A Path to Anonymization:
Through the Telecom, Health, AI, CCPA and Other Lenses

Tim Tobin
Partner,
Hogan Lovells

Ilya Vasilenko,
Global Risk & Compliance / DPO,
Starmind AG

Kelsey Finch
Policy Counsel,
Future of Privacy Forum

Khaled El Emam
Canada Research Chair, University of Ottawa, Founder & CEO, Privacy Analytics

Privacy + Security Forum
Oct 15, 2019

STARMIND

Legal Standards for Deidentification/Anonymization

- HIPAA
  No restrictions on use or disclosure of de-identified health information.
  - Removal of specified identifiers; or
  - Expert certification
- GDPR, Art. 26
  "data rendered anonymous in such a way that the data subject is not or no longer identifiable."
  Pseudonymized data is personal data
- U.S. Communications Act and Telecom Customer Records, 47 U.S.C. § 222 and regs
  - CPNI is individually identifiable information
  - "aggregate customer information" means collective data that relates to a group or category of services or customers, from which individual customer identities and characteristics have been removed.
- CCPA
  "Deidentified" means information that cannot reasonably identify, relate to, describe, be capable of being associated with, or be linked, directly or indirectly, to a particular consumer" if there are technical safeguards and business processes to prevent reidentification and protect against its inadvertent release, plus no attempts to reidentify.
- FTC Section 5 guidance
  reasonable steps to deidentify + public commitments to de-identify and no attempts to re-identify; contractually prohibit 3rd party cos. from re-identification

STARMIND

Assessment of the data processing

Purpose / Ethics

What is the purpose of processing?
Is it ethical?

- Purpose shall be documented.
- A group of different stakeholders shall decide whether the processing is ethical.

Method

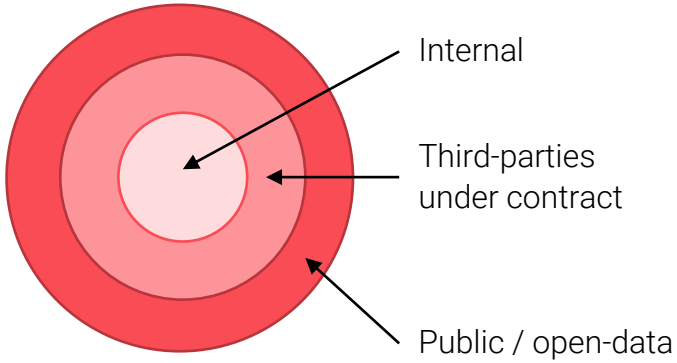
Is the method secure enough?

- Risks (incl. risks of re-identification) shall be assessed.
- Security measures shall be implemented.

Context

How and under which terms will the data be disclosed?

The broader the audience, the higher the risk of re-identification.



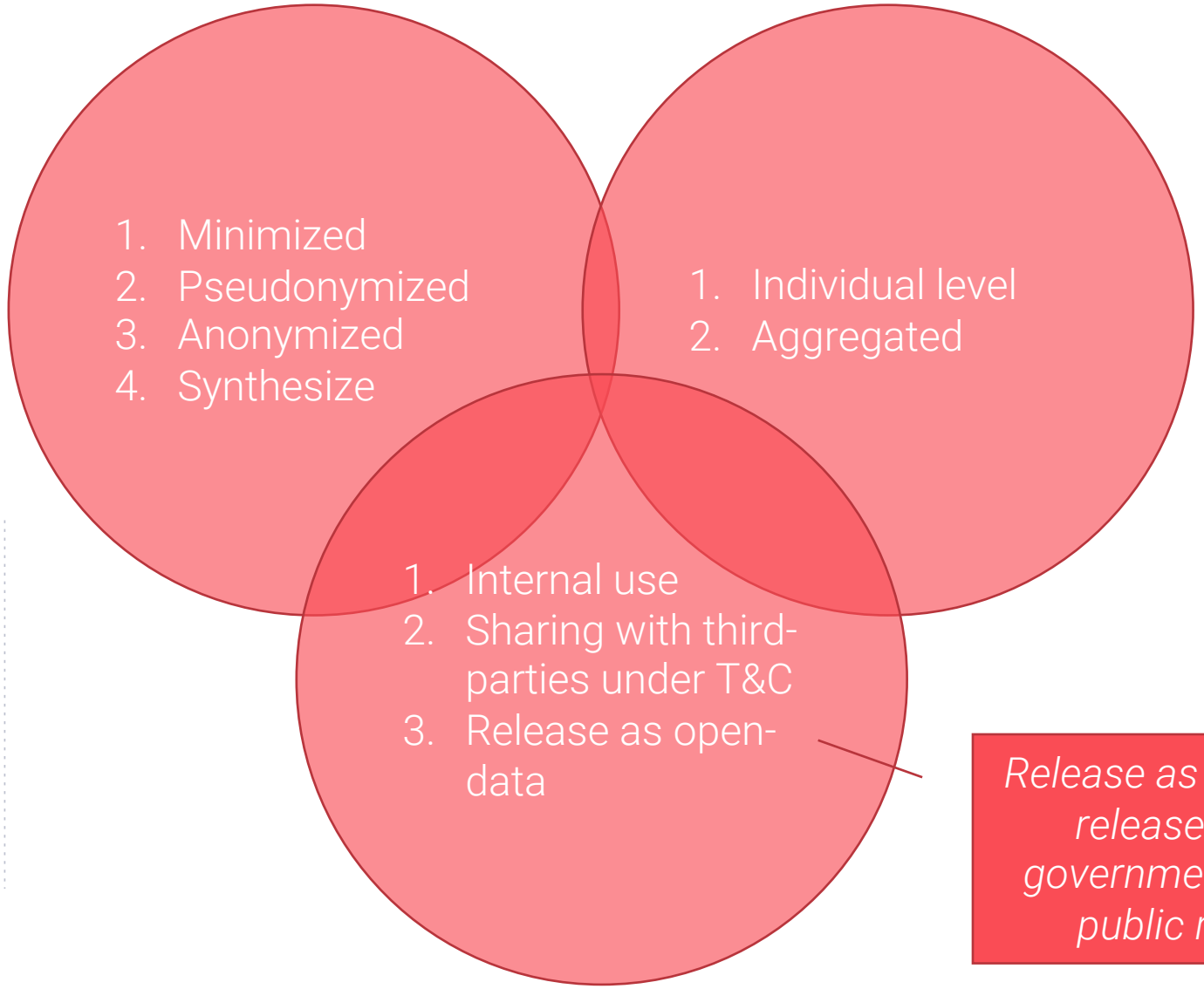
Other data sets – existing and future.
CPU capacity – now and future.

Example: The data shows that, in a certain region, people tend to have a certain disease more often.

- You can use this info to advise people to do more regular checks.
- You can also give this information to an insurance company so that they can raise the insurance police costs for these people.

STARMIND

Three degrees of data processing



Release as open data includes release the data to the government (because of the public right of access)

Open data challenge:

No risk: release a single number – e.g. population of the US is 327mil.

Potentially very high risk: release highly granular statistics – e.g. number of people per zip code and gender and age bucket – may be easy to identify

STARMIND

Anonymization – a new Swiss army knife!

Personal data

Anonymization

Anonymous data – freedom from data protection!!! YESSS!

GDPR introduced a great incentive to anonymize the data (or claim that it is) – because anonymous data is out of scope of GDPR.

In theory it is a great idea, but neither DPAs nor companies have required capacities and skills within their organization to assess that the anonymization method, the purpose and the context are set correctly.

Go for a use-case driven approach!

What does normally go wrong?

”Oh, cool! We can collect so much data! Let’s anonymize it and see how we can use it!”

What would work much better:

- What question would you like to answer?
- Which insights would you like to gain?

- Which data do you require for that?

- Where and how can you collect the data?

- How do you ensure compliant processing (and anonymization)?

Goal may be also ”monetization” and “future exploratory analysis”! – but in such a rather open case you would still have a discussion which type of analysis would be allowed and which not.

STARMIND

50 shades of processing

First questions to ask:

- Do you really need all this data?
- Do you really need the real data?

Minimization / Pseudonymization	Anonymization (EU) De-identification (US)	Synthetic data
• Easy to build in-house. • Less risky than original data, but requires a similar level of protection.	"data is either not fully anonymous or useless" – the level of anonymity is a function of the risks and the corresponding safeguards.	Same metrics as original data, but the data is completely generated. Many fields of applicable like testing, trend analysis & etc.

UK Anonymization framework (see our reading list)
ISO/IEC 20889:2018(en) - Privacy enhancing data de-identification terminology and classification of techniques

We also recommend IAPP Vendor report for the list of companies offering services in this field:
https://iapp.org/resources/article/2019-privacy-tech-vendor-report/

STARMIND

Blitzscaling

Blitzscaling – an aggressive, all-out program of growth that prioritizes speed over efficiency, even in an environment of uncertainty.

Works both for the entire company as well as for individual products! The world and the business speeds up. Blitzscaling becomes THE tool for companies to succeed.

The motto “Move fast and break things” is like jumping off the cliff and assembling the plane on the way. This worked for Facebook 5 years ago but will not work now – privacy landscape has changed completely.

It became as important as never before to be proactive as privacy professional to explain the CHANGED rules of the gravity (=privacy) BEFORE the company jumps of the cliff (with a new product).

STARMIND