Etzion, D. & Aragon-Correa, J.A., 2016. Big data, management, and sustainability: Strategic opportunities ahead, *Organization & Environment*, 29 (2), 3-10.

BIG DATA, MANAGEMENT, AND SUSTAINABILITY: STRATEGIC

OPPORTUNITIES AHEAD

Dror Etzion Desautels Faculty of Management McGill University 1001 Sherbrooke Street West Montreal, Quebec H3A 1G5, Canada <u>dror.etzion@mcgill.ca</u>

J. Alberto Aragon-Correa

Surrey Business School University of Surrey Faculty of Arts and Social Sciences, Guildford GU2 7XH, United Kingdom

j.aragon@surrey.ac.uk

BIG DATA, MANAGEMENT, AND SUSTAINABILITY: STRATEGIC OPPORTUNITIES AHEAD

Abstract:

We contend that big data and management for sustainability are very good bedfellows, in that many of the affordances big data provides are naturally aligned with sustainability concerns (e.g. multidimensional nature, collective actions, smart allocation of resources, efficiency priority). Notwithstanding this promising stepping off point, and the enticing analytical opportunities that an abundance of data will generate, we provide some reflections on big data and the most promising avenues of research it might inspire in the field of management and sustainability. In the first part of our essay we explore what managers can do with big data to reinforce organizational sustainability and how different operational, strategic, and corporate activities are affected in this process. In the second part we focus on what big data allows researchers to explore and examine, ranging from sustainability job descriptions through environmental metrics to industry transformation. We conclude by advocating for strong theoretical orientation in research on and with big data.

Keywords:

Big data, organizations and natural environment, sustainability, management, strategy.

It's alluring, and perhaps even joyous to engage with the linguistic creativity and metaphorical hyperbole spawned by the ubiquity of bits and bytes in our lives. Big Data, The Internet of Things, and the Deep Web have captured the imagination not only of Silicon Valley, but also of other industries, governments and non-profits all developing novel offerings that employ vast troves of data. Many successful and well-known firms focus a substantial part of their business around big data or use big data to enhance competitive advantage, including Netflix, Uber, Facebook, Google, and a host of insurance providers, among others. In general, media outlets, bloggers, activists and scientists regularly provide deep dives and fascinating visuals that help us to see trends distilled from massively sized and unimaginably complex information. Even the most interesting fodder for investigative reporting increasingly originates with the release of large, unfiltered data dumps that must be undergo intense analysis in the pursuit of meaning.

The data revolution is here, clearly. Happily, its potential has not been lost upon people actively working on sustainability challenges and opportunities. Software provided by SAP and its competitors in the enterprise planning space manages vast information flows for increasing operational efficiencies, but these tools provide SAP clients direct sustainability benefits in terms of energy efficiency, risk management and sustainability reporting, all of which are modules that SAP offers its clients. Another example is IBM's "Smarter Cities" initiative which aims to help city governments position their cities as attractive through reducing traffic jams, improving waste collection, and better managing water. Less time in traffic provides a better lived experience for residents, and efficient use of water implies lower operating costs, but (as any sustainability researcher will be quick to point out) these are win-win solutions in terms of the sustainability benefits they confer.

Our goal in this introductory essay is not to elaborate at length what big data is or isn't (see George, Haas, & Pentland, 2014; McAfee, Brynjolfsson, Davenport, Patil, & Barton, 2012 and others for a more detailed descriptive review), but a few important attributes nonetheless might bear repeating. Paradoxically, there is a wide consensus that "big" is not the defining feature of big data (Constantiou & Kallinikos, 2015). It is well accepted that big data differs from "regular" data along four dimensions, "4 V's": volume, velocity, variety, and veracity on data (e.g. Abbasi, Sarker, & Chiang, 2016). Big data can be conceptualized as a firehouse of unstructured and potentially heterogeneous information from various sources with different levels of integrity. Therefore, when it comes to lots of data, more isn't just more, "more is different", as proclaimed by Wired magazine (Anderson, 2008), And, as a consequence, "the developments with which big data is associated establish a new and distinctive context for data generation and use" (Constantiou & Kallinikos, 2015: 48). It's not so much what the data are; it's what the data enable or create. The real prize with big data is to generate insights that are unattainable with smaller amounts of data.

A second important attribute of big data is that it overcomes limitations of traditional sampling. Often, with big data the entire population is too vast to be handle even with high-powered computers, and a sample is preferred (Varian, 2014), but in general, big data enables less inference, inclusion of more heterogeneous perspectives, automatized collection of information, and more certainty. Because pattern is paramount, big data venerates correlation and eschews causality, with some proponents going so far as to suggest that big data renders theory obsolete. Others counter by stating that in fact we need more theoretical perspectives in order to better understand the complex reality captured and generated by big data. For researchers, who traditionally live or die by describing the world through lenses of theory, big

data presents a bewildering context, because it may require theoretical thinking about an antitheoretical, or agnostic context.

In the sections below, we provide rudimentary thoughts about some questions and opportunities that big data presents for research on organizations and the environment, or sustainability more broadly. We structure the paper in two main parts: the first part explores what we can do with big data - what it allows us as researchers to explore and examine; the second part focuses on what big data might do to us – what transformations big data might trigger, and the effects that those transformations might have. We examine these first and second order effects on research through lenses that are in turn descriptive of the available opportunities and challenges of big data and prescriptive of the research needed on a topic of growing importance in business and other organizations.

Operational and Strategic Implications of Big Data in terms of Sustainability

As regards sustainability, the most obvious first order effects of big data are on organizational and inter-organizational operations. Functional areas in the firm may simultaneously better attain their objectives and generate less environmental impacts by using big data from different sources. First, smart sensors and the seamless communication between them enable real time adjustments of systems to maximize efficiency, by optimizing the delivery of components only to when and where they are needed, e.g. from the forest product upstream in the supply chain all the way downstream to the paper trays in customers' office printers. In general, big data may provide more realistic and timely sales forecasts and, as a consequence, lead to the reduction of stocks and waste, thereby saving energy and resources. In non-business domains, the real time utilization of big data may also generate more opportunities for effective monitoring of environmental situations. For instance governments and NGOs can be alerted to

suspected deforestation activity via satellite imagery, drone reconnaissance and detection sensors planted in sensitive areas. More and faster knowledge allows better reaction and adjustment to real world happenings.

Second, in addition to these real-time applications, big data can be mined to align our resource usage with the multidimensional forces of our behaviors, markets, and resources. Domestic smart energy meters may help individual consumers reduce costs and energy consumption, but the big environmental payoff might come on the supply side if utility companies that will be better able to predict demand, allowing them to more sparingly employ carbon-intensive peaking power plants during periods of high demand (or, in an ideal scenario, retire them entirely). Agricultural practices can also become more efficient by turning on irrigation systems when a combination of factors happen including weather forecast, the *in situ* soil moisture drops below a certain level, temperature, and wished time to harvest considering the demand in the market. In general, an improved understanding of patterns of resource utilization yields better predictive capability, allowing organizations to allocate efforts and resources in a less reactive manner, reducing buffers and their inherent wastefulness.

A third, perhaps more tenuous link between supply, demand, and sustainability afforded by big data is the notion of mass customization, or the production of individually tailored goods and services to consumers, at scale. Conceptually at least, products designed and produced to accurately fulfill consumer specifications will yield outputs that are pared down to what users really want. These products, in turn, may attract a premium, may be designed more durably and will then perhaps be safeguarded and utilized for longer lifespans than typical goods and services. Similarly, big data may help create a more compelling value proposition for environmental differentiation strategies, by better adjusting product characteristics to the

willingness of customers to pay for these offerings. Credibility, a key constraining factor for many differentiation strategies (Reinhardt, 1998), can be enhanced through big data enabled chains of ownership, be it certified lumber or humanely raised chickens. In this way, big data may generate more opportunities for "green" customers to have their demands and preferences fulfilled.

Finally, another domain where big data is creating opportunities to improve daily operations around sustainability is risk management. For example, weather prediction is a notoriously data intensive activity whose reliability increases with larger quantities of information. Prediction of extreme events resulting from climate change inexorably improves as big data becomes increasingly easier to generate, store and process, allowing government agencies to better respond to catastrophes and insurance companies to better manage risk. Boutique and generalist consultancies are also translating sustainability related data analytics into service offerings for their clients, on issues relating to climate disruption, water management, and energy provision.

New business opportunities for management and sustainability

Contrary to many other technological developments in human history, the "big data revolution" was born not only digital but also "environmentally conscious". Many new companies formed by the opportunities big data creates highlight the environmental benefits inherent in changing patterns of activity. In other words, big data creates a large set of opportunities for business innovation around sustainability. It is not our intention here to provide an exhaustive description of the sustainable corporate opportunities for firms using big data, but rather would like to highlight several exemplars.

One of the most exciting ways to use big data is to trigger behavioral change towards sustainability. Various experiments and real-world experiences showcase the potential of big data presentation (e.g. comparing patterns of individual energy demands with peers) to generate changes in efficiency behaviors. These effects are supported by increasingly robust evidence (Asensio & Delmas, 2015) and are the main value proposition of oPower, a company that mediates the relationship between utilities and their customers, providing tailored nudges designed to unobtrusively encourage consumers to reduce energy consumption. Put simply, sustainability can be improved measurably through the use of big-data nudges (Thaler & Sunstein, 2008).

The rise of the "sharing economy" is often linked with the growing opportunities to collect and get benefits from managing multiple sources of data connected with some particular interest, and illustrates particularly well how big data and environmental benefits may come together in new business activities. As evidenced by platforms and apps such as Airbnb, RelayRides, Getaround, or Blablacar, "collaborative consumption" is a positive development for several reasons. Owners make money from underused assets and users pay less than they would if they bought the item themselves, or turned to a traditional provider such as a hotel or car-hire firm. However, there are important environmental benefits too and sharing economy companies often highlight them.

Renting a car when you need it rather than owning one, for example, means fewer cars are required and fewer resources must be devoted to making them. Car-sharing companies are successful to the extent they can understand the patterns in vehicle usage and availability that big data provides, allowing them to cost-effectively manage their stock, schedule maintenance, and deal with disruptions while providing high quality service that their customers demand and

expect. Avoiding massive building of mostly empty apartments in beaches or natural sites through shared utilization of people with similar preferences but different "downtime" calendars also generates clear environmental benefits that were more difficult to achieve before a good amount of related data was available through the internet.

Finally, we also want to highlight how big data is creating profitable business opportunities for renewable (or less polluting) energy. The possibility of putting a solar panel on your own roof and selling power back to the collective grid has generated a growing interest in solar energy in the world (together with much cheaper solar cells from China). Wind turbines too are becoming more and more popular. This increased popularity will yield its best sustainability returns however only when energy supply and demand across regions, countries, and even continents is better matched. Electric vehicle networks, the charging stations that form the backbone of unlimited mobility for electric car owners, currently have a relatively small number of users and modest infrastructure. Yet as the smart grid gradually comes to life, enabling decentralized energy provision and two way energy transmission (grid-to-user and user-to-grid) their position at the interface between utilities and end users will allow them to capture the economic benefits generated by price fluctuations available to them as wholesale purchasers of power from utilities. The real payoff will arrive when the network operators are able manage all the necessary data to store energy in the batteries of stationary cars and become arbitrageurs in electricity markets – assuming of course that they get the underlying analytics right.

Research opportunities on big data, management, and sustainability

In addition to these first order research outcomes that can accrue as we behold the uptake of big data applications, as academics we may be able to shed light on some important, but perhaps less apparent, second order effects. One can point to some likely similarities with the

dramatic shifts that followed the emergence of the Internet itself. At its core, the Internet is simply an easily accessible platform that enables quick and cheap information transfer. But over the span of two decades it has essentially destroyed industries (travel agents; music distribution; encyclopedias), altered accepted communication practices (from face-to-face and telephone to email and texting), transformed human behavior (online shopping; networking; addiction) and accelerated mega-trends (globalization; the gig economy). In this context, the urgency of sustainability problems and the dramatic emergence of big data opportunities should promote growing interest in joint research opportunities. We provide a description of some possibilities below.

Big data, sustainability, and corporate activities. Given the pace of current and expected developments that the big data revolution is unleashing, sustainability researchers of all stripes can easily feast upon the excitement, innovation, novelty and entrepreneurial fervor in this space. As highlighted above, many new businesses may emerge based on a value proposition that connects big data and sustainability. Researchers might want to pay specific attention to the effects of these offerings in the corporate portfolio of activities. Case studies can celebrate the fascinating insights that big data practitioners generate, and which can yield major sustainability benefits, often in unexpected places with surprisingly little effort. Scholars in the field of entrepreneurship have here an exciting assemblage of opportunities for analyzing new and innovative developments. More broadly, researchers in the field of organizations and natural environment will enjoy the opportunity to analyze how certain sustainable business activities may gain or reinforce their competitiveness when they are combined with big data opportunities (e.g. Cohen & Kietzman, 2014).

Big data, sustainability, and theoretical paradigms. Interestingly, the main theoretical paradigms to study organizations and sustainability (i.e. institutional theory and the resource-based view) have so far not provided good illustrations about how big data might be integrated in their conceptual toolkit. This state of affairs will probably change, for example when researchers will examine the industry convergence likely to occur as a result of big data. In 2011, the website Greenbiz launched an initiative named VERGE, based on the premise that four currently distinct industries - energy, information, buildings and vehicle – will increasingly be tethered together, in no small part because of technological innovation made possible by big data. Think smart cities in which smart cars and autonomous taxis roam the streets, all consuming minimal resources at maximal efficiency. Which industries will flourish and which will wither if indeed this transformation unfolds? Which companies will win and which will lose, and what will be the core competencies for succeeding in this space? Seminal work in institutional theory, particularly around the evolution of fields (Hoffman, 2001), can inform researchers tracking these industrial changes, and perhaps help predict how shakeouts will unfold.

Alternatively, big data may necessitate different theoretical lenses to push the sustainability research agenda forward. One example is the marriage of big data with social psychology, behavioral economics, and knowledge management, as illustrated above in the context of energy efficiency. Other opportunities are likely to be present.

Big data and better integration of environmental models and management research. In addition to industry convergence, big data is also likely to render interdisciplinary boundaries increasingly fuzzy as scientists and engineers, entrepreneurs and incumbents, recognize the business opportunities that climatology, remote sensing, hydrology and other scientific domains are uncovering. One immediate implication is that management of water and ecosystems

becomes more closely linked to "our" flavor of management. What do you call the business domain in which climate modelers harness big data to predict the frequency and severity of hurricanes, which are then transformed into financial instruments called catastrophe bonds that enable insurers to sell risk to institutional investors? Research that integrates entrepreneurial theory with sustainability is likely to benefit greatly from examining contexts such as these. At the same time, could we expect the integration of better environmental forecasts with technical and market models in the firm? How and what competitive implications might a more sophisticated integration of business and environmental models generate?

Big data and sustainability competitive advantage. A key question for researchers is how sustainability opportunities are identified, envisaged and framed. Considering that big data may bring new, heterogeneous, and multidimensional perspectives on corporate environmental impacts, will the focus remain on efficiency improvements – the tried and true framings that equate lower operational costs sustainability with outcomes? Will pursuit of sustainability be increasingly linked to risk reduction and management? More intriguingly, will sustainability be tied to new, hitherto unfamiliar, aspects of business value? In particular, will Big Data expand the scope of current business propositions by "unlocking value" from the atmospheric, terrestrial and oceanic commons, and the ecosystem services they provide? As big data makes more things analyzable, and consequently more commensurable and tightly linked to value, the language, analogies, and rhetoric used by both proponents and opponents will be amenable to analysis by researchers.

Big data and implementing sustainability in organizational departments. The industrial and disciplinary metamorphoses that big data is likely to trigger will also cause dramatic upheavals for people engaged with sustainability at work. One possible outcome - and an

ambivalent fantasy of many sustainability executives - is that their roles will become redundant. Smart cities, cars and grids have sustainability "baked in", one can argue, and therefore sustainability won't need to be championed. But perhaps this prediction is overly optimistic. It might, however, be true that sustainability managers will become much more active in developing business strategy, participating in product development, and interfacing with disciplinary specialists. Successful sustainability personnel might require a different skillset, with greater numerical and data visualization competencies, and overall greater research capabilities. More and more people in the workforce might strive to position themselves as sustainability experts, even if they do not have the environmental science credentials that have traditionally served as a disciplinary launchpad for the sustainability workforce. Such developments might herald an opportune moment to employ network methodologies, which have so far only been used rarely in sustainability research.

From a different perspective, managers in different departments may now obtain more powerful tools to implement sustainability in their different functions. It is particularly clear that big data provides opportunities for more sustainable operations and supply chain management by saving energy and resources and simultaneously satisfying demands and regulations. However, these new opportunities come with challenges. The interorganizational nature of the supply chains (or even certain market operations) and the fragmented nature of customers will also call for shared utilization of data for different firms and different departments in the same firm. How should big data be shared? Are environmentally sensitive data easier or more difficult to share? How might regulation affect the utilization of big data for sustainable progresses in firms?

Big data and measuring sustainability in organizations. A relatively safe bet is that sustainability reporting will become increasingly data driven, employing a wider array of real-

time, data-rich entryways into exploring organizational sustainability performance. Bloomberg, Thomson Reuters, and other purveyors are already showing a growing interest to offer more specific sustainability oriented metrics (e.g. see also Hahn et al., 2015 for a review on carbon disclosure). Big data will likely generate more opportunities to get more environmental and social data from firms and simultaneously gaining in precision and time opportunity (e.g. Barnaghi et al., 2015). So far, data from sensors or wearables have not been regularly incorporated in management research, but they can be particularly useful in the field of management and sustainability. For example, several mobile apps demonstrate the very real possibility of collecting real time air quality data from multiple smartphone users to generate pollution indexes which can be more robust than some official statistics. And if governments can easily access and analyze large amounts of raw data about firm environmental performance from multiple sensors, will regulators abandon "command and control" approaches, turning instead to flexible, output based rules? Clearly, different and more precise ways to measure sustainability in organizations will provide appealing opportunities for management scholars.

But when thinking about sustainability metrics, it is often helpful to reconsider the catchiest of dictums about analytics - that what gets measured gets managed. In a world of big data more will get measured and it is logical to presume that more will consequently get managed. More of what, though? More of what is measurable is a reasonable answer. More of the same, but better. The sequitur, of course, is that what doesn't get measured doesn't get managed. Big data might end up focusing our attention increasingly on a smaller subset of topics which are amenable to increasingly granular measurement. Like the data-driven exercise demons obsessing over their activity trackers, nutrition trackers and other life hacks to the (possible) detriment of other pathways towards wellbeing, researchers and managers pursuing sustainability

run the risk of being fixated on a set of data-rich internal indicators while remaining oblivious to broader, systemic deterioration in their environment. Big data might make organizational sustainability more self-absorbed, and might make the sustainability solutions organizations devise increasingly technicist.

Data, big data, and papers in this issue

While this regular issue of *Organization & Environment* was not intended to illustrate the orientation of this collaborative editorial, it is particularly interesting to see that some ideas and features of our regular contributors are congruent with our discussion of big data above.

On one hand, two papers are particularly related with theoretical clarification and discussion of relevant concepts and frameworks in the field of sustainability management. It illustrates well the importance of theoretical perspectives for gaining a more fruitful perspective in our field. We do not feel that big data will change this need in the future. In fact, we argued earlier for more robust theoretical perspectives integrating the new realities emerging from big data. Dyllick and Muff recognize a "big disconnect" between a growing interest in sustainability in major companies and the real state of the planet. They contend that claimed business sustainability is not always real and they develop a typology of business sustainability with a focus on effective contributions for sustainable development. Waistell expands this perspective by assessing the extent to which environmental aesthetics can promote corporate sustainability. He concludes that environmental aesthetics can promote corporate sustainability. He concludes that environmental aesthetics. Both papers provide recommendations for a more multidimensional understanding and metrics of business sustainability. We hope that big data might help this enriched perspective in the future.

On the other hand, three papers use different data to answer relevant research questions. Interestingly, none of them are using big data to provide a more panorama, their authors have made an excellent effort to provide a robust empirical analysis and we believe that these works illustrate how using big data in sustainability research will not be necessarily easy even when good reasons justify the utilization of enriched data. Firstly, Perrault and Clark examine how environmentally concerned shareholder activists vary in their status and reputation, and how these differences affect firm responsiveness to their concerns. They use a sample of 420 resolutions concerning the natural environment in the period 2004 to 2008 to show that firms respond positively to shareholder activists' high status (a desirable characteristic) and also to their reputation to threaten the firm (an unfavorable characteristic). Environmental activism is clearly one of the activities that can benefit from big data availability and we need to know more about if/how environmental activists get environmental data.

Secondly, Prime and Cater empirically examine the relationships between environmental proactivity, life cycle stages, competitive advantage, and industry on a sample of 155 Australian firms. Their results show that the construct of the organizational life cycle is significantly related to environmental proactivity. Surprisingly, they find environmental proactivity was positively related to competitive advantage not only in the innovative stages but also in the conservative ones. These findings reveal how combining organizational information and external information (i.e. cycle stages in the industry) can help to better understand the relationship between environmental proactivity and competitive advantage, but the difficulties of generating proper metrics and data are also highlighted.

Finally, Meyer, Cross, and Birne use qualitative data in one organization to examine the processes that occur to foster support for a new green initiative among people with diverse

values and perceptions. They find that framing was particularly important, but framing contests occurred due to variation in how individuals cognitively connected different frames together. The results are particularly compelling because they successfully show how frame alignment between managers and stakeholders can reinforce corporate green actions. In this context, we wonder if it could be possible in the future to use wearables collecting information about feelings and reactions of people in green issues and sustainable initiatives.

Final conclusions

It is tempting to think of big data as an empirical cornucopia for researchers in the field of management and environment. With so much data available, measuring so many different things, anything can be measured and correlated with anything else. Like others, however, we caution against blindly giving in to this temptation. Theorization is what researchers bring to the table, not unalloyed number crunching proficiency. And number-crunching proficiency is not enough for grappling with sustainability, because sustainability is a systemic, complex, interdisciplinary, and evaluative problem (Ferraro, Etzion, & Gehman, 2015) that requires as much good theorization as we can muster. It is at the meeting of one of the world's most intractable problems with the world's biggest datasets that researchers can truly put themselves to the test – in the theories they develop, the methods that they employ, the conclusions that they draw, and most importantly, the impact that they have.

REFERENCES

- Abbasi, A., Sarker, S., & Chiang, R. H. 2016. Big Data Research in Information Systems: Toward an Inclusive Research Agenda. *Journal of the Association for Information Systems*, 17(2): 3.
- Anderson, C. 2008. The End of Theory: The Data Deluge Makes the Scientific Method Obsolete, *Wired*.
- Asensio, O. I., & Delmas, M. 2015. *The Dynamics of Behavior Change: Evidence from Energy Conservation*. Paper presented at the Academy of Management Proceedings.
- Barnaghi, P., Bermudez-Edo, M., & Tönjes, R. (2015). Challenges for Quality of Data in Smart Cities. *Journal of Data and Information Quality*, 6 (2), 6.
- Constantiou, I. D., & Kallinikos, J. 2015. New games, new rules: big data and the changing context of strategy. *Journal of Information Technology*, 30(1): 44-57.
- Cohen, B., & Kietzmann, J. (2014). Ride on! Mobility business models for the sharing economy. *Organization & Environment*, 27(3), 279-296.
- Ferraro, F., Etzion, D., & Gehman, J. 2015. Tackling Grand Challenges Pragmatically: Robust Action Revisited. *Organization Studies*, 36(3): 363-390.
- George, G., Haas, M. R., & Pentland, A. 2014. Big data and management. *Academy of Management Journal*, 57(2): 321-326.
- Hahn, R., Reimsbach, D., & Schiemann, F. (2015). Organizations, Climate Change, and Transparency Reviewing the Literature on Carbon Disclosure. *Organization & Environment*, 28(1), 80-102.
- Hoffman, A. J. 2001. *From Heresy to Dogma: An Institutional History of Corporate Environmentalism.* Stanford, CA.: Stanford University Press.
- McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D., & Barton, D. 2012. Big Data: The Management Revolution. *Harvard Business Review*, 90(10): 61-67.
- Reinhardt, F. L. 1998. Environmental Product Differentiation: Implications for Corporate Strategy. *California Management Review*, 40(4): 43-73.
- Thaler, R. H., & Sunstein, C. R. 2008. *Nudge: Improving Decisions About Health, Wealth, and Happiness*: Yale University Press.
- Varian, H. 2014. Big Data: New Tricks for Econometrics. *Journal of Economic Perspectives*, 28(2): 3-28.