

Pragmatic Statistical De-Identification: Smooth Sailing, Choppy Waters, And Lessons Learned From Large-Scale Implementations

Lauren Groebe, Associate General Counsel – Privacy, Cybersecurity & Digital, Baxter International

John House, Executive Director, Science and Technology, Premier Applied Sciences

Brian Rasquinha, Associate Director, Solution Architecture, Privacy Analytics (Moderator)



Pragmatic Statistical De-Identification: Smooth Sailing, Choppy Waters, And Lessons Learned From Large-Scale Implementations

Statistical approaches to de-identifying data, whether HIPAA's Expert Determination method or other statistical assessments of identifiability, often form the core of an effort to manage identifiability. However, building a project, program, or business function around de-identification isn't always straightforward. In this panel discussion, we'll discuss experiences and lessons from standing up data de-identification initiatives from the contrasting perspectives of in-house legal, business leadership, and external privacy solution provider.

We will discuss:

- How to maximize the impact of statistical analysis of identifiability with strong business processes
- Effective strategies for communicating the utility, impacts, and limitations of a statistical analysis of identifiability
- Experiences balancing business needs and privacy constraints when approaching de-identification projects

Statistical de-identification: silver bullet, right?

Statistical de-identification: silver bullet, right?

- Need clarity on outcomes of a HIPAA Expert Determination (ED)
 - Documentation of de-identification “Recipe”
 - Guardrails around use cases where the ED applies: recipients, environments, purposes, contractual and process requirements, data transformations, and data linkages
- Organizations need to create alignment on the impact
 - Not a silver bullet
 - Data transformations are usually anticipated by arms-length stakeholders
 - Less obvious: impacts and restrictions on security controls, contract requirements, pass-through terms

What's the impact of the data use-case?

What's the impact of the data use-case?

- HIPAA ED assesses risk based on the *anticipated recipient's* ability to re-identify the data.
 - Modulated by the privacy and security controls in place in the destination environment
 - Also modulated by the motives and capacity of the recipient to attempt re-identification
 - If the recipient changes, these parameters can change too!
- Data linkage is also a key factor, with increasing prevalence in health data projects, particularly driven by AI
- Under ED, a de-identified dataset linked (at the person-level) with another de-id dataset **does not** necessarily result in a de-identified output
 - A given patient may have an increased set of identifiers associated with their record, which in turn can increase re-identification risk
- ED enables more customizability, and often higher data detail, but may have unanticipated workflow implications as compared to Safe Harbor

How important is internal/external communication?

How important is internal/external communication?

- Stakeholder communication can be high impact for both the HIPAA ED project itself as well as benefitting from the outcomes
- During the project:
 - Strong knowledge of data, use cases, controls, and recipients can reduce time spent chasing information on-demand
 - For changing situations, boxing in what modes of variation might be expected can increase the value of the ED if it can cover what-ifs
- After the project:
 - Stakeholders may not anticipate non-transformation recommendations
 - Helping arms-length groups get ahead of these impacts can reduce surprises and streamline activities downstream of the ED itself
 - Especially true when data sharing is external to the organization, where communication may be reduced and less formal

Thank you!



Lauren Groebe

Associate General Counsel –
Privacy, Cybersecurity &
Digital, Baxter International



John House

Executive Director, Science
and Technology, Premier
Applied Sciences



Brian Rasquinha

Associate Director, Solution
Architecture, Privacy Analytics
(Moderator)