The 4M Roadmap

A higher road to profitability by using big data for social good

Brennan Lake
2023

Contents

Executive Summary

Private-sector firms are under increased and opposing pressures to either adopt responsible business practices that benefit society, or to avoid the social arena altogether in favor of an absolutist stance on shareholder primacy. Companies that generate and collect big data are uniquely positioned to mobilize privately held data assets for the public good. However, poorly-executed social-impact initiatives can fall short of lofty goals, and executives may be rightfully concerned about opening access to strategically valuable and privacy-sensitive datasets.

By designing Data-for-Good programs that generate returns for the firm and society alike, managers can de-risk responsible data sharing while improving overall competitiveness. To achieve this, we encourage managers to follow the **4M Roadmap**, a Data-for-Good model that focuses on four key pillars of value-creation: **Mission**, **Messaging**, **Methods**, and **Monetization**. For each of the four M's, we provide in-depth analysis and explore real-world case studies to help managers launch responsible data-sharing programs that create sustained impact. Accordingly, we distill the primary objectives of each of the four M's, while providing key recommendations for successful implementation:

MISSION: Leverage social impact as a competitive differentiator, and a driver of talent acquisition and retention

Recommendations:
- Identify causes that align with organizational competencies
- Engage team members across all functional areas of an organization
- Collaborate with external stakeholders, including clients, suppliers, and regulators

MESSAGING: Increase commercial leads while creating a positive brand narrative

Recommendations:
- Partner with academia, civil society, and public-sector entities to demonstrate how your data is used for public benefit
- Engage data journalists focused on topical use cases relevant to the public interest
- Translate Data-for-Good case studies into lead-generating collateral

METHODS: Foster innovation ecosystems that expand and improve upon existing products.

Recommendations:
- Identify lead users who can innovate upon your dataset
- Validate your data by working with academia and trusted subject matter experts
- Translate customer pain points into research questions that ultimately optimize your data assets and commercial products

MONETIZATION: Expand your total addressable market while generating returns consistent with corporate values.

Recommendations:
- Leverage Data-for-Good outputs as proof points for commercial prospects
- Nurture public-sector and research partnerships to open new verticals that benefit the public and generate revenue

Introduction

As the private sector faces increasing demands from consumers, communities, and governments to address major societal challenges and market failures, businesses are forced to reconsider management practices that seek to reward shareholders at any cost. Instead of viewing this as an obstacle, forward-thinking businesses recognize the potential to improve profitability by taking a broader approach to value creation - one that embraces community development, environmental sustainability, and social equity as drivers of economic growth, rather than externalities to be dealt with.

While businesses have historically promoted views of shareholder primacy, interest groups such as the Business Roundtable have recently issued formal statements on the 'purpose of the corporation' that embrace principles of stakeholder capitalism.[1] More than just a public relations exercise to improve corporate image, a shift towards stakeholder-based value creation can lead to superior returns. Indeed, quantitative research from indexing firms like MSCI and Just Capital have shown that companies with higher environmental, social, and governance performance are associated with higher profitability, lower systemic risk, and market outperformance.[2,3]

Within the big data ecosystem, the "Data-for-Good" movement seeks to unleash the power of big data to solve some of the most critical challenges of our time - from climate change and social equity, to economic development and public health. In response, data-rich firms are increasingly launching public-benefit programs to become more purpose-driven, while improving their reputation in the process. However, even if Corporate Social Responsibility programs nobly align corporate values with social needs, CSR initiatives are often relegated to niche pet projects that do not fully realize the potential for the private sector to serve as a force for good.

There is abundant opportunity to go beyond passive CSR initiatives in the tech industry by proactively exploiting the untapped value that lies within private-sector big data assets. Managers should approach Data-for-Good programs with a deliberate focus on creating shared value[4] for both society and the firm. Such an approach aligns social-impact initiatives with overall business objectives, which in turn improves the likelihood of sustained success over the long term.

Within the broader big-data ecosystem, a focus on shared-value has the added benefit of de-risking social impact programs for those companies that are understandably hesitant to engage in open data-sharing activities. Moreover, in the face of politically motivated criticism of ESG initiatives as "woke capitalism" and virtue signaling, well-executed Data-for-Good initiatives demonstrate that creating measurable impact for society can also create outsized returns for shareholders.

In order to foster the development of sustainable and profitable Data-for-Good programs across the private sector, we propose the adoption of the 4 M's of shared value creation in big data, a roadmap that guides firms to intentionally embed corporate value creation into social-impact initiatives.

Mission — Creating measurable social impact, while instilling a sense of purpose in stakeholders and aligning with corporate values

Messaging — Developing compelling, brand-positive narratives that reinforce corporate messaging, while illustrating the social impact of big data an easily digestible way

Methods — Fostering innovation ecosystems with early adopters and lead users to innovate and add value to big-data products and services

Monetization — Driving financial sustainability and returns that are consistent with the program's values

We illustrate the importance and impact of the 4 M's by exploring real-world case studies of private-sector companies that incorporate public-benefit objectives into their business models, both through standalone Data-for-Good programs and as core components of commercial products and services. Drawing from key stakeholder interviews with Data-for-Good program managers at Microsoft, LinkedIn, Spectus, and Tomorrow IO, our case studies represent a diverse group of companies across multiple industries and varying stages of growth.

We provide key recommendations to guide managers in the development of prosperous Data-for-Good programs. We also provide a brief overview of key governance and technical mechanisms that assist managers in implementing ethical data collection, and responsible data sharing practices. Finally, we offer a customizable decision-support tool to assist Data-for-Good managers in prioritizing public-benefit initiatives based on the value generated across the 4 M's, and in keeping with overall business objectives.

Disclosure: The author is an employee of Spectus, and is head of the Spectus Social Impact Program

The 4M Roadmap

In evaluating the business case for Data-for-Good programs, we identify four distinct yet interrelated drivers of business value: Mission, Messaging, Methods, and Monetization. Here we explore how Data-for-Good programs influence each of the four domain areas, and detail the specific benefits and value each provides to the firm. In the following sections, we dive into each of the four M's and offer case studies and key benefits of each 'M' when put into practice.

The four M's are defined as followed:

MISSION
Definition: An organization's shared purpose centered around creating positive social impact for society through responsible data sharing

Objectives: Create measurable social impact while instilling a sense of purpose among key stakeholders

Key Benefits: Attracting and retaining top talent; differentiation within industry

MESSAGING
Definition: The public goods, content, and brand equity developed through social impact programs

Objectives: Develop compelling, brand-positive narratives that reinforce corporate messaging while broadening relevancy of product

Key Benefits: Earned media; commercial lead generation; brand equity; peer-reviewed case studies

METHODS
Definition: The intellectual property generated by engaging lead users and early adopters

Objectives: Foster innovation ecosystems with lead users to innovate on product offering

Key Benefits: Low-to-no cost R&D; data and product validation

MONETIZATION
Definition: Financial returns from converting social impact use cases into commercial opportunities

Objective: Generate returns consistent with corporate values while expanding customer base

Key Benefits: Top-line revenue; increased Serviceable Obtainable Market

The 4M Roadmap can be used at multiple stages of a program's development. Firms interested in launching a Data-for-Good program may be reasonably concerned about potential risks of sharing data in public venues. By using this roadmap, managers can better evaluate a program's potential costs and benefits as compared to business as usual. Likewise, managers of existing Data-for-Good programs - who may feel pressure in the face of cost cutting, leadership change, and M&A events - can use the roadmap to optimize their programs, and articulate benefits to key stakeholders.

In order to quantify the value created across each of the 4 M's, managers can apply a basic value perception formula: **value = perceived benefits / cost**. Compared to business as usual, a well managed program should result in increased perceived benefits in the numerator, and a predictable fixed cost in the denominator (i.e. Data-for-Good managers' salaries and general expenses). The following table elucidates the incremental value of Data-for-Good programs across the functional areas of an organization:

COST BENEFIT OF DATA FOR GOOD PROGRAMS V.S. BUSINESS AS USUAL				
Value	**Functional Areas**	**Traditional cost inputs**	**Data for Good Program cost inputs**	**Data for Good Program Benefits**
Mission	Human Resources	**Variable recruiting costs**		**Recruitment incentive & retention**
Messaging	Marketing & Public Relations	**Variable pay-per-click advertising cost; fixed PR retainer for media placements**		**Earned media & inbound lead generation**
Methods	Research & Development	**Fixed Data Science costs; Variable PhD sponsorship costs**	Fixed D4G OPEX (personnel)	**Outsourced R&D and no-cost PhD sponsorship**
Monetization	Business Development & Sales	**Fixed outbound prospecting and BDR costs; variable internal data validation costs for RFP**		**Increased conversion rates, RFP assistance, direct program revenue**

Mission

"Employees expect their jobs to bring a significant sense of purpose to their lives. Employers need to help meet this need, or be prepared to lose talent to companies that will." **-Naina Dhingra, McKinsey & Company**

Among the 4 M's, Mission is perhaps the most ubiquitous and prominently articulated raison d'etre for corporate Data-for-Good programs. This comes as little surprise since they are, by definition, mission-driven programs designed to confront societal challenges by responsibly harnessing the power of big data. Whether from internal inspiration or external requests for data, many companies have come to recognize that their data assets hold tremendous untapped social value. In turn, companies create Data-for-Good programs. However, firms must go beyond simply espousing aspirational notions of creating social impact within a program's mission statement. Rather, they should proactively embed public-benefit objectives - and strategies for achieving them - into their core business planning processes, while also considering how the fulfillment of mission-based objectives can benefit the firm.

A cynical view may question the motives of setting mission-oriented goals that intentionally benefit the firm. A more pragmatic outlook, however, recognizes that pursuing mission and business-oriented goals in tandem can lead to broadened and sustained social impact. This is especially true when there is support for - and engagement with - Data-for-Good initiatives from the bottom up. Indeed, engaging stakeholders across the organization in the design, planning, and implementation of Data-for-Good initiatives can lead to better outcomes both in terms of mission fulfillment and in enterprise value through improved employee recruitment, retention, satisfaction, and productivity. Moreover, a strong Mission-oriented program will ultimately benefit multiple teams within an organization, from HR and marketing to legal and public-affairs departments.

Talent Acquisition and Retention

As of the time of writing, the jobs site Glassdoor lists 'Data Scientist' as the third-ranked job in the United States in terms of demand, eclipsed only by 'Front End Engineer' and 'Java Developer' - all highly relevant positions within Big Data organizations.[5] To attract and retain top talent in an ever-competitive recruitment environment, firms would do well to create opportunities for engagement with truly purpose-driven work.

According to a statistically significant 2016 survey of 1,020 professionals ages 20+, 76% of those surveyed strongly consider a company's social responsibility commitments when deciding whether to accept an employment offer, and 64% would reject an offer outright from a firm without a compelling CSR focus.[6] 75% of those interviewed even expressed a willingness to take a pay cut to transition to a more socially responsible company.[7] Employers can expect to see improved retention by placing higher importance on social-impact initiatives, with 83% of those surveyed saying they would be more loyal to companies that facilitate opportunities to contribute to such initiatives.[8]

A cross-company effort

In order to enjoy benefits from improved recruitment and employee loyalty, organizations must engage in participatory practices when designing and implementing Data-for-Good programs. An executive may have an inspired idea for a given initiative, but without critical input and support from employees, Data-for-Good programs run the risk of becoming du jour pet-projects and fizzling out.

The data on employee sentiment supports this notion, with 89% of survey respondents expressing a desire to be involved in public-benefit programs by providing ideas, potential solutions, and feedback.[9] Moreover, 88% of respondents think it is important for employers to transparently share social-impact goals, progress, and achievements as they occur.[10] And while big data firms may choose from any number of CSR initiatives - like donation matching - employee sentiment favors corporate-led programs, with 83% favorability as compared to sporadic activities, like service trips or after-hours service opportunities.[11]

Data-for-Good programs are well-positioned to deliver value on these fronts, since they require organizational participation across a broad range of functional areas and skill sets. To ensure the data and insights that a program shares are relevant and actionable, the input of data scientists, engineers, and technical operations managers is of paramount importance. To incorporate responsible and ethical data sharing principles into programs, deep analysis and solutioning is required from legal and privacy departments. The sourcing of high-quality project and partnership opportunities requires business development and relationship management skills. Finally, to communicate results, marketing and communications teams must be engaged. In short, Data-for-Good programs are all-hands affairs that leverage the same technical, data, and human infrastructures required by commercial products. Accordingly, Data-for-Good programs are more naturally aligned to a firm's core business practices than many one-off CSR projects.

Multiple teams beyond Human Resources will also benefit from the firm's improved reputation and brand value that result from a well-executed Mission. As detailed in the subsequent sections of this report, Marketing and Sales teams benefit from a wealth of collateral and proof points to support the pre-sales process; Data Science and Engineering teams benefit from engagement with subject matter experts; and Legal and Public Affairs teams will be better positioned to demonstrate the benefits of responsible data sharing to the public sector and special-interests groups.

Case Study
Microsoft AI For Good

Overview

Building upon its corporate mission to *"to empower every person and every organization on the planet to achieve more"*, Microsoft's AI for Good program provides nonprofits, research institutions, and scientists with Microsoft AI and Machine Learning tools to tackle some of the world's largest challenges - from climate change and health inequities to digital skills development and broadband access.

Microsoft's AI for Good program engages teams from all functional areas across the organization, empowering AI for Good specialists to leverage their skills towards purpose-driven work. Additionally, by collaborating with leading research institutions and public and private organizations who utilize Microsoft's AI, Big Data, and Machine Learning products, Microsoft facilitates real-world applications and use cases for its solutions, leading to downstream innovation and upstream generation of Methods and product improvements.

In order to learn more about Microsoft's AI for Good program, a key-stakeholder interview was conducted with Microsoft's Kevin White, Senior Director, Program Management and Data Analytics, AI for Good.

The program

Microsoft's AI for Good program works directly with impact-oriented organizations through partnerships and grant-based engagements. These partnerships with NGOs, research institutions, and governments apply Microsoft's AI technologies across five main areas including: AI for Earth, AI for Health, AI for Accessibility, AI for Humanitarian Action, and AI for Cultural Heritage.[12] Microsoft's CSR and giving programs (including AI for Good) are largely run in-house by Microsoft Philanthropies. Backed by a financial commitment from the company, the program has full autonomy over the projects it pursues and the organizations with which it partners.

Leprosy Intelligent Image Atlas

One such project, launched under its AI for Health initiative, utilizes Microsoft's machine learning capabilities to scan anonymized dermatological images to identify early signs of leprosy (Hansen's Disease).[13] Developed in partnership with the Novartis Foundation and the Oswaldo Cruz Foundation in Brazil, the Leprosy Intelligent Image Atlas seeks to curb the nearly 200,000 annual diagnoses of Leprosy, which often occur in the global South, primarily in Brazil, India, and Indonesia.[14] Caused by a slow-growing bacteria, Leprosy is curable in its early stages, thereby preventing the disease's impairing later stage effects.

By training a Machine Learning model with images of skin exhibiting lesions caused by leprosy, the Leprosy Intelligent Image Atlas can predict the probability that a given dermatological image displays signs of the disease's early stages. In driving the deployment of this technology, Novartis and the Oswaldo Cruz Foundation are providing Hansen's Disease patients with a convenient and discreet way to be tested remotely, without the need to travel long distances to a clinic. In addition to supporting these two foundations and its constituents, Microsoft has also open-sourced the machine-learning algorithms that power the Atlas by making the code openly available on GitHub.[15] Of important note from a governance perspective, Microsoft makes clear that its open tools are intended for research purposes only, and are not to be used for medical decision making until clinical performance becomes better established.

Microsoft SmartNoise & Open Differential Privacy

As Microsoft employs AI to tackle the world's greatest challenges, they recognize the importance of preserving the privacy of users who entrust their data to others. In order to minimize privacy risks, Microsoft and academic researchers helped pioneer the development of Differential Privacy (DP) - the application of statistical noise to datasets to obfuscate individual data points without altering the overall accuracy of aggregate insights.[16]

In addition to applying its DP innovations to its own products, Microsoft created Smartnoise, an open-source DP toolkit developed in partnership with Harvard's OpenDP initiative. Going a step further than simply open-sourcing the methodology, Microsoft also launched an early-adopter accelerator program to foster the adoption of DP for social-impact use cases within the public, private and nonprofit sectors. As a result, the program's participants are now applying SmartNoise to several impactful use cases. For example, the Urban Institute is pioneering the use of DP to enable greater access to otherwise restricted public-sector datasets for use in urban development.[17] Another participant, Crisis Ready, is applying DP to geospatial datasets in order to provide emergency managers with privacy-preserving insights on human mobility during natural disasters and pandemics.[18]

Creating shared value:

As with most corporate social responsibility programs, the purpose of Microsoft's AI for Good program is to mobilize company resources towards its Mission of creating positive social impact. In turn, the AI for Good program creates tremendous value for Microsoft by creating an environment conducive for purpose-driven employee engagement, and the innovation of new Methods.

Microsoft AI for Good and related programs have permanent staff, and they also attract participation from Microsoft employees across a multitude of disciplines who benefit from engaging in socially impactful work. Additionally, the AI for Good program creates quantifiable value through the vast codebase generated by the program, and the inherently translational nature of machine learning algorithms with both social and commercial applications.[19] For instance, while the SmartNoise Early Adopter Accelerator program was initially launched to enable widespread data-sharing for social good without sacrificing user privacy, Microsoft in parallel benefits from collecting proof points on how their approach to Differential Privacy stands up to multiple real-world applications. This provides a virtuous feedback loop which can better optimize how DP is applied internally to products that already utilize differential privacy, such as Windows, LinkedIn, and Office.[20]

In articulating the Mission and Methods created through the AI for Good program, it is noteworthy to mention that there is an agnostic stance towards the potential Messaging and Monetization value that may be generated directly or indirectly through the program. Most importantly to Microsoft, provisioning access to broadband internet leads to more equitable access to educational resources, financial inclusion, and workforce opportunities.

Likewise, in considering the potential for the AI for Good program to generate brand-positive Messaging value, the Microsoft team noted that its initiatives are not designed to dominate the company's communications and PR cycles. Indeed, as noted by the team, when there is Messaging generated through the program, its greatest value is derived through internal communications and presentations, as an opportunity to inspire and engage Microsoft and employees and future AI for Good volunteers.

Messaging

"Storytelling allows readers to connect to key findings at a deeper level by understanding where data came from, why it is important, and what visualizations indicate, even if the readers lack a true grasp of data literacy." **-Brady D. Lund, The International Journal of Information**

When executed properly, Data-for-Good programs create positive brand narratives for technology firms. By effectively communicating the positive impacts achieved through altruistic applications of big data, purpose-driven firms can naturally articulate their Mission to broad audiences through relatable use cases. In turn, effective Messaging can serve both as an engine for lead generation, while also de-risking the adoption of data for prospective commercial clients who may be less familiar with business applications for new and innovative data sources.

When executed in the absence of a clear Mission and concrete programs designed to achieve social-impact goals, however, Corporate Social Responsibility PR attempts will ring hollow as mere virtue signaling, and may even lead to cynicism among the firm's stakeholders and the general public. By rooting Messaging firmly within the context of how specific Data-for-Good projects are delivering on a company's Mission, firms can authentically demonstrate how they create Shared Value.

Crafting a positive brand narrative
Beneath the buzzwords, the big data assets that power machine learning and artificial intelligence applications can be esoteric and uninspiring for the general public. In order to turn otherwise obscure datasets into the foundation for a compelling narrative, firms should identify and pursue relevant social-impact use cases for public-facing Data-for-Good projects. For example, MasterCard Advisors' use of aggregated transaction data-

to inform hedge fund investment decisions, while alluring for asset managers, is not relatable or meaningful for the general public.[21] On the other hand, MasterCard's Center for Inclusive Growth uses the exact same transaction data to tell impactful stories about economic development and social equity.

In one study, the Center for Inclusive growth examined aggregated transaction data of low-income communities in Chicago, revealing that residents often had to travel long distances to purchase everyday items, placing costly time and travel burdens on already disadvantaged communities.[22] In another study, researchers leveraged MasterCard data to identify a correlation between the growth of low-level crime and the closure of small businesses in Baltimore and Oakland.[23] By using big data for emotionally compelling and socially beneficial use cases, MasterCard is able to contextualize its data in human terms, while sharing stories that reinforce positive narratives about its brand and CSR efforts.

Aligning Cause with Competency

In order for Data-for-Good programs to create effective brand narratives, it is critical that a firm's Mission is closely aligned with its core competencies. As articulated by Knowles et. al in the 2022 Harvard Business Review piece, *"What is the Purpose of your Purpose"*, a cause-competency gap *"occurs when the connection between the nature of your business and your espoused cause is not obvious—a danger for even highly successful companies. For instance, a difficulty currently facing the platforms Facebook and Google is that their advertising-driven business models are perceived to be increasingly at odds with their stated missions: 'to build community and bring the world closer together' and 'to organize the world's information and make it universally accessible and useful,' respectively."*[24] By aligning a Data-for-Good program's Mission with its core competencies, firms can not only better deliver on public-benefit objectives, but they can also positively reinforce messaging around the firm's defining strengths.

LinkedIn's Economic Graph program offers an instructive example of effective alignment between cause and competency. With a vision to *"create economic opportunity for every member of the global workforce"*, LinkedIn's Economic Graph lays out an ambitious yet actionable vision that is aligned with its core competencies of matching employers' demand for talent with relevant workforce candidates.[25] The Economic Graph program realizes this vision by making its data openly available to policymakers, multilateral organizations, and civil society groups, providing decision makers with a holistic view of metrics such as workforce confidence, labor flows, and skills demand.

As a member of the Development Data Partnership, a public-private-partnership launched by the World Bank, LinkedIn provides vital insights for policymakers looking to modernize their workforce. For example, a labor bureau from a developing nation can request access to LinkedIn's Economic Graph to determine which skills are most in-demand within their labor market, and in turn initiate workforce-development programs to educate for those skills.

In addition to creating measurable impact, the Economic Graph program provides LinkedIn with an appealing narrative around the reuse of commercial data for the public good. Such a narrative can not only create a favorable impression with the general public through marketing materials, but it also serves as an opportunity to educate the public sector on the value of responsible data collection and reuse.

Engaging in Data Journalism

Companies in the big-data ecosystem are naturally adept at telling stories with data. Indeed, their bottom lines are often dependent on the ability to effectively collect, structure, and visualize data to inform business decision-making. Similarly, as traditional print media organizations shift to online-first content creation, they face growing competitive pressure to present information through captivating and interactive media.

The emergence of data journalism is a growing practice that marries data supply with the media's demand for rich, data-driven visualizations and storytelling. Major publications like the New York Times, Wall Street Journal, and The Guardian, as well as smaller organizations like FiveThirtyEight and Vox, have dedicated Graphics teams for creating interactive data journalism media.

Firms can gain significant PR exposure through mutually beneficial data partnerships with media organizations. For example, Biobot - a company that collects data on the presence of viral pathogens in wastewater - amassed a large following during the pandemic by sharing its COVID-19 early-warning data via public dashboards. Media outlets such as the Wall Street Journal[26] and the New York Times[27] released multiple features complete with data visualizations illustrating the effectiveness of wastewater monitoring as an early warning system during public health emergencies. In addition to being named one of Time Magazine's 100 Most Influential Companies of 2022,[28] Biobot gained significant public and private sector contracts, demonstrating that open data sharing does not necessarily lead to cannibalization of business opportunities.

Case Study
Cuebiq & Spectus.ai

Overview

Spectus is a geospatial data company that develops privacy-enhancing technologies for the analysis of human mobility data. Through its Data Cleanroom product, Spectus offers a Platform-as-a-Service model that allows data scientists to query granular mobility data on Spectus' servers, while receiving aggregate statistics in return. Along with its parent company, Cuebiq, Spectus supports commercial use cases related to real-estate site planning, transportation, finance, and advertising measurement.

Since its founding in 2016, Cuebiq has dedicated an entire business unit to creating societal value through the novel and ethical use of mobility data. Originally conceived as a data philanthropy program, Cuebiq's Data-for-Good initiative evolved into a core component of the company's value proposition. With the launch of the company's new business unit, Spectus rebranded its Data-for-Good program as "Spectus Social Impact", with a focus on creating value for society, while also serving as an engine of growth for the company through research & development, public relations, and lead generation.

The Program

The Spectus Social Impact program democratizes access to otherwise costly mobility data insights for use in academic research and public policy, with a focus on five primary topic areas:

Disclosure: The author is an employee of Cuebiq, and head of the Spectus Social Impact Program

14

- Equitable Urban Development
- Disaster Response
- International Development
- Public Health
- Geospatial Data & Privacy Innovation

Spectus has developed a hybrid approach to data philanthropy. The Social Impact program provides pro-bono aggregate datasets for qualified academic research and humanitarian initiatives, while providing at-cost access to its Data Cleanroom platform, which offers researchers the opportunity to access more granular data within a strictly controlled, on-premise environment. Additionally, Spectus allocates five-to-ten pro-bono accounts for its Data Cleanroom platform for organizations that demonstrate the potential for creating measurable public benefit, and developing open-source solutions for the geospatial community.[29]

The World Bank's Disaster Management MobilKit
The World Bank's Global Facility for Disaster Reduction and Recovery (GFDRR) is an international-development program designed to assist developing countries better understand, prepare for, and recover from natural disasters.[30] When natural disasters like hurricanes and earthquakes strike, policy makers require information to understand which communities are most affected. By utilizing smartphone-derived human mobility data, decision makers can empirically measure what percentage of a community has evacuated, where people evacuated to, and how long communities are displaced. The World Bank's GFDRR team was granted pro bono access to Spectus data to analyze such evacuation behaviors in the aftermath of the Puebla earthquake, which struck central Mexico in 2017. Based on their analysis, researchers were able to offer guidance on building community resilience, while also developing MobilKit, an open source toolkit for rapid analysis of human mobility data for natural disasters.[31] Building upon this work, the World Bank was also able to leverage Spectus data for global analyses on the effects of Non Pharmaceutical Interventions - like social distancing - during the early stages of the COVID-19 pandemic.[32]

Measuring Effectiveness of The Ad Council's Vaccination Campaigns
The federal government has invested billions of dollars to make COVID-19 vaccines accessible to the general public throughout the pandemic. In order to overcome vaccine hesitancy, the Department of Health and Human Services funded a massive public-service announcement (PSA) campaign through the Ad Council, the US's largest nonprofit advertising agency.[33] However, once the Ad Council's "It's up to you" campaign launched, there was a lack of data on whether the campaign was actually driving visits to vaccination centers. In partnership with Johns Hopkins University, the Ad Council, and Bloomberg Philanthropies, Spectus launched a program to evaluate the effectiveness of vaccination PSAs by measuring visits to single-purpose vaccination sites among cohorts of devices that viewed the Ad Council's campaign.[34] Researchers at JHU found that the campaign was ultimately effective at driving visits to vaccination counties across most counties in the US.

Creating Shared Value
Spectus creates Mission value externally through both the production of public goods (i.e. publications and open-source datasets), as well as through its active role in bridging the gap between research and policy. By actively engaging policy makers across NGOs, multilaterals, and the public sector, Spectus ensures that research outputs are disseminated to decision makers with the capability to take action upon research findings. The program also generates Mission value internally by engaging multiple functional areas across the organization, such as data science, engineering, marketing, privacy, and legal teams. In its recruitment practices, Spectus' HR team cites the Social Impact program as a key driver of interest among top talent.

With "Geospatial Data & Privacy Innovation" as one of its five key focus areas, Spectus Social Impact intentionally embeds the generation of Methods value into the program's overall strategy. As a result, Spectus fosters an innovation ecosystem where methodologies and algorithms developed by academic research clients are now being reused not only by other social-impact clients, but commercial clients as well. Given the company's relatively small size (roughly 100 employees globally), the engagement of lead users is an effective way to augment internal R&D capacity while also engaging world-class data scientists with subject matter expertise across a number of geospatial-related domains.

Spectus is also able to translate the public goods produced through its engagement with researchers into Messaging value. The inclusion of Spectus data in academic publications and data journalism articles not only contributes to inbound lead generation, but it also assists in the validation of data quality and its applicability to novel use cases. As demonstrated in subsequent sections, such steps to validate and demonstrate the utility of a firm's data can directly support Monetization efforts.

The latter is critical in attracting business from commercial clients who otherwise may be hesitant to use newer 'alternative' datasets in business decision making. Indeed, by independently verifying data quality in peer-reviewed publications, and demonstrating the applicability of alternative data in the mainstream media, data providers like Spectus are able to de-risk the data acquisition process for new business users.

Methods

"Even with the best in-house talent and innovation process, companies cannot come up with the best ideas on their own. That's why a complete innovation ecosystem requires collaboration with external innovators" -Alexander Osterwalder, Harvard Business Review

For big data firms to remain competitive, they must continually innovate and improve upon Methods for collecting, processing, and applying data in ways that meet ever-evolving demand. Large corporations may have the resources to invest in internal innovation and R&D teams or to higher external consultants, but smaller organizations may not have the luxury of diverting focus from sales-driven development roadmaps to focus on R&D. No matter the size or resources at hand, organizations can benefit from embedding innovation objectives into their Data-for-Good programs. By fostering innovation ecosystems centered around Data-for-Good initiatives, firms can benefit from a wealth of new Methods and proof points provided by external lead users and innovators.

In the Oxford Handbook of Innovation management, Autio and Thomas define Innovation Ecosystem as "a network of interconnected organizations, organized around a focal firm or a platform, incorporating both production and use-side participants, and focusing on the development of new value through innovation."[35] Data-for-Good programs naturally follow this model, wherein big-data firms form collaborative structures with NGOs, academia, and the public sector centered around privately held big-data assets (the 'focal platform') to create value for all stakeholders and society at large.

Researchers at New York University's Governance Lab ('GovLab') have developed an extensive typology and community of practice around data-centric innovation ecosystems, termed 'Data Collaboratives' by the GovLab.[36] According to the GovLab, Data Collaboratives convene siloed data holders and subject matter experts from various disciplines, thereby bridging supply gaps and focusing data analysis towards the development of innovative solutions to social challenges.[37] Among the various structures detailed in the GovLab's Typology (see below), Methods which enable active collaboration between data suppliers and users - Data Pooling, Research & Analysis Partnerships, and Prizes & Challenges - are especially conducive to creating upstream value for data suppliers.[38]

	Open Access	Restricted Access
Independent Use	**Public Interfaces**	**Trusted Intermediary**
Cooperative Use	**Data Pooling**	**Research and Analysis Partnership**
Directed Use	**Prizes & Challenges**	**Intelligence Generation**

NYU GovLab's Data Collaboratives Matrix[39]

Collaborative Methods Generation

Data science Prizes & Challenges help firms crowdsource novel approaches, algorithms, and Methods for solving entrenched challenges by providing citizen data scientists with incentives whether monetary, purpose-driven, or both. DataDriven, a nonprofit organization, provides a fitting example as it hosts data-science competitions that simultaneously create measurable impact.[40] By working with sponsors such as Mathworks, Microsoft, Facebook, and others, DrivenData provides participants with challenge statements, along with the rich datasets and tools to solve them. In turn, corporate sponsors not only help tackle societal challenges, but they can also gain new Methods centered around their technologies. Businesses can then internally apply such learnings for social and business applications, alike.

Facebook has sponsored multiple DrivenData competitions designed to reduce the spread of misinformation and hateful content through social media - a perennial challenge for social-media platforms. In one competition, Facebook provided participants with imagery data consisting of over 1 million original and manipulated copies of images.[41] Participants then use this data to build models that identify manipulated copies, with the end goal of automating the content moderation process, and making social media safer and more trustworthy platforms for users.

In another DrivenData competition sponsored by Facebook, titled "Hateful Memes", users are provided with memes in the form of images and strings of text.[42] Users must build a model for predicting whether a given meme is hateful or not. In both cases, Facebook benefits from crowdsourcing solutions - in the form of machine-learning algorithms - to major challenges that threaten not only Facebook, but also society at large: the spreading of misinformation and hateful content.

Data-for-Good programs and data science competitions do not necessarily have to be run by 'traditional' big-data companies. Indeed, as firms adopt digital transformation and automation strategies, they produce troves of big-data assets in the process. Take Schneider Electric, for example, a French multinational that develops software-

Additional resources: Explore the GovLab's 9R's Worksheet for establishing the business case for data collaboration and re-using data in the public interest. https://businesscase.opendatapolicylab.org/

solutions for improving energy efficiency and sustainability. By sponsoring five DrivenData competitions, Schneider Electric was able to incentivize citizen data scientists to build algorithms that forecast energy consumption, detect anomalies in usage, and optimize demand-side strategies.[43] In addition to benefiting from these Methods by incorporating them into their own product suite, Schneider Electric open-sources all of the winning algorithms in order to drive innovation and sustainability in the energy sector.

Much like the Prizes & Competition model, Research & Analysis Partnerships also drive interaction between data suppliers and users, but with much closer engagement. Rather than crowdsourcing solutions from citizen scientists, Research & Analysis Partnerships bring data suppliers together with leading subject matter experts in a given field to contribute to scholarly literature, and to directly support civil-society programs and public-sector policymaking. This can often be a sequential process. Once partnerships with academia generate peer-reviewed Methods, public-sector and civil-society decision makers can more confidently adopt such Methods and datasets to inform programs and policy. In turn, firms benefit from having their data sources vetted and validated, which can shorten sales cycles with discerning commercial prospects.

As an apt example of this process, take Cuebiq's response to the COVID-19 pandemic. Since 2017 Cuebiq (the parent company of Spectus) has launched bilateral data collaboratives with academic researchers through its Data-for-Good program. Once the COVID-19 pandemic emerged as a global health threat, it convened a data collaboration with 15 of its closest research partners from institutions including Oxford University, Johns Hopkins, Northeastern, Columbia, and others.[44] Equipped with a granular yet privacy-preserving mobility dataset, along with collaborative platforms such as Slack, the researchers were able to co-design basic metrics for understanding the effect of the pandemic and non-pharmaceutical interventions on human behavior. Once the Methods for generating these metrics were published in peer-reviewed journals, public sector entities including CDC, the City of New York, and the State of Rhode Island could make more informed decisions about how to utilize aggregated mobility data as an additional resource in their toolkit for managing the pandemic and economic recovery.

Lead User Design

When Data-for-Good programs foster innovation ecosystems with external stakeholders they may benefit from lead-user engagement. Lead users are early adopters whose particular needs force them to innovate or adapt solutions far in advance of the general market.[45] NASA provides abundant examples of the lead-user phenomenon, where the need for innovations that could withstand the extreme conditions of outer space led to the development of technologies that are ubiquitous today, such as the portable computer, wireless headphones, and camera phones.[46] Going beyond passive observation of this phenomenon, MIT professor Eric von Hippel developed the Lead User Methodology to assist product developers in actively identifying lead users, and engaging them in the design process.[47]

Data-for-Good programs are ripe environments for engaging Lead Users. With roughly 60% of the world's data produced in the past three years,[48] the market for the use and reuse of alternative datasets is relatively nascent. Certain industries like advertising and finance have been quick to adopt and incorporate big-data into their decision making processes. Others like agriculture, architecture, engineering, construction, and the public sector have lagged behind.[49] By developing Research & Analysis partnerships with leading subject matter experts who also possess advanced data-science competencies, firms can effectively practice Lead User design and kickstart the uptake of big data within sectors where it is currently underutilized.

Microsoft's AI for Good program is particularly effective at identifying, engaging, and supporting lead users in the innovation process. Take the SmartNoise accelerator program, for example, which actively recruits organizations with an acute need for utilizing potentially privacy-sensitive big-data for social good. The Urban Institute, for instance, is an exemplary lead user, whose leading research in social and economic policy greatly benefits from public-sector data, which is often difficult to access. With technological support and mentorship from Microsoft, the Urban Institute was able to develop a prototype for privacy-secure querying of confidential federal-government data.[50] In open-sourcing this technology, the Urban Institute and Microsoft's work stands to benefit both civil society and the private sector alike.

The Lead User Methodology is also helpful for stress-testing products in extreme environments, which are often hotbeds for innovation driven by necessity. In partnership with OpenStreetMaps, Microsoft's AI Assistant software was used to map informal settlements and buildings in Uganda and Tanzania where, like in many emergent nations, public census and mapping fall behind actual growth and development.[51] This is particularly problematic in the wake of natural disasters, when emergency managers struggle to deliver aid in unmapped areas.[52] By combining OpenStreetMap's Tasking Manager technology with Microsoft's Computer Vision AI, the initiative was able to map over 18 million previously uncharted buildings.[53] Since Bing Maps was originally developed in the United States and Canada, it faced new challenges in mapping informal settlements with different structural footprints, and a paucity of training data. However, stress testing the product in developing regions ultimately led to more robust mapping methodologies and datasets, which are now openly available on GitHub.[54]

Case Study
Tomorrow.io

Overview

Tomorrow.io is a software-as-a-service platform that helps governments and private-sector companies navigate challenges related to climate and extreme weather events.[55] By developing innovative ways to collect and deliver hyper-localized weather data, Tomorrow.io allows users to analyze past weather patterns and forecast future conditions in order to optimize operations and decision making across multiple industries, such as Agriculture, Transportation, and Supply Chain management.

With a vision of becoming a truly global company, Tomorrow.io is committed to ensuring that it does not leave emerging and developing markets behind. To realize this vision, Tomorrow.io has spun off an independent yet closely aligned nonprofit organization, TomorrowNow.org, which leverages the same core technology and weather data to help vulnerable populations build resilience to climate change.[56] By adapting its data and core technologies to serve users in extreme and underserved environments, Tomorrow.io and TomorrowNow.org collectively create a virtuous cycle of product innovation, while also laying the groundwork for future growth in emerging markets.

In order to learn more about the program, a key stakeholder interview was conducted with Georgina Campbell Flatter, Executive Director of TomorrowNow.org.

The program

As the pace of climate change accelerates, governments and private sector entities incur increasing costs from climate and weather related losses, with the US alone on track to spend up to 10.5% of GDP on climate change related costs by the end of this century.[57] Tomorrow.io seeks to provide the public and private sectors with situational awareness and operational intelligence in the form of hyper-localized weather data and forecasting models in order to better plan for and respond to climate and weather-related challenges.

While innovations in the weather space have traditionally been led by new scientific discoveries, Tomorrow.io also incorporates user-centered design to create solutions in response to sector-specific pain points. This upstream innovation process includes investing in observation - like satellite, radar, micro links and cell tower signaling - while also developing new modeling to forecast risk related to flood, fires, drought, and other disasters. By packaging granular weather data and forecasting within a context and industry-specific SaaS platform, Tomorrow.io is able to service a diverse customer base, from Uber and Ford to JetBlue and the US Air Force.

With an acute awareness that those who contribute the least amount of greenhouse gas emissions are also the most vulnerable to the effects of climate change, Tomorrow.io is dedicated to ensuring that its products and data can serve those whose needs are most urgent. To further this Mission, it spun off an independent 501(c)(3) nonprofit organization, TomorrowNow.org, whose sole focus is bringing user-centric weather intelligence solutions to partner communities, governments, and civil society organizations.[58] With grant funding from philanthropies like the Gates Foundation, and free software licensing and skills-based volunteering from Tomorrow.io corporate, TomorrowNow.org provides weather intelligence to stakeholders in East Africa to drive agricultural decision making and to support natural-disaster preparedness and response.

Project Locust

Locusts present one the most critical threats to livelihoods and food security across the African continent. With weather events such as heavy rainfall and wind determining the size, severity and directionality of locust swarms, access to timely weather data and forecasting can aid local stakeholders in limiting destruction to crops.[59] However, accurate and real-time weather data is often too coarse or simply nonexistent in many agricultural communities in East Africa. In response to this challenge, TomorrowNow.org launched Project Locust - a locust monitoring tool built with technology provided by Tomorrow.io, and developed through a process of participatory design.[60]

With funding from the Gates Foundation, TomorrowNow.org runs workshops with local farmers in East Africa to better understand what information is most vital to drive situational awareness, preparedness and response in the face of locust threats. By partnering with Digital Green, an NGO that develops technology solutions for smallholder farmers, the Locust Project aims to develop an early warning system to provide automated alerts, customized recommendations, and a channel for ground-truth validation and feedback on forecast accuracy. Delivered via the Tomorrow.io dashboard or SMS text messages, early warnings can be hyper-localized and recommendations can be context aware, such as guiding stakeholders to harvest or spray crops before the forecasted time of a swarm's arrival.

Creating Shared Value

While TomorrowNow.org's sole purpose is to create positive social impact, its symbiotic relationship with Tomorrow.io generates both social and business value through a virtuous cycle of innovation, product feedback, and employee engagement. As a result, Tomorrow.io benefits from value across each of the 4 M's.

First, Tomorrow.io derives Mission-driven value by virtue of its relationship with TomorrowNow.org. Tomorrow.io employees have opportunities to directly support TomorrowNow.org's initiatives like Project Locust through skills-based volunteering during working hours. Even those employees who do not directly engage with volunteerism can find purpose in developing the underlying technology that drives climate resilience for vulnerable communities.

Second, Tomorrow.io also benefits from product feedback and proof points produced by deploying the technology in novel ways within new markets. Indeed, by supporting farmers whose livelihoods and food security depend on the ability to adapt to mercurial weather conditions, TomorrowNow.org is effectively engaging in Lead-User Design processes. By receiving feedback on what works, and where the technology and data fall short, Tomorrow.io also gains invaluable insights into the performance and applicability of their product under extreme conditions.

Third, Tomorrow.io benefits from strong positioning as an innovative solution to the truly existential threat of climate change. A cynical view may hold that the company's commercial clients in the agricultural and transportation industries are also contributors to the underlying causes of climate change. However, by licensing its technology for free and offering pro bono expertise towards initiatives that support those communities most vulnerable to climate change, Tomorrow.io creates brand-positive Messaging that elevates the company above such critiques. Moreover, its decision to spinoff a wholly independent non-profit organization further insulates their public benefit work from corporate decisions.

Finally, whether intentional or not, Tomorrow.io's focus on driving adoption of its products in emerging and underserved markets may ultimately benefit the company's bottom line in the long term. By licensing its technology and providing pro bono services to TomorrowNow.org - whose programs are then implemented with grant-backed funding - Tomorrow.io is reducing market risk and fostering a community of potential consumers within emerging markets, where it may consider doing business in the future.

Monetization

"Shared value focuses companies on the right kind of profits - profits that create social benefits rather than diminish them" -**Mark Kramer and Michael Porter, FSG**

Data-for-Good programs that create Mission, Messaging, and Methods value are naturally positioned to generate financial returns. Here we summarize the key ways that each of the previous three M's can indirectly support profitability, while providing guidance towards successful implementation. Finally we demonstrate how Data-for-Good projects can directly earn revenue by converting pro-bono projects into paid engagements.

While pursuing monetization strategies it is important for managers to remain focused on their program's overarching Mission and public-benefit objectives. Striking a balance between Monetization and Mission can take various forms. For instance, in cases where it is operationally feasible, Data-for-Good programs may consider providing pro-bono data and insights to under-resourced organizations that create outsized impact. During times of crisis, such as public health emergencies and natural disasters, Data-for-Good programs should be ready to launch initiatives that effectively support the public interest. Companies that act ethically and with integrity will inevitably be well-positioned to improve enterprise value through their conscientious efforts.

Monetization by Design

Mission

A well-executed Mission can increase profitability by lowering HR costs, while also improving organizational competitiveness through differentiation. On the cost side, Data-for-Good programs can help draw top talent while reducing turnover. To achieve this, Data-for-Good managers must engage teams from multiple functional areas in order to create a shared purpose across the organization. In the early stages of a program's development, any employee should have the opportunity to identify relevant causes, use cases, and partners for the budding program. In later stages, there should be well-defined roles for each team to contribute their expertise. To institutionalize such efforts, leadership should consider allowing employees to dedicate a percentage of their time to supporting Data-for-Good initiatives.

On the revenue side, the organization should take each opportunity to include relevant references to its Data-for-Good initiatives in pre-sales processes and Requests-for-proposals. Where applicable, firms should also identify opportunities for commercial clients to partner on specific Data-for-Good initiatives where cause and value propositions align. Additionally, Legal and Public Affairs teams should engage the public sector and external special-interest groups to solicit feedback on how Data-for-Good programs can create measurable value for society while mitigating big-data and artificial-intelligence risks related to bias, inequality, and consumer privacy. Such efforts will go a long way in aligning public interest with corporate actions, thereby reducing overall regulatory risk.

Messaging

Effective messaging strategies focused on earned media and lead generation can drive profitability for Data-for-Good programs. On the cost side, content produced through partnerships can provide valuable source material for journalists. For instance, academic publications are often promoted in press conferences hosted by the journals in which they are published. Data collaborations with the public sector can tell compelling stories about how big data can inform policy making. Moreover, sharing datasets directly with the press can lead to visually captivating data-journalism features in digital media. Together, these outputs can generate earned media, which represents significant savings in paid advertising and public relations fees.

As with many elements of messaging and journalism, relevancy and timing is paramount. For instance, the development of a groundbreaking methodology to enhance the utility and privacy of transaction data might be noteworthy from a scientific perspective, but the topic may be considered less newsworthy to a journalist. On the other hand, if the study coincides with a notable consumer-data breach at a well-known retailer, its relevance makes it more newsworthy.

Data-for-Good programs can embed predictability into their media strategies by focusing on topics that are relevant on a national or global scale, while also building long-term partnerships with the press. For example, a North American satellite imagery firm could launch a project to map levels of destruction following natural disasters. Instead of waiting for the next disaster to strike, the program could collaborate with emergency managers to develop a proof-of-concept based on past hurricanes. Several months before the North American hurricane season begins, the program manager could present their proof-of-concept to relevant journalists to better prepare for the next reportable disaster event.

By leveraging earned media and data journalism partnerships, Data-for-Good programs can generate significant commercial pipeline activity. For an organization's mainstream commercial business, data journalism casts a wider net and attracts business prospects with easily accessible collateral. This approach can capture the interest of prospects who are not reached through traditional targeted advertising campaigns. Additionally, a

mainstream media article's relevance and context can make complex big-data topics more approachable for business users who otherwise would not consider utilizing alternative data for decision-making. Indeed, beyond simply generating commercial leads, effective media strategies have the potential to expand the serviceable available market for an organization's products and services.

Methods

Engaging lead users and subject-matter experts through research partnerships supports company-wide innovation efforts without increasing R&D costs. Likewise, by translating customer pain points into opportunities for improving and validating a firm's data products, Data-for-Good programs can reduce friction in the sales process while enriching an organization's overall product offering.

In order to produce novel Methods through Data-for-Good partnerships, it's critical to identify partners that are highly competent in data science. Researchers - whether from academia, private institutes, or even well-equipped multilateral institutions and NGOs - make an ideal partner profile for the development of innovative methodologies. Data-for-Good managers should engage internal Data Science and Product teams to understand product roadmap gaps and R&D aspirations. The resulting R&D 'wish list' can be used to solicit proposals from researchers, who in turn can conduct independent or joint research projects to develop open-source methodologies that enrich the company's product offering.

Likewise, managers can work with sales teams to identify key obstacles in the sales process that cause hesitation among prospects. For example, prospects may be concerned about data bias, privacy, and robustness, among other issues. By coordinating with external researchers to independently assess such issues, Data-for-Good programs can compile a preponderance of proof points and external validation for their data products. Companies can also proactively work with researchers to remedy any deficiencies revealed in the evaluation process. The resulting body of peer-reviewed research offers transparency for commercial prospects who can make more informed procurement decisions, which effectively shortens the sales cycle.

Monetizing Data-for-Good Partnerships

In addition to the positive influence of Mission, Messaging, and Methods on a firm's profitability, Data-for-Good programs can also directly earn revenue. Actively exploring earned-revenue opportunities not only supports the overall financial sustainability of Data-for-Good programs, but it can also unlock entirely new verticals that firms may decide to pursue. Here, we focus on two channels for directly monetizing Data-for-Good programs: research partnerships, and public sector contracts.

Research Partnerships

Partnerships with academia, along with public and private research institutions, are wellsprings for social impact, innovation, and publicity. But in addition to creating Mission, Messaging, and Methods value, research partnerships can also evolve into commercial relationships. For researchers, big-data assets provide a compelling value proposition: an empirical data source at scale facilitates evidence-based research, scholarly publication, and grant-based fundraising.

In order to create lasting, productive research partnerships, managers should accommodate the needs of researchers and the idiosyncrasies of research funding cycles. For example, if a research group is using a certain dataset for the first time (or if the dataset itself has not been applied to a specific use case) it is advisable to collaborate with researchers through a low-to-no cost model in order to develop a proof-of-concept. Such pilots are well positioned to attract grant-based funding. Indeed, collaborating in the research design process will often result in the firm actively participating in grant proposals, which can total in the millions-of-dollars through vehicles such as the National Science Foundation in the United States, and the European Union's European Research Council.

Public-Private Partnerships

Just as researchers require empirical evidence to inform their studies, so too do public-sector policy makers rely on data-driven insights to inform policy decisions, evaluate existing programs, and guide public investment. From municipal planning departments seeking to create more equitable environments, to multilateral institutions implementing sustainable infrastructure programs in emerging markets, public sector entities are increasingly embracing both public and private big-data assets, with some governmental agencies even creating Chief Data Officer roles.[61]

But how to break into a notoriously bureaucratic vertical? As with academic partnerships, it is advisable to start small with pro-bono engagements that deliver meaningful insights for public stakeholders. While such pilots could entail working directly with public-sector agencies, they could also consist of collaboratives involving relevant NGOs and academic researchers seeking to put their results into practice by actively bridging the research-policy gap. When initial pilots bear positive results, public-sector entities can more easily justify commercial procurements.

Finally, in building revenue-generating partnerships with academia and the public sector, it is always essential to develop pricing models that take into account the financial sustainability of Data-for-Good programs, as well as the partner organization's budget constraints and ability to pay.

Data Sharing Ethics and Responsibility

Opening access to big data sources - especially those derived from human subjects - can present potential ethical and privacy risks that must be considered in order to create a net-positive social impact and maintain brand safety. In short, even the most well-intentioned Data-for-Good initiatives can jeopardize its Mission if it does not take appropriate steps to ensure that data is collected ethically and shared responsibly. We provide key considerations for implementing governance mechanisms and technical interventions designed to facilitate ethical data provenance and responsible, privacy-preserving data sharing.

Ethical Data Collection

If a company's Data-for-Good program shares human-derived data, it should ensure that such data was collected with the informed consent of the individuals generating the data. As part of the collection processes companies should adopt universal data-protection standards by adhering to the most comprehensive regulations within the jurisdictions where they operate, such as the General Data Protect Regulation (GDPR) in the European Union, and the California Consumer Privacy Act (CCPA) in the United States. Companies should collect the minimal amount of data that satisfies downstream usage requirements, and consider additional regulations that may apply to certain classes of metadata, namely health, geolocation, ethnicity, and Personally Identifiable Information (PII).

Companies should consider providing as much control as possible to individuals from whom they collect data. If human-derived data is monetized, companies should also explore ways that allow individuals to share in the downstream value created by their data. Companies such as Pogo[61] and Enigma Data[62] provide applications that allow users to determine what aspects of their data may be shared with specific entities downstream, while also providing financial incentives and direct cash payments to users in exchange for their permissioned data. By recording immutable records of consent and data usage on distributed ledgers, companies like EnigmaData are able to demonstrate ethical data provenance while also delivering direct benefits to the individuals generating data.

Responsible Data Sharing

Once firms decide to expand access to their ethically sourced data for public benefit, it is imperative to implement governance and technical mechanisms to preserve user privacy. It is also critical to ensure that data is not used for unethical purposes, or for projects that could present risks to vulnerable populations. Beyond ethical objections, data misuse can lead to adverse business outcomes and increased risk. Governance mechanisms primarily consist of contractual obligations between the company providing data and the entity receiving data. It is advisable for managers to work with legal counsel to incorporate data governance practices into data use agreements. Key governance considerations include:

- Defining limited, specific use cases for which the data may be used, and specific use cases that are prohibited for data analysis. For example, a company with transaction data may stipulate that their data may be used to analyze food deserts in low-income communities, but may prohibit use of the same data to infer creditworthiness of individuals in the dataset, even if the latter was intended for academic purposes.

- Creating a time-bound engagement so that entities do not have perpetual access to data after the conclusion of a project term.

- Prohibiting certain entities, like law enforcement agencies, from receiving data.

- Signing data-sharing agreements with institutions (e.g. universities) as signatory, rather than with individuals (e.g. individual students), and defining a list of authorized personnel.

- Identifying and transparently communicating any inherent bias present in a dataset so that users can calibrate models and address issues of bias accordingly in publications, policy briefs, or other public goods.

Beyond contractual remedies designed to promote responsible data sharing practices, firms can also implement technical mechanisms to safeguard user privacy and limit harm for situations in which governance remedies fall short. Technical privacy interventions include:

- Reducing the granularity of data through aggregation and the use of privacy-enhancing technologies like differential privacy and synthetic data generation. When implementing such solutions, it is important to evaluate the tradeoffs between privacy and the utility of data for public-benefit use cases.

- Sharing relevant data insights while maintaining raw, granular data "on-premise." Emerging technologies and applications such as data cleanrooms, federated learning, homomorphic encryption, and multi-party computation offer compelling solutions that facilitate the generation of granular insights without having to physically transfer granular data assets from one server to another.

While the primary focus of this paper is to assist managers in creating shared value through Data-for-Good programs, an in-depth overview of technical and governance guidelines for responsible data collection and sharing will help managers avoid pitfalls and minimize privacy risk. For more detailed discussions and guides on responsible data sharing, Join The GovLab's Data Stewards Network to engage in thought-provoking conversations on professionalizing responsible data sharing. Benefit from renowned experts' insights on the risks and rewards of data collaboratives. For guidance and tools, visit the GovLab's Data Collaboratives initiative at https://datacollaboratives.org/.

Conclusion & Recommendations

By focusing on Mission, Messaging, Methods, and Monetization, Data-for-Good programs can benefit society while also strengthening a firm's bottom line. Firms should consider opening access to privately held datasets through Data-for-Good programs that rally teams around a unified mission, create intellectual property, support inbound marketing efforts, and unlock new sources of revenue. Specifically, managers should adopt these four key practices towards creating shared value:

Galvanize stakeholders

Engage internal and external stakeholders to identify issues and causes relevant to the firm. Causes should align with the organization's core competencies, inspire stakeholders, and mobilize employees into a human flywheel of support for the program. Managers should work with key employees to ideate how the firm's data can move the needle on the cause at hand, and to determine roles and responsibilities for all functional areas. Once the program's Mission and value propositions take shape, external facing teams should engage prospective partner organizations and interested parties to launch Data-for-Good initiatives, and to build Data Collaboratives.

Generate leads while elevating brand

Partner with civil society organizations whose causes align with your mission, and develop inspiring narratives and case studies that capture the public interest. Engage data journalists to create visually compelling stories with your data in the mainstream media. In both cases, focus on use cases and verticals that align with the products and services you offer, and leverage outputs as collateral to generate quality leads for commercial prospects.

Match technical needs with innovation opportunities

Coordinate with internal Data Science, Engineering, and Innovation teams to map out needs, competency gaps, and opportunities for technical innovation. Once identified, Data-for-Good managers can seek out and build partnerships with lead users in academia and research organizations. Depending on what is required - such as data validation, algorithm development, or data processing innovation - managers can incentivize external partners with open access to data while also embedding IP sharing requirements within data licensing agreements.

Open new verticals and expand addressable market

Translate Methods and Messaging benefits into a Data-for-Good monetization strategy. Harness the outputs of research partnerships that validate your data to win new business, and translate novel data processing methodologies into new products and services to retain and upsell existing commercial clients. Prioritize Data-for-Good projects that explore novel use cases for your data. Incorporate learnings and outputs from partnerships to create minimally viable products for new verticals and client types, like academia and the public sector.

The 4M Roadmap details the rationale and value proposition for opening access to privately held data, while also offering key recommendations and tools to guide managers in the design and implementation of prosperous Data-for-Good programs. In a forthcoming publication, "The D4G Playbook" we will provide a stepwise action plan for managers to rapidly and responsibly build and launch Data-for-Good programs that benefit society and the company alike.

The 4M Compass
Prioritizing Impact and Profitability

The 4M Compass

Data-for-Good programs that provide highly sought-after big-data assets often face a deluge of demand from organizations looking to gain access to empirical datasets to support their decision making and research objectives. Blindly pursuing all data-access requests could lead to operational headaches, data governance issues, and even commercial product cannibalization. In response to this obstacle, we developed the 4M Compass, a modular decision-support tool that guides managers towards Data-for-Good projects with the highest likelihood of creating shared value.

The 4M Compass is an open-source tool that helps managers comparatively evaluate project opportunities based on their likelihood of meeting objectives across the 4 M's. By utilizing the 4M Compass, managers are able to determine how relevant each of the 4 M's are to their overall objectives, while also defining the specific metrics by which projects are measured. After applying weights to the 4 M's and their corresponding metrics, managers can score specific project proposals, and compare scores across multiple proposals to improve strategic decision making. As a dynamic spreadsheet-based tool, the 4M Compass allows managers to modify and refine the relative importance it places on each criteria and metric under evaluation. Accordingly, as firms and their Data-for-Good programs evolve, the tool can be calibrated to reflect current needs and priorities.

Accessing the Tool

The 4M Compass can be accessed as a Google Sheet through this **LINK**. In order to modify the tool for your own purposes, you can duplicate the Google Sheet and assign yourself as the owner. To do this, first ensure you are signed in to a Google account, as indicated in the image below:

Once you have signed in to your Google Account, you may make a private copy of the tool, as demonstrated in the images below:
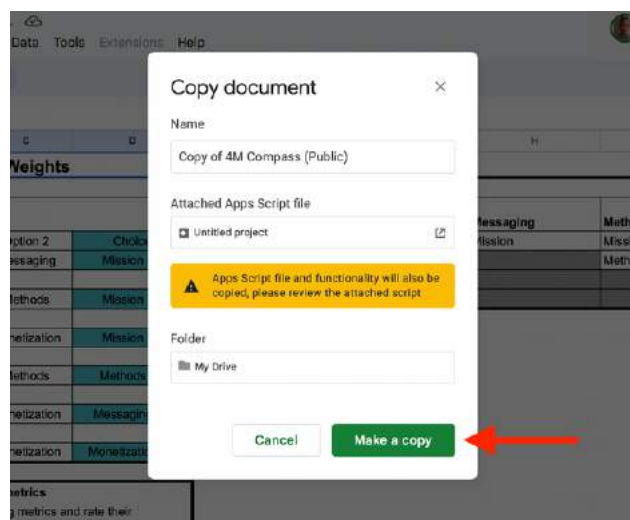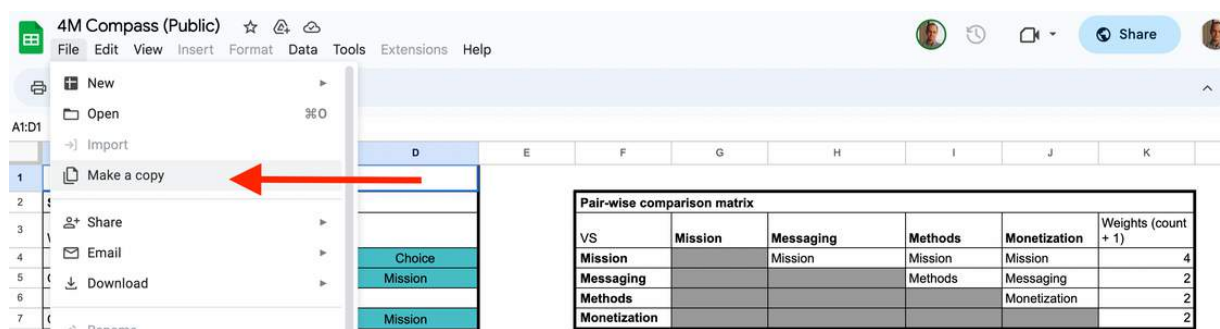
Now that you have your own private copy of the 4M Compass, you can begin to score projects by following the 5 steps detailed below:

1. Weighing the 4M Criteria
2. Selecting Metrics
3. Weighing the Metrics
4. Scouring Projects
5. Interpreting Results

If you find yourself in need of assistance with the tool, please contact Brennan Lake at **hbl36@cornell.edu**

Step 1: Weighing the 4M Criteria

The 4M Compass was designed as a weighted-decision matrix, which allows managers to weigh each M criteria based on its relative importance to the firm's overall strategic objectives. For example, a Fortune 500 company looking to engage employees and improve its reputation may weigh Mission and Messaging criteria higher than Monetization or Methods. Weights may also be adjusted over time. For instance, a start-up with a thriving Data-for-Good program serving academia may shift its focus towards Methods and Monetization to meet R&D and profitability goals.

Users can determine the relevance of each M criteria by completing a pairwise ranking of each criterion against another. For example, the following image demonstrates a selection that primarily favors Mission and Messaging:

Creating the Weights

Section 1 - Weights for the 4M Criteria

Which of these Criteria is most important?

	Option 1	Option 2	Choice
Q1	Mission	Messaging	Mission
Q2	Mission	Methods	Mission
Q3	Mission	Monetization	Mission
Q4	Messaging	Methods	Messaging
Q5	Messaging	Monetization	Messaging
Q6	Methods	Monetization	Methods

Step 2: Selecting Metrics

While the 4M Roadmap is universally applicable to any Data-for-Good program, the metrics by which programs are measured will differ from firm-to-firm across industries. Take the Methods criteria: An electronic medical records firm may create metrics related to developing innovations in privacy-safe computation, but such metrics would be irrelevant to firms that do not collect privacy-sensitive data. Accordingly, the 4M Compass allows users to create their own metrics under each M criteria, while also offering suggested metrics that are widely applicable. The below example illustrates three suggested metrics for the Monetization criterion, along with two additional metrics to be completed by the user. As with weighing Criteria, the best knowledge source for selecting metrics will be from internal functional areas that have their finger on the pulse of pressing and long-term needs across each M criteria.

2.d - Monetization metrics	
D1	Will this project financially break even?
D2	Will this project generate margin?
D3	Can the stakeholders related to this project become paying clients, or directly refer paying clients?
D4	Add a metric...
D5	Add a metric...

Step 3: Weighing the Metrics

Once metrics have been selected for each M criteria, it's time to determine the relative importance of each metric. We employ a numerical scale of 1-5 for each of the weight alternatives, with a lexical equivalent for each number, as illustrated below.

Section 2 - Weights for the underlying metrics

Within each Criteria, add up to 5 underlying metrics and rate their importance

Critical	5
Very Important	4
Nice to Have	3
Marginal	2
Irrelevant	1

The lexical reference helps provide context for the importance of each metric. Take a Fortune 500 technology firm launching a Data-for-Good program that opens access to recruitment and interview data. Their mission is to make hiring processes more equitable in the United States and to support regulation in that regard. It is helpful for their Data-for-Good program to be financially sustainable, but they do not treat the program as a primary driver of revenue, and they do not view the program's stakeholders as potential commercial clients. Accordingly, they may rank the social impact and public policy metrics as "critical", the financial sustainability metric as "nice to have", and the stakeholder-to-client conversion metric as "irrelevant". Again, it is helpful to work with leadership or other program sponsors to collectively determine the importance of each metric.

2.a - Mission metrics

		Importance
A1	How well does the project align with the D4G program's thematic use cases / focus areas?	Very Important
A2	Does the project have the potential to create a positive social impact?	Critical
A3	Will this project create Public Goods (publication, case study, open-source dashboard, etc)?	Nice to Have
A4	Does this project have the potential to affect policy making?	Very Important
A5	Add a metric...	

2.d - Monetization metrics

		Importance
D1	Will this project financially break even?	Nice to Have
D2	Will this project generate margin?	Marginal
D3	Can the stakeholders related to this project become paying clients, or directly refer paying clients?	Irrelevant
D4	Add a metric...	
D5	Add a metric...	

Step 4: Scoring Projects

Now that the metrics and their relative importance have been established, it is time to score the projects! Scoring the performance of certain project components can be straightforward. For instance, if there is an opportunity to provide data for a grant-backed research project whose primary objective is publication, it will be easy to assign the highest-possible score for publication and cost-recovery metrics. Other metrics, such as quantifying social impact, can be trickier to score. To contextualize metric performance, we again use lexical references that correspond to numerical scores: Poor, Marginal, Average, Very Good, and Outstanding.

Recall the tech company seeking to drive the adoption of more equitable hiring practices. If they were to receive a pro-bono project proposal related to the improvement of federally funded STEM education programs in underserved communities, it would not necessarily fit squarely within their Mission's focus area. However, it could reasonably be determined that the company's data could help orient STEM programs towards improving the hireability of students and graduates for tech internships and entry-level positions. In this case, managers may justifiably score the "A1" metric below as "Average" rather than "Poor".

	Criteria and Metrics		Performance
			Scoring values
	Mission		
A1	How well does the project align with the D4G program's thematic use cases / focus areas?		Poor (~ 0% - 20%)
A2	Does the project have the potential to create a positive social impact?		Outstanding (~ 80% - 100%)
A3	Will this project create Public Goods (publication, case study, open-source dashboard, etc)?		Very Good (~ 60% - 80%)
A4	Does this project have the potential to affect policy making?		Average (~ 40% - 60%)
A5			Marginal (~ 20% - 40%)
	Messaging		
B1	Would the company and its stakeholders be proud to see the subject of this work as a headline in the media?		Poor (~ 0% - 20%)
	Would this project's public outputs be appealing to commercial clients		Not Applicable

36

Such a program and partnership could make for good PR and an excellent case study to demonstrate how the firm is actively seeking to improve equity in hiring. Moreover, the fact that it is funded by the federal government could provide direct avenues for public-sector stakeholders to take action upon the program's outcome. Accordingly, the STEM proposal would likely receive positive scores across Mission and Messaging Criteria, while scoring lower on Methods and Monetization, given its pro bono nature and lack of intellectual property generation. Since Methods and Monetization - and their corresponding metrics - are weighted low in this example, the negative scoring impacts are limited, with a total score assigned to the project of 65.34% below.

	Criteria and Metrics	Performance	Weights		Score			
		Scoring values	Weight by metric	Criteria weight	Metric Weighted Score	Criteria Score	Weighted Criteria Score	Criteria Potential MAX
	Mission							
A1	How well does the project align with the D4G program's thematic use cases / focus areas?	Marginal (~ 20% - 40%)	4		120			
A2	Does the project have the potential to create a positive social impact?	Outstanding (~ 80% - 100%)	5		450			
A3	Will this project create Public Goods (publication, case study, open-source dashboard, etc)?	Outstanding (~ 80% - 100%)	5	4	450			
A4	Does this project have the potential to affect policy making?	Outstanding (~ 80% - 100%)	5		450			
A5		Not Applicable	0		0	77.37	309.4736842	360
							85.96%	
	Messaging							
B1	Would the company and its stakeholders be proud to see the subject of this work as a headline in the media?	Outstanding (~ 80% - 100%)	4		360			
B2	Would this project's public outputs be appealing to commercial clients and prospects?	Very Good (~ 60% - 80%)	2		140			
B3	Will this project lead to a publicly shareable output (e.g. case study, report, publication, media article) within 6 months of project launch?	Outstanding (~ 80% - 100%)	5	3	450			
B4	Will the outputs of the project be relevant for media outlets?	Very Good (~ 60% - 80%)	3		210			
B5		Not Applicable	8		0	82.86	248.5714286	270
							92.06%	
	Methods							
C1	Will this project lead to the development of new methods and tools for working with our data?	Poor (~ 0% - 20%)	1		10			
C2	Will this project help validate the quality of our data publicly?	Poor (~ 0% - 20%)	2		20			
C3	Will the methods developed solve pain points for commercial clients?	Poor (~ 0% - 20%)	1	1	10			
C4	Will this project improve areas such as computational efficiency or data privacy?	Poor (~ 0% - 20%)	1		10			
C5		Not Applicable	8		0	10.00	10	90
							11.11%	
	Monetization							
D1	Will this project financially break even?	Poor (~ 0% - 20%)	2		20			
D2	Will this project generate margin?	Poor (~ 0% - 20%)	1		10			
D3	Can the stakeholders related to this project become paying clients, or directly refer paying clients?	Poor (~ 0% - 20%)	1	2	10			
D4		Not Applicable	8		0			
D5		Not Applicable	8		0	10.00	20	180
							11.11%	
					Overall score (%)	65.34%		

Step 5: Interpreting Results

While 65% might seem low at first glance, context is key. Since the 4M Compass is designed as a comparative evaluation tool, it is helpful to benchmark performance. If there were another proposal that performed equally well across Mission and Messaging metric, but also resulted in profit margin and new clients (as detailed in the figure below) it would score around 82%, outperforming the STEM example.

	Criteria and Metrics	Performance	Weights		Score			
		Scoring values	Weight by metric	Criteria weight	Metric Weighted Score	Criteria Score	Weighted Criteria Score	Criteria Potential MAX
	Mission							
A1	How well does the project align with the D4G program's thematic use cases / focus areas?	Marginal (~ 20% - 40%)	4		120			
A2	Does the project have the potential to create a positive social impact?	Outstanding (~ 80% - 100%)	5		450			
A3	Will this project create Public Goods (publication, case study, open-source dashboard, etc)?	Outstanding (~ 80% - 100%)	5	4	450			
A4	Does this project have the potential to affect policy making?	Outstanding (~ 80% - 100%)	5		450			
A5		Not Applicable	0		0	77.37	309.4736842	360
							85.96%	
	Messaging							
B1	Would the company and its stakeholders be proud to see the subject of this work as a headline in the media?	Outstanding (~ 80% - 100%)	4		360			
B2	Would this project's public outputs be appealing to commercial clients and prospects?	Very Good (~ 60% - 80%)	2		140			
B3	Will this project lead to a publicly shareable output (e.g. case study, report, publication, media article) within 6 months of project launch?	Outstanding (~ 80% - 100%)	5	3	450			
B4	Will the outputs of the project be relevant for media outlets?	Very Good (~ 60% - 80%)	3		210			
B5		Not Applicable	0		0	82.86	248.5714286	270
							92.06%	
	Methods							
C1	Will this project lead to the development of new methods and tools for working with our data?	Poor (~ 0% - 20%)	1		10			
C2	Will this project help validate the quality of our data publicly?	Poor (~ 0% - 20%)	2		20			
C3	Will the methods developed solve pain points for commercial clients?	Poor (~ 0% - 20%)	1	1	10			
C4	Will this project improve areas such as computational efficiency or data privacy?	Poor (~ 0% - 20%)	1		10			
C5		Not Applicable	0		0	10.00	10	90
							11.11%	
	Monetization							
D1	Will this project financially break even?	Outstanding (~ 80% - 100%	2		180			
D2	Will this project generate margin?	Outstanding (~ 80% - 100%	1		90			
D3	Can the stakeholders related to this project become paying clients, or directly refer paying clients?	Very Good (~ 60% - 80%)	1	2	70			
D4		Not Applicable	0		0			
D5		Not Applicable	0		0	85.00	170	180
							94.44%	
					Overall score (%)	82.01%		

On the other hand, if there were a project that promised exceptional revenue without achieving high scores on the higher-weighted Mission and Messaging criteria, it would perform comparatively poorly as compared to the STEM proposal. For example, take a proposal from a well-heeled nonprofit consulting organization that sought paid access to employment data for government lobbying efforts, wherein all outputs would be confidential. While such a project could cover costs and may very well support the Data-for-Good program's policy objectives, it would perform poorly across the highly weighted Messaging category. As detailed in the results below, the project would score around 49%. This would lead to the counterintuitive, yet entirely appropriate result of a revenue-generating project scoring lower than the pro-bono STEM opportunity.

Criteria and Metrics		Performance	Weights		Score			
		Scoring values	Weight by metric	Criteria weight	Metric Weighted Score	Criteria Score	Weighted Criteria Score	Criteria Potential MAX
	Mission							
A1	How well does the project align with the D4G program's thematic use cases / focus areas?	Marginal (~ 20% - 40%)	4		120			
A2	Does the project have the potential to create a positive social impact?	Outstanding (~ 80% - 100%)	5		450			
A3	Will this project create Public Goods (publication, case study, open-source dashboard, etc)?	Poor (~ 0% - 20%)	5	4	50			
A4	Does this project have the potential to affect policy making?	Outstanding (~ 80% - 100%)	5		450			
A5		Not Applicable	0		0	56.32	225.2631579	360
	Messaging						62.57%	
B1	Would the company and its stakeholders be proud to see the subject of this work as a headline in the media?	Poor (~ 0% - 20%)	4		40			
B2	Would this project's public outputs be appealing to commercial clients and prospects?	Poor (~ 0% - 20%)	2	3	20			
B3	Will this project lead to a publicly shareable output (e.g. case study, report, publication, media article) within 6 months of project launch?	Poor (~ 0% - 20%)	5		50			
B4	Will the outputs of the project be relevant for media outlets?	Poor (~ 0% - 20%)	3		30			
B5		Not Applicable	0		0	10.00	30	270
	Methods						11.11%	
C1	Will this project lead to the development of new methods and tools for working with our data?	Poor (~ 0% - 20%)	1		10			
C2	Will this project help validate the quality of our data publicly?	Poor (~ 0% - 20%)	2	1	20			
C3	Will the methods developed solve pain points for commercial clients?	Poor (~ 0% - 20%)	1		10			
C4	Will this project improve areas such as computational efficiency or data privacy?	Poor (~ 0% - 20%)	1		10			
C5		Not Applicable	0		0	10.00	10	90
	Monetization						11.11%	
D1	Will this project financially break even?	Outstanding (~ 80% - 100%)	2		180			
D2	Will this project generate margin?	Outstanding (~ 80% - 100%)	1		90			
D3	Can the stakeholders related to this project become paying clients, or directly refer paying clients?	Outstanding (~ 80% - 100%)	1	2	90			
D4		Not Applicable	0		0			
D5		Not Applicable	0		0	90.00	180	180
							100.00%	

Overall score (%)	49.47%

As noted previously, as an organization's priorities evolve, the 4M Compass can be calibrated over time to adjust criteria and metric weights. Likewise, metrics can be added, removed, or modified. A project that scores relatively low one year may score higher once weights and metrics shift. For this reason, it is always advisable to keep track of past project proposals, and to continuously engage prospective partners over time. Finally, high-demand Data-for-Good programs will invariably come across multiple worthy project proposals. When deciding which projects to undertake, managers should use the 4M Compass as a guide to point them towards their goals. Ultimately, managers must determine which paths they will take to achieve them.

About

Acknowledgements

Citations

1) https://www.businessroundtable.org/business-roundtable-redefines-the-purpose-of-a-corporation-to-promote-an-economy-that-serves-all-americans

2) https://www.msci.com/esg-101-what-is-esg/esg-and-performance

3) https://justcapital.com/investor-resources/just-alpha-research/

4) https://hbr.org/2011/01/the-big-idea-creating-shared-value

5) https://towardsdatascience.com/is-data-science-still-a-rising-career-in-2021-722281f7074c

6) https://www.conecomm.com/research-blog/2016-millennial-employee-engagement-study

7) Ibid

8) Ibid

9) Ibid

10) Ibid

11) Ibid

12) https://www.microsoft.com/en-us/ai/ai-for-good

13) https://www.cdc.gov/leprosy/index.html (Note: Hansen's Disease is an alternative name for leprosy. While leprosy is a scientifically recognized term, the word "leper" is considered to be derogatory. The terms "person afflicted with leprosy" or "Hansen's disease patient" are generally preferred).

14) https://news.microsoft.com/europe/features/health-in-your-hands-how-data-and-ai-are-empowering-patients/

15) https://github.com/microsoft/leprosy-skin-lesion-ai-analysis

16) https://blogs.microsoft.com/ai-for-business/differential-privacy/

17) Disclosure: The author participated in the Smartnoise Early Adopter Accelerator program via Cuebiq's Data-for-Good program. Cuebiq also partners with Crisis Ready through unrelated initiatives.

18) Ibid

19) https://www.microsoft.com/en-us/research/group/ai-for-good-research-lab/downloads/

20) https://blogs.microsoft.com/ai-for-business/differential-privacy/

21) https://www.mastercardservices.com/en/solutions/payments/mastercard-advisors-consulting

22) https://www.mastercardcenter.org/insights/data-driven-insights-to-build-connected-communities

23) https://www.mastercardcenter.org/insights/data-exploration-neighborhood-crime-local-business

24) https://hbr.org/2022/03/what-is-the-purpose-of-your-purpose

25) https://economicgraph.linkedin.com/about/

26) https://www.wsj.com/articles/to-track-covid-19-surges-scientists-are-studying-sewage-11640082603

27) https://www.nytimes.com/2020/12/08/nyregion/covid-testing-sewers.html

28) https://time.com/collection/time100-companies-2022/6159393/biobot-analytics/

29) https://spectus.ai/social-impact/

30) https://www.gfdrr.org/en/feature-story/about-us

31) https://olc.worldbank.org/content/informing-disaster-management-through-human-mobility-data-case-study-and-toolkit

32) https://documents1.worldbank.org/curated/en/609961608007759863/pdf/Uncovering-Socioeconomic-Gaps-in-Mobility-Reduction-during-the-COVID-19-Pandemic-using-Location-Data.pdf

33) https://www.adcouncil.org/press-releases/the-ad-council-and-covid-collaborative-reveal-its-up-to-you-campaigns-to-educate-millions-of-americans-about-covid-19-vaccines

34)https://www.prweb.com/releases/cuebiq_partners_with_the_ad_council_and_covid_collaborative_to_measure_covid_19_vaccine_education_initiative/prweb17934582.htm

35) Erko, Autiou, and Thomas, Llewellyn "Innovation Ecosystems, Implications for Innovation Management?" The Oxford Handbook of Innovation Management 2014 (LINK)

36) https://datacollaboratives.org/existing-practices

37) Ibid

38) Ibid

39) https://datacollaboratives.org/static/files/existing-practices-report.pdf

40) https://www.drivendata.org/about/

41) https://www.drivendata.org/competitions/79/competition-image-similarity-1-dev/

42) https://www.drivendata.org/competitions/64/hateful-memes/

43) https://www.drivendata.org/competitions/

44) https://datacollaboratives.org/cases/cuebiq-covid-19-data-collaborative-providing-functional-and-trusted-access-to-mobility-data-for-transforming-current-and-future-epidemic-monitoring-and-response.html

45) https://evhippel.mit.edu/

46) https://www.jpl.nasa.gov/infographics/20-inventions-we-wouldnt-have-without-space-travel

47) https://evhippel.mit.edu/papers/section-1/

48) https://www.statista.com/statistics/871513/worldwide-data-created/

49) https://medium.com/@gocodal/why-are-these-industries-so-slow-to-adopt-new-technologies-a2d0325a9223

50) https://www.urban.org/sites/default/files/publication/104869/a-privacy-preserving-validation-server-prototype.pdf

51) https://www.microsoft.com/en-us/ai/ai-for-humanitarian-action-projects?activetab=pivot1:primaryr8

52) https://blogs.bing.com/maps/2019-09/microsoft-releases-18M-building-footprints-in-uganda-and-tanzania-to-enable-ai-assisted-mapping

53) https://blogs.bing.com/maps/2019-09/microsoft-releases-18M-building-footprints-in-uganda-and-tanzania-to-enable-ai-assisted-mapping

54) https://github.com/Microsoft/CNTK

55) https://www.tomorrow.io/

56) https://www.tomorrownow.org/

57) https://www.nber.org/papers/w26167

58) https://www.tomorrownow.org/about/

59) https://earth.org/data_visualization/is-climate-change-causing-locusts-swarm-resurgence/ & https://www.jstor.org/stable/4238?origin=crossref

60) https://www.tomorrownow.org/blog/announcing-project-locust-on-time-climate-action-for-farmers-by-farmers/

61) https://www2.deloitte.com/us/en/insights/industry/public-sector/chief-data-officer-government-playbook.html

62) https://www.joinpogo.com/

63) https://enigmadata.net/