Principles for Revenue Models of Data Stewardship

Aditi Ramesh* and Astha Kapoor

Aapti Institute | aditi@aapti.in

Abstract

The balance of societal good, market innovation, and individual rights forms the core of questions on data governance. These issues merit further research – to standardize the objectives, rules, and governance of data sharing across a variety of contexts and use cases. The value of a data steward, or an intermediary who works on behalf of the users and entities to manage data and its sharing is increasingly apparent.

One of the biggest concerns, given data is an intangible and valuable asset, is how business models can be structured so that stewards remain incentivized to serve the interests of individuals and communities, and do not use the data for their own ends. Data is a complex resource, and to understand potential revenue models of a steward, this paper studies resources or assets which embody the complexity of data and the accompanying ethical issues of protection.

In this paper, we root our analysis three theoretical framings: data as commons, data as labor, and data as an asset. With these framings in mind, we examine real world examples of "stewardship" to extrapolate principles for revenue models of a data steward.

Keywords – data governance; data stewardship; revenue models; commons governance

1 Introduction

There is little doubt that the global discourse on effective data governance mechanisms is evolving and, in fact, gaining pace. The societal value of data is clear – it can be variously used to develop medicines, track rainfall, or manage traffic. This aspect is being increasingly realized in the context of the COVID-19 pandemic, when availability of reliable data for contact tracing or pertaining to public health systems has been critical for government response. Efficient, trustworthy and equitable processes for collection, analysis, management and sharing of data need to be built so that more data can be deployed in public interest.

Data is a complex resource – making it difficult to create responsible sharing mechanisms. There is friction between data protection, ensuring that individual rights to privacy are safeguarded, and data sharing, which unlocks the value of data. Businesses regard data as intellectual property in many instances, which adds to the complexity of sharing. It is thus imperative to scrutinize sustainable mechanisms to enable the sharing of data while at once safeguarding rights and enhancing individual agency.

A data steward enhances accountability of platforms, user control over their data, and, consequently, trust in the processes of data sharing, use and analysis (Manohar et al,, 2020). It also allows for multi-stakeholder involvement and the safeguarding of data in the interests of individuals and communities. A steward, when successful, can fundamentally reimagine the way in which data is collected and controlled. It can restore the agency of individuals and communities and make them active stakeholders in their data lives.

A steward must find independent streams of revenue so that it is not co-opted by technology companies, data acquirers, or state bodies, and continues to serve and protect the rights of individuals and communities while opening up data for societal benefit. This is a complicated task, and the business models are likely to reflect this complexity while following ethical principles of fairness, integrity, decency, and sustainability (Hagenbuch, 2015). We argue that fiscal independence and consequently a robust business model is a prerequisite for a "good" data steward, one that actively seeks collaborations through data sharing but does so in a rights-first manner and one that can build technologies, protocols and standards to share data while ensuring that acquirers do not misuse it.

We argue that a steward's revenue generating structure is intricately linked to its function of safe, responsible sharing of data and the ways in which it imagines its relationship with individuals and communities. In thinking about "real world applications" of stewardship, we root our analysis in the following theoretical framings that help understand how data and the relationships around it are imagined.

1. *Public goods stewardship:* This framing imagines data as collectively owned and governed as "commons", and examines the stewardship of public goods such as land or trust ports
2. *Stewards for collective bargaining:* This model examines the framing of "data as labor", and the role of the data steward as an entity that works on behalf of the data subjects and can help negotiate more equitable data rights with technology companies
3. *Financial intermediaries as steward:* Finally, we imagine data as an asset that can be deployed for the benefit of individuals and communities, and examine financial intermediaries such as investment advisors

It is important, here, to note that our focus is on revenue models, and not business models. We seek to understand how models for data stewardship can generate income and are not concerned with broader business issues such as acquiring customers, and costs of running a data steward*. The Open Data Institute's work on "Designing Sustainable Data Institutions" serves as a starting point for us, in thinking about earned revenue models (Dodds et al., 2020). However, we acknowledge that long-term sustainability of a data steward will rely on how earned revenue compares to operating costs.

In this paper, we dissect the nuances of each framing, drawing from real-world examples. For instance, to understand governance of public goods, we examine Scotland's trust ports and Alaska's Permanent Fund Dividend to investigate the applicability of revenue models to data stewardship. In the following sections, we undertake similar analysis of union contributions, and financial commissions. In the final section, we round off the research paper to extract principles for data stewards.

2 Public good stewardship

As David Bollier says, the commons constitute a social system to share wealth, such that people can control, manage and distribute resources such as data (Edwards, n.d.). The commons approach to data governance draws the following parallels between data and public goods resources: First, data is a resource which many people can and do use simultaneously, and for varying purposes. It is a non-depletable and non- competitive resource, i.e. the ability of one entity to draw value from data does not interfere with another's. Second, data is a resource that is

more valuable when packaged together rather than siloed or broken down into individually owned chunks. It benefits from network effects, and the network accrues greater value as more people join in (Panfil & Hagopian, 2019).

By taking a commons lens to data, we can develop principles for revenue models of stewardship that distribute benefits widely, without commodifying or privatizing the 'goods'. In the following sections we explore examples of public goods stewardship to further our understanding of revenue structures to extrapolate to data, if it were governed with the commons approach.

2.1 Scotland: trust ports

Scotland's coastal ports are home to vast amounts of economic activity, from renewable energy generation to cruise activities to fishery management. Trust ports, legally recognized by the state, operate with a stakeholder-board governance structure. It is the responsibility of the board to govern, safeguard the port, and manage operations while finding a balance to accommodate the interests of all stakeholders. The stakeholders of a trust port can be port users, the local community, local and regional authorities, related interested groups, or local and regional businesses.

Unlike typical board-stakeholder structures, these stakeholders do not have direct financial investment in the port in the way that shareholders do in a private company. The stakeholders' primary powers lie in their ability to keep the board accountable in a responsible manner and monitor the port's performance. As a result, the primary aim of the trust port "is not the production of profit for shareholders but the best use of the assets they manage in order to secure that asset for future use." (Modernising Trust Ports, 2009) Because stakeholders pay harbor dues, they are allowed to scrutinize the performance of the board through performance indicators when necessary. Trust ports operate like commercial businesses, seeking to generate a surplus that is then invested back into port maintenance, development, or operations.

2.2 Alaska: permanent fund dividend

Set in place by the Alaskan government in 1982, the Alaska Permanent Fund Dividend (PFD) mimics the

---

* We will deal with costs of running a steward in subsequent writing. This paper focuses on earned sources of revenue and intentionally omits other revenue sources such as grants, donations and tax breaks to ensure that only independent models are analyzed. We do recognize that the idea of data stewardship is fairly nascent and may require grant support in the initial stages.

However, for models to be successful over time and deliver on their commitment of responsible intermediation and data sharing, independent revenue models are essential.

universal basic income model, giving each resident a portion of the ~$66 billion fund (Feloni, 2019). The state deposits royalties it receives from natural reserves (mineral, oil, gas, etc.) into the fund annually. This money is then managed and invested by the board members in domestic and global stock, bonds, and private equity; earnings from interest are then distributed to residents. The beneficiaries in this format are the citizens themselves, who are the recipients of the annual dividends of around $1,200 (subject to change based on returns), which is considered a basic income by the Alaskan government. The money is thus redistributed to the people whose resources are being exploited ("The Board of Trustees", 2019).

The PFD mirrors the independent, fiduciary duties of a data trust, in which trustees are bound by a fiduciary obligation of undivided loyalty and exercise rights on behalf of the trust's beneficiaries (Delacroix, 2019). This model also follows the model of dividend payments for technology companies that is being suggested actively.

2.3 Community land trusts

Community Land Trusts (CLTs), typically not-for-profit organizations, acquire land and lease it to families or individuals, allowing them to enforce restrictions on the use and affordability of such housing (Semuels, 2015). CLTs are not a trust by legal designation (and as a result, do not have a duty of care) but do have a similar structure; they are governed by a board of directors, comprising individuals selected to represent varying community interests and constituencies.

CLT membership is open to anyone who leases the land and resides within its geographic area. CLTs cover the cost of their operations in a variety of ways. Many rely on grants in their early years (either from private sources like foundations or through public sector funding). At some point, CLTs begin generating revenue internally through ground lease fees, lease re-issuance fees, membership fees, and fees for services. These fees are used to cover their stewardship responsibilities, "especially the cost of monitoring and enforcing the occupancy, eligibility, and affordability controls that encumber a CLT's housing" ("FAQs about CLTS", 2007).

CLTs serve as worthy models for governance structure and stakeholder representation, involving public representatives or trustees, which allows for inputs related to community interests. CLTs also highlight the potential for varying types of fee-to-entity structures, from service fees to interest accumulation.

2.4 Revenue structure: monetizing the commons

We recognize that the design of a steward and governing principles will likely inform its revenue structure. In thinking about data, using the commons approach, we apply the stakeholder-board format shared by many "common-pool resource" management models to contemplate revenue models for data stewardship. In the three models discussed above, the steward is generating revenue through the management and monetization of the resource (port, oil and land/water). However, there are two key differences – in redistribution of value, and in mechanisms for stakeholder participation. Alaska's PFD redistributes value to the people, it defines profits made from natural resources as entitlements of the community; in Scotland, any profits made from the port is invested back in maintenance of the asset. With regard to stakeholder participation, in Scotland, stakeholders pay a fee which entitles them to participate in decision-making on issues of the trust; in the case of the CLT, assets are more integrated with the community, and the steward comprises community members.

Monetization of the asset can only work if the community is actively involved in decision-making. This engagement can be functionalized through legal and social mechanisms. Without this crucial system of accountability, stewards should not be able to generate revenue by selling the data or else they will risk replicating the unequal and exploitative systems of existing platforms and will become yet another data holder.

When appraising stewardship through the lens of the commons, it is clear that learnings from these models cannot be applied to all instances. We imagine that a steward that will imbibe some of these principles will be more "public" in nature and in cases where data is clearly a commonly owned goods, there will be aggregating of data sets to address challenges in areas such as urban planning or transportation.

3 Union models

3.1 Why union models for data?

Currently, user data is treated as capital, a by-product of the current data economy. Consumers of data are able to generate profits from the lack of competition for data-suppliers. Some hypothesize that competition may come with the individuals' ability to collectively bargain with technology companies over payments for their contributions. (Ibarra et al., 2017). There is a need to make people aware of their own value so that they demand fairer

compensation from technology companies (Posner, n.d.). Re-thinking the data economy through a *data as labor* framework allows us to imagine a revenue structure for stewardship that can be built upon these rights-based, agency-centered principles.

In addition, in typical structures of data sharing, the burden of consent usually falls on the individual – who also often lacks bargaining power – to understand how their data is acquired and used by technology companies. Stronger mechanisms for data governance can give people the right to stipulate how their data is used, without requiring them to take ownership themselves (Tisne, 2020).

The nature of data is also such that it is more valuable in the aggregate, and negative externalities of how an individual's data is shared may have implications for society (Ruhaak, 2019). Information extracted from individuals can reveal private information about entire groups and communities, highlighting limitations in our current understanding of individual-centered notions of privacy (Reviglio, 2020). Many scholars are increasingly advocating approaches to data governance that involve communities and collective action.

Unions may appear a relic of the past, relevant for factory floors but not in an increasingly digitized workforce. However, worker mobilization in the context of technology, especially among tech workers, is growing (Tarnoff, 2020). There is a realization that worker solidarity, representation and negotiation are more critical than ever. Interestingly, data, its production, safety and use have become a significant part of this movement, and ideas of data justice are fundamental to worker justice. With this backdrop, data stewards, structured to enable collective bargaining, become crucial. Data stewards could resemble modern unions that can apply to thinking about long-term sustainability of similar institutions for data – networks with a wider set of institutions to push reform-centric innovation and involve workers in shaping technological and social systems (Hoerr, 2014).

3.2 Adapting the union revenue structure for data

The revenue structure of a union is simple. Union members contribute dues on a regular basis to a committee. This committee is elected by members of the union and works on behalf of the broader coalition to advocate /draft a contract of issues that are important to the workers of the union ("What is a union?", 2017) The committee meets with the representatives from employment management to negotiate these terms and then comes to an understanding. In this way, workers' interests are safeguarded. By paying union dues, members pool their resources to achieve

tangible benefits such as fair wages and adequate representation for the collective ("International Union", n.d.).

A data steward following similar principles, as we imagine it, would collect annual or quarterly fees from a broader pool of citizens, and advocate how the data should be used – with technology companies, businesses, and societies. Profits would be used to sustain the committee, and also be redistributed to support members. As a result, stewards are able to remain independent, working on behalf of individuals (as dues ensure some liability and a trust relationship between representatives and data owners), while interfacing with third parties to accomplish these necessary duties.

The union model for stewardship, similar to the public goods steward in some sense, is ideal in instances where collectives are coming together to govern data. As with public good stewards, fee structures can appear differently, dependent on the nature of the steward, and require further consideration. The fees-to-entity structure, however, allows individuals to keep the steward accountable to their needs, and ensures the steward is kept independent from other monetary interests. The idea of "membership" is also important here, as it keeps incentives tied to a broader context and allows for decisions to be made as a collective and with community interests in mind.

4 Financial intermediaries

A fundamental promise of data stewardship is that it can unlock value for society, while giving individuals more control over their data. Through this control, people can decide how their data is used, protect their privacy and ensure transparency and accountability of governments and platforms (Cañares, 2020). Data is an asset which needs to be leveraged in the interest of people. Therefore, we now consider the revenue model of investment advisors as a blueprint for data stewardship.

Financial advisors are located between the users and the third parties, managing the funds (financial assets). Advisors manage assets on behalf of users and provide advisory services on how best to optimize the asset. This function and relationship can be extrapolated to data stewards, imagined as "data advisors" for users and working with them to ensure that data is unlocked in the service of users.

4.1 Investment advisors and fee structures

Investment advisors, legally designated financial intermediaries, owe clients undivided loyalty and may not engage in activity that conflicts with a client's interest without the latter's consent. Investment advisors must provide suitable advice to their clients, ensure that there is no conflict of interest and maintain an arm's length between advisory and any other activities (Collins, 2010).

While the fee structures for financial intermediaries can vary, the two most common models are fee-only and commission-based. Commission based advisors earn income from products sold (for example, by selling insurance or mutual funds.) These advisors, usually financial services companies who sell investment products, are incentivized by the number of products sold, making their fiduciary responsibility vague. They also do not have to disclose their conflict of interests. Given the incentives of commission-based advisors are not structured to serve only the interests of clients, this model is not preferred.

The fee-only fiduciary, on the other hand, is paid directly by the client, and not through commissions for selling certain investment or insurance products. Advisors are expected to conduct a thorough review of all investments, disclose conflict of interest, and ensure that all actions are in service of the best interest of the client.

4.1 Structural principles for data stewardship

The model for financial intermediaries makes apparent the intimate link between revenue model and intent. The challenge in applying this model lies in ensuring that data stewards remain true to their purpose of restoring individual agency. The fee-only structure of a legally backed entity such as an investment advisor allows the intermediary to remain accountable.

A fee structure that aligns money with interests needs to be further explored for data stewardship. While some lessons can be drawn from financial intermediaries, this model also provides important warnings for the ways the intermediary-user relationship can become extractive if not carefully considered. However, beyond revenue generation, financial intermediaries provide important lessons for stewardship, such as the value of professionalism in the management of assets such as data. In this way, an individual can receive expert advice and feedback over the management and allocation of their data (as compared to a union model, in which representatives may/may not be professionally trained). This is also helpful in understanding how data stewardship can be made sustainable – through additional value-adding services that are desirable for individuals and companies.

Account Aggregators, the working model for which is currently being developed in India, is a steward that operates as an exchange layer for data, serving as a centralized consent engine. AAs communicate instructions initiated by the user to transfer their data from one fiduciary to another. Though claimed to be an architecture for data empowerment, the revenue structure of AAs is unclear. In order to remain independent, they can consider a fees-only model to align the interests of the user and remain true to the purpose of a data steward.

5 Principles for revenue models of a data steward

From the respective framings of data as commons, labor, and assets, we use this section to pull up principles for revenue models of data stewards. The function and intent of a data steward are inextricably linked to its revenue model. For example, if a data steward is placed to negotiate for better rights on behalf of users, the best revenue structure is likely a fees-to-entity model similar to that of a union.

5.1 Responsibility

A "responsible" revenue structure refers to one that maintains fiscal independence from third-party interests and allows a steward to be in service of individuals/communities. A fee-to-entity structure, for example, allows both public good stewards and financial intermediaries to act in the best interests of stakeholders. We, however, realize the difficulty in creating an independent steward, which often requires moving data outside of public or democratic control to an external entity. As a result, this principle cannot exist in a vacuum and must be tied to others such a legal regulation and accountability.

5.2 Legal regulation

We also believe data stewards should obtain some recognition in law. Many real-world examples of stewardship, such as trust ports in Scotland or investment advisors, are recognized as legal entities (valid or worth of consideration by some external body). It is important to note, in these cases, legal recognition does not necessitate involving state actors in decision-making. Rather, it emphasizes fiduciary responsibility and cements other principles such as accountability. In India, for example, the Personal Data Protection Bill notes a fiduciary responsibility and duty of care towards individual users/data principals. Data stewards in India, such as

Account Aggregators, should also be tied to these principles. There are other models of stewardship, however, that are not legally backed, such as Community Land Trusts. Here, a governance structure that involves varying community and individual interests serves a similar function of endorsement.

5.3 Value distribution

The concept of data stewardship was imagined to benefit communities, society, and the individual; value generated from these models should be distributed accordingly. Trust ports, for example, operate to benefit the community at large, acting on behalf of all stakeholders in the operation. Similarly, the value gained from the operation of a steward should be distributed accordingly to the data suppliers in the ecosystem.

5.4 Accountability

In all three examples, legal and social mechanisms for accountability ensure the steward stays true to their purpose and allows for decision-making systems that involve data suppliers. We believe these mechanisms are a critical design element in the revenue structure of a data steward to avoid replicating existing extractive systems of data commodification.

5.5 Sustainability

A revenue model for a data steward should strive for sustainability, which is rooted in its long-term value proposition. The following can be considered in the design of a steward to drive sustainability:
1. Value adding services: such as additional user protection or professional advice and guidance (as we see in the case of financial intermediaries)
2. Plans for future development: As technological innovations and challenges arise, a sustainable steward should always remain forward-looking, thinking about how to enhance/adapt to the frequently changing environment around data.
3. Diversification of revenue models: as the needs of a steward change, so should its approach to generating revenue

6 Conclusion

Given that data stewardship is an opportunity to overhaul existing data governance practices, there is a need to consider revenue models that do not replicate the current imbalances in the data economy. Stewards need to generate revenue independently and have structures for distributing value amongst the public.

While these models and principles may be imperfect, they do attempt to paint a picture of what a successful revenue model that co-exists with other thoughtful design principles can look like. We must strive to build models of data sharing that can be sustainable, while protecting individual rights. We hope this will serve as a starting point for further research and policy discussion on data stewardship and its design.

Acknowledgements

References

1. Manohar, S., Kapoor, A., & Ramesh, A. (2020, February 5). Understanding data stewardship: Taxonomy and use cases. Retrieved from https://www.aapti.in/blog/data-stewardship-a-taxonomy

2. Hagenbuch, D. (2015, January 16). The 4 Pillars of Ethical Enterprises. Retrieved from https://www.entrepreneur.com/article/240035

3. Dodds, L., Szász, D., Keller, J. R., Snaith, B., & Duarte, S. (2020, April). Designing sustainable data institutions. Retrieved from http://theodi.org/wp-content/uploads/2020/04/OPEN_Designing-sustainable-data-institutions_ODI_2020.pdf

4. Edwards, S. (n.d.). That's another fight. Big data as a commons. Retrieved 2020, from http://www.remourban.eu/News--Events/News/ThatS-Another-Fight-Big-Data-As-A-Commons.kl

5. Panfil, Y., & Hagopian, A. (2019, September 05). A Commons Approach to Data Governance. Retrieved 2020, from https://www.newamerica.org/weekly/commons-approach-to-data-governance/

6. *Modernising Trust Ports [second edition]* [PDF]. (2009, August 10). Gov.uk.

7. Feloni, R. (2019, February 17). Nearly everyone living in Alaska gets about $2,000 a year from the state's $65 billion fund. We asked 9 Alaskans how they spend it. Retrieved from https://www.businessinsider.com/alaskans-spend-permanent-fund-dividend-2019-2

8. The Board of Trustees. (2018, November 09). Retrieved July 28, 2020, from https://apfc.org/the-board-of-trustees/.

9. Delacroix, S., & Lawrence, N. D. (2019). Bottom-up data Trusts: Disturbing the 'one size fits all' approach to data governance. *International Data Privacy Law*. doi:10.1093/idpl/ipz014

10. Semuels, A. (2015, July 06). Affordable Housing, Always. Retrieved from https://www.theatlantic.com/business/archive/2015/07/affordable-housing-always/397637/

11. *Frequently Asked Questions About Community Land Trusts* [PDF]. (2007). Burlington: Burlington Associatesin Community Development LLC.

12. Arrieta Ibarra, I., Goff, L., Jiménez Hernández, D., Lanier, J., & Weyl, E. (2017, December 29). Should We Treat Data as Labor? Moving Beyond 'Free'. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3093683

13. Posner, E., & Weyl, E. G. (n.d.). Data as Labor. Retrieved July 28, 2020, from http://radicalmarkets.com/chapters/data-as-labor/

14. Tisne, M. (2020, April 02). It's time for a Bill of Data Rights. Retrieved from https://www.technologyreview.com/2018/12/14/138615/its-time-for-a-bill-of-data-rights/

15. Ruhaak, A. (2019, February 18). Data protection and individual consent: Why your privacy is about all of us. Retrieved from https://labourlist.org/2019/02/data-protection-and-individual-consent-why-your-privacy-is-about-all-of-us/

16. Reviglio, U., & Alunge, R. (2020). "I Am Datafied Because We Are Datafied": An Ubuntu Perspective on (Relational) Privacy. *Philosophy & Technology*. doi:10.1007/s13347-020-00407-6

17. Tarnoff, B. (2020, May 09). The Making of the Tech Worker Movement. Retrieved from https://logicmag.io/the-making-of-the-tech-worker-movement/full-text/

18. Hoerr, J. (2014, August 01). What Should Unions Do? Retrieved from https://hbr.org/1991/05/what-should-unions-do

19. What is a Union? (2017, November 30). Retrieved July 28, 2020, from https://afscmeatwork.org/union-hall/what-union

20. International Union of Operating Engineers Local No. 955. (n.d.). Retrieved July 28, 2020, from https://www.oe955.com/

21. Cañares, M. (2020, February 03). What do we mean by data empowerment? Retrieved from https://medium.com/data-empowerment/what-do-we-mean-by-data-empowerment-f842ef9880b

22. Collins, J. M. (2010, September). *A Review of Financial Advice Models and the Take-Up of Financial Advice* (Working paper). Retrieved 2020, from Center for Financial Security website: https://www.rand.org/content/dam/rand/www/external/events/2010/11/18/review-financial-advice-models.pdf