

Big Argumentation?

Daniel Faltesek

Oregon State University, USA, danielfaltesek@gmail.com,
<http://www.danielfaltesek.com/>

Abstract: Big Data has a long history. Public concern regarding the mass diffusion of data has appeared repeatedly with computing innovations, in the formation before Big Data it was most recently referred to as the information explosion. In this essay, I argue that the appeal of Big Data is not a function of computational power, but of a synergistic relationship between aesthetic order and a politics evacuated of a meaningful public deliberation. Understanding, and challenging, Big Data requires an attention to the aesthetics of data visualization and the ways in which those aesthetics would seem to depoliticize information. The conclusion proposes an alternative argumentative aesthetic as the appropriate response to the depoliticization posed by the popular imaginary of Big Data.

Keywords: Big Data, Information Explosion, Analytics, Cultural Logic of Computation, Aesthetics

1. Introduction

On March 29, 2012 President Obama called for an all hands on deck for a multi-departmental project on Big Data (Khalil 2012). With a budget in the hundreds of millions of dollars for cross-departmental collaboration to learn about the Government and to facilitate information exchange, and public-private partnerships the project would demonstrate the commitment of the Federal government not just to a set of practices but to the term. Departments ranging from the National Security Agency to the Veterans Administration are set to receive the benefits of enhanced computing capacity and information exchange. Exchange and processing capacity are concrete: they involve dump trucks, hand trucks, machines, wire, and cable. Big Data on the other hand means something very different, something more than that the Federal Government is going to make some really big computers, really fat wires, and really nifty fibre optics. Just as the Obama administration describes Big Data in the future tense, as something yet to come into existence, the National Security Agency has been using advanced computational systems for sometime to collect and process vast quantities of data. At least now those efforts have an adequate brand name. Big Data is really big, but it cannot be quantitatively measured. Big Data has enough cache to drive advertisements during prime-time television, and even during football games. According to IBM, Big Data can make a better planet¹. Big Data in the hands of your investment manager can explain unrelated world events and render big profits². Big Data can do anything you ask, and thanks to slick visualisations it looks great. Even the future of communication research seems to depend on Big Data in the form of the turn toward the quantitative or computational humanities. There is more data out there than ever before, with the total volume of information expanding like a classic power law distribution. If the promises are true, we will have the capacity to “drill down” to find valuable pressurized insights that will enrich the public like a geyser of oil.

Big Data is a metaphor used to describe not just a set of corporate promises, but a way of thinking about information and politically organizing efforts to use information. Big Data is

¹ This link leads to the main page for the IBM campaign:
http://www.ibm.com/smarterplanet/us/en/?ca=v_smarterplanet

² This is a claim made in commercials by T. Rowe Price in commercials, bibliographies for their claims are unavailable.

functionally a brand name for a conception of the relationship between society, technology, and politics. Manovich (2011) notes that the term, for all its popular press appeal, provides little in the way of clarity. For Manovich the novelty of a large dataset was never particularly appealing, the ways that large data sets might be managed for further analysis is. In a similar vein boyd and Crawford are right to caution researchers that Big Data discourses have under-theorized epistemology and research ethics. Although Big Data is a problematic term it carries a great deal explanatory power with the public, and further is the term through which arguments about large datasets are made. Big Data may be bigger today than ever before, but the idea of the bulk collection and processing of information is old, centuries old. Running alongside the discourse of data expansion for decades has been the question of the validity of insights gathered from very large datasets. At stake in the meaning and critique of Big Data is the political joint between methods for the collection of information and the process by which that information is interpreted.

This essay is concerned with the ways in which the promise of Big Data has been sustained despite of repeatedly under-delivering on the promise of computation. Big Data does this by acting as if data-driven approaches to thought exist outside the discursive rules established for arguments. The vacuousness of Big Data is an aspect of strength in this case, it provides a sense of wonderment, and an argumentative strategy. In this essay, I consider the inferential relationships that shape computational projects as prior to the act of computing, arguing that there is an aesthetic logic to the use of inferences that affords political power to a particular style of argument, or to use Galloway's appropriation of Jameson, I will outline the "allegories of control" inherent in the discourse of Big Data (2012, 99). People compute things for a reason, and the relationship between those reasons and computational projects should be subject to scrutiny. This essay uses the term aesthetics to refer to the distribution of the sensible or that ideas are never really separate from the form they are presented in (Ranciere 2004):

"Aesthetics refers to a specific regime for identifying and reflecting on the arts: a mode of articulation between ways of doing and making, their corresponding forms of visibility, and possible ways of thinking about their relationships (which presupposes a certain idea of thought's effectivity)" (Ranciere 2004, 10).

The meaning of a design, institution, or dataset cannot be separated from the way in which it was represented. This is a profoundly political conception of aesthetics, which seems only appropriate for an investigation of the symbiotic relationship between data collection, analysis, visualization, and political power.

First, the claim to uniqueness or difference in the project of Big Data will be juxtaposed with historical resources on the collection and computation of large datasets. The purpose of this section is not to be a history of Big Data, but to argue for the continuity of the promises of Big Data with the promises of the Information Explosion of the mid-twentieth century. These claims are important because they shed light on the arguments that people would like to make with data that they might like to treat as coming from the data itself. Second, I will engage a number of different aesthetic strategies used to manage the introduction and appreciation of large datasets through a reading of the analytics or inferential paradigms employed in several data visualization projects. In this section of the essay the relationship between aesthetic representations of data and seemingly objective argument evolves, starting with the use of data as an anchor to justify visualization and ending with the visualization being an end in itself for old or inadequate data. Finally, the conclusion of this essay suggests that an aesthetic of argument may provide conceptual resources for addressing the political power of Big Data.

2. The Historical Big Data

From the standpoint of information collection, Big Data has a long history. Driscoll has effectively traced the history of the idea that we call Big Data to the early part to the late nineteenth century (Driscoll 2012). From the time of Herman Hollerith on the sure volume of information available for computation has been increasing, with punch cards allowing basic computing functions for a fifteen million-person census in 1900, and one billion posts on social media every two days (Twitter 2011)³. Early twentieth century descriptions called the raw amount of data being produced by mechanical computing systems an information explosion. Herman Hollerith and the punch card system marked the beginning of the explosion. The key element for the explosion to begin was a paper card that could withstand the forces of moving through a mechanical computer. Punch card computing allowed a much more efficient census and the refinement of thinking at the level of the population (Heide 2009).

With the end of the Second World War, the question of what to do with incredible masses of data became quite real. Science magazine regularly dealt with matters related to the information explosion, and how we might solve “crises crisis,” a phrase describing the difficulty in connecting advances in mathematical thinking in the early twentieth century with mid-century computing power (Green 1964). Explosion discourse was so prevalent in fact that Simon (1968) began an opinion article in the journal of Management Science with the caution that the term had already entered overuse. What Simon was arguing was that was an increase in the amount of information available was secondary to having something meaningful to do with that information. Bar-Zakay (1970) described it aptly for the Rand Corporation, “the ‘information explosion’ should be termed, in my opinion, ‘infantile digestion’” (7). Bar-Zakay’s argument was made in the context of developing an infrastructure for international technology transfer, it was not enough to simply provide a great deal of data from the information explosion to a country, but that countries would need to develop a coherent approach to digesting and assimilating that data into something useful. This is much akin to what information theory would see as the distinction between signal and noise, when enough information is made to flow at the same time with little analysis, it is as bad, if not worse than no information at all.

In *1985: Corporate Planning Today for Tomorrows World Market* (1967), the writers of the now defunct Business International research company were convinced that success in the future world of business would require sophisticated data management. What is truly curious about the research in 1985, is the idea of the management crisis – that there simply are not enough good managers as the “scarcest resource of all...” (Business International Corporation 1967, 116). The promise of Big Data is not that it might empower individuals to make better corporate decisions but that it could empower visionary CEOs to orient the organization to larger contexts. The most pressing insight, and what the Business International understands is that the use of data is a political decision made with in an organization. Several decades in it seems that the information explosion may be the slowest moving explosion in history.

The historical trajectory of data production is often taken for granted. In an interview with Cukier, a major data scientist, The Economist noted that the amount of data being routed through information systems is increasing exponentially, and that this is no panacea. The most pressing problem presented by Cukier is that the publication of the inferences derived from large datasets (and the inferences themselves) can lead to negative effects, such as the use of statistics in creating the financial crisis of 2008. If a computer is programmed to treat all mortgages, bonds, or credentials as equal it will. No matter how far equipment advances, garbage in, garbage out. Information touches every area of life – but it is unclear if it ever can become more than a trailing indicator. This stands in opposition to the vision to many popular accounts of Big Data where the sure presence of data will combine with seemingly natural inferences made by business people to produce a different and apparently better and more profitable world (Beath et. al., 2012). The rhetorical strategy in works from Harvard Business

³ Due to a number of factors the total number of social posts per day is not available, as of Spring 2011, Twitter had over two-million items per day. <http://blog.twitter.com/2011/06/200-million-tweets-per-day.html>

Review or Sloan Management Review should be understood in that context – the inferential insights that the potential audience member should have are trained in during the process of mastering business administration.

Business side analyses of Big Data and visualization exist in a particular argumentative world, where the relevant test for the appropriateness of the inference is be profit for the corporation and in the event of a dividend, value for the shareholders (Boltanski and Thévenot, 2006)⁴. Just as Big Data might have offered solutions to the crisis of crises in the late 1960s, data offers seemingly endless solutions to contemporary problems. IBM's motto, "Solutions for a Smarter Planet" is a continuation of this discourse– refiguring power grids, enhancing aggregate demand, and many other tasks that their systems purport to accomplish are profoundly public, political, and solvable by a corporation with a powerful enough computer⁵.

3. Inadequate Analytics

An enduring issue for public sphere theory is the role of the image (Finnegan and Kang, 2004; Peters 1993). Images in the worst form appear as mere spectacle intended to distract the masses from more pressing concerns. In their most positive form they provide critical evidence from a first person perspective in the form of witnessing distant events. A basic distrust of surface appearances inspire hermeneutic, psychoanalytic, and critical approaches– yet as Gladwell adroitly argued in the context of breast exams and the Iraq war, publics rely on vision when other sense experiences could provide them with better information (Gladwell 2004). It is this flux between the appearance and reality that requires Rancière's use of the term aesthetics. Big Data appears not as a mere visual representation of information, but as the output mechanism of a sophisticated computational process. It is not that Big Data means that there is a pretty picture, but that there is an inference which can find order in chaos, resolve excessive noise in a signal, or find a logical, rational way to resolve what would otherwise be a messy aesthetic, moral, and political issue. Data is ingested, stored, analysed, and visualized. Visualization and ingestion are already understood to be rhetorical and political problems – the inferences made by Big Data projects seem to get something of a pass.

Visuality and the visual output of inference processes is curious in that sitting between the two worlds of visual aesthetic representation and advanced mathematics offers protection from what would seem to be direct political questions. This may be true for the simple reason that the corporate bend of most Big Data projects has eschewed thorny moral and aesthetic questions to this point. Big Data is not telling people how to live their lives or what they should enjoy and thus it is deactivated as a matter of political contestation. If this is the case, if the Big Data remains concealed within mixed technical-public-private realms of business it will either wither away under its own triviality, or it will become a vehicle where by important political questions are smuggled out of the public realm and into corporate governance.

The response by information activists in this regard has come in the form of tactical media. Raley argues that the act of making power relations visible, through projects like "They Rule" is a form of political action (Raley 2009)⁶. "They Rule" is an interesting project that allows users to form a menu of select individuals and see their relationships to a variety of corporation, think-tank, and non-profit boards (On 2011). The inference that the user derives from the site is that a linkage between boards implies a unity of purpose. The user can see a display of what they already know, the capitalists are in league with each other. In this sense, the network imaginary offered by the interface makes the fantasy of an agency for society real, an instantiation of what Lacanians refer to as the Big Other (Žižek 1997). It is in this

⁴ This approach contends that the rules for argumentative settings are reflexively determined, the relevant test is thus different in worlds of commerce, government, fame, and family. Big Data has been supposed to find meaningful answers in each of these worlds (Boltanski and Thévenot 2006).

⁵ See note 1.

⁶ This is the address for the visualization itself – Josh On, "The Rule," (2011). <http://www.theyrule.net/>. The analysis being discussed: Rita Raley, *Tactical Media*, University of Minnesota Press, 2009. Because the website has undergone extensive versioning, my commentary refers to the 2011 version, which seems to be similar to the 2009 version of which Raley writes.

sense that a network of thousands of individuals who may never speak to each other, rule. The artist is not alone in crafting this insight – the ideas both of commonality in ownership structure and ownership to institutional action are common in media studies and are critical to the work of political economists like Herman and McChesney (2001). Allowing users to discover the relationship has a very different affective potential than simply telling them about it. The key insight for tactical media and visualization scholars is that their artistic representations are always political. Raley's conclusion – activists need to structure their forms of data around fundamentally sound argumentative inferences. Their tactical work will be read through politics because it is marked as a form of politics.

Investment in representation comes full circle in the form of political representation games. ProPublica's "Free The Files" game is fascinating in that it deploys the metaphor of making publically available files visible for the purposes of securing free labour (Shaw et. al. 2012)⁷. Click the link of a muscular cartoon tearing his shirt asunder to begin a hulking clerical adventure. The text prompt for the game asks users to help find "dark money" which has been "locked" in television station file cabinets around the country. The underlying operation of the game depends on accessing a highly user-friendly Federal Communications Commission database that serves all of these supposedly locked away documents. Users are not breaking into a television station or really making anything more free than it already is. In playing they are logging information about publically accessible PDF files ingested from the web servers of the FCC into a database. This does add a valuable dimension to the data – it enhances the capacity to search and analyse data on a national scale, and that is a useful function in itself. Critical to the operation of this project is that users accept the idea that money corrupts politics and that journalists exposing influence peddling will change the public's perception of political influence.⁸ The interactivity driven by a particular aesthetic logic -- the player is unlocking information that will be used to enhance a set of inferences to which they already agree. What these cases indicate is that interfaces and Big Data tasks are separate – and that the interface provides a logic that runs along side a more complex, more political process of argument. The data secured and visualized by these projects is not intended to make a new argument, but to provide a new aesthetic appeal to an established approach.

4. Popular Analytics

Many basic visualization products do little more than count and represent information. While this may be useful, it is unclear why this matters aside from making information more comprehensible. If it were simply a matter of making counts clearer, there are many issues of public concern that would have been resolved years ago. Conduit metaphor fantasies have danced in the minds of communication theorists for millennia. Consider Yau's beautiful images of the nation's air transportation system (Yau 2010)⁹. The visualizations are good, if not great. The use of colour and line against the negative space of the undifferentiated territory of the upper atmosphere is gorgeous. At the same time, they should definitely not be used to make air traffic control policy. The inference to be made here is that airlines fly from particular cities and that some patterns look like some others. There is neither an additional layer of context or analysis, nor does there need to be. The solutions one might find for air traffic problems will be beautiful, but they will not necessarily be accurate.

Common inferences involve correlation and generalization. Other sorts of inferences are less common. It is unclear how a sign argument might be represented by in visualization, or how the difference between correlation and causation can be effectively parsed. For example it would be easy to infer from a geographic visualization that there is some determinate or at least influencing factor presented by location, even if that is mere coincidence. Similarly a representation of change over a fixed time would seem to imply a process evolving over time

⁷ <https://projects.propublica.org/free-the-files/>.

⁸ This is a central assumption in most independent journalism arguments.

⁹ <http://flowingdata.com/2010/04/15/explorations-of-real-world-traffic/>.

where time might be a factor in change, rather than a structural fact. What is striking about analytics and inferences is the lack of depth provided by depth producing technologies.

Oracle and IBM go out of their way in promotional materials to point out that a vast database is not useful without some analytics to parse through it – it doesn't matter how many insights you have, if they are all oversimplifications. To use IBM's own language from their online advertisement: "We pair our rich portfolio of capabilities with industry-leading services delivered by 9000 dedicated analytics consultants who can help you accelerate time to value and deliver breakaway results."¹⁰ At this point the major insight of Big Data is not that there is anything new to be thought, but that we might finally have a technology that can allow any number of old ways of thinking to have the imprimatur of new technology. Without an increasingly robust set of argumentative inferences, Big Data is just more wine in the same jug.

Returning to Yau, his most popular inferences involve space, time, proportion, correlation, distribution, and comparison. While this might appear to be an exhaustive list at first, there really are only a few basic inferential relationships at stake none of which seem to have any natural connection to normative recommendations. Much like how contemporary image editing software relies on the structural metaphor of layering, Big Data relies on a foundational layer of inferential authority to make itself viable. It seems most likely that this functions as Golumbia (2008) describes the cultural logic of computation, where the act of being computed provides a measure of authority. Spreadsheets have become a form of reactionary office politics. In more concrete terms, the reliance on space, time, proportion, correlation, and distribution provides an important form of political authority. These inferences are not designed to speak to normative questions. Normative insights come from some other place, which for the business data user is the businessperson. For the employee of Facebook, their mathematical intuition, and for the government from some place other than politics.

5. Beautiful Analytics

In contrast to what data science might suggest there is an important recognition of the role of aesthetics in mathematical thinking. Aesthetic preferences play a vital role in shaping mathematical processes and thinking (Sinclair 2004). Eisenberg and Dreyfus (1986) provided a specific vocabulary for the aesthetic qualities of mathematical thinking with a few dominant criteria: conciseness, simplicity, clarity; with lesser aesthetics of: structure, power, cleverness, and surprise. The first three values are clearly inherent in Big Data projects. Analytical systems pulling through a sea of data would be designed to produce basic results that are concise, simple, and clear. Simplicity is not simple, though and conciseness all too often relies on a strong appreciation for context. A concise summary of available data about aircraft departures and arrivals tells you very little, aside from that some planes were moving around. Structure and power provide additional aesthetic problems, as they would retroactively ground themselves – the stronger a distinction that an analytic would claim to make the greater its aesthetic appeal. Structure in itself is an intervention into the relationship between signal and noise in the data set. Deciding what is meaningful information to measure is an important decision. Even in the act of naming variables to be measured there are power dynamics which The tendency for data scientists to cross from the social sciences into the humanities becomes far clearer: basic insights about the geographic distributions of populations might be beautiful, a chart with an insight about the good life is powerful.

It is useful to be disabused of the notion that aesthetics are limited to pop-visualization from the outset. Normative business decisions have aesthetic grounds just as much as normative governmental decisions. Simon for example was calling for a new era of rationality that could cut through the noise and lead to a better form of management. Data has become such an active aesthetic term that it sought after to the point of being "sexy" as two authors declared in a recent issue of the Harvard Business Review, data scientists today are akin to Wall Street "quants" of the 1980s and 1990s (Davenport and Patil 2012). The role of quants

¹⁰ IBM. Smarter Analytics. Original date unknown, accessed January 13, 2013. <http://www.ibm.com/analytics/us/en/>

in high frequency trading, subprime mortgages, and the financial crisis is conveniently forgotten. There is no danger if the proof is attractive enough. This provides an explanation a critical bound on rationality: if the aesthetic vision of a company is predicated on maintain a beautiful warehouse they will deploy their analytics to decrease inventory, if the aesthetic vision involves less mess, they will purchase pre-cut meats, even if this decreases margin (Anderson 2008). Money does not eclipse humanity. The justification for Big Data as provided by IBM is that business leaders simply do not trust the information that they are being provided. Given that the sensors and other data systems that the leaders are using to collect their information is likely unchanged through the purchase of a database product, it is unclear if their lack of trust is really about the data or a lack of faith in their co-workers.

Passion is political, and the seemingly passionate use of mathematics has become a central strategy for some politicians like Paul Ryan or David Cameron, austerity is foremost an aesthetic (Krugman 2013)¹¹. Mathematical reasoning, or at least the trappings of that reasoning, can be performed as an identity. The proofs provided by politicians are often concise, clear, and simple, but they lack in analytic power or complete structure. Making the case for austerity, for example, depends on using graphics that have very little bearing on economic considerations, but a certain moralistic inflection. In the governmental context this becomes even more pressing. Agamben's use of Benjamin is pressing here in that that the creation of the authority for government and the creation of the temporality in which governance occurs depend on the aestheticized performance of emergency (Agamben 2004). On one side is the promise of Big Data that it might make abundance possible, on the other the claim that it could resolve crises. When not conducted in the aesthetic frame of an emergency, this is often known as cost-benefit analysis.

Expert policy analysis on the basis of cost-benefit analysis has become increasingly political. Ackerman and Heinzerling's (2005) critique of the cost-benefit tradition can be understood aesthetically: we know the price of everything because that data is discrete, stable, and powerful. Instead of seeing two values, which are commensurable, the costs are understood in very real terms, while benefits appear ethereal. In the process of calculation, representations are reality. Scholars have argued that often costs are computed in strong physical terms while benefits of government programs are nebulous and are calculated in far less generous terms. Cost-benefit analysis is a form of aesthetic activity. The value of the environment or democracy is a much more nebulous concept and thus often in cost-benefit analysis appears to be of limited importance. Concrete external costs with require few adjectives will win the day, even if that calculation is specious at best. An aesthetically charged impact like a nuclear detonation, even for being highly unlikely, sets the parameters an argumentative world. The aesthetic practices by which a world is configured determine the outcome of analysis.

In the 1960s, the study of the information explosion the crises crisis was an attempt to argue that information science might provide a way to deal with real high risk events for human kind through computational power. What Big Data does so effectively is to replace the empirical world with an abstract world of analysis. In a circular way, the costs of Big Data are self-concealing. This view of Big Data as coming from nowhere has been particularly important in popular theorizing that would suppose that the technological infrastructure for data processing has no environmental impact. Tomlinson and Silberman (2012) argued in their critique of Shirky's cognitive surplus, that the production of the computing capacity that makes mass data collection and processing possible without substantial negative impacts the environment and people. Omitting the full cost of data production makes it possible to envision a beautiful world of free knowledge production, unfettered by the reality that the negative externalities of technology are politically allocated to those on the margins.

Schlag's (2002) theory of legal aesthetics offers a way of reading for inferences in legal texts. Most distinctions that are made in legal briefs are those which make a direct tie be-

¹¹ The important idea here is that the impression of mathematical sophistication is the key. This is nothing new, public relations are normal means, what is interesting is that performing mathematical identity has become a public relations strategy (Krugman 2012).

tween a core insight about the law and a structured insight about what is inside and what is outside of that core boundary. Progressive scholarship at various points has taken on an energy aesthetic, using the idea that the distinctions are already in motion to justify the creation of new protections or the extension of larger inferential leaps between positions that those in the core framework. His other categories, identity and dissociation, are even more person-centred. The purpose of Schlag's aesthetics is to demonstrate that the story of an ongoing synthesis of law on the basis of rationality (economic or otherwise) is a fiction designed to cover for a complex system of decisions about the production of knowledge. Pretty law is clear law, at least clear in the sense that the distinctions abide by the judicial poetics. Jurisprudential authority depends on a series of metaphorical relationships grounded either through the use of bright-lines. In what would be a frightful turn for the strict Habermasians, legitimization is an aesthetic practice.

But what does any of this have to do with data? The inferential relationships involved in the use of Big Data are political. IBM and Oracle caution their clients that the quality of their analytics, not the ingestion, management, or visualization systems employed drives their results.¹² As a spokesperson for radical inductivism and cyber-utopianism, Anderson (2008) wrote the slogan of a post-argumentation society, "With enough data, the numbers speak for themselves." This is false not only because the data really can't speak for itself, but that the underlying relationships he relied on to make his statement – the supposed effectiveness of Internet advertising and the identification of new genomic sequences really are not that useful. Data becomes a form of temporal bracketing – from the moment of collection forward we can consider the past to be a science project, absolving ourselves of the long running political struggle over the creation of worlds. Big Data at the United States Supreme Court would not preclude the writing of a history paper as an opinion and it would not provide a meaningful application of terms on a polyvocal, polysemic, contested discursive terrain. The aesthetic politics of Big Data that are ideally suited for visualization are not suited for integration into the aesthetic domains of politics, law, international relations, ethics, or almost anything that should be understood through democratic contestation at least. That is the key point for disagreement – it is not that we are distracted by images, but that we use images in a complex performance designed to evacuate antagonism. One would pronounce an argument as coming from "Big Data" to absolve their position from the dirty, thorny world of the political. They display it with a complex animation to prove the difference within their distinction. This is the digital variation of Ranciere's miscount – if the inevitable failure of the mathematics of democratic process that create the conditions by which the political is realized, the best strategy to maintain order would be to aestheticize the miscount to the point that the democratic moment can not be realized (Ranciere 2004).

6. The Information Implosion?

In this essay, I compared the historical concept of Big Data with the information explosion, and found that a similar set of concerns have been resonant for decades: there is more data than ever before, but the inferences made are quite similar to those that have come before. Causation, correlation, and t-tests stay the same. Counting and mapping still dominate our representation and thinking systems. The fundamental mistake is that similar sets of aesthetically attractive inferences are assumed to do anything different with a larger dataset. Aesthetically, the inferences that are popular in Big Data projects seem to be the same inferences that have produced crises, and precluded political participation in any number of contexts. Politicizing the inference, or at least marking it overtly, offers a site for intervention into the depoliticized façade of data. Even if it might be interesting to think of challenges for data analysis as new epistemological and ethical matters, they are very similar to those encountered in argumentative visualization in the past. Much to the chagrin of pop information theorists like Anderson, numbers will not speak for themselves, there will always be a need to represent, organize, and persuade. In the age of Big Data, criticism needs to consider why

¹² The sales pitch presented is for business-to-business services and consulting, not for an autopoietic process, this is described at length earlier in this article.

certain kinds of arguments are persuasive in their contexts, that art known as rhetoric. What has changed, and what critical scholars must adapt to, is the function of mathematical aesthetics in public culture. This means that instead of diagnosing arguments as possessing fallacies or being out of their appropriate world, scholars should directly indicate that an information artist has found noise – not signal, and that a complex inference is just as important as sample size. Perhaps the most important implication of Big Data is to make the boundary between quantitative and qualitative research more permeable, which for qualitative scholars means more data and the chance to use the conceptual resources of the computational aesthetic. Permeability also reveals that power relations are strategically reversible, computation can be deployed for alternative ends, and those ends can be subject to criticism and democratic deliberation.

Visualization is a method for analysis in as much as it can help arguers see relationships and inferences they might have otherwise missed, but as a complete methodology, visualization depends on the conventional argumentative inference. Visualizations exist within a legitimation context with clear aesthetics, and surely exist to serve some argument. Data scientists then are also information artists, and in any event, they are not neutrally representing the world – they are making a meaningful contribution to an argument in a rich context.

Rather than treating the domain of data visualization and Big Data projects as a new domain that requires a new approach to criticism, this essay suggests that the vocabularies of the areas that Big Data would suppose to influence should be brought to bear on those projects. Big Data does not go far enough in embracing what data driven argumentation can be – a force for modelling the richness and complexity of human relationships. Inherent in the view of Big Data as argumentation is the rhetorical reality that the political is created through human agreements. Unfortunately, the other side of this approach is all too clear. Mathematical aesthetics are too powerful, too beautiful, and too seductive to fade into the background. The challenge then is not to ask questions at the margins of data-intensive projects, but to engage in rigorous argument about the political dimensions of the projects, from their inception, to their attractive public interfaces. In this sense, the representation of arguments as Big Data visualizations is a powerful argumentative resource. Instead of increased information opening up access, the combination of visualization as argument and the promise of data analysis technologies all too often compact context without the dimensions of argument. The stakes here are quite high. If Big Data successfully sublimates the political dimension of argument with visualizations of databases, almost every argumentative world could be affected. The discourse of Big Data requires a response with Big Argumentation.

References

- Ackerman, Frank and Lisa Heinzerling. 2005. *Priceless: On Knowing the Price of Everything and the Value of Nothing*. New York: New Press.
- Agamben, Giorgio. 2008. *State of Exception*. University of Chicago Press.
- Anderson, Chris. 2008. The End of Theory. *Wired*.
<http://www.simulex.it/Materiali/Articoli/Chris%20Anderson-The%20end%20of%20theory.doc>.
- Bar-Zakay, Samuel N. 1971. Policymaking and Technology Transfer: The Need for National Thinking Laboratories. *Policy Sciences* 2 (3): 213–227.
- Beath, Cynthia, Irma Becerra-Fernandez, Jeanne Ross, and James Short. 2012. Finding Value in the Information Explosion. *MIT Sloan Management Review*. Accessed June 19, 2013.
<http://sloanreview.mit.edu/the-magazine/2012-summer/53409/finding-value-in-the-information-explosion/>.
- Benjamin, Walter. 2008. The Work of Art in the Age of Its Technological Reproducibility. In *The Work of Art in the Age of Its Technological Reproducibility and Other Writings on Media*. Cambridge: Harvard University Press.
- Boltanski, Luc and Laurent Thévenot. 2006. *On Justification: Economies of Worth*. Princeton University Press.
- boyd, danah, and Kate Crawford. 2011. *Six Provocations for Big Data*. Accessed August 2, 2013.
http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1926431&download=yes.

- Business International Corporation. 1967. *1985: Corporate Planning for Tomorrows World Market*. New York.
- Davenport, Thomas, and D.J. Patil. 2013. Data Scientist: The Sexiest Job of the 21st Century - Harvard Business Review. *Harvard Business Review*. Accessed January 20 2013. <http://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century/ar/1>
- Dreyfus, Tommy and Theodore Eisenberg. 1986. On the Aesthetics of Mathematical Thought. *For the Learning of Mathematics* 6 (1): 2-10.
- Driscoll, Kevin. 2012. From Punched Cards to 'Big Data': A Social History of Database Populism. *Communication+* 1 (1): 4.
- Finnegan, Cara A. and Jiyeon Kang. 2004. 'Sighting' the Public: Iconoclasm and Public Sphere Theory. *Quarterly Journal of Speech* 90 (4): 377-402.
- Galloway, Alexander. 2012. *The Interface Effect*. Malden: Polity.
- Gladwell, Malcolm. 2004. *The Picture Problem*. *The New Yorker*, 74-81.
- Golumbia, David. 2008. *The Cultural Logic of Computation*. Cambridge: Harvard University Press.
- Green, John. C. 1964. The Information Explosion-Real or Imaginary? *Science* 144 (3619): 646.
- Heide, Lars. 2009. *Punched-card Systems and the Early Information Explosion, 1880-1945*. Johns Hopkins University Press.
- Hermann, Edward and Robert McChesney. 2001. *Global Media: The New Missionaries of Global Capitalism*. London: Continuum.
- IBM. *Smarter Analytics*. Accessed January 13, 2013. <http://www.ibm.com/analytics/us/en/>
- Krugman, Paul. 2012. *A Public Service Reminder: Paul Ryan Is a Con Man*. *The Conscience of a Liberal*. Accessed November 19, 2013. <http://krugman.blogs.nytimes.com/2012/11/19/a-public-service-reminder-paul-ryan-is-a-con-man/>.
- On, Josh. 2011. *They Rule*. Accessed January 25, 2013. <http://www.theyrule.net/>.
- Manovich, Lev. 2011. *Trending: The Promises and the Challenges of Big Social Data, Debates in the Digital Humanities*. Minneapolis: University of Minnesota Press. Accessed January 24, 2013: http://www.manovich.net/DOCS/Manovich_trending_paper.pdf
- Peters, J. 1993. Distrust of Representation: Habermas on the Public Sphere. *Media, Culture & Society* 15 (4): 541-571.
- Raley, Rita. 2009. *Tactical Media*. Minneapolis: University of Minnesota Press.
- Rancière, Jacques. 2006. *The Politics of Aesthetics*. London: Continuum.
- Rancière, Jacques. 2004. *Disagreement: Politics and Philosophy*. Minnesota: University of Minnesota Press.
- Schlag, P. 2002. The Aesthetics of American Law. *Harvard Law Review* 115: 7-8.
- Shaw, Al, Jeremy Merrill, and Amanda Zamora. 2012. Free the Files: Help Pro-Publica Unlock Political Ad Spending. *ProPublica*. Accessed January 25, 2013. <https://projects.propublica.org/free-the-files/>.
- Simon, Herbert. 1968. The Future of Information Processing Technology. *Management Science* 14 (9) (May): 619-624.
- Sinclair, Nathalie. 2004. The Roles of the Aesthetic in Mathematical Inquiry. *Mathematical Thinking and Learning* 6 (3): 261-284. Twitter. 2011. *200 Million Tweets per Day*. Accessed January 25, 2013: <http://blog.twitter.com/2011/06/200-million-tweets-per-day.html>
- Yau, Nathan. 2010. Explorations of Real World Traffic. *FlowingData*, Accessed January 25, 2013. <http://flowingdata.com/2010/04/15/explorations-of-real-world-traffic/>
- Yau, Nathan. 2008 "Physical Graphs as Critique on American Culture." *FlowingData*, Accessed January 25, 2013 <http://flowingdata.com/2008/08/20/physical-graphs-as-critique-on-american-culture/>
- Yau, Nathan. 2011. *Visualize This: The FlowingData Guide to Design, Visualization, and Statistics*. New York: John Wiley & Sons.
- Žižek, Slavoj. 1997. The Big Other Doesn't Exist. *European Journal of Psychoanalysis*. Accessed January 25, 2013 <http://www.lacan.com/zizekother.htm>

About the Author

Daniel Faltesek

Assistant Professor of Social Media in the School of Art and Communication