# geographies of (con)text: language and structure in a digital age

2017-11-25 16:11:19  andrew

## Abstract

This paper puts forward the concept of 'geographies of (con)text' to critique the metaphors and materialities of 'the digital', concentrating on the physical constructs and constraints of language on the web. A landscape of words as opposed to a landscape of code (Thrift & French, 2002), language-as-data becomes material in ways very different from both print and spoken word; its physicality represented in bits, bytes and circuitry, and its limits and variations mediated and governed by the processes which order, sort, move and index it. By virtue of their reproducibility and enhanced means of dissemination, digitised words can have paratextual – and often political – agencies and excesses beyond their linguistic function. Using examples of online search, dictionaries and translation, the paper will imagine how context as a kind of space might be produced, constructed and limited, how competing actors contribute tactically (de Certeau, 1984) to the (in)visibility and (im)mobility of the linguistic data in the searchable database, and how these actors negotiate the conflicting interests of money, efficiency and truth (Lyotard, 1984) in the geo-linguistic spaces of the web. With a geography of (con)text thus imagined, the mathematical and binary logics that construct and mediate the language within it are also clearly exposed. The paper goes on to discuss how creativity and originality might be restricted by ongoing processes of quantification and monetisation of language, before concluding that digitised language falls somewhere in the middle of a structuralist/post-structuralist critique; being at the same time both free from and constrained by the geographies of context.

## Introduction

What happens when words become data? Now that almost all language is in some way digitised, 'data-ised' and submitted to computation, the ownership, control and organisation of linguistic and other data can have a profound, yet not always intentional or indeed predictable, impact on social, cultural and political discourse, and on personal freedoms and securities. As well as significant power and profits for those who harvest and process it, the 'data-isation' of language brings the risk of serious collateral damage. By virtue of their reproducibility and enhanced means of movement and dissemination, words-as-data can have paratextual agencies and excesses beyond their linguistic function; the granular configurations and distortions of data can be instantly and exponentially magnified to a global geopolitical and discursively significant scale. But in contrast to this apparent post-modern untethering of language from its locational and referential functions, language that appears non-normative, or somehow unexpected, such as poetry or other creative variations of text, resists computation, and as a result has its movements through digital space restricted. It becomes suspect; penalised for its originality, or its unmarketability, as it fights its way through search engines, firewalls and new libraries of spam.

This paper explores this apparent contradiction in linguistic agency, examining digital geo-linguistic spaces of data as 'geographies of (con)text' – as in 'within text' – or the way words are ordered, or literally 'woven together' and processed within the corpus of linguistic data on the web. Once reduced to data, language loses the linear, or narrative order it might have had on the printed page or in speech. It is deconstructed in the process of digitisation, becoming part of the fluid pool of decontexualised data which flows through the spaces of the web, commoditised by the companies that increasingly control and mediate those spaces through advertising platforms and social networks. Language-as-data produces new contexts, in both a linguistic and physical sense. Processed on huge scales, digitised data has a wide geographical reach, so when it is reconstructed back into what we might call ordinary language (for example via search results, online translations or dictionaries), the manner in which that data has been stored, ordered and moved becomes extremely important.

Imagining context in this way facilitates the study of language in digital spaces from both material and theoretical angles. Firstly, it becomes necessary to think about how context as a kind of space might be produced, and the structures, biases and lacunae of the data within it. Secondly, we can start to think about the movement of data though those spaces. How is language-as-data made retrievable, searchable, (in)visible, and exploitable by a range of different actors, from tech giants such as Google and Apple, to individual web users, advertisers, activists and politicians, and how their competing motives and agencies change the contextual makeup of that space. Not only is the ownership of language online a lucrative business (as we see from Google advertising platforms AdWords and AdSense), but as Lyotard suggests, 'context control'[1] in the production of knowledge is a powerful means for the legitimisation and therefore dominance of information. The physical and epistemic context of language reduced to data is perhaps at once liberating and restrictive. When reduced to data, words carry an entirely new set of values, motives, and meanings. They may become trapped within the logic of the digital economy, but can also become magnified to new scales and agencies by virtue of algorithmic systems of processing and interpreting. In conceptualising the de/re/constructions of linguistic data in this way, the paper therefore considers both the physical and theoretical implications of the geographies of (con)text.

## What do I mean by (con)text?

Much of the work in the nascent field of 'digital geographies'[2] has to date been focused on how digital technologies affect and produce geographies[3], or how space has become intrinsically linked with code[4]. Yet more work has been done on algorithms and their work in

which these algorithms work in the hope of opening up thought to the material and theoretical spaces of calculation . . . . . . . .ch the data and algorithms interact, and in particular when the data is representing language. Even if such space is considered to be primarily topological, as is the case with vector space, the spatial relations within it are still important. As Till Straube contends, 'the various articulations of relational systems making up the layers found within digital infrastructures should be taken seriously as spaces proper'[7]. Although the context of words on pages have always been linked to their economic value (as a book, a telegraph, a newspaper article, for example), the manner in which data-ised words sit in relation to other words on web pages has taken on epistemic, economic, and even political possibilities unimaginable within the relative constraints of print capitalism. New kinds of capitalism mediate the contextual spaces of data. Forms of digital, platform, semantic, or linguistic capitalism[8] have created new workers, owners, users, products and commodities, and their currencies are based predominantly on data. But even if we think of (con)text as an abstract space, the generation of value by and within that space lends it a 'very real social existence'[9] wherein those who dominate have corresponding agency in 'real' space. As Shaw and Graham have shown, there is great power (and great money) in the control of informational flows by large technology companies such as Google. Of course, the 'text' of (con)text does not refer only to words, in fact its Latin root comes from texere (to weave). If we think how the weaving together of different threads determines different patterns, meanings and values in textiles, then the use of (con)text as a metaphor for the abstract space of linguistic data becomes even more relevant and politically significant. Often regarded as a forerunner of digital computer programming, and also the catalyst for new sets of labour relations,

technologies such as the Jacquard loom not only revolutionised the weaving of silk and other materials, but also had wide-scale societal and political impacts similar to those of the digital revolution. Just as big data technologies – and the patterns and narratives they produce – have become contentious issues today, the automation of the textile industry in nineteenth century Europe saw widespread resistance from workers facing unemployment and exploitation. In an era of AI and machine learning, data, and its (con)text, has become a similarly controversial raw material.

This paper thus sets out to develop a critique of 'the digital' that concentrates on the physical constructs of language on the web. Rather than the 'landscapes of code' imagined by Thrift and French[10], the paper seeks to imagine a 'landscape of words', or a 'geography of (con)text'. A geography of (con)text can be imagined as the relational makeup of an ever-changing body of text, through and amongst which paths and trails of significance are intentionally or unintentionally traced, woven, diverted or created, producing dynamic contextual spaces, in and with which different actors, with differing motives, continuously engage. The paper will begin by thinking about bodies of linguistic data as relational, yet often problematic and incomplete spaces. It will then discuss how competing actors contribute to the (in)visibility and (im)mobility of this data on the web; how it moves and is moved around, exchanged and manipulated. I find it useful here to imagine these movements as maneuvers, played out like de Certeau's everyday tactics in attempts to (re)gain power over or subvert certain narratives, controlling conduits of meaning, or perhaps 'poaching'[11] in lexical spaces. As de Certeau writes, 'on the blank page, an itinerant, progressive, and regulated practice – a "walk"- composes the artifact of another "world" that is not received but rather made'[12], but 'walking' through digital space is not necessarily such a progressive or liberating linguistic experience. The mathematical and often binary logics that make and mediate the language within this space set paths from which it is hard or inadvisable to deviate; systemic processes which are only complicated and exacerbated by the commodification of linguistic data by technology/media companies such as Google, Facebook or Apple. It is therefore to the dominant structures of (con)text we turn first, before considering the other actors within these spaces, their tactics, motives and intentions, and then finally thinking about the unintentional – or collateral consequences of maneuvers within this space which are made perhaps inevitable by the technologically magnified scale and reach of language-as-data.

### Language spaces

It is estimated that only 5% of the data which exists on the web is indexed by commercial search engines. The machinations of the deep web and the dark web are beyond the scope of this paper, but suffice it to say, there are still over 100,000,000 gigabytes of data for search algorithms to work on[13]. This is what I refer to as the searchable database[14]. The data that makes up the searchable database is, however, not necessarily democratically spread or representative of society, and neither is its visibility (or findability) upheld by any explicit or implicit norms of 'fairness'. Factors such as categorisation, language, tagging and indexing make some data easy to find, and some not, and the algorithmic systems managing the data are also always influenced by the motives, skills and potential prejudices of the flesh and blood programmer. The linguistic data which makes up the searchable database is also not an accurate reflection of all the analogue data produced in the world – as many scholars have noted, there are significant weightings and gaps in the technological, linguistic, geographic or social agency of some groups or areas which means that the 'searchable database' is a distinctly hierarchical and undemocratic dataset before commercial ranking algorithms have got anywhere near it. The corpus of linguistic data that exists in this digital space is far from inclusive. What can be contributed to the searchable database is highly contingent and can exclude groups of people on multiple social-economic levels. Despite the oft-touted democratic and analytic qualities of big data, as Kate Crawford writes, although 'data are assumed to accurately reflect the social world… there are significant gaps, with little or no signal coming from particular communities'[15]. Those with better technology, better skills, better social standing, better connectivity and with access to better platforms, are the main authors of the searchable database, and it is this privileged data that is used by the wealthiest and most powerful technology companies to search for answers to the world's questions. As organic search results, auto-completions and auto-suggestions are based on a mixture of previous queries and the data already existing on the web, the volume and structure (including any bias) of this data becomes crucially important in how these results are algorithmically extracted and reconstructed.

### Searching

The surprising, confusing, and even distressing results sometimes produced by search engines are well documented, and can have a significant impact on the wider cultural and political discourse, reinforcing stereotypes[16], marginalizing the less powerful or wealthy, or spreading 'fake news' and even potentially influencing election results[17]. A personal example of this stemmed from a search for the phrase *wives and girlfriends sexist*, which I had typed in Google after hearing 'wives and girlfriends' being used problematically at a security briefing. I had been expecting to find some cultural criticism about the portrayal of footballers' wives in the press and media, or even some more nuanced critique of the sexual semantics of the phrase – something along the lines of Cynthia Enloe's 'womenandchildren'[18]. What I was not expecting was for Google's search algorithm to 'correct' the word *sexist* to *sexiest* (see Figure 1) and for my results to consist entirely of references to the top ranking *sexiest* and *hottest* WAGs in its index.

Google  wives and girlfriends sexist 🔍

Web    Videos    Images    News    Shopping    More ▾    Search tools

About 1,080,000 results (0.33 seconds)

**#BRAZIL2014: Top 10 Sexiest Wives & Girlfriends Of ...**
www.360nobs.com/.../brazil2014-top-10-sexiest-wives-girlfriends-of-fo... ▾
21 Jun 2014 - 360nobs presents the **sexiest** WAGS(**Wives And Girlfriends**) at the 2014
World Cup. Enjoy!! 10)Kimberly Crew: Beautician Kimberly Crew is ...

**Top 10 Sexiest Soccer Wives and Girlfriends | NUVOtv |**
www.mynuvotv.com/blog/**sexiest-soccer-wags-2014** ▾
Soccer is the most popular sport on the planet. That's why it's no surprise that many of
the athletes who play it live the kinds of lives usually reserved for.

**101 Sexiest Soccer Wives and Girlfriends | Bleacher Report**
bleacherreport.com/.../772545-101-**sexiest-soccer-wives-and-girlfriends** ▾
24 Jul 2011 - 101 **Sexiest** Soccer **Wives and Girlfriends**. 561.5K. Reads. 18.
Comments. The exotic **women** who grace International stadiums are often seen ...

**AOL.com Article - The hottest wives and girlfriends of World ...**
www.aol.com/article/2014/06/18/...**wives-and-girlfriends**.../20915552/ ▾
18 Jun 2014 - The hottest **wives and girlfriends** of World Cup players. ... Click through
the gallery above to see some of the **sexiest** WAGs and let us know your ...

**Top 10 Sexiest NFL WAGs (Wives And Girlfriends) - YouTube**
www.youtube.com/watch?v=kq8Jfp3KHO0 ▾
11 Dec 2013 - Uploaded by 10ReasonsWhyShow
▶ 8:20   They have it all: the money, the physique, the cars and especially
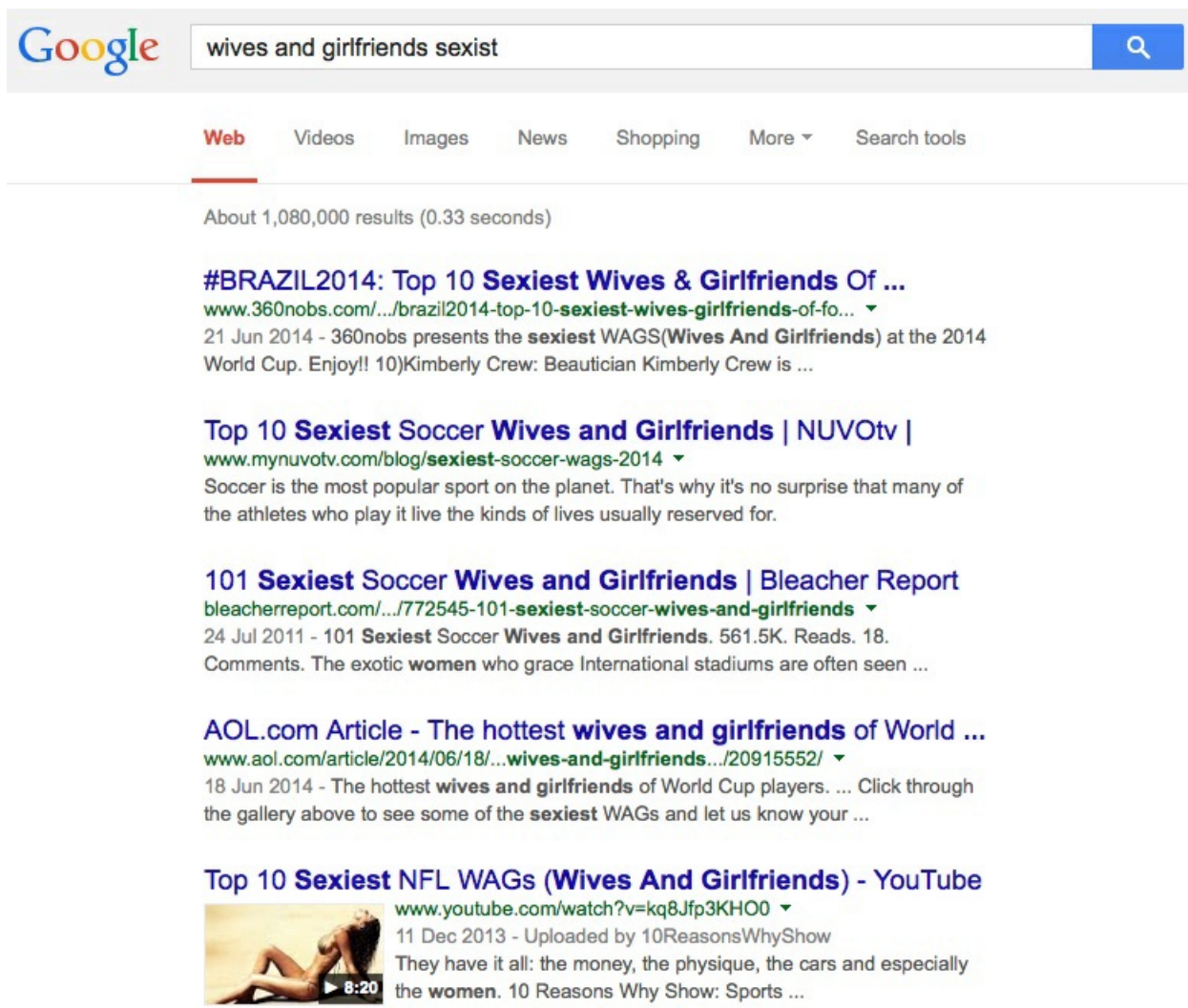the **women**. 10 Reasons Why Show: Sports ...

Figure 1. Google search *wives and girlfriends sexist*, screenshot February 20, 2014

Apparent anomalies and glitches such as these have tended to be held up popularly either as proof of some kind of inherent sexism/racism within the system, or as sad but inevitable reflections of society. However, the algorithmic decisions that generate them are not based on any cultural or semantic knowledge as such, but on the mathematical logics of search technologies, on vectors, on marked up, decontextualized language, and on the analysis of big data by 'keyness and co-location', a method of corpus linguistics which in the field of Critical Data Analysis has been criticized because it 'omits essential qualities of actual language use'[19]. In the pool of data in which the algorithm works, the word *sexiest* is statistically more likely to be linked to the phrase *wives and girlfriends* than the word *sexist* – the associations perhaps compounded by the disproportionate volume of sexualised content or sensationalised news copy in the database, by algorithmic reproduction of erroneous typos or synonyms, and by the machinations of the linguistic marketplace.

Due to the black-boxed nature of the proprietary technology which operates commercial search engines such as Google, it is almost impossible to reverse engineer the search algorithm, as many scholars have pointed out[20]. However, it is possible to understand why such apparent anomalies occur by studying the development of the information retrieval systems on which search technology is based, although some of the more scientific literature in this area can itself be fairly uncritical. In order to combat natural language processing (NLP) problems such as synonymy and polysemy[21], most search engines now use semantic searching based on Vector Space Analysis, an algebraic method of measuring the spaces between words in a large corpora data set. While Vector Space Models differ, their use in language corpora is based on the assumption that 'words that are closer together in the vector space are semantically closer in some sense'[22]. In this way, one recent study claims to have revealed the 'human' biases in semantics found in language corpora using a process of word embeddings. Caliskan-islam et al.'s 2016 report used various already acknowledged stereotypical word associations and compared them to the results of vector space analysis on data collected from the web, finding a correlation between human bias and prejudice and that found in the machine learnt results. The analysis revealed, for example, that 'flowers are significantly more pleasant than insects, and insects more unpleasant than flowers', musical instruments are more pleasant than weapons, and that European American names are more pleasant that African American names[23]. Similar results were found in another recent report that used a corpus of text from Google News as its training set[24]. Bolukbasi et al. found that 'the closest word to the query BLACK MALE returns ASSAULTED while the response to WHITE MALE is ENTITLED TO'[25]. Although these reports stress the importance of the contextual ordering of words in the production of meaning – even in non-absolute topological space – and show how a result like the *wives and girlfriends* example can have occurred due to a reflection of loaded or incomplete linguistic corpora, they are not without their problems. There is an important distinction to be made between the way algorithmic models work and the data they work upon, although the distinction is often ignored. Lev Manovich writes that algorithms and data are 'two halves of the ontology of the world according to a computer'[26], and as I have explained above, how the database is produced and constructed, and the relationships between the data within it is crucially important to the outcome of search results. Caliskan-islam et al.'s report notes that the statistical machine-learning model they used 'knows' the properties of flowers and insects 'with no direct experience of the world, and no representation of semantics other than the implicit metrics of words' co-occurrence statistics that it is trained on'[27]. As discussed above, given the

incomplete, loaded and skewed nature of the data search engines can work on, co-occurrence alone is no basis for absolute knowledge. Their model does not 'know' the properties of flowers any more than the Google algorithm 'knows' women are sexy and not sexist; what it 'knows' is how the properties of flowers are represented in particular body of data, which is not the same thing.

## Translating

The systemic problems faced by search data and algorithmic processing can be seen also in online translation, where similar vector space analysis is deployed to produce the most accurate predictions of corresponding words in different languages. Although its accuracy has improved over time (and there are of course other online translation tools), Google Translate has always been infamous for its amusing and intriguing inaccuracies and anomalies, a good example of which is when in early 2016 the word *Russia* began to be synonymized with *Mordor* when translated from Ukrainian to Russian. Reports at the time suggested either that the translations had somehow been manipulated by Ukrainian hackers to ridicule Russia[28], or were caused by a 'bug' or an 'automated error' in the algorithm[29]. Google's own response to the incident simply reiterated what has long been known – that Google Translate does not actually translate from language to language in a semantic way, but relies on the corpus of data available on the web in a particular language to provide most likely predictive matches based on the frequency and proximity of words and phrases in vector space. Google put out statements explaining how its translation tool 'looks for patterns in hundreds of millions of documents to help decide the best translation for you…. automatic translation is very difficult, as the meaning of words depends on the context in which they're used'[30]. As the context for the translate algorithms is the same indexed dataset used by search algorithms, then there is always the possibility that 'translations' will pick up on cultural or regional anomalies, or stories most recently disseminated. As several sources pointed out, there are a significant number of references to Russia and Mordor in Ukrainian language popular literature, news and satire[31]. Such is the strength of association between the two terms, that if the words *Russia* and *Mordor* become interchangeable in the limited quantity of Ukrainian text which Google Translate has to work on, then 'glitches' like this will happen. It isn't a 'bug' or an 'error', in fact the algorithms are doing exactly what they are programmed to do. Of course there is the possibility that the anomalous glitches could have been deliberate, and similar ones could easily be replicated with the appropriate amount of effort; it is after all relatively easy to 'play around' with Google, as some Google Map hackers[32] and Googlebombers[33] have shown, although presumably it is becoming harder to do so anonymously as more and more Google services require logging in. Whether or not it was a concerted tactical attempt to manipulate and subvert the system, or whether the data available to the translate algorithms was already loaded with cultural references which might cause mistranslations, the (de)contextual ordering of words on a digital level in this case was quickly magnified to a geo-political and potentially dangerous diplomatic level. A similar thing happened in 2013 when researchers contacted US intelligence officials after noticing that Google Translate had begun 'translating' the cod-Latin placeholder text Lorem Ipsum into 'apparently geopolitical and startlingly modern' English words and phrases such as China, NATO and Internet[34]. Rather than uncovering some kind of spy communication network, however, this was simply an insight into how translation technology projects vector representations and 'word neighbors' from a source 'language space' (in this case Lorem Ipsum), to target 'language space' (English), thereby predicting the most likely translation[35]. The incident therefore reveals far more about the types of website which use Lorem Ipsum than it does about covert operations. The web only holds a finite quantity of real Latin text and corresponding translations, just as it only holds a finite amount of Ukrainian text and translations, so when the Lorem Ipsum text is used as a placeholder which corresponds with the content of a website which can be viewed in different languages, in effect diluting the word-stock of that language, then the algorithm knows no better than to match the 'Latin' word to the English one. As Lorem Ipsum tends to be used as placeholding text on the multi-lingual websites of official organisations, governments, diplomatic pages or multinational businesses, so the 'translations' will reflect the language used on those sites.

## Defining

Online dictionaries also suffer from algorithmic predictions of meaning based on large corpora data sets. Although Samuel Johnson was criticized for stamping his own personality on some of the definitions in his 1755 *Dictionary of the English Language*[36], the objectivity of an equally popular but more modern, online dictionary was recently also called into question when it started to use the phrase *rabid feminist* as a contextual usage example for the word *rabid*. Oxford Dictionaries Online, the digital version of the Oxford Dictionary of English which (amongst other things) supplies Apple products with their built-in British-English dictionary[37], responded to criticism of this apparently sexist word pairing by explaining that 'our example sentences come from real-world use and aren't definitions'. Their website explained how the phrases are drawn from 'a vast bank of more than 1.9 million example sentences (around 38 million words) of real English, extracted from the world's newspapers and magazines, academic journals, fiction, and blogs'[38].

# rabid
/ˈrabɪd, ˈreɪ-/ ◀))

*adjective*
adjective: **rabid**

1. having or proceeding from an extreme or fanatical support of or belief in something.
   "a rabid feminist"
   *synonyms:* extreme, fanatical, overzealous, over-enthusiastic, extremist, violent, maniacal, wild, passionate, fervent, diehard, uncompromising; intolerant, unreasonable, illiberal, bigoted, prejudiced, biased, partisan, one-sided; *informal* raving, gung-ho; *informal* swivel-eyed; *literary* perfervid
   "a rabid anti-royalist"
   *antonyms:* moderate, liberal, half-hearted

Figure 2. Oxford Dictionary example of *rabid*, screenshot October 17, 2016

Just like Google's search and translation tools, the Oxford Dictionary's examples are seemingly just a reflection of the linguistic data that exists online, which as I explained above, is not necessarily representative, democratic, or untainted by the technologies and commercial interests which mediate it. Unlike in Johnson's day, definitions based on digital data are calculated mathematically, so their potential inaccuracies or controversies are based on the accumulated mass of available data, rather than on the qualitative bias of one person or small group. The phrases generated as likely usage examples will therefore mirror the most likely pairings or orderings of already existing words based on the semantically irrelevant factors of frequency and proximity. It would therefore seem (rightly or wrongly) that the word *feminist* is topologically closer to the word *rabid* than any other word in the corpus of available linguistic data,

including, we must assume, the word *dog*, which is an interesting linguistic development in itself.

What can be drawn from the examples of online search, translation and definitions, is that when language becomes data – due to the nature of data processing, language technologies and digital networking – it becomes volatile and invasive; its effects spreading more widely and more quickly, than the printed word. It can perpetuate stereotypes and inequalities, confirm biases, create diplomatic incidents, and – with the increasing ubiquity and indispensability of the technologies that employ it – has a perhaps unprecedented impact on web users. The Oxford Dictionary has what is presumably a very lucrative contract with Apple for the supply of words for its database of definitions and examples, all of which are made visible at the touch of a screen though the millions of Apple devices worldwide. With examples such as 'rabid feminist', the Oxford Dictionary is in effect 'beaming a sexist lexicography straight to students' iPads'[39]. While there may be no tactical agenda for Apple beyond a convenient and cost-effective business agreement, the physical reach and volume of their products, coupled with their reliance on algorithmically mediated samples effectively puts an extraordinary and unintended epistemic power in the hands of another private company. As useful and ubiquitous as the tools provided by companies such as Apple and Google are, what we can see from this is that the way big tech companies use mathematical methods to extract search results, translations and definitions from large corpora can have serious side effects. But there are other actors in this space; activists, politicians, academics, advertisers, and even perhaps the hand of the market are all part of these contextual maneuvers.

## Language games
Recent debates around privacy and encryption have highlighted the specific nature of the interests of the state, civil society and commercial companies in the visibility or invisibility of linguistic (and other) online data. Text can be hidden or redacted for private or political purposes – through the use of passwords, paywalls or even through site architecture such as sitemaps and robot.txt. Such was the case, for example, in 2003 when it was reported that the White House had hidden webpages which referenced Iraq from being indexed by the webcrawlers[40]. Privacy activists have suggested ways to deliberately obscure personal data. Brunton and Nissenbaum's users' guide to 'Obfuscation' calls for the 'deliberate use of ambiguous, confusing, or misleading information to interfere with surveillance and data collection'.[41] They see obfuscation as a tactic, or a 'weapon of the weak'[42] to be deployed from within the surveillance society using 'everyday' language as a subversive tactic against the powerful.[43] Criminal activities such as phishing attacks often vary a web address or keyword by one letter in order to exploit the inadvertent user who may have misspelled a search query, or clicked on the wrong link, thus directing them to a fake site. 'In the lexicon of the World Wide Web' write Matt Fuller and Andrew Goffey, 'such typos are the homonyms and synonyms, the words that allow a user to pass over into another dimension of reference'[44]. Companies or individuals can guard the portals and spaces of sites by buying up domain names or keywords. Spam emails often rely on algorithmically generated jumbles of words that need to make just enough sense, not to fool humans (at least in the first instance), but rather, to fool firewalls. Language in online spaces can also be disguised in order to deceive commercial copyright or plagiarism software. Evidence of the resulting linguistic mutations this embeds in the database is clearly visible. Tarleton Gillespie uses the example of the misspelling of 'Britny Speers' so that people searching for illegal downloads could find recordings, 'but the record industry software could not.'[45] On the flipside, however, it is sometimes impossible to search for a specific word without being autocorrected, which may seem to give the platform doing the 'correcting' a more 'strategic' method of 'producing, tabulating and imposing'[46] power within the contextual space of the search, yet, as I will suggest, power balances within geographies of (con)text are constantly being challenged by actors on all levels, and it is hard to imagine an unconnected strategic vantage point within that context.

Data-ised language in this sense becomes either a tool, or means of access or restriction; its linguistic function as a means of human communication becoming secondary to its use and exchange value for other purposes. But as well as being used as a tactical tool, language can in itself have its movements restricted as it passes through digital spaces. Non-normative, creative language can be actively 'criminalised', not only by Google, but also by firewalls or anti-virus software which blocks text that does not conform to whatever 'natural language' databases or templates it has been taught. The condensed, carefully worded nature of poetry can sometimes be construed by algorithms as key-word stuffing; a tactical manipulation of the search rankings which is frowned upon by Google, and which can lead poetry websites to be 'buried' in the search rankings[47]. Artist Sophie Mayer noticed how the poetry magazine *Poems in the Waiting Room* had to issue advice to potential contributors, asking them to send their submissions in the body of an email rather than in an attachment, as the attached poetry might be mistaken for a virus by the spam filtering software[48]. All this starts to beg the question – just how communicable is poetic or creative language in a digital age? Is it less communicable than a virus? How restricted – or restrictive – is language when it becomes data? And further to this, what is the function (and fate) of language if it is written for, or increasingly by in (the case of algorithmic content generating) machines or algorithms for the sole consumption of other machines or algorithms? As the method for this type of exchange is all acted out mathematically, then the onward effects on non-mathematic language becomes little more than collateral damage; the leftovers of a linguistic power-grab perhaps, or a post-modern language game[49].

## Language markets
Poetry is also an interesting lens through which to think about another major player in the landscape and mobility of online language – the advertising industry. As part of a critique and artistic intervention into the commodification of language by Google's AdWords platform, I have been calculating the suggested bid prices of poems which have been fed through Google's keyword planner[50]. As is commonly known, AdWords is the system by which advertisers bid for keywords and phrases in order to secure the top spots on the search engine results page. When somebody searches for a word on Google, an auction takes place, and the advertiser with the highest bid at that time wins the right to use that word in an advert. Google then earns the price of the winning bid each time the advert link is clicked on. Using popular poems as examples allows me to explore the change in the terms on which words are valued as they move through the search engine. As they become monetised, their value becomes determined not by aesthetics or poetics, but by how much money the words will make Google in adverts. Words become primarily carriers of economic value; stripped of any poetic value they may have had when they entered into the search bar. In the poem *Daffodils*, for example, the words *crowd*, *cloud* and *host* are given relatively high suggested bid prices not because of Wordsworth's Cumbrian vision, but because they are revenue generating search terms for advertisers of *cloud computing*, *crowd funding*, and *web hosting*.

Daffodils
by William Wordsworth

SALE

19th May 2016          12:27 PM
BATCH #: CRC-32
APPR #: D25CE1BA

| 1 | i | £0.28 |
| 1 | wandered | £0.00 |
| 1 | lonely | £0.91 |
| 1 | as | £0.42 |
| 1 | a | £0.46 |
| 1 | cloud | £2.12 |

SUBTOTAL:          £4.19
TAX:                    N/A
TOTAL:               £4.19

APPROVED

THANK YOU FOR SHOPPING AT GOOGLE

CUSTOMER COPY

Figure 3. example of {poem}.py: a critique of linguistic capitalism, Pip Thornton (2016).

In this way language becomes infused with the logics and values of what might be termed a linguistic marketplace; its grammars, patterns and frequencies formed not through generations of linguistic happenstance, but as advertisers compete for lucrative keywords, and (as we have seen earlier) as machines learn and magnify algorithmic glitches. Words have gained a currency detached from, for example, their narrative function. Not only do their prices fluctuate according to apparently neoliberal free-market competition, but there are also many distortions to this market due to centralised interference and control from Google in the form of regulations, censorship and other linked applications. The data-isation and monetisation of language in this process has a subsequent effect on the density and frequency of certain words within the searchable database, privileging both paid-for words and phrases, and also organically optimised language in the search results. To paraphrase N. Katherine Hayles, language does not emerge unscathed from its encounter with code[51]. Indeed, data-ised language tells other, paratextual stories as it moves and is moved through digital space. It gains economic value as it draws eyes to adverts and negotiates spaces controlled by 'cognitive rent'[52] or 'lexical squatting'[53], and as we have seen, it becomes weighted with the residue of the wider pool of decontextualized, yet hierarchised data, or 'tracts of knowledge'[54]. This residue turns tools such as translate and search into 'authoring device(s)'[55], and turns (machine) translations into new linguistic art forms[56]. But the stories told and the paths forged in this manner are always products of the structures from which they originate and through which they flow. Even the vectors which quantify data-ised language do not just point to, or connect things, but are carriers of meaning – they

add to the story. According to Phil Garnett, we too have become vectors in the digital assemblage[57], or as Mackenzie Wark puts it:

'Both the flâneur and the facebooker are voluntary wanderers through the signage of commodified life, taking news of the latest marvels to their friends and acquaintances'[58]

But Wark's flâneur is very much the Benjaminian flâneur of the arcades, not the modern day dérive-drifter whose goal is to subvert, not admire, the spectacle. It is surely impossible to drift through digital space today without picking up the 'signage of commodified life' in the form of 'likes', 'shares' and click-bait, whether voluntary or not. As de Certeau identified, with the ubiquity of the mass media 'instead of an increasing nomadism, we thus find a "reduction" and a "confinement": consumption, organised by this expansionist grid takes on the appearance of something done by sheep progressively immobilized and "handled" as a result of the growing mobility of the media as they conquer space.'[59] There is of course much more to be said about the changing values of data-ised language and the circulation of linguistic capital, but that is for another paper.

**The prison house of digitised language**

In conclusion, it is apparent that language is an integral element of digital space, and one way to conceptualise this is through thinking about the composition of data-ised language as '(con)text'. Just as in other spaces, there are social and political factors at play, and power is exerted and subverted by competing actors. Following de Certeau, the manipulation and movement of language within this produced space can therefore be read as kinds of tactical maneuvers which 'make use of the cracks that particular conjunctions open in the surveillance of the proprietary powers'[60]. We can see in the decontextualised database, or, indeed, in such a simple action such as cutting and pasting what Derrida called the 'citational graft' of a sign; the capacity to put quotation marks around any word or phrase and move it elsewhere, thus 'break(ing) with every given context, engendering an infinity of new contexts in a manner which is absolutely illimitable'.[61] As Ming Lim suggests, in the mechanics of the search engine industry we can see that there really is nothing outside of the (con)text:

In this space, signs truly refer only to other signs and all pretense at 'presence' or 'essence' or 'authenticity' is no longer necessary. Marketers who use SEM [Search Engine Management.] are now subject to the peculiar laws of an architecture which have barely begun to be theorised[62]

But Derrida's illimitable contexts were imagined before the structures and restrictions of modern digitised language were fully realised. In this context, words are not 'free' to be eternally deferred, but become exponentially re-infused with the residue of the dominant structures, market forces, biases and stereotypes that make up the corpus of the searchable database. Inextricably linked with code[63], their meaning has been 'arrested'[64]; chained to the past by algorithmic association and given a record it is impossible to expunge. Followed around by 'semantic escorts'[65] which have the power to decide who or what can and cannot be sexist, (un)pleasant, rabid or racist, they have in effect been 'reconstructed'; squeezed through binary systems which force an either/or logic on words even as they squirm to get away. And neither is digitised language free from an organising metanarrative. Words-as-data, already (re)constructed by virtue of the way they are stored and transmitted are also monetised within the system of 'linguistic capitalism'[66]. As Franco Berardi states, 'today the economy is the universal grammar traversing the different levels of human activity'.[67]

But what of de Certeau's tactics? Who benefits from the structural and economic restrictions of language-as-data? If, as Gunnar Olsson believes 'it is in the interest of social cohesion to impoverish language'[68], then the power wielded by the big technology companies such as Google, Apple and Facebook becomes even more terrifying. In terms of '(con)text' – that is the geo-linguistic space of the web – it might perhaps be easy to think of these big technology companies as strategic actors in de Certeau's sense of an institution which has established itself in 'a *place* that can be delimited as its own and serve as the base from which relations with an *exteriority* composed of targets or threats (customers or competitors)… can be managed'[69]. But it is more complicated than that. As the content and the currency of the web is made up of interactive, interconnected, data-ised language, none of the actors (including us as users) involved in the kinds of (con)textual maneuvers I have described can ever act from a completely strategic (air gapped) vantage point. While there are different levels of access to other resources that shape digital infrastructures, such as source code, what we might call 'natural' language – which has been deconstructed and decontexualised into data – has a liquid quality that permeates through the spaces of the web, in some way making users of us all. It may seem that with their control over the mediation of language, the big technology companies have the upper hand, but they have, to a certain extent, created something over which they no longer have complete strategic control. What does seem to be the case, however, is that like the spaces of consumerism described by De Certeau, language-as-data always seems to have an ulterior, paratextual motive and, perhaps because of the mathematised logic of code, can never just be language. So just as we are trapped within the structures of a digitalised discourse, so are we too involved to be able to look in from the outside from a strategic vantage point. 'There is' de Certeau says 'no longer an elsewhere.'[70] We are truly coopted; practicing out everyday lives within what Jameson might have called the prison house of digital language[71]. Yet despite its structural constraints, the manner in which linguistic data is disseminated online can, as I have shown, have far reaching effects untethered to and unpredicted by its topological environment or physical location. In this way, it could be concluded that digitised language falls somewhere in the middle of a structuralist/post-structuralist critique; being at the same time both free from and constrained by the geographies of context.

**Bibliography**

Amoore, Louise. "Cloud geographies: Computing, data, sovereignty." *Progress in Human Geography* (2016).

Ash, James, Kitchin, Rob, and Leszczynski Agnieszka. 'Digital Turn, Digital Geographies?' *Progress in Human Geography*, 10.1177/0309132516664800 (2016).

Becker, Konrad, and Felix Stalder. "Deep Search." *Politik des Suchens jenseits von Google. Innsbruck/Wien/Bozen: StudienVerlag* (2009).

Beer, David, 'Power through the Algorithm? Participatory Web Cultures and the Technological Unconscious.' *New Media & Society*, 11 (2009): 985–1002.

Berardi, Franco. *The Uprising : On Poetry and Finance*. Cambridge: The MIT Press, 2012.

Bolukbasi, Tolga, Chang, Kai-Wei, Zou, James, Saligrama, Venkatesh and Kalai, Adam. 'Quantifying and Reducing Stereotypes in Word Embeddings.' 1606.06121 (2016).

Brunton, Finn and Nissenbaum, Helen. *Obfuscation : A User's Guide for Privacy and Protest*. Cambridge: The MIT Press, 2015.

Caliskan-islam, Aylin, Bryson, Joanna J. and Narayanan, Arvind. 'Semantics Derived Automatically from Language Corpora Necessarily Contain Human Biases.' 1608.07187 (2016).

De Certeau, Michel. *The Practice of Everyday Life*. Berkeley: University of California Press, 1988.

Crawford, Kate. 'The Hidden Biases in Big Data', *HBR Blog Network 1* (2013).

Cresswell, Tim. *Geographic Thought : A Critical Introduction*. Wiley-Blackwell, 2013.

Derrida, Jacques. *Limited Inc*. IL: Northwestern University Press, 1988.

Dourish, Paul. 'Algorithms and Their Others: Algorithmic Culture in Context.' *Big Data & Society*, 3, No.2 (2016).

Elmer, Greg. 'Robots.txt : The Politics of Search Engine Exclusion.' In *The Spam Book: On Viruses, Porn, and Other Anomalies from the Dark Side of Digital Culture*, edited by Jussi Parikka and Tony D. Sampson, 217-227. New Jersey: Hampton Press, 2009.

Enloe, Cynthia. 'Womenandchildren: Making Feminist Sense of the Persian Gulf Crisis.' *The Village Voice*, 25, No. 9 (1990).

Fairclough, Norman. *Language and Power*. Routledge, 2013.

Feuz, Martin, Matthew Fuller, and Felix Stalder. "Personal Web searching in the age of semantic capitalism: Diagnosing the mechanisms of personalisation." *First Monday* 16, no. 2 (2011).

Fuchs, Christian, and Vincent Mosco. *Marx in the age of digital capitalism*. Brill, 2015.

Fuller, Matthew, and Goffey, Andrew. 'Toward an Evil Media Studies.' In *The Spam Book: On Viruses, porn, and Other Anomolies from the Dark Side of Digital Culture*, edited by Jussi Parikka and Tony D. Sampson, 141-159. New Jersey: Hampton Press, 2009.

Garnett, Phil, 'Vectorising the Human' (forthcoming).

Gillespie, Tarleton. 'The Relevance of Algorithms.' In *Media Technologies: Essays on Communication, Materiality, and Society*, edited by Tarleton Gillespie, Pablo J. Boczkowski and Kirsten A. Foot, 167-193. Cambridge: The MIT Press, 2014.

Graham, Stephen D. N. 'Software-Sorted Geographies.' *Progress in Human Geography*, 29, no. 5 (2005): 562–580.

Hayles, N. Katherine. *My Mother Was a Computer: Digital Subjects and Literary Texts*. University of Chicago Press, 2005.

Hess, Aaron. 'Reconsidering the rhizome: A textual analysis of web search engines as gatekeepers of the internet.' In *Web search: Multidisciplinary perspectives*. Edited by Amanda Spink and Michael Zimmer. Springer Science & Business Media (2008): 35-50.

Hoy, Dan. 'The Virtual Dependency of the Post-Avant and the Problematics of Flarf: What Happens when Poets Spend Too Much Time Fucking Around on the Internet.' *Jacket Magazine*, 29 (2006).

Jameson, Fredric. *The Prison-House of Language: A Critical Account of Structuralism and Russian Formalism*. Princeton University Press, 1974.

Kaplan, Frederic. 'Linguistic Capitalism and Algorithmic Mediation.' *Representations* 127, no. 1 (2014): 57–63.

Kitchin, Rob and Dodge, Martin. *Code/space: Software and Everyday Life*. Cambridge: MIT Press, 2011.

Kitchin, Rob. 'Thinking Critically about and Researching Algorithms.' *Information, Communication & Society* 20, No.1 (2017): 1-16.

Krebs, Brian. *Lorem Ipsum: Of Good & Evil, Google & China*. Accessed September 11th, 2017. https://krebsonsecurity.com/2014/08/lorem-ipsum-of-good-evil-google-china/.

Kwan, Mei-Po. 'Algorithmic Geographies: Big Data, Algorithmic Uncertainty, and the Production of Geographic Knowledge.' *Annals of the American Association of Geographers*, 106, no.2 (2016): 274-282.

Langville, Amy N. and Meyer, Carl D. *Google's PageRank and beyond: The Science of Search Engine Rankings*. Princeton University Press, 2011.

Lefebvre, Henri. *The production of space*. Vol. 142. Blackwell: Oxford, 1991.

Lim, Ming. 'Postmodern Paradigms and Brand Management in the 'Search' Economy.' *International Journal of Internet Marketing and Advertising* 5, No. 1-2 (2008): 4-16.

Lyotard, Jean-Francois. *The Postmodern Condition: A Report on Knowledge.* Manchester University Press, 1984.

Mayer, Sophie. '(Re)Search/Destroy? Or, Mistaking Poetry for Viruses.' *Archive of the Now*. Accessed October 19, 2016. http://www.archiveofthenow.org/pages/researchdestroy-or-mistaking-poetry-for-viruses.

Merrifield, Andy. "Amateur urbanism." *City* 19, no. 5 (2015): 753-762.

Metahaven. 'Peripheral Forces: On the Relevance of Marginality in Networks.' In *Deep Search: The Politics of Search Engines beyond Google*, 185-197. Edited by Konrad Becker and Felix Stalder. Strange Chemistry, 2009.

Mikolov, Tomas, Quoc V. Le, and Ilya Sutskever. "Exploiting similarities among languages for machine translation." *arXiv preprint arXiv:1309.4168* (2013).

Nabugodi, Mathelinda. 'Pure Language 2.0: Walter Benjamin's Theory of Language and Translation Technology.' *Feedback*. (2014). Accessed 30 October, 2016. http://openhumanitiespress.org/feedback/literature/pure-language-2-0-walter-benjamins-theory-of-language-and-translation-technology.

Neyland, Daniel. 'On Organizing Algorithms.' *Theory, Culture & Society* 32, no.1 (2014): 119-132.

Pasquale, Frank. *The Black Box Society*. Harvard University Press, 2015.

Pasquinelli, Matteo. 'Google's PageRank Algorithm: A Diagram of Cognitive Capitalism and the Rentier of the Common Intellect.' In *Deep Search: The Politics of Search Engines beyond Google*, 152-162. Edited by Konrad Becker and Felix Stalder. Innsbruck: Studien Verlag, 2009.

Rogers, Richard. 'The Googlization question, and the inculpable engine.' In *Deep search: The Politics of Search Engines beyond Google*, 173-184. Edited by Konrad Becker and Felix Stalder. Innsbruck: Studien Verlag, 2009.

Shaw, Joe, and Mark Graham. "An Informational Right to the City? Code, Content, Control, and the Urbanization of Information." *Antipode* (2017).

Srnicek, Nick. *Platform capitalism*. John Wiley & Sons, 2016.

Straube, Till. 'Stacked Spaces: Mapping Digital Infrastructures.' *Big Data & Society* 3, no.2 (2016): 2053951716642456.

Thornton, Pip. '{poem}.py: a critique of linguistic capitalism'. Accessed September 11, 2017. https://linguisticgeographies.com/2016/06/12/poem-py-a-critique-of-linguistic-capitalism/

Thrift, Nigel and French, Shaun. 'The Automatic Production of Space.' *Transactions of the Institute of British Geographer*s 27 (2002): 309–335.

Wark, Mckenzie. 'Benjamedia.' *Public Seminar* (2015). Accessed on 30 October, 2016. http://ww.publicseminar.org/2015/08/benjamedia/#.WFB6yfP0wwI.

Winseck, Dwayne. 'Netscapes of Power: Convergence, Consolidation and Power in the Canadian Mediascape.' *Media, Culture & Society* 24, no.6 (2002): 795–819.

Zook, Matthew A., and Graham, Mark. 'The Creative Reconstruction of the Internet: Google and the Privatization of Cyberspace and DigiPlace.' *Geoforum* 38 (2007): 1322–43.

**Acknowledgements**

**Notes**

1. Jean-Francois Lyotard, *The Postmodern Condition: A Report on Knowledge*. Manchester University Press, 1984, 46-7. ↵
2. James Ash, Rob Kitchin and Agnieszka Leszczynski. 'Digital Turn, Digital Geographies?' *Progress in Human Geography*, 10.1177/0309132516664800 (2016). ↵
3. for example Graham, Stephen D. N. 'Software-Sorted Geographies.' *Progress in Human Geography*, 29, no. 5 (2005): 562–580 ; Matthew A. Zook and Mark Graham, 'The Creative Reconstruction of the Internet: Google and the Privatization of Cyberspace and DigiPlace.' *Geoforum* 38 (2007): 1322–43. ↵
4. Rob Kitchin and Martin Dodge, *Code/space: Software and Everyday Life*. Cambridge: MIT Press, 2011. ↵
5. Mei-Po Kwan, 'Algorithmic Geographies: Big Data, Algorithmic Uncertainty, and the Production of Geographic Knowledge.' *Annals of the American Association of Geographers*, 106, no.2 (2016): 274-282; Paul Dourish, 'Algorithms and Their Others: Algorithmic Culture in Context.' *Big Data & Society*, 3, No.2 (2016); Daniel Neyland, Neyland, Daniel. 'On Organizing Algorithms.' *Theory, Culture & Society* 32, no.1 (2014): 119-132; Tarleton Gillespie, 'The Relevance of Algorithms.' In *Media Technologies: Essays on Communication, Materiality, and Society*, edited by Tarleton Gillespie, Pablo J. Boczkowski and Kirsten A. Foot, 167-193. Cambridge: The MIT Press, 2014. ↵
6. Louise Amoore, 'Cloud Geographies: Computing, Data, Sovereignty.' *Progress in Human Geography*, 10.1177/0309132516662147 (2016), 11. ↵
7. Till Straube, 'Stacked Spaces: Mapping Digital Infrastructures.' *Big Data & Society* 3, no.2 (2016): 2053951716642456, 6. ↵
8. Srnicek, Nick. *Platform capitalism*. John Wiley & Sons, 2016; Fuchs, Christian, and Vincent Mosco. *Marx in the age of digital*

*capitalism*. Brill, 2015; Feuz, Martin, Matthew Fuller, and Felix Stalder. "Personal Web searching in the age of semantic capitalism: Diagnosing the mechanisms of personalisation." *First Monday* 16, no. 2 (2011); Kaplan, Frederic. "Linguistic capitalism and algorithmic mediation." *Representations* 127, no. 1 (2014): 57-63. ↵

9. Merrifield, Andy. "Amateur urbanism." City 19, no. 5 (2015): 753-762. See also Lefebvre, Henri. *The production of space*. Vol. 142. Blackwell: Oxford, 1991, 27. ↵

10. Nigel Thrift and Shaun French, 'The Automatic Production of Space.' *Transactions of the Institute of British Geographers* 27 (2002): 309–335, 309. ↵

11. This is an idea inspired by the chapter 'Reading as Poaching' in Michel De Certeau, *The Practice of Everyday Life*. Berkeley: University of California Press, 1988, 165-176. ↵

12. Michel De Certeau, *The Practice of Everyday Life*. Berkeley: University of California Press, 1988, 134-5. ↵

13. https://www.google.co.uk/insidesearch/howsearchworks/crawling-indexing.html [accessed 16th October 2016. ↵

14. although not a database in the strict sense of the word, this phrase attempts to describe the quantity of indexed data on which search algorithms are able to act. ↵

15. Kate Crawford, 'The Hidden Biases in Big Data', *HBR Blog Network* 1 (2013). ↵

16. Other examples include autocompletions changing *she invented to he invented*, 'Google: Did You Mean: "He invented"?', accessed October 16, 2016, http://www.michaelzimmer.org/2007/05/09/google-did-you-mean-he-invented/ and the 'UN Women ad series reveals widespread sexism', accessed October 16, 2016, http://www.unwomen.org/en/news/stories/2013/10/women-should-ads ↵

17. In the aftermath of the 2016 US presidential election, both Google and Facebook have been forced to acknowledge the potential impact of politically biased fake news stories or advertising click-bait hosted on their platforms. https://www.ft.com/content/2910a7a0-afd7-11e6-a37c-f4a01f1b0fa1. [Accessed October 11, 2016. ↵

18. Cynthia Enloe, 'Womenandchildren: Making Feminist Sense of the Persian Gulf Crisis.' *The Village Voice*, 25, No. 9 (1990). ↵

19. Norman Fairclough, *Language and Power*. Routledge, 2013. ↵

20. For example see Frank Pasquale, *The Black Box Society*. Harvard University Press, 2015 and Rob Kitchin, 'Thinking Critically about and Researching Algorithms.' *Information, Communication & Society* 20, No.1 (2017): 1-16. ↵

21. Amy N. Langville and Carl D. Meyer, *Google's PageRank and beyond: The Science of Search Engine Rankings*. Princeton University Press, 2011, 6. ↵

22. Aylin Caliskan-islam, Joanna J Bryson and Arvind Narayanan, 'Semantics Derived Automatically from Language Corpora Necessarily Contain Human Biases.' 1608.07187 (2016), 1–14. ↵

23. ibid, 3. ↵

24. Tolga Bolukbasi et al., 'Quantifying and Reducing Stereotypes in Word Embeddings.' 1606.06121 (2016). ↵

25. ibid, 41. ↵

26. Lev Manovich, 'Database as Symbolic Form', 'Convergence: The International Journal of Research into New Media Technologies', quoted in 'The Relevance of Algorithms.' In *Media Technologies: Essays on Communication, Materiality, and Society*, edited by Tarleton Gillespie, Pablo J. Boczkowski and Kirsten A. Foot, 167-193. Cambridge: The MIT Press, 2014, 169 ↵

27. ibid, 3. ↵

28. http://now.howstuffworks.com/2016/01/13/mordor-google-translate. Accessed 19 October, 2016. ↵

29. http://www.theguardian.com/technology/2016/jan/07/google-translates-russia-mordor-foreign-minister-ukrainian. Accessed 19 October, 2016. ↵

30. http://www.theguardian.com/technology/2016/jan/07/google-translates-russia-mordor-foreign-minister-ukrainian. Accessed 19 October, 2016. ↵

31. See for example http://www.e-ir.info/2015/05/05/russia-as-ukraines-other-identity-and-geopolitics/. Accessed 19 October, 2016 ; http://www.bbc.co.uk/news/world-europe-34037743. Accessed 19 October, 2016, and Leonid Bershidsky, Putin's Russia, Tolkien's Mordor: What's the Difference? https://www.bloomberg.com/view/articles/2014-12-11/putins-russia-tolkiens-mordor-whats-the-difference. Accessed 19 October, 2016. Bershidsky also notes that a translator of Tolkien's text suggesting that Saurian was based on Josef Stalin – a suggestion refuted by Tolkien himself, as well as Putin's apparent resemblance to Sauron, and mentions a rumour that "a design company was planning to light up an enormous Eye of Sauron over a Moscow skyscraper" in 2012. ↵

32. http://i100.independent.co.uk/article/something-strange-is-going-on-with-google-maps–gJxfcnXWhg. Accessed 19 October, 2016. ↵

33. http://www.searchenginepeople.com/blog/incredible-google-bombs.html. Accessed 19 October, 2016. ↵

34. Krebs, Brian. *Lorem Ipsum: Of Good & Evil, Google & China*. Accessed September 11th, 2017. https://krebsonsecurity.com/2014/08/lorem-ipsum-of-good-evil-google-china/. ↵

35. Mikolov, Tomas, Quoc V. Le, and Ilya Sutskever. "Exploiting similarities among languages for machine translation." *arXiv preprint arXiv:1309.4168* (2013). ↵

36. A British Library webpage states how Johnson's definition of the word oats is 'very rude to the Scots. He defines the word as "A Grain, which in England is generally given to horses, but in Scotland supports the people"'. Accessed October 19, 2016. http://www.bl.uk/learning/langlit/dic/johnson/1755johnsonsdictionary.html. ↵

37. For a comprehensive 'storyfied' report of the controversy see https://www.buzzfeed.com/laurasilver/go-on-call-me-a-rabid-feminist?utm_term=.blPw5NR2j#.goWOEVG1J. ↵

38. https://en.oxforddictionaries.com/help. Accessed January 22, 2016. The content of the site has now changed. ↵

39. Derek Attig @bookmobility, Twitter.com, January 22, 2016. ↵

40. Greg Elmer, 'Robots.txt : The Politics of Search Engine Exclusion.' In *The Spam Book: On Viruses, Porn, and Other Anomalies from the Dark Side of Digital Culture*, edited by Jussi Parikka and Tony D. Sampson, 217-227. New Jersey: Hampton Press, 2009. ↵

41. Finn Brunton and Helen Nissenbaum, *Obfuscation : A User's Guide for Privacy and Protest*. Cambridge: The MIT Press, 2015, 1. ↵

42. ibid, 55. ↵

43. Michel De Certeau, *The Practice of Everyday Life*. Berkeley: University of California Press, 1988, 30. ↵

44. Matthew Fuller and Andrew Goffey, 'Toward an Evil Media Studies.' In *The Spam Book: On Viruses, porn, and Other Anomolies from the Dark Side of Digital Culture*, edited by Jussi Parikka and Tony D. Sampson, 141-159. New Jersey: Hampton Press, 2009, 153. ↵

45. Tarleton Gillespie, 'The Relevance of Algorithms.' In *Media Technologies: Essays on Communication, Materiality, and Society*, edited by Tarleton Gillespie, Pablo J. Boczkowski and Kirsten A. Foot, 167-193. Cambridge: The MIT Press, 2014, 184. ↵

46. Michel De Certeau, *The Practice of Everyday Life*. Berkeley: University of California Press, 1988, 30. ↵

47. Dan Hoy, 'The Virtual Dependency of the Post-Avant and the Problematics of Flarf: What Happens when Poets Spend Too Much Time Fucking Around on the Internet.' *Jacket Magazine*, 29 (2006). ↵
48. Sophie Mayer, '(Re)Search/Destroy? Or, Mistaking Poetry for Viruses.' *Archive of the Now.* Accessed October 19, 2016. http://www.archiveofthenow.org/pages/researchdestroy-or-mistaking-poetry-for-viruses. ↵
49. Jean-Francois Lyotard adapted Wittgenstein's idea of language games (*Philosophical Investigations, Oxford Blackwell*, 1953) in *The Postmodern Condition: A Report on Knowledge*. Manchester University Press, 1984. ↵
50. Pip Thornton, '{poem}.py: a critique of linguistic capitalism'. Accessed September 11, 2017. https://linguisticgeographies.com/2016/06/12/poem-py-a-critique-of-linguistic-capitalism/ ↵
51. N. Katherine Hayles, *My Mother Was a Computer: Digital Subjects and Literary Texts*. University of Chicago Press, 2005, 39. ↵
52. Matteo Pasquinelli, 'Google's PageRank Algorithm: A Diagram of Cognitive Capitalism and the Rentier of the Common Intellect.' In *Deep Search: The Politics of Search Engines beyond Google*, 152-162. Edited by Konrad Becker and Felix Stalder. Innsbruck: Studien Verlag, 2009, 157. ↵
53. Matthew Fuller and Andrew Goffey, 'Toward an Evil Media Studies.' In *The Spam Book: On Viruses, porn, and Other Anomolies from the Dark Side of Digital Culture*, edited by Jussi Parikka and Tony D. Sampson, 141-159. New Jersey: Hampton Press, 2009, 153. ↵
54. "Rather than connecting through random intersecting points of the rhizome… the rhizome has grown trees." Aaron Hess, 'Reconsidering the rhizome: A textual analysis of web search engines as gatekeepers of the internet.' In *Web search: Multidisciplinary perspectives*. Edited by Amanda Spink and Michael Zimmer. Springer Science & Business Media (2008), p.35. ↵
55. Richard Rogers, 'The Googlization question, and the inculpable engine.' In *Deep search: The Politics of Search Engines beyond Google*, 173-184. Edited by Konrad Becker and Felix Stalder. Innsbruck: Studien Verlag, 2009, 176. ↵
56. Mathelinda Nabugodi, 'Pure Language 2.0: Walter Benjamin's Theory of Language and Translation Technology.' *Feedback*. (2014). Accessed 30 October, 2016. http://openhumanitiespress.org/feedback/literature/pure-language-2-0-walter-benjamins-theory-of-language-and-translation-technology. Accessed 19 October 2016. ↵
57. Phil Garnett, Vectorising the Human (forthcoming). ↵
58. Mckenzie Wark, 'Benjamedia', *Public Seminar*, 2015. ↵
59. Michel De Certeau, *The Practice of Everyday Life*. Berkeley: University of California Press, 1988, 165. ↵
60. ibid, 37. ↵
61. Jacques Derrida, *Limited Inc, Context*, 1988, 12. ↵
62. Ming Lim, 'Postmodern Paradigms and Brand Management in the "Search" Economy', *International Journal of Internet Marketing and Advertising*, 2008, 10. ↵
63. Hayles, N. Katherine. *My Mother Was a Computer: Digital Subjects and Literary Texts*. University of Chicago Press, 2005, 16. ↵
64. "To arrest the meaning of words once and for all, that is what the Terror wants", quoted in Michel De Certeau, *The Practice of Everyday Life*. Berkeley: University of California Press, 1988, 165. ↵
65. Metahaven, 'Peripheral Forces: On the Relevance of Marginality in Networks.' In *Deep Search: The Politics of Search Engines beyond Google*, 185-197. Edited by Konrad Becker and Felix Stalder. Strange Chemistry, 2009, 189. ↵
66. Frederic Kaplan, 'Linguistic Capitalism and Algorithmic Mediation.' *Representations* 127, no. 1 (2014), 1. ↵
67. Franco Berardi, *The Uprising : On Poetry and Finance*. Cambridge: The MIT Press, 2012, 158. ↵
68. Gunnar Olsson, 'Of Ambiguity', in Ley, D. and Samuels (eds), *1978: Humanistic geography: prospects and problems*. Chicago: Maaroufa Press (1978), 110, quoted in Tim Cresswell, *Geographic Thought : A Critical Introduction*. Wiley-Blackwell, 2013, 189. ↵
69. Michel De Certeau, *The Practice of Everyday Life*. Berkeley: University of California Press, 1988, 36. ↵
70. ibid, 40. ↵
71. Fredric Jameson, *The Prison-House of Language: A Critical Account of Structuralism and Russian Formalism*. Princeton University Press, 1974. ↵