



# Mozilla's Data Futures Lab

## Building a Human-Centric Data Economy

PRESENTED BY

## Siegel Family Endowment

We are a foundation focused on understanding and shaping the impact of technology on society.

### ABOUT SIEGEL

Siegel Family Endowment employs an inquiry-driven approach to grantmaking that is informed by the scientific method and predicated on the belief that philanthropy is uniquely positioned to address some of the most pressing and complex issues facing society today. Our grantmaking strategy positions us to be society's risk capital. We support high quality work that will help us derive insights to timely questions and has high potential for future scale. Our focus is on organizations doing work at the intersection of learning, workforce, and infrastructure. We aim to help build a world in which all people have the tools, skills, and context necessary to engage meaningfully in a rapidly changing society. Siegel Family Endowment was founded in 2011 by David Siegel, co-founder and co-chairman of financial sciences company Two Sigma.

### OUR FOCUS ON INFRASTRUCTURE

The internet and emerging technologies have changed the way we engage with one another and our institutions, and reshaped elements that underpin our civil society. The choices and values driving the building and maintenance of our infrastructure shape our future. It's essential that we rethink how to define, design, govern, and fund it. We apply our [multidimensional framework](#) for infrastructure to deliver positive community outcomes and address the urgent challenges facing American society.

Strong communities require just and equitable infrastructure – and just and equitable infrastructure require engaged communities. Yet, as a result of excluding the social and physical spheres, inequities persist in the design of digital services and accessibility of those services. How might we increase the meaningful participation and agency of communities to shape the development, deployment, use, and stewardship of the digital infrastructures upon which they rely?

### COMMUNITY-DRIVEN DIGITAL INFRASTRUCTURE

Inequities persist not only in access to effective digital services, but in the quality of peoples' experiences online and the ways in which those experiences shape offline communities and physical spaces. These disparities – in connectivity, reliable information, knowledge, opportunities, resources, government services, and more – result in a lack of connectedness and participation in vibrant, healthy communities. We prioritize practices that empower communities in co-creating sustainable digital, social, and physical infrastructures centered on their unique values, needs, and aspirations.



## KEY TAKEAWAYS

- **Testing and scaling data models centered on human agency requires an alternative ecosystem of support** beyond traditional venture-capital financing and accelerators. Building community and offering wraparound services are critical ingredients, along with cohort-based incubation, and partnerships that involve a range of stakeholders and voices to nurture new ideas and tools.
- Creating a fairer data economy **requires experimentation**. Systems-change work of this sort is deeply needed, but it also takes time, patience, and the possibility that an approach might not work or scale as expected.
- Clear outcomes are not always immediately visible in systems-change work. This makes it vitally important that systems-oriented organizations work to **evaluate the long-term impact of the approach on the larger ecosystem and community** and the penetration of the values in the larger world, rather than the short-term outcomes of individual projects alone.
- Technical projects aimed at developing new modes of thinking and redistributing power **require a new financing ecosystem**. **Philanthropic organizations are vital partners, but impact investors and aligned capital** also offer important expertise, aligned incentives, the ability to bring stakeholders together to build community, and flexible funding to invest in new models for data stewardship.

## About Grantee

[Data Futures Lab](#), a project of [the Mozilla Foundation](#), served as an experimental space for developing new approaches to data stewardship between 2021-2024. It provided funding, project incubation, scaffolding for collaboration, convening around emerging ideas, and a place to workshop approaches to data governance. By bringing together innovators, capital, and community, DFL hoped to give greater control and agency to people, rather than corporations or systems. The project has since ceased as a standalone initiative, and its success has inspired the broader Mozilla Foundation to take its approach to funding and tailored support so that more promising sociotechnical projects can flourish over a longer term.

## Executive Summary

FOR MANY PEOPLE, a desire for better approaches to data stewardship is a brief, passing thought. Perhaps a customer begins to worry after a healthcare company announces a hack of customer data. Maybe someone tries to opt out of cookies on a website only to give up, exasperated at the multiple screens and clicks required. Perhaps a consumer wonders why an online restaurant reservation requires a birthday. It's at such moments that we often think, "There must be a better way." These concerns are only heightened with developments in AI fueling new uses for our data.

The goal of Mozilla's Data Futures Lab was to build tools and communities designed to offer such better ways—ways that give greater control and agency to people. In so doing, the project sought to advance policies, approaches, and tools that prioritize data co-creation, ownership, and stewardship in ways that benefit a diverse and inclusive public.

The project provided funding, held convenings, and built coalitions to incubate tools that address common data stewardship challenges—for example, data governance or data donation. Technologists came to the work from different vantage points and in service of distinct constituencies. Examples of data donation projects that reflect these different constituencies include the following:

- A [project](#) to allow users to donate anonymized data to benefit diabetes research
- A [project](#) to create a trustworthy data space for users to decide when, how, and with whom to share mobility and socioeconomic data with third parties under ethical terms
- A [project](#) to leverage disparate data to build an interactive data tool to understand broadband inequalities in the United States

The Lab aimed to invest in approaches that look at problems in new ways; that could serve and elevate individual communities in important ways; that could be applied in different contexts to benefit more people; or that could create a deeper sense of community and connection among stakeholders. Like its commercial counterparts, Data Futures Lab didn't expect that all of its projects would successfully spin off from its incubator—though some certainly did. Instead, its ultimate goal was to nurture a human-centered data stewardship *ecosystem*—to change data stewardship from a passing thought of a consumer or a check-the-box exercise of a corporation, in order to shift power in product development, data stewardship decision-making, and data usage.

## Core Elements: What Makes the Program Work?

DATA FUTURES LAB emerged from the belief that the old rules for data governance weren't working. To be sure, the landscape had improved in small ways. New regulations made it easier for users to opt-out of having their data sold to third parties. Technological developments in areas such as AI raised the profile of data stewardship questions in the public consciousness. And tech-based collective actions by workers and social activists offered new ways of organizing. Despite these developments, the power dynamics shaping data stewardship had budged little though. As [an early announcement](#) of Data Futures Lab put it, "We are left to play by industry rules."

The solution? Bring together disparate movements, people, and organizations to build solutions and community that can rebalance power in favor of people. Mozilla Program Officer Miguel Morachimo sums up the work of his organization this way: "The projects that we support are part of a new wave of bets that tried to intervene at the root cause. They're thinking about this question, , 'How can things can be done differently?'"

### ► CREATING SPACE FOR EXPERIMENTATION

FOR DECADES, [Mozilla](#) has developed community and products designed to promote a free, fair, and open internet. Yet for all of Mozilla's success, by the late 2010s, there was a sense among fair internet activists that corporations were winning the war for control of public data. Mozilla's [landscape analysis](#) revealed that many innovators were stymied by lack of resources for experimentation about data stewardship for the public good. Morachimo recalls, "There was a vacuum for collective rights. Companies were extracting and profiting from the data." Mozilla's community saw the need not just to push for better transparency but to address the fundamental reasons why the problem existed in the first place. "There was a need to come up with alternatives, to experiment with different data arrangements," Morachimo recalls.

It was that need for alternative approaches and the development



Shayne Longpre of the Data Provenance Initiative, a Data Futures Lab grantee, speaks at MozFest House Amsterdam



and deepening of a larger ecosystem committed to public-centered data governance that led to the creation of Data Futures Lab as an experimental space. Experimentation occurs in R&D functions within companies, but the motives are different, says former Lab Director Champika Fernando. “Our goal [was] to fund projects that are experimenting with pieces of data stewardship or data governance that the private sector is not incentivized to take on.” As examples, Fernando points to community-building initiatives, practices of government, and other areas where there might not be an opportunity to generate significant revenue or profit.

The Lab model also recognized that iteration and testing are necessary elements of the process. Fernando said, “We tried to create a space where we can test the limits of the idea and how practical it is.” There are few such spaces within the data stewardship community, but Data Futures Lab’s staff say that resourced laboratories such as their own are necessary if a people-centered approach to internet data is to take root.



Former DFL Lab Director Champika Fernando and MozFest Curator Zeina Abi Assy at MozFest House

## DFL PARTNER PROFILE

### The Drivers Cooperative



“**THE PROBLEMS** of the gig economy are well known,” Erik Forman says, citing a long list. Workers labor for piece-rate pay, denying them access to a steady hourly wage. They lack basic worker protections. In the rideshare industry, gig workers must supply their own vehicles and take care of maintenance themselves. In short, Forman says, “The ridesharing companies have successfully displaced all the risk of doing business onto the actors with the least resources—the drivers.” In New York City, those drivers are overwhelmingly immigrant and people of color. “Driving an Uber is the employment option of last resort for New Yorkers,” Forman, a union organizer, says.

It was a desire to change the power imbalance between vulnerable drivers and wealthy, venture capital-backed ridesharing companies, that led Forman to co-found The Drivers Cooperative. The project is structured as a driver-owned cooperative, in which the profits from rides are returned

to the drivers, rather than to for-profit ridesharing companies and their investors. It currently operates in New York City and is expanding to Colorado and Minnesota.

Forman says that signing up workers to become part of The Drivers Cooperative was the easy part. Far more difficult was developing and sustaining the technological platform in ways that were consistent with the needs and goals of the myriad drivers who comprised the cooperative’s ownership. Forman recalls, “After we incorporated as a co-op, then began the white knuckle experience of managing a tech startup without anything close to adequate resources.”

Data Futures Lab has proved an important resource as The Drivers Cooperative confronts these challenges. As a member of the 2022 Data Futures Lab cohort, The Drivers Cooperative delved deep into learning about how the ridesharing industry used data. Forman found that drivers were less concerned about data ownership, a key concern of other projects at Data Futures Lab; the types of data that ridesharing companies typically had access to were not especially sensitive and companies had not tried to monetize those data through sale.

Instead, drivers’ primary concerns were income and employment stability—issues that most drivers initially didn’t see as related to data, but which were in fact intimately linked to the role of data in rideshare platforms’ business models. Forman says that Data Futures Lab was helpful as The Drivers Cooperative worked to illustrate the connections between

those income-related concerns and the larger data economy.

Adopting a cooperative governance model has allowed The Drivers Cooperative to lift up those priorities in its technology. For-profit ridesharing companies use data in pricing algorithms that limit the amount of money that drivers will see, even if those companies now need to abide by minimum wage laws in New York. In contrast, The Drivers Cooperative created a pilot “Economic Security Program” that established a custom algorithm in its payment system for collective benefit, guaranteeing drivers a \$30 minimum wage per hour gross of expenses.

In addition, The Drivers Cooperative set up an elected review board to adjudicate customer service disputes, instead of relying solely on algorithms. It promoted from within, creating a pathway for some drivers to develop skills as data analysts who can analyze calls, generate reports, create dashboards, and work on algorithms to ensure that drivers make their daily minimums.

Forman emphasizes that this work is very difficult. But he hopes that it will serve as the basis for other worker-responsive models that develop ways to harness data for the benefit of those who have been shortchanged by low-wage employers. “I would define the entire project we’ve done as a prototype,” Forman reflects. “We did a lot of learning that can illuminate a future effort with more resources.” He sees the connections that The Drivers Cooperative developed through Data Futures Lab as a key component of that ongoing work.

## Unlike a traditional venture-capital financed accelerator, building community is an explicit goal at Data Futures Lab, rather than a side benefit of incubation.

Fernando and Morachimo say there were significant challenges to cultivating an experimental, laboratory space. The Lab could only support a few projects at a time. There were inevitable failures. Solutions might serve only limited numbers of users or might be employed only in narrow use cases. There was a constant need for more resources—time, money, and expertise among them. But the end result was profound: a growing community of innovators who are designing data stewardship practices, products, and approaches for the public good.

### ► BUILDING COMMUNITY TO CHANGE ECOSYSTEMS

WHEN ASKED WHAT distinguishes Data Futures Lab from other initiatives to support the development of data stewardship tools, Fernando didn't skip a beat—community. “We [had] the ability to fund technology projects, but we also [had] a strong network and

community with the technical know-how to support these projects at their inception and early development,” Fernando said. In other words, wraparound support and connections were just as important as the checks that the Lab issued.

From the beginning, Data Futures Lab developed a theme-based cohort model focusing on topics such as data donation or data governance. That approach ensured that selected projects had a community of builders with similar interests, solving similar problems, and facing common challenges—even if details, areas of expertise, models, and audiences differ between them. In addition, Data Futures Lab offered wraparound support, including formal mentorship for the leads of projects in the incubation cohort, as well as learning sessions around specific topics of shared interest. The Lab also drew on experts and practitioners through its parent organization Mozilla to advise project leads on how to develop an open source community and to provide consultations on technical aspects of development.

#### DFL PARTNER PROFILE

### POSMO



**ROGER FISCHER**, a product manager and entrepreneur, remembers that the Silicon Valley tech scene of the 2010s first appeared “shiny and positive.” It was only when Fischer pulled back the curtain on the data practices that were fueling emerging technologies that he began to see “the dark side.” Numerous companies were developing autonomous vehicles—a race that Fischer describes as “a coming tsunami.” Yet, companies were doing little to protect the reams of mobility data that were created in the process.

Fischer and his colleague Lea Strohm had a different vision for their start up, whose goal was to collect mobility

behavior—that is data about foot traffic, bikes, bus routes, transportation infrastructure, and other areas—in order to lay the foundations for sustainable cities. “We realized that if we wanted to create a mobility tracking startup, we needed to do it right from the get-go.” “Doing it right” meant that both the data producers and the data that they created needed to be protected. Out of that philosophy came POSMO, an ethical data commons for mobility data that Fischer and Strohm created to reward individual contributions fairly and equitably.

The POSMO DataMarket is composed of mobility data collected via two products: Mobility Tracking and VelObserver (Citizen Science). POSMO is structured as a cooperative, with members deciding the entities that they will share their data with. Buyers of the data include cities, universities, and nonprofit organizations that are developing sustainable transportation and mobility solutions, combating climate change in the process.

As a member of the 2023 Data Futures

Lab cohort, POSMO has been able to expand its reach and connections within the data community. Fischer and Strohm found sessions on fundraising especially valuable, and the connection to Mozilla has raised the profile of POSMO in Europe and internationally. With the guidance and support of Data Futures Lab, POSMO has also developed and published ethics criteria for data intermediaries; strengthened its internal governance process; and conducted research about what motivates people to participate in data projects.

There have been pivots and bumps along the way—most notably, a switch in the way that data is collected—but POSMO's founders believe that their association with Data Futures Lab has given them a boost in the work to come. Fischer says, “Sometimes a funder takes a transactional relationship, where you provide something for them. With Data Futures Lab, it was completely different. It really lifted the spirits, to know that you were part of a group working toward the same ends.”

## DFL PARTNER PROFILE

## Te Hiku Media



**WHEN HE TALKS** about his work at [Te Hiku Media](#), Keoni Mahelona uses a comparison that immediately resonates with Indigenous audiences. “Data equals land. They’ve taken our land and now they’re back for our data,” Mahelona says. Like the land that was taken from the Māori people of present day New Zealand, the data that Te Hiku Media preserves, protects, and makes accessible digitally is equally precious: the first-person stories of Māori elders and other community members in their native language. In fact, these stories

are so precious that they are referred to as “treasures” in the Māori tongue.

“We were never going to put our treasures in a place like YouTube,” Mahelona says, citing the inability of external vendors to make use of data in their own Māori language processing models. “We knew how important it was that we put our treasures online in the right way, in terms of the copyright and ownership of the data.” That insight led Mahelona and his colleagues to develop their own platform for housing the interviews and stories, as well as a natural language processing tool to recognize and transcribe speech in the Māori language.

Te Hiku Media received an Infrastructure Grant from Data Futures Lab in 2022 to complete documentation for the [Kaitiakitanga indigenous data license](#) and digital media platform. The project provides examples and resources for other Indigenous communities to write their own data license, enabling communities to maintain guardianship over their data

while making it accessible online.

“Asking permission to use or cite data is a form of respect that is an important part of Indigenous culture,” Mahelona says, referencing his upbringing in Hawai‘i. “That’s what the license allows.”

Interest in the data license has grown, spurring a movement of communities interested in responsible and empowered stewardship of their own data. The support of Data Futures Lab has driven much of that interest, Mahelona says. “People are hearing our story and I don’t think they would have heard our story without Mozilla,” Mahelona reflects. In turn, Mahelona says that Te Hiku is having an impact on how people perceive what data is. He likens Te Hiku’s vision for data to the environment. “In the Indigenous way of thinking, our environment is not an infinite resource from which we can extract value, to enrich ourselves. We personify the environment and treat it as we would want to be treated. It should be the same for data.” That is what Te Hiku is advocating for.

Unlike a traditional venture-capital financed accelerator, building community was an explicit goal at Data Futures Lab, rather than a side benefit of incubation. “We wanted the non-monetary component to be the biggest selling point,” Morachimo says. That means that funded projects had an opportunity to deepen and expand the data stewardship ecosystem, rather than simply serving as one node within it. Fernando said, “The work that they’re doing is not seen as in-competition with other folks here in the Lab or in the community. The goal is to continuously be open and sharing, which just naturally lends itself also to building up the ecosystem and community around it.”

Leaders of funded projects cite the community-building aspect of the Lab as one of its most important features. Roger Fischer, a co-founder of Posmo, an ethical data broker, said, “Data Futures Lab and Mozilla are excellent at building networks and making connections. I also enjoyed the exchange with others in the cohort.” Erik Forman, co-founder of The Drivers Cooperative agrees. “One of the best things about participating in the Data Futures Lab was having a community of practice to connect with.”

### ► REDEFINING SUCCESS

DATA FUTURES LAB’S barometer for success was different from a traditional venture capital-backed incubator. For one, the Lab wasn’t focused on merely maximizing monetary return on its investments. Instead, Data Futures Lab was more interested in finding sustainable

models for human-centered data stewardship. As Morachimo put it, “We [were] not assuming that every grantee will become a market leader.” He continues, “We were hoping that the tools that grantees developed and the practices that they established can be important models and precedents for others.”

DFL grantees also experimented with new models for data that are generating important proof points and templates for others. [Te Hiku Media](#), a small non-profit radio station on New Zealand’s most northern tip, is one such example. Te Hiku used 300 recorded hours of spoken Māori language to build language technology tools, including speech-to-text transcription using open source tools, such as Mozilla’s DeepSpeech. Te Hiku’s leaders were concerned about placing interviews and stories of Indigenous people in their natural language on sites that could use the data to train AI models or for other extractive ends. In response, Te Hiku created [a new data license](#), guided by its deep values in community and collective ownership, ensuring that anything created using its data is owned by and directly benefits the Māori people.

In another example, [POSMO](#), another Data Futures Lab grantee created a new model as an ethical data broker. The project asks people to contribute their data about their daily movements in a city and established a data market for third parties to access those data. The project created an ethical review board to evaluate all incoming requests for data against the agreements that people who donated their data made and the values that the project espouses.





Lawrence Adu-Gyamfi of GhanaNLP presents at the Data Futures Lab Showcase at MozFest House Kenya

The money earned from selling those data to third parties is given partly to the cooperative and can be redistributed to the contributors according to their participation. Regardless of whether this particular grantee is successful as a venture in its own right, it is offering a new, people-centered approach to data donation and brokering that can be leveraged by others.

### ► BUILDING AN ALTERNATIVE FINANCING ECOSYSTEM

TECHNICAL PROJECTS—corporate or otherwise—are notoriously labor- and time-intensive, and often take a long time to complete. Technical projects that center equity and agency are no exception—and also require staff with an understanding of and connections to multiple communities and constituencies. In addition, Data Futures Lab’s efforts to develop a larger community and ecosystem around data stewardship in the public interest required outreach and coalition-building across locations, areas of expertise, institutions, and interests. Big societal shifts—including the redistribution of power around data and the infrastructure to support that change—require new and alternative types of financing and investment that can meet the moment.

To be sure, philanthropic organizations were vital partners for Data Futures Lab. Flexible, multi-year funding from philanthropy allows Data Futures Lab to invest in projects and approaches that might not work as expected, but that could further the field in other

ways. Beyond money, Data Futures Lab projects have benefited from the introductions that philanthropic partners have made—to other funders, to nonprofit organizations, to potential users, to corporate partners, to policymakers, and to thinkers in the field. In addition, philanthropic organizations have a unique ability to support convenings that bring together diverse voices around data stewardship challenges. Such convenings have spurred collaborations between Data Futures Lab projects and partners and resulted in pivots to serve different audiences and needs. For example, Data Futures Lab connected [The Drivers Cooperative](#), [StopClub](#), and [The Drivers Seat Cooperative](#)—three organizations committed to creating a fairer system for drivers in the ridesharing economy—to discuss the aspects of their data insights could be integrated into StopClub’s expansion to the United States.

Still, Data Futures Lab recognized that there is a need to engage impact investors and mission-aligned capital holders to further the impact that philanthropic dollars have initially unlocked. Fernando reflected, “There is a tremendous amount of capital to support initiatives that are desired by the public. Think about something like the [Y-Combinator](#), which is designed to help founders make something people want.” Fernando continued, “If we can get those investors to recognize data agency-related projects as central to realizing their missions, we could substantially further positive data stewardship practices and deepen the entire ecosystem.” In the coming years, Mozilla hopes to build this pipeline of investors to sustain projects that were initially supported by philanthropy.

# Impact

DATA FUTURES LAB'S aim to both promote change within existing systems and to cultivate alternative ecosystems was hard to quantify, especially when it came to near-term outcomes. To address these challenges, Data Futures Lab tried to examine long-term change in the field; connections and relationships established through convenings and partnerships; and whether and how new data stewardship projects in the public interest emerged. Fernando explained, "Even if that one particular organization [didn't] go on to succeed, the outcome we want is to learn from the experiment they did so that others who are trying things can benefit."

While the Lab views the whole as greater than the sum of its parts, Data Futures Lab has developed a set of proxy measures to help capture its collective impact:

- Data Futures Lab supported **dozens of projects** to develop open-source tools that enable a more fair and trustworthy data economy. These projects represent a range of locations, languages, institutions, approaches, and people. Projects are also at various stages of development, offering an opportunity for less-seasoned innovators to learn from those who are further along in their journeys.
- Many funded projects have gone on to be **used by thousands of users**, changing the way that they engage with and control their own data. For example, [The Drivers Cooperative](#) offers an alternative to Uber, Lyft and other for-profit ride sharing services, delivering money back to drivers and collecting data as a protection for both riders and drivers. The cooperative now has [over 9000 drivers in New York City](#) and is also expanding to Denver and other cities.
- [Data Futures Lab's public events](#) exposed thousands of people to new voices and perspectives in the data governance

world. Initiatives that are trying to do things differently in the data economy take a single-sector approach—mobility, health, or labor rights, to name a few—and are unaware that there may be other, related initiatives in other sectors. Data Futures Lab's public events offered a space for people and projects to come together and build community across sectors. For example, Data Futures Lab [hosted a speaker series](#) featuring projects and people exploring ways to make more equitable data ecosystems in the era of generative AI. The speakers and participants are now in touch and are in a position to collaborate.

- Read more in [DFL's Impact Report](#) (June 2024).

## Next Steps

**Mozilla's** work to imagine and build better data futures is built on community, and there are numerous ways for a range of stakeholders to become engaged with this work:

- Philanthropic foundations and mission-aligned investors can reach out to **Mozilla** to begin a conversation about partnering to support a new data economy.
- Technologists seeking support to workshop new approaches for data stewardship can check out opportunities for funding, community, mentorship, and programming.
- Anyone interested in data stewardship can engage with the Data Futures Lab by attending its public events and **those of related Mozilla programs like [Common Voice](#)**.
- Reach out to the **Mozilla team** to learn more about the project's work and to see other ways to get involved.

### DFL by the Numbers

**\$1.5 Million** in grants made between 2021-2024

323

Registered community members who participated in our monthly community calls on data governance topics

383

Responsible Data Stewardship projects submitted to our open calls

20

Infrastructure and prototype projects funded

19

Community calls hosted with researchers, practitioners, industry partners, and civil society organizations

3

Showcases that spotlighted projects addressing local data stewardship challenges

### DFL IMPACT REPORT

To learn more and contact Siegel Family Endowment, visit [www.siegelendowment.org](http://www.siegelendowment.org)