

Saukko, Paula. (forthcoming, 2017) Shifting metaphors in direct-to-consumer genetic testing: From genes as information to genes as big data, *New Genetics and Society* (version accepted for publication)

Shifting metaphors in direct-to-consumer genetic testing: From genes as information to genes as big data

This article analyses shifts in metaphors in direct-to-consumer genetic testing, analysing the websites and select media coverage of the nutrigenetic testing company Sciona (2000-2009) and the personal genome service 23andMe (2006--). Sciona represented genes and communication through the classical metaphor of information; genes coded for disease, and this information was transmitted from the expert company to the consumers. 23andMe represented genes and communication through a new metaphor of big data; genes were digital data or a resource that was browsed, correlated with other data, uploaded and retrieved across lay customers, websites and companies. In terms of understanding health 23andMe tests and research still cast genes as coding for disease to be mitigated by lifestyle change and targeted drugs. However, rendering genes digital data or resources changed their social and economic meaning; genes could be circulated, shared and traded, which legitimised 23andMe's business model of consumer genetics and private biobanking.

This article examines metaphors used by two direct-to-consumer (DTC) genetic testing companies; the nutrigenetic testing company Sciona (2000-2009) and 23andMe, which doubles up as a consumer genetics company and private biobank with over one million (23andMe, 2015) DNA samples and customers (2006--). I will argue that the companies represented genes and communication between the company and its customers through the two distinct ontological metaphors of information and big data.

There is a significant body of literature on metaphors used in science. It has been noted that scientists use metaphors to popularise their findings (Nelkin, 1994) but, as pointed out by Lakoff and Johnson (Lakoff & Johnson, 2008), metaphors are not innocent but typically legitimate specific agendas, such as bids for funding scientific projects, such as the Human Genome Project (Nerlich & Hellsten, 2004), or the creation of markets, such as the creation of data-driven markets for health (Ruckenstein & Pantzar, 2015).

Nelkin's classic analysis of metaphors of genes in science and popular culture argued that genes were frequently cast as the "master code," which scientists could "decipher" and "read" to predict human "fates" as if through a "medical crystal ball" (Nelkin, 1994). The historical ramifications of this deterministic vision of genes as the "blueprint" or "book of life" was studied by Kay (Kay, 2000), who noted that the metaphor of information underpinning Francis Crick's formulation of the "central dogma" in genetics according to which DNA codes for RNA, which codes for proteins originated from Shannon's classical theory of communication (Shannon & Weaver,

1949). Shannon, an engineer in the Bell telephone company, envisioned communication in terms of information transmission from source to receiver, with the aim to analyse and improve the accuracy of this transmission periled by entropy and noise. Kay (Kay, 2000) chronicled how the interpersonal and institutional connections between Shannon and his co-author Warren Weaver and the early geneticists in the Rockefeller Foundation facilitated the transfer of the transmission model from communication to life sciences.

The central dogma has been criticised for overplaying the role of genes as the origin of life processes (Keller, 2009). Moss has argued that the idea that genes predict the development of disease applies to rare monogenic conditions, such as Huntington's chorea, but tends to get conflated with the vast majority of diseases, which are polygenic, caused by complex interactions between environmental factors and genes (Moss, 2004). Shannon's model of communication in terms of linear transmission was also discredited for envisioning audiences as passive recipients (Hall, 1980). As of late, research has argued that current Web 2.0 technologies, such as Wikipedia, have ushered a more participatory communication, where all users can take part in knowledge creation and dissemination (O'Reilly, 2007).

The metaphor of genes as blueprint has, however, persevered and was still mostly used in scientific and media reporting on the Human Genome Project, even if scientists found that the human genome had surprisingly few genes, raising doubts about whether single genes could contain the programs for traits (Nerlich & Hellsten, 2004). New scientific developments, such as epigenetics, have complicated the

central dogma and metaphors associated with it. Landecker and Panofsky (2013) have argued that epigenetics represents a more sociological understanding of genes, as it is interested in how social forces (pollution, nutrition, mothering) affect gene expression. Stelmach and Nehrlich (Stelmach & Nerlich, 2015), indeed, observed that UK media coverage of epigenetics came up with new metaphors, such as framing genes as “switches,” which could be turned on and off, and in terms of music, where epigenetic processes “played” genes differently.

When closely reading the 23andMe’s website and interviews with its co-founder Anne Wojcicki, I noticed that genes were often referred to as data, raw data and being abundant; often genes were cast as data by the use of verbs, so genes were browsed, googled, uploaded, retrieved, collected and crowdsourced. Thus, genes were not represented as information or a code to be deciphered or cracked by scientists but cast as abundant natural or raw digital resource that could be searched, analysed and pooled for different purposes by scientists, lay customers and companies. These references to abundant, raw data that could be analysed by everyone for correlations, repeated key premises of the emergent discourse on big data (Kitchin, 2014).

A methodological point is due here. Most research on genes and metaphors refer to what Lakoff and Johnson call conceptual metaphors -- so genes are *like* a book (Lakoff & Johnson, 2008). Understanding genes in terms of information transmission or big data are not necessarily conceptual metaphors. Information transmission can be seen as a historical discourse (Kay, 2000), but information transmission and big

data can also be defined as ontological metaphors, which seek to render a vague phenomenon or experience a discrete object that has specific aspects, can be referred to and possibly quantified, reasoned and acted upon (Lakoff & Johnson, 2008).

The nature and implications of the information transmission metaphor are well rehearsed in terms of leading to a gene-centric understanding of human behaviour and health (Keller, 2009). Big data, however, is an emergent concept with alleged ontological characteristics. Social scientists have identified and critically examined key alleged aspects of big data. The first aspect of big data resonates with the colloquial meaning of data as “facts and statistics collected together for reference or analysis” (Oxford Living Dictionaries, 2016), which implies data are raw material from which analyses or interpretations can be made. The promise of big data is argued to be that digital devices enable the collection of previously unforeseen quantities of naturally occurring data, such as clicks, which are not marred by researcher bias (Kitchin, 2014). Second, big data supposedly renders theories obsolete, as the sheer quantity of data can reveal correlations, which emerge straight out of data (Kitchin, 2014; Ruppert, Law, & Savage, 2013). This aspect of big data is encapsulated in the slogan “correlation is everything” proclaimed by Chris Anderson, former editor in chief at Wired magazine (Anderson, 2008), and big data is seen as capable of revealing surprising correlations between diverse or heterogenous data (Kitchin, 2014; Ruppert et al., 2013). The third alleged feature of big data are that it can be interpreted by anyone with tools available online, yielding to multiple interpretations (Ruppert et al., 2013).

Lakoff and Johnson note that ontological metaphors do not seem metaphors at all, rather the aspects of the phenomenon simply seem like its natural or self-evident qualities (Lakoff & Johnson, 2008). So, it seems as if genes simply code for traits or constitute big data to be browsed for insights and correlations by whomever. The way in which ontological metaphors naturalise certain qualities of a phenomenon render them close to what Barthes calls myth, a second order of signification, where connotations and denotations cohere around a core belief system or ideology (Barthes, 1972), and big data has been argued to have qualities of myth (Boyd & Crawford, 2012) and ideology (Van Dijck, 2014).

However, metaphors and ideologies are not uncontested. Recent research on 23andMe has pointed out how the company mobilises the rhetoric of sharing, open access and gift to legitimise extracting DNA and survey data from customers to be sold for private gain (Harris, Kelly, & Wyatt, 2016; Van Dijck & Poell, 2016).

O’Riordan has noted how 23andMe frames its customers as publics who read and write about their genomes, whilst obfuscating how the focus on DNA constrain these practices (O’Riordan, 2013), and Kragh-Furbo and Tutton have observed how the company has turned customers’ saliva into a promissory substance, which is easy to circulate in the bioeconomy (Kragh-Furbo & Tutton, 2017). I agree with the observations of these authors. My contribution to this on-going scholarly discussion is to contextualise the rhetoric used by 23andMe within the body of literature on genes and metaphors as well as in the history of DTC genetic testing in the past 17 years. This contextualisation allows me to pinpoint continuities between the two

apparently quite different companies in that the way in which both of them represented genes was underpinned by the metaphor of transmission of information, whereby genes code for risk of disease, which can be deciphered by scientists and mitigated by lifestyle change or targeted drugs. However, 23andMe also mobilised the metaphor of big data to render genes a natural or raw digital resource, which could be circulated, shared and taken up by devices across numerous sites transcending institutional boundaries (Ruppert et al., 2013) to legitimise its construction of a private biobank with over a million DNA samples and customers. The analysis that follows will show how the metaphors used by the two companies to make sense of genes and their relationships with their customers are -- confusingly enough -- both fundamentally similar and fundamentally different.

Methods

This article derives from a larger project, which started in 2006 and examined the marketing and regulation of nutrigenetic testing companies (Saukko, Reed, Britten, & Hogarth, 2010). The project expanded to include the 23andMe's personal genome service in 2008.

Sciona and 23andMe

I chose to focus on Sciona, as it was the first DTC genetic testing company in the UK and was the foci of regulatory debates both in the UK and later in USA. The company was established in 2000 in London, moved to USA in 2004 and ceased to exist in 2009. Sciona sold nutrigenetic tests for a limited number of gene variants, purportedly associated with diet, and offered genetically tailored advice on nutrition

and micronutrients. The tests were mostly sold online, but the results and advice were communicated through a brochure sent to customers by post. Sciona was criticised for making scientifically invalid predictions and discussed by regulatory bodies, including the UK Human Genetics Commission (Human Genetics Commission, 2003) and the US Government Accountability Office (US Government Accountability Office, 2006).

23andMe was founded in 2006 by Linda Avey and Anne Wojcicki, ex-wife of Sergey Brin, founder of Google. The company offered a “personal genome service” (PGS), and due to the development of increasingly efficient microarray technologies (Baker, 2013) was able to test for a large variety of genes (a total of 254 in 2013—as counted from the test itself), associated with health, traits, drug response and ancestry. 23andMe also sold its test online and uploaded the results onto an interactive platform. From the start 23andMe invited its customers to take part in research and sold customers’ DNA and survey responses to researchers and private companies. In 2015 23andMe announced that it had genotyped a million customers (23andMe, 2015).

23andMe was also criticised by the GAO (US Government Accountability Office, 2010), and in 2013 the US Food and Drug Administration (FDA) banned 23andMe from selling health-related tests for failing to provide evidence to support its marketing claims (Food and Drug Administration, 2013). After the ban the company continued to sell ancestry tests and what it termed raw data, a selection of genotyped single nucleotide polymorphisms (SNPs or ‘letter changes’), which

customers could upload onto online workaround software, such as Promethease, which linked the SNPs to disease risks.

In 2014 23andMe launched a revamped health-related genetic test in the UK and Canada for being a carrier for a selection monogenic diseases and for some polygenic “risk factors.” In 2015 23andMe relaunched FDA approved health-related genetic tests in the USA for being a carrier for certain monogenic diseases. In 2017 the company gained FDA approval for tests for ten polygenic risk factors.

Material and Analysis

The material analysed consists of old webpages of Sciona, which we retrieved using the Wayback Machine in 2006 in addition to saving the 2006-2008 webpages as pdf's. We also purchased the Sciona test in 2006 and received a 86-page brochure containing the results and advice. We purchased the 23andMe US/UK test in 2008 and bought the UK test in 2015, which gave access to two different sets of online test results and associated interactive features, which both remained active as of 2017; the US test bought in 2008 remained largely unchanged since 2013, when FDA banned the test, whereas the features incorporated into the UK test have slightly changed between 2015 and 2017. We also uploaded the UK test results onto Promethease in 2015 and received a report. The material analysed include the Sciona websites 2001-2008, the 23andMe online results 2008 and 2015 and the Sciona (2006) and Promethease (2015) reports.

This material was complemented with a small selection of media articles. I focused on UK media coverage, as the tests were bought from the UK, and I searched the Lexis Nexis database for articles on the two companies in national newspapers. The database identified 34 articles on Sciona and 435 on 23andMe, and I chose to analyse a small selection of articles covering key moments in the companies' lifespans. I selected the one and only article published when Sciona was launched in 2001 and four from 2003 when the company was scrutinised by the Human Genetics Commission. Again, I selected the one and only article covering 23andMe's launch in 2007 in USA, nine articles from 2014 when the company launched its test in the UK and nine articles focusing on research supported by 23andMe on skin ageing and depression between 2015 and 2017. I also analysed an article from 2013 in MIT Technology Review on Promethease and two interviews with Anne Wojcicki in PLoS Genetics and tech magazine Inc. published in 2015 when the company gained its first FDA approval. The sample does not seek to be representative of the media coverage of the companies, it simply complements the main analysis of the websites and genetic test results by instances where public media and the company representatives describe, criticise, defend or market the tests at key points.

I initially started doing a thematic analysis of the websites of the two companies to ascertain similarities and differences. This is when I noted the metaphors of information transmission and big data that repeated in the material, which refocused my analysis. The analysis was inductive, initially following the constant comparative method (Glaser, 1965); as mentioned I first read the websites and identified how the ontological metaphors (Lakoff & Johnson, 2008) of information transmission and big data repeated. I then read the material and literature on information transmission and

big data back and forth, highlighting instances where the two metaphors were articulated in the text in relation to genes and communication between the company and its customers.

In what follows I will first analyse how Scion's website, results and media coverage represented genes and communication between the company and its customers through the metaphor of information transmission. Second, I will analyse how the 23andMe website and media coverage and Promethease report articulated the metaphors of information transmission and big data. I will focus on how the metaphor of big data was articulated through (i) representing genes as data, (ii) casting genes as data correlated with diverse other data and (iii) the way in genetic data was represented as analysed by diverse actors from lay customers to other companies and software.

Scion and Information Transmission

On an early version of its website in March 2001 Scion's described genes as "governing" metabolism:

"...your genes govern how your metabolic pathways digest and dispose of nutrients and toxins within your body (Scion, 2001)

This idea of genes as governing metabolism was translated into nutritional advice in the test results. So, customers were encouraged to adjust their diet to offset any defects in their genes. This is illustrated by the recommendations given based on the test results for six genes associated with “detoxification” in the customer report we received from Sciona in 2006:

“You have variations in your genes important for antioxidant defences which may lead to less efficient removal of free radical damage from your body. ... Increase your consumption of foods high in antioxidants such as vitamins A, C and E.”

Thus, on Sciona’s website genes were cast as “governing” metabolism, implying the idea that genes code for beneficial or harmful traits, such as more or less efficient removal of free radicals, which could be offset by taking specific micronutrients.

The news report on Sciona’s launch represented genes as “responsible for everything,” including “diseases you are susceptible to” (Cooke, 2001). The article went on to describe how in the test “your DNA is checked to see if your genes have adapted to cope with a certain diet,” continuing that if “your distant ancestors were Eskimos” you might “need to eat more oily fish” (Cooke, 2001). In this media report genes were cast particularly deterministically (being “responsible for” disease) and evolution was evoked to tell a story of how variations in genes related to diet were due to adaptation and, therefore, we should eat the diet of our ancestors.

The metaphor of information transmission also shaped how Sciona envisioned communication about nutrigenetics. In March 2001 the company presented itself as seeking to educate its customers and to ensure the faithful transmission of accurate scientific knowledge:

“ ... because much of the science and be impenetrable to the layperson and the debates sensationalist. The service will include a publishing component that educates the market by providing access to the underlying scientific principles, research and trials.”

In these early texts Sciona adopted a fairly paternalistic view of its customers as needing educating by the company. This understanding was reinforced by, for example, an extensive “glossary” of scientific terms, such as “polymorphism,” related to genetics included in the webpages.

The gist of Sciona’s marketing was that its genetics based nutritional advice would “motivate” behaviour change, undergirded by the psychological version of the transmission metaphor: stimulus (genetic risk information) triggering response (behaviour change). As illustrated on its question and answer section in 2004:

“... the people who will find the Sciona service of considerable value are those who need an extra bit of motivation to maintain healthy habits or those who find it easier and more realistic to focus on areas particularly important for them ...”

Further, the notion of genes and communication in terms of information transmission not only informed symbolic metaphors in marketing and regulatory texts but also the material structure of Sciona’s digital platform. The company’s website offered information for its customers and the test results were sent directly to them. The website had no interactive features that would allow customers to give feedback, have a conversation or tinker with their genetic results or data. The Sciona website, thus, conformed to what has been termed Web 1.0 design, predicated on transmitting authoritative information from sender to receiver (O’Reilly, 2007).

Starting in 2002 and particularly in 2003 Sciona came under attack from the UK Human Genetics Commission for selling tests that were not scientifically valid. Media reporting of these debates used the metaphor of “health horoscopes” (Barnett, 2003; Thornton, 2003) to describe the tests, evoking the domain of superstition, so that the tests were not associated with scientific predictions of future but illegitimate fortune telling. A news report lamented that “I thought I could find out if I had an obesity or depression gene” but “all I got was healthy eating advice” (Thornton, 2003), whereas another news article debated the “ethics” of knowing “how you will die and when,” referring to the tests as “genetic time bombs” (Barnett, 2003). Whilst both news reports evoked critical metaphors of horoscopes and time bomb in relation to genes,

they still confirmed the underlying ontological metaphor of genes as transmitting information, just communicating disappointment that the nutrigenetic tests did not live up to this ideal or that such predictions might raise ethical issues.

Browsing Genes

23andMe's website and its media coverage frequently cast genes in a different way as data, using verbs such as browse, explore, search or google. This was illustrated by the only UK piece of news in the Lexis-Nexis database on 23andMe's launch (in the USA) in 2007, which was titled "Search engine aims to Google your genes" and reported:

Google has expanded its mission to lay bare the world's information by investing in a company set up by its co-founder's wife that lets users trawl their genetic profile online (Blakeley, 2007)

In this article genes were something that could be digitally searched (googled) or fished with a net (trawled), implying genes were bits of data or a natural resource, which could be explored. Furthermore, the article implied that the data or resource could be exploited, referring to Google "laying bare" this information and googling customers' genes, hinting at invading and making profit out of private data. Similarly, the reference to "online genetic profiles" played with the association between genetic profiles and social media profiles associating genes with the domain of the digital.

This idea of genes as a digital resource to be explored and exploited mapped onto the discourse and debates on big data (Kitchin, 2014).

23andMe's website represented genes as data most explicitly in the sections encouraging customers to analyse their "genome." The company offered its customers what it described as "raw data" or "uninterpreted raw genotype data, including data that is not used in 23andMe reports," which referred to single nucleotide polymorphisms (SNPs) identified by the company's microarray. The references to "raw" and "uninterpreted" made the data seem unprocessed or natural, harking back to the alleged premise of big data as being naturally occurring and, thus, free of researcher bias (Ruppert et al., 2013).

Throughout the 23andMe website the company used verbs, such as "navigate," "browse," "upload," "share" and "compare," in conjunction with genes, representing them as data. Similarly, the Promethease software described itself as a service to "retrieve" information about the genome. Metaphors typically operate across domains (Lakoff & Johnson, 2008; Nerlich & Hellsten, 2004), and the verbs used illustrated the way in which genes were likened not to information (Kay, 2000) but to digital data. Whereas information is typically interpreted as having one accurate meaning to be deciphered by experts, digital data is cast as dispersed or "circulated, shared and taken up by devices and across numerous websites" (Ruppert et al., 2013). This rhetorical move, casting genes as mobile, digital data legitimised and reflected the way in which 23andMe moved customer DNA out of the laboratory and

the clinic to be circulated, shared and traded in the digital domain to enable its dual business model of consumer genetics and live private biobank.

However, in other parts of 23andme's website the representation of genes was underpinned by the classic, deterministic information transmission metaphor, the genetic risk profiles of 2008 stating that "your test results indicate you are at increased risk for atrial fibrillation based on genetics" and adding that "the heritability of atrial fibrillation is estimated to be 62%." However, the risk profile for lung cancer estimated its "heritability" to be 8-14%, continuing that "environmental factors," such as smoking, contributed "more to differences in risk than genetic factors." In these cases genes were represented as playing a bigger (62%) or smaller (8-14%) role in causing polygenic disease, together with environmental factors. Nevertheless, even if the role of genes was estimated to vary, genes were still cast as coding for disease, which could be prevented with lifestyle changes (quitting smoking).

Understanding genes in terms of information transmission was also evident in the coverage of the 2014 UK launch of 23andMe's health-related tests. The Sun reported how the test let users

... have their DNA screened for genes associated with inherited conditions, such as Alzheimer's, cancer or diabetes. Kits claim to reveal which conditions you are likely to develop (Earle & Quinton, 2014).

In this instance genes were “screened,” which implied expert, medical knowledge, which “revealed” whether an individual will develop cancer or Alzheimer’s. The idea of genes revealing your destiny repeated the same notion of genes containing the code, which determined future, as discussed in previous research (Nelkin, 1994). However, in another piece of news of the same launch the tests were described to “allow users to both browse the raw code of their genome and use tools to investigate their genetic makeup” (Gibbs, 2014), again representing genes as digital data to be explored.

Overall, 23andMe’s website and media coverage shifted between describing genes through the ontological metaphors of information and big data. Representing genes as digital data to be browsed bestowed them with qualities of being uninterpreted or raw, which obfuscated the fact that no data is ever raw but always curated (Bowker, 2000), which has been discussed in relation to the big data discourse (Kitchin, 2014). In the case of 23andMe this curating was exactly what the company sold, turning saliva into SNPs using a specific microarray technology. Thus whilst the use of verbs, such as browsing, gave a sense of limitless exploration, it was restricted to investigating a digital dataset containing a list of selected SNPs, which obviously was a very limited and specific lens through which to explore health, as has been observed (O’Riordan, 2013).

Furthermore, and perhaps more importantly, when genes were described as digital big data or resource they could be moved around, downloaded, uploaded, retrieved, navigated, pooled and shared. This rhetorical move legitimated the collection of

customers' DNA to be sold on to researchers and commercial companies, as genes were represented as a digital resource circulated to facilitate discovery by lay users and commercial partners alike rather than being a fateful code cracked by scientists in a laboratory.

Genes and Correlations

Another key feature of the discourse on big data is the idea that its sheer volume enables the correlation of diverse or heterogenous data, facilitating new discoveries and insights (Kitchin, 2014). 23andMe's platform fostered viewing genes as data that could be correlated with other forms of data; this view was not only communicated through text but also by the way in which the platform was structured.

Prior to the FDA ban in 2013 23andMe offered its customers "genetic risk profiles," which synthesized information from several gene variants. Thus, a customer could be told they were at an "increased" genetic risk for a specific health condition, such as type 2 diabetes. 23andMe did not offer preventive advice directly based on the risk profiles (i.e. based on your genes you should eat this or avoid that). Rather, next to the profiles the company linked to "resources," which would lead to organisations, such as American Heart Association, offering information on the disease, risks and prevention. Thus, the genetic risk profile for type 2 diabetes would link to a BMI calculator, family history tool and various diabetes risk calculators, which would calculate users' risk based on questionnaires on e.g. waist circumference or

consumption of simple carbohydrates as well as offer preventive advice. These links invited customers to correlate their genetic test results with their BMI, family history and so on.

The polygenic risk profiles were withdrawn after the FDA ban in 2013. However, the company continued to offer links to similar resources and other data in conjunction with the genetic risk factors in the UK 2015 test. Further, customers could continue to get information on the associations between their gene variants and alleged susceptibility for many different diseases by uploading their raw data onto software, such as Promethease, which explicitly described itself as linking data or “connecting a file of DNA genotypes to scientific findings cited in SNPedia” (an open access database containing SNPs). The Promethease report linked individual gene variants to multiple conditions and traits. For example, the gene variant rs1800497 was linked to avoidance of errors, ADHD, alcohol dependency, postoperative nausea, obesity and effectiveness of a smoking cessation drug (Buprion). Similarly, conditions were linked to multiple genes; so that one could find multiple genes associated with bipolar disorder, and the report also described possible gene interactions. Thus, Promethease did not envision a single gene coding for a single trait. Rather, the service cast genes as part of a complex, even confusing, cluster of data, where multiple disparate conditions were associated with multiple genes and vice versa.

The way in which customers were invited to participate in research on the 23andMe website also encouraged thinking about associations between genes and other data. On the front page after login on the 2015 UK platform customers were asked to

answer short questions for research, such as whether one is a morning person or a night person or how close one lives to a farm or an industrial plant. These questions insinuated that there was a potential association between one's genes and being a morning person or living close to farms or factories.

In 2017 "featured content" appeared on the front page of the 2008 account, combining the customer's genetic test results ("Your genes put you at **lower** risk for migraines") and her responses to research surveys ("YOU SAID I have not had migraines") as well as survey data from CureTogether -- a platform for people with chronic conditions crowdsourcing data to find cures and acquired by 23andMe in 2012 to source phenotypic data (Empson, 2012). The section described how "we tapped the wisdom" of "the CureTogether community" and identified triggers (stress 85%, alcohol 42%, bananas 6% etc.) and treatments (Tylenol®, dark room, wet towel around your head etc.) for migraines, continuing that "symptom data and treatment data" can "powerfully" reveal that those who "experienced dizziness with their migraines were three times more likely to have a negative reaction to Imitrex®." So, the juxtaposition of genetic test results, one's own survey answers and crowdsourced data on symptoms, triggers and treatments for migraine created a sense of it being possible to gain "powerful" insights on, for example, drug response, by correlating these different data.

The same idea of finding new insights by correlating genetic and survey data occurred in the media coverage of research using 23andMe's data. For example, a study funded by Olay (a cosmetics brand owned by Procter & Gamble) on skin

ageing was reported to have found out that “secret” of “ageing well” was “down to lifestyle,” noting that genetics can “give you a head start” but other factors that “really” “shape how you look” included taking a multivitamin, sleep and using a moisturiser (Allen, 2017). Another study, funded by Pfizer, on depression was reported to have found “17 genetic variations that raise the risk of depression” and could possibly “lead to new treatments.” The article also noted that each genetic variant contributed only to a “minuscule” increase in the risk and so “huge sample sizes” were needed to “spot” them, highlighting “the value of data from genetics companies” (Davis, 2016).

These news reports did not represent genes deterministically but described genes as playing only a small role in skin ageing or development of depression, which led the other article to rehash the discourse on big data referring to how “huge” quantities of data was needed to “spot” these “minuscule” risks. However, genes were still seen through the metaphor of information transmission, as they were represented as coding for risk, even if small. The more important thing, however, was the way in which the data brought to bear on skin ageing and depression dovetailed the commercial interests of the funders, illustrating how moisturisers help with youthful looks or pinpointing genetic variants that might aid in the development of new antidepressants in the potentially commercially lucrative situation where more than 1 in 10 Americans are taking medications for depression but the development of new drugs has stalled since the invention of selective serotonin reuptake inhibitors in the 1980s (Block, 2015; Centers for Disease Control and Prevention, 2011).

Overall, 23andMe often cast genes through the ontological metaphor of big data, whereby they were seen as one data among many, and where correlation is everything (Anderson, 2008). The juxtaposition or correlation of different data that allegedly contributed to health and illness -- from genes to waist circumference, dark rooms and use of moisturiser -- created an impression of diverse data and surprising correlations born “straight out of data” (Kitchin, 2014; Ruppert et al., 2013).

However, even if genes were cast as part of heterogeneous big data they were not seen through the epigenetic metaphors (Stelmach & Nerlich, 2015) as modifiable by environmental triggers. Rather, genes were represented through the information metaphor as coding for (smaller or bigger) risk for disease or other trait, such as premature skin ageing. This representation of genes was essentially not much different from Sciona’s view of them as coding for risk for lifestyle-related diseases.

Perhaps more importantly, although the data correlated with genetic data appeared diverse, it mostly focused on individual behaviour or indicators typically associated with such behaviour, such as waist circumference, stress, sufficient sleep and alcohol consumption. The data collected in the commercially funded studies also reflected the funders’ interests, such as including data on use of cosmetics or seeking to aid in the development of pharmaceuticals. The way in which 23andMe represented genes as big data to be correlated with limitless plurality of other data obfuscated the narrow, commercially informed range of data collected and correlated.

Multiple Interpretations

Following from above, another key aspect of the discourse on big data is the idea that the vast amounts of data becoming available online can be analyzed by anyone with digital devices (Ruppert et al., 2013). Thus, the way in which communication between 23andMe and its customers was represented through the ontological metaphor of big data implied that genetic data could be analyzed by a wide variety of actors, yielding multiple interpretations.

The 23andMe platform provided its customers many such “digital devices” to enable them to analyse their genetic data. The platform enabled customers to “connect” with alleged “relatives” and “share” and “compare” “genomes” with them, and early on news coverage noted that the company “echo the style of social-networking sites” (Blakeley, 2007). The ability to upload one’s raw data enabled customers to analyze or compare their SNPs alone or together. It was commonplace in the 23andMe user fora for customers to solicit others to share their SNP data and disease data to search for correlations to corroborate putative theories of specific genes being associated with disease, mimicking the company’s research.

However, on other sections of the platform 23andMe articulated its relationship with its customers in terms of accurate information transmission. The genetic test results included illustrative graphs on relative risk and the ‘Genetics 101’ video on the front

page of the website both conforming to the ideal of accurate information transmission from experts to lay customers.

Further, 23andMe enabled its customer not only to explore or browse their raw data on its website but also to upload it onto other software. MIT Technology Review article on one of the software, Promethease, noted that “DNA information is essentially digital. That means it can plug and play anywhere” (Regalado, 2014); the same idea was espoused by Wojcicki in a PLoS Genetics interview:

We have an API [Application Program Interface], a standard tool in the tech world to enable people to pull down their data. That way, Promethease and all these others can build tools on top of the infrastructure that we have. We try to encourage use for art or music or other things that we just aren't doing. It's your data—you can do all kinds of things—explore it! We want to foster that industry of others bringing your data to life in ways we don't, and we've had other research groups use data form medical information (directly from our clients)” (Gitschier, 2015).

The software customers could use to “plug and play” their raw data included services which interpreted associations between gene variants and disease susceptibility (Promethease, Interpretome), turned the raw data into music (DNAMelody) or for nutritional advice (NutraHack) and tracked potential relatives through comparing DNA (GEDMatch), just to name a few. Whereas the metaphor of information

insinuated that there was one accurate expert interpretation of genes, the metaphor of big data cast DNA as open for multiple interpretations or to be “played” or “brought into life” differently. The musical metaphor of “playing” DNA differently has been observed to be used in epigenetics (Stelmach & Nerlich, 2015), however, in this instance it was not nature or environmental triggers but software that was “playing” DNA differently; the latter metaphor implied that data was dead or inert until it was brought to life by software, further imputed it with powers of vitality and creativity.

Further, 23andMe envisioned genetic data as part of a broader discourse on digital health or Quantified Self (Ruckenstein & Pantzar, 2015), whereby individuals continuously collect, store and analyse a variety of data on themselves, as explained by Wojcicki:

Your health care is no longer about the episodic visit to your doctor, where you have this once-a-year assessment of random vitals," she says. "It's about the continuous stream of you.(Bercovici, 2015).

This interview projected a future vision, whereby doctors had been surpassed as providers of health knowledge, and individuals, streaming data, using digital devices, to monitor their health.

This vision of self-monitoring and data collection for health also folded back onto 23andMe's goal of collecting and trading customers' DNA and survey data. In 2017 a vignette appeared on top of the genetic test results, stating "23andMe Research Discoveries were made possible by 23andMe members who took surveys," implying that by answering the surveys (available by clicking on the vignette) customers could add to the "discoveries" making up the results (even if these were based on scientific research on genetic associations published worldwide). Similarly, the feature combining customer's genetic test results, survey answers and the CureTogether crowdsourced data on migraines invited customers to contribute to the CureTogether database and 23andMe research. Thus, customers and the company were represented as jointly creating, streaming or "crowdsourcing" diverse data, which could directly enable individuals to know more about themselves; this has been noted by Prainsack and Tutton (Tutton & Prainsack, 2011), who have observed that 23andMe invited its customers to take part in research not for altruistic but for "entrepreneurial" reasons to gain knowledge about themselves.

Overall, imagining the communication between 23andMe and its customers through the ontological metaphor of big data represented them jointly producing and analysing data with the aid of diverse software, yielding multiple interpretations. However, as has been noted, the promise of big data as being analysable by everyone is hollow, as companies, such as Google or Procter & Gamble, have privileged access to the big data sets and the computational skills to analyse them (Boyd & Crawford, 2012). Lay consumers can analyse their genetic and other data; yet their explorations are constrained by the platforms, data and software made

available to them mainly by 23andMe and other companies to encourage them to consume the services and to produce data to be sold on.

Discussion

Sciona - one of the first DTC genetic testing companies – represented genes through the classical metaphor of information as coding for lifestyle related diseases and transmitting this expert knowledge to consumers so that they could offset any deficiencies in their genetic make-up with specific foods and supplements.

23andMe's website and its media coverage also evoked the metaphor of information to represent and communicate about genetic risk factors. However, the company and media also cast genes and the relationship between the company and its customers through the novel ontological metaphor of big data. So, genes were represented as an abundant resource, which could be browsed and correlated with diverse other data to reveal new insights. Further, in keeping with the big data discourse (Kitchin, 2014) this browsing and correlating was represented as something that lay customers and scientists could both do. Representing DTC genetic testing through the ontological metaphor of big data gave a sense of limitless possibilities in terms of the amounts and diversity of data to be collected and correlated and the kinds of interpretations or insights to be yielded. Yet, on closer examination, the kinds of data correlated with 23andMe's genetic data and the kinds of discoveries made highlighted that 23andMe's tests and the research the company

supported still represented genes in terms of coding -- in a bigger or smaller way -- for illness or and wellness to be mitigated by lifestyle or targeted drugs.

However, the ontological metaphor of big data was also different from the classical notion of genes as information. Metaphors operate cross domains (Lakoff & Johnson, 2008), and rendering genes digital big data transfers them from the closed, expert domains of the laboratory and the clinic into digital media, within which data can be moved about, interpreted, shared and traded by any user. The ontological metaphor of big data, thus, legitimised 23andMe's business model, both in terms of selling genetic tests direct to consumers and customers' DNA and other data to other companies, as genes were like any digital data to be browsed and shared.

Research on metaphors and genes has often noted that the information metaphor leads to a deterministic notion of genes, as if containing the programme for life (Nelkin, 1994). Scholars have suggested that genes should be understood in more "contextual" manner, which would acknowledge that they interact with their biological, environmental and social context (Kay, 2000). Recent metaphors of genes as music or as part of an ecology identified in scientific and media coverage of epigenetics and the Human Microbiome (Nerlich & Hellsten, 2009; Stelmach & Nerlich, 2015) suggest that more interactive notions of genes are emerging. The way in which the ontological metaphor of big data was articulated on 23andMe's website and media coverage did not, however, suggest that genes are influenced by other factors, even if other factors were acknowledged to play a role in determining wellness and illness. The metaphor of big data, thus, did not represent biological

processes or what genes “can do” (Moss, 2004) in a new way, but it represented social processes or what can be done with genes in a novel way by rendering them a digital resource to be circulated, shared and sold.

Critical social scientists have observed that 23andMe mobilises concepts associated with the participatory or democratic potential of digital media, such as open source, being part of a community and sharing, to legitimise collecting customers’ DNA and other data to be sold for profit (Harris et al., 2016; O’Riordan, 2013; Van Dijck & Poell, 2016). The analysis of how the company mobilised the metaphor of big data illustrates how 23andMe legitimised its operation by casting genes not as a special code to be cracked by the cognoscenti but as any digital data resource to be circulated among the many.

In the end, policy conversations around DTC genetic testing have been underpinned by the information transmission metaphor seeking to guarantee the accuracy of the tests and that consumers accurately understand them (Food and Drug Administration, 2013). 23andMe, which has quickly become a major player in consumer genetics and private biobanking raises the possibility that these kinds of companies begin to not only influence the way in which lay customers perceive their health but also shape health research. In this situation there is a danger that data --- such as data on social and environmental inequalities -- which does not fit the agenda of entrepreneurial individualism, biotech and profit making gets increasingly downplayed and ignored in preventive healthcare and research (Reardon, 2013). So, the political questions raised by 23andMe extend beyond consumer protection and

towards critical issues vis a vis health and data both highlighted and hidden (Lakoff & Johnson, 2008) by the ontological metaphor of big data: What kinds of data are considered important or become data in the first place? (Bowker, 2000). Whose data is being collected, who does the analytical and interpretive work that turns the data into legitimate knowledge and for what purpose?

References

- 23andMe. (2015, June 18). The power of one million. Retrieved from <https://blog.23andme.com/news/one-in-a-million/>
- Allen, V. (2017, March 3). Secret to ageing well down to lifestyle, not just genes. *Mail Online*
- Anderson, C. (2008). The end of theory. *Wired Magazine*, 16(7), 16-07.
- Baker, S. (2013, January 14). Next-generation sequencing vs. microarrays. Retrieved from <http://www.genengnews.com/gen-articles/next-generation-sequencing-vs-microarrays/4689>
- Barnett, A. (2003, January 19). New gene 'horoscope' predicts our life -- and death. *Observer*
- Barthes, R. (1972). *Mythologies*. 1957 Hill & Wang.
- Bercovici, J. (2015, October 21). The billion dollar company that not even the feds could stop. *Inc*,
- Blakeley, R. (2007, May 24). Search engine aims to google your genes. *The Times*
- Block, J. (2015, December 2). Markets for antidepressants, anti-anxiolytics do decline. *Psychiatry Advisor*,
- Bowker, G. C. (2000). Biodiversity datadiversity. *Social Studies of Science*, 30(5), 643-683.

Boyd, D., & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society*, 15(5), 662-679.

Centers for Disease Control and Prevention. (2011, October). Antidepressant use in persons aged 12 and over: United states, 2005-2008. Retrieved from <https://www.cdc.gov/nchs/products/databriefs/db76.htm>

Cooke, A. (2001, September 9). Health: Can gene testing change your life? *Daily Mirror*

Davis, N. (2016, August 1). Scientists discover 17 genetic variations that raise the risk of depression. *Guardian*

Earle, C., & Quinton, M. (2014, December 5). Home DNA test can spot fatal illness early ... so why are scientists sceptical? *The Sun*

Empson, R. (2012, July 11). 23andMe nabs first acquisition, nabs CureTogether to double down on crowdsourced genetic research. Retrieved from <https://techcrunch.com/2012/07/11/23andme-first-acquisition/>

Food and Drug Administration. (2013). Warning letter to 23andMe inc. Retrieved from <http://www.fda.gov/iceci/enforcementactions/warningletters/2013/ucm376296.htm>

Gibbs, S. (2014, December 2). DNA screening test 23andme launches in UK after US ban. *Guardian*,

- Gitschier, J. (2015). Your data to explore: An interview with anne wojcicki. *PLoS Genet*, 11(10), e1005548.
- Glaser, B. (1965). The constant comparative method of qualitative analysis. *Social Problems*, 12(4), 436-445.
- Hall, S. (1980). Encoding/decoding. In S. Hall, D. Hobson, A. Love & P. Willis (Eds.), *Culture, media, language* (pp. 128-138) HUTchinson.
- Harris, A., Kelly, S., & Wyatt, S. (2016). *CyberGenetics: Health genetics and new media* Routledge.
- Human Genetics Commission. (2003). *Genes direct ensuring the effective oversight of genetic tests supplied directly to the public.* (). London:
- Kay, L. E. (2000). *Who wrote the book of life?: A history of the genetic code* Stanford University Press.
- Keller, E. F. (2009). *The century of the gene* Harvard University Press.
- Kitchin, R. (2014). Big data, new epistemologies and paradigm shifts. *Big Data & Society*, 1(1), 2053951714528481.
- Kragh-Furbo, M., & Tutton, R. (2017). Spitting images: Remaking saliva as a promissory substance. *New Genetics and Society*, , 1-27.
- Lakoff, G., & Johnson, M. (2008). *Metaphors we live by* University of Chicago press.
- Moss, L. (2004). *What genes can't do* The MIT Press.

- Nelkin, D. (1994). Promotional metaphors and their popular appeal. *Public Understanding of Science*, 3(1), 25-31.
- Nerlich, B., & Hellsten, I. (2004). Genomics: Shifts in metaphorical landscape between 2000 and 2003. *New Genetics and Society*, 23(3), 255-268.
- Nerlich, B., & Hellsten, I. (2009). Beyond the human genome: Microbes, metaphors and what it means to be human in an interconnected post-genomic world. *New Genetics and Society*, 28(1), 19-36.
- O'Reilly, T. (2007). What is web 2.0: Design patterns and business models for the next generation of software. *Communications & Strategies*, (1), 17.
- O'Riordan, K. (2013). Biodigital publics: Personal genomes as digital media artefacts. *Science as Culture*, 22(4), 516-539.
- Oxford Living Dictionaries. (2016,). Definition of data in english. Retrieved from <https://en.oxforddictionaries.com/definition/data>
- Reardon, J. (2013). On the emergence of science and justice. *Science, Technology, & Human Values*, 38(2), 176-200.
- Regalado, A. (2014, October 19). How a wiki is keeping direct-to-consumer genetics alive. Retrieved from <http://www.technologyreview.com/featuredstory/531461/how-a-wiki-is-keeping-direct-to-consumer-genetics-alive/>
- Ruckenstein, M., & Pantzar, M. (2015). Beyond the quantified self: Thematic exploration of a dataistic paradigm. *New Media & Society*, , 1461444815609081.

Ruppert, E., Law, J., & Savage, M. (2013). Reassembling social science methods:

The challenge of digital devices. *Theory, Culture & Society*, 30(4), 22-46.

Saukko, P. M., Reed, M., Britten, N., & Hogarth, S. (2010). Negotiating the boundary

between medicine and consumer culture: Online marketing of nutrigenetic tests.

Social Science & Medicine, 70(5), 744-753.

doi:10.1016/j.socscimed.2009.10.066

Shannon, C. E., & Weaver, W. (1949). *The mathematical theory of information*

University of Illinois Press.

Stelmach, A., & Nerlich, B. (2015). Metaphors in search of a target: The curious case

of epigenetics. *New Genetics and Society*, 34(2), 196-218.

Thornton, J. (2003, April 10). DIY gene tests under fire. *The Sun*

Tutton, R., & Prainsack, B. (2011). Enterprising or altruistic selves? making up

research subjects in genetics research. *Sociology of Health & Illness*, 33(7),

1081-1095.

US Government Accountability Office. (2006). *NUTRIGENETIC TESTING tests*

purchased from four web sites mislead consumers. ().

US Government Accountability Office. (2010). *DIRECT-TO-CONSUMER GENETIC*

TESTS misleading test results are further complicated by deceptive marketing

and other questionable practices. ().

Van Dijck, J. (2014). Datafication, dataism and dataveillance: Big data between

scientific paradigm and ideology. *Surveillance & Society*, 12(2), 197.

Van Dijck, J., & Poell, T. (2016). Understanding the promises and premises of online health platforms. *Big Data & Society*, 3(1), 2053951716654173.