Preserving Privacy in NLP Applications

Pieter Luitjens
CTO
Private AI

Privacy is a Real Thing Now

Heightened awareness around privacy

GDPR and the patchwork of global legislation

Largest 1 day loss in a stock in US history was because of privacy*

*Source: https://www.investopedia.com/investing/biggest-singleday-market-cap-drops-us-stocks/

What About NLP?

- The primary concern in NLP applications is Personally Identifiable Information (PII)
- Covers direct identifiers like name, phone number, BSN
- **Personal data is broader than you think**
- GDPR:
  - “Personal data is any information that relates to an **identified or identifiable living individual.** Different pieces of information, which collected together can lead to the identification of a particular person, also constitute personal data.”
- Also covers indirect identifiers like political beliefs, sexual orientation

Source: https://ec.europa.eu/info/law/law-topic/data-protection/reform/what-personal-data_en

Identity & Attribute Disclosure

SLATE

RADIATES WHOLESOME PERMANENCY

future tense

A South Korean Chatbot Shows Just How Sloppy Tech Companies Can Be With User Data

BY HEESOO JANG

APRIL 02, 2021 • 2:19 PM

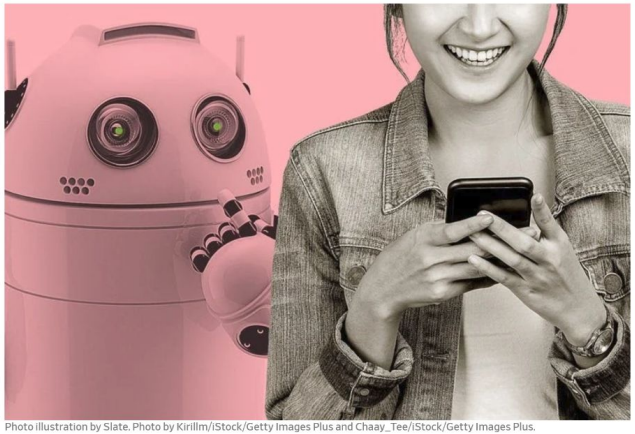

Photo illustration by Slate. Photo by Kirillm/iStock/Getty Images Plus and Chaay_Tee/iStock/Getty Images Plus.

> It also soon became clear that the huge training dataset included personal and sensitive information. This revelation emerged when **the chatbot began exposing people's names, nicknames, and home addresses** in its responses."

68% of enterprise data goes unused *

Among the top 5 reasons:
Making collected data usable *

*Source: https://www.seagate.com/our-story/rethink-data

So what can you do about it?

Solution 1: Differential Privacy

Pros:

- Mathematical guarantee

Cons:

- Basically like hitting your model with a sledgehammer - expect a large loss in accuracy
- Needs privacy expertise - pretty hard to find right now
- Hard to convince non-technical stakeholders and data custodians it works

Source: https://ec.europa.eu/info/law/law-topic/data-protection/reform/what-personal-data_en

Solution 2: Synthetic Data

PRODUCTION DATA:

Dear Sir,

A truck hit my car at the Eaton Center on March 13. My policy number is 049305. Am I eligible for a claim?

Martha McEwan
647-954-3456

SYNTHETIC DATA:

Hello,

My truck parked at the parking lot was rear-ended and got scratched. How do I make a claim?

Best,
Maarten

Pros:

- No added complexity to your ML systems
- Can be used as data augmentation

Cons:

- How do you ensure that the synthetic data is similar to the original data?
- Similar to DP, requires privacy expertise and difficult to convince non-technical stakeholders it works

Private AI: Redaction or De-Identification

PRODUCTION DATA:

Dear Sir,

A truck hit my car at the Eaton Center on March 13. My policy number is 049305. Am I eligible for a claim?

Martha McEwan
647-954-3456

DE-IDENTIFIED DATA:

Dear Sir,

A truck hit my vehicle at the [LOCATION_1] on [DATE_1]. My policy number is [PERSONAL_NUMBER_1]. Am I eligible for a claim?

[NAME_1]
[PHONE_NUMBER_1]

Pros:

- PII is usually not important
- Inherently explainable
- No need to change your ML systems
- Easy to use

Cons:

- Sometimes involves accuracy degradation in downstream applications
- Language-specific

Private AI's Redaction System

Multi-lingual

Supports 42 languages

Runs on-prem

So your data never leaves your premises

Built for scale

Processes 4.5B requests per month

Redaction & the GDPR

> Personal data that has been rendered **anonymous** in such a way that the individual is not or no longer identifiable is no longer considered personal data. For data to be truly anonymised, the anonymisation must be irreversible."

- Needs to be very good at finding PII to be useful - regexes don't cut it
- Private AI's system is built on the latest transformer architectures (not BERT) by a team of 20 people

Source: https://ec.europa.eu/info/law/law-topic/data-protection/reform/what-personal-data_en

Private AI: Synthetic PII

PRODUCTION DATA:

Dear Sir,

A truck hit my car at the Eaton Center
on March 13. My policy number is 049305.
Am I eligible for a claim?

Martha McEwan
647-954-3456

SYNTHETIC DATA:

Dear Sir,

A truck hit my car at the Grand Inn
on August 23. My policy number is 812342.
Am I eligible for a claim?

Paul Koehler
661-650-9773

Pros:

- Mostly the original data
- Reduces re-identification risk
- Reduces accuracy loss in downstream applications

Cons:

- Compute heavy

Summary

- ☑ There is no silver bullet for all applications
- ☑ Redaction and synthetic PII are a great match for NLP
- ☑ The best systems implement a range of techniques

Thank You!

Come visit us

pieter@private-ai.com

@_PrivateAI

PRIVATE AI