AI and the Better Deal for Data

Major Questions Essays

Jim Fruchterman, May 6, 2025

Note: In the following essay we reference the eight Better Deal for Data™ commitments originally proposed in our April 2024 white paper. In December 2025, we released the BD4D™ Commitments as a set of seven refined commitments.

Introduction

AI is the hottest tech topic in society today. Data and AI are of course intimately linked. You can't have AI without training data, preferably of high quality and a lot of it. However, the great majority of humanity, and most social problems, are not well represented in the data being used to train today's AI tools. One of the long term goals of the Better Deal for Data (BD4D) is to help remedy this gap – to make more comprehensive and representative data available while protecting data subjects from abuses. The goal of this paper is to explore this issue and offer proposals for feedback around the crucial question: **how can organizations work with AI technology under the Better Deal for Data commitments?**

Data comes first. You can easily have data without AI, but not the other way around. So, our first concern is the challenge of collecting data without that data being appropriated for AI uses by third parties. Next, we will discuss how to use data ethically to create AI models: what will organizations committing to the BD4D do differently when it comes to AI. Finally, we will talk about the promise of AI based on ethically sourced data for making social impact, including good and bad examples of applying AI.

This working paper is one of a series we are publishing on major questions we are being asked about implementing the Better Deal for Data. It is not intended to be the final word on where the Better Deal for Data standard version 1.0 ends up on these questions. Instead, these Major Questions papers are encouraging us to explore each specific question, and think through our initial answers. We are hoping to engage in more extensive debates over any controversial points.

Collecting Data without Appropriation

AI models are often trained with data generated for a reason other than creating AI. The content of books, news stories, images, and Wikipedia articles was created to meet other needs, long before it was used to help create the current generation of Large Language Model (LLM) AI tools. This reuse is a subject of considerable controversy in society, as content creators sue (or simply complain) about the appropriation of their intellectual property without permission or compensation. One particular issue is when AI tools spit out articles or images which are strikingly similar to an original creation, seemingly egregiously violating copyright.

Many of the Better Deal for Data commitments address these controversies. Commitment One dedicates the use of the data to social good outcomes, rather than for private profit. Commitment Four makes "selling" data off-limits. Commitment Six requires sensitive data to be treated confidentially. Commitment Seven talks about anonymization of data when it is being reused for research. Should we expand that to include the reuse of data for AI? We think the right answer is probably yes. The commitments are explicitly creating an alternative to the current commercial norm of surveillance capitalism, where data is regularly reused in ways that surprise and disappoint the subject of the data.

Let's imagine a social change organization committed to BD4D is collecting data as part of their programs while providing needed support (money, food, housing, counseling, healthcare and so on) to people in need. The org is not planning on using the data for AI at the time it is being collected. In this case, the challenge is preventing the data from being reused, scraped, or stolen by third parties training an AI model, ranging from big tech companies to hackers. The org should protect this data appropriate to the sensitivity of the data.

The challenging thing is how easy it is for well-meaning organizations to enable problematic data sharing without knowing. Employees might share the content of a sensitive text message or email with a commercial generative AI tool for help writing a response. The mailing list software your team uses might keep track of your stakeholders by name and build profiles of them, likely then pulling those profiles into AI algorithms for targeting. Sometimes it is nearly impossible to avoid some data sharing. It's hard to have a website today without a tool like Google Analytics, and simply having a link to Facebook on your website (often a requirement to reach your donors or clients) causes at least some data to be shared.

Not all data is sensitive, of course. Some data might be openly shared as a matter of course, such as weather data from a community weather station, which is critical to the development of weather forecasting algorithms. For many nonprofits, the use of a website analytics tool may not pose any significant problem to the people visiting your website. The goal of the Better Deal for Data is to be practical, and we don't want the perfect to be the enemy of the good.

That being said, context matters. A specific data practice might be fine in one application, and might put stakeholders at risk in a different situation. For example, this author was once responsible for developing a secure human rights documentation tool with strong encryption that

kept the data safe from prying eyes. At the same time, we actively discouraged activists in Tibet from using our tool. That wasn't because we thought it was insecure, it was because the use of a secure tool could be observed by watching Internet traffic and traced back to the source, even if the content could not be read.

As you collect data from the communities you serve, you have a responsibility under the Better Deal to consider the welfare of those communities, both individually and collectively. You are generally in a better position to evaluate the risks the data will be exposed to in the real world context where you operate. Consider threat models, and think about how your actions might inadvertently expose confidential data to an industry dedicated to collecting as much data as possible. These responsibilities don't fall just on leadership: every team member who touches this data needs to be aware of the promises being made to the communities you serve.

AI Under the Better Deal for Data

Many organizations are sitting on a large amount of data collected over years. As AI applications proliferate, the question of reusing this data to create an AI solution will come up more and more often. The first question is whether the proposed AI solution is actually in the interest of the community being served. It isn't always easy to assess this, but the BD4D does operate under a "no unpleasant surprises" rule of thumb: if many of your community members would be unhappy with the application, or find it hard to imagine why it helps them, it probably is not a use case that fits under the Better Deal for Data. There are plenty of data uses which are legal but do not fit under an ethical use initiative such as BD4D.

The first set of positive use cases are around internal program improvement, assuming your organization has the capacity to apply the relevant tech tools. For example, using AI to better diagnose a disease or a learning challenge, figure out which program options are likely to work for a specific client, or to predict when a piece of equipment might fail or what spare parts should be kept in stock in a hospital, are all reasonable examples of using AI to glean insights from a trove of historical data. The data doesn't need to leave your organization, and it's being used to deliver better outcomes. Improved efficiency or program outcomes are typically textbook examples of properly using technology for social impact.

Of course, you have to watch out for applying new technology that makes things worse. Using AI to complete administrative tasks, for example, can increase capacity, but if it degrades client experience or even outcomes, replacing human staff would not be desirable. There are plenty of illustrative examples of this in the field.

Over time, we should expect organizations to pool their data with other mission-aligned groups to create shared datasets for training better and more powerful AI models that are beneficial to a social good field of practice. We hope that BD4D makes ethical data sharing across organizations

more practical, by setting a common approach to such efforts. Given that more data generally improves the performance of AI models, this could be an important step for a given field.

The BD4D does put some constraints on such data sharing exercises. For data sets containing detailed information about individuals (personally identifiable information, or PII for short), it is a requirement to strip out that PII before sharing as part of anonymizing the data. Beyond that, there needs to be a requirement that anybody accessing data sourced under the BD4D also commit to those principles, including agreeing to not try to reverse engineer the anonymization process, since the data available in the world is often large enough to make figuring out the identity of an individual straightforward even in heavily redacted data. For example, having the age of everyone in a family and the street name where they live might uniquely identify that family in the United States, because there are many street names which are unique to a single town. In addition, no algorithm to scrub out PII from unstructured text is 100% accurate.

The further removed from personal information the original data is, the more likely that this kind of reverse engineering of privacy efforts becomes effectively impossible with an AI model. For example, an AI image recognition model for detecting malaria parasites or spotting a tumor in a lung would not be likely to pose a privacy risk, and is a beneficial example of a use case that would be improved with more source data. At the same time, if these images were accompanied by detailed DNA data and/or complete medical records, a much higher degree of care needs to be exercised to ensure privacy of the medical patient.

Finally, you need to test your compliance with the Better Deal for Data as it applies to data and AI. Have you confirmed that private datasets cannot be accessed by unauthorized parties? Before you deploy an AI tool, have you tested it extensively to confirm that it works as advertised, and that errors have been minimized? Are the guardrails sufficient to prevent the worst impacts? Have you tested it to see if PII can leak through?

Using AI for Social Impact

The nonprofit sector is abuzz with the potential of AI, and is busy navigating the dual challenges of having a positive social impact while using the data of their community members ethically and responsibly. We include all kinds of AI as part of this conversation, both traditional predictive AI (is this a picture of a malaria parasite?) as well as the latest generative AI conversational tools.

Predictive AI requires source data to function: it needs many examples of pictures of blood samples that both contain and do not contain malaria parasites to create a tool that can spot malaria. The public health field has long been a leader in the ethical use of medical data to advance diagnostics and science, while protecting the identities of the individuals contributing to these datasets. An exciting recent example of dataset generation is Karya, which pays speakers of languages not well-represented in linguistic datasets to generate samples of their spoken

language. These samples are used to improve the performance of speech recognition tools in these neglected languages, and retain a royalty right for the person who originated the sample. This is an example of a win-win data scenario for a nonprofit: the people creating the samples are well-compensated and AI technology advances for these important languages.

Tech Matters, our organization, creates open source software for crisis response helplines. Conversational chatlogs and recordings, which are some of the sensitive data we hold for these helplines, belong to the helplines. We received permission to use the data to try using AI to summarize text conversations, saving helpline counselors from needing to write up these summaries, and engaged with helplines in user research to design an AI data entry tool that would be useful. As a first step, we ran the chat conversations through an anonymization step to remove personally identifiable information before using commercial LLM models for summarization, under an agreement that the LLM vendor would not retain any of the data. Then we provided the summaries to counselors for review before saving them as part of the helpline's records, which are kept as part of the organization's records, with the same standards of data protection for other case data.

Our last good example is an increasingly popular AI use by nonprofits, which is Retrieval-Augmented Generation ("RAG"). This uses a database of vetted content, such as training videos, or common questions and answers, to focus an AI-driven chatbot to constrain its responses to be modelled after the vetted content. This generally provides the needed guardrails to make generative AI tools useful to consumers without making up incorrect answers.

There are of course many examples of AI use by nonprofits which fail to meet the commitments embodied in the Better Deal for Data. Selling confidential data to a for-profit has been done, and is explicitly not permitted under BD4D. Unleashing an unconstrained chatbot on people in need is also an irresponsible action, as such chatbots are likely to give false advice, as in the famous example of the National Eating Disorders Association, which fired its human counselors and replaced them with Tessa, the AI chatbot which was caught telling helpline users the exact opposite of good advice.

Conclusion: What's Missing?

The Major Questions papers are intended to explore the big issues we hear about from the many collaborators who have contributed their data use cases, their feedback, and their support to the Better Deal for Data. The foregoing distills our initial thinking about the intersection of AI, data, and nonprofits. At this point, we would like to get even more feedback from the community, especially:

● What additional issues come to mind about this subject?
● What did we get wrong?

- What examples do you have of nonprofit data use which should be inside or outside the Better Deal for Data, or are simply puzzles to consider?

When it comes to AI and ethical data use, we in the nonprofit sector are learning about problems as we try new things, and see good and bad examples. We hope that the Better Deal for Data provides some guidance as the field moves forward, by placing the interests of the people served by organizations at the center of our AI tech applications.

We are looking forward to many new questions and ideas as we work together to craft a usable Better Deal for Data.