Evaluation of the Social Science One – Social Science Research Council – Facebook Partnership

Drs. Amy O'Hara and Jodi Nelson

December 2019

Contents

Introduction

Background

Launched in April 2018, a partnership between Facebook, Social Science One (SS1), the Social Science Research Council (SSRC) and 7 foundations sought to make Facebook data available to researchers to study the impact of social media on democracy and elections.

The project was developed at a critical juncture in Facebook's history. Although the social media company had been under scrutiny for its data sharing practices since 2011, unease about Russia's ability to use Facebook to influence democracy heightened in 2017. Some of Facebook's strongest critics believe that the company's algorithms and the fake news in its news feeds unduly influenced the results of the 2016 presidential election and the Brexit vote. Russian efforts to manipulate elections in the Netherlands, France and Germany exacerbated public apprehension that the social network could no longer be trusted.[1]

By the Spring of 2018, Facebook CEO Mark Zuckerberg faced calls to explain why he had not done more to protect user privacy and whether the company was in violation of its 2011 consent decree with the Federal Trade Commission (FTC).[2] In March, he was called to testify before Congress. Soon thereafter, news broke that Cambridge Analytica – a data firm that worked with Donald Trump's 2016 presidential campaign – had inappropriately accessed the personal data of roughly 87 million users without their permission.[3]

As that news broke, Facebook Vice President Elliot Schrage approached Hewlett Foundation president Larry Kramer to discuss what Facebook could do to avoid a repeat of the 2016 election interference and to support independent research. Kramer galvanized interest across peer foundations to support a new industry-academic partnership providing independent researchers with access to Facebook data.[4] The other foundations joined the effort, aware of both the urgency of pending elections and the unique opportunity presented by the political pressure on Facebook.

The industry-academic partnership featured a new organizational model to offer academics extensive privacy-preserving data from Facebook with no pre-publication approval requirements. Facebook and the foundations believed that pooled funding from seven ideologically diverse philanthropies to support this model would help ensure that research results on social media's impact on democracy was independent and credible.

The idea was based on a concept developed by Professors Gary King and Nathaniel Persily, two social scientists who had been conferring with different teams at Facebook on increased researcher access to the social media company's data. As described in their seminal article, "A New Model for Industry-Academic Partnerships," the key obstacle to advancing important

---

[1] https://www.nytimes.com/2017/10/30/technology/facebook-google-russia.html

[2] The FTC charged Facebook with deceptive practices over its tracking and sharing of user data, violating consumer privacy. While Facebook admitted no wrongdoing, they entered the 2011 consent decree, agreeing to begin a "comprehensive privacy program" and to have a third-party conduct audits every two years for the next 20 years to certify its program is effective. https://www.ftc.gov/news-events/press-releases/2012/08/ftc-approves-final-settlement-facebook

[3] https://www.nytimes.com/2018/04/04/technology/mark-zuckerberg-testify-congress.html

[4] Jeffrey Mervis describes the context and conversations that took place among the key Facebook, foundation and academic individuals in "Privacy concerns could derail unprecedented plan to use Facebook data to study elections," Science on line, September 24, 2019.

research using social media data was twofold: one, researchers could not gain access to social media data without signing non-disclosure agreements that sacrificed their independence; and two, many of the most pressing research topics were inherently provocative and against company self- interests.[5] As a result, companies were not motivated to give researchers both access to their data and freedom to publish without their approval. The "standoff" King and Persily describe is that insiders to the social media company have access to the data and other information necessary to do research but not the independence to do credible research, while researchers outside the company have independence and credibility, but not access.

King and Persily's solution to this dilemma was to create a model that ensures that the company's interests are protected, data are secure, user information is kept private, and researchers maintain independence. Their model featured a commission of academics who would forgo the right to publish in return for complete access to the company's data and knowledge of its data systems, policies, platforms and practices. With insider access to the company and expertise in the relevant academic field, the commission would be able to identify important research questions that could be answered with specific and privacy-preserving subsets of the firm's data.

Expectations for the project were high. Interviews revealed partner aspirations to "[enable] the world to benefit from the data that Facebook has collected and stored in order to answer hard social and policy questions;" "liberate data from companies;" "set a precedent for academic freedom;" and "awaken researchers all over the world because they would get access to new data." Nothing of this nature had been tried before with Facebook and many of the funders interviewed characterized the project as an experiment without a roadmap. At the time of the project's launch, they took a leap of faith that the model presented by King and Persily provided a viable guide.

All parties spent significant time and resources on project implementation during the next 16 months. By all accounts, there were more challenges than any one of them could have anticipated. As the evaluation report describes, Facebook's failure to release data as originally planned stymied the project's ability to achieve its primary goal of producing independent, high quality research. As of November 2019, Facebook has released three data sets, none of which meets the parameters agreed upon by the partners in July 2018. Multiple stakeholders have expressed that Facebook remains committed to releasing a dataset as close to the originally specified data as possible given privacy constraints.

Given the novelty of the partnership and the complexity of providing data access at this scale in an environment where privacy concerns are paramount, the Hewlett Foundation commissioned an external evaluation[6] to document and reflect on the progress made, and to share the lessons learned. This assessment focuses on the structure of the project, not the quality of the data made available. The evaluation focuses on the duration of the project, from March 2018 through November 2019, though additional data sets and research reports will accrue over time.

---

[5] King, Gary, and Nathaniel Persily. (2019), "A new model for industry-academic partnerships." Cambridge Mass. https://gking.harvard.edu/partnerships

[6] We prepared this evaluation knowing that the foundation intended a public audience. We comply with the foundation's standard editorial guidance for such evaluations. This includes adherence to its general rule about openness and transparency in evaluation documents, as well as two exceptions for documents that are shared with the public, namely, "information that [the Foundation has] an ethical or legal duty to keep confidential (e.g. staffing changes at a grantee that have not yet been made public) or situations where sharing information publicly could cause material harm to a grantee, such as criticism of an individual organization's work." We have thus not included information of this sort in the following document.

The report includes four sections:

- **The Introduction** describes the background and evaluation design.
- **The Project** describes the project set up, partners and activities, including a timeline that provides visibility into the regulatory and media context.
- **Evaluation Findings** shares the evaluation team's primary observations, derived from the interviews and document review conducted during the evaluation period.
- **Takeaways and Considerations** provides the evaluation team's additional impressions, offered to inform the partners' decisions about next steps.

Evaluation Design

The Hewlett Foundation identified the evaluation's primary audience, objectives and questions.

The primary audience of the evaluation includes the Hewlett Foundation, its peer funders and the public writ large. The evaluation objectives were to document and share lessons learned in the SS1-SSRC-Facebook project in order to inform stakeholders interested in the relationship between social media and democracy and, in particular, how to secure privacy-protected data for use in independent research. To meet these objectives, we needed to ask and answer basic questions about what happened, why it happened and what lessons can benefit similar projects and the public good moving forward.

Given the objectives and retrospective nature of the evaluation, we used qualitative methods that included:

- Review of all available project documents provided to the evaluation team by the Hewlett Foundation, SSRC, SS1, and Facebook;
- Review of publicly available press releases, organization and project descriptions as well as relevant media articles;
- Interviews with 20 internal stakeholders, all directly involved in the project (note that not all funders were interviewed) and identified by the Hewlett Foundation;
- Interviews with 7 outside experts with direct experience with industry-academic collaborations.

The relatively small number of stakeholders interviewed, and the short period of project implementation presented limits on the evaluation team's ability to trace the project's trajectory and underlying dynamics. To strengthen our evidence base, we consulted outside experts for their insights on comparable models and reviewed philanthropic literature for relevant lessons and findings.

The speed, high stakes, newsworthiness and number of players involved contribute to the reality that there are competing perspectives about how the project unfolded. We note the Rashomon quality to our evaluation and press surrounding this project: people's positions can and do shape their views and experiences.[7] With this and the quick timeline for this evaluation in mind, we hope the following pages share patterns and insights to inform similar endeavors.

---

[7] Rashomon is a famous Japanese film known for a plot device that involves various characters providing subjective, alternative and contradictory versions of the same incident.

The Project

The Set Up and Partners

On April 9, 2018, Facebook CEO Mark Zuckerberg announced the project on his blog:

> "Today we're taking another step – establishing an independent election research commission that will solicit research on the effects of social media on elections and democracy. The goal is both to get the ideas of leading academics on how to address these issues as well as to hold us accountable for making sure we protect the integrity of these elections on Facebook. To do this, we're working with foundations across the US to set up a committee of academic experts who will come up with research topics and select -- through a peer review process -- independent researchers to study them. We'll give those researchers access to our resources so they can draw unbiased conclusions about Facebook's role in elections, including how we're handling the risks on our platform and what steps we need to take before future elections. They'll share their work publicly, and we won't require our approval to publish."[8]

The announcement came amidst growing concerns that Facebook data could be used to alter global political outcomes. Facebook, and social media companies more generally, were accused of simultaneously oversharing and under-sharing data. They were oversharing with developers[9] whose handling of the data was neither monitored nor regulated, while under-sharing with qualified researchers who could study the relationships between social media, disinformation and democracy.

Facebook had a history of partnering with academic researchers before the project, sharing data with those who agreed to investigate predefined topics or proposed their own research questions but agreed to work within the company.[10] As noted by King and Persily, such arrangements did not produce fully independent research. Moreover, data access was granted to these researchers on an ad hoc basis, without separate, robust monitoring and logging systems.

As details of the Cambridge Analytica story emerged, it became evident that Facebook had not been paying enough attention to the different ways that users' data could be leaked and at what cost to society. In parallel, many funders, including the Hewlett Foundation, had been supporting analyses on the relationship between democracy and digital disinformation. It was clear to funders and scholars that the proprietary data in Facebook's platform was crucial to understand social media's role in recent and upcoming elections. Hewlett president Larry Kramer describes the project accordingly:

---

[8] https://www.facebook.com/zuck/posts/10104797374385071
[9] Facebook had opened an application programming interface (API) to user data in 2010, originally offering developers data for users and their social network. By 2014, the Open Graph API no longer included data on users' friends.
[10] https://www.bloomberg.com/news/articles/2018-06-14/if-you-re-a-facebook-user-you-re-also-a-research-subject

"This is a critical first step toward a deeper understanding of how social media is used to sow distrust and spread disinformation that threatens American democracy – and what we, as a society, can do about it to create a healthier discourse…This commitment by Mark Zuckerberg, and Facebook, reflects the need to take responsibility for how the platform is used. We recognize, given our own institutional heritage, that Silicon Valley leaders with high ideals who pledge and maintain an enduring commitment to the public interest can make a profound and long-lasting contribution to society."[11]

Personal relationships drove a quick mobilization of resources and partners. Harvard professor Gary King knew Elliot Schrage, then Facebook Vice President of Communications and Public Policy. Schrage knew Hewlett President Larry Kramer, the former dean of Stanford Law School, who also knew Nathaniel Persily, a Stanford professor and an expert on election commissions. Facebook had approached King and Persily separately to discuss the possibility of studying the 2016 election as the Cambridge Analytica scandal broke. King and Persily proposed the academic-industry partnership as an alternative to working as insiders with bespoke access to Facebook data. With Facebook and Hewlett engaged, when Kramer reached out to foundation peers to join the Hewlett Foundation in support of the idea, nearly $10 million was raised within a week from The John S. and James L. Knight Foundation, Laura and John Arnold Foundation, the Alfred P. Sloan Foundation, The Charles Koch Foundation, The Democracy Fund and Omidyar Network.

This ideologically diverse group of foundations shared a common sense of urgency that the public pressure on Facebook presented a unique opportunity to act quickly since Facebook had an incentive to increase transparency before more elections were impacted. As articulated by the Omidyar Network's Mike Kubzansky, "we are realistic about the range of motivations of the players at the table, but also know that systemic change does not come without a few uncertain bets. We believe this is a critical step forward at a critical juncture and will remain vigilant for the rest of the journey to ensure that the outcome moves us closer to a tech industry that embraces its fundamental responsibilities to society."[12] Table 1 below shares funders' quotes, published in relevant press releases.

Table 1: FOUNDATION PRESS[13]

John S. and James L. Knight Foundation	"Social media is now where many go for news. We can't understand our democracy without opening the hood and taking a look. This first, serious and independent effort to do that is an exciting opportunity to look deep inside the data and operations of the world's largest social network," said Alberto Ibargüen, Knight Foundation president.
Democracy Fund	"We believe that independent funding of this research is critical, and hope that the program will help the public and policymakers better understand

[11] https://hewlett.org/newsroom/hewlett-knight-koch-foundations-with-other-funders-will-support-independent-research-on- facebooks-role-in-elections-and-democracy/

[12] https://www.omidyar.com/blog/partnership-more-responsible-tech-assessing-facebook%E2%80%99s-role-elections

[13] https://hewlett.org/newsroom/hewlett-knight-koch-foundations-with-other-funders-will-support-independent-research-on- facebooks-role-in-elections-and-democracy/; https://www.democracyfund.org/blog/entry/df-on-support-independent-analysis-of-facebook-role-in-elections; https://www.omidyar.com/blog/partnership-more-responsible-tech-assessing-facebook%E2%80%99s-role-elections; https://www.arnoldventures.org/newsroom/laura-and-john-arnold-foundation-supports-effort-to-study-social-medias-impact-on-elections/

	how Facebook is shaping our elections, social fabric, and democratic life" said Democracy Fund's Tom Glaisyer.
William and Flora Hewlett Foundation	"We are not naive, and we understand the risks in working closely with a for-profit company whose business model may be threatened by the results of the research. We believe the protections that have been put in place, with Facebook's support, adequately ensure both the importance of the questions to be asked and the independence of the work that attempts to answer them" said Larry Kramer, Hewlett Foundation president.
Omidyar Network	"The potential topics for analysis – misinformation; polarizing content; promoting freedom of expression and association; protecting domestic elections from foreign interference; and civic engagement – have huge implications not only for the tech industry, but for our society at large. … We hope that this will set a model – leveraging independent research based on actual data to improve products – that other social media platforms and tech companies can follow."
Laura and John Arnold Foundation	"Social media's arrival in the political universe has brought with it both promise and uncertainty. We believe that this phenomenon needs study. We don't know where the data will take us but we do know that the examination is of vital importance."

Facebook and the foundations believed that the King and Persily model was an experiment worth trying. Facebook agreed to provide access to data systems and work through policies and protocols to enable research while preserving users' privacy. The election research commission would identify important research questions and also serve an accountability role, acting in the public interest to call out Facebook if it failed to fulfill commitments to support independent research.

King and Persily established SS1 as a Limited Liability Corporation (LLC) on May 30, 2018, hosted the new organization at Harvard's Institute for Quantitative Social Science, and described it as follows: "Given the scope of our work at hand, which extends beyond the role that Facebook and social media plays in elections, we have officially named this effort "Social Science One". The "commission" is a group of academics inside SS1 dedicated to this project."[14] The model as presented by King and Persily (2019) is depicted in Figure 1.

---

[14] https://socialscience.one/faq/wheres-election-commission-election-research-commission-research-commission-thing-0

Figure 1: ACADEMIC INDUSTRY-PARTNERSHIP MODEL
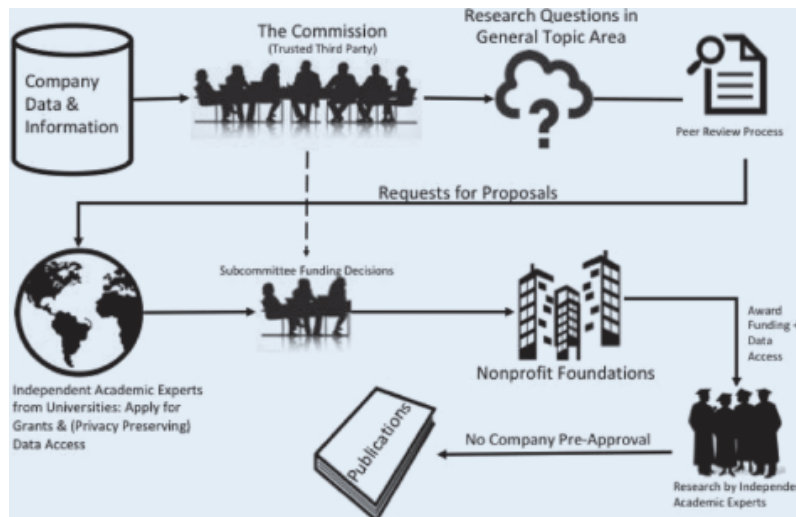
Source: King and Persily (2019*)*

The co-chairs of SS1 had independent, successful careers prior to forming SS1, but had never worked together. Persily had deep experience with commissions and elections. King was an innovator in quantitative social science. Both had expressed interest in expanding access to Facebook data for researchers: Persily and other academics had engaged with Facebook's policy research team and King had discussed data access with Mark Zuckerberg.

As the initiative took shape, the funder group turned to SSRC to administer the project. They wanted a more diverse steering committee and appreciated SSRC's reputation for transparency, independence and credibility. For nearly a century, SSRC has pursued a mission to "mobilize necessary knowledge for the public good by supporting scholars worldwide, generating new research across disciplines and linking researchers with policymakers and citizens."[15] With a strong reputation for running effective peer review processes and academic awards, SSRC was a logical choice to bring missing administrative capacity. As with the foundations and SS1, SSRC had a standing interest in understanding the social impact of media and technology.[16]

Project Budget

As described in project press releases, the foundations would pay the expenses of social media and democracy research conducted by an independent and diverse set of scholars, and Facebook would grant these scholars access to proprietary data that had met the company's "new, heightened security around user privacy."[17]

---

[15] http://ssrc.org/
[16] Launched in April 2017 and completed in 2018, SSRC task force report "To Secure Knowledge" was led by many of the same academics involved in the SS1-Facebook partnership. The report details the changing institutional arrangements and impact of data and technology on 21st century social science research, and suggested paths to address these changes (https://s3.amazonaws.com/ssrc-static/tsk/SSRC+To+Secure+Knowledge.pdf).
[17] https://hewlett.org/newsroom/hewlett-knight-koch-foundations-with-other-funders-will-support-independent-research-on- facebooks-role-in-elections-and-democracy/

The work was funded by a 4-month, $1 million planning grant and a subsequent 12-month $9.5 million follow-on grant to SSRC.[18] The table below presents the respective contributions of participating foundations, pooled and managed by SSRC in order to protect the independence of all research decisions. The project was designed to grow and scale as additional datasets were made available by Facebook and as additional funders joined the collaboration. For example, the Children's Investment Fund Foundation (CIFF) pledged $1 million support in December 2018 to fund research projects outside of the United States.

Table 2: FOUNDATION CONTRIBUTIONS

FUNDER	ORIGINAL COMMITMENT	FUNDS RECEIVED
The William And Flora Hewlett Foundation	1,950,000	1,950,000
The Charles Koch Foundation	1,000,000	1,000,000
The Knight Foundation	1,900,000	950,000
The Laura and John Arnold Foundation	1,650,000	650,000
The Alfred P. Sloan Foundation	500,000	500,000
The Children's Investment Fund Foundation	1,000,000	500,000
The Democracy Fund	150,000	150,000
Omidyar Network[19]	1,350,000	-
Total Funds	9,500,000	5,700,000

Project Goals

The project's primary goal was to produce scholarly research on elections and democracy by providing selected researchers with $50,000 grants and access to Facebook data. The results outlined in SSRC contract documents focus on
the research process and publications to be generated after Facebook and SS1 had successfully connected researchers to privacy-protected data.

SSRC documents define an ambitious set of outputs on the research process, given the relatively short period of the project. These include:
- 12 full rounds of rolling review, each able to handle 80-100 applications
- Standing peer review panel of 40 members with additional pool of up to 100 reviewers
- 90-200 research grants
- Grantee workshop for 90 + grantees

$5.26 million of the original approved budget of $9.5 million will have been spent by December 31, 2019. The budget is presented in Figure 2.

---

[18] $9.5 million is the original 12-month SSRC budget. The $1 million 4-month planning grant preceded the yearlong grant. In late 2018, the Children's Investment Fund Foundation pledged an additional $1 million, bringing the total amount pledged to support the project to $11.4 million.
[19] The Omidyar Network pledged $1.35M to support the project, making the funding conditional on outputs delivered, specifically the publication of 10 white papers expected to be produced six months after the project started. As the project hit unexpected road bumps, the outputs were not achieved and the funding was not allocated.

Figure 2: SSRC BUDGET, SEPTEMBER 2018-DECEMBER 2019
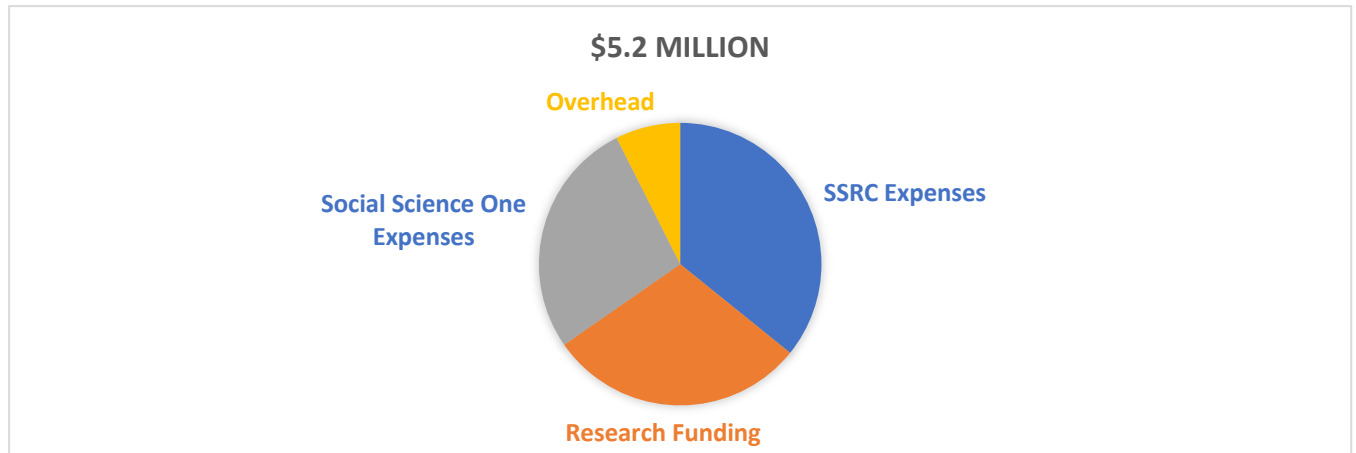
Project Implementation

Project implementation occurred from June 2018 to August 2019 and included four core activities. The timeline in Figure 3 places project activities on the right and external events on the left to convey the potential influence outside pressures may have had on the different partners.

Timeline

The political and regulatory context in which the project was implemented created a sense of urgency among the principals. We note relevant events on the timeline on page 13 and define them below.

Zuckerberg's April 2018 announcement of the Election Research Commission was immediately criticized in the press for its lack of specifics and approach. The General Data Protection Regulation (GDPR) took effect on May 25, 2018. The following month, California passed the Consumer Privacy Act (CCPA), the first in the US to reflect GDPR-like conditions. When it becomes effective in January 2020, the CCPA will grant California residents new rights regarding their personal information and impose data protection duties on relevant businesses in California, just as GDPR did for the European Union.

In July 2018, the SSRC and SS1 announcement of the Facebook URL Shares RFP was met with a sharp warning from the Electronic Privacy Information Center (EPIC), a non-profit public interest research center. They sent SS1 a letter noting what they believed to be violations of FTC and GDPR consent requirements and personal data protections. EPIC advised against the project.

In October 2018, Facebook signed on to the Code of Practice on Disinformation developed by a working group of the European Commission to combat disinformation, fake accounts, and bots in online platforms. Internally, Facebook began development of a privacy-preserving research tool that would apply differential privacy to the URL Shares data.

In November 2018, funders, SSRC, and Social Science One were informed that data would not be ready by year's end. Another funder (CIFF) joined the initiative in December 2018, expecting to fund non-US research in later rounds of grantmaking. However, reacting to data delays, the partners had suspended proposal review in November.

In January 2019, SS1 released its first update since the RFP launch in July. In their blog, they described their achievements and the delays in data access. They explained that the path forward would include the delivery of URL Shares data in stages, reported that project grantees would have immediate access to Facebook's CrowdTangle and Ad Library APIs, and announced development of linked Facebook-survey datasets.

In February 2019, the non-profit Mozilla Foundation and 37 civil rights and non-profit organizations from around the world sent an open letter[20] to Facebook asking for public access to the Ad Archive API. They wanted the API open before April 2019 so that developers could create tools before EU Parliamentary elections. Facebook responded, making Ad Library available in March. Mozilla criticized this release for not including all ad data, for lacking information on targeting criteria, and for putting constraints on researchers.

In April 2019, Facebook, Social Science One, and SSRC announced the first awardees of the Social Media and Democracy Research Grant program, where more than 50 researchers from 25 academic institutions across 8 countries were chosen through a competitive peer review process organized by the SSRC. In June, after finalizing legal agreements with each of the researchers Facebook hosted an onsite training to acquaint researchers with the Research Tool they would leverage to access Facebook data. Facebook had worked from August 2018 to June 2019 to build out and conduct security tests of the system. Researchers at this point were provided access to a URL training dataset, Crowdtangle, and the Ad Library API.

By Summer 2019, Facebook's delays in providing the originally announced URL dataset motivated the foundations to assess the project's status and prospects. In July 2019, the FTC announced a $5 billion fine against Facebook for violating its consent decree in the Cambridge Analytica scandal.

August 2019 marked the end of the SSRC-Funder Committee project's grant period. Project wind down activities began at SSRC. SS1 released two blog posts: an update on the project and a joint statement with Facebook expressing gratitude for project support from the funders and SSRC and clarifying that future data access would involve only SS1 and Facebook.

Near the end of September, an SS1 blog announced a new industry-academic partnership between Gary King and Microsoft. In October, SS1 announced development of new peer review processes for the URL Shares Light data for future requestors, and the funders decided to support the Cohort 2 grantees who had been selected through SSRC's peer review process.

---

[20] Other open letters from the Mozilla Foundation have asked tech companies to improve privacy practices, as with Venmo, and to halt political ads in the run-up to the UK General Election.
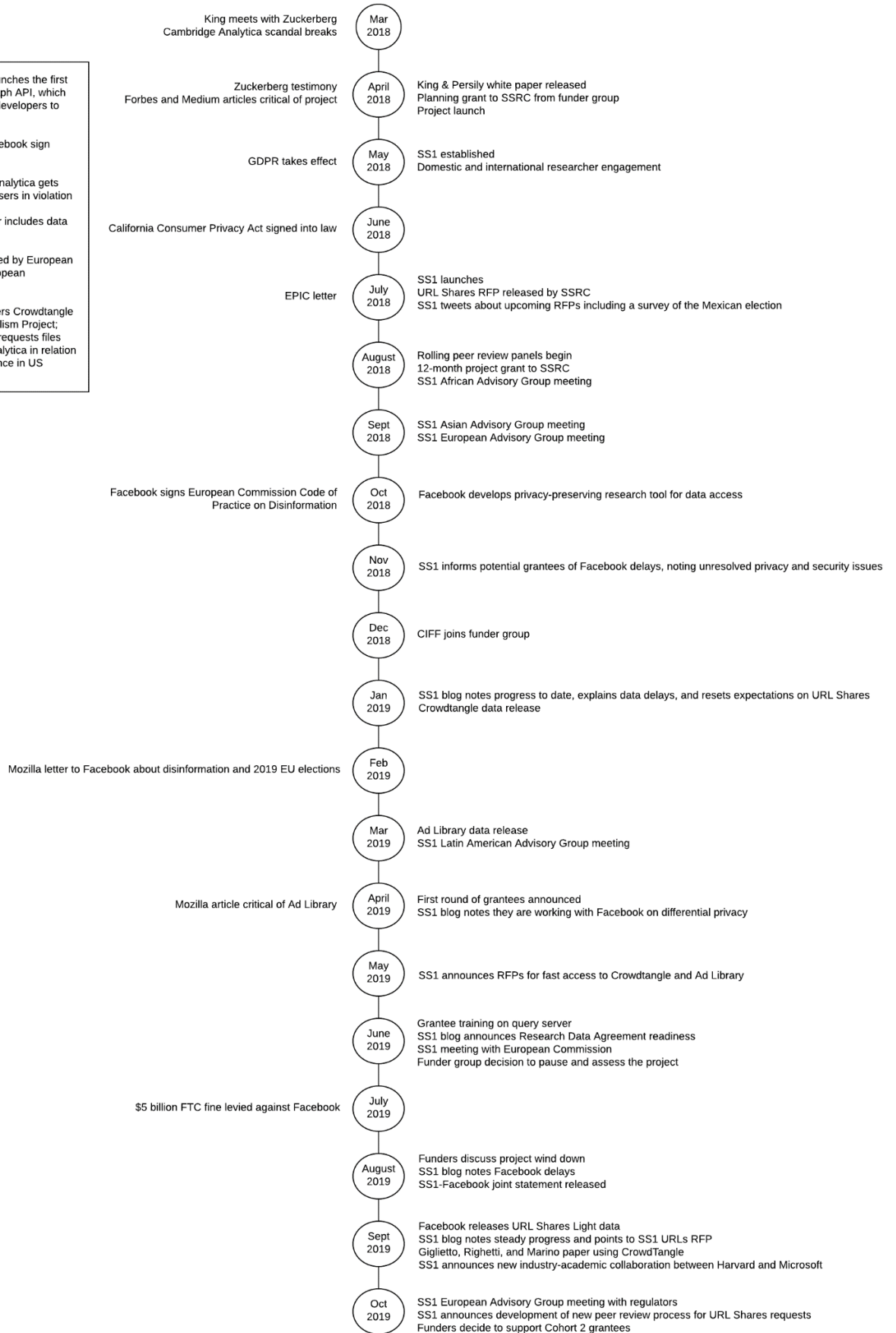
Context box (left):

2010: Facebook launches the first version of Open Graph API, which allows outside app developers to access user data

2011: FTC and Facebook sign consent decree

2014: Cambridge Analytica gets data on Facebook users in violation of terms of use; Graph API no longer includes data on user friends

2016: GDPR adopted by European Parliament and European Commission

2017: Facebook offers Crowdtangle data through Journalism Project; US special counsel requests files from Cambridge Analytica in relation to Russian interference in US presidential election

Timeline:

Mar 2018
Left: King meets with Zuckerberg / Cambridge Analytica scandal breaks

April 2018
Left: Zuckerberg testimony / Forbes and Medium articles critical of project
Right: King & Persily white paper released / Planning grant to SSRC from funder group / Project launch

May 2018
Left: GDPR takes effect
Right: SS1 established / Domestic and international researcher engagement

June 2018
Left: California Consumer Privacy Act signed into law

July 2018
Left: EPIC letter
Right: SS1 launches / URL Shares RFP released by SSRC / SS1 tweets about upcoming RFPs including a survey of the Mexican election

August 2018
Right: Rolling peer review panels begin / 12-month project grant to SSRC / SS1 African Advisory Group meeting

Sept 2018
Right: SS1 Asian Advisory Group meeting / SS1 European Advisory Group meeting

Oct 2018
Left: Facebook signs European Commission Code of Practice on Disinformation
Right: Facebook develops privacy-preserving research tool for data access

Nov 2018
Right: SS1 informs potential grantees of Facebook delays, noting unresolved privacy and security issues

Dec 2018
Right: CIFF joins funder group

Jan 2019
Right: SS1 blog notes progress to date, explains data delays, and resets expectations on URL Shares / Crowdtangle data release

Feb 2019
Left: Mozilla letter to Facebook about disinformation and 2019 EU elections

Mar 2019
Right: Ad Library data release / SS1 Latin American Advisory Group meeting

April 2019
Left: Mozilla article critical of Ad Library
Right: First round of grantees announced / SS1 blog notes they are working with Facebook on differential privacy

May 2019
Right: SS1 announces RFPs for fast access to Crowdtangle and Ad Library

June 2019
Right: Grantee training on query server / SS1 blog announces Research Data Agreement readiness / SS1 meeting with European Commission / Funder group decision to pause and assess the project

July 2019
Left: $5 billion FTC fine levied against Facebook

August 2019
Right: Funders discuss project wind down / SS1 blog notes Facebook delays / SS1-Facebook joint statement released

Sept 2019
Right: Facebook releases URL Shares Light data / SS1 blog notes steady progress and points to SS1 URLs RFP / Giglietto, Righetti, and Marino paper using CrowdTangle / SS1 announces new industry-academic collaboration between Harvard and Microsoft

Oct 2019
Right: SS1 European Advisory Group meeting with regulators / SS1 announces development of new peer review process for URL Shares requests / Funders decide to support Cohort 2 grantees

Activities

The SS1-SSRC-Facebook partnership involved four core activities, detailed below.

Activity 1: Creating the data intermediary organization
King and Persily formed the independent commission SS1 in May 2018. Because SS1 did not have its own staff, team members from Harvard's Institute for Quantitative Social Studies supported their efforts, under their director, Gary King. Available evidence indicates that SS1 held frequent meetings with Facebook; engaged legal counsel to develop and negotiate multiple contracts that protected the interests of the company, researchers, and their institutions; created a public online presence to help drive awareness of the project among academics; and formed 11 expert committees to provide advice and maintain strong connections with global academic communities.

Activity 2: Building the research infrastructure
SSRC and SS1 conducted academic outreach and external communications and also established new research protocols and processes to facilitate research. SS1 co-chair Nathaniel Persily traveled to several countries to raise project awareness with the academic community and to invite researchers to SS1 committees. Through these committees, SS1 assembled subject matter expertise[21] and perspectives from Africa, Asia, Europe, Latin America, and North America. Both organizations developed innovative peer and ethical review processes for this project.[22]

SSRC and SS1 jointly developed and released the URL Shares Request for Proposals (RFP) in July 2018. Rolling proposal review began in August 2018. Review panel meetings were held in September and October 2018, but were suspended in November 2018, noting a drop off in submissions amid growing uncertainty about data readiness. At the beginning of the project, SSRC planned to coordinate a standing peer review panel of 40 members with each member participating in 3 rounds of review before final selections were made. SSRC also expected to mobilize a pool of up to 100 ad-hoc reviewers, and to hold monthly in-person review meetings, anticipating that each round could accommodate between 80-100 applications.

In May 2019, SS1 released RFPs for data-access-only to Crowdtangle and Ad Library data through new review processes. These were external to the process designed and administered by SSRC for URL Shares that was associated with grant funding. SS1's request process required only a 300-word proposal and promised a two-week turnaround; applicants needed principal investigator status at their institution and a scope of study focused on elections and democracy.

Activity 3: Creating data access
The uncharted nature of this project relates to the many unknowns of providing research access to Facebook data in a technically efficient and legal manner. Evidence provided from Facebook

---

[21] These committees included research design, privacy and security, disinformation and election integrity, pollination, political advertising, and civic engagement.
[22] SSRC developed and announced its Social Data Research Review Framework, a peer review architecture uniquely designed for the study of social media to guide "the processes used to determine both who should be granted Facebook data access and project funding and what they are allowed to do with them." SS1 developed its Pre-Peer Review process for the research thus produced with the goal of reducing the time that papers with flaws are on the web prior to peer review and publication. While SSRC's RFP involved grant funding as well as data access, SS1 RFPs were designed to provide data access only.

and interviews with project principals describe the iterative approach Facebook teams used to test whether different datasets and access modes satisfied all necessary requirements. For example, an early course correction involved a shift away from custom dataset requests from researchers to an approach providing a dataset that could satisfy many research inquiries.

During the project's duration, Facebook made three data products available for academic research: Crowdtangle, Ad Library and URL Shares Light. Crowdtangle and Ad Library were arguably not new – earlier versions were already public facing and had cleared dissemination protocols within Facebook. However, CrowdTangle in particular was not offered to the academic research community only to the journalism community until January 2019. The data at the heart of the project was defined in the July 2018 RFP as the Facebook URL Shares dataset, and at the time of this evaluation, only the "Light" version was available to researchers.[23]

Crowdtangle, a tool for content discovery and analytics, had been freely available to news industry professionals through the Facebook Journalism Project since 2017. Crowdtangle lets users explore the spread of and reactions to social media content. Facebook expanded access to academic researchers during the course of this project.

In March 2019, Facebook released Ad Library, an expansion of its Ads Archive offering that had launched in May 2018. Ad Library lets users see "who was pushing what messages and how much they were paying to do it," according to Matthew Rosenberg in the New York Times.

While Ads Archive only included ads related to politics or policy issues, Ad Library includes all active ads about anything, as well as inactive political and issue ads. Facebook made access to the data easier and offered the data through an API, as noted above with the Timeline. Upon initial release, the API was difficult to use; researchers reported bugs to Facebook until the bug reporting page broke. The tool's usability challenges were reported in the New York Times and Facebook addressed the bugs. Researchers indicate that the API now works well enough for their purposes, though there are still some structural issues.

Activity 4: Distributing research monies
Two cohorts of grantees were selected from submissions to the URL Shares RFP. Their approved projects were funded through the pooled resources.

SSRC expects to disburse $1.38 million for grantees selected in cohorts 1 and 2 by the end of December 2019. These teams had immediate access to CrowdTangle and Ad Library and were given synthetic data to develop code in advance of the URL Shares Light release.

At the outset of the project, funders expected that SSRC would be responsible for peer review and financial administration, and SS1 would lead the facilitation of data access with Facebook. During the project, Facebook and SS1 identified data sources (CrowdTangle and Ad Library) for offer outside the SSRC sponsored research and peer review approach. By the end of the project, SS1 and Facebook shifted the request process for URL Shares away from SSRC, maintaining a path for data access without the $50,000 grants advertised in the July 2018-September 2019 project period. SS1's August 27, 2019 blog[24] noted the completion of the one-

---

[23] The URL Shares data are described here: https://s3.amazonaws.com/ssrc-cdn2/5c5dbe0b2d752.pdf. The URL Shares Light dataset excluded demographic information and exposures data, as described in SS1's January 9, 2019 blog.

[24] The August 27, 2019 blog reads: "We are grateful for the initial support from the Social Media and Democracy Foundation Funders and the Social Science Research Council to our project. Their efforts over this planned one year term have been integral to setting this important work in motion and we thank them. As the organizations responsible

year project and confirmed that SS1 and Facebook would continue working together on URL Shares access.

Evaluation Findings

We have developed a theory of change to specify the roadmap for data access from Facebook and independent research on the impact of social media on democracy. We use this as an analytic framework to develop the findings that follow. We present six findings, noting how actual conditions deviated from the expected course, what caused the deviation, and the effects we observed.

Figure 4: THEORY OF CHANGE
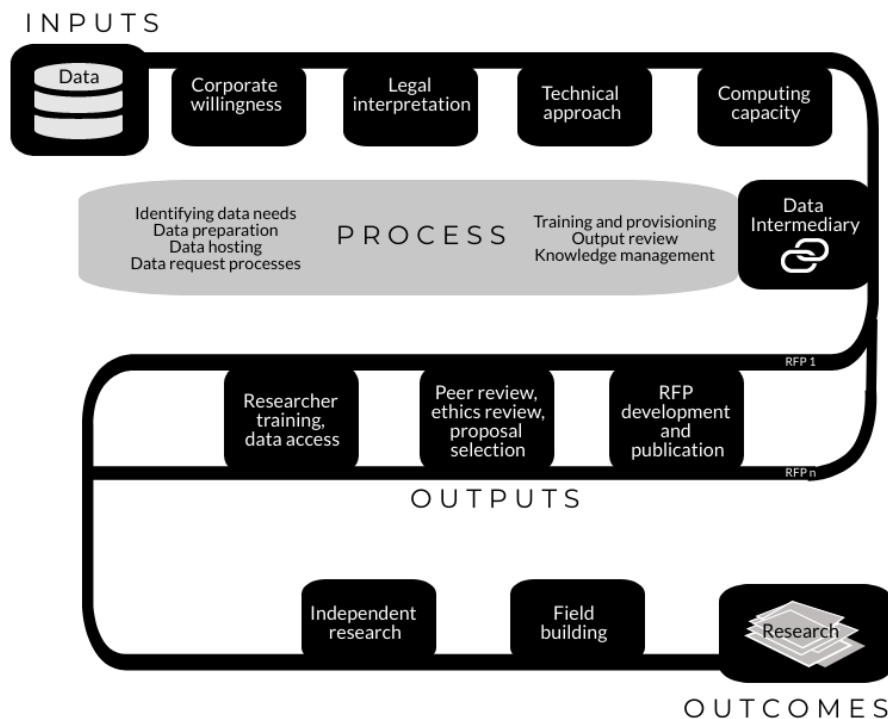
① Bold leadership at a critical time created an unparalleled partnership

This project began with bold action and ingenuity on the part of the Hewlett Foundation, Facebook and academic researchers. Across stakeholders interviewed, people shared a common view of the imperative to lead within their respective organizational contexts. Foundation directors and program officers put their reputations on the line to advocate for the approval of quick, large grants outside of normal grant cycles and in a politically charged environment; Facebook pursued a new type of partnership with researchers, invested in new staff, technology and infrastructure; and leading academics gave up their right to publish in order to play a liaison role between Facebook and peer researchers.

---

for initiating this project and managing the infrastructure for independent academic study and privacy-preserving data access, we look forward to continuing and expanding our efforts.

The partnership was unparalleled in large part because it marked the first time Facebook agreed to independent research on a potentially provocative topic: how social media – and therefore Facebook itself – influences elections. Prior Facebook research partnerships focused mostly on product development and the company retained the right to review and suppress publication of findings.[25]

② Facebook did not deliver data access within the project's timeline

Facebook's willingness and ability to provide data access are both fundamental to the project's theory of change. Despite its repeated commitments to do so, Facebook failed to provide URL Shares data access to researchers during the project's duration. SS1 and SSRC initially believed that Facebook would provide data by Fall 2018, then mid-2019, but Facebook approved access to the URL Shares Light data after the project's planned completion date.

By the Facebook team's own admission, delivering the URL Shares dataset to researchers in a secure and privacy-preserving way was much more difficult than originally envisioned. Part of the complexity arose because Facebook did not have the technical infrastructure to provide such an enormous amount of data to researchers in a privacy-preserving manner. It's helpful to consider that Facebook reported 2.4 billion monthly active users in its first quarter 2019 report. These users get news and updates from their friends, see and react to content, upload pictures, and are part of groups. The URL Shares data include more than 60 billion public shares of posts on the platform (only including those shared with public privacy settings). The most prominent URL in these data has more than 25 million public shares. The initial URL data set was to include data from January 1, 2017 to February 19, 2019. Preparing this data for grantees required processing more than 50 terabytes per day for just the shares and interaction metrics. Facebook applied filters and aggregated the data to reduce it to just over 16 gigabytes of information in 32 million rows, containing key aggregates and statistics of interest. Despite this data reduction strategy, files are still incredibly large.

In addition to finding ways for researchers to physically access the data, Facebook had to interpret both American and European privacy regulations. In July 2018, Facebook had not yet figured out how to protect users' data in a way that still allowed for credible research. In Fall 2018, Facebook began to pursue differentially private data options with the help of academic and industry consultants.[26] But differential privacy is a relatively new technique and computer scientists are still perfecting its methods. Facebook had to develop, test and reject various approaches before settling on the tabular release method for URL Shares Light in Fall 2019.[27]

In short, Facebook failed to devise data access plans that clearly satisfied GDPR and FTC legal requirements for data access during the project. As a result, the originally advertised URL Shares data were not made available on schedule, causing frustration among researchers and funders and skepticism in the media over Facebook's true dedication to the project.

---

[25] Weise, Karen and Sarah Frier, "If you're a Facebook User, you're also a research subject," Bloomberg News, June 14, 2018 https://www.bloomberg.com/news/articles/2018-06-14/if-you-re-a-facebook-user-you-re-also-a-research-subject

[26] Differential privacy is a mathematical approach to protect privacy by slightly altering statistics. It involves a tradeoff of accuracy for privacy protection. Facebook worked with computer scientists to determine how to add an appropriate amount of noise to their user data to protect personal data in a way that preserved accurate research results. This approach would address concerns that platform users had not provided explicit consent for research uses, as their data would no longer be identifiable. This would prevent violation of Facebook's 2011 FTC consent decree and also address GDPR personal data concerns by reducing the ability for any user to be singled out in the research data.

[27] For example, Facebook worked with a contractor for over a year on an approach that was ultimately rejected.

③ Social Science One mobilized the field but did not influence data delivery

The intermediary model is central to the project's theory of change or strategy to connect researchers with Facebook data. SS1 was expected to function as a liaison or data intermediary, mobilizing the academic community and translating its research questions to the complex data sets within Facebook's control. With inside access to the company's data and policies, SS1 was expected to help motivate Facebook to provide data access.

Interviews and documents suggest that SS1's co-chairs pressured the company to release the data through both informal correspondence and external influence. SS1 updated the funder group almost every month on Facebook data preparations, noting their weekly, often daily, contact with the company. Through regional advisory meetings, SS1 brought Facebook to the table with scholars internationally. SS1 Co-Chair Persily held 12 meetings in 10 countries during the grant period, providing peer scholars with an opportunity to discuss relevant topics and to explain to regulators for the need for clearer data access guidelines.

It is difficult to assess SS1's influence over the timing of Facebook's data releases – the most important output of the project. SS1 struggled to facilitate data access because neither they nor Facebook was prepared for the legal and technical issues encountered once the project started: they had no plans or contingencies for these obstacles and delays. Interviewees note that SS1's initial "in" with Facebook was not as robust as they had been led to believe. From all reports, SS1's ideas and suggestions were taken under advisement by Facebook but were neither binding nor sufficiently influential.

④ Investing in research was premature given the uncertainty of data access

All partners operated on the mistaken belief that data access would be easier and quicker to provide. With 20/20 hindsight, it's easy to see that the technical and legal complexities should have been sorted within Facebook before SS1 and SSRC sought research proposals and awarded grants. Motivated by their grant commitments and a desire to create momentum respectively, SSRC and SS1 pursued proposals from the research community prematurely, before data preparation and access methods were established at Facebook.

The July 2018 release of the project's RFP for URL Shares occurred before company decisions involving privacy and security were resolved and before the infrastructure for data sharing was built. One interviewee described the timing of the first RFP's release as the "decisive failure" of the project. From our analysis, we find that the mistaken belief that Facebook could deliver quickly on its commitment to create privacy-protected data was at the heart of these premature investments of time, money, and reputation across the academic and philanthropic organizations involved. At the launch of the project, none of the parties anticipated the challenges and complications to come, specifically involving compliance with GDPR and the FTC consent order.

As a result, SSRC never realized their ambitious grantmaking plans, having expected 12 full rounds of rolling review, each able to handle 80-100 applications, and 90-200 research grants in total.

⑤ The project's governance created challenging dynamics

The funders did not leverage their resources with good governance. More specifically, there is evidence that they missed three opportunities to build the structure needed for an effective collaboration.

A missed opportunity for the funders to align on values and build trust
The Hewlett Foundation's success in mobilizing the resources to create a philanthropic collaboration was not matched by what one interviewee called the "hard work" to build an alliance based on mutual trust and shared values. "We skipped over discussing values and governance in the rush," noted one interviewee. Experts in philanthropy and successful collaboration note the importance of the "startup" phase of collaborative projects – "When one of the great potential benefits of collaboratives—as well as one of the potential challenges—is the diversity of interests, knowledge, and viewpoints, it's essential to push past politeness and surface the advantages of different approaches from the beginning."[28]

Instead of an intentional, in-person project launch meeting designed to create this kind of mutual understanding and alignment, the funder group co-chairs used email, conference calls and bilateral discussions with individual principals to communicate and solve problems among a diverse set of partners. Funder committee co-chairs put time and effort into the project, but its complexity required a different quality of attention. Interviews revealed that a degree of mistrust grew among the project principals at SS1 and SSRC. Regardless of the root cause of mistrust, we believe that frustration developed into conflict in the project due in part to the funder group's failure to prioritize an early alignment on values and ways of working.

A missed opportunity for the funders to set roles and responsibilities with the principal partners
Weak governance was also evident in the confusing, duplicative and unclear division of roles and responsibilities among the principal partners, SSRC and SS1. Despite details written in the project's founding Memorandum of Understanding (MOU) documents below, evolving and often conflicting roles were identified as problematic by all stakeholders interviewed. "We did it all so quickly and there were differences in the MOUs that left undefined some roles. It wasn't clear who had final responsibility."

Table 4: FOUNDING DOCUMENTS

FOUNDING DOCUMENT	ROLES & RESPONSIBILITIES	DECISION MAKING	SIGNATORIES
Facebook-Funder MOU	Independent scholarly committee to identify, pursue, and fund research projects. Committee has right to publicly report on whether Facebook has provided enough access and data within the parameters described.	Committee includes King, Nelson and Persily. All decisions of the Committee require consensus (i.e., a unanimous vote of all Committee members).	Funders, Facebook

---

[28] https://ssir.org/articles/entry/how_philanthropic_collaborations_succeed_and_why_they_fail

Funder Collaborative MOU with SSRC	SSRC legally responsible for managing and carrying out the activities of the Project and has "full authority and control over all funds." SSRC will engage the work of SS1; SSRC will collaborate with SS1 on RFPs; SS1 will review data and define and write RFPs with SSRC, peer review process and proposal selection and communications.	Steering committee includes Nelson, King and Persily SS1 Co-Chairs King and Persily responsible for decisions with input from commission members. SS1 co-chairs make final decisions on proposals	Funders, SSRC
Facebook-Researcher Agreement	The Social Science Research Council ("SSRC") will administer the Research Projects in accordance with the terms of a separate agreement between Facebook and SSRC.	No mention of the committee or Social Science One	Facebook, 17 grantees and their institutions

All groups labored on these and other agreements: the funder group co-chairs worked for months to produce and agree upon the SSRC-funder agreement; SSRC developed researcher agreements and subcontracts with SS1 co-chairs; and SS1 negotiated agreements with Facebook for their own work and for research data access. Yet none of these documents was fully coordinated across all parties.

For example, the founding MOU between SSRC and all funders listed SSRC President Alondra Nelson and SS1 co-chairs Gary King and Nathaniel Persily as principals leading the project, but the document was not signed by either of the SS1 co-chairs or Facebook. Without agreement on project control issues, the principals became increasingly bogged down in disagreements that undermined the project team's ability to adapt to the unexpected challenges presented by Facebook's delays.

A missed opportunity to clarify decision-making rights and predictable process
Disagreements among the principals over who made final decisions and "owned" the project were noted by all interviewees as a sign of the project's dysfunction. The perceived dilemma was twofold: on the one hand, SSRC was asked (and paid) to administer the project, including running the RFP and peer review processes in concert with SS1; on the other hand, SS1 developed the industry-academic partnership model, had expertise in elections and was the only partner with regular access to and visibility into what was happening within Facebook. Although the founding documents referred to consensus as the decision-making rule to be used by the principals, each assumed their role was to act as final decision maker.

The resulting bottlenecks in decision making increased inefficiency and frustration among all the partners. Principals reached out to funders bilaterally to identify and discuss problems but lacked a more formal or structured way to resolve disagreements or mediate disputes.

⑥ The foundations did not follow rigorous grantmaking or management practice

Like the impact that weak governance had on the collaboration, weak grant design and management practices also contributed to inefficiency. Although they work together in other instances and collaborations, interviewed funders noted that they did not follow their typical grantmaking practices in this project.

Evidence that suggests a lack of rigorous grant design and management includes:

- No common name for the project and the use of different terminology by funders (the SSRC-SS1-Facebook project), SSRC (The Social Data Initiative), SS1 (a partnership with Facebook), Facebook (Election Research Commission);
- Poor communication and milestone setting, as all partners continued to create and finance the infrastructure for research, regardless of Facebook's failure to release original data as planned or develop other datasets for release;
- Little attention to value for money as funds supported duplicative activities[29] by the primary and sub-grantee organizations;
- Continued public communication that progress was afoot when almost all stakeholders interviewed described the project as failing in its most important output – the release of data that would be used to generate new research;
- No evidence of clear expectations set for SS1 by SSRC, indicated by a lack of an agreed upon MOU between the two principal organizations and lack of agreed upon milestones or regular project reports on the outputs produced for money spent.

While understandable, the arms-length approach that the project's funders took to this project was arguably a mismatch for the high-stakes experiment.

Takeaways and Considerations

Taken together, the evaluation findings suggest that this project was not set up to succeed for several reasons. Many of the stakeholders we interviewed shared their view that Facebook's inability to provide access to the data originally defined in the July 2018 RFP was a clear indication of the project's failure. Others saw success in the progress made on access terms, legal agreements, ethical review process, and data preparation.

With the breadth and depth of activities implemented through this project – global outreach, dozens of scholars on committees, enormous engineering and policy investment by Facebook – focusing on results in a 20-month period seems unrealistic. We join others in hoping that results will accrue for years, and we advise funders and observers to take a longer view.

While players may change over time, with Facebook seeking more or different data intermediaries to study other topics, this project demonstrated that the company would invest in processes to support independent research. Lessons learned may be applicable to other private sector companies with data valuable for research. Or this project's lessons may be overcome by events if the regulatory environment shifts. Regardless, breakthroughs have been made and more independent research is coming that would have been impossible without the growing pains experienced here.

Much work remains to be done. We offer two takeaways we hope will inform the project's stakeholders as they reflect on this experience and decide whether and how to continue investing in this important domain.

---

[29] Both organizations had communications and legal contracts, and both conducted proposal review and research management activities for selected applicants.

Takeaway 1: Outside learning is always instructive

The most common refrain in our interviews was that this project was "new, unchartered and unprecedented" and that there was therefore no roadmap for the partners to use to achieve their hoped-for results. While industry-academic partnerships in other industries may not be entirely replicable, our first takeaway is that there is always something instructive to gain from outside evidence and experience.

There are many data intermediary models for academics to use sensitive, restricted data from private companies for independent research. We share examples of intermediaries with a range of organizational structures in Appendix III. Across these examples, we find two lessons for interested stakeholders to consider when designing future projects. First, intermediaries need agency to provide data to researchers. This project lacked agreements with explicit terms about when, whether, and how Facebook and SS1 would interact. Second, small steps and demonstration projects permit companies and intermediaries to test and refine procedures before scaling. Stakeholders in this project ambitiously anticipated hundreds of research grants with a global reach while lacking a set of indicators that would show Facebook's progress and a plan to learn intentionally, test and then expand.

There are multiple pathways to support academic research, ranging from in-house approaches like LinkedIn to a multi-firm collaboration like the Health Care Cost Institute. The Partnership for AI is an even broader multi-firm collaborative that was mentioned during an interview: this partnership is a credible, independent organization balancing industry and research interests exploring how artificial intelligence technologies can improve the quality of people's lives and can be leveraged to help humanity. With more than 90 funders (including many tech firms and non-profits), it is unlikely for individual companies to skew its research agenda. The volume and complexity of Facebook's data is a non-trivial challenge, but such examples suggest that access paths are possible even in highly regulated contexts.[30] We describe roles for data intermediaries and examples of their different organizational structures in Appendix III.

Takeaway 2: "Big bets" require more planning, structure and management

The use of large, multi-million-dollar gifts to advance major social change is an increasingly popular trend and target of study in philanthropy. The role of philanthropy as "society's risk capital" is manifest more and more, as foundations use big bet competitions to mobilize non-governmental organizations, academics and private sector partners to come up with their best ideas to solve some of the world's most pressing issues.[31]

The funding collaborative to create an industry-academic partnership that would generate both data access and independent research on social media's impact on democracy is such a big bet. All funders interviewed noted the risk and uncertainty surrounding their investment. Our final takeaway is that this type of uncertainty and risk taking should trigger more attention to structure, alignment and leadership than is typical of business as usual grants. We find outside

---

[30] In addition to these intermediary models, the role of the federal government could be explored. The National Science Foundation's Social, Behavioral, and Economic Sciences Directorate could provide a model of partnership: its research administration and rigorous compliance structure could make it an attractive process intermediary. Also, the possibility of drawing social media data under the strong protections of the Confidential Information Protection and Statistical Efficiency Act (CIPSEA), with its stringent fine and prison penalty provisions, could be desirable and technically viable through the Federal Statistical Research Data Center network.

[31] https://ssir.org/supplement/unleashing_philanthropys_big_bets_for_social_change

evaluations and case studies that have reached similar conclusions about the need to invest in good governance, suggesting that high stakes collaborations may require more not less structure to succeed.[32]

Grantcraft is an opportunity for funders to create structure to support partners who are managing uncertain and high stakes collaborative projects. The reliance on the King and Persily model without obvious consultation of outside models or experts with contrarian views seems to have limited the foundations' consideration of alternative approaches. To match the experimental nature of the endeavor, a fit for purpose approach to designing, measuring progress and managing this grant might have included:

- A deliberate consultation of experts with different views on the model to assure that other industries and experiences were considered before implementing this one;
- Clear criteria for what it meant to "test" or "pilot" the approach defined in the King and Persily article with a meaningful dialogue among partners and funders about how to adapt timelines and spending as the testing progressed;
- An agile planning and project management approach that centers on learning, iteration and adjustment rather than a focus on the production of a research process and products as the main metrics of progress;
- Prioritizing consistent, transparent internal communication rather than investing in external communication expenses and activities.

Concluding Remarks

Judging by the incessant media coverage, the story of how Facebook provides access to its data, protects people's privacy and contributes to the understanding of its own impact on society is far from over. This evaluation was neither designed nor conducted to assess Facebook's commitment or the many factors that influence its behavior.

The evaluation was also not designed to conduct a comprehensive analysis. Operating with a small sample and quick timeline, we heard varied perspectives about the project's critical challenges. We conclude our evaluation with a set of questions that may prove useful for the funders and principals to consider before deciding next steps, and for outside stakeholders to think about before investing in similar efforts.

1. How can funders support development of open standards and open source tools to enable secure, responsible data access? Many innovations are project or data source specific or held as a company's intellectual property. What incentives can spur development of practical applications?

2. We were struck by the trust the foundations and Facebook seemed to bestow upon individuals from elite academic institutions to solve a challenging societal problem without evidence that they considered alternative approaches or developed careful plans. Do these indications of an elite philanthropy need to be reflected upon and addressed?

---

[32] "ClimateWorks Foundation: Lessons in Leadership and Learning," February 2016, PFC Social Impact Advisors; Powell, Alison, Susan Wolf Ditkoff & Kate Hassey, "Value of Collaboration research Study: Literature Review on Funder Collaboration," https://www.bridgespan.org/bridgespan/Images/articles/how-funder-collaborations-succeed-and-fail/bridgespan-value-of- philanthropic-collaboration-study-literature-review.pdf

3. One of the notable outputs of Social Science One's activities is the mobilization of academics across many countries, who share an interest in using data generated through social media to study important questions. Can philanthropy help to build this community and contribute to its cohesion and effectiveness as an advocate for new standards for data access?

We hope the findings of this report add to other evaluations, case studies and articles that highlight the value that intentional leadership, strategic clarity and governance can play in motivating successful collaboration. We find the need to invest in these requirements to be even greater when funders identify a "big bet" where the roadmap to the finish line is unclear.

By way of a final conclusion, we offer these six recommendations for the evaluation's audience:

1. In multi-stakeholder collaboration, leverage bold leadership, resources and partnership with good governance, insisting on clear roles and responsibilities, decision making rights, timelines and contingency plans.
2. Verify company procedures for secure data access that meet legal requirements before committing to a multi-partner project.
3. Defend the nascent international coalition of scholars seeking transparency and expanded access to social media data.
4. Prepare data access protocols and contract templates before users are solicited and chosen.
5. Assure grant design, management and measurement practices match the experimental nature of a 'big bet' investment; go small and slow to achieve successful scale in the long run.
6. Seek outside and contrarian views of organizational models and use external learning to weigh, test and choose among alternative paths.

As of January 2020, Facebook and Social Science One have released the URL Shares data set. The release includes exposure data: views and clicks for each URL that are broken down by month, country, age, gender, and in the U.S., by political page affinity. This new differentially private table contains more accurate information than the previous release while still protecting individual privacy.

Appendix I: Sources

Arnold Ventures. (April 10, 2018). "Laura and John Arnold Foundation Supports Effort to Study Social Media's Impact on Elections." [Press release]. Retrieved from https://www.arnoldventures.org/newsroom/laura-and-john-arnold-foundation-supports-effort-to-study-social-medias-impact-on-elections/.

Borchers, Callum. (October 5, 2018). "After Facebook Infamy, Crimson Hexagon Is Ditching 'The Coolest Company Name in The World'." Retrieved from https://www.wbur.org/bostonomix/2018/10/05/facebook-infamy-crimson-hexagon-ditches-name.

Brown, Eileen. (January 17, 2017). "Crimson Hexagon stores one trillion social media posts for customer analytics." ZDNet. Retrieved from https://www.zdnet.com/article/crimson-hexagon-stores-one-trillion-social-media-posts-for-customer-analytics/.

Burns, Tobias. (June 5, 2015). "JPMorgan Chase Institute: Think Tank or Industry Advocate?" U.S. News and World Report. Retrieved from https://www.usnews.com/news/articles/2015/06/05/jpmorgan-chase-institute-think-tank-or-industry-advocate.

Callahan, David. (October 4, 2017). "Banking Against Urban Poverty: Inside JPMorgan Chase Philanthropy." Inside Philanthropy. Retrieved from https://www.insidephilanthropy.com/home/2017/10/4/banking-against-urban-poverty-inside-jpmorgan-chase-philanthropy.

Daniels, Alex. (April 9, 2018). "Facebook Gives Scholars Access to Data to Study Its Role in Elections." Chronicle of Philanthropy. Retrieved from https://www.philanthropy.com/article/Facebook-Gives-Scholars-Access/243070.

Edelson, Laura, Shikhar Sakhuja, Ratan Dey, and Damon McCoy. (February 12, 2019). "An Analysis of United States Online Political Advertising Transparency." Retrieved from https://arxiv.org/pdf/1902.04385.pdf.

European Commission. (October 16, 2018). "Roadmaps to implement the Code of Practice on disinformation." Retrieved from https://ec.europa.eu/digital-single-market/en/news/roadmaps-implement-code-practice-disinformation.

Federal Trade Commission. (August 10, 2012). "FTC Approves Final Settlement with Facebook." [Press release]. Retrieved from https://www.ftc.gov/news-events/press-releases/2012/08/ftc-approves-final-settlement-facebook.

Gaynor, Martin. (2012). "Introducing the Health Care Cost Institute" Health Management, Policy and Innovation, 1(1): 35-36.

Glaisyer, Tom. April 9, 2018). "Democracy Fund and Omidyar Network Support Independent Analysis of Facebook's Role in Elections." [Blog post]. Retrieved from https://www.democracyfund.org/blog/entry/df-on-support-independent-analysis-of-facebook-role-in-elections.

Goldman, Paula. (April 19, 2018). "A Partnership for More Responsible Tech: Assessing Facebook's Role in Elections." [Blog post]. Retrieved from https://medium.com/positive-returns/a-partnership-for-more-responsible-tech-assessing-facebooks-role-in-elections-9d51d755305a.

Grind, Kirsten. (July 20, 2018). "Facebook Suspends Analytics Firm on Concerns About Sharing of Public User-Data." Wall Street Journal. Retrieved from https://www.wsj.com/articles/facebook-probing-how-analytics-firm-shares-public-user-data-1532104502.

Hewlett Foundation. (April 9, 2018). "Hewlett, Knight, Koch Foundations, with other funders, will support independent research on Facebook's role in elections and democracy." [Press release]. Retrieved from https://hewlett.org/newsroom/hewlett-knight-koch-foundations-with-other-funders-will-support-independent-research-on-facebooks-role-in-elections-and-democracy/.

Hotham, Tristan (@TristanHotham). "A thread on @SocSciOne's @crowdtangle through which all analysis of public @facebook data will be operated after 4th Sept. Overtly restrictive, it warrants the death of masters/undergrad FB study, endangers non-political study, and promotes the illegal scraping of data." August 14, 2019, 3:10 PM. Tweet.

Hutchinson, Andrew. (January 30, 2019). "Facebook Expands CrowdTangle Analytics Platform Access to Academics and Researchers." Social Media Today. Retrieved from https://www.socialmediatoday.com/news/facebook-expands-crowdtangle-analytics-platform-access-to-academics-and-res/547140/.

King, Gary, and Nathaniel Persily. (2018). "A new model for industry-academic partnerships." Cambridge Mass. https://gking. harvard.edu/partnerships.

King, Gary, and Nathaniel Persily. (2019). "A new model for industry-academic partnerships." Cambridge Mass. https://gking. harvard.edu/partnerships.

King, Gary, and Nathaniel Persily. (2019). "A new model for industry-academic partnerships." PS: Political Science and Politics. Publisher's Version Copy at http://j.mp/2q1IQpH.

Knight Foundation. (April 9, 2018). "Hewlett, Knight, Koch Foundations, with other funders, will support independent research on Facebook's role in elections and democracy." [Press release].

Retrieved from https://knightfoundation.org/press/releases/independent-research-on-facebook-role-in-elections-and-democracy.

Kozlowska, Hanna, Dave Gershgorn and Sarah Todd. (March 29, 2018). "The Cambridge Analytica scandal is wildly confusing. This timeline will help." Quartz. Retrieved from https://qz.com/1240039/the-cambridge-analytica-scandal-is-confusing-this-timeline-will-help/.

Kubzansky, Mike. (April 19, 2018). "A Partnership for More Responsible Tech: Assessing Facebook's Role in Elections." [Blog post]. Retrieved from https://www.omidyar.com/blog/partnership-more-responsible-tech-assessing-facebook%E2%80%99s-role-elections.

Leetaru, Kalev. (April 12, 2018). "Is Facebook's New Academic Initiative Even More Frightening Than Its Own Research?" Retrieved from https://www.forbes.com/sites/kalevleetaru/2018/04/12/is-facebooks-new-academic-initiative-even-more-frightening-than-its-own-research/#605832bd478d.

Lerner, Josh, and Leslie Jeng. (May 2016). "Making Private Data Accessible in an Opaque Industry: The Experience of the Private Capital Research Institute." American Economic Review: Papers and Proceedings 106, no. 5: 157–160.

Mervis, Jeffrey. (September 24, 2019). "Privacy concerns could derail unprecedented plan to use Facebook data to study elections." Science Magazine. Retrieved from https://www.sciencemag.org/news/2019/09/privacy-concerns-could-derail-unprecedented-plan-use-facebook-data-study-elections.

Mozilla. (April 29, 2019). "Facebook's Ad Archive API is Inadequate." [Blog post]. Retrieved from https://blog.mozilla.org/blog/2019/04/29/facebooks-ad-archive-api-is-inadequate/.

Overly, Steven. (April 9, 2018). "Facebook to create independent election commission to study social media's impact on democracy." Retrieved from https://www.politico.com/story/2018/04/09/facebook-social-medias-impact-on-democracy-508776.

PFC Social Impact Advisors. (February 2016). "ClimateWorks Foundation: lessons in Leadership and Learning." Retrieved from https://hewlett.org/wp-content/uploads/2017/11/ClimateWorks_Foundation_Case_Study.pdf

Powell, Alison, Susan Wolf Ditkoff and Kate Hassey. (October 2018). "Value of Collaboration Research Study – Literature Review on Funder collaboration." The Bridgespan Group. Retrieved from https://www.bridgespan.org/bridgespan/Images/articles/how-funder-collaborations-succeed-and-fail/bridgespan-value-of-philanthropic-collaboration-study-literature-review.pdf

Powell, Alison, Susan Wolf Ditkoff, and Fay Twersky. (July 10, 2019). "How Philanthropic Collaborations Succeed, and Why They Fail." Stanford Social Innovation Review. Retrieved: https://ssir.org/articles/entry/how_philanthropic_collaborations_succeed_and_why_they_fail.

Rosenberg, Matthew. (July 25, 2019). "Tool Facebook Built to Fight Disinformation Doesn't Work as Advertised." New York Times. Retrieved from https://www.nytimes.com/2019/07/25/technology/facebook-ad-library.html.

Roush, Wade. (November 12, 2008). "How Crimson Hexagon Translates the Blogosphere's Babel into Wisdom." Retrieved from https://xconomy.com/boston/2008/11/12/how-crimson-hexagon-translates-the-blogospheres-babel-into-wisdom/."

Serrato, Ray (@raymserrato). "The restrictions on access are particularly puzzling because other @crowdtangle partners - news outlets, media startups, etc. - don't seem to have these kinds of prerequisites imposed on them, though the data they have access to is no different from that offered by @SocSciOne." August 15, 2019, 5:42 AM. Tweet.

Social Science Research Council. (2018). "To Secure Knowledge: Social Science Partnerships for the Common Good." Report retrieved from https://s3.amazonaws.com/ssrc-static/tsk/SSRC+To+Secure+Knowledge.pdf.

Staples, Kirsten. (November 1, 2018). "Facebook offers CrowdTangle free to news industry." News & Tech. Retrieved from https://newsandtech.com/in_this_issue/facebook-offers-crowdtangle-free-to-news-industry/article_3122235c-f1a1-11e8-8357-eb9789ea67ab.html.

Sullivan, Mark. (March 28, 2019). "Facebook expands ad transparency beyond politics: Here's what's new." Fast Company. Retrieved from https://www.fastcompany.com/90326809/facebook-expands-archive-ad-library-for-political-ads-heres-whats-new.

Weise, Karen and Sarah Frier. (June 14, 2018). "If you're a Facebook User, you're also a research subject," Bloomberg News. Retrieved from https://www.bloomberg.com/news/articles/2018-06-14/if-you-re-a-facebook-user-you-re-also-a-research-subject.

Wylie, Christopher. (October 4, 2019). "How I Helped Hack Democracy." New York Magazine Book Excerpt. Retrieved from http://nymag.com/intelligencer/2019/10/book-excerpt-mindf-ck-by-christopher-wylie.html.

Zuckerberg, Mark. (April 9, 2018). In Facebook post. Retrieved from https://www.facebook.com/zuck/posts/10104797374385071.

Appendix II: Interview Participants

David Barkan	Independent Consultant	Barkan Consulting
Mark Bastos	Associate Professor	City University of London
Jesse Blumenthal	Director	Charles Koch Institute
Kelly Born	Funder Chair and Program Officer (former)	Hewlett Foundation
Sarah Drinkwater	Director	Omidyar Network
Tom Glaisyer	Managing Director	Democracy Fund
Sam Gill	Vice President for Communities and Impact	John S. and James L. Knight Foundation
Daniel Goroff	Funder Vice Chair and Vice President and Program Director	Alfred P. Sloan Foundation
Elizabeth Hess	Director	IQSS, Harvard University
Gary King	Professor and Co-Chair Social Science One	Harvard University, Social Science One
Larry Kramer	President	The William and Flora Hewlett Foundation
Vidya Krishnamurthy	Director of Communications	Hewlett Foundation
Solomon Messing	Research Scientist	Facebook
Chaya Nayak	Public Policy Research Manager	Facebook
Alondra Nelson	President	Social Science Research Council
Nathaniel Persily	Professor and Co-Chair Social Science One	Stanford University, Social Science One
Jason Rhody	Director	Social Science Research Council
Elliot Schrage	Vice President	Facebook
Daniel Stid	Director	Hewlett Foundation
Paul Waters	Senior Associate	Democracy Fund

Appendix III: Alternative Models

Data intermediaries are organizations that facilitate data sharing and access between data holders and researchers. They typically handle sensitive or restricted data from multiple sources, often as a service across many institutions. Data intermediaries have standard data request and review processes to offer efficient and responsible data access.[33]

Data intermediaries may identify and pursue data sources, sponsoring data collection where necessary; develop and manage agreements; enforce negotiated terms of use; ingest and harmonize data; regularly assess the adequacy of their governance and data models; act as a trusted third party to link data; coordinate screening, training, and monitoring of researchers; coordinate output review; gather tools and models that make analysis more efficient; and provide technical assistance to data holders.

Sound governance in a data intermediary requires clarity in scope and authority. Many intermediaries establish a board of directors, scientific advisory board, and policy board to work with their leadership. They also establish clear agreements with their host institutions (especially with regard to security, liability, and conflict of interest policies). Intermediaries are transparent about how they address legal, privacy, and security issues involving data.

The table below shows examples of six data intermediaries existing in for-profit and non-profit entities. They are funded in a variety of ways. Descriptions of these organizations follow.

MODEL	DATA TYPE	PLACEMENT	FUNDING	PUBLICATION RIGHTS
YODA	Clinical trials	University	Companies, feds, and foundations	Yes
IRIS	University HR and contract data	University	Users and foundations	Yes
PCRI	Private capital data from companies and aggregators	Non-profit	Foundations	Yes
HCCI	Healthcare costs from major insurers	Non-profit	Companies, users, and foundations	Yes
LinkedIn	Employment history and networks from platform users	Within company	Company	Yes
JPMC Institute	JP Morgan Chase consumer, business, and investor accounts	Within company	Company	No external academic access

Intermediaries need the agency to represent private company partners
The Yale University Open Data Access (YODA) Project is an independent intermediary that makes decisions regarding the design of the data request process, the criteria for access, approval or rejection of requests, and secure data access. Johnson & Johnson (J&J) joined

---

[33] The Administrative Data Research Facilities (ADRF) Network has been building support and momentum for data intermediaries to support academic research in the social sciences. A summary of intermediary interviews can be found at https://repository.upenn.edu/cgi/viewcontent.cgi?article=1002&context=admindata_reports.

YODA in 2014 and is one of its largest industry partners. YODA has full agency to facilitate access to J&J clinical trial data for pharmaceutical, medical device, and consumer products. Before YODA lists any trial as available, J&J confirms the location of the data and the electronic format, as well as any agreements with collaboration partners to ensure each trial listed can be shared. Like the Facebook project, YODA evolved through firm partnership, and a steering committee comprised of an independent group of leaders in the fields of clinical research and biomedical ethics. Unlike this project, YODA sought input on the process and governance structure from others in industry, regulators, and even the general public.

The Institute for Research on Innovation and Science (IRIS)[34] is an intermediary at the University of Michigan's Institute for Social Research. IRIS connects researchers to data from 35 university systems to facilitate understanding of the public value of research. IRIS curates sensitive personnel and student information from university human resources, sponsored projects, and procurement systems, and manages the research request and data access processes, offering virtual enclave access for the university data and enabling linkages to government data through the Federal Statistical Research Data Center network. IRIS demonstrates that multiple access paths are possible, and that an intermediary can represent the interests of the data generating institution when negotiating access with the federal statistical system.

The non-profit Private Capital Research Institute (PCRI) was established to improve data access for academic research on how venture capital or private equity firms invest in companies. PCRI has been building a comprehensive and authoritative private capital database since 2010 to support independent research. PCRI manages the data preparation, research request, and data access processes. Like this Facebook project, researchers have access to a query server with anonymized data; they may not download or view individual data entries.

Intermediaries start small and scale up
The Health Care Cost Institute (HCCI) is an independent, nonprofit entity providing data on US health care costs. Founded in 2011, HCCI built a unified, standardized database through industry partnership with Aetna, Humana, Kaiser Permanente, and (until this year) United Healthcare. During its early years, HCCI facilitated a small number of independent research projects using custom constructed databases for each project, in order to initiate research and develop experience and knowledge with the data. Unlike the Facebook project, HCCI started small and learned from those efforts. HCCI harmonizes and curates data from the private insurers and now has a partnership to offer federal health insurance data, licensing data to universities, government research agencies, and non-profit organizations.

The LinkedIn Economic Graph provides an example of a research partnership program run by a company. LinkedIn manages an internal process to select researchers to study the company's complex network data compiled in the Economic Graph. Launched in 2017, the Economic Graph Research Program demonstrates that a technology company can support independent academic research. They provide access to aggregated or de-identified data through a monitored environment, on a secure network, with the approval of LinkedIn, but the results of the analyses can be published without company approval. They selected ten teams from the 200 submissions to their first call for proposals. They note that results produced by the research

---

[34] IRIS is one of the original ADRF Network intermediaries. Their comprehensive governance approach is described on their website (https://iris.isr.umich.edu/about/governance/). Their founding documents state the IRIS vision, guiding principles, organizational structure, and many other governance details.

teams at Indiana University and MIT prompted the company to form longer partnerships with their institutions.

The JPMC Institute was launched in 2015 as "a global think tank dedicated to delivering data-rich analyses and expert insights for the public good," according to their website. However, unlike traditional think tanks, it is not a non-profit but instead a division of JPMC, obtaining extracts of account data for research purposes. JPMC Institute studies are quicker and deeper than many other economic research efforts because the company's data are available without the significant lags that government source suffer and offer the depth of connections across 70 million retail consumers, 2.5 million small businesses, and 44,000 institutional investors. Like the Facebook project, a company with rich consumer data is trying to produce valuable insights while heeding the expectations and requirements of its customers and regulators. However, JPMC Institute does not have an external research request process that permits independent research.