



E-BOOK

PRIVACY-PRESERVING ANALYTICS FOR HEALTHCARE

Using the LeapYear Platform and Differential Privacy to unlock value from sensitive data

INTRODUCTION

This e-book covers a broad range of applications for privacy-preserving reporting, analytics, and machine learning in the healthcare ecosystem. Healthcare providers, payers, pharmacy benefit managers (PBMs), and the surrounding technology industries provide key services to communities and individuals to keep people healthy and live longer and more productive lives. Each of the players in this ecosystem deals with many types of sensitive data, including patient outcomes, treatment costs, and drug efficacy patterns. Increasingly, companies are looking to generate additional value from their datasets—both through internal use and third-party partnerships.

One of the central challenges limiting the use of healthcare data is patient privacy. It is a well-known fact that [traditional methods](#) of protecting privacy, such as data masking and anonymization, can be easily circumvented to compromise protected health information (PHI). Moreover, these approaches involve modifying or redacting data, thus creating inefficiencies in data use and reducing the value of information.

Therefore, healthcare companies require a new way to make datasets available to their partners, to third-parties, and across internal silos. For the first time, the LeapYear platform enables healthcare companies to share and extract value from sensitive data sets while providing ironclad privacy guarantees to their customers and partners. The platform fundamentally changes how companies can work with and leverage their data, achieving results such as:

- [A top-5 payer achieved a 50% reduction in the number of people who had direct access to sensitive data and eliminated exposure risks of re-identification attacks on their data.](#)
- [A payer was able to leverage previously inaccessible data \(social determinants of health\) in their internal modeling efforts without introducing additional risk.](#)
- [Multiple research hospitals are able to share data with each other for medical research, with improved efficiencies and reduced patient exposure risk.](#)
- [A payer commercialized patterns in their data, such as benchmarks and trends, which are learned from combining data across multiple employers.](#)
- [A pharmacy benefits manager \(PBM\) and a population health manager pursuing direct commercialization of healthcare data recognized nine-figure business opportunities.](#)

This e-book begins with an overview of the LeapYear technology and then is organized by use cases for the healthcare ecosystem. Each section outlines the challenges created by various regulatory and internal data protection frameworks, which create barriers to effective data use, and exhibits how LeapYear customers have overcome them using the LeapYear platform in order to drive significant business value.

THE LEAPYEAR PLATFORM

LeapYear is the world's [first platform](#) for differentially private reporting, analytics and machine learning. The platform embeds mathematically proven privacy into every computation enabling analysts and data scientists to generate insights from data without exposing the data itself. The platform is based on the mathematical standard of [differential privacy](#), which enables companies to compute on data across business lines, geographic boundaries, and organizations while preserving locality, confidentiality, and value.

Continue →

Using LeapYear, several top healthcare companies have been capable of the following:

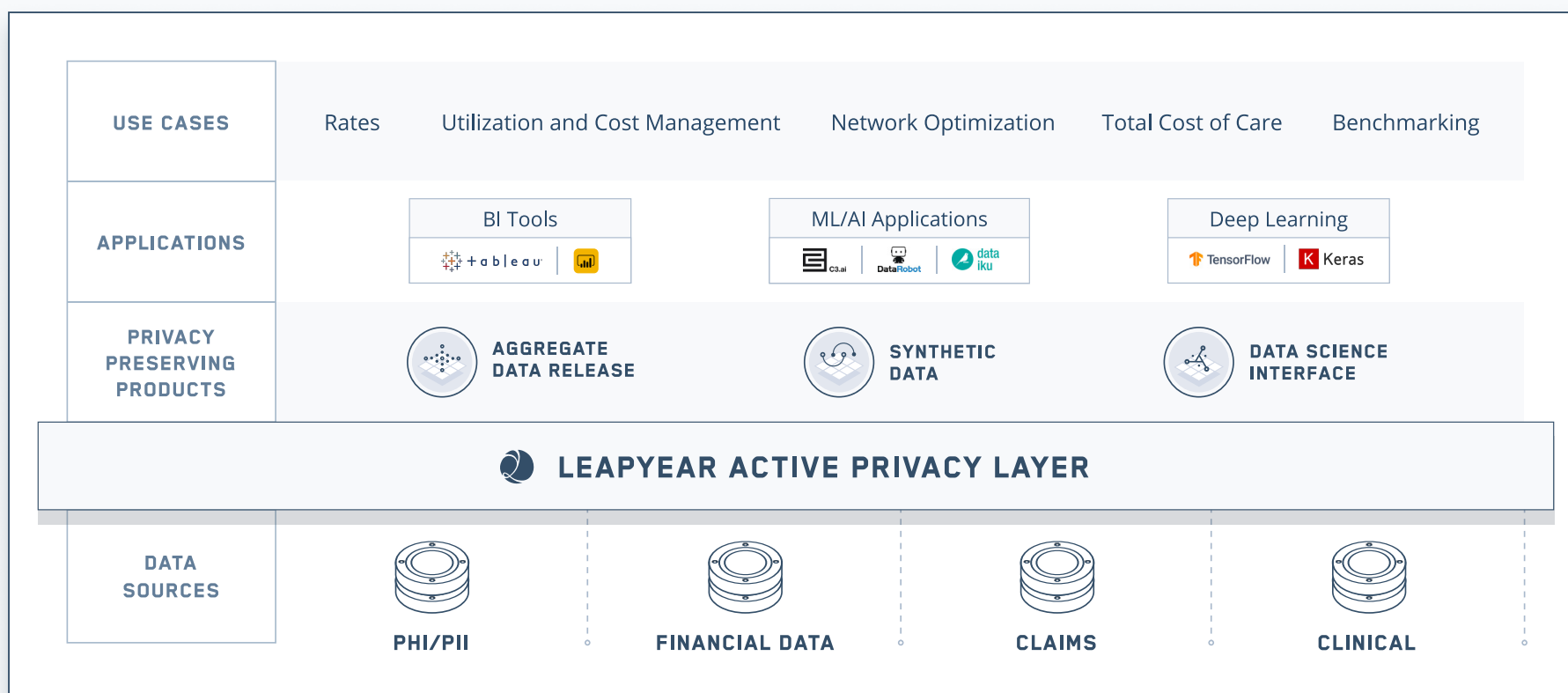
- Reducing privacy breach risk.
- Enabling analytics across internal silos, partners, and third-parties.
- Increasing the resolution and granularity of existing datasets as they are shared.
- Addressing patient concerns of privacy and confidentiality.
- Optimizing their business with datasets that were previously inaccessible.
- Developing entirely new information-based revenue streams.

Unlike traditional approaches which use pre-defined rules to modify or redact information from a dataset, LeapYear’s platform dynamically introduces privacy based on the context of each computation. The ability to react dynamically to each computation eliminates the need to make assumptions about data use and dramatically simplifies data preparation and governance schemas.

LeapYear is designed to embed seamlessly into the analytics ecosystem, providing a privacy-preserving computational layer for applications across:

- Reporting and business intelligence
- Statistical analysis
- Application testing and development
- Data science and engineering
- AI, machine learning, and deep learning

LeapYear’s technology is based on decades of research and 100’s of academic papers, and has been evaluated by world experts in differential privacy and various privacy regulations, including HIPAA, GDPR, and CCPA. LeapYear is deployed in production, at multi-petabyte scale, across global 1000 financial institutions, healthcare companies, and insurers.




USE CASE:**BROADEN DATA ACCESS WHILE REDUCING RISK AND INEFFICIENCY**

VALUE PROPOSITION: Reduce the number of people who have direct access to PHI data. Eliminate costly and time consuming data preparation.

DESCRIPTION: LeapYear diminishes a major risk point in many healthcare organizations: reducing the number of people who have direct access to PHI data without reducing their ability to use that data for their day-to-day work.

Often, a large number of people inside an organization have direct access to sensitive datasets for use in analysis. However, each person who has access to protected data is a risk point, as even de-identified data can be used to aggregate information and reverse engineer information about individuals. In most cases, however, data is used to generate statistical observations and insights about the underlying populations. This is important, as it means that the analyst does not actually require direct read access to the individual records. What they really need is a method to securely and privately generate insights on the data. LeapYear provides a new way to work for analysts of PHI data: gain full access to data to generate aggregates, statistics, and powerful models without needing to take possession or even see the underlying records.

LeapYear also dramatically improves the “time to value” for each dataset. As a system, LeapYear alters the way companies prepare and use data for analytics, shifting from a “certify each dataset” to a “certify the system” mentality. This table summarizes the three major challenges encountered in preparing data for use and how LeapYear addresses each issue.

CHALLENGE	TRADITIONAL METHODS	 LEAPYEAR
Prepare data for use	Companies build complex, time consuming processes to de-identify or aggregate data.	No extra prep is required, all fields are secured with mathematically proven privacy automatically.
Grant access to data	Companies build complex processes to understand why access to sensitive data is needed, who can approve such requests and finally grant direct access to data.	LeapYear ensures no direct access to data is needed, and that every query is mathematically proven private, eliminating the need for complex data access processes.
Maximum data value and veracity	Most methods to prepare data for use involve removing or aggregating fields, which results in stripping out information content (think signal!) and potentially introducing spurious relationships.	Analysts gain full granular access to use data for analytics, ensuring maximum value and veracity of each data asset.

► IMPACT:

A top-5 payer achieved a 50% reduction in the number of people who had direct access to sensitive data and eliminated exposure risks of re-identification attacks on their data. Moreover, prior to LeapYear, 40% of the payer’s de-identified records were at risk of being re-identified. LeapYear eliminates this risk.

► IMPACT:

A payer was able to leverage previously inaccessible data in their internal modeling efforts without introducing additional risk. Specifically, the payer was able to provide data scientists access to Social Determinants of Health data as well as PHI elements, such as date of admission, procedure dates, gender, zip code, and date of birth. Typically, these fields are restricted due to privacy concerns and HIPAA regulation; however, with LeapYear, the payer could use these fields for analytics while protecting patient privacy.

USE CASE:**PRIVATE AND SECURE THIRD-PARTY DATA SHARING****VALUE PROPOSITION:**

Significant improvement in patient privacy and data control.

DESCRIPTION:

There are many opportunities for healthcare companies to share their data with third-parties to support cooperative research, facilitate business development, or explore pathways for value-based care programs. Payers may want to explore a cross-payer analysis of outcomes, costs, and patient engagement to increase overall improvements in population health. Similarly, healthcare providers want to enable collaborative research between hospitals on data sets across various patient populations and geographies.

These initiatives for sharing data, however, are subject to stringent regulatory frameworks. Moreover, to make them truly useful, the datasets need to be readily available in near real time, with access to important (but sensitive) fields. The following table outlines four key data sharing challenges and how the LeapYear platform addresses each issue:

CHALLENGE	LEAPYEAR CAPABILITY	OUTCOME
Ensure patient / information privacy.	Differential privacy across the entire analytics workflow.	Mathematically proven privacy—holds even as the data updates and is joined with other datasets
Include sensitive fields for key research.	Differentially private calculations on all fields including the most sensitive.	Improved signal in datasets for analysis and model-building
Maintain data control.	API driven access to data.	Third-parties cannot exfiltrate or reconstruct any data, including anonymized data
Speed of data access/use.	Every field, every row is protected automatically.	No time-consuming data operations processes to use and share data. For example, no need to define which fields are sensitive

[Continue →](#)

Solving these four challenges means the data owner can safely, and with maximum value, share data with a third-party. Since the LeapYear platform provides mathematically proven privacy, data owners can enable detailed granular access to data to generate statistical insights for their partners, all while maintaining compliance. Additionally, in the best-case scenario, partners can achieve self-service access for analytics; this is extremely powerful as it simplifies the need to “guess” at what someone wants to analyze, and instead quickly allows partners to capture value from this data access.

► **IMPACT:**

Multiple research hospitals are able to share data with each other for medical research. In the past, traditional approaches to de-identifying data were reducing data quality for research, took time to implement (specifically in cases where data from multiple providers were joined, requiring recertification), and created known privacy risks. With LeapYear, all the data could be used, including sensitive fields, with a privacy platform that is fully automated and mathematically proven.

► **IMPACT:**

A payer often has employer clients who are interested in understanding data patterns, such as benchmarks and trends, which can be learned from combining data across multiple employers. In this use case, there was a privacy and a confidentiality concern: the insights disclosed from the combined dataset must protect patient data as well as the confidentiality of employers. With LeapYear, the payer’s analysts were able to safely combine data from across employers to provide insights while ensuring mathematically proven privacy and confidentiality.

USE CASE:

DATA COMMERCIALIZATION

VALUE PROPOSITION: Generate new revenue streams from data assets.

DESCRIPTION: Companies throughout the world are looking to healthcare data to generate competitive differentiation in the market. For example, in financial services, vertical asset managers are interested in using alternative data sources from healthcare-related organizations to drive their strategies for portfolio growth. Another example comes directly from healthcare and pharmaceutical companies, in which each party is looking for access to data that can help improve outcomes, streamline costs, and optimize the time-to-market for new therapies. In many cases, these same companies are willing to purchase access to data sets in order to drive their critical business initiatives.

Typically, owners of healthcare data understand that opportunities exist to commercialize their information, but they face various challenges on their journey to maximize the value of their data assets, such as:

- The use and distribution of personal data is strictly limited by privacy regulations (international, national, and state-level)
- Client or patient confidentiality, data use agreements, infosec policies, and reputational risk further restrict monetization of data
- Datasets that have been anonymized or masked are substantially less valuable and can be easily compromised to exploit PII and other confidential data
- Developing commercial relationships with buyers of data can take years

Because of the core properties of the LeapYear platform, data owners are increasingly comfortable providing broader access in monetization channels to previously limited or aggregated datasets. This access may even include self-service functionality, drastically increasing value for the data consumer.

Combined, the four fundamental tenets of the LeapYear platform enable the rapid protection of (and remote access to) sensitive data assets for monetization opportunities.



Figure 1. The four key components of the LeapYear platform enable rapid, safe and private commercialization of sensitive data.

It should be noted that LeapYear is never a custodian or broker of data, and it does not generate any analysis or insights based on the data. LeapYear is purely a technology platform that facilitates the secure access and analysis of data assets between a data owner and any parties with whom the data owner is transacting.

In addition to providing the technology platform, LeapYear also has partnerships with trusted organizations that provide a commercial vehicle for data monetization. One example of such is a global financial institution that has developed a data brokerage business line. The institution has the relationships, infrastructure, and processes to automate the marketing, sales, and distribution of datasets to data buyers in the financial services community.

Through these partnerships, LeapYear is able to provide healthcare companies a turn-key solution for safely commercializing their data.

► **IMPACT:**

A large pharmacy benefits manager (PBM) and a population health manager are pursuing the direct commercialization of healthcare data through LeapYear and its partners. Historically, these organizations were not comfortable commercializing their data at all or were commercializing only a limited sample of their data through traditional de-identification (which could be re-identified). Through LeapYear, these organizations found a technology platform and commercial vehicles to recognize nine-figure business opportunities.

SUMMARY

Healthcare companies, including payers, providers, PBMs, and population health platforms, are stewards of highly sensitive medical data. They have a responsibility and regulatory obligation to protect this data. At the same time, the healthcare ecosystem needs to leverage data to reduce costs, increase efficiency, and improve patient outcomes. Traditional methods, such as anonymization and data masking, are insufficient — although they meet compliance requirements, they do not protect patient data and they also reduce the value of a dataset. These approaches create critical security risks and eliminate opportunities for innovation. The LeapYear platform changes this, and has been used across the entire healthcare ecosystem to bring significant value to customers for optimizing the protection and use of their sensitive data.