Stefaan G. Verhulst

Today's Rembrandts in the attic: Unlocking data for social value

Twenty years ago, Kevin Rivette and David Kline wrote a book about the hidden value contained within companies' underutilized patents. These patents, Rivette and Kline argued, represented "Rembrandts in the Attic" (the title of their book). Patents, the authors suggested, shouldn't be seen merely as defensive tools but also as monetizable assets that could be deployed in the quest for profits and competitive dominance. In an interview given by the authors, they referred to patents as "the new currency of the knowledge economy."

We are still living in the knowledge economy, and organizations are still trying to figure out how to unlock under-utilized assets. But today the currency has shifted: today's Rembrandts in the attic are private and public sector data. And its value resides in finding ways to responsibly re-use that data to inform public interest decisions.

At The GovLab, an action-oriented think tank located within the Engineering School of NYU, we are dedicated to unleashing the societal value of data to improve decision making in the public interest. If there is an overarching theme that emerges from our work, it is about the value of *re-using data for improving people's lives*. In recent years, several countries have witnessed the rise of an open data movement, and a growing number of private organizations have taken steps to release or made accessible previously siloed data sets. Despite occasional trepidation on the part of data holders, our research has repeatedly shown that such efforts can be value-enhancing – both for data holders and for society at large. Better and more transparent re-use of data is arguably the single most important measure we can take to unleash the full possibilities of data. Toward that end we should consider four immediate steps:

1. Develop new participatory methodologies to identify and measure the value of data

The first step required to fulfill this potential is for all stakeholders to arrive at a better understanding of just what we mean by value. Today there exists widespread consensus that data is valuable. Despite such agreement, however, there exists no equally accepted method for calculating or estimating the value of data. Such a consensus must be arrived at through a broad process of consultation that involves data holders and users from all sectors, as well as policymakers, researchers and academics, and civil society or other groups representing the public interest. In particular, it will be required to collectively determine what are the questions we seek to answer by re-using data (See The 100 Questions initiative) and subsequently acquiring a social license to re-use data to answer these questions through citizen assemblies (See The Data Assembly).

2. Develop enabling ecosystems and collaborative frameworks to move from extraction to co-creation of value

Unlike physical assets, data goods are non-rivalrous and intangible, which means that they can be shared without depriving their original holders of benefit. The process of maximizing under-utilized data assets will therefore often involve arriving at new institutions and frameworks to enable data collaboration and what we call "co-creation of value." This concept of co-creation is not new and various experts have called for the creation of new institutions to facilitate it in different sectors. In her book, _The Entrepreneurial State_, University College London Professor Mariana Mazzucato argues that such a framework is necessary to bring the public and private sectors together to spur innovation.

Drawing on the analogy with patents (those earlier "Rembrandts in the attic"), it is worth, in this context, pointing out the dangers and risks of _not_ sharing. While patents can be competitive assets for companies, they also often block innovation and prevent true competition from emerging. In much the same way, data hoarding can result in broader societal and monetary losses. These losses may ultimately rebound on the data holders themselves, who fail to benefit from missed-out innovations or breakthroughs.

3. Innovate with new data collaborations and re-use conditions

In order to enable sharing, we need new structures that foster partnerships and more collaborative approaches. The old model of single-ownership is outdated and no longer conducive to maximizing the value of data assets. Several structures have been proposed, including data co-ops, data commons and (our preferred term at the GovLab) data collaboratives.

Data collaboration can take many forms. In our typology, we generally focus on two defining variables: engagement and accessibility. The first variable, **engagement**, refers to the degree to which the data supply and demand actors co-design the use of corporate data assets. We find that collaboration is often independent, in that the private-sector holder has little to no involvement in data re-use, cooperative, in that data suppliers and data users work together, and directed, in that the data holder seeks a specific product. The second variable, **accessibility**, is the extent to which external parties can access private data. Within it, we find that data is either open access, in that there are few restrictions on who can see it, or restricted, in that only pre-selected partners received unfettered access.

4. Identify and nurture data stewards

As data collaboratives and other similar structures gain increasing validity, it is becoming clear that new human and institutional roles will be required to foster them (and more generally to encourage a culture of sharing). In our work at the GovLab, we have identified a key role within data holding organizations for what we call _data stewards_. As the European Commission's High-Level Expert Group on Business-to-Government Data Sharing recognizes, these individuals or teams empowered to

proactively initiate, facilitate, and coordinate data sharing are essential to using cross-organizational and cross-sector data toward the public interest.

Data stewards are individuals or groups who manage data within organizations, and whose specific remit is to foster collaboration and sharing, with an eye to maximizing both societal and monetary value. Among other responsibilities and roles, data stewards can identify under-utilized data that may have potential value; locate and foster partnerships to help unlock that value; and ensure a responsible framework that balances potential benefits of sharing against possible risks such as harms to privacy or security.

To conclude: Re-using data is a vital step toward generating social value in data. Yet we are only beginning to understand the trade-offs involved in re-using as well as the institutional frameworks and structures that can encourage it. The four points outlined above represent a start, but we need a rigorous assessment into what's already being done, and more experimentation to push the frontiers of what's possible. We need, above all, a more creative and innovative mindset that can help organizations dust off the cobwebs from their hidden Rembrandts, in the process allowing them – and society at large – to maximize the monetary and public good value of our ever-growing data streams. Toward that end, looking forward to this workshop.