



Envisioning the power of data analytics

David Beer

Department of Sociology, University of York, York, UK

ABSTRACT

It could be argued that the power of data is located in what they are used to reveal. Yet we have little understanding of the role played by the emerging industry of data analytics in the interpretation and use of big data. These data analytics companies act as intermediaries in the digital data revolution. Understanding the social influence of big data requires us to understand the role played by data analytics within organisations of different types. This particular article focuses very specifically upon the way in which data and data analytics are envisioned within the marketing rhetoric of the data analytics industry. It is argued that to understand the spread of data analytics and the adoption of certain analytic strategies, we first need to look at the projection of promises upon that data. The way that data and analytics are imagined shapes their incorporation and appropriation into practices and organisational structures – what I call here the *data frontiers*. This article draws upon a sample of 34 data analytics companies in order to explore the way in which data analytics are envisioned within that increasingly powerful industry.

ARTICLE HISTORY

Received 18 July 2016

Accepted 25 January 2017

KEYWORDS

Data; big data; data analytics; power; capitalism

Introduction

The dominant contemporary tropes about data, with their focus upon the influential social presence of *big data*, could easily distract us from the machinations of power that are at play. Indeed, one of the arguments I will make in this piece is that it is this ability to conjure such hype or ‘cyberbole’ (Woolgar, 2002) that is central to facilitating the spread of data-led practices throughout the social world. The sense that there is an overwhelming deluge of data today, however historically accurate this may be (see Beer, 2016), means that when it comes to the so-called data revolution the power is firmly in the hands of those who are able to interpret or tell stories with the data. In short, the power of data is located in what they are used to reveal – giving a unique primacy to those who are in a position to engineer those revelations. The data themselves come to life and begin to have consequences when they are analysed and when those analyses are integrated into social, governmental and organisational structures. Yet we have so far given very little attention to the powerful role – both technical and rhetorical – played by the emergent industry of analytics that has come to fill this analytical space of social ordering.

Clearly, there has been a drastic escalation in the data available to organisations of various types (see Kitchin, 2014). As a result, the powerful intermediary role played by the

analytics industry has taken on greater significance. Within notions of a ‘data deluge’ or a ‘data revolution’, we have tended to miss out these intermediaries and the role they play in realising the transformations around big data. As we go about ‘socialising big data’ (Ruppert et al., 2015), we will need to bring these intermediaries into focus. That is to say that we have given very little attention to the influential role that is played by those who are locating value, narrating and attaching meanings to that data. Given that we know that these data are not ‘raw’ (Gitelman, 2013), we have perhaps given too little attention to the data analyst and to the products and services of the data analytics industry. A powerful new assemblage of human and non-human actors now combine as *data intermediaries* within this apparently overwhelming accumulation of data. This analytic assemblage performs a central part in how data is captured, presented, interpreted and then incorporated into organisational structures and decision-making.

As this might suggest, the analytics industry’s role is likely to be varied and complex, not least because its actors are both creating *as well as* responding to these visions of data. As the availability of new types of data has grown, and as the vision of a data deluge have been conjured in the powerful discourse of the data industry, the role of finding something comprehensible to say about this mass of accumulating data has taken on greater significance (as discussed in boyd & Crawford, 2012). The result is that we have seen a powerful new industry of data analytics emerge. If, as I have suggested, the power of data is in the insights that they are used to garner, then this data analytics industry is powerful in shaping what is said, made visible or known through data. It is crucial then to explore the growing analytical and interpretive power of this emergent industry as its reach spreads out, at different scales, across different social spheres. Only then, building upon the recent work on the practices of data mining by Kennedy (2016), can we begin to understand the role of data or big data in the social world. To understand big data is to understand how data are analysed and the types of analytic strategies that are being deployed. In addition to this, this article suggests, it is also important to explore the seductive visions of data analytics that enable these analytics to spread out into organisations of different types as well as into individual lives.

This industry of analytics, as we will see, positions itself as providing solutions to the data deluge. It provides, we are told, opportunities for anyone to see things with the data. It provides services for a kind of data led ‘telling about society’ (Becker, 2007) through intuitive and accessible analytics packages. As such, the interpretive role of the data analytics industry now needs to take centre stage if we are to fully understand the way that power is exercised or deployed through data – ranging from larger to smaller scale providers. This article is part of such a project. It aims to illuminate the discourse that surrounds the analysis of big data and to show the powerful ways that the data analytics industry frames and projects the analytical prowess of data. We are not then just to be concerned with practices but with the way in which this industry cultivates a particular type of vision of data and its possibilities. This will allow us to see not just what is being done with the data, it will also allow us to see how those data analytics are presented, to what end, the type of organisation they are attached to and the sort of imagined futures they use to facilitate the embedding of data analytics within everyday organisational life. The growing power of the data analytics industry, I suggest, is not just in what they

do but in what they promise – in a similar way that expectations and promises play a central part in the ‘dynamics’ and ‘momentum’ of developments in science and technology (Brown & Michael, 2003). The allure and seductive envisioning of the possibilities of data is a key facilitator for the adoption, incorporation and spread of data-led processes. These visions might be seen to be part of the broader imaginary upon which organisational practice is based (see Knorr Cetina, 1994, pp. 7–8).

With these aims in mind, this article uses the public marketing materials from 34 data analytics companies (see the [Appendix](#)) in order to explore how data analytics are imagined or envisioned across the sector (further discussion of the sector breakdown of the companies covered can be found in Beer, 2017). As I have suggested, taking such an approach enables us to understand *the conceptual making-up of data and data analytics* – which is part of understanding how these data circulate into and become part of the social world. The point here is that the promises that are being made act as the incentives to the expansion of ‘datafication’ (Van Dijck, 2013). These are the *data frontiers*. By which I mean that these are the boundary lines at which data-informed processes reach their limits. Data frontiers are what we might think of as the limits of datafication. Datafication has edges or boundaries. Those invested in this industry are inevitably seeking to push these boundaries back. It is at these frontiers that these visions of the power of data analytics do their work. They are active in ushering-in the expansion and intensification of data within organisational and social structures. These visions of the power of analytics persuade, enact and produce notions and ideals that are designed to force back those data frontiers. If we are to understand the expansion and intensification of data as an active component of the social world, then we need to look at the work that is being done at these data frontiers. However, these are not simple boundary lines, rather these frontiers work in two ways. First, we have the entry of data analytics into previously uncharted and new territories. These are social territories into which data analytics is yet to spread. And, second, we have frontiers at which data is already active but where there is the pursuit of the further enhancement and escalation of those data processes. This is where the analytics industry is attempting to intensify data use. The more powerful the envisioning of data analytics’ powers, the more porous these data frontiers are likely to be.

As such, to understand the influence of the data analytics industry is to look at the promissory visions they create around data and its analysis. To this end, this piece focuses upon a selection of data analytics companies as typical instances within an emergent and burgeoning industry. It is in the marketing materials of these companies and the presentation of services and solutions that they provide that data are most actively being made-up. The article uses this sample of 34 analytics companies to look at how data analytics, as both practices and processes, are being constructed in the discourse and rhetoric of those who are facilitating its spread and diffusion. This is not to say that these visions of the power of data analytics will always equate directly to the realities of organisational data use, but it is to say that these visions will feed into those organisational practices, perceptions and structures. We need to see the promises that are being made in order to understand how these particular types of data analysis, and the thinking behind them, work their way into the structures in which we live. Data and data analytics do not somehow exist outside of this projection of its promises. Or as Ang (1985, p. 83) famously put it, in a very different context, ‘a life without imagination does not exist’.

Locating the data analytics industry

The aims of this project required the identification of a series data analytics organisations whose activities could then be examined. The sample was created by first searching on Google for three different combinations of terms (the Google search was used as it was imagined that this is likely to be where organisations start when they are looking for analytics expertise). These search terms were: (1) data analytics companies; (2) data analytics organisations and (3) data analytics solutions. These were felt to be the most appropriate terms that were likely to be used by organisations who were looking to try to locate data analytics services. The sample was then created using search terms 1 and 3, which created the most useful and extensive lists of the type of organisations that were being sought. Two different approaches were then used to create a list of organisations that varied in type. The top 10 for search term 1 included two recent magazine articles that provided overviews of a range of data analytics companies. *NetWorkWorld*¹ and *Forbes*² magazine had both published articles on big data companies ‘to watch’. These lists were used to create a list of data analytics companies that were in some way notable in the industry. I used these two magazine lists, visiting the websites of the named analytics companies, and included within my sample any companies that described themselves as providing data analytics (see the [Appendix](#)). It was felt that some supplementary examples beyond those contained in the magazine articles were needed. I used search term 3 to locate a further six data analytics providers. To do this, I simply selected the first six companies that were listed in my Google search that in some way identified themselves as providing data analytics (excluding those that had already been included in my sample as a result of being named in one of the magazine lists). This created a sample of 34 data analytics organisations of different types, ranging from consultancy to software package providers.³ The sample of data analytics organisations is listed in the [Appendix](#).

Once the sample had been established I then explored the marketing materials as published on the public websites of each of the organisations in the sample. In each case, and in line with the aims of this article, I looked specifically for instances in which the data analytics services provided were being described, rather than any other types of services. This enabled me to focus specifically upon how these data analytics were being presented in the materials provided by each company. As well as focusing upon content concerning data analytics, I looked for three more specific things to help to guide the analysis. First, I looked at the types of services and solutions that were on offer. This was to explore the different types of data analytics that were being presented and to see what types of problems or opportunities these solutions were said to be offering. Second, I looked at the scale and scope of those services and solutions. That is to say I looked at the size of opportunities that were presented, the notions of scale that were being discussed, the inclusion of any geography or size of data set, and the type of coverage that was being offered. Third and finally, I also focused attention on the promises, hopes and expectations that were linked to the data and its analysis. What was it said that they offer? What changes and transformations were data analytics said to bring about? These were the key points of focus that were used to navigate the resources and provide some direction in extracting materials that spoke to the core aims of this project.

The data analytic imaginary

From exploring the materials in the way described above, a number of themes occurred frequently in the content. Six themes were found to have particular prominence in the materials. Each theme emphasises a property of data analytics that might be seen to provide a seductive introduction to its use. This is, of course, to be expected. We are looking here at marketing materials. We would expect them to attempt to sell the features and benefits of data analytics to an imagined customer. To reiterate my earlier point though, this requires us not to see this simply as an exercise in promotion but as a series of attempts to instigate, facilitate and afford the expansion of data-led processes of evaluation, judgment and decision-making. This is the rhetoric aimed at oiling the spread of big data and the type of calculative judgments, ordering and evaluation that it brings. Seeing the detail of how big data analytics are constructed in sales pitches to organisations will help us to understand more clearly how big data analytics spread across sectors, how they are understood and in what form they come to embed themselves in our everyday lives.

What emerged from this was a very specific *data analytic imaginary* in which the data analytics were presented as *speedy, accessible, revealing, panoramic, prophetic* and *smart*. The rest of the article explores the details of the features of this data analytic imaginary.

Speedy

If there is one overriding message of the visions conjured by this set of materials, it is that data and data analytics are speedy, quick, fast and rapid (for a detailed discussion of this see Beer, 2017). They do not hang around dwelling over possible insights, they are instant and continuous. We are told that they allow organisations to respond to their data in ‘real time’. Thus, these are data analytics that allow the viewer to be *in the moment* and to react without delay or hesitation to the changing scenes that they see unfolding. This is data mobility in real time. The insights produced by the data then are seen to be a representation of the world as it unfolds – rather than being a reflective process of looking back. There is no delay or gap in the knowledge being produced, rather these data analytics are depicted as providing continuous encounters with the actually existing world as it is in that moment.

There is an interesting temporality at play here. The data produce insights into those moments, but this is also an ongoing and continuous process. The vision is of an always switched-on presence of analytics in the background, leading to moments of punctuation in which a response, decision or reaction is needed. The quicker the analytics, so the logic goes, the better and more useful it is. Quick analytics lead to responsive, flexible and successful organisations, it is implied.

Part of this speed is tied into the envisioning of the power of these analytics. These are presented as being an invitation to ‘Supercharge’ (Appendix, ref 14) or ‘turbocharge’ (Appendix, ref 10) your analytics. These analytics are not only described as fast, they are also powerful or mighty – the sense is of them having horse-power under the bonnet. The use of engine metaphors is telling here, with the analytics machine being like a suped-up high-performance engine. With extra power, torque and superior revolutions per minute, at least that is what is conveyed. The usual ‘metaphors of big data’ (Puschmann & Burgess, 2014) are not usually far away.

The result is an intensive form of knowledge that arrives in real time ... continuously. This runs in parallel with broader notions of a kind of jam-packed, inescapable and real-time experience typical of 'intensive culture' (Lash, 2010). A flow of knowledge and insights gushes, yet the complexity of the data is simplified to enable its quick and rapid usefulness. This is to quickly find accurate insights without hesitation, despite the massive deluge. As is illustrated by the passage suggesting that this is: 'Speed at scale. Trillion rows in 3 seconds, billions in less ... provides access to 100% of the raw event data with the speed to easily ask a series of questions in seconds, without the consequence of being wrong' (Appendix, ref 7). Another describes it as a combination of 'SCALE, SPEED, AGILITY' (Appendix, ref 25). Again, we see claims about the speediness with which insights are drawn from the huge accumulating data as it is harvested. Enabling a quick response. In another instance this is described as closing the gap between the gathering of data and its use in making decisions: 'achieve rapid insight into action across your organisation, closing the gap between transactions, data preparation, analysis, and action – all with analytics' (Appendix, ref 34). Closing this gap is seen to be an acceleration of the working practices of organisations, who are then able to react nimbly and instantly to those insights. With their being no gaping fissure between data capture and analysis. The claim is that by drawing on such analytics you can:

Break the speed limit at your desk. Ready. Set. Done. Platfora's in-memory query engine and massive parallel processing architecture let you crunch petabytes of data at the speed of thought—your thoughts, that is. (Appendix, ref 28)

Again we see the combination of scale or volume with the speed of the extraction of analytical insights. The vision of the breaking of speed limits, reaching new speeds of insight, whilst at your desk is particularly telling.

Taken together, these 34 examples describe speedy analytics in which masses of data can continuously be drawn upon to inform and enlighten organisational processes and decision-making. The image is of a more intensive organisation that reacts to real-time knowledge about themselves. The gap between data and knowledge then is depicted as closing. These are speedily accumulating big data being rapidly analysed. What we are seeing here are attempts to provide solutions that are seen to be able to cope with broader notions of cultural speed-up and acceleration (as discussed in Tomlinson, 2007; Wajcman, 2015). The idea being that companies that use data analytics can accelerate their practices and keep-up with what is perceived to be an accelerating world and set of competitors.

Accessible

Alongside speed the accessibility of these analytics is a second key theme. A dominant idea here is that knowledge can be gleaned without any real technical know-how. Thus, these are instant analytics both in the sense that the data analytics produce instant results, but they are also instant in the sense that users can readily interpret and understand the analytics that they are encountering. The data might be incomprehensible but the analytics are intuitive and can therefore be easily accessed and understood. The software is doing the work for you. The software provides the analyses from which *anyone* can draw inference. Thus, the data is rendered accessible and the analysis requires little training or expertise. The software becomes the expert intermediary in the data analysis relationship that is

being envisioned here. The growing role of the expert and the consultant has been noted in some important recent contributions (Amoore, 2013, p. 6; Davies, 2014, p. 30). The message conveyed in these marketing materials though is that you can become your own expert and consultant, with the help of these easy to use, proactive and intuitive software packages. The software turns the user into their own analytics expert.

Simplification of complexity is the key underlying message. They are products and services designed to make your data accessible the words ease, easy and easily appearing frequently. The visualisation of data is one particular area of focus in the accessibility of the data. Here, the work being done by the visualisation (see Kennedy, Hill, Allen, & Aiello, 2016) is to render data amenable to instant and simplified analytical insights that are manageable by anyone. These visualisations are presented as being a translational device for enabling the untrained eye to easily extract knowledge from the data. The visualisations are presented as being one means by which these devices are making data manageable, comprehensible and instant. Again, this is seen to be part of the by-passing of technical skill. To communicate this type of empowerment and accessibility terms like 'self-service' (Appendix, refs 3, 7, 11, 12, 23 and 33) and 'do-it-yourself' (Appendix, ref 28) are frequently used. Becoming an analyst and taking on these self-service software enables the user to 'Take control of your data' (Appendix, ref 5). This suggests that these are solutions that enable organisations to engage in their own data analytics so as to empower them in their data usage. This is a set of data analytics then that are designed to give the impression of by-passing the expert third party and are instead about making everyone their own data analytics specialist. The industry of analytics appears to be premised upon its ability to turn anyone into a data analyst.

In terms of access then, what we have here is something close to what Graham (2004) has called the 'dreams of transcendence' or the 'anything-anywhere-anytime-dream'. This is most directly articulated in this claim: 'Your business is changing and you need an easy, visual way to explore your data ... [this] enables you to analyze any data, anytime, anywhere' (Appendix, ref 33). The data analytics then are seen to be accessible at anytime or from anywhere, thus the temporality and spatiality of data analytics are not anchored to organisational buildings or to fixed working hours. The analytics are accessible both in terms of time and space. The analytics may be 'all in one place' (Appendix, ref 28), with everything reduced to a single package, but this place is mobile and transient.

Revealing

Similarly, trust is an issue in the coverage of the insights that the data offers. The insights are depicted as being trustworthy and accurate. The industry attempts to tap into and promote what Porter (1995) has referred to as a 'faith' or 'trust' in numbers. Data analytics are presented as being about objectivity and efficiency in the production of insights (see also Beer, 2016). These revelations are said to be limitless in their flexibility, there are no boundaries to data or to what they facilitate for organisations.

The implicit theme here is that there is a raft of untapped potential to be found in the data, it is this potential that can be unlocked through analytics. Analytics are presented as being the means by which 'hidden' (Appendix, refs 5, 12 and 22) value might be unearthed or where new types of value might be tapped. As it is put in one instance, there is a potential 'gold mine' (Appendix, ref 16) of information. This then is about creating or extracting

hidden value. The images are of masses of wasted unused data that could be potentially lucrative. The analytics packages and solutions are said to provide opportunities to salvage this unused data waste.

The revelations are not seen to end there though. Rather, tapping into these hidden insights is said to provoke the newly informed analyst to ask new questions and seek new insights. Again, as with the notion of self-service, these analytics packages are seen to provide opportunities for self-training and for radical culture-shifts within organisations. This is about organisations being charged with taking control of their data. The suggestion is that the insights produced will motivate and inspire those who use it to seek more and more analytical insights and to ask new questions about their organisations. These analytics then are designed for the 'curious' ([Appendix](#), refs 27 and 28), those who want to 'look closely at what others ignore' ([Appendix](#), ref 27) or who want to 'solve really hard problems' ([Appendix](#), ref 28). Analytics provoke or stimulate 'curiosity' ([Appendix](#), ref 29). Data analytics, in this formulation, are said to breed and satisfy curiosity and the pursuit of hidden value. This then is an invitation to dig. As such the impression or connotation is that these data analytics, far from producing passivity, will make the user less dependent on others in understanding their business – with new activities and ways of thinking burgeoning from those revelations. The idea appears to be that insights breed a desire for more insights.

Panoramic

As might then be expected, these data analytics are depicted as being all-seeing. They offer a kind of prosthetic eye with which to see the data that is accumulating. Data analytics are presented as almost omnipotent. They are inescapable and comprehensive in their scope, vision and sight. Nothing escapes this prostheticised and technologically enhanced vision. These data analytics, it is suggested, allow their user to 'see around corners and into the future' ([Appendix](#), ref 28). Data analytics shine a light on blind-spots, those parts of the organisation hidden around corners suddenly become visible. Nothing is in the shadows. Data analytics are said to provide a '360-degree view' ([Appendix](#), ref 12). Hence data analytics are described as having a kind of panoramic view in which nothing is outside of the knowledge that is produced from the data. The view from that position is envisioned as comprehensive, despite the scale and density of the data. A variety of terms are used to account for the way that the whole is made visible in this way. It would seem that it is by bringing together different forms of data that these analytics are able to expand the scope of their sight. The focus here is upon the ability to draw together different data streams in order to increase the coverage of the analytics. In this regard terms like 'integrated' or 'integration' ([Appendix](#), refs 1, 13, 25, 30 and 34), 'blending' ([Appendix](#), refs 17, 23, 24, and 33), and 'harmonization' ([Appendix](#), ref 17) are used to suggest the seamless use of variegated data to produce a coherent and comprehensive set of insights.

As well as providing a panoramic view of the internal workings of organisations, these data analytics are also described as providing a perfect panoramic view of the exterior context. The organisation using these analytics is imagined as being able to fully understand their position within the world of capitalist competition. This external panorama is constructed through the data, so that the organisation might make use of its position and of its understanding of its global competitors. The emphasis is upon the use of international

data through which such analyses can be performed, and through which comparative forms of understanding might develop. The emphasis here is upon data analytics turning companies into world-players, or to enhance their position on the world stage. This is to take advantage of what is described as ‘this hyper-connected world, with data volumes constantly increasing’ (Appendix, ref 34). Another instance is the observation that ‘the world runs on data’ (Appendix, ref 32). The result is that data analytics allow for this to be drawn upon to ‘explore the world’s public data’ (Appendix, ref 5). Data analytics then are envisioned as providing a panoramic view of the interior and exterior world.

Prophetic

The visibility provided by data analytics, as has already been hinted, is not just about the moment in which it occurs. They are said to act in real time and provide a comprehensive view of the internal and external conditions, but beyond this these data analytics are also said to have prophetic properties. Data analytics open up a world in which it is possible to anticipate what will happen, and respond accordingly. Here in this rhetoric we see what Amore (2013) has called ‘the politics of possibility’, the use of imagined futures to make current decisions. As well as providing a vision of the moment in real time, these are data analytics that are strategic and that have an eye on the future. They are about seeing into the future to protect and maintain value and a competitive edge. We are told that we might anticipate rather than respond through these analytics. Espeland and Sauder (2007) speak of ‘reactivity’ to data, where we change our behaviours in response to the data being gathered, here we see anticipation being ramped up and reactivity being folded into imagined futures.

An underlying narrative here is that data analytics are a necessity for progress. It is what intelligent forward-thinking organisations are seen to do – data analytics are conjured as being the desirable direction for all organisations. They aim to predict, forecast and bring the future into the present. This is even described as being part of a ‘New industrial revolution’ (Appendix, ref 16), which we might interpret as a suggestion that there is a marked epochal change brought about by data and its analysis. A key aspect of this then is obviously the ability to predict behaviours. As it was put in one instance, ‘the next generation of analytics will let you see patterns for predicting future behaviors, not just analyzing those in the past’ (Appendix, ref 12; see also ref 13). Seeing patterns and making predictions are a key part of these messages. Not just seeing the past but also then using patterns to see the future.

In this regard, the emphasis of these data analytics pitches is upon a combination of having a clear sense of the future whilst also being flexible enough to respond to current changes. Or as it is put: ‘Future-proof your organization without ever getting locked in’ (Appendix, ref 14). Being able to predict the future is seen to be of extreme value in remaining competitive or in giving an edge. The analytics industry draws in potential customers with seductive leading questions such as ‘What if you could accurately predict your customer’s behavior?’ (Appendix, ref 8). The idea is that by predicting you are able to anticipate what people will want and shape your business accordingly, thus protecting its future value. The emphasis here is upon visions of strategic and intelligent thinking – a kind of predictive intelligence emerges in this rhetoric. For example, one package is said to provide ‘predictive intelligence to help your agents be more productive and

focus where customers need them most' (Appendix, ref 10). It is conveyed that prediction and predictive intelligence enable enhanced value extraction. In another instance this is described as 'deep predictive analytics' (Appendix, ref 11). Indeed, this seems to be an attempt to go beyond the frequently mentioned 'predictive analytics' (Appendix, ref 31), so as to suggest that there is a depth to the insights that might be generated by the 'predictive capabilities' (Appendix, ref 13) of these data analytics.

Data analytics then are presented not just as enabling future sight, but also are able to bring those desired futures into existence. This then is about using data to see and then manipulate possible futures through current action and decisions. It is about making decisions informed by those imagined futures. The implicit logic of adapting and incorporating such methods is that it is the best or only way to remain competitive and to ensure a safe future. Without such analytic expertise the impression is that the result will be that such organisations will be left behind. The analytics industry is built upon this idea that data should be used to your advantage. Thus companies need comprehensive and holistic analytics that capture everything, or 'an end to end data solution' (Appendix, ref 23), to facilitate data-led decision-making. Embracing analytics is positioned as forward-thinking.

Smart

The above shows that the predictive capacities of these data analytics are often associated with a latent intelligence that resides in these systems. They can learn in order to predict. The notion of learning is important here. The analytics are not passive, rather they are presented as being intelligent and active devices that are able to learn, adapt and develop in the insights that they actively produce. Thus they respond to particular needs, learning what is required, whilst also learning from the data that is accumulating.

Indeed, the word 'smart' is used a good deal to evoke this latent intelligence and learning power. This deals with both the smartness of the analytics and the smartness of the insights that they create – and then the implied and inherited smartness of the organisations that use them. For instance, we are told that these analytics bring 'Smarter answers to big data questions' (Appendix, ref 26) and 'Better Analytics, Smarter Decisions', (Appendix, ref 34). A big part of the smartness described is associated with the type of 'machine learning' or 'machine intelligence' that resides within these analytics packages. With one claiming that it:

deploys world-class machine learning algorithms to help you predict your customer's monetary value. Our proprietary algorithms are self-learning so they automatically improve and adjust throughout your customer's lifetime. (Appendix, ref 8)

Here, the algorithms take on the thinking, with the analytics being based on forms of machine learning. These algorithms do not need training, rather they do the learning for themselves, they are a 'self-learning' technology able to adjust discoveries. This is said to create improved results, as it is put in one instance: 'Machine intelligence delivers better outcomes' (Appendix, ref 18). Again there are connotations of progress woven into this, that data analytics provide the only sensible future for social ordering, governance and decision-making processes. Also there is the connotation that the thinking, the difficult work of the analytics, can be done by the analytics software. This reinforces those

earlier notions of accessibility, with the self-training algorithms helping to provide self-service analytics to the untrained user. The software is training itself so that the interested user does not need to. This creates a set of questions around agency, with the data analytics software often seen to be doing much of the thinking (which relates to debates about the relations between data and agency raised very recently by Kennedy, Poell, & van Dijck, 2015).

Indeed, the theme of self-learning and machine intelligence are woven through many of the instances in this sample. With the data analytics seen to be responsive as they learn about the data and the demands of their use. We get claims about the analytics such as 'backed by self-learning algorithms that tune for actual heavy query patterns' (Appendix, ref 1) or 'Advanced machine learning algorithms learn as you go, allowing you to focus on insights' (Appendix, ref 3). In this second instance, we see again the point that the software does the analytical thinking so that the human actor is able to focus on the outputs of those analytics. Again, little technical knowledge is needed, this is presented as 'out-of-the-box Artificial Intelligence tools' (Appendix, ref 8). There is 'built-in intelligent data inferencing' (Appendix, ref 17). The image is of ready-made and thinking technologies that carry the burden of technique, know-how and method. It is ready to do the thinking for you.

The consequence of this is that these thinking technologies are seen to find and locate value in a way that the human alone is unable to do. As this passage illustrates, where the claim is that a particular analytics package 'leverages advanced machine learning algorithms to create significant business value and insights correlating internal and external data for any Enterprise' (Appendix, ref 8). This type of 'machine learning' (also in the Appendix, ref 10), is used to find value through algorithmic means. The algorithms are depicted here as active, autonomous and learning components within the software. The algorithms are described as being the means by which this learning takes place and are able to adapt to the data and to the types of insights that are desired. The software package is not then a complete product, but rather it is seen to be an organic system that actively evolves in its powers over time. As this instance suggests: 'Our proprietary algorithms are self-learning so they automatically improve and adjust throughout your customer's lifetime' (Appendix, ref 8). This image is of a responsive and changing technology that hones its own analytics. This 'self-learning' means that these packages 'update continuously without manual intervention' (Appendix, ref 10). The result, it is suggested, is 'blending machine learning with human intelligence' (Appendix, ref 24). It is not that humans are cut-out of the equation, but rather that they work alongside these thinking technologies. These are presented as 'an intelligence engine that uncovers hidden insights in data and supports automated decision making' (Appendix, ref 5). Yet this does not mean that the human actors are described as passive. Rather we frequently see the learning and thinking analytics being presented as an aid to human decision-making. The software even takes on a kind of anthropomorphic presence, rendering it easier to deal with and again emphasising its intuitiveness and ease of use. As this excerpt demonstrates: 'their experience feels less like it was produced by a machine, and more like it's coming from a friend' (Appendix 1, ref 19). In this case, the suggestion is that the interactions between human and machine are facilitated and made seamless by the machine behaving like a human presence. The result is that these new forms of agency or intelligence in machine

learning and algorithms are presented in terms of their ability complement, extend and blend with human agency.

Conclusion

This article has explored the rolling-out of a logic, a way of thinking, alongside the rolling-out of a set of practices. As I have outlined, a very particular rationality emerges here – a rationality that promotes quick and accessible know-how that enables all-seeing predictive and smart decision-making. A world in which anyone can be their own data analyst. It is hoped that this article has provided some insight into how it is that data analytics are envisioned. These are powerful promises that are active in shaping and pushing back *data frontiers*– expanding both the reach and intensity of data-led processes. The logic is hard to resist. Who would not want to follow the logic of the six characteristics of the data imaginary that I have pulled out in this article? The seductive allure becomes clear to see; it also begins to suggest how these characteristics might become a part of how data analytics are subsequently folded into organisations. The emergent industry of data analytics – providing both solutions and a compelling rationale – is powerful in its intermediary role. It is an industry built upon the idea that the accumulation of data needs a response, and that the only logical response is to use as much data as possible.

We have seen the rise of what might be thought of as a new set of *data intermediaries*, an assemblage of human actors, code, software and algorithms that are active in shaping the circulation and integration of new forms of data. These data intermediaries are active in building the infrastructure central to what Pasquale (2015) has called the ‘black box society’– a society with a range of underpinning algorithmic processes in which we have little knowledge or understanding of the secrets it holds or the knowledge it obfuscates. Yet, the power of the data analytics industry is not just to be found in what it helps organisations to know or say with their data. Rather, a significant part of the power of the data analytics industry is in how it actively envisions data and data analytics. It is here that data gains its compelling edge. It is here that data analytics are presented as a competitive necessity. It is here that data analytics are imagined and thus where their incorporation is instigated. So, it is not enough to say that data analytics are being sold in this marketing rhetoric, rather we need to look at the detail of exactly how data and data analytics are being made-up in this marketing speak. The power of the data analytics industry is two-fold: it is to be found in both the way that the data analytics are envisioned as well as in the actual data solutions that are then deployed on the ground. Without the former, we do not have the latter. It is the compelling rationale that ushers in these practices.

The data analytics industry is inevitably attempting to theorise and imagine a future that potential customers will be drawn towards. The result is that they conjure up a range of possible futures for those organisations so as to pull-them into datafication. The shock of the new is tempered by reassuring images of a successful and desirable future and a progressive organisational way of life. Data and their analytics are presented as being a powerful ongoing and permanent presence, giving constant insights that are always there. Those insights are then readily available in easy to consume forms that require little technical expertise – leading to judgment without know-how. These analytics reveal hidden value in the data, they shine a light on organisations and show things that were previously invisible. They enable the future to be seen, and for an imagined future to be part of

the present decisions that are taken. They see everything, in detail, nothing escapes their sight. Their vision is omnipotent and sharp. And then finally, data analytics are smart. They are the smart thing to be involved with, giving the edge over competitors. But these systems are also themselves smart, they blend machine and human intelligence to enhance insight. This is the data imaginary in action.

What we have then is a powerful logic that gives some explanation as to why data analytics spread so rapidly through organisations and through the social world more broadly. The data imaginary is charged with pushing back the data frontiers. This imagery is an active presence that oils those circulations and shapes their pathways. To understand the role of big data in the social world requires us to understand the role of the data analytics industry as both an active presence in the use of data and as a promoter of data-led thinking. In these visions of the power and promise of data analytics we are looking at the cutting-edge of broader processes of data centred forms of capitalism, competition and calculative logics.

Notes

1. This article was published in *Networkworld* on the 13 August 2015 and is available here <http://www.networkworld.com/article/2970498/big-data-business-intelligence/13-big-data-and-analytics-companies-to-watch.html>.
2. This article was published in *Fortune* on the 13 June 2014 and is available here <http://fortune.com/2014/06/13/these-big-data-companies-are-ones-to-watch/>.
3. The materials drawn upon in this sample were all accessed between the 25 November and the 2 December 2015.

Disclosure statement

No potential conflict of interest was reported by the author.

Notes on contributor

David Beer is Reader in Sociology at the University of York. His most recent book is *Metric Power* [email: david.beer@york.ac.uk].

References

- Amoore, L. (2013). *The politics of possibility: Risk and security beyond possibility*. Durham: Duke University Press.
- Ang, I. (1985). *Watching dallas*. London: Routledge.
- Becker, H. (2007). *Telling about society*. Chicago, IL: University of Chicago Press.
- Beer, D. (2016). *Metric power*. Basingstoke: Palgrave Macmillan.
- Beer, D. (2017). The data analytics industry and the promises of real-time knowing: Perpetuating and deploying a rationality of speed. *Journal of Cultural Economy*, 10. doi:10.1080/17530350.2016.1230771
- boyd, d., & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society*, 15, 662–679. doi:10.1080/1369118X.2012.678878
- Brown, N., & Michael, M. (2003). A sociology of expectations: *Retrospecting prospects and prospecting retrospects*. *Technology Analysis & Strategic Management*, 15, 3–18. doi:10.1080/0953732032000046024

Davies, W. (2014). *The limits of neoliberalism*. London: Sage.

Espeland, W. N., & Sauder, M. (2007). Rankings and reactivity: How public measures recreate social worlds. *American Journal of Sociology*, 113, 1–40. doi:10.1086/517897

Gitelman, L. (ed.) (2013). “Raw data” is an oxymoron. Cambridge, MA: MIT Press.

Graham, S. (2004). Introduction: From dreams of transcendence to the remediation of urban life. In S. Graham (Ed.), *The cybercities reader* (pp. 1–30). London: Routledge.

Kennedy, H. (2016). Post, mine, repeat: Social media data mining becomes ordinary. Basingstoke: Palgrave Macmillan.

Kennedy, H., Hill, R., Allen, W., & Aiello, G. (2016). The work that visualisation conventions do. *Information, Communication & Society*, 19, 715–735. doi:10.1080/1369118X.2016.1153126

Kennedy, H., Poell, T., & van Dijck, J. (2015). Data and agency. *Big Data & Society*, 2(July–December), 1–7. doi:10.1177/2053951715621569

Kitchin, R. (2014). *The data revolution*. London: Sage.

Knorr Cetina, K. (1994). Primitive classification and postmodernity: Towards a sociological notion of fiction. *Theory, Culture & Society*, 11, 1–22. doi:10.1177/026327694011003001

Lash, S. (2010). *Intensive culture*. London: Sage.

Pasquale, F. (2015). *The black box society: The secret algorithms that control money and information*. Cambridge, MA: Harvard University Press

Porter, T. M. (1995). *Trust in numbers: The pursuit of objectivity in science and public life*. Princeton, NJ: Princeton University Press.

Puschmann, C., & Burgess, J. (2014). Metaphors of big data. *International Journal of Communication*, 8, 1690–1709. doi:1932–8036/20140005

Ruppert, E., Harvey, P., Lury, C., Mackenzie, A., McNally, R., Baker, S. A., ... Lewis, C. (2015). Socialising big data: From concept to practice. CRESC Working Paper no. 138. Retrieved from <http://www.cresc.ac.uk/medialibrary/workingpapers/wp138.pdf>

Tomlinson, J. (2007). *The culture of speed*. London: Sage.

Van Dijck, J. (2013). Datafication, dataism and dataveillance: Big data between scientific paradigm and ideology. *Surveillance & Society*, 12(2), 197–208.

Wajcman, J. (2015). *Pressed for time*. Chicago, IL: University of Chicago Press.

Woolgar, S. (2002). *Virtual Society? Technology, Cyberbole, Reality*. Oxford: Oxford Univetrstity Press.

Appendix. The sample of data analytics companies

Reference number	Organisation name	Organisation URL
1	Arcadia Data	http://www.arcadiadata.com
2	Cazena	https://www.cazena.com
3	DataHero	https://datahero.com
4	DataTorrent	https://www.datatorrent.com
5	Enigma	http://enigma.io
6	Experfy	https://www.experfy.com/
7	Interana	http://www.interana.com
8	Neokami	https://www.neokami.com/
9	Mapr	https://www.mapr.com
10	Wise.io	http://www.wise.io
11	Paxata	http://www.paxata.com
12	Informatica	https://www.informatica.com
13	Syntasa	http://syntasa.com
14	Actian	http://www.actian.com
15	Tableau	http://www.tableau.com
16	Sight Machine	http://www.sightmachine.com
17	Clear Story Data	http://www.clearstorydata.com
18	Ayasdi	http://www.ayasdi.com
19	Wibi	http://www.wibidata.com

(Continued)

Continued.

Reference number	Organisation name	Organisation URL
20	Tamr	http://www.tamr.com
21	Trifacta	https://www.trifacta.com
22	Cloudera	http://www.cloudera.com
23	Datameer	http://www.datameer.com
24	Premise	http://www.premise.com
25	Palantir	http://www.palantir.com
26	Teradata	http://www.teradata.co.uk
27	Splunk	http://www.splunk.com
28	Platfora	http://www.platfora.com
29	Avalon	http://www.avalonconsult.com
30	Das	http://www.dasconsultants.com
31	CSC	http://www.csc.com/big_data
32	Avanade	http://www.avanade.com
33	Oracle	https://www.oracle.com/solutions/business-analytics/index.html
34	SAP	http://go.sap.com/uk/solution/analytics.html