Stanford PACS
Center on Philanthropy and Civil Society
—
Digital Civil Society Lab

Workshop Summary

Nonprofit Data Governance

By Lucy Bernholz and Rob Reich
January 2017

We begin with a simple claim: digital assets should have the same status as human and financial resources. All three need to be managed with integrity toward organizational mission. Any smart organization already does this with human and financial resources. Adding digital assets to the list requires rethinking – and redesigning – the "tech stack"– from the software code that powers their constituent management systems through to the board policies that address managerial responsibilities and organizational liability. We are at the beginning of reinventing the nonprofit organization to become purpose-built institutions designed to dedicate and protect digital and analog resources toward public benefit.

Nonprofit corporations differ from their commercial counterparts through a legal tweak to the rules of ownership. They are defined by a so-called "nondistribution clause," a governance requirement that all revenue above costs be used for mission rather than as profit to shareholders or owners. Nonprofits can earn revenue, including revenue that exceeds costs, resulting in a "profit." But this profit may not be distributed beyond the mission of the organization. .This governance innovation is the legal mechanism that facilitates the public trust that these nonprofit organizations will use their financial resources to produce social benefits.

Today's challenge is to figure out how these same organizations can and should govern digital resources for public benefit. Digital assets are more varied than financial resources, they come from a wider variety of constituents (including beneficiaries), and they are often exchanged over platforms and software that put third-party interests between donors and recipients. Digital data are also subject to an ever shifting mosaic of rules depending on whether they represent children, medical, educational, consumer protected, legal clients, journalistic sources, or financial information. Given the characteristics of networked digital data and the expanding realms of protected information, the wisest strategy is to design protective systems as defaults.

Let's be clear. This is a far cry from where we are today, in which most software systems presume the desire and capacity to collect, aggregate, and analyze unencrypted and barely protected data, frequently with nothing more than an obscure and unread Terms of Service agreement. For civil society organizations that largely lack the capacity or the scale to benefit from data analytics, the risks of these TOS defaults tend to outweigh any benefits.

The problem goes deeper. Almost all civil society organizations are dependent on software designed for the for-profit marketplace, for a world in which money is to be made from the collection and aggregation of digital data. The data defaults of most such software amounts to "collect all and hold forever." But for nonprofits and foundations, we argue that the default setting should be, "Collect little and destroy as soon as

Continued

possible." Commercial firms and public agencies trying to govern digital data also must attend to the ways that networked binary code differs from financial or human resources, but the data defaults that lean toward profit or surveillance are, at least, in line with those sectors' broader purposes. Civil society is different.

When we picture the nonprofit or philanthropic organization of the 21st century, we assume it will be using donated resources – financial, human, *and digital* – to achieve its public purpose. For financial and human resources we have models of good practice, from internal financial checks to external auditors, legal guidelines that include nondiscrimination laws and human resource managers. But when it comes to managing and governing digital data we are in uncertain and sometimes unchartered territory. Some organizations must follow strict guidelines for managing certain types of data. Organizations that manage medical or student data, for instance, will be familiar with strict regulatory guidelines, software requirements, training protocols, and privacy restrictions. Such nonprofit organizations know that using digital data well requires more than just the right software; it also demands personnel training, organizational policies, and board liabilities.

All organizations using digital data now need to consider how their dependence on this resource changes their technological, managerial, and governance practices. Civil society organizations are further challenged because these dependencies adhere the entire (theoretically independent) sector to governments and commercial actors in ways not yet fully acknowledged. Digital ties between nonprofits and governments are persistent and pervasive. They include nonprofit dependence on public data sources for program work and advocacy, an unintended consequence of open data, open government, and transparency initiatives. More broadly, government surveillance of the internet and wireless spectrum means that all digital communication, including that of civil society organizations, is swept into state-monitored systems, potentially eliminating independence and privacy. The same dynamic is at play for nonprofits and foundations using commercial software, cloud storage, or platform services in their default modes. Civil society, as a space free from government or commercial monitoring, needs its own regulatory standards that will protect its independence. This requires carrying into digital environments – or crafting anew – normative, legal, organizational, and technological mechanisms that will bound the associational, expressive and privacy rights that give rise to civil society. This is a challenge not just to the resilience of individual organizations, but to the democratic systems that depend on a flourishing independent sector.

Individual nonprofits and foundations now run on digital data – from emails to home addresses, performance measurements to programmatic information on beneficiaries (often vulnerable populations), evaluation data and donor information. Some domains are ahead of others, but in general we are at the early stages of designing practices and policies for managing and governing digital data safely, ethically, and effectively for public benefit.

The challenges faced by nonprofits and foundations stemming from the characteristics of digital data and infrastructure include:

- Defining processes for donating digital assets where multiple owners have many rights;
- Developing mechanisms to respect the intent of digital data donors; particularly the data donated by beneficiaries or constituents;
- Designing organizational practices that will allow voluntary digital participation and data withdrawal (given the known limits of "notice and consent" processes);

- Building mechanisms for public oversight that enable the open and accountable application of algorithmic and data-driven decision making tools;
- Continuously adjusting to the ever-shifting nature of privacy incursions made possible by aggregated digital data from multiple sources;
- Inventing organizational governance policies and practices that include data stakeholders.

The Digital Civil Society Lab at Stanford's Center on Philanthropy and Civil Society has been investigating these questions for the past year in a series of workshops with nonprofit organizations, scholars, policy makers, and digital data experts. We have investigated the rise of governance practices custom-built for digital resources, including ethical review panels, independent data policy boards, open source hubs, and contracted relationships between nonprofits, community members, and commercial data owners. We've looked at the roles being played by what we've labeled trusted data intermediaries, organizations such as ArtStor or LearnSphere that aggregate digital data from other providers and negotiate the use thereof for a variety of different end-users. And we have examined the existing "tech stack" that nonprofit organizations and foundations are using to identify shared "pain points" that might indicate areas for collective action.

These different elements are connected but hard to synchronize. Creating a coherent tech stack for mission-driven organizations requires aligning software defaults, operating practices, organizational governance, and public reporting and oversight expectations. It requires the coordination of software, legal, and organizational codes.

Digital data challenges for nonprofits

The questions before us are not those of data science or analytic methodology. They are instead a mix of internal challenges about resource stewardship and external challenges shaped by the political economy of civil society. The internal, or organizational challenges, have both managerial and governance components. They involve decisions, tactics, and strategies for protecting and stewarding digital resources in line with a nonprofit's or foundation's particular mission. The external challenges arise because the vast majority of software and digital infrastructure used by nonprofits and foundations are commercial products and government surveilled systems. The nonprofit sector has largely had to compromise its values, often unwittingly, to fit the default commercial offerings of these digital platforms and tools, and has only occasionally been able to leverage any collective power to develop and maintain digital tools that align with its values. Even then, the sector's reliance on commercial and public digital infrastructure compromises the sector's cherished sense of independence.

The organizational challenges of digital data, both managerial and from a governance standpoint, are distinct from the programmatic or analytic challenges of using data for performance measurement or evaluative purposes. In an ideal situation, operational questions will be answered in ways that facilitate these performative tasks. In the real world, the desire to collect and use data pushes the operational challenges to the fore.

Nonprofit organizations in many domains, those with regulated data procedures and those without, face a number of unanswered questions. These require sector-wide consideration, as the tactical answers to these questions will shape practice and policy for many organizations. Our work identified the following such cross-cutting challenges:

Mission based practices across the digital data lifecycle

Philanthropy as we know it has been designed around analog assets for which ownership claims are clear. The owner of a resource donates it, relinquishing control, to another enterprise that then dedicates the resource to mission. Digital data, on the other hand, are almost always contested property. Most digital data is created via processes that involve multiple potential owners – the person whose information is involved, the institution that provides the software being used (the school, nonprofit, clinic, etc.), and the software or platform vendor whose tools are being used (e.g., Salesforce, Blackbaud, Mailchimp, Facebook, Google). In addition, the network providers (telecommunications companies and ISPs) will also have a claim to data that travels over their networks.

As a result, every digital data entry – from an email or a text message conversation to the photos on websites the donor records in an online database – has multiple potential "owners." Each entry certainly has many different points of access, even if ownership is left unclear. Very few nonprofits have a clearly communicated ownership position, vis-à-vis the different types of data they collect from their stakeholders. In most cases, the norms of ownership are being set by software vendors and platforms, with little consideration from the nonprofits and even less from the person identified in the email or photo.

While nonprofits have processes for managing the voluntary donation of time or money, the dynamic for data donations are structurally different, beginning with the fact that most such "donations" are intermediated by software, meaning that there is no longer a one-to-one relationship between the donor and the organization but a triangulated relationship involving software platform companies or data brokers. This abundance of data owners can be a positive attribute. For example, it provides numerous opportunities to directly engage the individuals represented in the data in governing the use thereof, creating new forms of participation, power, and leadership. Civil society organizations can use digital data as a catalyst to create more pluralistic governance models.

Doing so will require stepping back from analog notions of ownership as a set of exclusive rights and finding new ways to manage a multiplicity of owners with a multiplicity of rights. There are many models here, including those that undergird the commons and those that manifest in some traditional knowledge practices. There are also digital and technological contributions that are being used, ranging from alternative licensing regimes to digital rights management software that privileges individuals' relationships to their data. What we know for sure is that civil society has served a unique function in the analog age of helping to make "private resources public." If it is to continue to do so in the digital age it needs to engage head on with questions of digital ownership.

Ownership and rights to use plays out over a different lifecycle for digital goods than for analog ones. Digital practices for "notice and consent" or "permissioning" are difficult to create or effect, especially for downstream use of collected data. Part of the challenge – and perhaps part of the solution for nonprofits – is to think about consent as a value differentiator, build it in as part of program design, and shift away from the compliance-based approaches with which most of us are familiar. Nonprofits and designers in both the

medical and legal fields are taking on the challenges of digital consent. There are numerous experiments underway to reconceive of consent in digital environments. Two key players in this space are the Stanford Legal Design Lab and Sage Bionetworks (working with donated data for medical research)

Nonprofits and civil society organizations need to change how they think about consent. Social media, remote sensor data, massive aggregation, and dependence on third parties all but eliminate meaningful "consent moments." They also make it very difficult for individuals to "opt out" even if they've had a chance to "opt in," thus changing the meaning of what might have started as voluntary participation. In many digital contexts, a nonprofit organization's best efforts at consent will be constrained by the terms of service of the software platform, rendering them possibly useless. Even this reality allows for more explicity exchange than the shift to remotely collected passive sensor data coming from internet connected devices, building monitors, vehicle sensors and other sources. These are pervasive, passive data collection situations. These environments offer no "log on" moment at which consent can be requested, nor any way that one could opt in (or out). Nonprofits can make decisions about using such "Internet of Things" devices within their own facilities (thermostats, building security systems, robotic delivery systems) and in the choices of technology gadgets (tablets or phones with microphones and cameras) they deploy in the field. But the challenge runs further than that, for nonprofits to engage in the political and technological battles that will determine how pervasive these sensors will become, what kinds of warnings or options individuals are allowed, and where they may simply be inappropriate.

Nonprofits and foundations also need data practices that navigate between two values that sit in constant tension – privacy and openness. Voluntary associational life in democratic society thrives in this ever-shifting passage way. Individuals need places to participate privately, without being monitored or compelled by external forces. And associations need to communicate openly both to advance their missions and to be accountable to the public. Democratic societies struggle to strike the right balance for civil society organizations; the nature of digital data and the numerous new relationships that come with it adds new dimensions to this tension. For example, a donor's right to make an anonymous charitable gift could be handled as private information within an organization, withholding the name of anyone involved while still publicly reporting the receipt of funds. Today, the ecosystem of digital records (payment processors, political activity, organizational purchases, online databases of nonprofit financial information) make it harder to ensure that such transactions aren't identifiable. Today an individual's right to associational privacy is arguably as challenged by the ready access to triangulated digital purchasing information, web search information, or social network analysis than by the content of organizational mailing lists.

But this tension between privacy and openness extends far beyond the known issue of donor anonymity. The nature of digital storage ties privacy practices to a nonprofit's security capabilities. Here we see just how thoroughly good data management practices implicate the entire organizational structure. Good digital security is not one person's job. It requires everyone in the organization to practice good digital hygiene, vendor contracts to be negotiated through the lens of organizational mission, decisions about cloud storage and remote access to be viewed from the perspective of beneficiary vulnerability, and board policies that plan for data breaches or government demands, strategic risk modeling, and organizational communication practices that account for volunteers who use privately owned phones or tablets. Managing digital resources well requires assessing staff capacity to analyze digital data, setting access controls that include staff and board, and assessing the organization's dependence on third-party systems and their data policies. In short, like all good governance and management actions, it requires a commitment of time and money.

The persistence of digital data throughout civil society organizations requires leaders to think about primary, secondary, and long term uses of digital data. It's tempting to take information collected for one purpose and hold onto it in anticipation of new opportunities. This is certainly the digital data message that currently dominates – collect more and hold on to it. But nonprofits and foundations may find that this doesn't hold for them – out of consideration of the trust with which the original information was provided, the organization's own capacity for analysis, and the liability costs of storing and protecting that information. Essentially, digital data and its affordances may or may not align with the mission or organizational integrity of civil society activities. The value or costs of digital data, unlike financial resources, change at each step in the information lifecycle – from collection to storage, analysis, access, to retention and deletion decisions. Organizational assessments regarding digital data need to account for this entire lifecycle.

Every organization faces the challenge of knowing what data it has to govern. Even public libraries, which have expertise in data management and privacy issues, have to expand their "data inventories" in the digital age to include not just patron records but also the video footage captured on their closed caption security systems. Similarly, hospitals are having to account for "smart" devices – from the televisions in patient rooms to the robot delivery systems – to ensure that they are neither collecting nor leaking sensitive medical information. And while an increasing percentage of data is entirely digital, nonprofits also have a lot of information collected or held in analog form, either initially (e.g. meeting sign-in sheets) or forever (e.g. historic donor records). Data management and governance policies need to extend forward and backward in time, crossover multiple interest holders, and, perhaps most important, be understood and used by everyone at the organization. A set of policies is insufficient, what is needed are both useful practices and known processes for updating and advancing that practice.

From organization to sector

There are no easy answers. Many of these challenges will be most effectively and efficiently met by collective action focused on setting sector norms and practices. Too many associations and nonprofits lack the legal, technical, or organizational skills to take on these challenges individually. Several practices are already being shared across the sector, from the use of ethics review panels to the rise of trusted data intermediaries. We identified several additional opportunities including:

- Creating a set of sector-wide norms and contact defaults for negotiating with third-party software vendors.
- Adapting existing codes of practice (e.g. the Library Bill of Rights or the professional ethics resources held at the Illinois Institute of Technology).
- Using sector-wide resources such as the Berkman-Klein Center's work on Intellectual Property and licensing for foundations
- Demonstrating practices for clear, goal-enhancing alignment between privacy protecting actions and mission accomplishment.
- Amplifying and accelerating the use of existing, community-proven resources such as the Responsible Data Forum's tools.
- Collective, ongoing investment in open source technology stack (partially exists, some needs to be built) for civil society enterprises. Stack needs to be seen as in and invested in as independence-protecting infrastructure for funders and nonprofits.

Options for organizations and the sector

Building organizations – and a sector – that uses digital resources for public purposes is an opportunity to define civil society for the present and future. There are three distinct levels of challenge to doing so: first, organizations can equip themselves to harness data to mission; second, the sector as a whole can develop norms, rules, and tools to facilitate this organizational change; and, third, civil society can lead all of society by creating privacy-protecting norms of practice where, so far, markets and governments have failed to do so. This third level is a unique opportunity for civil society organizations given their role in dedicating private resources to public purpose and the values that undergird their independence in democracies.

In short, this is a sector-defining moment. Seizing it requires first recognizing the degree to which civil society is currently dependent on commercial and government provided digital infrastructure, tools, and norms. This can be seen when one looks at the common "tech stack" on which today's civil society organizations depend. Figure One, below, is an adaptation of the common model that shows "five levels of the network" to focus in on those levels over which organizations have some control:

Figure One: Network levels from Organizational Perspective[1]

LEVEL OF NETWORK	Examples (illustrative)
End user devices	Phones, tablets, desktops
Enterprise software	CRM, Document sharing, social media choices, word processing, databases, open source alternatives
Enterprise hardware	PC, Apple, open source hardware
ISPs, web hosts, cloud storage	Amazon, Electric Embers, RiseUp, Google
Network access	Telecommunications services, backbone providers

Civil society organizations that seek to align their data models with their missions must consider how their choices – both technological and legal – carry through each level indicated above. The values that drive their mission and shape their organization can be reflected in the technical and legal choices they make. These values – and the choices they engender – will then inform the particular software, hardware, and network vendors with whom the organization works. Figure Two shows a hypothetical set of relationships between an organizastion's values and its technical and legal considerations.

[1] See Xplane's images of the levels of network governance, ICANN, [citation]

Figure Two: Aligning organizational values with technical and legal considerations

Organizational values	**Technical choices (Illustrastrative)**	**Legal Considerations (Illustrative)**
Institutional control over data (allowing permissioning for various constituencies)	Open source or highly customizable	Vendors' data policies; storage locations; network of software support (open source v proprietary)
Visibility into decision making – transparency and accountability	Open source software, algorithmic or analytic platform	Vendor's data policies, proprietary nature of code
Privacy protecting	End to end encryption; strong privacy protections from hosts; transnational protections	Location of servers, vendor commitment to privacy and transparency; national/regional data protection regimes
Informed and voluntary relationships with constituents	Interoperability standards	Data ownership and portability rules

In the digital age, where we are dependent on the encoded values represented in our hardware and software choices, an ethical tech stack involves aligning our organizational logic with our technology choices. To date, this type of alignment has been most widely practiced in subsets of civil society – rights protecting organizations, journalistic endeavors, and advocacy organizations have been on the forefront of doing so. There is much that the rest of civil society can learn from these trailblazers.

There is an additional set of steps necessary for all civil society organizations. Having aligned one's mission with one's technological and legal choices for digital tools, we must further align the governing and managerial logic that drives the sector. Whereas we have previously focused civil society organizations on managing and governing analog resources, these practices also need to be redesigned to account for the fluid, generative, intermediated, and pervasive nature of digital data.

Once those dependencies are understood we can articulate both best practices for organizations and reestablish the boundaries that provide civil society with its independence and resilience.

Nonprofits, voluntary associations, and philanthropic institutions can define themselves by their respectful, trustworthy use of digital data for public purpose. Doing so is critical to their ability to accomplish their missions, to retain and extend their role as trusted resource stewards, and to distinguish themselves from commercial and government actors.