t.max()-t.min())*2.*np.pi

month_ticks= []
month_labels = []

for index, row in month_starts.iterrows():
    m = mdates.date2num(index.to_pydatetime())
    mnorm = (m-t.min())/(t.max()-t.min())*2.*np.pi
    month_ticks.append(mnorm)
    month_labels.append(index.strftime('%b %y'))

ax.xaxis.set_major_locator(ticker.FixedLocator((month_ticks)))
ax.xaxis.set_major_formatter(ticker.FixedFormatter((month_labels)))
ax.xaxis.set_tick_params(pad=40, labelsize=25)

# Y/Radial Axis - remove ticks
ax.set_rticks([],[])
# Remove outer circle
ax.spines['polar'].set_visible(False)

for planet in planet_list:
    col_name = planet + "_dist"
    ax.plot(tnorm, future[col_name],
        linewidth=2,
        linestyle="dotted",
        dashes=(1, 1),
        color = '#7786FF')

plt.savefig("_temp/planets.png", facecolor = (0,0,0,0.), transparent=True)
```



Dec 19
Jan 20
Feb 20
Mar 20
Apr 20
May 20
Jun 20
Jul 20
Aug 20
Sep 20
Oct 20
Nov 20
Dec 20
Jan 21
Feb 21
Mar 21
Apr 21
May 21
Jun 21
Jul 21
Aug 21
Sep 21
Oct 21
Nov 21

In [27]:
```
# create plot
fig = plt.figure(figsize=(15,15))
ax = plt.subplot(projection='polar')
ax.set_theta_direction(-1)
ax.set_theta_zero_location("N")

# normalise theta axis and ticks
t = mdates.date2num(future.index.to_pydatetime())
tnorm = (t-t.min())/(t.max()-t.min())*2.*np.pi

# set min and max
import math
price_min = future["pred_price"].min()
axis_min = math.floor(price_min)

price_max = future["pred_price"].max()
axis_max = math.ceil(price_max)
ax.set_ylim(axis_min, axis_max)

peak_ticks= []
peak_labels = []
peak_values = []

for index, row in peaks_valleys.iterrows():
    m = mdates.date2num(row['date'].to_pydatetime())
    mnorm = (m-t.min())/(t.max()-t.min())*2.*np.pi
    #peak_ticks.append(mnorm)
    peak_labels.append("${0:.2f} ".format(row['pred_price']))
    peak_values.append(row['pred_price'])
    if row['direction'] < 0:
        c = 'r'
    else:
        c = (0,1,0)
    ax.plot([axis_min,mnorm],[axis_min,axis_max], linestyle = '-', color = c)

# theta ticks off
ax.xaxis.set_major_locator(ticker.FixedLocator((peak_ticks)))

# Radial ticks
if peak_values:
    peak_values.sort()
    peak_labels.sort()
    peak_labels = [peak_labels[0], peak_labels[-1]]
    peak_values = [peak_values[0], peak_values[-1]]

else:
    peak_values = [price_min, price_max]
    peak_labels = ["${0:.2f} ".format(x) for x in peak_values]

ax.set_rgrids(peak_values, labels=None, color='k', angle=0, ha="left", zorder=10)
ax.yaxis.set_tick_params(pad=25, labelsize=0, width=20)

ax.grid(color="k")

# plot block by block - green if start < end, red if start > end
for name, block in price_preds.groupby(['block']):
    if block.head(1)["pred_price"].values < block.tail(1)["pred_price"].values:
        c = (0,1,0)
    else:
        c = 'r'
    # normalise theta axis and ticks
    year = mdates.date2num(price_preds.index.to_pydatetime())
    t = mdates.date2num(block.index.to_pydatetime())
    tnorm = (t-year.min())/(year.max()-year.min())*2.*np.pi

    ax.plot(tnorm, block["pred_price"],
        linewidth=5,
        color = c, zorder= 1)

ax.spines['polar'].set_visible(False)
plt.savefig("_temp/pred_price.png", facecolor = (0,0,0,0.), transparent=True)
```



Combine layers into the final chart

In [28]:
```
from PIL import Image, ImageDraw, ImageFont
import os
```

In [29]:
```
 # dim of the matplotlib PDFs
width = 1080
height = 1400

center_x = width/2
center_y = (width/2) + (height - width)
```

In [30]:
```
preds = pd.read_csv('_temp/pred_price.csv', names = ["date", "price"])
first_date = preds.iloc[0]["date"].split(" ")[0]
first_day = datetime.datetime.strptime(first_date, "%Y-%m-%d")
first_price = round(float(preds.iloc[0]["price"]), 2)

last_date = preds.iloc[-1]["date"].split(" ")[0]
last_day = datetime.datetime.strptime(last_date, "%Y-%m-%d")
last_price = round(float(preds.iloc[-1]["price"]), 2)
```

In [31]:
```
def price_flag(date, price, side, up=None, diff=None):
        '''
        draw price boxes at the top
        date: a datetime
        price: float
        side: "L" or "R"
        up: Boolean
        diff: price difference float
        '''

        w = 400
        h = 140
        y = ((height - width)/2) - (h/2)

        if side == "L":
            x = center_x - w
        else:
            x = center_x
        if side == "L":
            if up == True:
                y -= 30
            else:
                y += 30

        if up == True:
            color = (0,255,0) # green
        elif up == False:
            color = (255, 0, 0) # red
        else:
            color = 'black'


        idraw.rectangle((x, y, x + w, y + h), fill=None, outline="black",)
        date_string = date.strftime("%Y-%m-%d")
        idraw.text((x + 10, y + 5), date_string, fill='black', font=operator_date)


        sign = ""
        if diff:
            if diff > 0:
                sign = "+"
            if diff < 0:
                sign = "-"
            price_diff = sign + "$" + str(round(abs(diff),2))
            idraw.text((x + 10, y + 150), price_diff, fill=color, font=operator_b_diff)


        price_tag = "$" + str(price)
        idraw.text((x + 10, y + 30), price_tag, fill=color, font=operator_b_price)
```

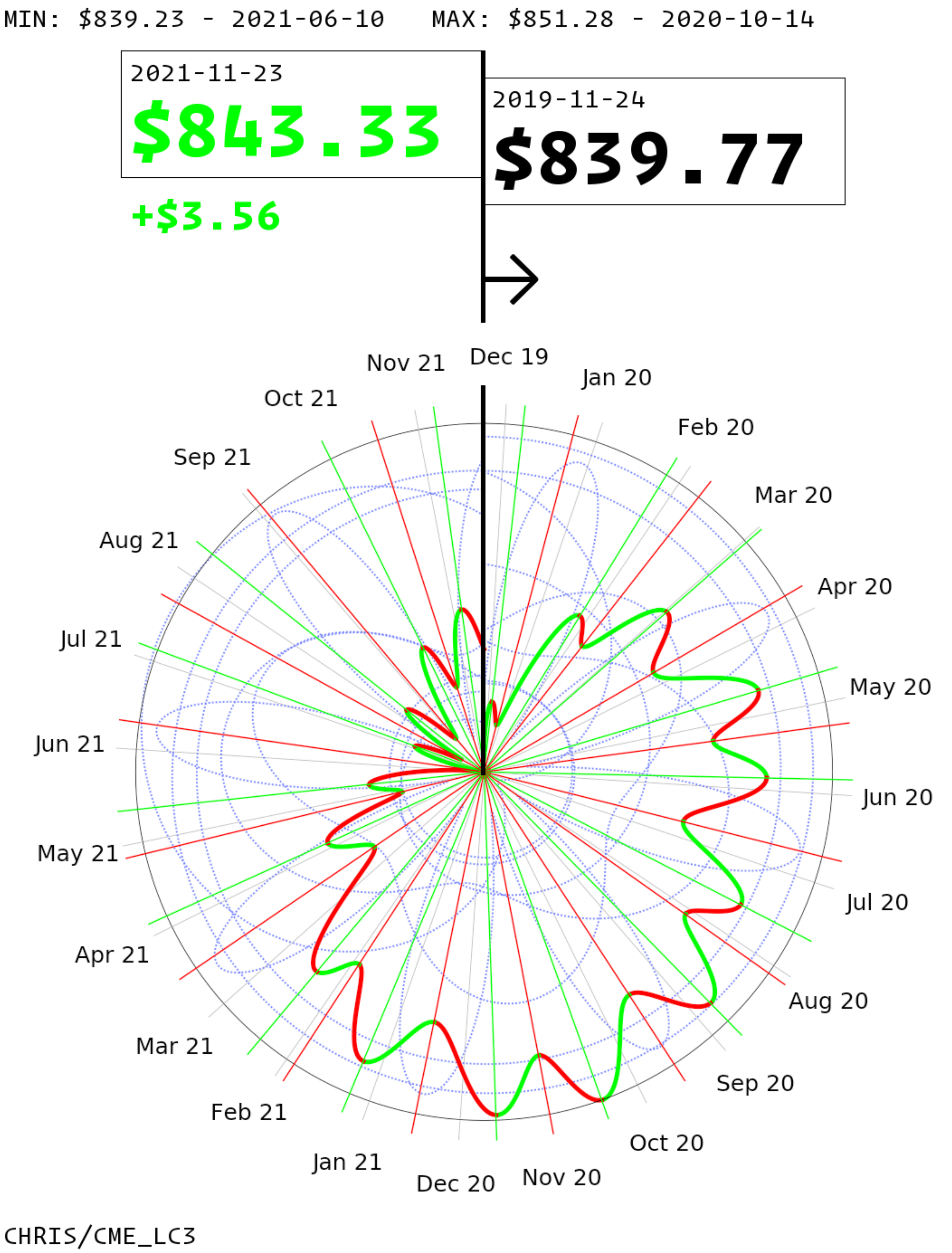
In [32]:
```
# %% fonts (proprietary, not included in the repository)
font_path = os.environ['TMA_home'] + '/_data/fonts/'
operator_date = ImageFont.truetype(font= font_path + "OperatorMono-Book.ttf", size=30)
operator_b_diff = ImageFont.truetype(font= font_path + "OperatorMono-Bold.ttf", size=50)
operator_b_price = ImageFont.truetype(font= font_path + "OperatorMono-Bold.ttf", size=90)
calibri = ImageFont.truetype(font= font_path + "Calibri.ttf", size=80)
```

In [33]:
```
# %% set up image
img = Image.new('RGBA', (width,height))
idraw = ImageDraw.Draw(img)

#background
idraw.rectangle((0,0,width,height), fill="white")


# %% Matplotlib Layers

planets_lyr = Image.open("_temp/planets.png")
preds_lyr = Image.open("_temp/pred_price.png")

for layer in [planets_lyr, preds_lyr]:
    #layer = layer.thumbnail((1080,1080), Image.ANTIALIAS)
    img.paste(layer, (-12,height-width), layer)


# %% Price Labels and center line

price_flag(first_day, first_price, "R")

is_up = first_price < last_price
price_flag(last_day, last_price, "L", up=is_up, diff=last_price - first_price)

# center line and arrow
up_offset = 60 if is_up else 90
idraw.line([(center_x, (height-width)+40),(center_x,up_offset)], fill='black', width=5)
idraw.line([(center_x, center_y),(center_x,430)], fill='black', width=5)

idraw.text((center_x-5,(height-width)-60), "→", fill='black', font=calibri)

# %% description
chart_info = open("_temp/CCC_title.txt").readlines()
chart_title = chart_info[4].rstrip()

idraw.text((10,height-50), chart_title, font=operator_date, fill='black')


# %% min and max
minmax = "{type}: ${price} - {date}"
min = preds.loc[preds['price'].idxmin()]
max = preds.loc[preds['price'].idxmax()]

min_str = minmax.format(type="MIN", price=round(min.price, 2), date=min.date.split(" ")[0])
max_str = minmax.format(type="MAX", price=round(max.price, 2), date=max.date.split(" ")[0])

idraw.text((10,5), min_str + "   " + max_str, font=operator_date, fill='black')
```

In [34]:
```
print(f"{query[0]} {query[1]} Chart")
print(f"{query[2]} Days, starting {chart_info[0]}")
print(f"Futures contract: {chart_info[4]}")
img
```

```
Full Soybean Chart
730 Days, starting 2019-11-21 00:10:15.412376

Futures contract: CHRIS/CME_LC3
```

Out[34]:



Buying/Selling opportunities

In [35]:
```
# Buying Low
peaks_valleys.loc[peaks_valleys['direction'] == -1]
```

Out[35]:

	pred_price	direction	date
1	840.708833	-1.0	2019-12-24 00:00:00+00:00
3	844.620229	-1.0	2020-02-09 00:00:00+00:00
5	845.922454	-1.0	2020-03-24 00:00:00+00:00
7	847.146554	-1.0	2020-05-09 00:00:00+00:00
9	846.235501	-1.0	2020-06-22 00:00:00+00:00
11	847.688958	-1.0	2020-08-04 00:00:00+00:00
13	848.372492	-1.0	2020-09-17 00:00:00+00:00
15	849.172357	-1.0	2020-11-01 00:00:00+00:00
17	847.990035	-1.0	2020-12-16 00:00:00+00:00
19	847.080302	-1.0	2021-01-29 00:00:00+00:00
21	843.683764	-1.0	2021-03-16 00:00:00+00:00
23	841.966451	-1.0	2021-04-27 00:00:00+00:00
25	839.226117	-1.0	2021-06-10 00:00:00+00:00
27	839.930071	-1.0	2021-07-22 00:00:00+00:00
29	840.535895	-1.0	2021-09-03 00:00:00+00:00
31	842.138410	-1.0	2021-10-18 00:00:00+00:00

In [36]:
```
# Selling High
peaks_valleys.loc[peaks_valleys['direction'] == 1]
```

Out[36]:

	pred_price	direction	date
0	841.486027	1.0	2019-12-07 00:00:00+00:00
2	845.453643	1.0	2020-01-27 00:00:00+00:00
4	847.557736	1.0	2020-03-02 00:00:00+00:00
6	849.083526	1.0	2020-04-21 00:00:00+00:00
8	848.961004	1.0	2020-05-27 00:00:00+00:00
10	849.182710	1.0	2020-07-19 00:00:00+00:00
12	850.404279	1.0	2020-08-25 00:00:00+00:00
14	851.281445	1.0	2020-10-14 00:00:00+00:00
16	851.097627	1.0	2020-11-19 00:00:00+00:00
18	850.044612	1.0	2021-01-08 00:00:00+00:00
20	848.182510	1.0	2021-02-12 00:00:00+00:00
22	845.052088	1.0	2021-04-05 00:00:00+00:00
24	843.077891	1.0	2021-05-12 00:00:00+00:00
26	841.600419	1.0	2021-07-05 00:00:00+00:00
28	842.495473	1.0	2021-08-11 00:00:00+00:00
30	843.879361	1.0	2021-10-01 00:00:00+00:00
32	844.777843	1.0	2021-11-07 00:00:00+00:00

Appendix F

Possibility Space

Exploratory project, produced as part of this research but not discussed in the thesis.

Draft 1

The following pages were produced for the *Methods of Intent* post-graduate seminar (14th March 2017) convened by Prof. Teal Triggs at the School of Communication, Royal College of Art.

Possibility Space

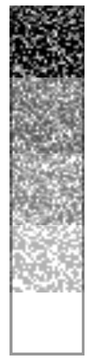
Speculative Design in the regime of machine futures

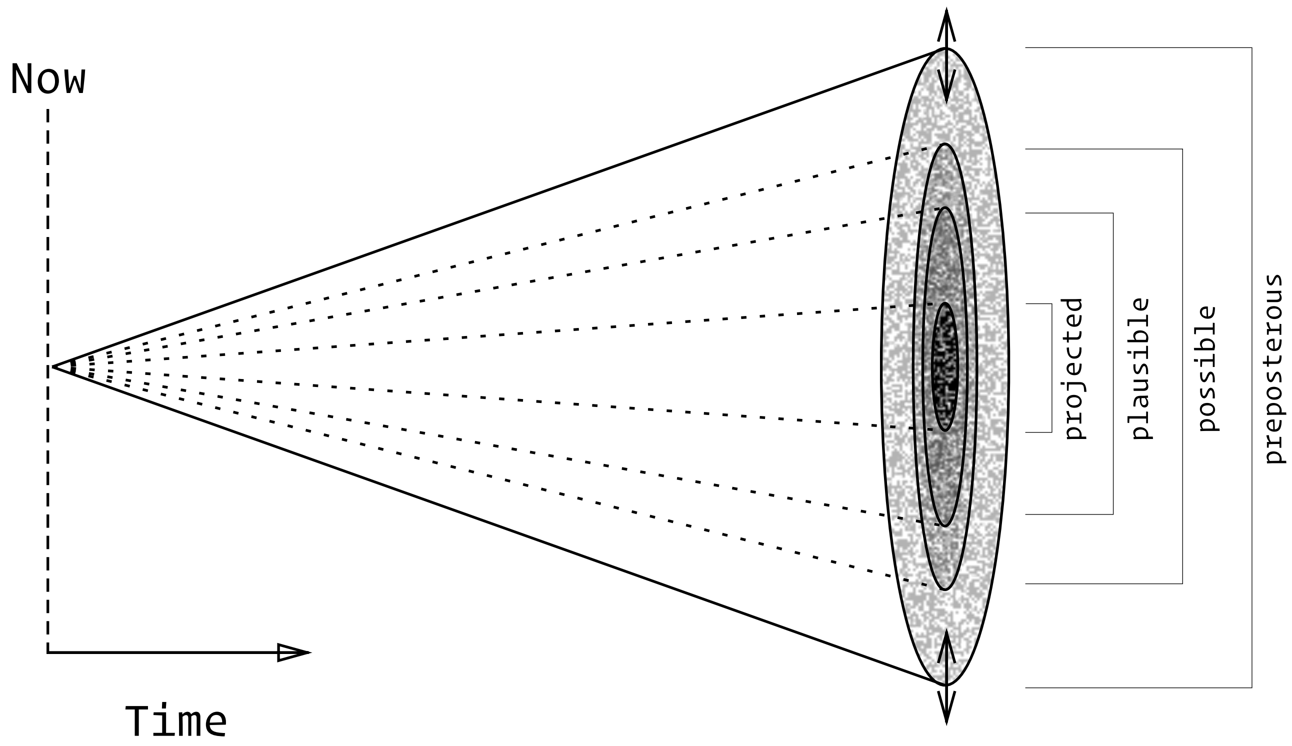
Draft 1.0

David Benqué
Methods of Intent
14/03/2017

Conical Futures

Possibility Space



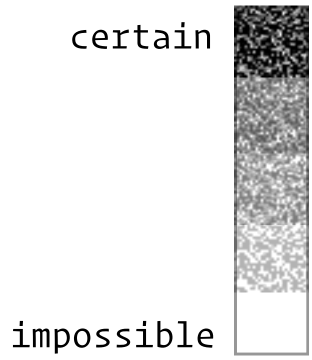


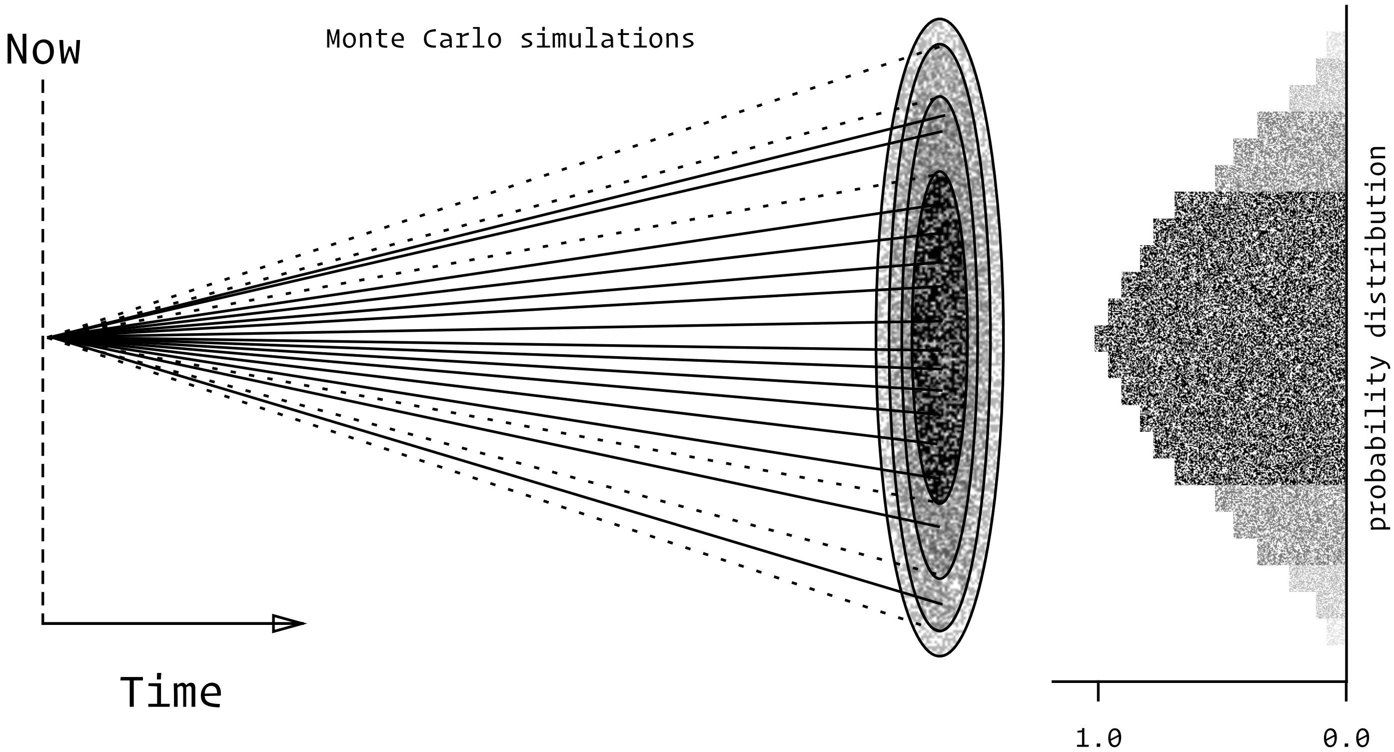
redrawn after:
Voros, J. (2003) ‘A generic foresight process framework’, *foresight*, 5(3), pp. 10–21. doi: 10.1108/14636680310698379.
who had redrawn after:
Hancock and Bezold (1993) ‘An Overview of the Health Futures Field’, WHO Consultation, July 19-23.

‘preposterous’ added in 2015:
Voros, J. (2015) On examining Preposterous! futures, The Voroscope, 28 December. Available at: https://thevoroscope.com/category/preposterous-futures/ (Accessed: 14 March 2017).

Probabilistic Futures

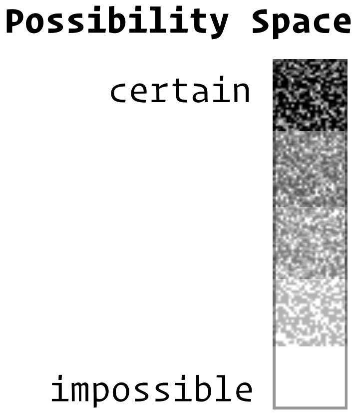
Possibility Space

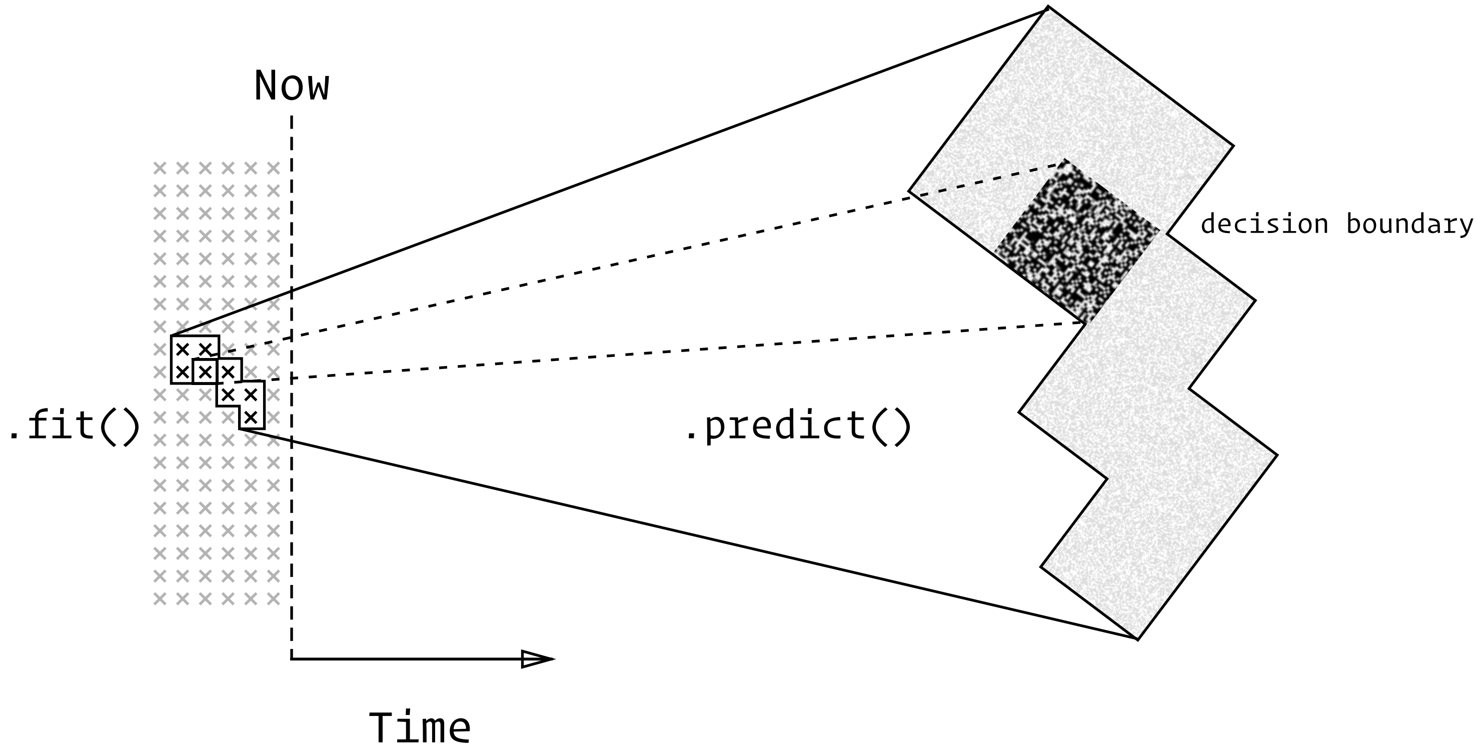





```
sample_size = 10000

for n in range(sample_size):
     rf = mc.random_future()
     if rf.end_point in projected:
          projected_futures.append(rf)
     elif rf.end_point in plausible:
          plausible_futures.append(rf)
     elif rf.end_point in possible:
          possible_futures.append(rf)
     elif rf.end_point in preposterous:
          preposterous_futures.append(rf)
     else:
          wildcards.append(rf)

plt.hist(futures)
```

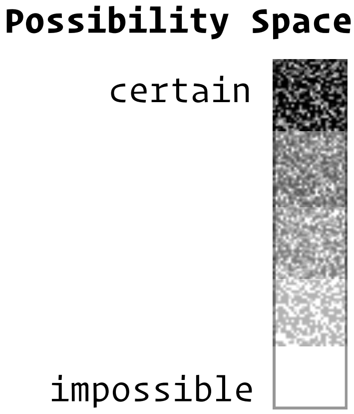
Machine Futures

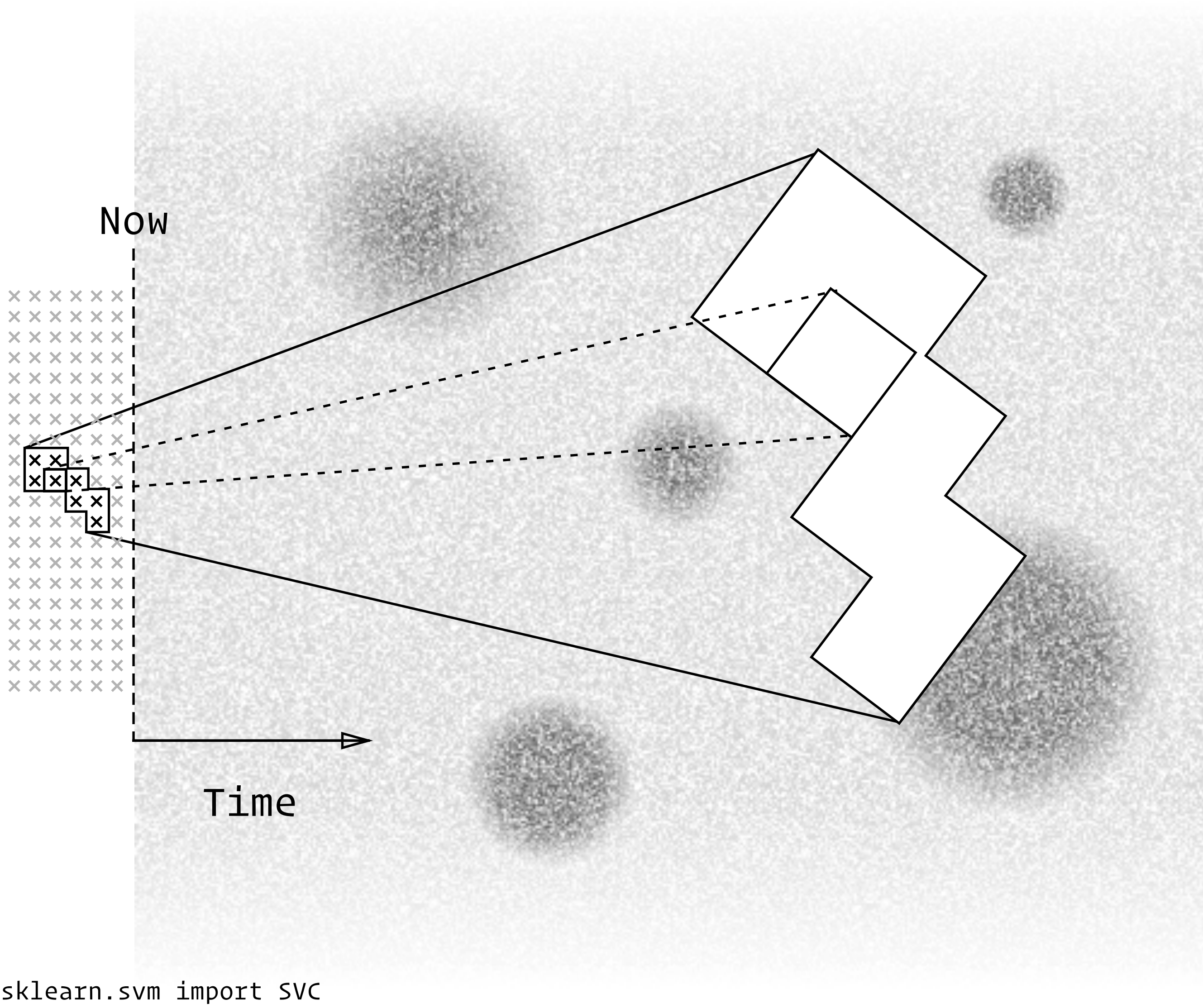





```
from sklearn.svm import SVC

clf = SVC(kernel="linear")
clf.fit(features_train, labels_train)
future = clf.predict(features_test)
```

Emancipatory Machine Futures






```
from sklearn.svm import SVC

clf = SVC(kernel="linear")
clf.fit(features_train, labels_train)
machine_futures = clf.predict(features_test)

for future in infinite_futures:
     if future not in machine_futures:
          interesting_futures.append(futures)
     else:
          pass
```

Benque, D. (2019) ‘World.predict(); Speculative design and machine futures’, PhD thesis, Royal College of Art, London UK.

Possibility Space - Draft 2

The following images are frames from animations produced for the *Uncertainty Playground* exhibition, part of London Design Festival 2017 at the London College of Communication.

The full animation is available at the following link:
https://gitlab.com/davidbenque/thesis/raw/master/06_Appendix/poss-space/Poss_Space_LCC_FULL.mp4

Exhibition Text:
This project revisits the notion of 'possibility space' as a diagrammatic depiction of possible futures.

The *futures cone* diagram, brought to design from future studies by Candy (2010), played a key role in defining speculative design as a practice to test possible futures, and discuss which might be a preferable one.

Our current moment is also saturated by attempts to count, compute, and predict the future using algorithms trained on large datasets. These systems also operate on possibility spaces, multi-dimensional arrays of data, and make predictions through regression or classification.

This project compares and/or conflicts these different versions of possibility space through experiments with representations and 'operations' (Bach *et al.,* 2016*).* Through these diagrams, the aim is to reflect on different modes of speculation; between a *firmative* position which aims to solidify and control the future, and an *affirmative* one which aims to open up new possibles (Uncertain Commons, 2013).

References:

Bach, B., Dragicevic, P., Archambault, D., Hurter, C. and Carpendale, S. (2016) 'A Descriptive Framework for Temporal Data Visualizations Based on Generalized Space-Time Cubes', Computer Graphics Forum, pp. 1–26. doi: 10.1111/cgf.12804.

Candy, S. (2010) *The futures of everyday life: Politics and the design of experiential scenarios.*

Uncertain Commons (2013) *Speculate This!* Duke University Press.

Figure F.1: *Possibility Space* - Title Card. animation still.

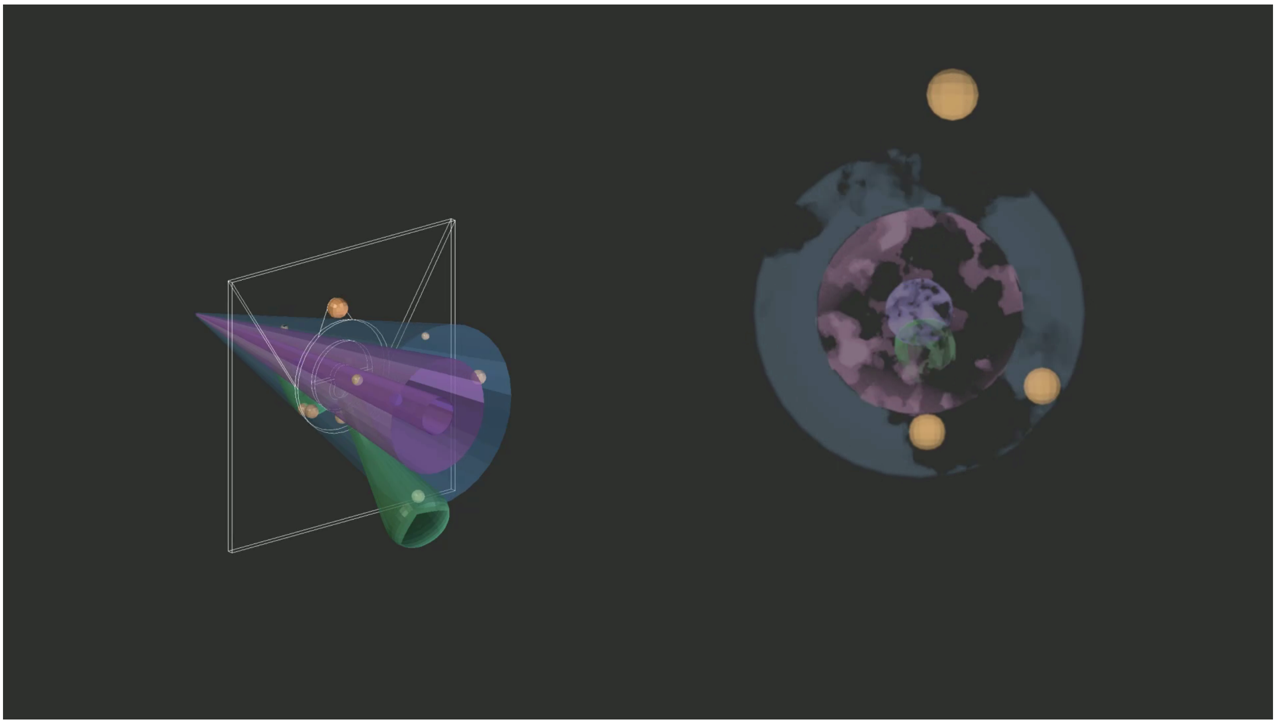


Figure F.2: Futures Cone; Reverse Time Slicing. animation still.

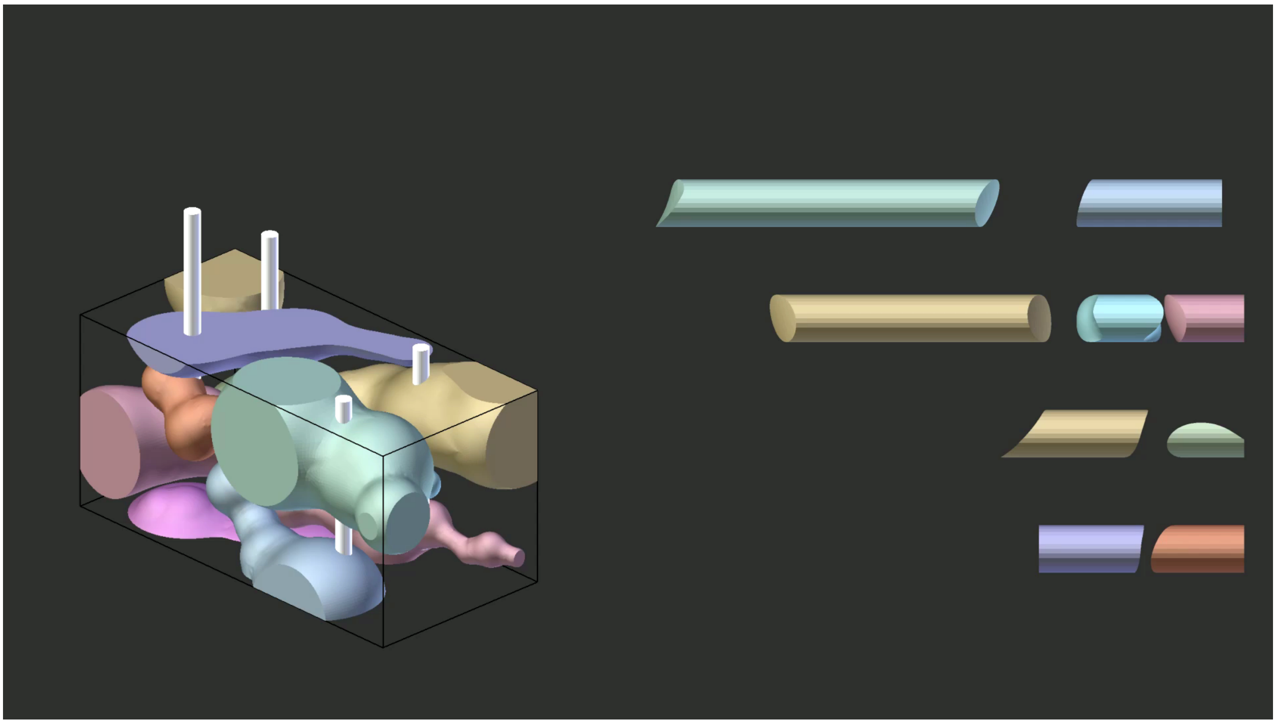


Figure F.3: Clusters; Repeated Drilling. animation still.

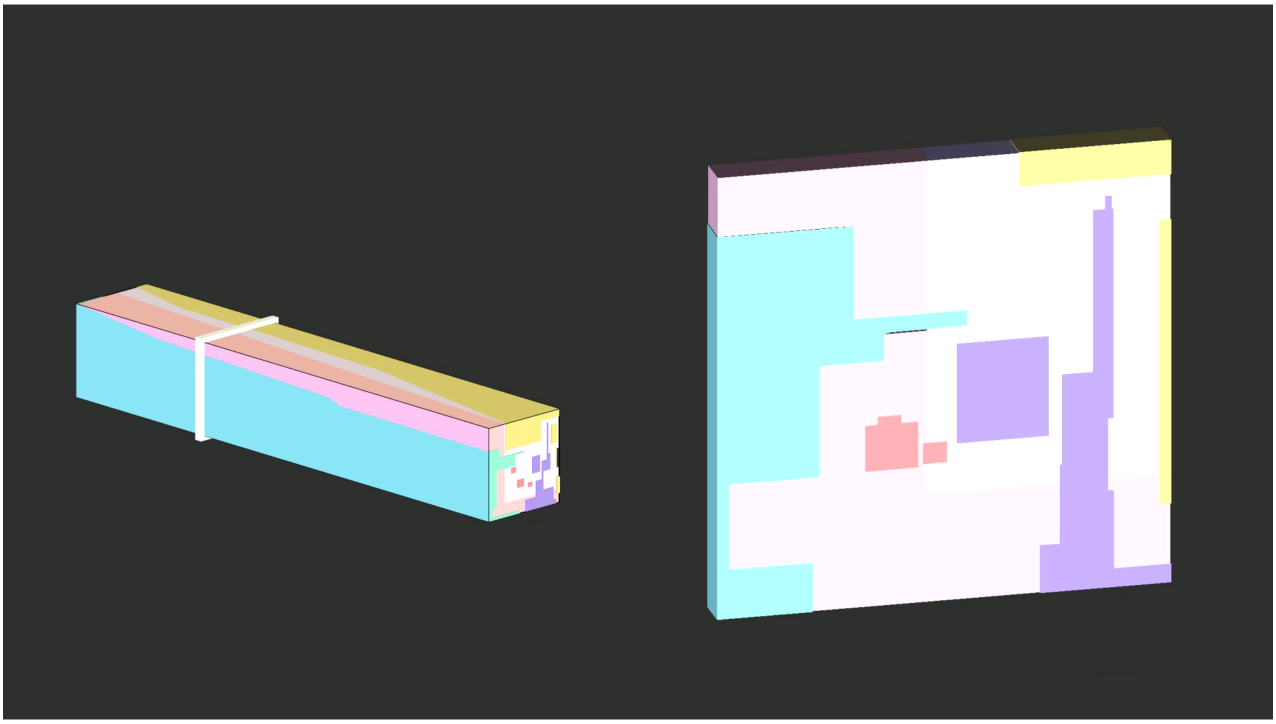


Figure F.4: Decision Surface; Time Shift. animation still.

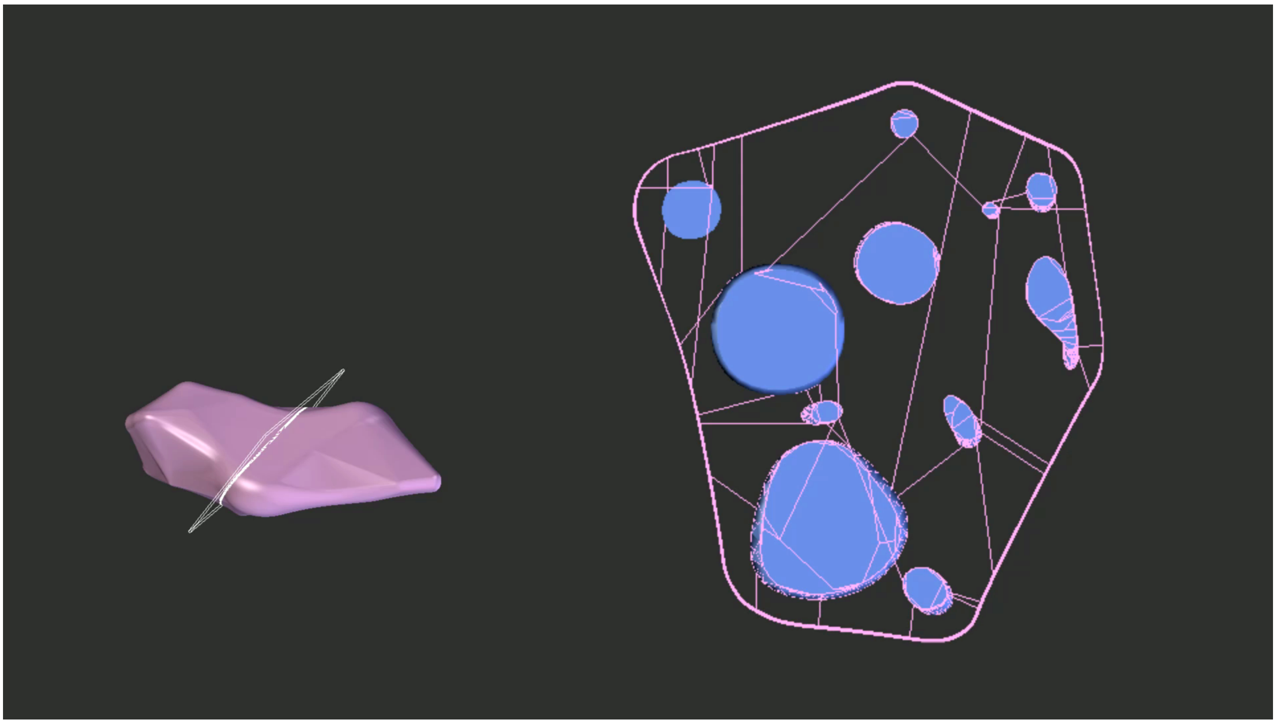


Figure F.5: Haruspex; Repeated Oblique Slicing. animation still.

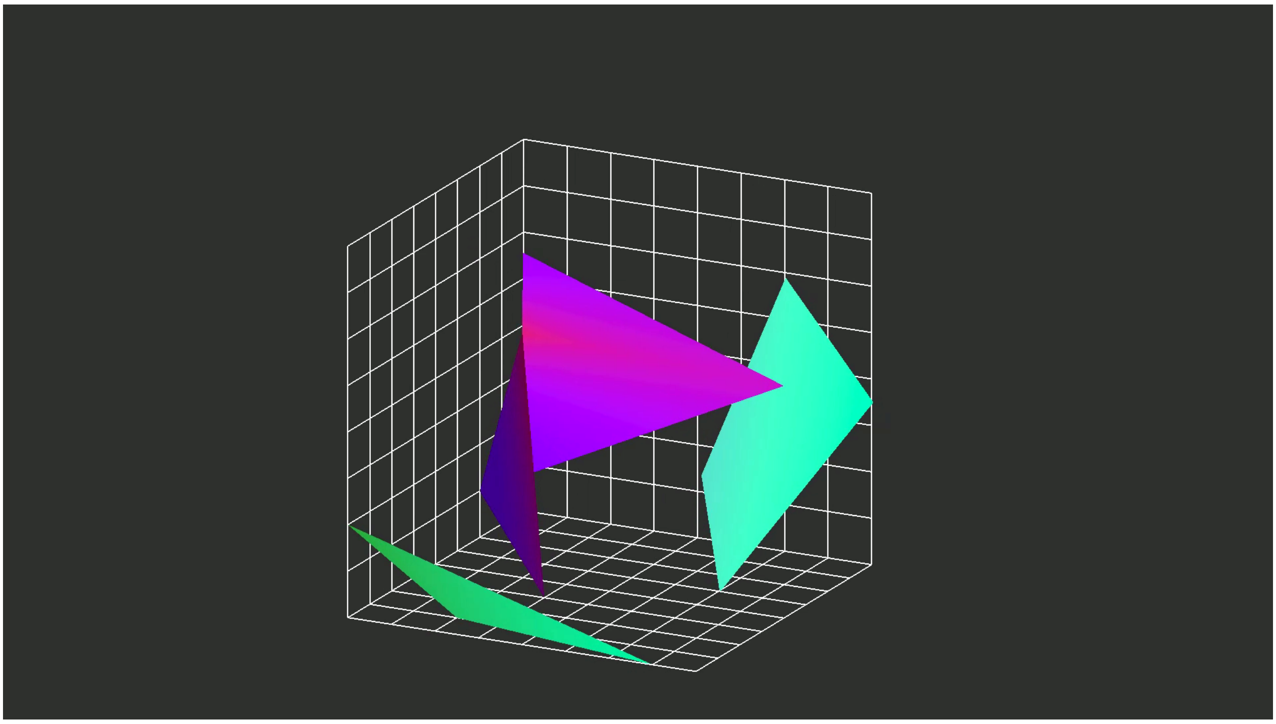


Figure F.6: Probability Space; Orthogonal Flattening. animation still.

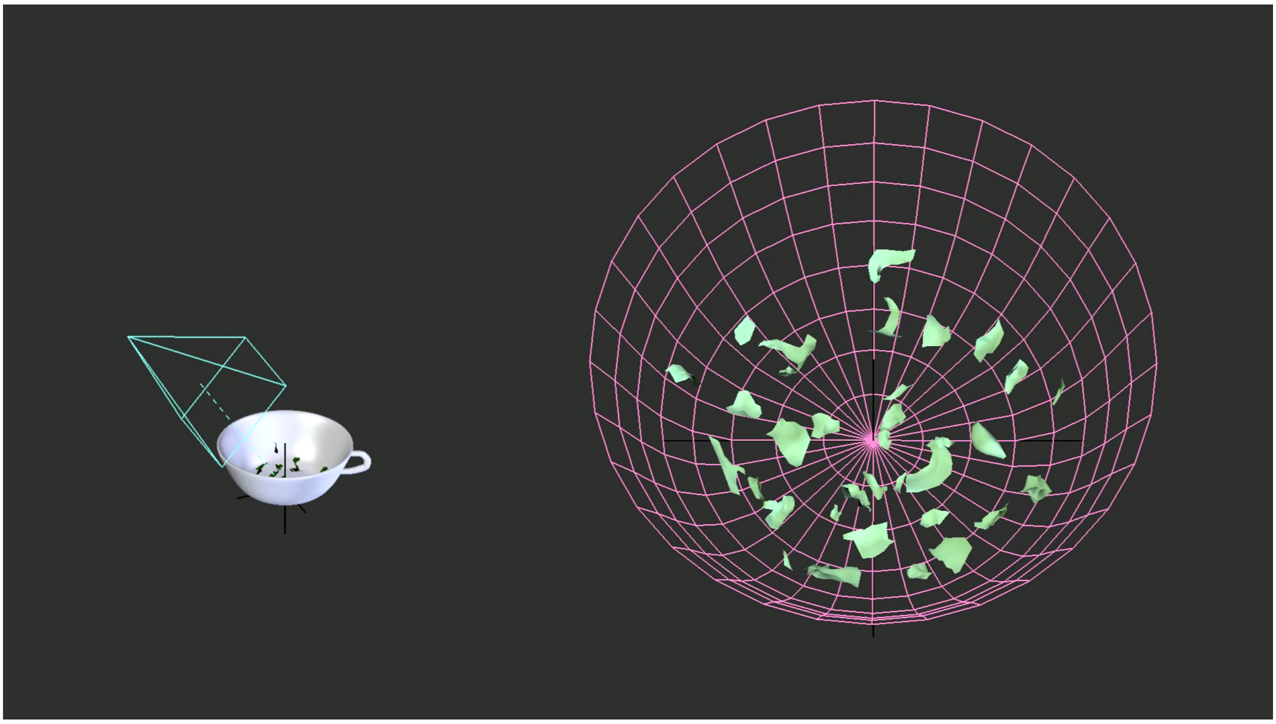


Figure F.7: Tea Cup; Oblique Flattening (orbit). animation still.

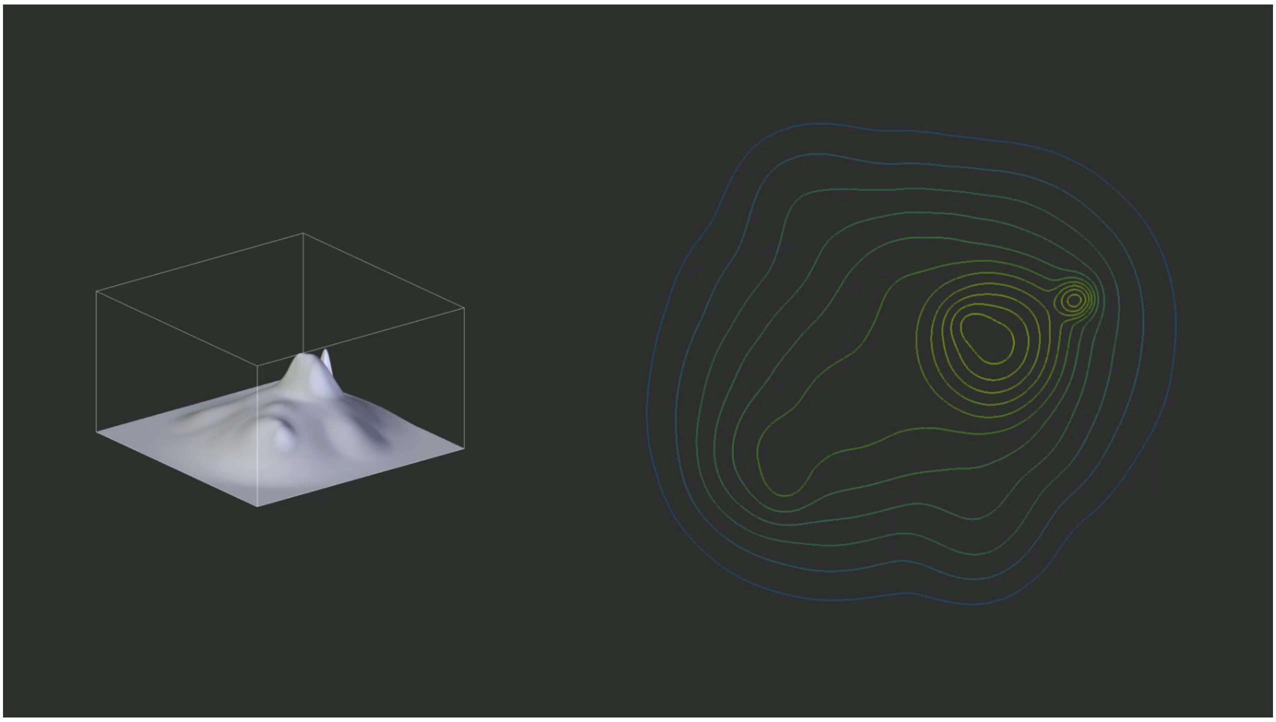


Figure F.8: Probability Distribution; Contour Slicing. animation still.

Appendix G

The Presage Range

Exploratory project, produced as part of this research but not discussed in the thesis.

Commissioned as part of Fiber Festival 2017 *Prima Materia; Alchemical Thinking And Making In Art Design And Music.*

Exhibtion text:
The PRESAGE™ Range is a series of fictional scientific machines that predict the future. These devices reference historical scientific theories and the imagery of lab-equipment catalogs to explore the aesthetics of mechanical objectivity.

C-Type prints, 70x70cm each

- Press Release -
12.05.2017

Aviato, the worldwide leader in predictive hardware solutions, is proud to release PRESAGE™, its latest line of cutting edge instruments. These combine time-tested science with the latest technology to deliver reliable prediction and model-checking in super-computing, genetics and climatology.

Aviato has long believed in the power, accuracy and security of on-premises hardware based solutions. PRESAGE™ brings these advantages to a new level by providing ground-truth models of unmatched reliability with the lowest possible latency. Our precision engineering will allow you to:
- Calibrate and prove vortex models faster than ever before (BOL930i).
- Predict cell-fate determination with unprecedented accuracy (WAD720i).
- Interact directly with composite atmosphere meta-models (RIC320i).

The PRESAGE™ range opens a new era in predictive hardware and demonstrates Aviato's commitment to innovation as the industry leader in the quest to extend knowledge and reduce uncertainty.

- - - END - - -

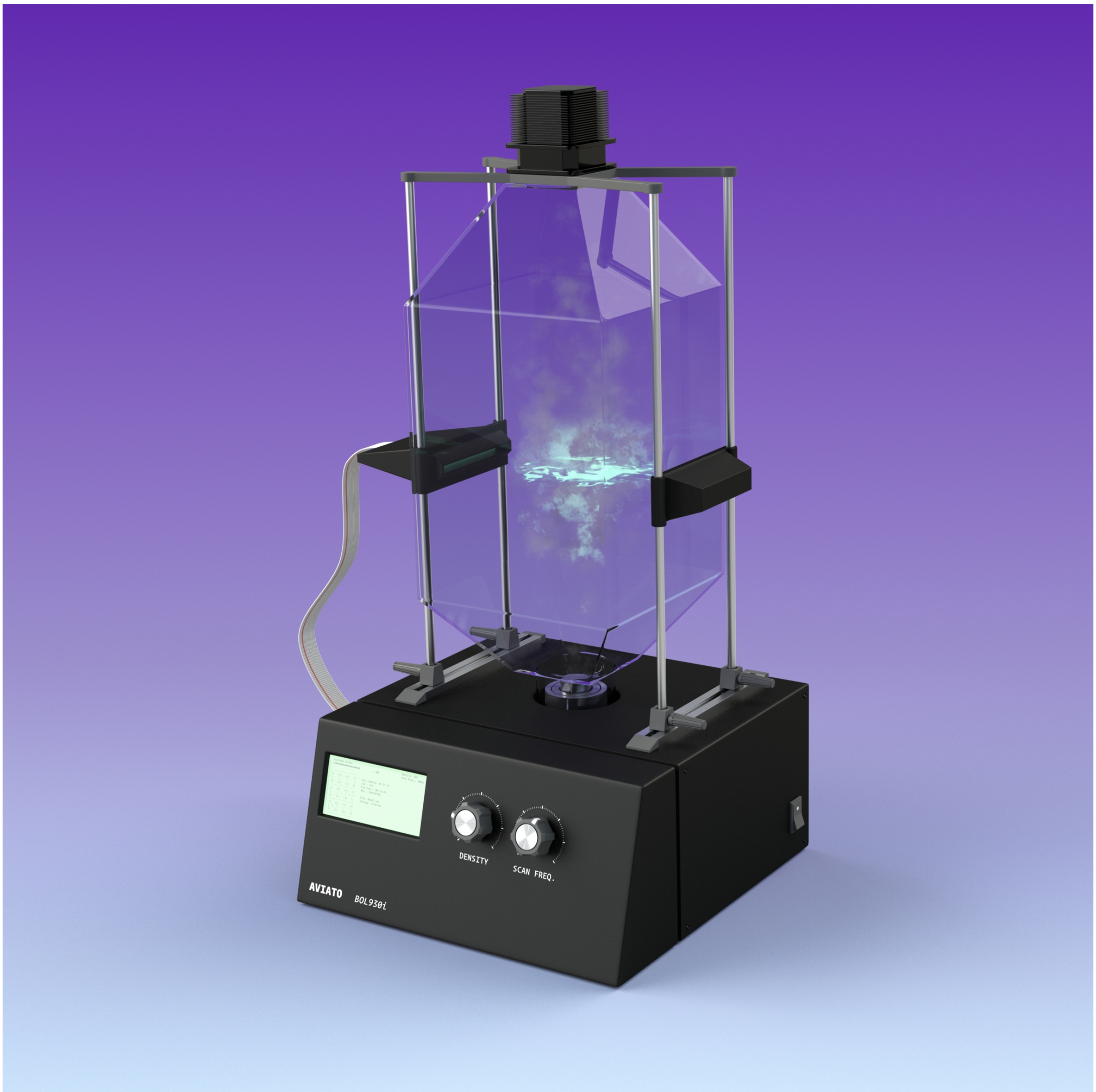


Figure G.1: Boltzmann Scanner - *BOL930i*
Features:
- Hardware based vortex lattice generator
- Save up to 12 Gflops at execution (Linpack benchmark)
- Fully replicable with multi-flow point-cloud logging

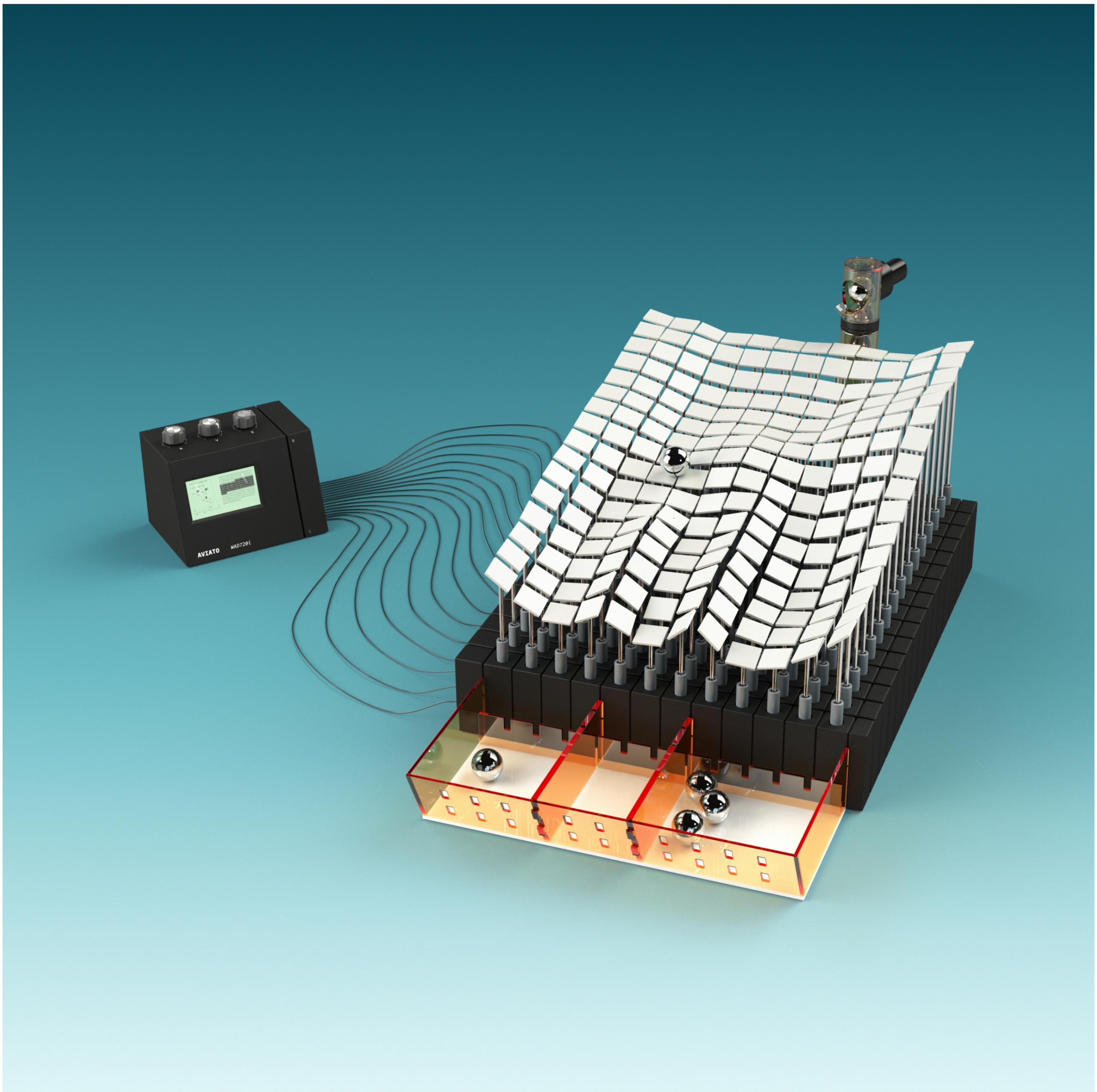


Figure G.2: Waddington Surface - *WAD720i*
Features:
- Input DNA sequences in: FASTQ, EMBL, FASTA or GenBank formats
- Modular design expands to your requirements
- Micro-hydraulics with 0.2 µm precision
- Combine with our Visio system for automated batch testing

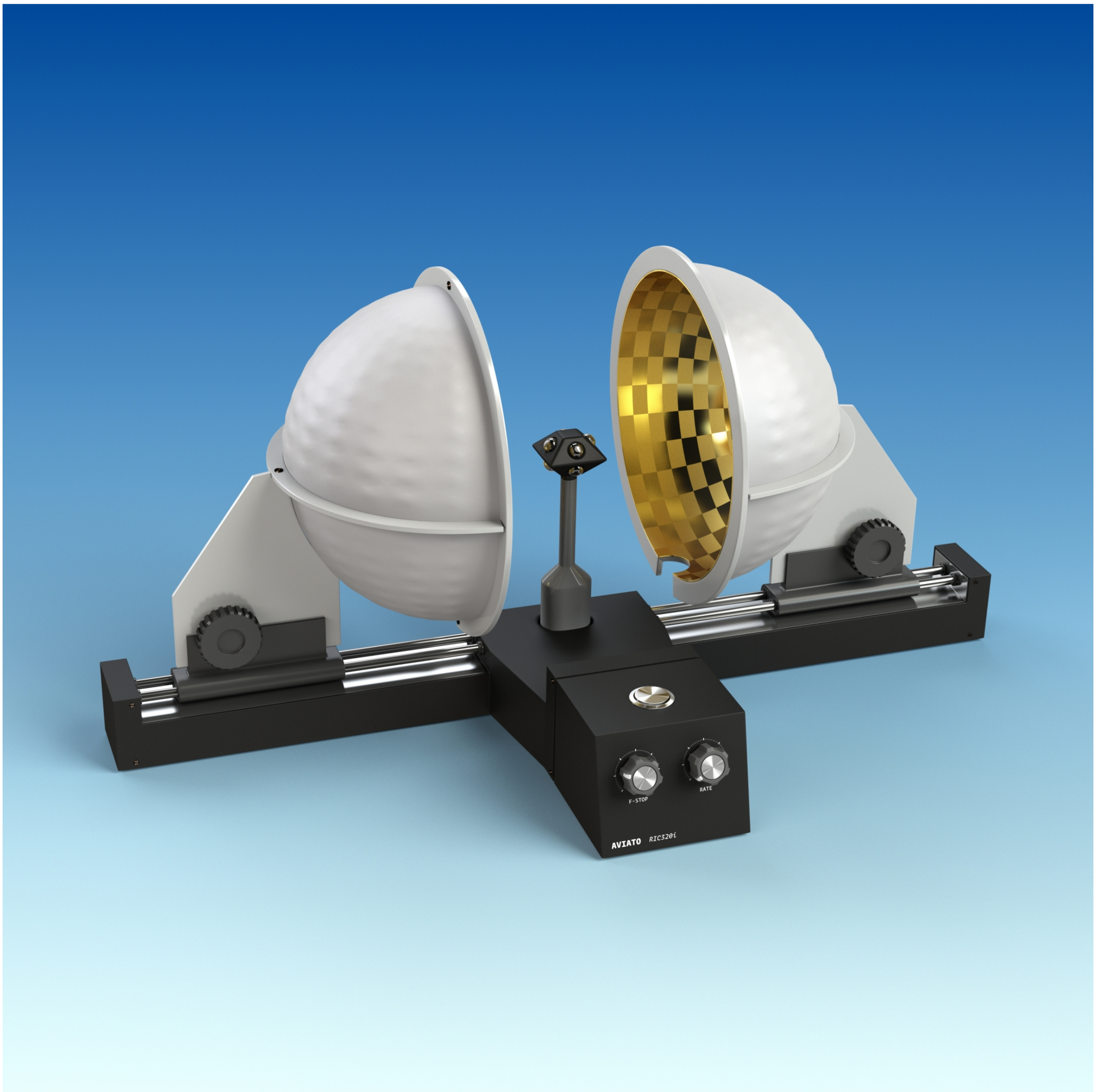


Figure G.3: Richardson Observatory - *RIC320i*
Features:
- Aggregate sources (IMMA, netCDF, and more) with unified granularity
- Interact directly with composite atmosphere meta-models
- Full 4K-360 resolution

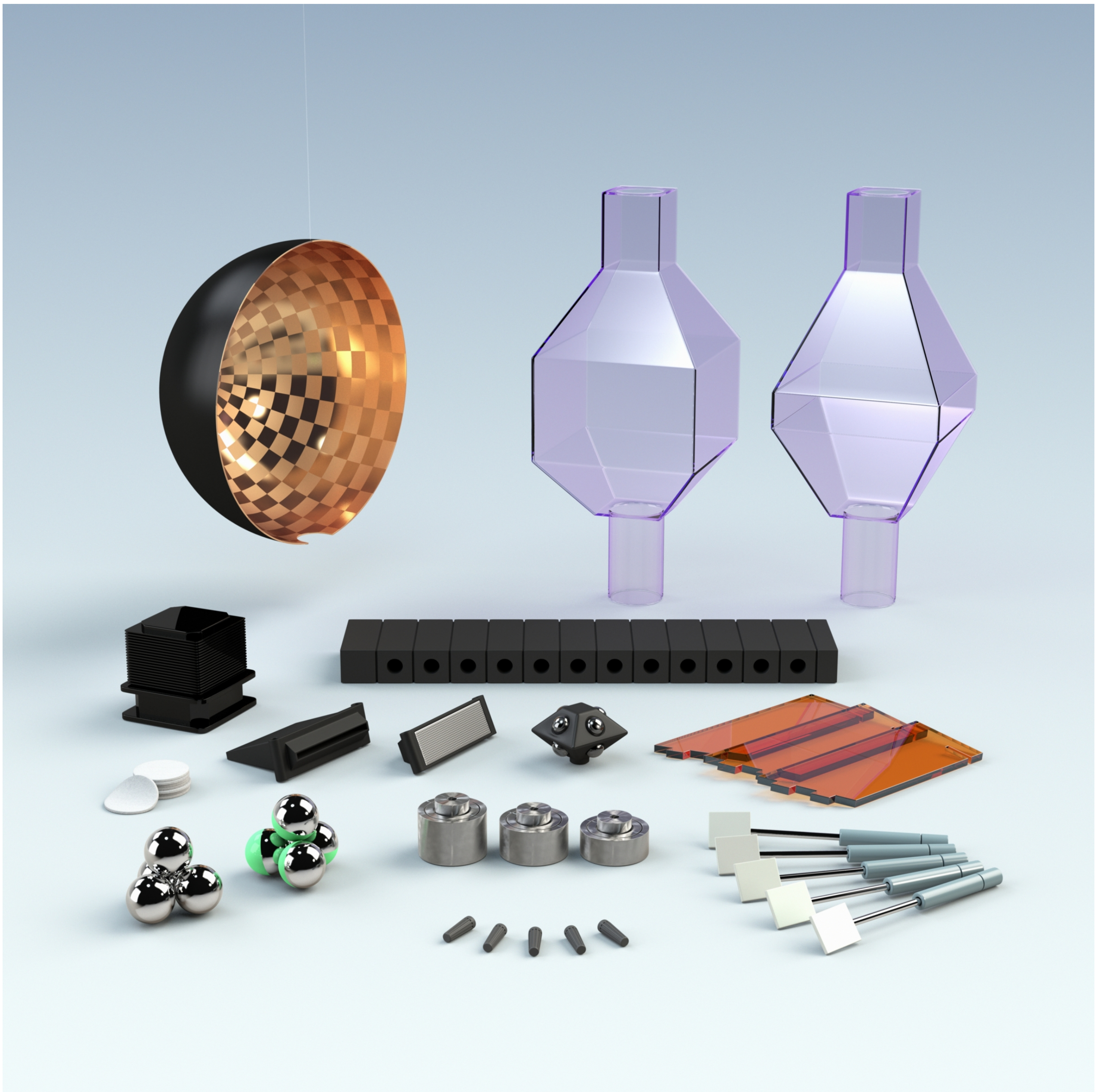


Figure G.4: Accessories

Appendix H

List of conferences, exhibitions, and events attended

05/11/2015: PhD by Design conference
Short presentation, contribution to *PhD by Design - Instant Journal #2 - Researching across difference* (Thomson et al., 2015).
Goldsmiths University of London, UK

07/2016: Microsoft Research PhD Summer School
Poster [fig.C.2] and presentation on the very early stages of *Diagrams of the Future.*
Microsoft Research Cambridge, UK

08/2016: 4S/EASST Science and Technology by Other Means
Presentation: Counting the Future (early version of *Diagrams of the Future*), see Benqué (2016, and fig.C.3).
Part of the closed session *Counting by Other Means*, convened by Alex Taylor (supervisor, Microsoft Research), and Sarah Kember (Goldsmiths) (see Taylor and Kember, 2016).
Barcelona, Spain.

11/2016: Datalmanach
Production of an almanac publication during a 4 day workshop "sprint" with students from a range of different courses.
S47 Research week
École supérieure d'Art et de Design, St. Etienne, France.

02-03/2017: Designing Decisions—Live Brief
5 week project for students of Goldsmiths BA design as part of their annual partnership with industry; in collaboration with Michael Golombewski and Helene Steiner from Microsoft Research Cambridge.

05/2017: Fiber Festival, Prima Materia; alchemical thinking and making in art, design and music.
Presentation: *The Monistic Almanac.*
Exhibition: *The Presage Range* (apendix G).
Looiersgracht 60, Amsterdam, the Netherlands.

07/2017: Data Fictions
1 day data visualisation workshop as part of the Critical and Speculative Design Summer School.
Convened by Ben Stopher and Tobias Revell.
London College of Communication, University of the Arts London (LCC, UAL).

16/09 to 20/10/2017: Uncertainty Playground
Exhibition: *Possibility Space* (appendix F).
London Design Festival 2017.
LCC, UAL, London, UK

24/09/2017: Very Very Far Away
Podcast: *Predicting the Future from the stars* interview with Dr. Roberto Trotta, reader in Astrophysics at Imperial College London.
`https://theairpump.davidbenque.com/predicting-the-future/`
Convened by Sitraka Rakotoniaina and Andrew Friend.
Digital Design Weekend at the Victoria & Albert Museum, London, UK

18/10/2017: Futurs Pluriels
Presentation: *Possibility Space* (appendix F).
Sciences Po, Paris, France.

03/11/2017: Kikk Festival, Invisible Narratives
Presentation: *The Monistic Almanac.*
Namur, Belgium.

10/11/2017: Anticipation 2017 Conference
Presentation: *The Monistic Almanac.*
Senate House, London, UK.

01-02/2018: (Re)Distributed Media: Leakage
Seeing-[:like]->a Diagram workshop for MA Graphic Media Design students. (Benqué, 2018c).
Convened by Paul Bailey.
LCC UAL, London, UK

10/02/2017: Information Experience Design seminar
Presentation: *Counting the Future; Unpacking prediction's "historical burden" through epic meta-diagramming.*
Royal College of Art, London, UK.

15/03/2018: Colossal Dust: Practices of Obsession and Investigation
Presentation: *The Flower and the Future (and the obsessive practice of investigation)* (Benqué, 2018b).
Curated by Marion Lagedamont and Rosie Allen.
This Happened London, UK.

14/04/2018: Navigating Finance and the Imagination
Presentation: *The Monistic Almanac*.
Collaborative, exploratory, theoretical walking tour of the City of London.
Convened by Dr. Aris Komporozos-Athanasiou and Dr. Max Haiven.

03/05/2018: Art, Materiality and Representation Conference
Presentation: *Speculative diagrams: plotting to reclaim algorithmic prediction* (Marenko and Benque, 2018).
Organised by the Royal Anthropological Institute at the British Museum in London, UK.

15/09 to 17/10/2018: Everything Happens So Much
Exhibtion: *The Monistic Almanac* printing station & *Architectures of Choice Vol.1: YouTube* (first experiments).
London Design Festival 2018.
Curated by Georgina Voss (Supra Systems Studio).
LCC UAL, London, UK.

30/10/2018 Fields of Communication: Nature · Culture · Technology
Presentation: *future.shape()* (PhD work in progress)
School of Communication, Royal College of Art, London, UK.

22-23/10/2018: Supra Systems Office Rites
Exhibition: *The Monistic Almanac* electional astrology personalised calendar printing station.
Curated by Georgina Voss (Supra Systems Studio).
Digital Design Weekend at the Victoria & Albert Museum, London, UK.

19-22/03/2019 Method & Critique; frictions and shifts in RTD - Research Through Design Conference
Presentation: *Speculative diagrams: Experiments in mapping YouTube* (Marenko and Benqué, 2019).
Exhibition: *Architectures of Choice Vol.1: YouTube* (traces).
TU Delft, the Netherlands

19/03/2019 Guest Lecture
Presentation: *Some Diagrams*.
Experimental Publishing Master (XPUB)
Piet Zwart Institute, Rotterdam, the Netherlands

23/05/2019: MiXiT Conference
Presentation: *Hyperland, a scam in n dimensions*
Lyon, France.

23-27/09/2019: Atlas des Recommendations
Workshop, *Architectures of Choice Vol. 2.*
École Nationale Supérieure de Création Industrielle (ENSCI)
Paris, France.

18/10/2019: Sonder les dispositifs numériques; Pratiques archéologiques en art et en design.
Presentation: *Diagrams of the Future.*
Study day convened by Vincent Ciciliato, Julie Martin, Anthony Masure, and Carole Nosella.
LLA CRÉATIS, Université Toulouse Jean Jaurès, France.

02/11/2019: IMPAKT Festival: Speculative Interfaces.
Presentation: *Architectures of Choice Vol. 1 YouTube*
Curated by Marloes de Valk.
Utrecht, the Netherlands.

Bibliography

Adams, Vincanne, Murphy, Michelle, and Clarke, Adele E (2009), ‘Anticipation: Technoscience, Life, Affect, Temporality’, *Subjectivity*, (28):pp. 246–265, doi:10.1057/sub.2009.18

Adorno, Theodor (1994), *The Stars Down to Earth*, Routledge, London, UK and New York, NY, USA

Agamben, Giorgio (2009), *What Is an Apparatus? : And Other Essays*, Stanford University Press, Stanford, CA, USA

Agrawal, Ajay, Gans, Joshua, and Goldfarb, Avi (2018), *Prediction Machines: The Simple Economics of Artificial Intelligence*, Harvard Business Review Press, Boston, MA, USA

Albright, Jonathan (2016), ‘What’s Missing From The Trump Election Equation? Let’s Start With Military-Grade PsyOps’, https://medium.com/@d1gi/whats-missing-from-the-trump-election-equation-let-s-start-with-military-grade-psyops-fa22090c8c17 (Accessed: 22/10/2019)

Albright, Jonathan (2017), ‘FakeTube: AI-Generated News on YouTube’, https://medium.com/@d1gi/faketube-ai-generated-news-on-youtube-233ad46849f9 (Accessed: 03/10/2018)

Amatrian, Xavier and Basilico, Justin (2012), ‘Netflix Recommendations: Beyond the 5 Stars (Part 1)’, https://medium.com/netflix-techblog/netflix-recommendations-beyond-the-5-stars-part-1-55838468f429 (Accessed: 22/10/2019)

Ananny, Mike and Crawford, Kate (2016), ‘Seeing without Knowing: Limitations of the Transparency Ideal and Its Application to Algorithmic Accountability’, *New Media & Society*, pp. 1–17, doi:10.1177/1461444816676645

Anderson, Chris (2008), ‘The End of Theory: The Data Deluge Makes the Scientific Method Obsolete.’, *Wired Magazine*, (16.07), http://www.wired.com/2008/06/pb-theory/

Anderson, Edgar (1936), ‘The Species Problem in Iris’, *Annals of the Missouri Botanical Garden*, 23(3):pp. 457–483, doi:10.2307/2394164

Andrejevic, Mark Bardeen (2018), 'Framelessness, or the Cultural Logic of Big Data', in *Mobile and Ubiquitous Media*, eds. Daubs, Michael and Manzerolle, Vincent, Digital Formations, pp. 251–267, Peter Lang Publishing, Germany

Angwin, Julia, Mattu, Surya, Larson, Jeff, and Kirchner, Lauren (2016), 'Machine Bias: There's Software Used across the Country to Predict Future Criminals. And It's Biased against Blacks.', *Propublica*, https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

Anthony, Boris (2017), 'Socio-Cultural Dynamics between the Eras of Commodity Economy: Pastorialism, Capitalism, Vectorialism', https://borisanthony.net/diagrams/vectorialism/

Aphek, Edna and Tobin, Yishai (1990), *The Semiotics of Fortune Telling*, John Benjamins Publishing Company, Amsterdam, the Netherlands

Aradau, Claudia and Blanke, Tobias (2017), 'Politics of Prediction: Security and the Time/Space of Governmentality in the Age of Big Data', *European Journal of Social Theory*, 20(3):pp. 373–391, doi:10.1177/1368431016667623

Archer, Michael D and Bickford, James Lauren (2007), *The Forex Chartist Companion*, A Visual Approach to Technical Analysis, John Wiley & Sons, Hoboken, NJ, USA, doi:10.1002/9781119197935

Arnold, Sarah (2016), 'Netflix and the Myth of Choice/Participation/Autonomy', in *The Netflix Effect: Technology and Entertainment in the 21st Century*, eds. McDonald, Kevin and Smith-Rowsey, Daniel, Bloomsbury Publishing USA

Ashworth, William J (1994), 'The Calculating Eye: Baily, Herschel, Babbage and the Business of Astronomy', *The British Journal for the History of Science*, 27(04):pp. 409–441, doi:10.1017/S0007087400032428

Auger, James (2012), *Why Robot?: Speculative Design, the Domestication of Technology and the Considered Future*, Ph.D. thesis, Royal College of Art

Auger, James (2014), 'Why Heart Attacks Could Be a Thing of the Past: Reading and Valuing Speculative Design', *Kunstlicht*, 35(3/4), http://tijdschriftkunstlicht.nl/35-speculari-why-heart-attacks-could-be-a-thing-of-the-past-james-auger/ (Accessed: 09/09/2019)

Auger, James and Loizeau, Jimmy (2015), 'Real Prediction Machines', http://www.auger-loizeau.com/projects/real-prediction-machines (Accessed: 09/09/2019)

Bardini, Thierry, Broye, Lionel, Citton, Yves, Galligo, Igor, Guez, Emmanuel, Guess, Jeff, Julien, Quentin, Krzywkowski, Isabelle, Lechner, Marie, Masure, Anthony, Pandelakis, Saul, Thibault, Ghislain, and Vargoz, Frédérique (2016), 'Manifeste Médiarchéologiste', http://wiki.pamal.org/wiki/Manifeste_M%C3%A9diarch%C3%A9ologiste (Accessed: 18/07/2019)

Barnes, T J (2013), 'Big Data, Little History', *Dialogues in Human Geography*, 3(3):pp. 297–302, doi:10.1177/2043820613514323

Barnes, T J and Wilson, M W (2014), 'Big Data, Social Physics, and Spatial Analysis: The Early Years', *Big Data & Society*, 1(1):pp. 1–14, doi:10.1177/2053951714535365

Bastian, Mathieu, Heymann, Sebastien, and Jacomy, Mathieu (2009), 'Gephi: An Open Source Software for Exploring and Manipulating Networks', in *International AAAI Conference on Weblogs and Social Media*, http://www.aaai.org/ocs/index.php/ICWSM/09/paper/view/154

Beck, F, Burch, M, Diehl, S, and Weiskopf, Daniel (2014), 'The State of the Art in Visualizing Dynamic Graphs', *Eurographics Conference on Visualization*, http://www.visus.uni-stuttgart.de/uploads/tx_vispublications/eurovis14-star.pdf

Beer, David (2019), *The Data Gaze : Capitalism, Power and Perception*, Sage, London, UK

Benqué, David (2016), 'Counting the Future', in *4S EASST 2016: Science & Technology by Other Means*, https://theairpump.davidbenque.com/counting-the-future/

Benqué, David (2018a), 'Cosmic Spreadsheets', in *Supra Systems #1*, ed. Voss, Georgina, London College of Communication, London, http://suprasystems.studio/downloads/book-chapters/Supra%20Systems%20Book_Chapter%2010_Benque.pdf

Benqué, David (2018b), 'The Flower and the Future', https://theairpump.davidbenque.com/the-flower-and-the-future/ (Accessed: 14/12/2018)

Benqué, David (2018c), 'SEEING-[:Like]->a DIAGRAM', https://theairpump.davidbenque.com/seeing-like-a-diagram/ (Accessed: 13/10/2018)

Benqué, David (forthcoming), 'Making an Almanac; Producing Predictions between Data Science and Astrology', in *Futures*, eds. Kemp, Sandra and Andersson, Jenny, Twenty First Century Approaches to Literature, Oxford University Press

Berman, Bob (2019), 'Solar Minimum Approaching — A Mini Ice Age?', https://www.almanac.com/news/astronomy/astronomy/solar-minimum-approaching-mini-ice-age (Accessed: 12/05/2019)

Bernhardsson, Erik (2013), 'Annoy; Approximate Nearest Neighbors in C++/Python Optimized for Memory Usage and Loading/Saving to Disk', Spotify, https://github.com/spotify/annoy (Accessed: 16/08/2019)

Bernhardsson, Erik (2015), 'Nearest Neighbor Methods and Vector Models – Part 1', https://erikbern.com/2015/09/24/nearest-neighbor-methods-vector-models-part-1.html (Accessed: 16/08/2019)

Berry, David (2011), 'The Computational Turn: Thinking about the Digital Humanities', *Culture Machine*, 12, http://sro.sussex.ac.uk/49813/ (Accessed: 19/02/2019)

Bertin, J (1967), *Semiologie Graphique*, Éditions de l'Ecole des Hautes Etudes en Sciences Sociales, Paris, France

Biddle, Sam (2016), 'Troubling Study Says Artificial Intelligence Can Predict Who Will Be Criminals Based on Facial Features', *The Intercept*, https://theintercept.com/2016/11/18/troubling-study-says-artificial-intelligence-can-predict-who-will-be-criminals-based-on-facial-features/ (Accessed: 04/09/2019)

Blanc, Julie and Haute, Lucile (2018), 'Publier La Recherche En Design : (Hors-)Normes, (Contre-)Formats, (Anti-)Standards', *Réel_Virtuel*, (6), http://www.reel-virtuel.com/numeros/numero6/sentinelles/publier-recherche-design#fn-note-15 (Accessed: 17/01/2019)

Blas, Zach (2013), 'Face Cages', http://www.zachblas.info/works/face-cages/ (Accessed: 25/01/2019)

Blegvad, Peter (), 'Imagined Observed Remembered', http://www.amateur.org.uk/imagined-observed-remembered/ (Accessed: 10/07/2019)

Boehnert, Joanna (2016), 'Data Visualisation Does Political Things', in *DRS2016: Design + Research + Society: Future-Focused Thinking.*, https://drs2016.squarespace.com/387

Boekmann, Catherine (2019), 'How We Predict the Weather', https://www.almanac.com/content/how-we-predict-weather (Accessed: 31/07/2019)

Bostock, Michael (2018), 'Scatterplot Matrix - Notebook', https://observablehq.com/@d3/scatterplot-matrix (Accessed: 17/08/2019)

Bostock, Michael, Ogievetsky, Vadim, and Heer, Jeffrey (2011), 'D3 Data-Driven Documents', *IEEE Transactions on Visualization and Computer Graphics*, 17(12):pp. 2301–2309, doi:10.1109/TVCG.2011.185

Bounegru, Liliana, Venturini, Tommaso, Gray, Jonathan, and Jacomy, Mathieu (2017), 'Narrating Networks. Exploring the Affordances of Networks as Storytelling Devices in Journalism', *Digital Journalism*, 5(6):pp. 699–730, doi:10.1080/21670811.2016.1186497

Box, George E. P. (1976), 'Science and Statistics', *Journal of the American Statistical Association*, 71(356):pp. 791–799, doi:10.1080/01621459.1976.10480949

boyd, danah and Crawford, Kate (2012), 'Critical Questions for Big Data', *Information, Communication & Society*, 15(5):pp. 662–679, doi:10.1080/1369118X.2012.678878

Bratton, Benjamin (2016a), 'On Speculative Design', *DIS Magazine*, http://dismagazine.com/blog/81971/on-speculative-design-benjamin-h-bratton/ (Accessed: 22/08/2018)

Bratton, Benjamin (2016b), *The Stack — On Software and Sovereignty*, MIT Press, Cambridge, MA, USA

Bree, P and Kessels, G (2013), 'Nodegoat: A Web-Based Data Management, Network Analysis & Visualisation Environment', LAB1100, http://nodegoat.net

Bridle, James (2017), 'Something Is Wrong on the Internet', https://medium.com/@jamesbridle/something-is-wrong-on-the-internet-c39c471271d2 (Accessed: 03/10/2018)

Broughton, P (1985), 'The First Predicted Return of Comet Halley', *Journal for the History of Astronomy*, http://articles.adsabs.harvard.edu/cgi-bin/nph-iarticle_query?1985JHA....16..123B&data_type=PDF_HIGH&whole_paper=YES&type=PRINTER&filetype=.pdf

Browne, Julian (2014), 'Graffeine', https://github.com/julianbrowne/graffeine (Accessed: 02/10/2019)

Bruckert, Romain (2018), 'Neo4j-Js-Ng2', https://github.com/adadgio/neo4j-js-ng2 (Accessed: 02/10/2019)

Bruno, Isabelle, Didier, Emmanuel, and Prévieux, Julien (2014), *Statactivisme: Comment Lutter Avec Des Nombres*, Zones Editions La Découverte, Paris, France

Brynjolfsson, Erik and McAfee, Andrew (2014), *The Second Machine Age : Work, Progress, and Prosperity in a Time of Brilliant Technologies*, W.W. Norton & Company, New York, NY, USA

Buckenham, George (2015), 'Cheap Bots, Done Quick!', https://cheapbotsdonequick.com/ (Accessed: 13/09/2019)

Burdick, Anne, Drucker, Johanna, Lunefeld, Peter, Presner, Todd, and Schnapp, Jefrey (2012), *Digital Humanities*, MIT Press, Cambridge, MA, USA and London, UK

Bureau des longitudes (1875), *Annuaire Pour l'an 1875*, Gauthier-Villars, Paris, France, https://gallica.bnf.fr/ark:/12148/bpt6k65393410 (Accessed: 03/01/2020)

Bureau d'Etudes (2004), 'World Government', https://bureaudetudes.org/2004/01/19/gouvernement-mondial-bureau-d%e2%80%99etudes-2004/ (Accessed: 29/07/2019)

Burgess, Matt (2016), 'This AI Learned to Predict the Future by Watching Loads of TV', *Wired UK*, https://www.wired.co.uk/article/algorithm-watch-tv-predict-future (Accessed: 08/07/2019)

Burrell, Jenna (2016), 'How the Machine 'Thinks': Understanding Opacity in Machine Learning Algorithms', *Big Data & Society*, 3(1), doi:10.1177/2053951715622512

Burrington, Ingrid (2014), 'Star Charts for Five Eyes', http://lifewinning.com/projects/star-charts-five-eyes/ (Accessed: 03/08/2018)

Cadwalladr, Carole and Graham-Harrison, Emma (2018), 'Revealed: 50 Million Facebook Profiles Harvested for Cambridge Analytica in Major Data Breach', *The Guardian*, https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election (Accessed: 22/10/2019)

Cairo, Alberto (2016), *The Truthful Art : Data, Charts, and Maps for Communication*, New Riders, San Francisco, CA, USA

Calheiros, Cécilia (2012), 'La fabrique d'une prophétie eschatologique par la cybernétique : le cas du Projet WebBot', *Raisons politiques*, 48(4):p. 51, doi:10.3917/rai.048.0051

Candy, S (2010), *The Futures of Everyday Life: Politics and the Design of Experiential Scenarios*, Ph.D. thesis

Carney, Scott (1999), *The Harmonic Trader*, HarmonicTrader.com, Tucson, AZ, USA

Cascone, Kim (2013), 'Errormancy', https://www.paradoxethereal-magazine.com/kim-cascone-errormancy/ (Accessed: 18/07/2019)

Chalmers, Matthew and Maccoll, Ian (2003), 'Seamful and Seamless Design in Ubiquitous Computing', Technical report, In Proceedings of Workshop At the Crossroads: The Interaction of HCI and Systems Issues in UbiComp. 2003

Chaslot, Guillaume (2016), 'Youtube-Explore', https://github.com/pnbt/youtube-explore (Accessed: 17/10/2018)

Chaslot, Guillaume, Montcoudiol, Adrien, Ledésert, Soline, Wielonsky, Nicolas, Bardolle, Frédéric, and Grac, Mathieu (2016a), 'AlgoTransparency', http://algotransparency.org/ (Accessed: 03/10/2018)

Chaslot, Guillaume, Montcoudiol, Adrien, Ledésert, Soline, Wielonsky, Nicolas, Bardolle, Frédéric, and Grac, Mathieu (2016b), 'Youtube Map', https://algotransparency.org/youtube.html (Accessed: 05/06/2019)

Cheney-Lippold, J (2011), 'A New Algorithmic Identity: Soft Biopolitics and the Modulation of Control', *Theory, Culture & Society*, 28(6):pp. 164–181, doi:10.1177/0263276411424420

Chernoff, Herman (1973), 'The Use of Faces to Represent Points in K-Dimensional Space Graphically', *Journal of the American Statistical Association*, 68(342):pp. 361–368, doi:10.2307/2284077

Choi, Hyunyoung and Varian, Hal (2012), 'Predicting the Present with Google Trends', *Economic Record*, 88(s1):pp. 2–9, doi:10.1111/j.1475-4932.2012.00809.x

Chollet, François (2018), *Deep Learning with Python*, Manning Publications Co., Shelter Island, NY, USA

Chrisman, Nicholas (2006), *Charting the Unknown : How Computer Mapping at Harvard Became GIS*, ESRI Press Independent Publishers Group (IPG) distributor, Redlands, CA, USA

Chung, Shing Tat (2012), 'The Superstitious Fund Project', http://shingtatchung.com/the-superstitious-fund-project/

Clifton, Brian, Lavigne, Sam, and Tseng, Francis (2017), 'White Collar Crime Risk Zones', *The New Inquiry*, 59: ABOLISH, https://whitecollar.thenewinquiry.com (Accessed: 21/08/2019)

Co – Star Astrology (), 'Co–Star: Hyper-Personalized, Real-Time Horoscopes', https://www.costarastrology.com/ (Accessed: 18/07/2019)

Compton, Kate, Kybartas, Ben, and Mateas, Michael (2015), 'Tracery: An Author-Focused Generative Text Tool', in *Interactive Storytelling*, eds. Schoenau-Fog, Henrik, Bruni, Luis Emilio, Louchart, Sandy, and Baceviciute, Sarune, Lecture Notes in Computer Science, pp. 154–161, Springer International Publishing

Cornelius, Geoffrey (2016), 'Chicane: Double-Thinking and Divination among the Witch-Doctors', in *Divination; Perspectives for a New Millennium*, ed. Curry, Patrick, Routledge

Corner, James (1999), 'The Agency of Mapping: Speculation, Critique and Invention', in *Mappings*, pp. 231–252, Reaktion Books

Covington, Paul, Adams, Jay, and Sargin, Emre (2016), 'Deep Neural Networks for YouTube Recommendations', in *Proceedings of the 10th ACM Conference on Recommender Systems*, RecSys '16, pp. 191–198, ACM, Boston, MA, USA, doi:10.1145/2959100.2959190

Cramer, Florian (2005), *Words Made Flesh*, Piet Zwart Institute, Rotterdam, the Netherlands, http://cramer.pleintekst.nl/00-recent/words_made_flesh/html/words_made_flesh.html

Crawford, Kate, Miltner, Kate, and Gray, Mary L (2014), 'Critiquing Big Data: Politics, Ethics, Epistemology', *International Journal of Communication*, 8:pp. 1663–1672, http://ijoc.org/index.php/ijoc/article/view/2167/1164

Cronon, William (1991), 'Pricing the Future: Grain', in *Nature's Metropolis*, W.W. Norton & Company, New York, London

Cubitt, Sean (2015), 'Data Visualization and the Subject of Political Aesthetics', in *Postdigital Aesthetics: Art, Computation and Design*, eds. Berry, David M. and Dieter, Michael, pp. 179–190, Palgrave Macmillan UK, London, doi:10.1057/9781137437204_14

Curry, Patrick, ed. (2016), *Divination; Perspectives for a New Millennium*, Routledge, London, UK

Dalton, Craig and Thatcher, Jim (2014), 'What Does a Critical Data Studies Look like, and Why Do We Care?', *Society & Space*, (12), http://societyandspace.org/2014/05/12/what-does-a-critical-data-studies-look-like-and-why-do-we-care-craig-dalton-and-jim-thatcher/ (Accessed: 20/08/2019)

Daston, Lorraine and Galison, Peter (2007), *Objectivity*, Zone Books, New York, NY, USA

Dat Foundation (), 'Dat.Foundation - A Community-Driven Web Protocol', dat.foundation (Accessed: 25/10/2019)

David McCandless (2012), *Information Is Beautiful*, Collins, London, UK

Davis, Erik (1998), *TechGnosis : Myth, Magic + Mysticism in the Age of Information*, Harmony Books, New York, NY, USA

Debaise, Didier and Stengers, Isabelle, eds. (2015), *Gestes spéculatifs*, Les Presses du réel, Dijon, France

Debatty, Regine (2016), 'Predictive Art Bot. A Call for Artworks That Interpret AI-Generated Concepts', https://we-make-money-not-art.com/predictive-art-bot-a-call-for-artworks-that-interpret-ai-generated-concepts/ (Accessed: 19/08/2019)

Deleuze, Gilles (1988), *Foucault*, University of Minnesota Press, Minneapolis, MN, USA

Deleuze, Gilles (1992), 'Postscript on the Societies of Control', *October*, 59:pp. 3–7, https://www.jstor.org/stable/778828 (Accessed: 23/11/2018)

Deleuze, Gilles and Guattari, Félix (1987), *A thousand plateaus: capitalism and schizophrenia*, University of Minnesota Press, Minneapolis, MN, USA

Deleuze, Gilles and Guattari, Félix (1991), *Qu'est-Ce Que La Philosophie*, Éditions de Minuit, Paris, France

Desrosières, Alain. (1998), *The politics of large numbers : a history of statistical reasoning*, Harvard University Press, Cambridge, MA, USA

Diakopoulos, Nicholas (2015), 'Algorithmic Accountability', *Digital Journalism*, 3(3):pp. 398–415, doi:10.1080/21670811.2014.976411

Didier, Emmanuel and Tasset, Cyprien (2013), 'Pour un statactivisme. La quantification comme instrument d'ouverture du possible', *Tracés*, (24):pp. 123–140, doi:10.4000/traces.5660

Digital Methods Initiative (2019), 'Post-API Research? On the Contemporary Study of Social Media Data', https://wiki.digitalmethods.net/Dmi/WinterSchool2020 (Accessed: 10/10/2019)

D'Ignazio, Catherine and Klein, Lauren F (2016), 'Feminist Data Visualization', *Workshop on Visualization for the Digital Humanities - IEEE VIS Conference*, http://www.kanarinka.com/wp-content/uploads/2015/07/IEEE_Feminist_Data_Visualization.pdf

Disnovation (2018), 'Predictive Art Bot', http://www.disnovation.org/pab.php (Accessed: 06/09/2019)

Domingos, Pedro (2015), *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World*, Basic Books, New York, NY, USA

Drucker, Johanna (2009), *SpecLab; Digital Aesthetics and Projects in Speculative Computing*, University of Chicago Press, Chicago, IL, USA

Drucker, Johanna (2011), ‘Humanities Approaches to Graphical Display’, *DHQ Digital Humanities Quarterly*, 5(1), http://www.digitalhumanities.org/dhq/vol/5/1/000091/000091.html

Drucker, Johanna (2014a), ‘Diagrammatic Form and Performative Materiality’, https://www.youtube.com/watch?v=VRgkkF97XCA&feature=youtu.be (Accessed: 01/10/2018)

Drucker, Johanna (2014b), *Graphesis: Visual Forms of Knowledge Production.*, Harvard University Press, Cambridge, MA, USA

Drulhe, Louise (2015), ‘Critical Atlas of Internet’, http://internet-atlas.net/ (Accessed: 08/10/2018)

Du Gay, Paul and Pryke, Michael (2002), *Cultural Economy: Cultural Analysis and Commercial Life*, Sage, London, UK

Dunne, Anthony (2005), *Hertzian Tales : Electronic Products, Aesthetic Experience, and Critical Design*, MIT Press, Cambridge, MA, USA

Dunne, Anthony and Raby, Fiona (2013), *Speculative Everything: Design, Fiction, and Social Dreaming*, MIT Press, Cambridge, MA, USA

Edwards, Paul N (2010), *A Vast Machine*, MIT Press, Cambridge, MA, USA

Eklund, Aron (2016), ‘Beeswarm: An R Package’, http://www.cbs.dtu.dk/~eklund/beeswarm/ (Accessed: 05/07/2019)

Elish, M. C. and boyd, danah (2018), ‘Situating Methods in the Magic of Big Data and AI’, *Communication Monographs*, 85(1):pp. 57–80, doi:10.1080/03637751.2017.1375130

Endor (2018), ‘AI-Powered Business Predictions’, https://www.endor.com/ (Accessed: 29/04/2019)

Eric Steven Raymond (2000), ‘The Cathedral and the Bazaar’, http://catb.org/esr/writings/cathedral-bazaar/cathedral-bazaar/index.html (Accessed: 08/10/2018)

Eriksson, Maria, Fleischer, Rasmus, Johansson, Anna, Snickars, Pelle, and Vonderau, Patrick (2019), *Spotify Teardown: Inside the Black Box of Streaming Music*, MIT Press, Cambridge, MA, USA

Ernst, Wolfgang (2011), ‘Media Archaeography; Method and Machine versus History and Narrative of Media’, in *Media Archaeology: Approaches, Applications, and Implications*, eds. Huhtamo, Erkki and Parikka, Jussi, University of California Press, Berkeley, CA, USA

Eubanks, Virginia (2019), *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*, Picador, New York, NY, USA

Evans, Jon (2015), 'When Facebook Knows You Better Than You Know Yourself', http://social.techcrunch.com/2015/10/24/when-facebook-knows-you-better-than-you-know-yourself/ (Accessed: 14/08/2019)

Farmers' Almanac (), 'Daily Calendar & Planner', https://farmersalmanac.com/calendar/ (Accessed: 17/07/2019)

Few, Stephen (2006), *Information Dashboard Design*, The Effective Visual Communication of Data, O'Reilly Media, Inc., Sebastopol, CA,USA

Fidora, Alexander and Sierra, Carles, eds. (2011), *Ramon Llull: From the Ars Magna to Artificial Intelligence*, Artificial Intelligence Research Institute, Barcelona, Spain

Fiennes, Sophie (2006), 'The Pervert's Guide to Cinema',

Fisher, R A (1936), 'The Use of Multiple Measurements in Taxonomic Problems', *Annals of Eugenics*, 7(2):pp. 179–188, doi:10.1111/j.1469-1809.1936.tb02137.x

Fisher, R.A. (1988), 'Iris Data Set', https://archive.ics.uci.edu/ml/datasets/iris (Accessed: 23/01/2019)

Fisherkeller, Mary Anne, Friedman, Jerome H, and Tukey, John W (1988), 'PRIM-9: An Interactive Multidimensional Data Display and Analysis System', *Dynamic Graphics for Statistics*, pp. 91–109, http://inspirehep.net/record/91655/files/slac-pub-1408.pdf

Folkner, William M, Williams, James G, Boggs, Dale H, and Kuchynka, Petr (2014), 'The Planetary and Lunar Ephemerides DE430 and DE431', *IPN Progress Report*, 42(196), ftp://naif.jpl.nasa.gov/pub/naif/generic_kernels/spk/planets/de430_and_de431.pdf

Forbes, Eric G. (1971), 'Gauss and the Discovery of Ceres', *Journal for the History of Astronomy*, 2(3):pp. 195–199, doi:10.1177/002182867100200305

Foucault, Michel (1972), *The Archaeology of Knowledge ; and, the Discourse on Language*, Pantheon Books, New York, NY, USA

Frayling, Christopher (1993), 'Research in Art and Design', *Royal College of Art Research Papers*, 1(1)

Frazee, Paul, Vancil, Tara, and Buus, Mathias (), 'Beaker Browser', https://beakerbrowser.com (Accessed: 25/10/2019)

Friendly, Michael (2006), 'A Brief History of Data Visualization', in *Handbook of Computational Statistics: Data Visualization*, eds. Chen, Chun-houh, Härdle, Wolfgang Karl, and Unwin, Antony, Springer-Verlag, Heidelberg

Friendly, Michael (2008), 'The Golden Age of Statistical Graphics', *Statistical Science*, 23(4):pp. 502–535, doi:10.1214/08-STS268

Friendly, Michael and Denis, Daniel (2005), 'The Early Origins and Development of the Scatterplot', *Journal of the History of the Behavioral Sciences*, 41(2):pp. 103–130, doi:10.1002/jhbs.20078

Fuchs, Christian (2017), 'From Digital Positivism and Administrative Big Data Analytics towards Critical Digital and Social Media Research!', *European Journal of Communication*, 32(1):pp. 37–49, doi:10.1177/0267323116682804

Fuller, Matthew (2018), 'Software Studies Methods', in *The Routledge Companion to Media Studies and Digital Humanities*, ed. Sayers, Jentery, Routledge, New York, doi:10.4324/9781315730479

Gansterer, Nikolaus (2017), *Drawing a Hypothesis*, Figures of Thought, de Gruyter, Berlin, Germany

Gardner, Martin (1958), *Logic Machines and Diagrams*, McGraw-Hill, New York, NY, USA

Gauquelin, Michel (1967), *The Cosmic Clocks: From Astrology to a Modern Science*, H. Regnery Co., Chicago, IL, USA, https://archive.org/details/cosmicclocks00mich (Accessed: 26/10/2018)

Gaver, Bill, Dunne, Tony, and Pacenti, Elena (1999), 'Design: Cultural Probes', *Interactions*, 6(1):pp. 21–29, doi:10.1145/291224.291235

Geuter, Jürgen (2017), 'Transcending the Algorithm-Data-Dualism', in *Theorizing the Web*, https://tante.cc/2017/04/23/transcending-data-algorithm-dualism/ (Accessed: 25/07/2019)

Gieseler Greenbaum, Dorian (2016), 'Arrows, Aiming and Divination: Astrology as a Stochastic Art', in *Divination; Perspectives for a New Millennium*, ed. Curry, Patrick, Routledge

Gigerenzer, Gerd, Porter, Theodore, Swijtink, Zeno, Daston, Lorraine, Beatty, John, and Kruger, Lorenz (1990), *The Empire of Chance: How Probability Changed Science and Everyday Life*, Cambridge University Press, Cambridge, UK

Gilbert, Annette, ed. (2016), *Publishing as Artistic Practice*, Sternberg Press, Berlin, Germany

Ginsberg, Alexandra Daisy (2018), *Better: Navigating Imaginaries in Design and Synthetic Biology to Question 'Better'*, Thesis, Royal College of Art, http://researchonline.rca.ac.uk/3231/ (Accessed: 18/09/2019)

Ginzburg, Carlo (1980), 'Morelli, Freud and Sherlock Holmes: Clues and Scientific Method', *History Workshop*, 9:pp. 5–36, www.jstor.org/stable/4288283

Gitelman, Lisa (2013), *Raw Data Is an Oxymoron*, The MIT Press, Cambridge, MA, USA and London, UK

Gitelman, Lisa (2017), 'Review of Archaeology of Algorithmic Artefacts, by David Link', *Computational Culture*, (6), http://computationalculture.net/review-of-archaeology-of-algorithmic-artefacts-by-david-link/ (Accessed: 28/11/2018)

Goatley, Wesley (2019), *Critical Data Aesthetics: Towards a Critically Reflexive Practice of Data Aestheticisation*, Ph.D. thesis, University of Sussex, http://sro.sussex.ac.uk/id/eprint/84681/

Goatley, Wesley and Revell, Tobias (2018), 'Augury', http://www.wesleygoatley.com/augury/ (Accessed: 05/08/2019)

Goodfellow, Ian, Bengio, Yoshua, and Courville, Aaron (2018), '2.1 Scalars Vectors Matrices and Tensors', https://hadrienj.github.io/posts/Deep-Learning-Book-Series-2.1-Scalars-Vectors-Matrices-and-Tensors/ (Accessed: 14/08/2019)

Google (2012), 'Google Books Ngram Viewer', http://storage.googleapis.com/books/ngrams/books/datasetsv2.html (Accessed: 06/09/2019)

Graphileon (), 'Graphileon', https://graphileon.com/ (Accessed: 02/10/2019)

Gray, Jonathan (2016), '"Let Us Calculate!": Leibniz, Llull, and the Computational Imagination', *The Public Domain Review*, https://publicdomainreview.org/2016/11/10/let-us-calculate-leibniz-llull-and-computational-imagination/

Grier, David (2007), *When Computers Were Human*, Princeton University Press, Princeton, NJ, USA

Grus, Joel (2015), *Data Science from Scratch: First Principles with Python*, O'Reilly Media, Sebastopol, CA,USA

Guy, Amy (2017), *The Presentation of Self on a Decentralised Web*, PhD Thesis, University of Edinburgh, http://dr.amy.gy/ (Accessed: 14/02/2019)

Hacking, Ian (1990), *The Taming of Chance*, Cambridge University Press, Cambridge, UK

Hagberg, Aric A., Schult, Daniel A., and Swart, Pieter J. (2008), 'Exploring Network Structure, Dynamics, and Function Using NetworkX', in *Proceedings of the 7th Python in Science Conference*, eds. Varoquaux, Gaël, Vaught, Travis, and Millman, Jarrod, pp. 11–15

Hall, Peter (2008), 'Critical Visualization', in *Design and the Elastic Mind*, ed. Antonelli, Paola, pp. 1–5, The Museum of Modern Art, New York, http://peterahall.com/pdfs/Hall_critviz.pdf

Hancock, T. and Bezold, C. (1994), 'Possible Futures, Preferable Futures.', *The Healthcare Forum journal*, 37(2):pp. 23–29

Hargittai, Eszter and Sandvig, Christian, eds. (2015), *Digital Research Confidential: The Secrets of Studying Behavior Online*, The MIT Press, Cambridge, MA, USA

Harris, Mark (2017), 'How Peter Thiel's Secretive Data Company Pushed Into Policing', *Wired Magazine*, https://www.wired.com/story/how-peter-thiels-secretive-data-company-pushed-into-policing/ (Accessed: 21/08/2019)

Hastie, Trevor, Tibshirani, Robert, and Friedman, Jerome (2001), *The Elements of Statistical Learning: Data Mining, Inference and Prediction*, Springer-Verlag, New York, NY, USA

Hayashi, Chikio (1996), 'What Is Data Science? Fundamental Concepts and a Heuristic Example', in *Proceedings of the Fifth Conference of the International Federation of Classification Societies*, pp. 40–51, Springer, Kobe, Japan

Hebron, Patrick (2016), *Machine Learning for Designers*, O'Reilly, Sebastopol, CA,USA

Huhtamo, Erkki (2011a), 'Dismantling the Fairy Engine: Media Archaeology as Topos Study', in *Media Archaeology: Approaches, Applications, and Implications*, eds. Huhtamo, Erkki and Parikka, Jussi, Univ of California Press

Huhtamo, Erkki (2011b), 'On the Archaeology of Imaginary Media', in *Media Archaeology: Approaches, Applications, and Implications*, eds. Parikka, Jussi and Kluitenberg, Eric, Univ of California Press

Huhtamo, Erkki and Parikka, Jussi, eds. (2011), *Media Archaeology: Approaches, Applications, and Implications*, University of California Press, Berkeley, CA, USA

Hunter, J. D. (2007), 'Matplotlib: A 2D Graphics Environment', *Computing in Science & Engineering*, 9(3):pp. 90–95, doi:10.1109/MCSE.2007.55

James, Robin (2015), 'Cloudy Logic', *The New Inquiry*, https://thenewinquiry.com/cloudy-logic/

James, Robin (2019), 'Financial Music', *American Book Review*, 40(6):pp. 13–14, doi:10.1353/abr.2019.0116

Jameson, Frederic (1988), 'Cognitive Mapping', in *Marxism and the Interpretation of Culture*, eds. Nelson, Cary and Grossberg, Lawrence, Macmillan Education, Basingstoke

Jones, Alistair (), 'Arrows', https://github.com/apcj/arrows (Accessed: 02/10/2019)

Jurgenson, Nathan (2014), 'View From Nowhere', *The New Inquiry*, https://thenewinquiry.com/view-from-nowhere/ (Accessed: 13/08/2019)

Katz, Anna (2019), 'Prime Member January Horoscope', https://www.amazon.com/primeinsider/tips/horoscope-jan19.html (Accessed: 24/04/2019)

Kazemi, Darius (), 'Tiny Subversions', http://tinysubversions.com/projects/ (Accessed: 06/09/2019)

Keeling, Charles D., Piper, Stephen C., Bacastow, Robert B., Wahlen, Martin, Whorf, Timothy P., Heimann, Martin, and Meijer, Harro A. (2001), 'Exchanges of Atmospheric CO2 and 13CO2 with the Terrestrial Biosphere and Oceans from 1978 to 2000. I. Global Aspects', https://escholarship.org/uc/item/09v319r9 (Accessed: 22/07/2019)

Kitchin, R (2014), 'Big Data, New Epistemologies and Paradigm Shifts', *Big Data & Society*, 1(1), doi:10.1177/2053951714528481

Kluyver, Thomas, Ragan-Kelley, Benjamin, Pérez, Fernando, Granger, Brian E, Bussonnier, Matthias, Frederic, Jonathan, Kelley, Kyle, Hamrick, Jessica B, Grout, Jason, Corlay, Sylvain, et al. (2016), 'Jupyter Notebooks-a Publishing Format for Reproducible Computational Workflows.', in *ELPUB*, pp. 87–90

Knoespel, Kenneth J (2001), 'Diagrams as Piloting Device in the Work of Gilles Deleuze', *Théorie - Littérature - Enseignement*, 19

Koch, Dieter and Treindl, Alois (2014), *Swiss Ephemeris; Computer Ephemeris for Developers of Astrological Software*, Astrodienst AG, Zurich, Switzerland, https://www.astro.com/swisseph/swisseph.htm (Accessed: 17/11/2019)

Kopf, Dan (2017), 'Meet the Man behind the Most Important Tool in Data Science', *Quartz*, https://qz.com/1126615/the-story-of-the-most-important-tool-in-data-science/ (Accessed: 10/08/2018)

LaFrance, Adrienne (2015), 'How The Old Farmer's Almanac Previewed the Information Age', *The Atlantic*, https://www.theatlantic.com/technology/archive/2015/11/how-the-old-farmers-almanac-previewed-the-information-age/415836/ (Accessed: 09/08/2018)

Landrum, Asheley (2019), 'YouTube as the Primary Propagator of Flat Earth Philosophy', in *Symposium at the American Association for the Advancement of Science (AAAS)*, Washington, D.C., http://www.asheleylandrum.com/uploads/4/7/8/3/47833717/landrum_2019_02.18_aaas_online.pdf

Landscheidt, Theodor (1973), *Cosmic Cybernetics; The Foundations of a Modern Astrology*, Ebertin-Verlag, Aalen, Germany

Landscheidt, Theodor (2003), 'New Little Ice Age Instead of Global Warming?', *Energy & Environment*, 14(2):pp. 327–350, http://www.schulphysik.de/klima/landscheidt/iceage.htm

Laplace, Pierre Simon (1814), *A Philosophical Essay on Probabilities*

Lapowsky, Issie (2015), 'How Facebook Knows You Better Than Your Friends Do', *Wired*, https://www.wired.com/2015/01/facebook-personality-test/ (Accessed: 14/08/2019)

Lavrenko, Victor (2015), 'IAML8.3 Examples of Overfitting and Underfitting', https://www.youtube.com/watch?v=dBLZg-RqoLg (Accessed: 17/08/2019)

LeCun, Yann, Cortes, Corinna, and Burges, Christopher JC (1998), 'The MNIST Database of Handwritten Digits', http://yann.lecun.com/exdb/mnist

Leeb, Susanne (2017), 'A Line with Variable Direction, Which Traces No Contour and Delimits No Form', in *Drawing a Hypothesis*, ed. Gansterer, Nikolaus, de Gruyter

Leese, Matthias (2016), "SEEING FUTURES': Politics of Visuality and Affect', in *Algorithmic Life: Calculative Devices in the Age of Big Data*, eds. Amoore, Louise and Piotukh, Volha, Routledge, London and New York

Levinovitz, Alan Jay (2016), 'How Economists Rode Maths to Become Our Era's Astrologers', *Aeon*, https://aeon.co/essays/how-economists-rode-maths-to-become-our-era-s-astrologers (Accessed: 09/08/2018)

Lewis, Paul (2018a), ''Fiction Is Outperforming Reality': How YouTube's Algorithm Distorts Truth', *The Guardian*, https://www.theguardian.com/technology/2018/feb/02/how-youtubes-algorithm-distorts-truth (Accessed: 17/10/2018)

Lewis, Paul and McCormick, Erin (2018), 'How an Ex-YouTube Insider Investigated Its Secret Algorithm', *The Guardian*, https://www.theguardian.com/technology/2018/feb/02/youtube-algorithm-election-clinton-trump-guillaume-chaslot (Accessed: 01/10/2018)

Lewis, Rebecca (2018b), 'Alternative Influence: Broadcasting the Reactionary Right on YouTube', Technical report, Data & Society, https://datasociety.net/output/alternative-influence/

Link, David (2010), 'Scrambling TRUTH: Rotating Letters as a Material Form of Thought', *Variantology*, 4:pp. 215–266, http://www.alpha60.de/research/scrambling_truth/DavidLink_ScramblingTruth2010_100dpi.pdf

Link, David (2016), *Archaeology of Algorithmic Artefacts*, Univocal, Minneapolis, MN, USA

Lohr, Steve (2015), *Data-Ism: Inside the Big Data Revolution*, Oneworld Publications, London, UK

Loukissas, Yanni (2019), *All Data Are Local : Thinking Critically in a Data-Driven Society*, The MIT Press, Cambridge, MA, USA

Macgregor, Jeff (1997), 'Moonrise On The Internet', *The New York Times*, https://www.nytimes.com/1997/08/22/opinion/moonrise-on-the-internet.html

Mackay, Robin (2017), 'Stages, Plots, and Traumas', in *Futures and Fictions*, Repeater Books, London, http://readthis.wtf/writing/stages-plots-and-traumas/ (Accessed: 08/08/2019)

Mackenzie, Adrian (2013), 'Programming Subjects in the Regime of Anticipation: Software Studies and Subjectivity', 6(4):pp. 391–405, doi:10.1057/sub.2013.12

Mackenzie, Adrian (2017), *Machine Learners : Archaeology of a Data Practice*, MIT Press, Cambridge, MA, USA

MacKenzie, Donald (1976), 'Eugenics in Britain', *Social Studies of Science*, 6(3/4):pp. 499–532, https://www.jstor.org/stable/284693 (Accessed: 25/04/2019)

Majaca, Antonia and Parisi, Luciana (2016), 'The Incomputable and Instrumental Possibility', *e-flux*, 77, http://www.e-flux.com/journal/77/76322/the-incomputable-and-instrumental-possibility/

Malone, Katie and Thrun, Sebastian (2015), 'Intro to Machine Learning', https://eu.udacity.com/course/intro-to-machine-learning--ud120 (Accessed: 30/01/2019)

Marenko, Betti (2015), 'When Making Becomes Divination: Uncertainty and Contingency in Computational Glitch-Events', *Design Studies*, 41:pp. 1–16, doi:10.1016/j.destud.2015.08.004

Marenko, Betti and Benque, David (2018), 'Speculative Diagrams: Plotting to Reclaim Algorithmic Prediction.', in *Art, Materiality and Representation*, Royal Anthropological Institute, British Museum, London, https://theairpump.davidbenque.com/speculative-diagrams-plotting-to-reclaim-algorithmic-prediction/

Marenko, Betti and Benqué, David (2019), 'Speculative Diagrams: Experiments in Mapping Youtube', in *Method & Critique; Frictions and Shifts in RTD*, TU Delft, doi:10.6084/m9.figshare.7855811.v1

Marenko, Betti and Brassett, Jamie (2015), *Deleuze and Design*, Edinburgh University Press, Edinburgh, UK

Masure, Anthony (2014), *Le design des programmes*, PhD Thesis, Université Paris1 Panthéon-Sorbonne, Paris, http://www.softphd.com/ (Accessed: 14/02/2019)

Masure, Anthony (2017), *Design et Humanités Numériques*, Esthétique Des Données, Éditions B42, Paris, France

Mattern, Shannon (2015), 'Mission Control: A History of the Urban Dashboard', *Places Journal*, doi:10.22269/150309

Maudet, Nolwenn (2017), *Designing Design Tools*, PhD Thesis, Université Paris-Saclay, https://designing-design-tools.nolwennmaudet.com// (Accessed: 14/02/2019)

McInerny, Greg J., Chen, Min, Freeman, Robin, Gavaghan, David, Meyer, Miriah, Rowland, Francis, Spiegelhalter, David J., Stefaner, Moritz, Tessarolo, Geizi, and Hortal, Joaquin (2014), 'Information Visualisation for Science and Policy: Engaging Users and Avoiding Bias', *Trends in Ecology & Evolution*, 29(3):pp. 148–157, doi:10.1016/j.tree.2014.01.003

McKinney, Wes (), 'Pandas: Python Data Analysis Library', https://pandas.pydata.org/ (Accessed: 22/07/2019)

McKinney, Wes (2010), 'Data Structures for Statistical Computing in Python', in *Proceedings of the 9th Python in Science Conference*, pp. 51–56, Austin TX

McMahan, H. Brendan, Holt, Gary, Sculley, D., Young, Michael, Ebner, Dietmar, Grady, Julian, Nie, Lan, Phillips, Todd, Davydov, Eugene, Golovin, Daniel, Chikkerur, Sharat, Liu, Dan, Wattenberg, Martin, Hrafnkelsson, Arnar Mar, Boulos, Tom, and Kubica, Jeremy (2013), 'Ad Click Prediction: A View from the Trenches', in *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*

McQuillan, Dan (2017), 'Data Science as Machinic Neoplatonism', *Philosophy and Technology*, 3(1):pp. 253–272, doi:10.1007/s13347-017-0273-3

Meyerson, Eric (2012), 'YouTube Now: Why We Focus on Watch Time', https://youtube-creators.googleblog.com/2012/08/youtube-now-why-we-focus-on-watch-time.html (Accessed: 15/12/2019)

Michael, Mike (2012), 'De-Signing the Object of Sociology: Toward an 'Idiotic' Methodology', *The Sociological Review*, 60(S1):pp. 166–183, doi:10.1111/j.1467-954X.2012.02122.x

Michael Guggenheim, Bernd Kräftner, and Judith Kröll (2017), 'Creating Idiotic Speculators; Disaster Cosmopolitics in the Sandbox', in *Speculative Research*, eds. Wilkie, Alex, Savransky, Martin, and Rosengarten, Marsha, pp. 145–162, Routledge, London and New York

Miller, Libby, Moore, Joanne, Howe, Tom, Cowlishaw, Tim, Onumonu, Anthony, Needham, Chris, Sedriks, Sacha, and Cooke, Henry (2017), 'Tellybox', https://www.bbc.co.uk/rd/projects/tellybox (Accessed: 07/06/2019)

Moatti, Alexandre (2017), 'Auguste Comte et l'institution scientifique : modalités et ressorts de son opposition et de ses critiques', *Bibnum. Textes fondateurs de la science*, http://journals.openedition.org/bibnum/1057 (Accessed: 05/01/2020)

Monjou, Marc (2014), 'La Recherche En Designet Sa Publicité', *Azimuts*, 40-41:pp. 53–62, https://revue-azimuts.fr/numeros/40-41/la-recherche-en-design-et-sa-publicite

Montgomery, Elliott P. and Woebken, Chris (2016), *Extrapolation Factory - Operator's Manual*, CreateSpace, New York, NY, USA

Morton, Timothy (2013), *Hyperobjects : Philosophy and Ecology after the End of the World*, University of Minnesota Press, Minneapolis, MN, USA

Mosco, Vincent (2014), *To the Cloud: Big Data in a Turbulent World*, Paradigm Publishers, Boulder, CO, USA - London, UK

Munster, Anna (2013), 'Networked Diagrammatism: From Map and Model to the Internet as Mechanogram', in *An Aesthesia of Networks: Conjunctive Experience in Art and Technology*, Technologies of Lived Abstraction, pp. 19–56, MIT Press, Cambridge, Mass

Murphy, John J (1999), *Technical Analysis of the Financial Markets*, A Comprehensive Guide to Trading Methods and Applications, New York Insitute of Finance, Paramus, NJ, USA

Nguyen, Christopher and Andreessen Horowitz (2015), 'A16z Podcast: Making Sense of Big Data, Machine Learning, and Deep Learning', https://a16z.com/2015/05/01/a16z-podcast-making-sense-of-big-data-machine-learning-and-deep-learning/ (Accessed: 10/09/2019)

Noble, Safiya Umoja (2018), *Algorithms of Oppression: How Search Engines Reinforce Racism*, NYU Press, New York, NY, USA

Nowak, Michael and Eckles, Dean (2017), 'Determining User Personality Characteristics from Social Networking System Communications and Characteristics', https://patents.google.com/patent/US9740752B2/en (Accessed: 22/04/2019)

Nowviskie, Bethany Paige (2004), *Speculative Computing: Instruments for Interpretive Scholarship*, Ph.D. thesis, University of Virginia, Charlottesville, USA, http://nowviskie.org/dissertation.pdf (Accessed: 22/04/2019)

NYS Division of Criminal Justice Services (2012), 'NYS COMPAS Probation Risk and Need Assessment Study', pp. 1–55, http://www.northpointeinc.com/downloads/research/DCJS_OPCA_COMPAS_Probation_Validity.pdf

Old Moore's Almanac (), 'Old Moore's Almanac Website', https://oldmooresalmanac.com/ (Accessed: 14/04/2018)

Old Moore's Almanack (2017), *Old Moore's Almanack 2017*, Foulsham Publishing, Marlow, UK

O'Neil, Cathy and Schutt, Rachel (2013), *Doing Data Science*, Straight Talk from the Frontline, O'Reilly Media, Inc., Sebastopol, CA,USA

Orion, Rae (2007), *Astrology for Dummies*, John Wiley & Sons, Hoboken, NJ, USA, second edition

Palmer, Oliver Matthew (2017), *Scripted Performances: Designing Performative Architectures through Digital and Absurd Machines*, Ph.D. thesis, Bartlett School of Architecture UCL, London

Paolillo, John C. (2018), 'The Flat Earth Phenomenon on YouTube', *First Monday*, 23(12), doi:10.5210/fm.v23i12.8251

Parikka, Jussi (2011), 'Operative Media Archaeology: Wolfgang Ernst's Materialist Media Diagrammatics', *Theory, Culture & Society*, 28(5):pp. 52–74, doi:10.1177/0263276411411496

Parikka, Jussi (2012), *What Is Media Archaeology*, Polity, Cambridge, UK - Malden, MA, USA

Parikka, Jussi (2019), 'Inventing Pasts and Futures; Speculative Design and Media Archaeology', in *New Media Archaeologies*, eds. Roberts, Ben and Goodall, Mark, pp. 205–235, Amsterdam University Press

Parisi, Lucianna (2012), 'Speculation: A Method for the Unattainable', in *Inventive Methods*, eds. Lury, Celia and Wakeford, Nina, pp. 232–244, Routledge

Parrish, Allison (2015), 'Exploring (Semantic) Space With (Literal) Robots', https://vimeo.com/134734729

Parrish, Allison (2017), 'Electronic Rituals, Oracles and Fortune-Telling: A Collective Bibliography', http://eroft.decontextualize.com

Parrish, Allison (2018), 'N-Grams and Markov Chains', https://github.com/aparrish/rwet (Accessed: 07/12/2019)

Pasquinelli, Matteo (2015a), 'Anomaly Detection: The Mathematization of the Abnormal in the Metadata Society', in *Transmediale*, https://www.academia.edu/10369819/Anomaly_Detection_The_Mathematization_of_the_Abnormal_in_the_Metadata_Society (Accessed: 12/12/2019)

Pasquinelli, Matteo (2015b), 'The Evolution of Machinic Intelligence (Map)', https://www.academia.edu/15275569/The_Evolution_of_Machinic_Intelligence_map_ (Accessed: 25/10/2018)

Pasquinelli, Matteo (2017), 'Machines That Morph Logic: Neural Networks and the Distorted Automation of Intelligence as Statistical Inference', *Glass Bead*, 1, http://www.glass-bead.org/article/machines-that-morph-logic/?lang=enview

Pasquinelli, Matteo (2019), 'Three Thousand Years of Algorithmic Rituals: The Emergence of AI from the Computation of Space', *e-flux*, 101, https://www.e-flux.com/journal/101/273221/three-thousand-years-of-algorithmic-rituals-the-emergence-of-ai-from-the-computation-of-space/ (Accessed: 24/06/2019)

Pedregosa, F, Varoquaux, G, Gramfort, A, Michel, V, Thirion, B, Grisel, O, Blondel, M, Prettenhofer, P, Weiss, R, Dubourg, V, Vanderplas, J, Passos, A, Cournapeau, D, Brucher, M, Perrot, M, and Duchesnay, E (2011), 'Scikit-Learn: Machine Learning in Python', *Journal of Machine Learning Research*, 12:pp. 2825–2830

Pentland, Alex (2014), *Social Physics: How Good Ideas Spread-the Lessons from a New Science*, The Penguin Press, New York, NY, USA

Pereira, Lídia (2019), 'Redirecting Responsibility: From Structural Insecurity to Systemic Depression', in *State Machines; Reflections and Actions at the Edge of Digital Citizenship, Finance, and Art*, eds. Colakides, Yianis, Garret, Marc, and Gloerich, Inte, Institute of Network Cultures, Amsterdam

Pérez, Fernando and Granger, Brian E. (2007), 'IPython: A System for Interactive Scientific Computing', *Computing in Science and Engineering*, 9(3):pp. 21–29, doi:10.1109/MCSE.2007.53

Peyton Jones, Simon (2016), 'How to Write a Great Research Paper', https://www.microsoft.com/en-us/research/academic-program/write-great-research-paper/ (Accessed: 02/07/2019)

Phan, Thao (2017), 'The Materiality of the Digital and the Gendered Voice of Siri.', *Transformations*, (29) Plummer-Fernandez, Matthew (2019), *The Art of Bots: A Practice-Based Study of the Multiplicity, Entanglements and Figuration of Sociocomputational Assemblages*, PhD Thesis, Goldsmiths, University of London, http://research.gold.ac.uk/26599/ (Accessed: 11/07/2019)

Porter, T M (1995), *Trust in Numbers: The Pursuit of Objectivity in Science and Daily Life*, Princeton University Press, Princeton, NJ, USA

Prévieux, Julien (2011), 'Atelier de Dessin - B.A.C. Du 14e Arrondissement de Paris', http://www.previeux.net/fr/works-atelierbac.html (Accessed: 17/08/2019)

PythonParseltongue (2017), 'FOREX Harmonic Pattern Scanning Algorithm Tutorial Series', https://www.youtube.com/playlist?list=PLFYcVF7huJR_4FRTZcR-rrSOpdTfoFnEI (Accessed: 22/07/2019)

Ramey, Joshua (2016), *Politics of Divination: Neoliberal Endgame and the Religion of Contingency*, Reinventing Critical Theory, Rowman & Littlefield International, London, UK

Research Through Design (2019), 'Method & Critique; Frictions and Shifts in RTD', https://www.researchthroughdesign.org/2019/index.html# (Accessed: 12/12/2019)

Rhodes, Brandon (), 'Skyfield; Elegant Astronomy for Python', https://rhodesmill.org/skyfield/ (Accessed: 19/07/2019)

Richardson, Rashida, Schultz, Jason, and Crawford, Kate (2019), 'Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice', Technical Report ID 3333423, New York University Law Review Online, Rochester, NY, https://papers.ssrn.com/abstract=3333423 (Accessed: 21/08/2019)

Rieder, Bernhard (2015), 'YouTube Data Tools', https://tools.digitalmethods.net/netvizz/youtube/ (Accessed: 10/07/2019)

Rieder, Bernhard (2016), 'RankFlow', http://labs.polsys.net/tools/rankflow/ (Accessed: 10/07/2019)

Rieder, Bernhard, Matamoros-Fernández, Ariadna, and Coromina, Òscar (2018), 'From Ranking Algorithms to 'Ranking Cultures': Investigating the Modulation of Visibility in YouTube Search Results', *Convergence: The International Journal of Research into New Media Technologies*, 24(1):pp. 50–68, doi:10.1177/1354856517736982

Rieder, G and Simon, J (2016), 'Datatrust: Or, the Political Quest for Numerical Evidence and the Epistemologies of Big Data', *Big Data & Society*, 3(1):pp. 1–6, doi:10.1177/2053951716649398

Robertson, Hamish and Travaglia, Joanne (2015), 'Big Data Problems We Face Today Can Be Traced to the Social Ordering Practices of the 19th Century.', blogs.lse.ac.uk/impactofsocialsciences/2015/10/13/ideological-inheritances-in-the-data-revolution/

Roose, Kevin (2019), 'The Making of a YouTube Radical', *The New York Times*, https://www.nytimes.com/interactive/2019/06/08/technology/youtube-radical.html (Accessed: 12/07/2019)

Roszkowska, Maria and Maigret, Nicolas (2015), 'Predictive Art Bot', http://predictiveartbot.com/ (Accessed: 19/08/2019)

Rouvroy, Antoinette (2011), 'Technology, Virtuality and Utopia: Governmentality in an Age of Autonomic Computing', in *Law, Human Agency and Autonomic Computing: The Philosophy of Law Meets the Philosophy of Technology*, eds. Hildebrandt, Mireille and Rouvroy, Antoinette, Routledge

Rouvroy, Antoinette (2013), 'The End(s) of Critique : Data-Behaviourism vs. Due-Process.', in *Privacy, Due Process and the Computational Turn.*, eds. Hildebrandt, M and De Vries, K, Routledge, London

Rouvroy, Antoinette (2016), 'L'art de Ne Pas Changer Le Monde', *La Revue Nouvelle*, (8), https://www.academia.edu/28370856/Le_gouvernement_algorithmique_ou_lart_de_ne_pas_changer_le_monde

Rudder, Christian (2015), *Dataclysm: Who We Are - When We Think No One's Looking.*, Broadway Books, New York, NY, USA

RYBN (), 'The Cybernetic Gambler's Hall of Fame', http://rybn.org/halloffame/ (Accessed: 03/06/2019)

RYBN (2015), 'ADMXI', http://www.rybn.org/ANTI/ADMXI/ (Accessed: 12/06/2019)

RYBN (2017), 'The Golem', http://git.rybn.org/rybn/golem (Accessed: 10/08/2018)

RYBN (2018), 'AAI Chess', http://rybn.org/human_computers/aaichess.php (Accessed: 05/09/2019)

RYBN (2019), 'Human Computers', http://rybn.org/human_computers/ (Accessed: 03/06/2019)

Sales, Tom (2011), 'Llull as Computer Scientist, or Why Llull Was One of Us', in *Ramon Llull: From the Ars Magna to Artificial Intelligence*, eds. Fidora, Alexander and Sierra, Carles, Artificial Intelligence Research Institute, Barcelona, Spain

Sandvig, Christian (2014), 'Seeing the Sort: The Aesthetic and Industrial Defense of "The Algorithm"', *Media-N Journal of the New Media Caucus*, http://median.newmediacaucus.org/art-infrastructures-information/seeing-the-sort-the-aesthetic-and-industrial-defense-of-the-algorithm/ (Accessed: 03/08/2018)

Sandvig, Christian, Hamilton, Kevin, Karahalios, Karrie, and Langbort, Cedric (2014), 'Auditing Algorithms: Research Methods for Detecting Discrimination on Internet Platforms', *Data and discrimination: converting critical concerns into productive inquiry*, 22

Sayers, Jentery, ed. (2017), *Making Things and Drawing Boundaries: Experiments in the Digital Humanities*, University of Minnesota Press, Minneapolis, MN, USA

Sayers, Jentery, ed. (2018), *The Routledge Companion to Media Studies and Digital Humanities*, Routledge, New York, NY, USA, doi:10.4324/9781315730479

Schmitt, Philipp (2020), 'Curse of Dimensionality', https://curseofdimensionality.science/ (Accessed: 18/05/2020)

Schmitt, Philipp, Dunne, Anthony, and Raby, Fiona (2018), *In Search of an Impossible Object*, Designed Realities Lab, The New School, New York, NY, USA

Schnapp, Jeffrey and Presner, Todd (2009), 'The Digital Humanities Manifesto 2.0', http://www.humanitiesblast.com/manifesto/Manifesto_V2.pdf

Schön, Donald (1983), *The Reflective Practitioner: How Professionals Think in Action*, Basic Books, New York, NY, USA

scikit-learn (), 'Nearest Neighbors Classification — Scikit-Learn 0.21.3 Documentation', https://scikit-learn.org/stable/auto_examples/neighbors/plot_classification.html (Accessed: 16/08/2019)

Seaver, Nick (2012), 'Algorithmic Recommendations and Synaptic Functions', *Limn*, (2), https://escholarship.org/uc/item/7g48p7pb

Seaver, Nick (2017), 'Algorithms as Culture: Some Tactics for the Ethnography of Algorithmic Systems', *Big Data & Society*, 4(2), doi:10.1177/2053951717738104

Seaver, Nick (2018), 'Captivating Algorithms: Recommender Systems as Traps', *Journal of Material Culture*, pp. 1—16, doi:10.1177/1359183518820366

Shapiro, Aaron (2015), 'The Next Big Thing In Design? Less Choice', *fastcodesign.com*, https://www.fastcodesign.com/3045039/the-next-big-thing-in-design-fewer-choices

Singleton, Benedict (2014), *On Craft and Being Crafty*, Doctoral, Northumbria University, http://nrl.northumbria.ac.uk/21414/ (Accessed: 23/07/2019)

Smith, Richard Joseph (2012), *The I Ching; a Biography*, Princeton University Press, Princeton, NJ, USA

Somers, James (2018), 'The Scientific Paper Is Obsolete', *The Atlantic*, https://www.theatlantic.com/science/archive/2018/04/the-scientific-paper-is-obsolete/556676/ (Accessed: 08/10/2018)

Søndergaard, Marie Louise Juul and Hansen, Lone Koefoed (2018), 'Intimate Futures: Staying with the Trouble of Digital Personal Assistants Through Design Fiction', in *Proceedings of the 2018 Designing Interactive Systems Conference*, DIS '18, pp. 869–880, ACM, New York, NY, USA, doi:10.1145/3196709.3196766

Srnicek, Nick (2012), 'Navigating Neoliberalism: Political Aesthetics in an Age of Crisis', *The Matter of Contradiction: Ungrounding the Object*, https://www.wdw.nl/files/Navigating_Neoliberalism_Political_Aesth.pdf

Srnicek, Nick (2017), *Platform Capitalism*, Polity, Cambridge, UK - Malden, MA, USA

Stahl, W A (1995), 'Venerating the Black Box: Magic in Media Discourse on Technology', *Science, Technology, & Human Values*, 20(2):pp. 234–258, doi:10.1177/016224399502000205

Stanley, Matthew (2013), 'Where Is That Moon, Anyway? The Problem of Interpreting Historical Solar Eclipse Observations', in *Raw Data Is an Oxymoron*, ed. Gitelman, Lisa, pp. 77–88, MIT Press

Stark, Luke and Hoffmann, Anna Lauren (2019), 'Data Is the New What? Popular Metaphors & Professional Ethics in Emerging Data Culture', *Journal of Cultural Analytics*, doi:10.22148/16.036

Stefaner, Moritz (), 'Truth & Beauty', https://truth-and-beauty.net/projects (Accessed: 13/08/2019)

Stengers, Isabelle (2005), 'The Cosmopolitical Proposal', in *Making Things Public. Atmospheres of Democracy*, pp. 994–1003

Stengers, Isabelle (2011), 'Comparison as a Matter of Concern', *Common Knowledge*, 17(1):pp. 48–63, doi:10.1215/0961754X-2010-035

Stephens, Zachary D., Lee, Skylar Y., Faghri, Faraz, Campbell, Roy H., Zhai, Chengxiang, Efron, Miles J., Iyer, Ravishankar, Schatz, Michael C., Sinha, Saurabh, and Robinson, Gene E. (2015), 'Big Data: Astronomical or Genomical?', *PLOS Biology*, 13(7):p. e1002195, doi:10.1371/journal.pbio.1002195

Stewart, John Q. (1948), 'Demographic Gravitation: Evidence and Applications', *Sociometry*, 11(1/2):pp. 31–58, doi:10.2307/2785468

Subramanian, Samanth (2017), 'Meet the Macedonian Teens Who Mastered Fake News and Corrupted the US Election', *Wired*, https://www.wired.com/2017/02/veles-macedonia-fake-news/ (Accessed: 22/10/2019)

Taylor, Alex and Kember, Sarah (2016), '4S Preview: Counting by Other Means', https://www.4sonline.org/blog/post/4s_preview_counting_by_other_means (Accessed: 20/08/2019)

Taylor, Sean J. and Letham, Benjamin (2017), 'Forecasting at Scale', *PeerJ Preprints*, doi:10.7287/peerj.preprints.3190v2

Thomas, Robert B. (1847), *The Farmer's Almanack*, Jenkins, Palmer & Co., Boston, MA, USA, https://archive.org/details/farmersalmanacca00thom

Thomas, Robert Bailey (1976), *The Old Farmer's Almanac*, Jenkins, Palmer & Co., Boston, MA, USA, http://archive.org/details/oldfarmersalman00thom (Accessed: 24/08/2018)

Thomson, Alison, Elzenbaumer, Bianca, and Franz, Fabio, eds. (2015), *PhD by Design - Instant Journal #2 - Researching across Difference (2015)*, https://www.academia.edu/19027082/PhD_by_Design_-_Instant_Journal_2_-_Researching_across_difference_2015_ (Accessed: 19/11/2019)

Toscano, Alberto and Kinkle, Jeff (2015), *Cartographies of the Absolute*, Zero Books, Lanham, MD, USA

Tufekci, Zeynep (2018), 'YouTube, the Great Radicalizer', *The New York Times*, https://www.nytimes.com/2018/03/10/opinion/sunday/youtube-politics-radical.html (Accessed: 01/10/2018)

Tufte, Edward (2001), *The Visual Display of Quantitative Information*, Graphics Press, Cheshire, CT, USA

Tukey, John W. (1977), *Exploratory Data Analysis*, Addison-Wesley Pub. Co, Reading, MA, USA

Uncertain Commons (2013), *Speculate This!*, Duke University Press, Durham, NC, USA

Van de Velde, Walter (2003), 'The World as Computer', in *Proceedings of the Smart Objects Conference*, Grenoble, France

van der Maaten, Laurens and Hinton, Geoffrey (2008), 'Visualizing Data Using T-SNE', *Journal of machine learning research*, 9(Nov):pp. 2579–2605

van der Walt, S., Colbert, S. C., and Varoquaux, G. (2011), 'The NumPy Array: A Structure for Efficient Numerical Computation', *Computing in Science Engineering*, 13(2):pp. 22–30, doi:10.1109/MCSE.2011.37

van Veen, Fjodor (2016), 'The Neural Network Zoo', http://www.asimovinstitute.org/neural-network-zoo/ (Accessed: 01/10/2018)

VanderPlas, Jake (2016), *Python Data Science Handbook*, Essential Tools for Working with Data, O'Reilly Media, Inc., Sebastopol, CA,USA

VanderPlas, Jake (2017), 'The Python Visualization Landscape', in *PyCon*, Portland OR, https://www.youtube.com/watch?v=FytuB8nFHPQ (Accessed: 18/08/2019)

Vega, Amador, Weibel, Peter, and Zielinski, Siegfried, eds. (2019), *DIA-LOGOS: Ramon Llull's Method of Thought and Artistic Practice*, University Of Minnesota Press, Minneapolis, MN, USA

Veloso, Pedro L.A. (2014), 'Cybernetic Diagrams: Design Strategies for an Open Game', *International Journal of Architectural Computing*, 12(4):pp. 379–397, doi:10.1260/1478-0771.12.4.379

Ventura, Joao (), 'Flatlib; Python Library for Traditional Astrology.', Flatangle, https://github.com/flatangle/flatlib (Accessed: 19/07/2019)

Venuturupalli Rao, Vyjayanthi, Krishnamurthy, Prem, and Kuoni, Carin, eds. (2015), *Speculation, Now!*, Duke University Press, Durham, NC, USA

Vernon, James (2018), 'The Making of the Neoliberal University in Britain', *Critical Historical Studies*, 5(2):pp. 267–280

Verteego (2017), 'Verteego Website', http://www.verteego.com/en/ (Accessed: 19/10/2017)

Vigen, Tyler (2015), 'Spurious Correlations', http://tylervigen.com/spurious-correlations (Accessed: 03/08/2018)

Villani, Cédric, Schoenauer, Marc, Bonnet, Yann, Berthet, Charly, Cornut, Anne-Charlotte, Levin, François, and Rondepierre, Bertrand (2018), *Donner un sens à l'intelligence artificielle*, Conseil national du numérique, Paris, France, https://hal.inria.fr/hal-01967551 (Accessed: 09/11/2019)

Vondrick, Carl, Pirsiavash, Hamed, and Torralba, Antonio (2015), 'Anticipating Visual Representations from Unlabeled Video', *arXiv:1504.08023 [cs]*, http://arxiv.org/abs/1504.08023 (Accessed: 08/07/2019)

Wark, McKenzie (2004), *A Hacker Manifesto*, Harvard University Press, Cambridge, MA, USA

Wark, McKenzie (2015), 'The Vectoralist Class', *e-flux*, 65, http://supercommunity.e-flux.com/texts/the-vectoralist-class/ (Accessed: 30/01/2019)

Weatherall, James (2013), *The Physics of Wall Street : A Brief History of Predicting the Unpredictable*, Mariner Books, Boston, MA, USA

Wegman, Edward J. (1990), 'Hyperdimensional Data Analysis Using Parallel Coordinates', *Journal of the American Statistical Association*, 85(411):pp. 664–675, doi:10.2307/2290001

Wilkie, Alex, Savransky, Martin, and Rosengarten, Marsha, eds. (2017), *Speculative Research*, Routledge, London, UK and New York, NY, USA

Wilkinson, Leland (1999), *The Grammar of Graphics*, Springer, New York, NY, USA

Willis, Roy G and Curry, Patrick (2004), *Astrology, Science and Culture*, Berg Publishers, Oxford, UK

Wu, Xiaolin and Zhang, Xi (2016), 'Automated Inference on Criminality Using Face Images', *arXiv*, https://arxiv.org/abs/1611.04135v1 (Accessed: 20/08/2019)

Zipf, George Kingsley (1942), 'The Unity of Nature, Least-Action, and Natural Social Science', *Sociometry*, 5(1):pp. 48–62, doi:10.2307/2784953