WHITEPAPER

Strategizing Data Privacy

Data privacy is a critical building block of trust that your enterprise needs

Table of contents

Overview

In our earlier papers, we discussed our **RAI policy and framework**. We have also talked about transparency, accountability, and fairness. Our RAI framework explicitly mentions the need for privacy and safety. These are far more critical now since the European Data Protection Board has released a statement on the new Trans-Atlantic Data Privacy Framework that will strengthen the EU-US privacy shield.

Fractal Responsible AI (RAI) framework

Foundational Principles	Encourage Behaviors	Enabling Toolkit
Transparency All deserve white-box decision choices	Human-Centricity Evaluate blases	Codify/Democratization Workbench for implementation
Accountability Human responsible	Contestability Newspaper test	Maturity Assessment Benchmarking and audits
Privacy & Safety Respect and protection from mis-use	Explainability Decision factors traceable	Thought leadership/Education Courses, Webinars, White-papers
Fairness Inclusion and equity as default	Adaptability Augment signals and update	Success stories/Case studies Modules, examples to learn from

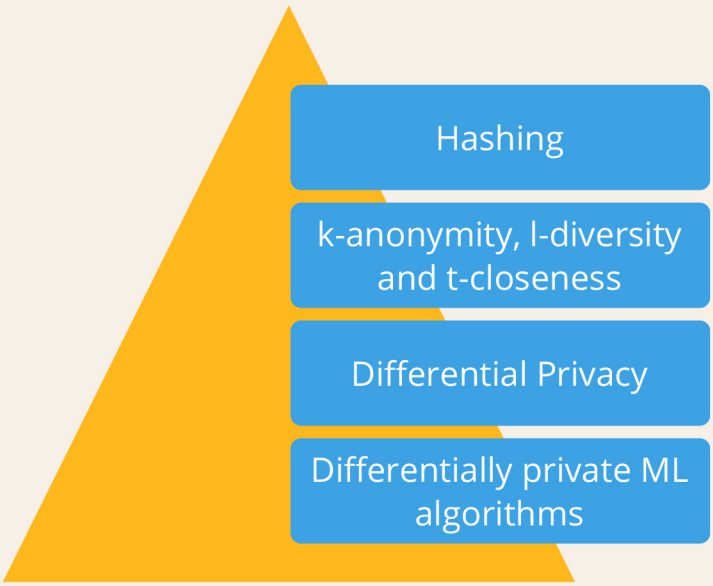
1. Why do we need privacy-preserving AI?

AI models are highly prone to privacy attacks. There have been many instances where training participants were revealed when attackers reconstructed images of their faces.

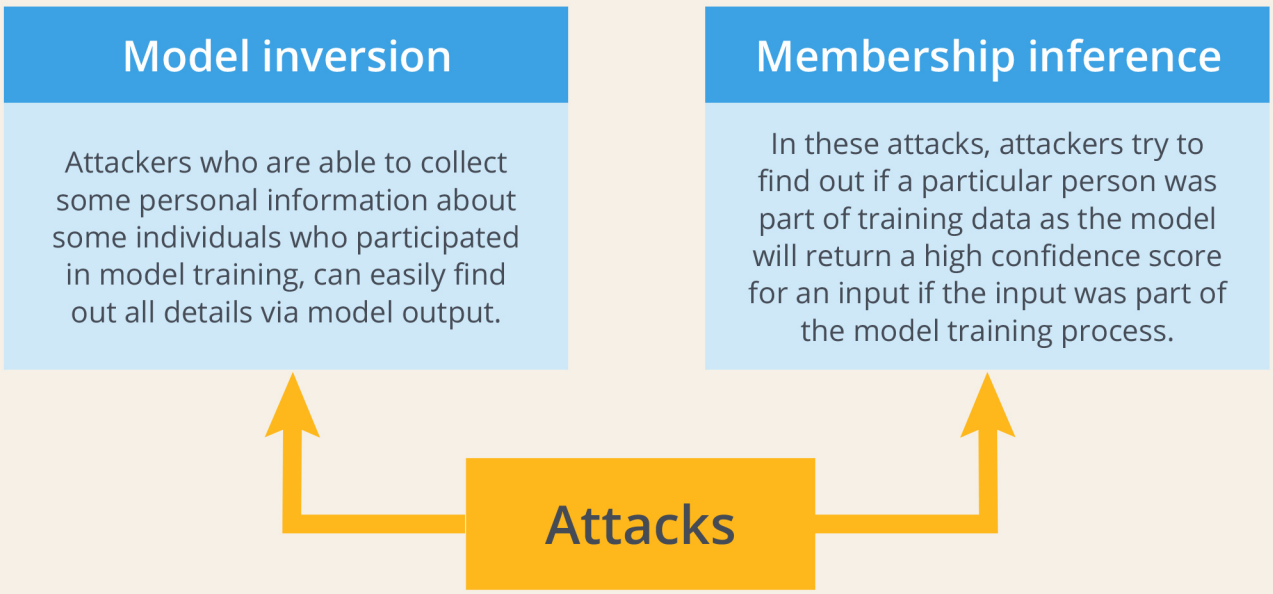
For example, in 2000, a computer scientist used anonymized hospital data with census records to identify the health records of the Governor of Massachusetts.

In another 2006 incident, a leading OTT platform released 100 million anonymized movie rating records (containing the unique subscriber ID, movie title, release year, and rating date) for an online competition to build an algorithm to predict movie ratings by a subscriber. Within two weeks of the data being released, two researchers were able to reverse-engineer the anonymized data released by the OTT platform to identify the subscriber, their viewing history, their political inclination, and other personally identifiable data.

It's prevalent for attackers to create permutations of data to be sent to the model for prediction, which in turn can be used to glean information about model parameters—or even reconstruct training data to a large extent.



Linear models are very dependent on the training data; even a small change in the training data can reveal a lot about the model parameters. Attacks on models can create problems for people using data in model training. For instance, if criminal records were used in some models and that information fell into the hands of attackers, it could cause problems for the participants. More shockingly, if a model is trained on a specific population (ethnicity, cancer patients of a particular geography, criminals from a particular county, etc.), attackers can use some of the revealed personal information to learn more details about a given individual used in the model.



Typically, attackers have two main objectives:

1. To weaken the model by skewing the model parameters and eventually compelling it to give outcomes the attacker wants to create a shadow model
2. Use training data to reveal the model parameters

Hence, it becomes critical to exhibit safety, reliability, transparency, and control. As we go along, we will see responsible AI becoming an increasingly effective way to understand data and keep it secure

2. Integration with RAI

Fractal combines privacy with other important concepts of responsible AI to make an AI system private and fair, such as adding noise that reduces the relationship between data points. Using variously tested and tried methods (e.g., the Differentially Private and Fair classification), the correlation between protected and unprotected features is lowered, thus ensuring fairness and privacy simultaneously.

> Here is one concept that makes data anonymized and secure for further validation

Data Clean Rooms and leveraging PII in a cookie-less world

A data clean room is a secure environment where Personally Identifiable Information (PII) is stored. This is first-party data obtained in its raw form from a CRM or other data sources like web browsing or transaction history. The data is first processed and anonymized, then stored in a protected, encoded manner. With the advent of a cookie-less world, data clean rooms can now use PII to create a single customer view using various data sets.

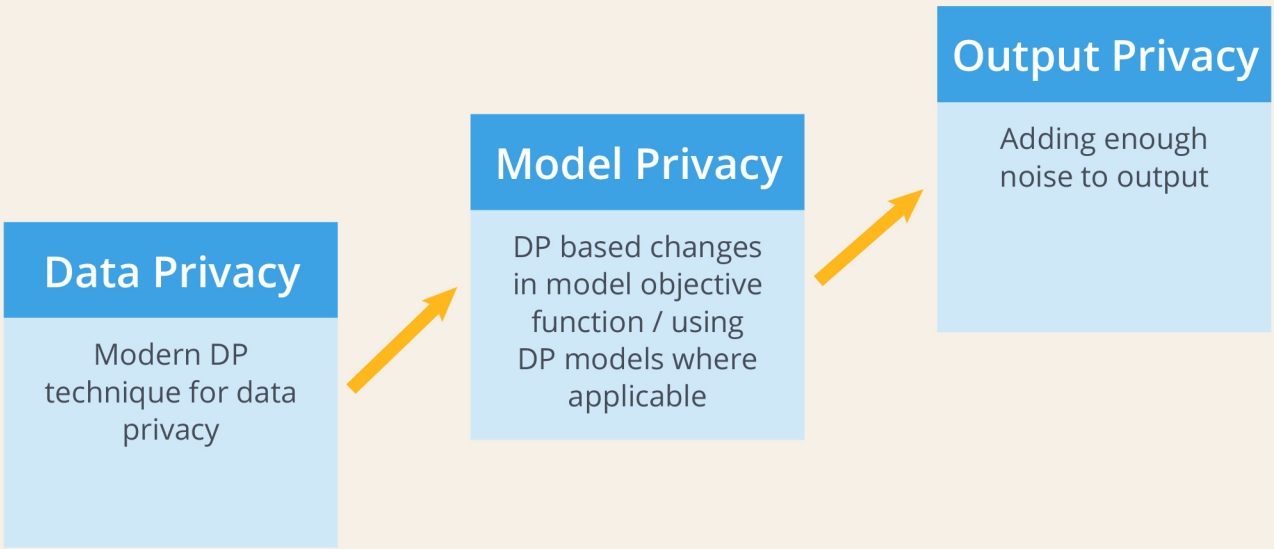
3. How do we plan to address privacy?

We have a 3-step approach to ensuring that an organization offers and enjoys security and privacy.

1. Discover

As clients define their privacy strategy, they must understand their current state of maturity compared to their goals. A structured framework (RAI Privacy cards) can help clients quantify their maturity, identify gaps, and take appropriate steps to meet their aspirations.

During the discovery phase, we seek to understand the technology and Machine Learning (ML) landscapes in which the client operates. Understanding the technology landscape is critical for understanding how and where privacy needs to be integrated. A privacy meter algorithm is also used to measure the riskiness of ML algorithms concerning privacy.



2. Analyze

In this stage, a thorough gap analysis is done to investigate the need for privacy and the risks associated with the current process. A model risk algorithm is used to determine how prone the models are to attacks and privacy breaches. Then, we consolidate a list of customized methods, modules, algorithms, and processes to plug any potential security or privacy holes, wrapping the models inside privacy-preserving techniques. Clients are also provided options for privacy-preserving strategies for future models.

3. Roadmap

After an in-depth analysis, the team assesses the resources available within the organization. It develops a roadmap of activities needed for preserving privacy. This includes a build-vs-buy comparison in which third-party tools are evaluated for addition to the existing AI toolkit. We also develop customized solutions for clients to fit their needs and privacy. The roadmap includes project plans, timelines, budgets, and outcomes to make economically feasible decisions. We also provide a monitoring dashboard to continuously analyze (and report) privacy needs across various metrics.

4. Use case

We implemented privacy and transparency for a use case on COVID Social Distancing Compliance with CCTV / drone video data. Few reusable modules were developed that kept an image private by evaluating the "objects" without revealing the identity of any without permissions, blurring all the detected objects so that any of the views on video streams would also protect the privacy of the objects. This maintained privacy by allowing identification of the objects with encrypted keys only.

Additional modules provided identity detection access only based on the keys that provided full transparency to the operations on identity retrieval.

5. Conclusion

Privacy today is one of the biggest pain-points for corporations. While working with various clients, we realized that there is no one-solution-fits-all approach that works. Thus, it is important to see privacy-preserving techniques from the lenses of data and algorithms combined with sensitivity and impact on the bottom line.

fractal dimension

We believe complex problems need to be looked at through multiple lenses simultaneously to be grasped. With the new lens new dimensions emerge, thus making complexity more evident and solvable.

How is Fractal Dimension set up to do it?

We identify complex and unstructured problem themes in the industry that are relevant. We invest in building expertise and a dimensionalized point of view around it.

We engage clients via 'slow-thinking' workshops and co-creation jams to curate our perspective for their problem. We invest in architecting an end-to-end state-change program.

We partner with client teams at Fractal to deploy cross-functional solutions and support them in helping clients realize value ROI.

Want to find out more on how our approach can help your business? Reach out today at dimension@fractal.ai

Our experts

Sagar Shah
Client Partner, Strategic Center

Akbar Mohammed
Lead Data Scientist,
Strategic Center

Sray Agarwal
Principal Consultant,
Strategic Center

Krutika Choudhary
Senior Consultant,
Strategic Center

Enable better decisions with Fractal

Fractal is one of the most prominent players in the Artificial Intelligence space. Fractal's mission is to power every human decision in the enterprise and bring AI, engineering, and design to help the world's most admired Fortune 500® companies.

Fractal product companies include Qure.ai, Crux Intelligence, Theremin.ai, Eugenie.ai & Samya.ai.

Fractal has more than 2,300 employees across 16 global locations, including United States, UK, Ukraine, India, and Australia. Fractal has consistently been rated as India's best company to work for, by The Great Place to Work® Institute, a 'Leader' by Forrester Research in its Wave™ on Specialized Insights Services, Computer Vision & Customer Analytics and as an "Honorable Vendor" in 2021 Magic Quadrant™ for data & analytics by Gartner.

Corporate Headquarters
Suite 76J,
One World Trade Center, New York,
NY 10007

Get in touch