



## Low-risk, high-utility data anonymization with Smart De-Identifier

As healthcare progressively adopts digital models, organizations need data de-identification capabilities to ensure data security, while preserving data utility for business insights. CitiusTech's Smart De-identifier enables fast and effective PHI anonymization, using proven methodologies and user configurations, to provide an optimal balance between risk and data utility.

- Structured & un-structured data
- Expert determination & safe harbor
- Risk & data utility metrics

## Introduction

Healthcare data has been increasing exponentially over the last few years, driven by greater digitalization, cloud computing and increased data sharing amongst healthcare ecosystems. Today, most healthcare organizations store large volumes of patient data in different formats, which need to be constantly secured. Organizations will need to leverage robust and standardized data de-identification techniques to:

- **Protect patient health data as per HIPAA, GDPR, EMA and Health Canada regulations**
- **Avoid penalties due to non-compliance of regulations / breach of patient data**
- **Set up secure processes to share and utilize high-value healthcare data for clinical innovations**
- **Continuously re-assess and mitigate the risk of re-identification for large volumes of historic data**

Healthcare organizations need effective tools for data anonymization to protect and utilize sensitive patient health data at reduced cost and scale



## Health Data De-identification: Barriers to Adoption

1

Traditional methods of data de-identification are either manual or rule-based, which incur high cost and effort

2

Diverse data types, formats, and interoperability standards make it difficult to take a standardized approach towards data de-identification

3

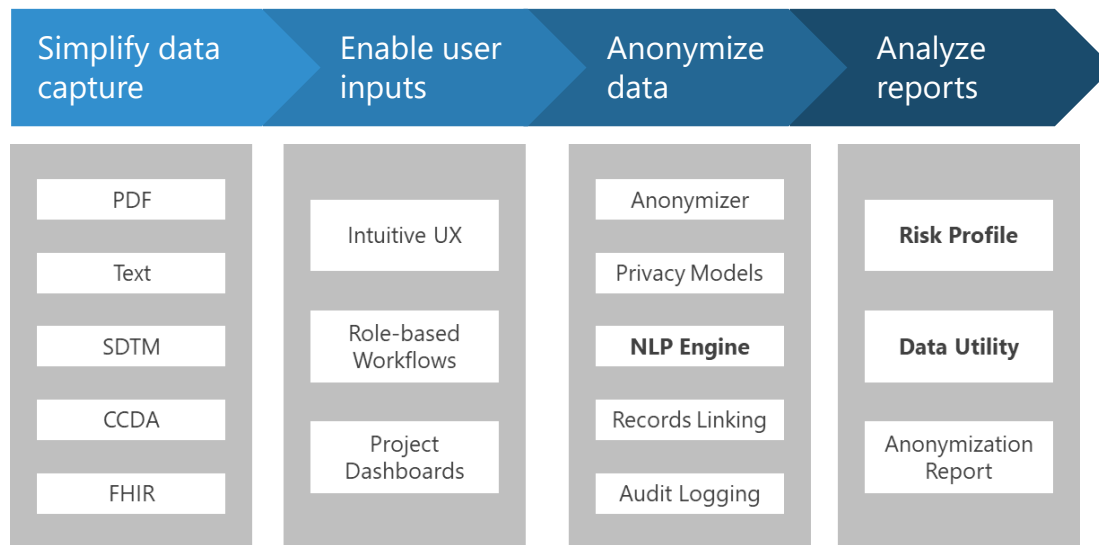
Difficulty in staying updated with multiple, and ever-changing rules and regulations across geographies

4

Lack of flexibility in existing solutions make it difficult to customize configurations and develop company-specific best practices

## CitiusTech Data De-identification Approach

CitiusTech understands the unique requirements for healthcare data de-identification. It provides effective and scalable anonymization capability for using and sharing anonymized data, enabling better data analytics, software development, and testing.



CitiusTech's data de-id services leverages healthcare domain expertise, accelerators, toolsets, together with engineering and data management practices and a broad partner ecosystem

## Smart De-identifier: Solution Overview

CitiusTech's Smart De-identifier is a microservice that enables fast, effective PHI anonymization using multiple methodologies. It leverages user-specific configurations and ML-based analytics with adherence to healthcare regulations.

Multiple De-id Techniques	Leverage user-configurable suppression rate, privacy criteria, K-anonymity, etc.
Optimized Outcomes	Compare effectiveness between multiple methodologies including Safe Harbor and Expert Determination.
Embedded Intelligence	Implement ML models and statistical methods for intelligent detection of high-risk PHI attributes and re-identification risk.
Customizable UI	Easy-to-use UI for user-specific configurations and detailed project analytics.
Extensive Data Support	Support for a wide variety of structured and unstructured input data e.g. CSV, XLS, SAS, XPT, PDF, TXT, etc.

## The Smart De-identifier Advantage

CitiusTech's Smart De-identifier combines healthcare expertise, data science proficiency, and statistical methods of de-identification to enable protection of sensitive health data while extracting high value insights.

- **White-boxed approach** with shareable code base, components and design for customer-specific configurations
- **Easy integration** with interface engines, existing workflow tools, ETL processes and installation options **on-premise or on cloud**
- In-built **compliance to EMA policy** by utilizing ARX core library as per policy guidelines
- Intelligent PHI detection and auto-suggestions on generalization of healthcare datasets
- Leverages CitiusTech's **healthcare data accelerators** for onboarding of additional data types (DICOM, Genomics) and **NLP accelerator** to extract key clinical entities

## About CitiusTech

With over 4,000 professionals worldwide, CitiusTech enables healthcare organizations to drive clinical value chain excellence, across integration & interoperability, data management (EDW, Big Data), performance management (BI / analytics), data science (predictive analytics, Machine Learning, AI) and digital engagement (mobile, IoT).

### CitiusTech Data Science Proficiency

The Data Science & Consulting Proficiency is a multidisciplinary team of data scientists, AI professionals, business analysts, statisticians, data architects, business process consultants, and clinical informatics professionals.

The 200+ member team has strong expertise in enabling clients to transform their customer experience, health outcomes and business operations through Data Science and AI.

CitiusTech's Embedded Intelligence Framework is an outcome-focused approach to drive operational excellence. It makes current processes smarter through continuous, incremental innovations and ensures minimal disruption to the organization's core business processes.

**\$205+ Mn**  
in revenue



**4,000+**  
healthcare IT professionals

**40 Mn+**  
lives touched

**69**  
NPS - highest in the industry!

**110+**  
healthcare customers



This document is confidential and contains proprietary information, including trade secrets of CitiusTech. Neither the document nor any of the information contained in it may be reproduced or disclosed to any unauthorized person under any circumstances without the express written permission of CitiusTech.