Data and Politics

Power and the Formatting of Constituents' Opinions

Jianing Fan

June 2022 · Unpublished essay

Introduction

These days, data has become a vital player in the world of politics, profoundly shaping how institutions make decisions in almost every aspect. Typically, when we think of power, we associate that concept with institutions, for they establish a set of rules that require specific actions, thus making them authoritative. By contrast, the connection of numbers and statistics to authority is often elusive, yet data exhibits its resemblance to institutions in that authorities depend on them to make decisions in the first place. In this respect, data possesses a subtle form of power that operates through and beyond institutions.

In this paper, drawing on my experience with my internship in a congressional office, I will explore how data helps the congressional office form its political opinions, potentially influencing the corresponding decisions. Specifically, the focus of this paper will be on how the format of data or data structure exemplified by the issue tally influences and exercises power in the process of tracking constituents' opinions and making corresponding policies in a Congressional office. Rather than being a neutral record, the issue tally embodies the judgments of the agents who create such a format and record each opinion, displaying a sense of coerciveness without apparent coercion when its categories constrain how constituents can be represented. As a result, the congressional office, perhaps the Congressman, might not even be aware of how these prior judgments shape the information on which a decision is based. To investigate this phenomenon, I will first examine the issue tally and the judgments involved in recording opinions, then question the concept of raw or unprocessed opinions vis-a-vis data structure, and provide a historical example of how data embodies subjectivity and predisposed biases. Finally, I will survey the broader implications and limits of my observations. Doing so will provide a lens through which we view and define power, understand how it functions, and reconsider our relationship with each other, society, and government.

The Issue Tally

One of the observations I made while working in a Congressional office is how it collects the opinions and voices of its constituents. In the office of Congressman Lieu, we use an "issue tally," a drop-down menu that records the issues and their numbers of pros and cons, to track the trend of the opinions of our constituents. For instance, an issue such as "providing more military aid to Ukraine" may receive four pro comments and one con comment. Thus, each opinion is matched with an issue and recorded as either approval or disapproval.

As a result, the format of the issue tally exercises power itself in that it limits the distinctions immediately visible to the Congressman in the numerical summary. Although we write phrases and sentences that attempt to describe an opinion as accurately as possible, categorizing issues in such a format can restrain the accuracy and authenticity of each opinion when the numerical summary is treated as sufficient on its own. For instance, an issue tally with "supporting Ukraine" can indicate providing weapons to Ukraine, providing economic relief to the country, having heads of the U.S. government visit Ukraine, and so on, but it is impossible to account for every individualized message and mark them as separate issues. Note that each aforementioned scenario fits into the broader category of supporting Ukraine, yet the constituent who advocates for economic relief might not agree with military aid. When such situations occur, the issue tally can misrepresent the constituents if approval of one form of support is read as approval of another. The problem is the loss of a politically meaningful distinction, rather than simplification alone.

The written phrases and sentences are therefore an important qualification. If the office consults them alongside the tally, it may recover distinctions that the numerical summary omits. My experience with recording opinions identifies this risk, but does not establish that the tally changed a particular vote or that the office disregarded the accompanying descriptions. Nor do the messages received by one office necessarily represent the opinions of the entire district.

Furthermore, the issue tally records each opinion in a binary manner, with pros and cons. However, constituents might be ambivalent towards a variety of issues, and their opinions on such could constantly shift as they are exposed to more knowledge about that particular policy.

In the Office of Congressman Lieu, it is the interns who regularly categorize and summarize the points made by constituents on phone calls and record the processed versions of opinions on the issue tally. This mechanism involves political actors besides the constituents (who are the source of authority in the politics of representation), such as the interns, since how they phrase the voices of the constituents can render slightly distinct political epistemological units to the Congressman, thus shaping the decision-making process. Furthermore, interns may approach the issue tally based on their political experiences and what appears salient to them, identifying labels and marks such as "Ukraine", "economy", "aid", and "military aid" and assigning each of them a different weight. The label with more weight may supersede other meanings implicated in the original message. Back to our previous example of "Supporting Ukraine", one with political experiences concerning more military prowess and warfare might consider this message as "providing military aid", while one with such experiences mainly focusing on the economic side would be likely to specify this opinion as "providing economic relief". For this reason, the recording of opinions also involves subjectification, despite the seeming objective nature of data. The Ukraine example illustrates how such a judgment could arise; it does not establish how any particular constituent's message was classified.

Formatting and Fastening

The observations provided in the previous section illustrate the argument that raw data do not exist in that the presentation of all statistics and numbers requires some sort of format or mode to display and inform, be it a drop-down menu, a pie chart, or a bar graph, and doing so inevitably necessitates a deliberate process of abstraction, undermining the rawness of such data and adding a sense of

subjectivity by selecting, abbreviating, reordering, and categorizing the numerical values and opinions (expressed in phone calls, emails, letters, etc.) in their "raw" forms.

In "*Raw Data" Is an Oxymoron*, Lisa Gitelman and Virginia Jackson write that, akin to photographic images, "[d]ata too need to be understood as framed and framing, understood, that is, according to the uses to which they are and can be put. Indeed, the seemingly indispensable misperception that data are ever raw seems to be one way in which data are forever contextualized — that is, framed — according to a mythology of their own supposed decontextualization" (Gitelman and Jackson 2013, 5-6). Contextualization, or formatting, a concept developed in Colin Koopman's work, thus becomes an intrinsic feature of data as they only become meaningful and applicable when organized (interpreted) in a specific way, making interpretation an imperative step. In this regard, a mandatory subjectivity has been imposed upon every form of data, including the issue tally in the office of Congressman Lieu. Abstraction alone does not make a record false. The question is whether the distinctions omitted by a format matter for the political use to which the record is put.

Since each individual opinion in the issue tally is formatted, the subject which each issue designates and describes is inherently delineated by the format. As Koopman states, "the subject of data is pinned down to the format's delimitations and definitions at the same time that it hustles us up in how efficiently we are handled by whatever agency or network we have been formatted to fit" (Koopman 2022, 351-52). In this sense, there is a process of fastening (a form of power) when each message is recorded in the issue tally, conceived in two aspects of which the first implies excludability of other possibilities and the second demonstrates the reproducibility and intensity of such power.

Because a format only allows particular choices, the subject of data is fastened to the available options. The drop-down menu of the issue tally in which only Pro and Con are included serves as a classic example. In this case, regardless of the fluidity of voters' preferences as a spectrum, there is no other alternative within this binary field, manifesting the restrictive nature of the power of fastening. The written description can preserve alternatives, but only if it remains connected to the tally when the opinion is interpreted. In the words of Koopman, "[w]e are pinned down to the…formats" (Koopman 2022, 351-52). This phenomenon reveals that when constituents' opinions are fastened to the data (the issue tally in this case) that describe them, the richness and sophistication of such opinions can be simplified into labels that recognize only a limited part of their meaning rather than adequately representing them, which should be the objective of representative politics epitomized by the House of Representatives.

During the 1920s and 30s, fastening obliged residents of neighborhoods to conform to the ostensibly objective standards of single-race evaluation, de facto classifying households of color through a single racial category and denying them alternative methods of identification within that format, so that residents are "accounted for by their race (but not occupation or education); their race is accounted for singularly (and without mixed-race categorization); their household's race is accounted for by a real-estate equivalent of the one-drop rule (and in a way that overcodes nonwhite races)" (Koopman 2022, 353). From this perspective, the second aspect of fastening, (quick) reproducibility, or acceleration, also comes into play. As residents are represented primarily through

racial data, such data reproduce and propagate themselves to a multitude of procedures and operations in connection with those data: property appraisal and credit evaluation.

In a congressional office, once opinions collected from the issue tally are incorporated into research, briefings, memos, notes, cosponsor blocks, and other information referred to in decision-making, the same categories can travel from legislative staffers to the congressmember. This comparison concerns the propagation of classifications; it does not equate the purposes or harms of an issue tally with those of racial redlining. It is also possible to consider techniques and methodologies involved in the decision-making process of the office analogous to an algorithm, and such an algorithm becomes a deployment of subjectivity (Noble 2018) under the purported objectivity of the "raw data" from which the decisions are made. For these reasons, data can exercise power within traditional institutions through the less visible work of classification.

Hence, institutions and infrastructures that rely on such information can augment and reciprocally reinforce data's attachment to formats (the first aspect). Koopman summarizes: " [f]astening was crucial to institutionalized redlining, yet more insidious was how it conveyed an exercise of power that traveled much more widely and deeply, in all kinds of ways that are not all reducible to institutional politics" (Koopman 2022, 354). Data politics, characterized by formatting and fastening, need not exercise power through direct coercion. It conveys a sense of coerciveness without apparent coercion: the constraint lies in how a voice becomes legible within the record, even when no one threatens or deliberately silences the speaker.

Implications for Representation

Due to the characteristics of the issue tally and its potential influence on politics at various levels and scales, it is possible that politics, at least on a congressional level, can become less representative while influence over information becomes more dispersed. In this sense, politics of decentralization can develop within politics of representation. We distinguish these dimensions of politics primarily by investigating who holds a significant portion of power. Regarding *politics of representation*, it is intuitive that the constituents, or those being represented, are the source of the representative's formal authority, and this mode of politics expresses a familiar democratic ideal in the United States. The principal-agent relationship is apparent in this mode, with the constituents being the principal and the elected congress member being the agent. The groups being represented are generally assumed to share some common interests or desires that their representatives advocate; in other words, the power of the congress member derives from the people whom he/she represents. Thus, *politics of representation* constitutes our common-sense epistemology of politics in general.

However, data complicate the entire ecology of this type of politics as each individual relies on them as inputs from which they form opinions and make decisions, yet such data are potentially biased and processed by the formats in which they are organized and presented. Not only is the office vulnerable to the issue tally's supposedly faithful representation of constituents, but in a larger sense, a multitude of the "raw opinions" are based on "raw data" from other sources, extending the problem of interpretation beyond the office itself. That is to say, even when constituents are developing their understanding and stances on specific policies that are not directly relevant to them (an example would be the war in Ukraine versus a municipal decision that plans to demolish one's house for a

freeway, with the former scenario as indirectly relevant while the latter as directly relevant), they have to gather information from different channels of data: reports from governmental agencies, news, social media, etc. Here, we are expanding the discussion from structured data to sources of information more broadly; this is a theoretical extension of the office example. As previously surveyed, reports based on statistics are subject to formatting and fastening, and news, analogously, is subject to other effects such as embedded journalism and framing that can inhibit objectivity and accuracy.

For these reasons, no one group fully controls every stage through which an opinion becomes available for political action. Responsibility for a particular classification may be difficult to trace when several actors rely on one another's records, engendering politics of decentralization in the limited sense of dispersed informational influence. In such politics, not even the constituents fully control how their opinions are represented; instead, influence over information extends to various seemingly remote entities, including but not limited to corporations, photographers, statisticians, media companies, government, and so on. It is not always certain which judgments shape the information that reaches a decision-maker, once again highlighting coerciveness without apparent coercion. Yet dispersion does not imply equality: those who design a format may retain considerable power over everyone required to use it. This decentralization concerns influence over the information available to the congressmember; it does not remove the formal authority of constituents or the responsibility of their representative.

In the age of algorithms and data, the traditional *politics of representation* can become more like a politics of simulation when decisions rely on formatted inputs that are treated as complete substitutes for constituents' opinions. Here, simulation describes a system that relies on components of reality to produce an outcome that encompasses features of reality yet obscures what has been omitted. All representation requires some mediation; the concern arises when the office can no longer distinguish the summary from the fuller opinions it is meant to represent. This is a risk suggested by the tally, rather than proof that congressional representation has already been replaced by simulation.

One alternative to the binary tally is to track the attitude towards an issue by saying something like: on a scale of 1-10, to what extent does the constituent support the issue, from 1 being least supportive to 10 being most supportive. Such practices would enable a potentially more holistic representation of the constituents' opinions. A scale would still require a clearly defined issue, however: measuring the degree of support for Ukraine does not resolve whether a constituent favors military aid or economic relief. Standardization also enables the office to hear and compare more messages than it could manage as isolated accounts. A practical response is to preserve that advantage while allowing mixed or uncertain positions and keeping the written descriptions connected to the totals. More categories alone will not guarantee accuracy; the office must also consider whether those categories answer the question for which the tally is being used.

Conclusion

These observations suggest that data exercises a peculiar kind of power through the everyday practices by which institutions make opinions available for political action. Formatting as a form of power exercised through data generalizes information presented and limits available choices to the

audience. Through this process, everything from constituents' opinions to an individual human being can be reduced to labels that carry some characteristics of the subject yet reduce and omit other traits, qualities, and connotations that are entailed. Nonetheless, formatting expedites the process of information propagation as labels such as "Black", "liberal", and "student" are immediately ready for organizations and individuals to put into practical usage such as conducting a census and making respective policies. Meanwhile, this accelerated propagation of information ties the subject to its labels, known as fastening.

Because of the judgments embedded in such data, politics of decentralization and politics of simulation become possible, challenging the conventional epistemology of power and politics. In the Office of Congressman Lieu, the issue tally that is used to track the opinions of constituents offers a concrete site for examining these possibilities. One possible way to pursue more accurate representation is to provide more available choices, ameliorating the restrictive authority of formatting, yet doing so still does not solve the problem that we are becoming the subjects of data and that data exerts power through the human actors and institutions that use it. This might push us to reconsider our relationship with each other and the government as data helps constitute the political ecology in which representation takes place.

Bibliography

Congressional Office Records. 2022. Office of Representative Ted Lieu. Internship materials.

Gitelman, Lisa, and Virginia Jackson. 2013. “Introduction.” In *“Raw Data” Is an Oxymoron*, edited by Lisa Gitelman, 1-14. Cambridge, MA: MIT Press.

Koopman, Colin. 2022. “The Political Theory of Data: Institutions, Algorithms, & Formats in Racial Redlining.” *Political Theory* 50 (2): 337-361. https://doi.org/10.1177/00905917211027835.

Noble, Safiya Umoja. 2018. *Algorithms of Oppression: How Search Engines Reinforce Racism.* New York: NYU Press.

Ribes, David, and Steven J. Jackson. 2013. “Data Bite Man: The Work of Sustaining a Long-Term Study.” In *“Raw Data” Is an Oxymoron*, edited by Lisa Gitelman, 147-166. Cambridge, MA: MIT Press.