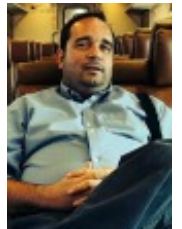


How data does political things: The processes of encoding and decoding data are never neutral.

 blogs.lse.ac.uk/impactofsocialsciences/2015/10/07/how-data-does-political-things/

10/7/2015

It is difficult to see the political structure of data, because data maintains a veneer of scientific objectivity. But data is inherently a form of politics, argues [Jeffrey Alan Johnson](#). Data does not just allocate material things of value, it allocates moral values as well. Data producers encode a state of the world at a given time, which is then decoded by data users to shape social practice. As such, a political theory of data, grounded in distributive and relational information justice, is necessary.



In what ways is data political? As data becomes increasingly the most important force in decision-making, whether in the public or private sectors, we should expect it to influence political practice. Certainly the growth of “big data” applications in campaigning and of transitions from descriptive to predictive uses of data in public policy analyses show that data influences political action, and issues related to data and information technology such as privacy and surveillance or the open data movement show that data is now a political issue. I wish to suggest here, however, that the political study of data requires more than simply considering it as a changing condition of more familiar modes of politics. Rather, data should be understood as a mode of politics itself.

Data as political practice

At its most basic, data is a political practice because it does one of the most basic political things: what Gabriel Almond formulated as “the authoritative allocation of values.” Data, the algorithms that use it, and the judgments based on it commonly allocate many valued goods. Performance data drives funding, population data determines the location of public services, and crime data directs police presence. But, like politics generally, data does not just allocate material things of value. It allocates moral values as well. Voting and demographic data were used to determine, under the U.S. Voting Rights Act of 1965, which of the states would be subject to pre-clearance of changes in voting laws—and to strike down that requirement in the 2013 [Shelby County v. Holder](#) decision of the U.S. Supreme Court—thus allocating authority between the federal and state governments.

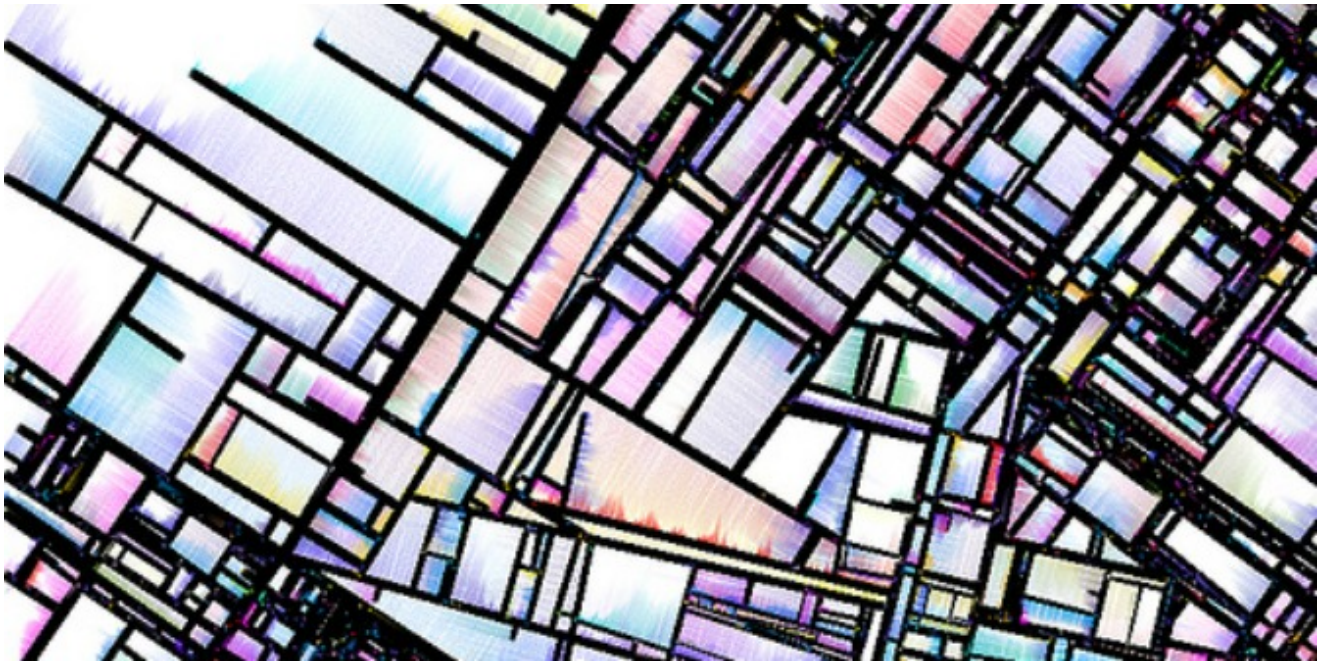


Image credit: [Substrate algorithm](#) by [Jonathan Lidbeck](#) (Flickr CC BY)

Such distributive views of politics reflect traditional understandings of political action focused on legislation as the paradigmatic political practice. But distribution is not the only function of politics. Political practices are as much about establishing and maintaining relationships among people and groups, most often relationships involving control of one by another. Data is no exception. In this meso-politics of data, the locus of political action shifts from the specific values of the data to the fact of that data's existence.

Data as political control

In their classic article "[Administrative Procedures as Instruments of Political Control](#)," Matthew McCubbins, Roger Noll, and Barry Weingast (whom political scientists have been calling "McNollgast" since long before Bennifer and Brangelina) argue that administrative procedures, for example those used in rulemaking by government agencies, act to secure legislatively preferred outcomes without requiring direct legislative intervention (for example by requiring that rulemaking take note of certain kinds of evidence likely to be provided by the interests that the legislature favors). The existence or content of data can do the same. To require specific kinds of data or to even insist on using data-driven decision-making is to ensure that the "right" decisions are made by the "right" people.

Texas Tech dean David Perlmutter noted that his institution's human resource management software would not likely allow it to post a job listing [not tied to a specific department or discipline](#). In effect, departments become privileged interests that can inhibit interdisciplinary scholarship through the design of the data management system for job searches. One could look to many other practices of political control—James Scott's notion of legibility, for instance, or Foucauldian notions of disciplinarity and governmentality.

Data as cultural product

But data is not only a form of political action; data, as a cultural product, is inherently political in itself though what we might call the meta-politics of data. In ways very similar to how Stuart Hall looked at television,¹ data can be understood as a form of communication in which a message – in this case, a state of the world at a given time – is first encoded by a producer and then decoded by an audience. Encoding data is a process of translation, in which data structures and social processes (what I have called [elsewhere](#) the "[translation regime](#)") select one data state to represent the world from among many incommensurable possibilities: the many possibilities for describing a person's gender identity are reduced to a gender binary by linking the data field to a validation table or made a

narrative by defining the field for storage of lengthy text values. This is a problem that institutional researcher Tod Massa has called “[counting to one](#),” that is, of determining what constitutes an instance of a value to be included in the database.

Those data states may be recoded several times, translating a data state into a new one (for instance by writing queries and joining tables or by incorporating data into models). Finally, they are decoded as data users extract data into a nexus of problems and interventions that will shape social practice, for example by interpreting an “Unspecified” value in a gender field as missing data rather than gender non-conformity when deciding how many unisex restrooms to build into a new facility. Much like a game of “Telephone,” the meaning structures at the end of the process are not necessarily identical to those at the beginning; indeed they may not bear much resemblance to each other at all.



Image credit: [Gender neutral toilet sign](#) (wikimedia, public domain)

These encoding and decoding processes are, of course, never neutral. Hall argues that the encoding and decoding processes are built around a “dominant social order” designed “to impose [a society’s] classifications of the social and cultural and political world” through the creation of “preferred meanings”¹; at the same time, decoding may take place from standpoints that affirm, negotiate, or oppose those meanings depending on the extent to which the audience’s own set of preferred meanings aligns with or deviates from that of the producer. This is true in data systems as well. A binary gender field that does not allow missing values is an imposition, by those with power to determine the data structure, of a certain model of gender normativity that has traditionally been the preferred meaning of gender in Western society. The user of the data may share this meaning as they decode the data, or it may become a site of contestation over competing meanings. In any of these cases, we see that data is, always already (to further abuse an already challenged term), a form of politics.

It is difficult, often, to see the political structure of data, because data maintains a veneer of scientific objectivity that protects it from challenge. But if, behind that veneer, data is a form of politics, then a political theory of data is necessary, which starts by working toward a theory of information justice. Corresponding to the modes of politics described above, such a theory considers both distributive and relational justice while promoting a kind of structural analysis that exploits the gaps between the claim of objectivity and the deeper subjectivity of data to bring about social change. Deliberative information architectures and open data are necessary, though not likely sufficient,

conditions of information justice. Such a theory might further integrate the range of ethical issues currently being explored in a more piecemeal fashion, ultimately allowing future societies to use information to address, rather than create, common challenges.

Reference

Stuart Hall, "Encoding/Decoding", in Meenakshi Gigi Durham and Douglas M. Kellner (eds.), *Media and Cultural Studies: KeyWorks* Rev. Ed. (Oxford: Blackwell, 2006). [[pdf](#)]

This is part of a wider series on the [Politics of Data](#). For more on this topic, also see Mark Carrigan's [Philosophy of Data Science](#) interview series and the Discover Society [special issue on the Politics of Data \(Science\)](#).

Note: This article gives the views of the authors, and not the position of the Impact of Social Science blog, nor of the London School of Economics. Please review our [Comments Policy](#) if you have any concerns on posting a comment below.

About the Author

Jeffrey Alan Johnson is the Assistant Director of Institutional Effectiveness and Planning at Utah Valley University, responsible for managing the UVU strategic planning process and supporting units at UVU with operational planning and assessment. He also serves at UVU resident "education futurist," keeping the campus leaders aware of trends in higher education. Jeff continues to maintain research interests in higher education policy, technology policy, and political theory, and teaches in the College of Humanities and Social Sciences at UVU. His research, including content from his current work on information technology and social justice, can be found at [SSRN](#). Follow Jeff on Twitter ([@the_other_jeff](#)) or his blog on higher education and public policy at [The Other Jeff](#). (The views expressed are strictly his own and do not reflect the views of UVU.)

- Copyright © The Author (or The Authors) - Unless otherwise stated, this work is licensed under a Creative Commons Attribution Unported 3.0 License.