



Anonymization Insights from 2025

Honz Slipka, MSc
Assoc. Principal Transparency Specialist

CERTARA 

Learning from EMA Policy 0070 Submissions



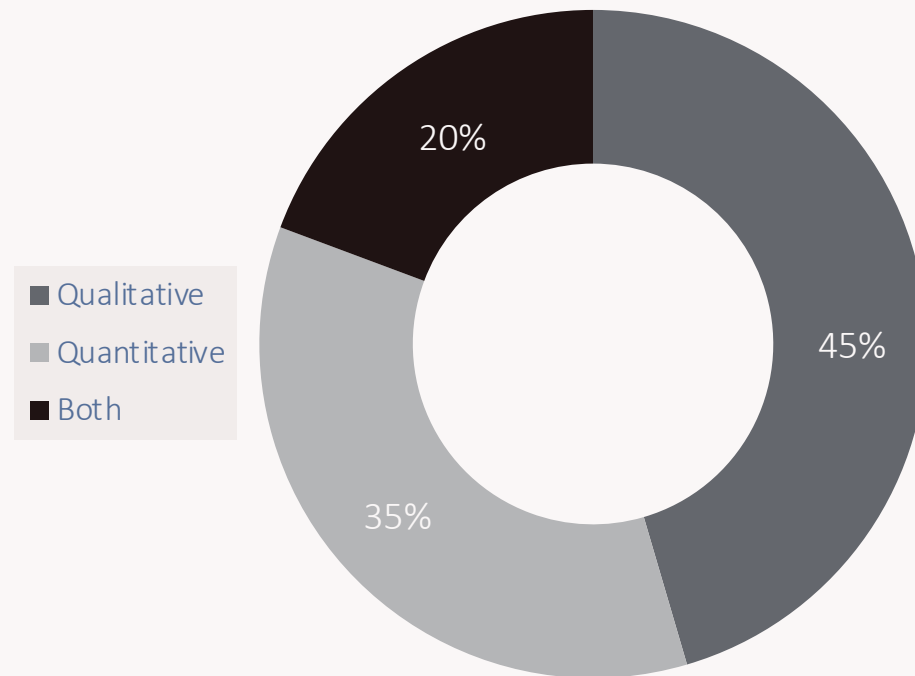
- Analysis of 2025 EMA Policy 0070 anonymization reports
- Dataset includes ~90 submissions and >2000 anonymization decisions
- Covers 26 EMA indirect identifier categories

Goal: identify patterns that improve efficiency, consistency, and regulatory acceptance

Evidence from real submissions can help standardize anonymization approaches and reduce unnecessary rework.

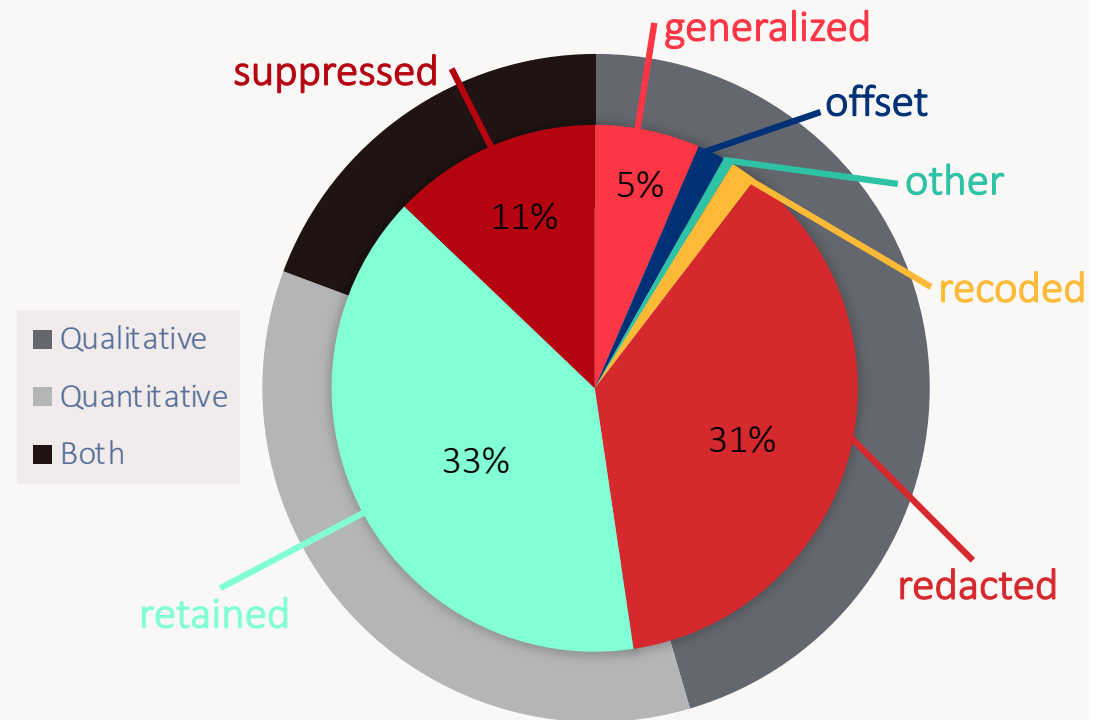
Structure of EMA Indirect Identifier Data

- EMA defines 26 categories of indirect identifiers
- Sponsors apply multiple anonymization techniques



Structure of EMA Indirect Identifier Data

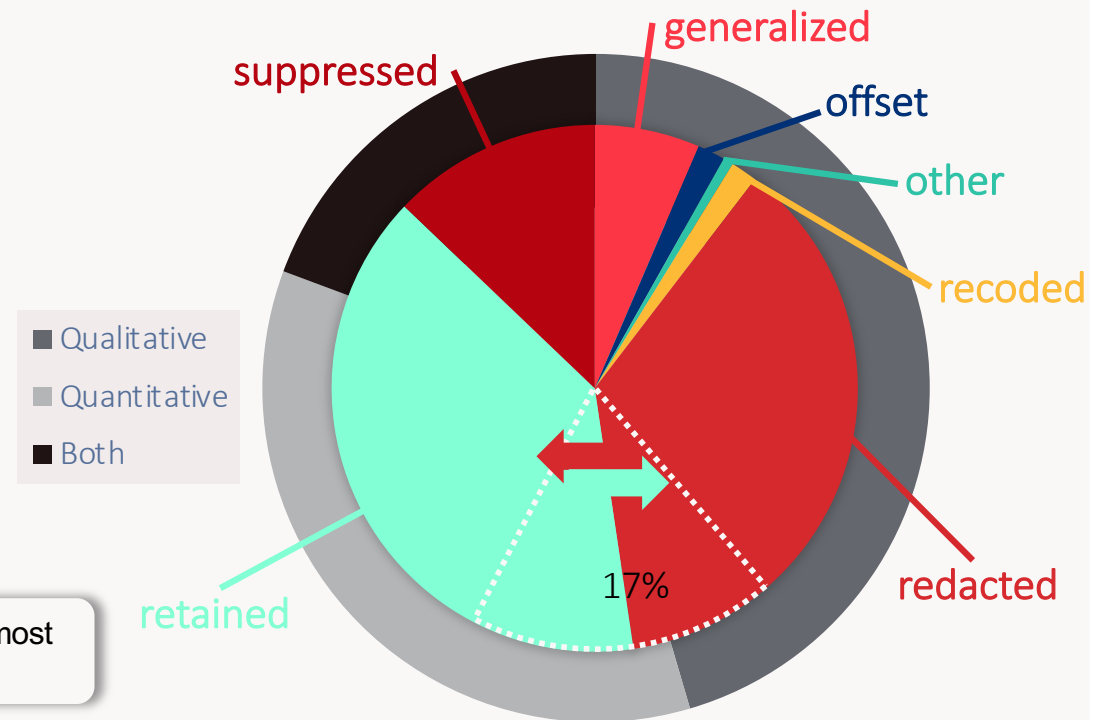
- EMA defines 26 categories of indirect identifiers
- Sponsors apply multiple anonymization techniques
- Most common techniques:
 - Suppression
 - Redaction
 - Generalization



Structure of EMA Indirect Identifier Data

- EMA defines 26 categories of indirect identifiers
- Sponsors apply multiple anonymization techniques
- Most common techniques:
 - Suppression
 - Redaction
 - Generalization

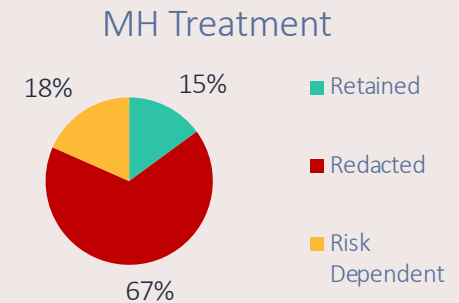
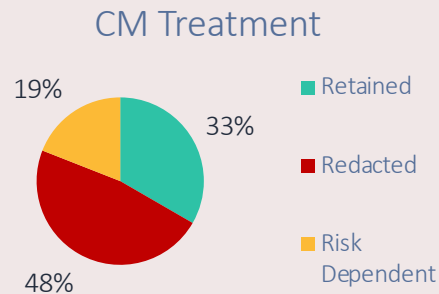
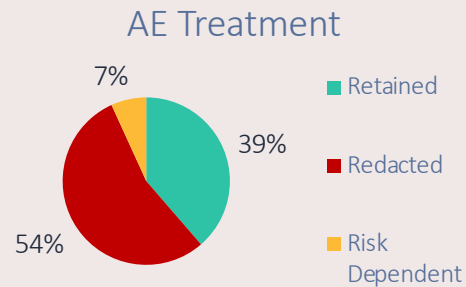
Regulators permit qualitative, prefer quantitative, but most real-world submissions rely on hybrid strategies.



Clinical Data with Highest Anonymization Activity

- These categories contain rich clinical narrative information
- They show the highest anonymization complexity
- Common approaches include:
 - generalization of clinical terms
 - suppression of rare conditions
 - selective redaction of identifying narratives

Narrative variables drive most anonymization complexity.



Variable Anonymization Across Other Identifiers

- Highly redacted variables concentrate on information with low data-utility
- Highly retained variables concentrate on low-risk information

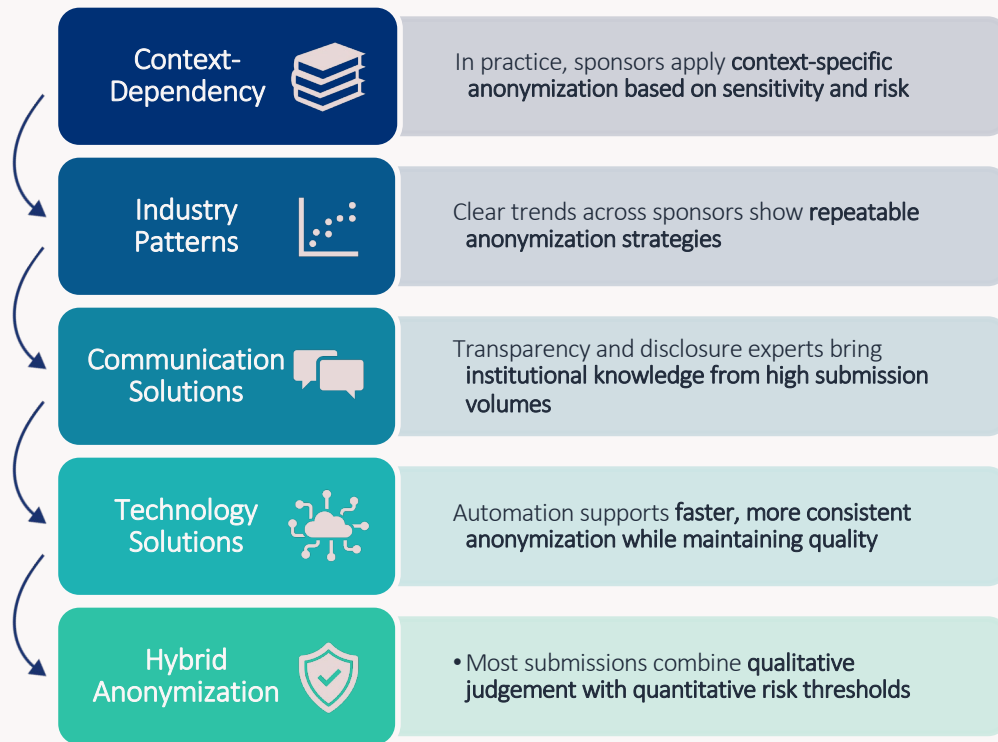


Variable Anonymization Across Other Identifiers

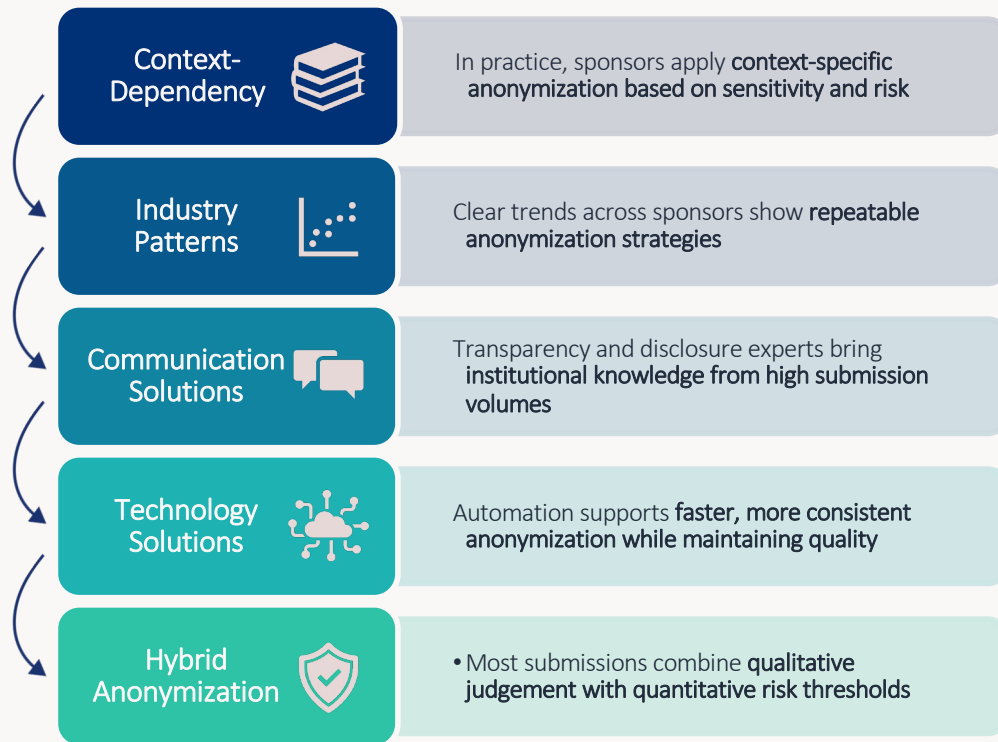
- Highly redacted variables concentrate on information with low data-utility
- Highly retained variables concentrate on low-risk information



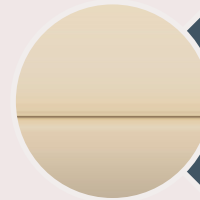
Key Takeaways for Future EMA Submissions



Key Takeaways for Future EMA Submissions



Measurable Impacts



EMA Policy 0070

- **Decreased PAF** to only administrative clarifications



EMA CTIS

- **Doubled capacity** between Jan '25 & Jan '26



Honz Slipka, MSc
Assoc. Principal Transparency Specialist