



EUROPEAN  
MEDICAL  
WRITERS  
ASSOCIATION



*The Resource for Medical Communicators*

## Clarity and Openness in Reporting: E3-based

**Webinar by Obaraboye Olude, Privacy Analytics**  
**“Statistical anonymization software for utility-preserving privacy protection in clinical documents for public disclosure”**

Hosted by Alison McIntosh, for The CORE Reference Team  
11 June 2025

# Agenda

- Recap: CORE Reference on LinkedIn
- Automated Anonymisation Processes
- External speaker: Obaraboye Olude, Privacy Analytics
- Questions from webinar attendees

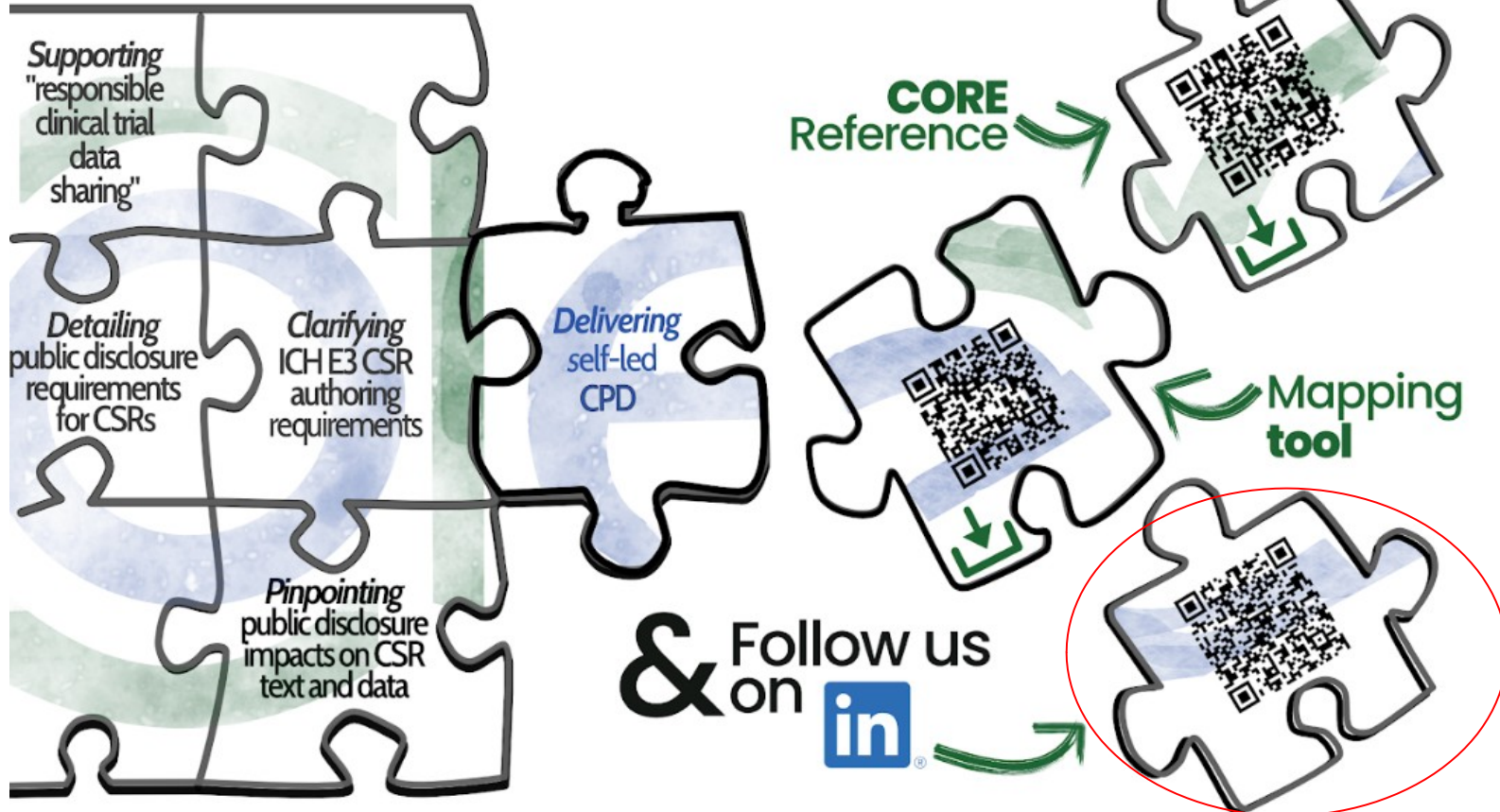
# CORE Reference Team

- Sam Hamilton – Chair
- Vivien Fagan
- Alison McIntosh
- Zuo Yen Lee
- Raquel Billiones
- Jen Bell

# LinkedIn



## An Open Access Resource to Support Authoring of Clinical Study Reports for Interventional Studies



# LinkedIn

<https://www.linkedin.com/company/the-core-reference-project>

Please share with your networks - forward, like, comment

# Automated Anonymisation Processes

# Commercial Tools & Services



Shadow – Data Anonymization Tool.



- Tools are operated by Sponsor/CRO under license or the vendor offers anonymisation as a service
- Sponsor/CRO should investigate and complete own cost analysis

# Webinar



External speaker: **Obaraboye Olude**, Privacy Analytics

“Statistical anonymization software for utility-preserving privacy protection in clinical documents for public disclosure”



# Statistical anonymization software for utility-preserving privacy protection in clinical documents for public disclosure



June 11, 2025

Obaraboye Olude MBBS,MPH  
Sr Advisor and Technical Manager, Clinical Trial Transparency

# Disclaimer

As with any attempt to estimate future events, the forecasts, projections, conclusions, and other information included herein are subject to certain risks and uncertainties and are not to be considered guarantees of any particular outcome.

All reproduction rights, quotations, broadcasting, publications reserved. No part of this presentation may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without the express written consent of IQVIA.

# IQVIA's Privacy Analytics

## Privacy Solutions



### Solutions combining proven technology and expertise

We deliver solutions (software and services) that empower our clients to use sensitive data safely and responsibly for innovation that benefits everyone.

## Pioneer & Leader



### Pioneer in data and document anonymization/redaction

We have informed 20+ regulatory guidelines and standards, anonymized thousands of datasets and millions of pages.

## An IQVIA Company



### A wholly-owned subsidiary of IQVIA

As an IQVIA company, we benefit from the global scale of 86,000+ employees in over 100 countries globally and offer an award-winning workplace.

# PA Platform for anonymization and redaction of clinical documents



## PA Platform Mirage module



### **Fine-tuned, domain-specific, large language model (LLM)**

Powers high-speed and high-accuracy detection of personal information



### **Statistical anonymization**

High-utility, industry best practice for maximum transparency



### **Redaction**

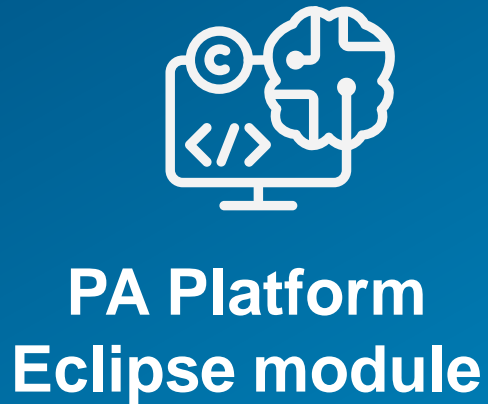
Fast and cost-effective way to handle unforeseen or time-sensitive requests



### **As-a-Service or License**

Benefit from the experience of our 100-strong team of experts or deploy Mirage on-premise

# PA Platform for anonymization of clinical trial datasets



## **Statistical Anonymization**

Flagship product for structured data anonymization with defensible proof of privacy protection



## **Tailor-Made for Trials Data**

Auto-classification for CDISC-compliant datasets, PHUSE-shifting



## **Flexible Transformation Strategies**

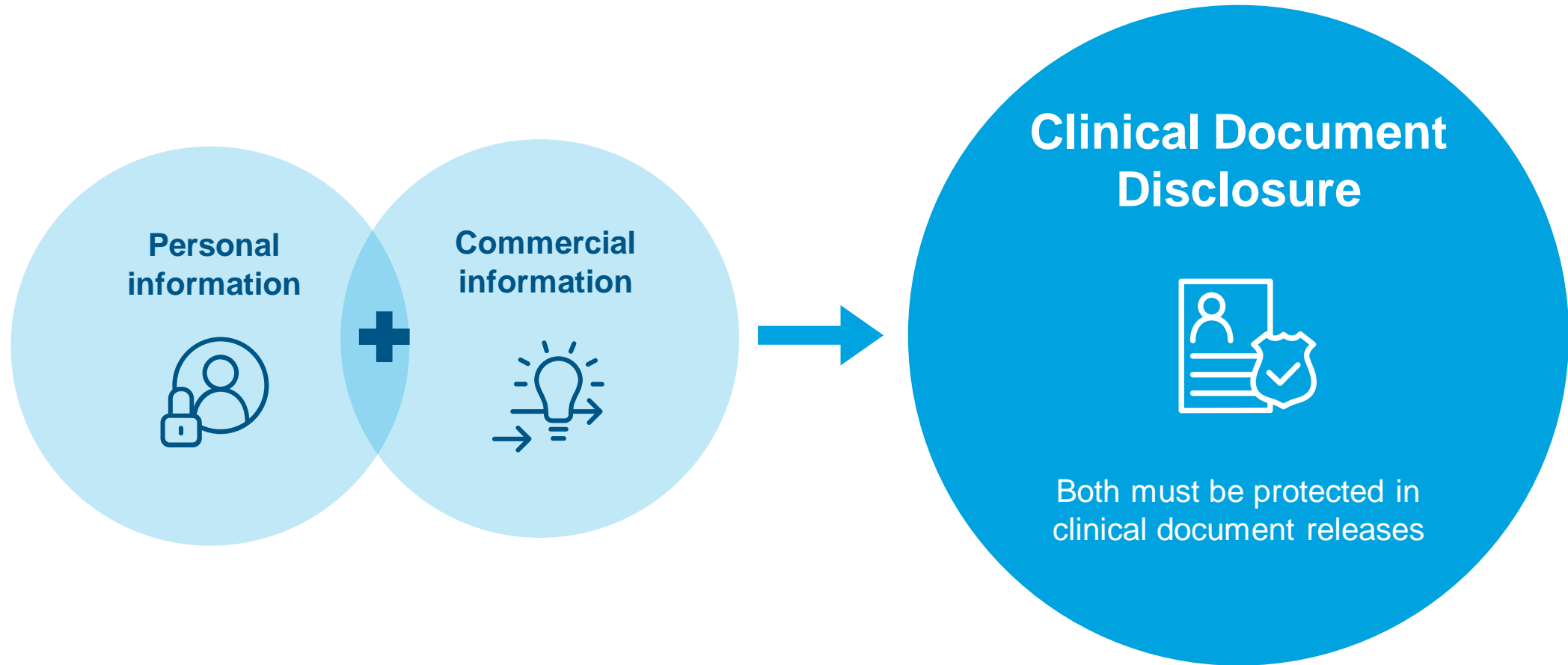
Optimizes utility to achieve sub-threshold re-identification risk for a given release context



## **As-a-Service or License**

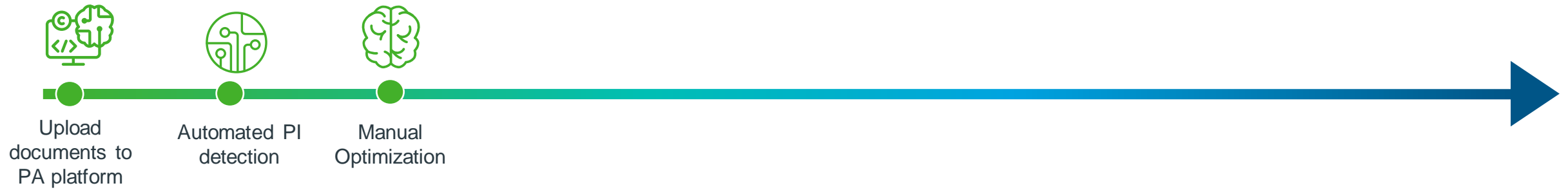
Benefit from the experience of our 100-strong team of experts or deploy Eclipse on-premise

# Sponsors must protect personal and commercial information when disclosing documents publicly under EMA Policy 0700, Health Canada Public Release of Clinical Information



A large, dark blue rounded rectangle with a teal rounded rectangle on top and a teal rounded rectangle on the bottom left, all on the left side of the slide.

# **Workflow for anonymization of clinical trial documents**



The screenshot displays the Privacy Analytics interface. The top navigation bar shows the path: **Projects** → **Demo** → **Trial 1** → **Tag Documents**. The main content area shows a document titled **Clinical Study Report**. The text of the report is as follows:

**Subject 122-330**

Subject 122-330, a 62-year-old white male with a history of diabetes, signed informed consent and was enrolled in Study 1B on 01-Jan-2010.

The subject's medical history includes hypertension, myopia, bronchitis, osteoarthritis, gastroesophageal reflux, carpal tunnel syndrome, seasonal allergies, irritable bowel syndrome, hyperlipidemia, and gallstones.


On the right side, there is a sidebar with a list of tags. The tags are:


- DATE**: 01-Jan-2010
- MEDICAL\_HISTORY**: hypertension
- MEDICAL\_HISTORY**: myopia

The interface also includes a toolbar with options like Zoom Out, Zoom In, Fit Width, and a search bar.





PRIVACY ANALYTICS  
an IQVIA company

Projects → Demo Trial 1 → Subject Profiles 

Subject	PII Type	Value	Document	Page
122-330	DATE	23-Sep-2011	Trial 1_Clinical_Study_Report.p	2
122-330	MEDICAL_HISTORY	bronchitis	Trial 1_Clinical_Study_Report.p	2
122-330	MEDICAL_HISTORY	carpal tunnel syndrome	Trial 1_Clinical_Study_Report.p	2
122-330	MEDICAL_HISTORY	diabetes	Trial 1_Clinical_Study_Report.p	2



PRIVACY ANALYTICS

Projects → Demo Trial 1 → Overview

L

H

Low Risk

A risk measurement process was run and it was found to be low risk for re-identification

SUBJECTS 10

Direct Identifier Risk

Result of 0.0092683 does not exceed risk threshold 0.02521282

Low

Quasi Identifier Risk

Result of 0.00294962 does not exceed risk threshold 0.09

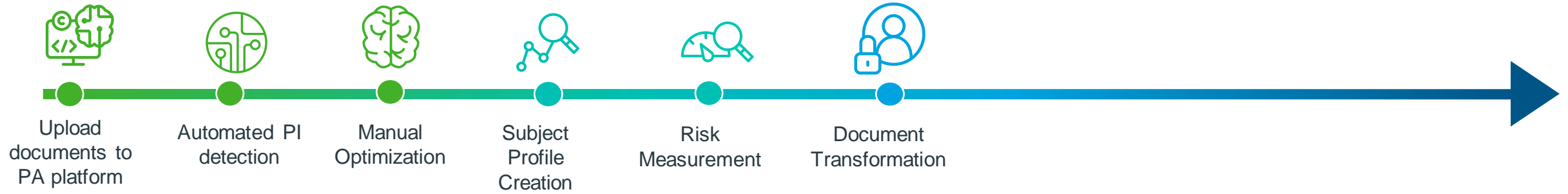
Low

PRIVACY ANALYTICS

Projects → Demo Trial 1 → Overview

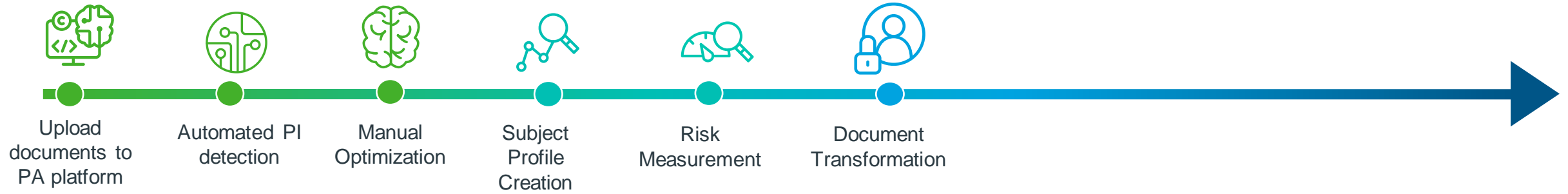
View by: Subjects PII Types

SUBJECT	RISK	
122-330	<div><div>L</div><div>H</div></div> Low Risk0.0880539	25
112-411	<div><div>L</div><div>H</div></div> Low Risk0.03787166	25



Subject 123-425 was a 32-year-old with a medical history of coeliac disease.

The subject was enrolled on 01-Jan-2010 and experienced an adverse event on 04-Jan-2010.



Subject 123-425 was a 32-year-old with a medical history of coeliac disease.

The subject was enrolled on 01-Jan-2010 and experienced an adverse event on 04-Jan-2010.

Pseudonymization with FPE

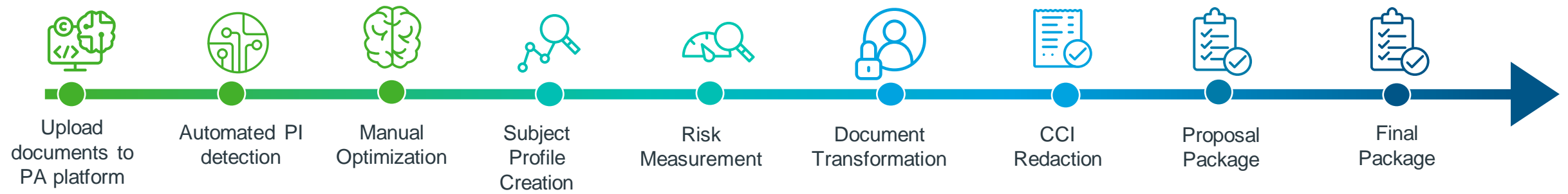
Generalization

Subject 627-889 was a 30-35-year-old with a medical history of autoimmune disorders.

Generalization (SOC)

The subject was enrolled on 03-Mar-2009 and experienced an adverse event on 06-Mar-2009.

Date offset



On the left side of the slide, there are several overlapping blue shapes. A large, dark blue rounded rectangle is the central element. Above it is a smaller, lighter blue rounded rectangle. Below it is a teal rounded rectangle. A thin, light blue line also curves around the bottom right of the large dark blue shape.

## **Demo - Statistical anonymization with sample file**



**Questions?**

**Thank you! For more information:**

**Obaraboye Olude MBBS, MPH**  
**Sr. Advisor, Clinical Trial Transparency**  
**[oolude@privacy-analytics.com](mailto:oolude@privacy-analytics.com)**

