



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Boundless riches

Citation for published version:

Fuller, M 2017, Boundless riches: Big data, the Bible and human distinctiveness. in *Issues in Science and Theology: Are we special?*. Issues in Science and Religion: Publications of the European Society for the Study of Science and Theology, vol. 4, Springer, pp. 181-189. https://doi.org/10.1007/978-3-319-62124-1_13

Digital Object Identifier (DOI):

[10.1007/978-3-319-62124-1_13](https://doi.org/10.1007/978-3-319-62124-1_13)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

Issues in Science and Theology

Publisher Rights Statement:

This is a post-peer-review, pre-copyedited version of a chapter published in Issues in Science and Theology: Are We Special?. The final authenticated version is available online at: https://link.springer.com/chapter/10.1007/978-3-319-62124-1_13

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



Boundless riches: Big Data, the Bible and Human Distinctiveness

Michael Fuller

Abstract

A number of scientific concepts have proved helpful to theologians in the resonances they offer with theological ideas, meaning that they can bring us novel images and insights to feed into the ways in which we think about God. The present author has suggested elsewhere that the phenomenon of Big Data – large, richly-complex datasets, which can be gathered from a plethora of sources through the medium of computing – offers a novel space for the interaction of theology and the sciences regarding ethical and hermeneutical issues. The present paper describes the phenomenon of Big Data, and introduces it as a new resource from the sciences with the potential to cast light on two areas of interest to theologians. First, the richness and variety of Big Data as a source of the various analyses which may be based upon it suggest interesting parallels with the richness and variety of the scriptures as a source of understandings of God. Second, in highlighting the way in which qualitative distinctiveness might arise from quantitative difference, Big Data may also have some important insights to offer regarding the question of human distinctiveness.

Keywords: analogy, analysis, Bible, big data, brain, computer, imagery, model

Introduction

It has been a common criticism of the dialogue of science and theology that the traffic within it has too often been one-way: that the findings of various scientific disciplines have offered more to the theologian than the reflections of theologians have offered to the sciences (cf. Southgate and Poole 2011: 29). As a possible example of traffic in the other direction, I have drawn attention elsewhere to the new form of science and scientific practice that is coming about with the advent of the very large datasets made possible by the modern mass usage of computers – the phenomenon known as ‘big data’; and I have suggested that in this new area of science, the insights of theologians may be of value to data scientists. In particular, the *hermeneutical* skills developed by theologians in their engagement with texts might be of value to those engaging with big datasets (Fuller 2105), since similar issues around meaning, context (of the dataset and of the interpreter), bias, and the role of insights from many disciplines may be involved in both cases. I have also suggested that the Churches, as organisations which are in a position to speak out on behalf of citizens over issues of ethical concern, might do so regarding the many *ethical* issues raised by the storage and use of data which is specific to individuals (Fuller 2016), not least where these concern information about a person’s medical history, or other information regarding which that person has a reasonable expectation of privacy.

The present paper offers the suggestion that the dialogue between theologians and those engaged with big data might be of benefit to theologians, too. Specifically, it puts forward two ways in which insights from data science might feed into theological thinking. First, the phenomenon of big data demonstrates that the extraction of information from

complex, multivariate sources allows for – indeed, necessitates – the assumptions that valuable information is there to be extracted, and that the use of a variety of interpretative strategies may be required: in this, it indicates that common methodological assumptions may underlie theological and scientific research strategies. Second, big data demonstrates that qualitative distinctiveness may arise from a purely quantitative phenomenon; and this offers us an analogy for thinking about human beings as qualitatively distinct from other species of which we are aware. On this basis, it is suggested that it is possible to answer the question ‘Are we special?’ in the affirmative.

What is meant by ‘big data’?¹

Although the expression ‘big data’ has been much contested – to the extent that O’Neil and Schutt (2014: 2) have commented of ‘big data’ and ‘data science’ that ‘These terms are so ambiguous, they’re well-nigh meaningless’ – big data has classically been thought of in terms of ‘three Vs’: volume, variety and velocity. (This understanding was introduced by Laney (2001): for a more thorough discussion of the distinctiveness of big data, see Kitchin (2014: 68ff).) In other words, it is used to refer to huge quantities of data, which is highly heterogeneous, and which is generated at enormous speed. With personal computers, smartphones, satnavs and all kinds of data-generating devices now in more or less constant use throughout the world, and with capacities for storage of the data which they generate growing ever more vast, enormous amounts of data can now be collected and stored in all kinds of contexts.

To give a few examples of how big data is generated and used:

- Some scientific research projects routinely generate vast amounts of data, for example the Sloan Digital Sky Survey, which is gathering astronomical information, or the experiments being carried out with the Large Hadron Collider at CERN, in Geneva (cf. Mayer-Schönberger and Cukier 2013: 7; Chen *et al.* 2014: 21-23).
- When people shop online, or use store cards, their purchasing habits can be monitored, stored, and used to make recommendations for future purchases. It is reckoned that something like one-third of Amazon’s sales are generated by this means (Mayer-Schönberger and Cukier 2013: 52), so it is clearly economically beneficial for retail companies to invest in the computing power to sift its customers’ purchasing records and to make such recommendations. Indeed, a significant industry has now grown up around the harvesting and sale of consumer data harvested from such sources (cf. CIPPIC 2006).
- Whenever people use a device that is linked to GPS systems, their location can be recorded and stored. The potential value of such information to security services is obvious, particularly if there are reasons to suppose that individuals are engaged in any kind of nefarious activity. There are also, of course, potential concerns about personal liberties raised by the covert storage and possible uses of this information.
- People’s medical records contain a great deal of information that will have been gleaned over the years concerning their state of health; and if they have ever been enrolled in clinical trials or medical research of any kind, these records might well contain information about various genetic markers and/or other factors in their make-up that have been germane to that research. These records constitute an enormous database which might usefully be scrutinised again and again for information as medical research becomes more and more refined. (They also give rise to a whole range of ethical issues regarding

¹ The word ‘data’ is, of course, a plural form. However it, and the expression ‘big data’, are generally (if not universally) taken to be singular in the literature, and that convention is followed here.

anonymisation, consent, safe storage of data, and so on (cf. Fuller 2016), although these will not concern us here.)

- All these examples concern data gathered with a particular focus or intention, be that scientific, medical, commercial, or political. But it is also possible to harvest datasets which are much more variegated and unstructured. Consider, for example, the dataset generated by gathering all the terms put into a search engine like Google over a 24-hour period. This might appear to be a completely random, unstructured dataset; but it is possible that useful information may be buried within it, if one asks the right questions of it. Interrogation of this data has been used in attempts to predict flu outbreaks, for example, although these attempts have so far been of limited usefulness (cf. Fuller 2015).

Characteristics of big data

With this background in place, what general conclusions may we draw regarding the nature of big data, and our relationships with it?

Characteristic of big data is that it is:

- Bewilderingly rich and varied in its quality. Because big datasets may be generated in an indiscriminating way, they will contain all kinds of information thrown together in an uncoordinated way (think of all those terms input into Google over 24 hours). A corollary of this is that big data is:

- Extremely difficult to understand and interpret. The sheer variety encompassed by large datasets makes their interrogation problematic: it might be possible to generate any number of different answers to questions from them, depending on the precise questions are asked (cf. Berman 2013: 145: ‘When the amount of data is sufficiently large, you can find almost anything you seek lurking somewhere within’). They will always contain more information than it is possible for us to extract from them.

Given all this, it would appear not to be stretching a point to say that big data may be considered to contain ‘boundless riches’ (cf. Eph 3:8, NRSV): there is so much that it is possible for us to mine from it – and yet also so much about it that exceeds our ability to grasp.

And not only this. It has been pointed out that with the accumulation of so much data, ‘something new and special is taking place. Not only is the world awash with more information than ever before, but that information is growing faster. The change of scale has led to a change of state. *The quantitative change has led to a qualitative one*’ (Mayer-Schönberger and Cukier 2013: 6, my emphasis). This observation appears to be a common one: it has similarly been noted that ‘It is becoming clear that big data have a number of inherent characteristics that make them qualitatively different to previous forms of data’ (Kitchin 2014: 79), and that ‘an intensification of quantitative differences allows for the articulation of qualitative difference’ (Raley 2013: 124). In other words, the sheer volume of data leads to a step change: a massively large dataset is qualitatively, not just quantitatively, different from a traditional, smaller one, and it requires different tools and different models for its handling and analysis.

Now, theological discourse has often been enriched by noting parallels between theological and scientific concepts (for example, in the application of the phenomenon of complementarity in both quantum physics and Christology (Polkinghorne 1994: 133-4), or the way in which top-down causation might be used to explore models for the interaction of God with the physical universe (Peacocke 1993: 157-160)). Are there ways in which the phenomenon of big data may similarly enrich this discourse? The observations above

concerning big data suggest two sets of resonances which are of considerable theological interest.

Theological resonances: (i) interpreting complex material

We may immediately note that the richness and complexity of big data means that there are likely to be many different approaches to large datasets, yielding different (and, maybe, even contradictory) understandings through their analysis of them. With such a dataset, we are dealing with something which may be many-layered, ambiguous, and difficult to interpret – many readings may lie within it, such that wisdom is required in surfacing one reading (with a concomitant suppression of others). In the same way, the biblical witness to God is richly varied, supporting a variety of interpretations: it may also be ambiguous and difficult to interpret, with multiple meanings lying within it; and there may be dangers inherent in surfacing particular readings and suppressing others. By way of illustration, consider the following.

Images used of God in the Bible are extraordinarily rich and varied, to the extent that that richness can only be touched on here. They can be elemental (a burning bush, Exodus 3:2: a refiner's fire, Malachi 3:2: light, Isaiah 60:20: wind and tongues of flame, Acts 2:2-3). They can relate to animals (a bear or a lion, Hosea 13:8: a mother hen, Matthew 23:37, Psalm 17:8), or to humans fulfilling particular roles (a mother, Isaiah 66:13: a shepherd, Psalm 23:1: a potter, Isaiah 64:8, Jeremiah 18:6: a judge, Psalm 9:8: a king, Jeremiah 10:10, Psalm 93:1, 97:1). They may be apophatic in character (for example, Elijah's encounter with God, in which God is *not* found in elemental forces of wind, earthquake and fire, but rather in 'a sound of sheer silence': 1 Kings 19:11-12). All these images – and many more besides – generate an enormously rich resource of ways in which people have thought about the person and the activity of God; and since they are images which rely to various extents on the emotional, physical and socio-political resonances which they present, all of them offer interpretative challenges. Moreover, such challenges will vary at different times and in different cultures: for example, the qualities of shepherds and kings which the biblical writers are attempting to convey will mean different things to the rural inhabitant of a first-century near-eastern monarchy and to the metropolitan inhabitant of a twenty-first century Western republic.

In addition, the Bible stresses the incomprehensibility of God. 'To whom then will you compare me, or who is my equal? says the Holy One. Lift up your eyes on high and see: who created these? He who brings out their host and numbers them, calling them all by name' (Is 40:25-26). The difficulty of a human being comprehending God is beautifully illustrated in the concluding chapters of the book of Job, in which God's questioning of Job lead to Job's eventual confession that 'I have uttered what I did not understand, things too wonderful for me, which I did not know' (Job 42:3). The task of the theological interpreter in speaking of God based on such richness is clearly a problematic one, requiring self-awareness, and a careful and thoughtful nuancing of what is said.

All of this is, of course, familiar to anyone engaged in the exegesis of biblical texts: as Thiselton (2009: 8) puts it, 'The interpreter of texts is not a neutral observer ... *Understanding* in the fullest sense demands *engagement and self-involvement*' (italics in original). It is crucial to note that *what is true of the interpreter of texts is true also of the interpreter of data*. 'Statistics appear to be objective and neutral, but they are seldom that in practice. Everything depends on their purpose and presentation' (Thiselton 2009: 218). As Kitchin (2014: 2) elaborates:

While many analysts may accept data at face value, and treat them as if they are neutral, objective, and pre-analytic in nature, data are in fact framed technically, economically, ethically, temporally, spatially, and philosophically. Data do not exist independently of the ideas, instruments, practices, contexts and knowledges used to generate, process and analyse them.

If theology can offer to data science an awareness of the importance of hermeneutics (cf. Fuller 2015), data science has the potential to offer to theology a confidence in the application of analytical tools, be they algorithms or hermeneutical methods, as a means of uncovering the truths that may lie hidden in the material being investigated.

Each of these disciplines also has the potential to caution the other over the claims that are being made. Just as big datasets offer a huge mass of material which requires careful, critical (and self-critical) reading, so the information we are able to obtain from the biblical witness presents us with richly varied, and also ambiguous and even paradoxical, understandings of God; and just as any one ‘take’ on the information contained in a big dataset will stand in permanent need of correction with an evolving interrogation and understanding of the dataset as a whole, any understanding about God taken from one part of the Bible requires careful reading and interpretation in the light of our dynamically-unfolding comprehension of the biblical witness as a whole. In the same way that partial (in either sense of the word) reading of big datasets can lead to distorted understandings, partial reading of the Scriptural witness to God can also lead to distorted understandings. At their worst, these may even lead to forms of what modern Western societies would dub religious extremism – for example, if texts which are understood to advocate discrimination against individuals or groups are held to outweigh texts which affirm the ending of such discrimination (e.g. Galatians 3:28), or the universal Law of Love (Matthew 5:43 par; Matthew 22:37-39 par).

Theological resonances: (ii) human distinctiveness

Can big data cast any light on the question of whether or not human beings are special, the theme of this symposium?

To be ‘special’ is often considered to be equivalent to being ‘unique’ in some way. But need that necessarily be the case? Consider one of Pascal’s *Pensées*: ‘Man is only a reed, the weakest in nature, but he is a thinking reed. ... all our dignity consists in thought. It is thought which must raise us up’ (Pascal 1966, 95).² Pascal does not claim that humankind is *unique* in possessing the capacity to think, but he does suggest that that capacity confers a particular *dignity*: a dignity which might be considered to distinguish us in some way from the rest of the created order (‘It is thought which must raise us up’). Now, as noted above, the phenomenon of big data has demonstrated how size and complexity can generate a qualitative as well as a quantitative distinction between large and small datasets, such that new forms of analysis and description become necessary when we think about big data. In the same way (it might therefore be argued), the size and complexity of the human brain – the organ responsible for our capacity to think – makes us qualitatively as well as quantitatively distinct from other species. (I am not here advancing as far as those who would speak of new and distinct *ontological* concepts emerging from simpler substrates – for example, of the mind as emerging from the brain (cf. Clayton 2004: 107 ff.). Rather, I simply wish to affirm the observation that, once a

² L’homme n’est qu’un roseau, le plus faible de la nature, mais c’est un roseau pensant. ... Toute notre dignité consiste donc en la pensée. C’est de là qu’il faut nous relever.

particular level of complexity is reached, new ways of approaching a phenomenon become *methodologically* necessary. This may or may not correspond with the emergence of new ontological concepts.)

Of course, the statement that human beings are qualitatively as well as quantitatively distinct from other species invites all sorts of criticisms, not least regarding the size and level of complexity a species' brain has to reach before we can make such a claim. (Similar questions have been raised regarding the point at which it becomes appropriate to speak of 'big data': 'Just how *big* is big? Or is it just a relative term?' (O'Neil and Schutt 2014: 2).) However, it is surely the case that human beings possess brains which can do far more than simply process information in a straightforward way ('if *x*, then *y*'). And, indeed, the way in which we are capable of interrogating big data might perhaps offer a paradigmatic example of the way in which our powers of thought might be the expression of something uniquely human.

Moreover, it is this capacity for thinking which allows us to make sense of so much that we encounter in the world around us. A very large dataset offers us an example of the way in which, confronted by something of a bewildering complexity, it is our capacity for inventive thinking which enables us to find ways of interrogating it and deriving meaning(s) from it. When we are similarly confronted by the masses of bewilderingly rich and varied information which we encounter in the world around us, we likewise assume that it is possible to derive meaningful information from it, and we endeavour to do so – this is not only the task of the scientist, but of every individual who seeks meaning in the situation in which she is located. In the same way, a Christian theologian might say, those who attempt to grapple with the idea of God must use the plethora of sources found in the Bible (and also those found in Christian tradition, and in their own experience) to come to an understanding of that God; and it is their distinguishing capacity of thought which means that they are capable of doing so.

We can have no idea of the depth of thought that is possible for chimpanzees, or horses, or even (of course) bats (cf. Nagel 1982) – let alone for extra-terrestrial beings, should they exist. Nevertheless, as we have seen, if we think of human beings as qualitatively distinct from other species on the basis of our quantitatively different brains, then even if we cannot think of such distinctiveness as a sign that human beings are *unique* without running the risk of being charged with anthropocentric chauvinism, we may still perhaps tentatively suggest that it is something that makes us *special*. Perhaps, too, our capacity to come to at least a partial understanding of the boundless, unfathomable, and ultimately incomprehensible being we call God (whilst holding in mind the revisability of any understanding we have, in the light of our continual re-interrogation of the data at our disposal) is another thing that makes us distinguish us from other creatures.

Big data, then, offers us a potentially valuable resource for theology. The analysis of large datasets offers a parallel between the work of data scientists and theological exegetes; and the way in which large datasets highlight the possibility for qualitative distinctiveness to arise from purely quantitative phenomena throws interesting light on the question of human specialness. Doubtless, further interesting resonances remain to be uncovered as dialogue is pursued between the practitioners of theology and data science.

Bibliography

Biblical references are to the New Revised Standard Version.

Berman, Jules J. 2013. *Principles of Big Data* (Amsterdam: Elsevier).

- Chen, Min, Shiwen Mao, Yin Zhang and Victor C. M. Leung. 2014. *Big Data: Related technologies, challenges and future prospects* (Heidelberg: Springer).
- CIPPIC 2006. *On the Data Trail: How information about you gets into the hands of organisations with whom you have no relationship. A Report on the Canadian Data Brokerage Industry*. Available at <http://www.cippic.ca/uploads/May1-06/DatabrokerReport.pdf>, The Canadian Internet Policy and Public Interest Clinic, Ottawa (accessed 18 August 2016).
- Clayton, Philip. 2004. *Mind and Emergence: From quantum to consciousness* (Oxford: Oxford University Press).
- Fuller, Michael. 2015. 'Big Data: New science, new challenges, new dialogical opportunities', *Zygon*, vol. 50 no. 3, pp. 569-582.
- Fuller, Michael. 2016. 'Some Practical and Ethical Challenges Posed by Big Data', in Jennifer Baldwin (ed.), *Embracing the Ivory Tower and Stained Glass Window: A Festschrift in Honor of Archbishop Antje Jackelen* (Heidelberg: Springer), pp. 119-127.
- Kitchin, Rob. 2014. *The Data Revolution: Big data, open data, data infrastructures and their consequences* (London: Sage Publications).
- Laney, Doug. 2001. '3D Data Management: Controlling Data Volume, Velocity, and Variety'. META Group report. Available at <http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf> (accessed 20 July 2016).
- Mayer-Schönberger, Viktor and Kenneth Cukier. 2013. *Big Data: A revolution that will transform how we live, work and think* (London: John Murray).
- Nagel, Thomas. 1982. 'What is it like to be a bat?', in Hofstadter, Douglas R. and Daniel C. Dennett (eds.), *The Mind's I: Fantasies and reflections on self and soul* (Harmondsworth: Penguin).
- O'Neil, Cathy and Rachel Schutt. 2014. *Doing Data Science* (Sebastopol CA: O'Reilly)
- Pascal, Blaise. 1966. *Pensées*, trans. A. J. Krailsheimer (Harmondsworth: Penguin).
- Peacocke, Arthur. 1993. *Theology for a Scientific Age (enlarged edition)* (London: SCM Press).
- Polkinghorne, John. 1994. *Science and Christian Belief: Theological reflections of a bottom-up thinker* (London: SPCK).
- Raley, Rita. 2013. 'Dataveillance and Countervailance', in Lisa Gitelman (ed.), *'Raw Data' is an Oxymoron* (Cambridge, MA: MIT Press).
- Southgate, Christopher and Michael Poole. 2011. 'Introduction', in Southgate, C. (ed.), *God, Humanity and the Cosmos* (3rd edition, London: T&T Clark).
- Thiselton, Anthony C. 2009. *Hermeneutics: An Introduction* (Grand Rapids: Eerdmans).