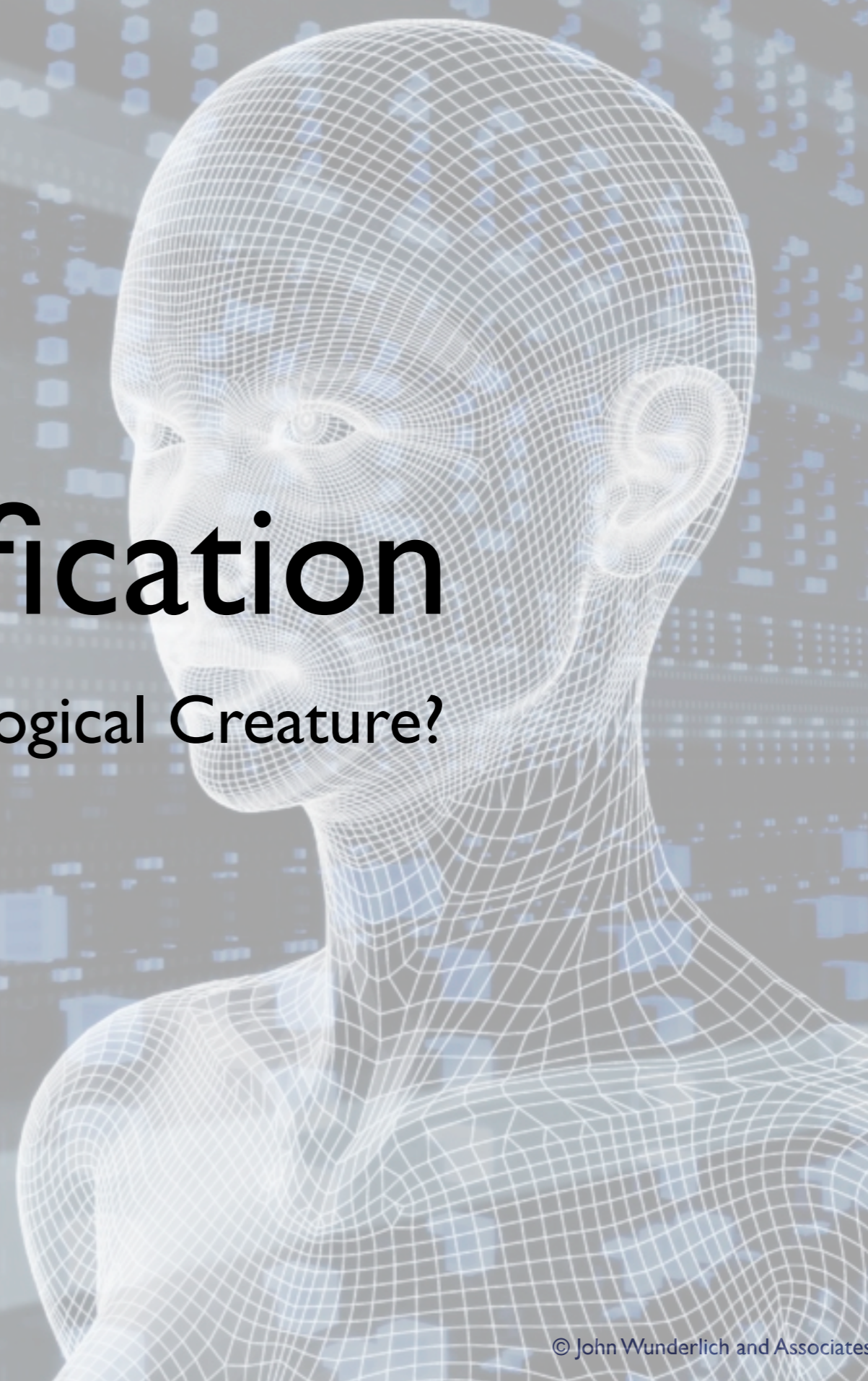




De-identification

Silver Bullet or Mythological Creature?





- What is de-identification?
- How does de-identification relate to access & privacy?
- Why isn't there a simple answer?



De-identification

- “identifying information” means information that identifies an individual or for which it is reasonably foreseeable in the circumstances that it could be utilized, either alone or with other information, to identify an individual. (PHIPA s. 4 (2))
- De-identified information, therefore, is the reverse



Access and Privacy

- Privacy
 - De-identifying data means that consent, use, and disclosure restriction no longer apply.
 - Demonstrates due diligences
- Access
 - Personal information (recorded information about an identifiable individual) can't be included in access requests
 - Minimize PIBs



Simple Answers

- Simple Field Redaction doesn't work
- Data linkage risk increases with time
- Making data both de-identified and useful is hard work



Re-identification Risks

Prosecutor Risk - In this scenario, an intruder wants to re-identify a specific person in a de-identified database.

Journalist Risk - In this scenario, the journalist does not care which individual is re-identified.

Marketer Risk - In this scenario, an intruder wants to re-identify as many individuals as possible in a database.

Definitions from Privacy Analytics white paper “De-Identification: Reduce Privacy Risks When Sharing Personally Identifiable Information”