CENTER FOR AI POLICY

Summary: Privacy Concerns & AI

Americans care about their privacy, but protecting it is hard. 94% of Americans are concerned about the privacy of their personal data, but struggle to protect it given arduous and lengthy processes designed to discourage them.

AI exposes your data in new ways. Not only will AI make identification cheaper and quicker at scale, it will likely continue to make use of new signals to further infer our characteristics or thoughts. One important development is recent progress in AI's ability to associate brain waves with thoughts; further advances in AI and brain-scanning technology could bring remarkable mind-reading capabilities.

Continued AI development could contribute to further use of our data for manipulation. The practice is not new, but AI stands to increase both the effectiveness and frequency of attempts to influence us by using our data. Whether by leveraging our characteristics (age, education, medical history, etc) or our personalities (e.g. introvert vs extrovert), this capacity won't just be used to get us to buy products, but will likely be expanded out for other manipulative means, like trying to influence our votes.

Our current system of privacy protection is failing. U.S laws largely follow the "notice-and-choice" approach, where companies inform people about their practices, and then people must consent to proceed. This practice is flawed because it gives people two unappealing choices: forfeit privacy rights, or forfeit access to the company's product, service, or website. A second problem is that, as AI's predictive capabilities improve and further vision is granted into our lives, privacy violations will increasingly come from inferences models make about us rather than especially sensitive data we've given up. Voluntary steps could protect our privacy, but thus far companies have more often chosen to subvert privacy protections rather than protect them.

Comprehensive data protection is needed. The bipartisan American Privacy Rights Act (APRA) would be a huge step forward in safeguarding Americans' privacy, requiring: transparency over how data is used, standards for data minimization, the option for consumers to access and delete their data, and more. Some changes could be made which would further strengthen the bill: government organizations should also be covered, opt-in consent for targeted ads would be an improvement over the option to opt-out. But even without these changes, this bill would greatly improve the current situation, and if passed would be a foundation upon which further privacy protections could be built. Americans are in support too, 80% approving of its major tenets.

Full Report: Privacy Concerns & AI

"Thanks to the growth of predictive analytics, algorithms and big data-mining businesses you can now look forward to a future that's made up of equal parts Orwell, Kafka, and Huxley." — Dr. Simon Moores

Introduction

94% of Americans are concerned about their data privacy[1], but most are unaware of how their data is being used and what they can do to protect their own privacy. Companies are partly to blame, as many intentionally put barriers in place to prevent understanding how data is collected and used.

AI is not only likely to exacerbate existing privacy issues, but also will further change the privacy landscape itself, where invasions of our privacy in the future might not look like loss of data but rather increasingly accurate algorithms which are able to infer our sensitive information.

In this report, we'll cover the current state of data privacy, how your data is abused, and how AI might further expose your data, finally taking a look at what policy options might be able to address the issue.

The Current State of Your Data

As you use different services online, nearly all of them will be collecting some form of data on you, which includes the usual social media apps (Instagram, Snapchat, Tik Tok, etc.) as well the websites you visit and the web browser you use (e.g. Google if you use Chrome). Google, for example, "knows more about you than your spouse" and stores:

- Every search you've ever made (including, until recently, searches made in Incognito mode which they only stopped using after a class action lawsuit).
- Every email you've ever sent or received, including those deleted.
- Every single location you've been to with Google Maps tracking enabled.
- Every photo you've uploaded into Google Photos, including those deleted.
- Every document you've made in Google Docs.
- Every IP address of the devices you've signed into Google with.

Social media apps keep a fair share too. Instagram, for example, stores:

[1] Privacy is an amorphous concept that has meant different things over time and still means different things to people as they currently use it. Here, we're generally focusing on privacy as an instrumental good, as **whatever is necessary to protect individual autonomy**, the ability to determine what information about you is public and what is not.

- Every message you've ever sent, even if deleted.
- Every comment you've ever made, even if deleted.
- Every post you've ever liked.

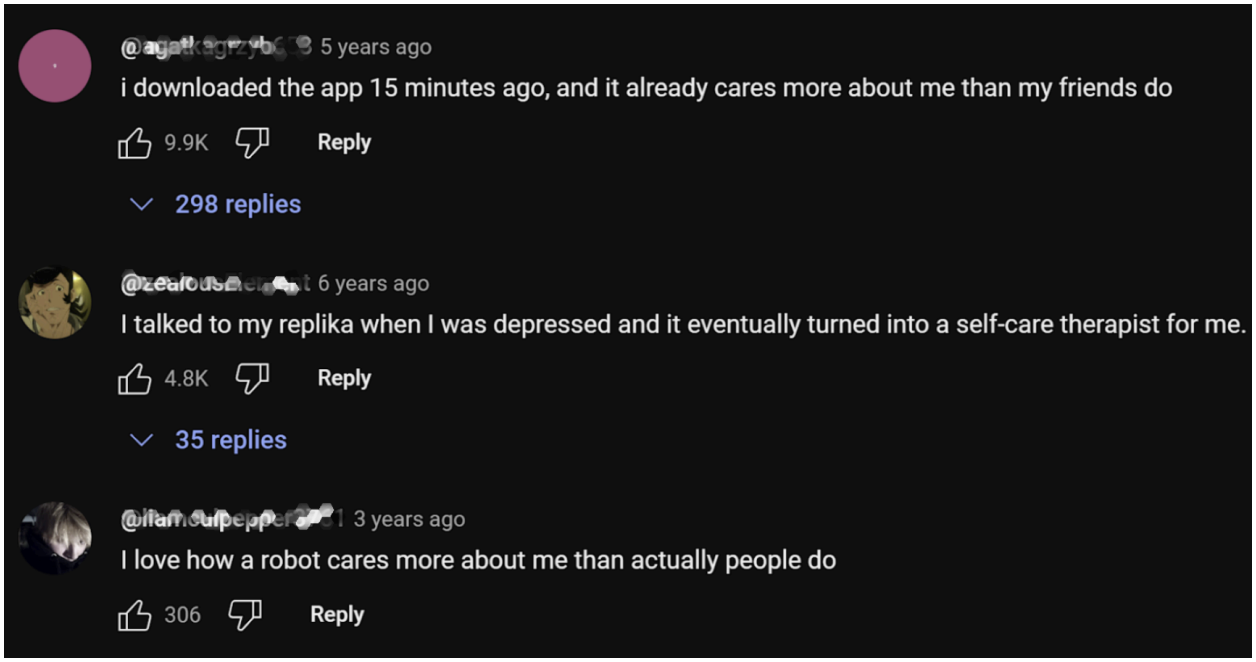
While this data can be used for any number of things, it's often used by businesses to more effectively target people for advertising. After it is collected, data is normally purchased by data brokers, companies which amalgamate different information sources into cohesive profiles on who you are. Acxiom, one of the largest data brokers, claims to collect over 3,000 data points, including your:

- Marital status
- Date of birth
- Homeowner status
- Income
- Ethnicity
- Occupation
- Education
- Media consumption
- Retail behavior
- Current health (e.g. Flo, a period and pregnancy tracking app, which sold information on pregnancies)
- Medical history (e.g. medications you're taking)
- Mental health issues (e.g. depression)
- Travel and Entertainment choices
- Interests and Behavior (e.g. gardening, fashion, theater, crafts)

AI Models Continue the Trend

AI companies are much the same. OpenAI, for example, keeps a log of **every single chat you've made on the platform**, a worrying prospect for unknowing users.

Worryingly, some interactions with AI are even more personal than simple requests for information. Millions of people worldwide are already conversing with "virtual AI companions", AI-powered avatars that many treat as friends, therapists, or even romantic partners. You may be tempted to dismiss the use as more in jest than serious, but some people really connect to these personas:



These conversations aren't informational but rather relational; people often reveal deeply personal information in the midst of creating what they see as a real relationship with these "AI companions". These aren't just heartfelt text messages, but voice notes, pictures, and even videos too, all now part of the companies' data.

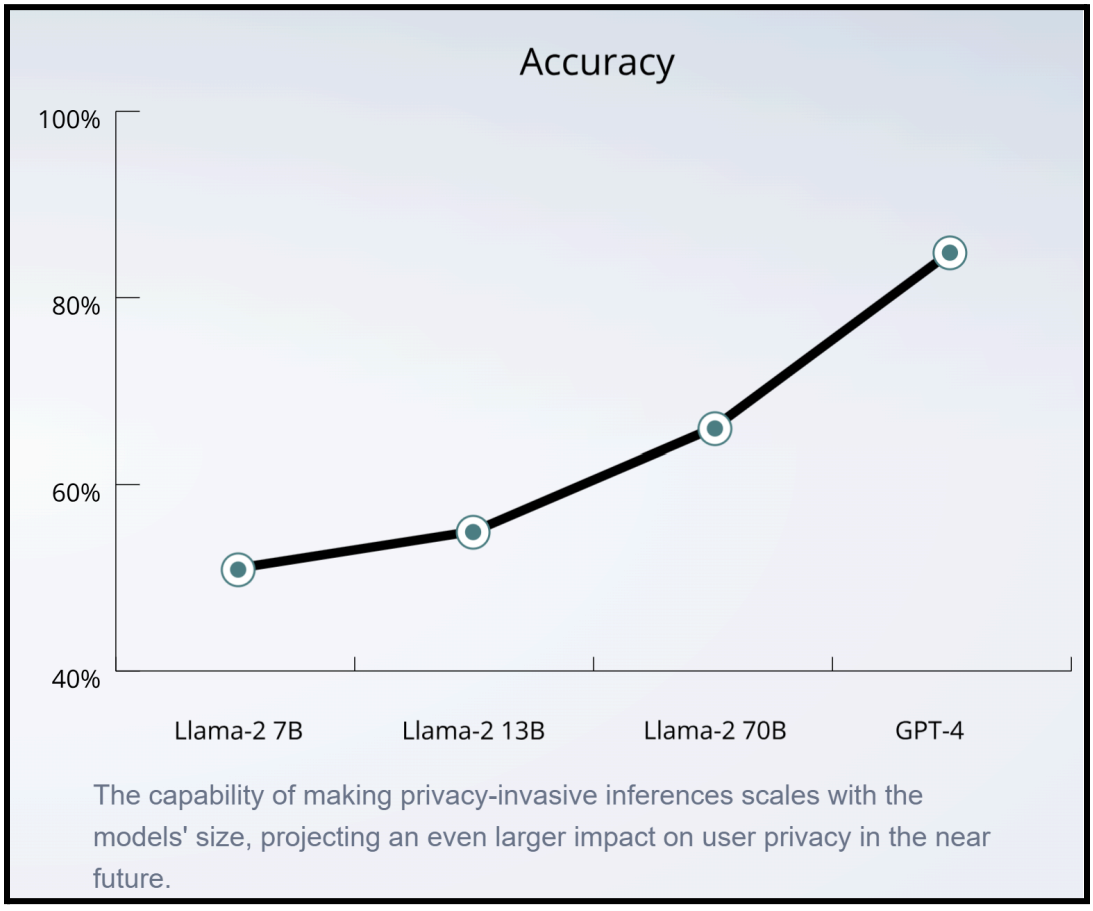
Be sure that, as with ChatGPT, your data is now a business asset which they'll likely "share", i.e. sell, to third parties. Some use your data to better promote their product, others use it to specifically market the app to you, which is especially worrying given the potential to keep users hooked, young users already finding the product "addictive". And even when they're not using your data for their own purposes, they're selling access to third parties to monitor and track you as you engage in the product, an average app giving access to **over 2,663 trackers in just the first minute**.

What's worse is that many apps don't even give you an avenue to delete your data, meaning that even if you've realized you sent sensitive information that you regret, you might have no power over its continued use. There's also the risk that a third party might gain unauthorized access, as many of these companies also don't seem to have deployed robust data security measures. The takeaway with these apps is that their privacy policies are a mess, and involve potentially quite sensitive data. Mozilla's privacy team has reviewed many apps and websites throughout its time, and the AI companion app Replika has received the worst rating for privacy they've ever given.

AI Exposes You in New Ways

AI doesn't just continue past trends though, it also exposes your data in new ways. In 2016, a father received a pregnancy-focused coupon book from Target. What at first seemed like a mistake turned into a realization his daughter was actually pregnant. Just from analyzing purchasing history, Target could both estimate the likelihood customers were pregnant and even predict a due date, then targeting advertisements to expecting mothers without them ever disclosing they're pregnant.

Your shopping history is not the only place your sensitive data might be exposed: simple texts and comments now put you at risk too. Research has shown that as AI models continue to grow, they are becoming increasingly capable of inferring aspects like location, age, sex, or place of birth from seemingly innocuous casual conversations you might have over social media like Instagram or Reddit, or talking with a chatbot.



The capability of making privacy-invasive inferences scales with the models' size, projecting an even larger impact on user privacy in the near future.

In the study, just by analyzing Reddit profiles GPT-4 was able to correctly guess: sex 98% of the time, location 86% of the time, marriage status 92% of the time, age 78% of the time, education level 68% of the time, occupation 72% of the time, place of birth

93% of the time, and income level 63% of the time. Such data may seem trivial, but research has found that nearly half of the US population can be uniquely identified from just a few attributes (e.g. location, gender, date of birth). AI's contribution is that it can scale these inferences, executing identifications roughly 100x faster than humans and 240x cheaper.

It's possible to go beyond these basic traits to infer even more sensitive pieces of information too. Previous work found it's possible to infer things like relationship status, religion, drug use, and political views at above-chance levels just from access to someone's Facebook data.

AI's ability to infer more about us from current data isn't limited to text either. GeoSpy, an AI designed to identify locations from nothing more than a photo, can do so with surprising accuracy. Polygraphs, used for lie detection, are not very accurate, but a recently developed AI model purports to tell truth from lies with 90% accuracy, improving from 73% accuracy in just a year.

Researchers have also recently begun training AI on human brain waves to associate thoughts with brain patterns, commencing the development of AI which can predict what we're thinking about (reconstructing our thoughts and mental images), again already with surprising accuracy. As this technology further improves and becomes easier to deploy we're likely to see more and more of it out in the world, researchers already working out potential use cases like predicting deception in CEOs.

AI has also been able to use signals in entirely different ways, making entirely different kinds of information potentially invasive. AI, for example, can now identify people in a building (including their posture and position) just from WiFi signals. Furthermore, AI can identify keystrokes of a nearby phone, or laptop on a zoom call, simply from their sound with over 90% accuracy, indicating a whole new realm of situations in which your passwords will be at risk.

How Your Data is Abused

Much harm flows from people just having access to your personal data, but more harm can come if that data is then abused.

Manipulative Advertising

The normal use for your data is to enable better targeting of ads, which are being deployed to increasingly more specific audiences. One of the main ways this occurs is by targeting groups created by data brokers, lists of people who share some common characteristic. After facing previous backlash, data brokers are much more PR aware in

how they name their target groups now, but some past categories made it clear what these groups were being used for, with target groups like: "Suffering Seniors", "PayDay Loan Central-Hispanic" or "Help Needed-I Am 90 Days Behind With Bills". Another category targeting elderly gamblers described the target group by saying: "These people are gullible. They want to believe that their luck can change." It's not just elders that are targeted either, children's data is gathered too, which is worrying as children are easily influenced by the ads they see.

Increasingly though we're seeing even more fine grained targeting, where your weaknesses can be identified at an individual level from other data about you. Facebook's Advantage+ which will automatically "optimize" your audience for you, focuses ads on those identified as most likely to make a purchase. One way to do this is by targeting psychological characteristics (e.g. introvert vs extrovert), because persuasion is more effective when tailored to people's specific psychological characteristics. For example, Facebook likes[2] can serve as indicators of deeper personality traits[3], and when people in a study were targeted on those traits, the likelihood they would purchase a product increased by 31 to 54%.

What's troubling is that this targeting is often unknown and subtle. Remember Target sending pregnancy ads based on predictions from items bought by customers? Well, after realizing that sending a full set of pregnancy ads out of nowhere might be creepy, they didn't stop the practice but rather got smart, **mixing in random items with those focused on pregnancy to make people think they weren't being targeted**. As an anonymous Target executive put it: "We found out that as long as a pregnant woman thinks she hasn't been spied on, she'll use the coupons. She just assumes that everyone else on her block got the same mailer for diapers and cribs. As long as we don't spook her, it works."

Manipulation Outside the Market

The ability to leverage data for abuse won't just be limited to getting you to buy more products. Your data will likely soon be used for spear phishing, where instead of scammers casting a broad net and sending out the same spam to everyone, they can leverage AI's ability to quickly infer your characteristics to personalize spam to each person, individually manipulating people at scale.

Another worrying potential is influencing elections. Cambridge Analytica collected the data of millions of Facebook users to profile them, later targeting them based on their personality. They raised some candidates from obscurity, and helped others all the way

[2] Among a range of other online sources (blogs, Twitter, Facebook, Instagram, etc.).
[3] E.g. liking content about "computers" indicated a person was likely to be introverted.

to presidential offices. A subsidiary helped make Brexit possible, supporting the Vote Leave campaign whose director said "we couldn't have done it without them". Research has shown similar worrying signs, one study finding messaging can affect real-world voting behavior, the study generating over 340,000 additional votes in the 2010 election.

Other Data Abuses

Outside the ability to leverage our data against us for manipulation, there are times too where just having data exposed is enough to cause harm in-and-of itself. A woman was murdered after a stalker bought access to her data and used it to find out where she lived. Scammers were able to take advantage of millions of older Americans and specifically target the most vulnerable using data obtained from a data broker. A priest was outed as gay by obtaining his data from his use of a dating app. Conversations with a mental health crisis hotline were sold to a company developing a customer service bot.

Problems With the Current System

Current Privacy Regulation is Lacking

The above abuses have been made possible by a broken US system for privacy protections. The US has taken what's called the "notice-and-choice" approach, where organizations *post notices* about how they collect and use your data and you then *choose* whether to do business with them. As many know, the problem with this system is that it's overly burdensome: one 2007 study estimated it would take **over 240 hours a year** for the average person to read the privacy policy of all the websites they visit.

Even when you actually read the privacy policies, that doesn't guarantee you'll be able to make an informed decision afterwards, especially when you don't understand the technology. As one privacy researcher puts it, "AI is far too vast and complicated for individuals to understand and to assess the impact on their privacy". Without knowing what the model will be used for, or how exactly your data will contribute, consent for these companies to use your data just becomes "a blank check to do nearly anything".

Companies Won't Protect Privacy On Their Own

To be clear, this privacy ecosystem hasn't just been created by lack of government intervention but also a lack of safeguarding privacy by big tech companies and data brokers. For data brokers' part, it's fairly damning that some have knowingly sold information on vulnerable people (like elders with dementia) to scammers.

For big tech companies, abuse starts at the beginning as various tools are already used to manipulate people into agreeing to share more data than desired, one study indicating such tactics make those targeted 4x more likely to take the desired action.[4]

Then, even when we've agreed to a policy that can often abruptly change with no notice. Google decided to change its privacy policy in the middle of the 4th of July Weekend to allow it to train on public documents without notice so no one would notice. Slack automatically opted everyone into allowing Slack to use their data to train their AI without any notice, as done earlier by Zoom. Slack and Zoom later walked back their policies, but only after they were caught and there was consumer backlash.

They even go further to signal that they care, but without actually taking the actions that would give force to privacy protections. As faces were scraped from social media sites by other companies and public pressure mounted, companies like Facebook, Twitter, Youtube and Venmo all came out to publicly chastise them for violating their terms of service, but took little action beyond that. Beyond chastisements and cease and desist letters, they **failed to take the technological precautions** that could have prevented this, **or the legal recourse[5] that would have actually shown they were serious** about protecting user's data and privacy.

Data brokers and big tech put responsibility on those training the models. Those training the models put responsibility on model users. In the end no one is actually indicated as responsible, and users are left to pick up the pieces. But again, this is by design, because you and your data are the product. These services aren't free, you just paid for them by giving up your privacy.

What Further AI Development Will Mean

These abuses are likely to only increase as AI is further developed, largely for two reasons: AI companies' need for further data will only increase over time, and AI and its predictive capability is redefining what privacy looks like.

Generally speaking, the quality of an AI model improves with the amount of data you add in. The problem is that these companies have already nearly exhausted publicly available sources of data. For now, they are exploring legal means of acquiring more data, signing multi-million dollar deals with publishing companies to train on their

[4] The same tactics come out in the end too, as opt-out processes are often hard to find and quite arduous.
[5] I.e. actually bringing litigation against Clearview AI for violating their Terms of Service.

catalogs, buying access to all data on Reddit, and even considering buying entire companies to access their data[6].

But these deals only go so far, and soon these options too will be exhausted. The next focus will likely be our conversations, as content platforms and messaging apps hold nearly half as much data as the entire web. Most terms of service currently prevent this from happening, but as mentioned before, companies are covertly changing these to make this possible. Google currently says they don't train on your Google Photos or Gmail, but they keep copies of all the data, and are likely only waiting for the right moment to begin training on these materials too, just as they did with training their AI on public documents.

Even if you think your own personal data is safe, there's still reason to be worried. AI stands to shift the privacy landscape to a focus on what inferences can be made about us, where even a model with none of our own data can still infer sensitive characteristics about us. Over time, more sources of data are likely to be used towards this end, the frequency and accuracy of inferences likely improving with the models. We might soon be in a world where anything could give up your privacy, where we stand to lose our opportunity of obscurity, of having aspects about us that are ours to know and free from exploitation.

To answer this risk and hope to maintain our privacy in the age of inferences, we'll have to decide what sort of uses of the technology we do and don't support as a society, and crafting legislation that draws holds companies accountable to respecting that line, which is where we'll turn next.

Policies Directed at the Issue

Meta recently decided it would start using all of your Facebook and Instagram data to train their AI model. In the EU, despite the process being quite arduous, users had the option to opt-out of this. In the US, users got no such option, and instead were powerless to protect their data from use by Meta, because Americans continue to go unprotected by data privacy laws that would give them meaningful control over their data and how it's used.

70% of countries have passed national data privacy laws, and it's far past time the US joined the list of countries actually taking the steps to protect citizens data from use and

[6] And as AI companies are willing to pay more and more for data, companies will begin to realize selling customer data can be a profitable second income stream, further broadening the number of companies collecting data on us.

abuse. The chance to do that currently exists in the American Privacy Rights Act (APRA).

APRA would secure many of the benefits Europeans currently have under GDPR: transparency over how data is used, standards for data minimization[7], further protection for sensitive data, and the option for consumers to access and delete their data, actions which are supported by over 80% of Americans.

We do still think there are still opportunities for improvement. To use someone's data to target ads at them should likely require opt-in consent rather than just requiring that companies provide the ability to opt out. Government organizations should likely be covered under the same APRA rules, perhaps with exceptions for national security. And notification of data breaches shouldn't be left to state laws which don't all offer adequate protection, better off required within a short timeframe as with the GDPR[8].

But APRA stands strong as it is, and would provide the bedrock from which further privacy measures can be built, taking a major step forward towards a better future for Americans' data privacy.

[7] This is the concept that data should not be collected unless the business can specify why it is necessary for their product or service.
[8] The GDPR requires notification within 72 hours "where feasible".