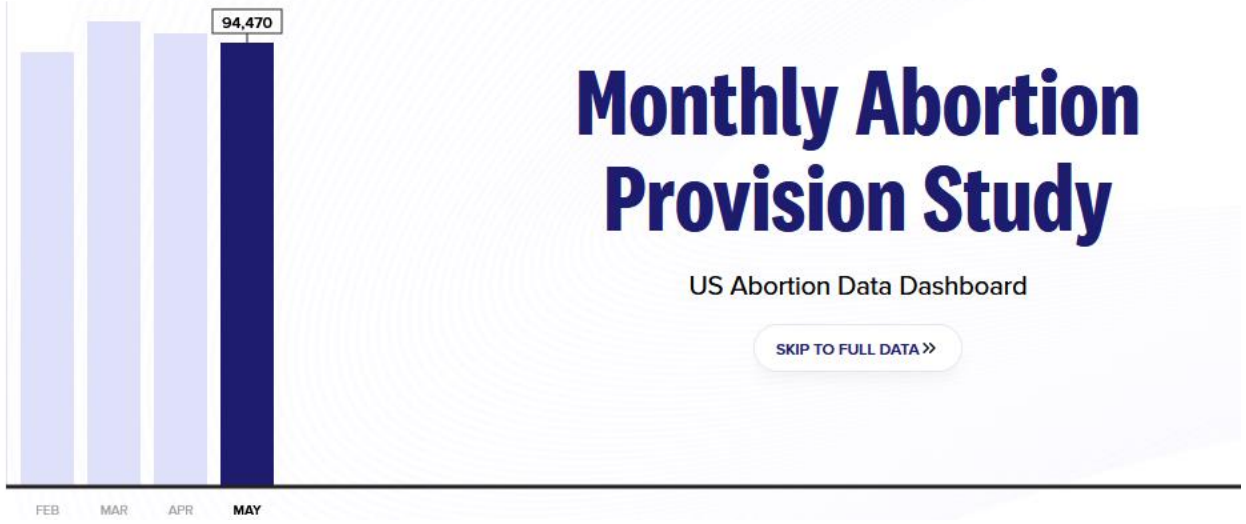
Should we use generative AI with sensitive data?

Isaac Maddow-Zimet, Data Scientist

guttmacher

Our work often involves sensitive data, and relies on trust



RESEARCH IN BRIEF

Perspectives on Sexual and Reproductive Health

Characteristics of abortion patients in protected and restricted states accessing clinic-based care 12 months prior to the elimination of the federal constitutional right to abortion in the United States

Rachel K. Jones | Doris W. Chiu

Research Division, Guttmacher Institute, New York, New York, USA

Correspondence
Rachel K. Jones, Research Division, Guttmacher Institute, 125 Maiden Lane, 7th Floor, New York, NY 10038, USA.
Email: rjones@guttmacher.org

Abstract

Background: As a result of the June 2022 decision of the United States (US) Supreme Court, as many as 24 states have, or are expected to, ban or severely restrict abortion. We provide baseline information about abortion patients living in different state environments prior to this decision.

Methods: We obtained surveys from 6674 women, transgender men, and other pregnancy capable individuals accessing abortion care at 56 facilities across the US between June 2021 and July 2022. The final analytic sample uses information from 5930 abortion patients to compare the demographic and situational profiles of those living in the 24 states likely to ban abortion (restricted) to those living in states where it is likely to remain legal (protected).

Online Medication Abortion Study

An innovative research initiative on online abortion access in the United States



Project Overview

Building on Guttmacher's decades of experience measuring abortion incidence in the United States and around the world, the Online Medication Abortion Study (OMAS) estimates the number

guttmacher

Challenges with data security

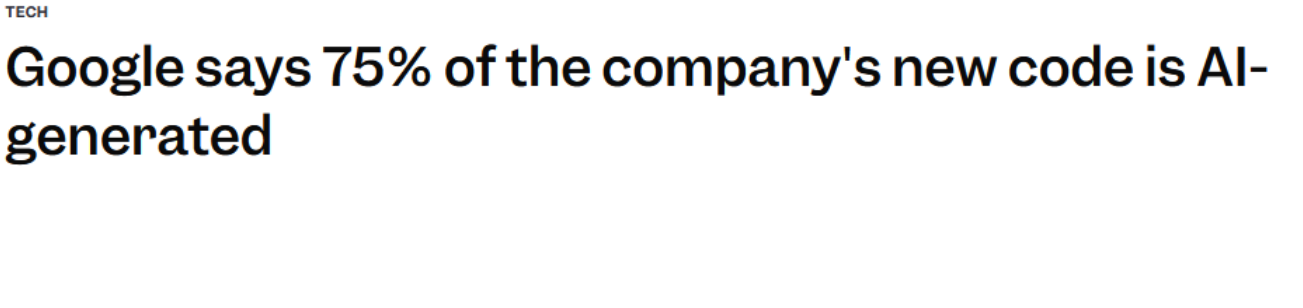
- Almost all frontier AI models do inference non-locally, on servers that aren't controlled by the researcher and which may not be IRB compliant
- Risk of data leakage unclear
- Even opting out of training relies on trust that ToS will be followed

But what about using LLMs or AI agents for code?



guttmacher

AI can be really useful for programming


TECH

Google says 75% of the company's new code is AI-generated


404

Software Developers Say AI Is Rotting Their Brains

EMANUEL MAIBERG · MAY 13, 2026 AT 9:00 AM

- Can supplement human code review
- Can help debug programs
- Can (sometimes) write pretty good code independently

guttmacher

But…code can have sensitive data too

```
1
2 # Study site in Birmingham identified some ineligible patients (getting
3 # STI testing, not abortion services, on September 8th, so we exclude from
4 # dataset)
5 analysis_data ← raw_data %>% filter(!(site == 73 &
6                                        service_type == 1 &
7                                        date_service ==
8                                          as.Date("09-09-2026",
9                                                  format = "%m-%d-%Y")))
```

(just an example, not real code)

guttmacher

…and can produce security risks

We Have a Package for You! A Comprehensive Analysis of Package Hallucinations by Code Generating LLMs

Joseph Spracklen
University of Texas at San Antonio

Raveen Wijewickrama
University of Texas at San Antonio

A H M Nazmus Sakib
University of Texas at San Antonio

Anindya Maiti
University of Oklahoma

Bimal Viswanath
Virginia Tech

Murtuza Jadliwala
University of Texas at San Antonio

Looking forward

guttmacher

Some paths forward

- Using local LLMs
- Sandboxing
- Developing code with synthetic data

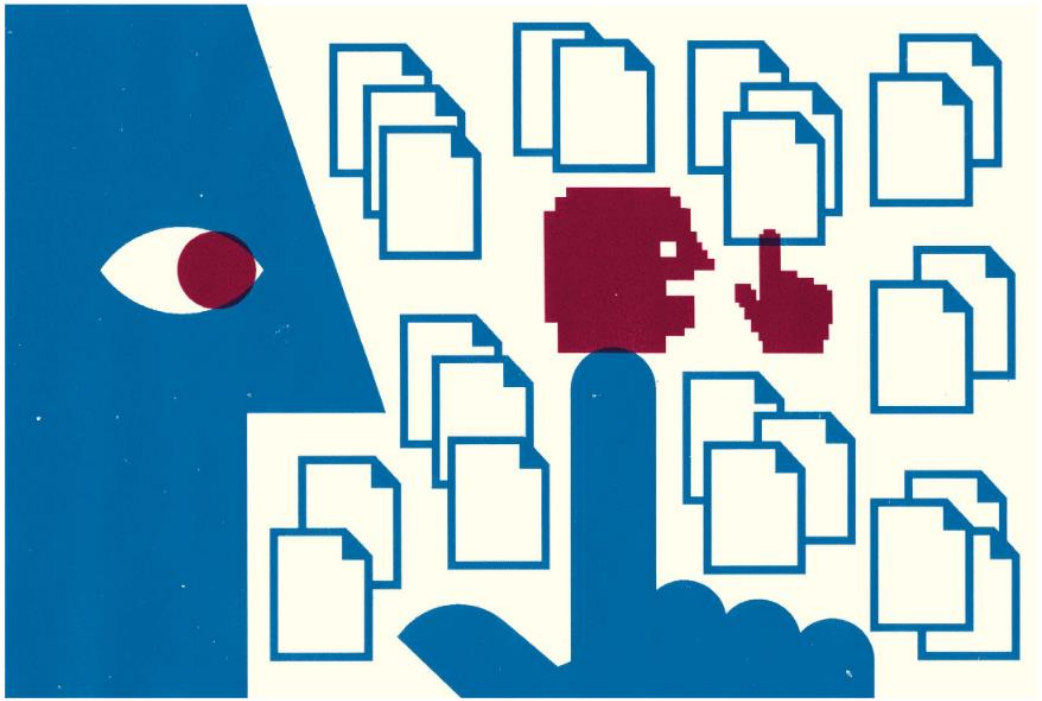


ILLUSTRATION: THE PROJECT TWINS

HOW TO SEARCH THE LITERATURE WITH LOCAL AI SYSTEMS

Concerns about privacy, reproducibility and cost are prompting scientists to use artificial-intelligence tools on their computers. **By David Brzostowicki**