

## **Data as Representation**

Sean Martin McDonald (Centre for International Governance Innovation) and Ben Gansky (Arizona State University)

### Abstract

When people share and use data they are, more often than not, making a representation. Here, we draw on the legally defined concept of representation - meaning an assertion of fact in a context, especially in rights-impacting contexts. While data also has other uses and contexts, the fundamental difference between data and other forms of exchange is that data is information - and the primary end goal of its exchange and use is the assertion of its relevance and value in a defined context. Understanding data as representation means that the exchange of data implies not only 'technical' claims (towards useability and veracity), but also political claims: that the underlying relationships and context that comprise the data supply chain and governance relationships are legitimate. The reason that people exchange, use, and rely on data is not "to own all the data sets," it's because of the characteristics of the information it represents.

A representation is the assertion that a data set not only contains a technically usable representation of the underlying information, but that the political representation that the underlying relationships and context that comprise the data supply chain and governance relationships are legitimate. The exchange and use of data, as an act, while containing a 'material' exchange, is indistinguishable from the political economy and power relationships that define the appropriateness of that exchange - including the rights that emanate from those relationships, regardless of their means of administration - or the intended impact of that data's use.

Historically, assessments of data - and the supply chains that underpin its aggregation and use - have focused on functional and qualitative characteristics, correlating data's assumptive value to its completeness, 'quality', and volume. As the public use of data to achieve governance outcomes has matured, both in theory and practice, it's increasingly clear that the value - and liabilities - created by data are increasingly defined by a different set of qualitative characteristics, like the diversity, legitimacy, and ethics of the underlying digital supply chain. This shift from useability to politically normative rubrics implies a reactivity to the rights and responsibilities emanating from the relationships encoded in data. Relationship models, as others have noted, are an integral component of the normative character of digitally intermediated relationships - most often articulated and enforced by an ecosystem of institutions. In particular, duty bearing professions - those that are regulated by public and private institutions - provide models for the way that we might govern digital representation relationships; they are realized by institutionally regulated, tangible, operational infrastructures designed to ensure the integrity, equity, and symmetry of power in inherently asymmetrical representation relationships. While there is a significant range of practice, both within and between duty bearing professions, there

are common governance design patterns that offer valuable guidance for those attempting to design integrity measures for data and digitally intermediated relationships.

Rather than attempt a universal theory of data valuation, this article uses the common characteristics of professional governance in duty bearing professions in order to propose a valuation framework for data - not as an object, but as a representation in context. This article proposes an approach focused in three parts: (1) the articulation and limitation of animating purpose - essentially, the rubric for determining the legitimacy of use in a context; (2) the relationships between the scope of representation, standards of expertise and care, and boundaries of available representative actions; and (3) the responsibility to support, if not provide, means of independent oversight and accountability. These broad dynamics won't address a commodifying approach to data governance - rather, they will provide ways to contextually assess the value and risks of the different approaches to building data supply chains, based on the highest-integrity models implemented in relevant contexts.