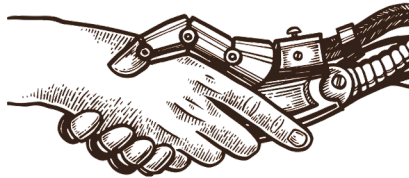


A methodology for reconciling computer science and legal approaches to privacy



Kobbi Nissim
Georgetown University
(on sabbatical @ Bocconi U, Milan)



W/ Micah Altman and Aloni Cohen

[Applied Algorithms for Machine Learning, Paris 2024]

Applied Algorithms for Machine Learning?



ML algorithms cannot be applied if they do not meet legal requirements

Do they?

Do ML systems meet legal privacy standard?

- We need to know!
 - A huge number of decisions with legal implication happen in computer systems
 - Even if only a small fraction of these decisions required human review, they would quickly overwhelm our judiciary or administrative systems
 - Need negligible error rate!

Hey, no worries! We have ...

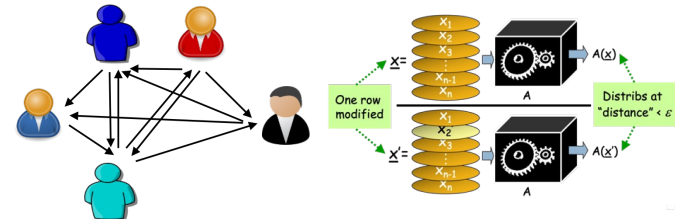
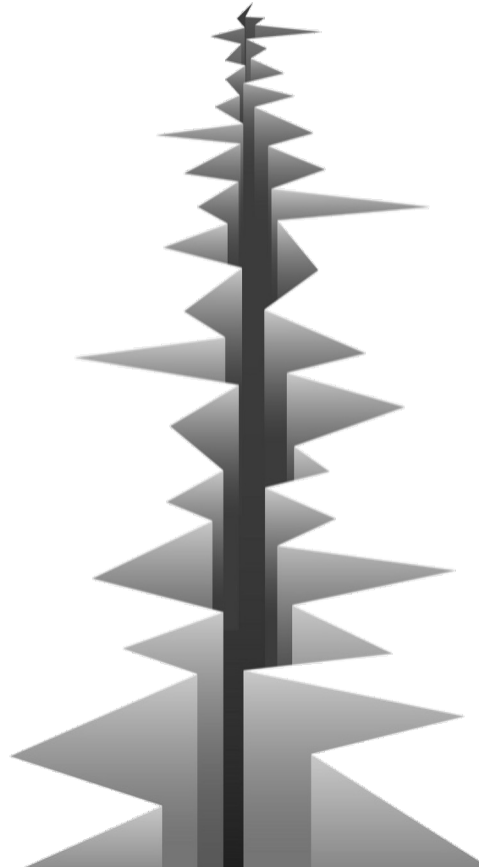
New laws/regulations

- General Data Protection Regulation
- California Consumer Privacy Act
- California Privacy Rights Act
- ...



Strong PETs

- Encryption
- Secure multiparty computing
- Differential privacy
- ...



Do machine learning systems meet
the requirements of legal privacy
standards?

Do we even understand the question?

Hard to reason on whether machine learning systems meet legal privacy standards

- Same words, different meaning:

- Legal and technical definitions of privacy protection have evolved in diverging ways [N, Wood 2018]

- How to map between technical concepts and normative expectations of privacy

- Different ways of arguing

- Differing/contrasting values:

- Mathematical rigor vs. flexibility

- Generality of protection afforded

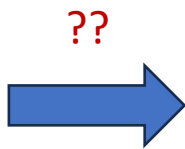
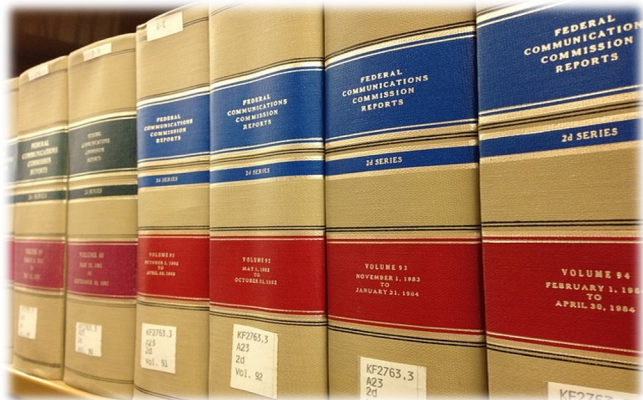
- Reactive vs. proactive

This talk

Elements of an approach to bridging between legal and computer science privacy formulations*

Benefits of a principled approach to regulation

Legal privacy standards under a computational lens

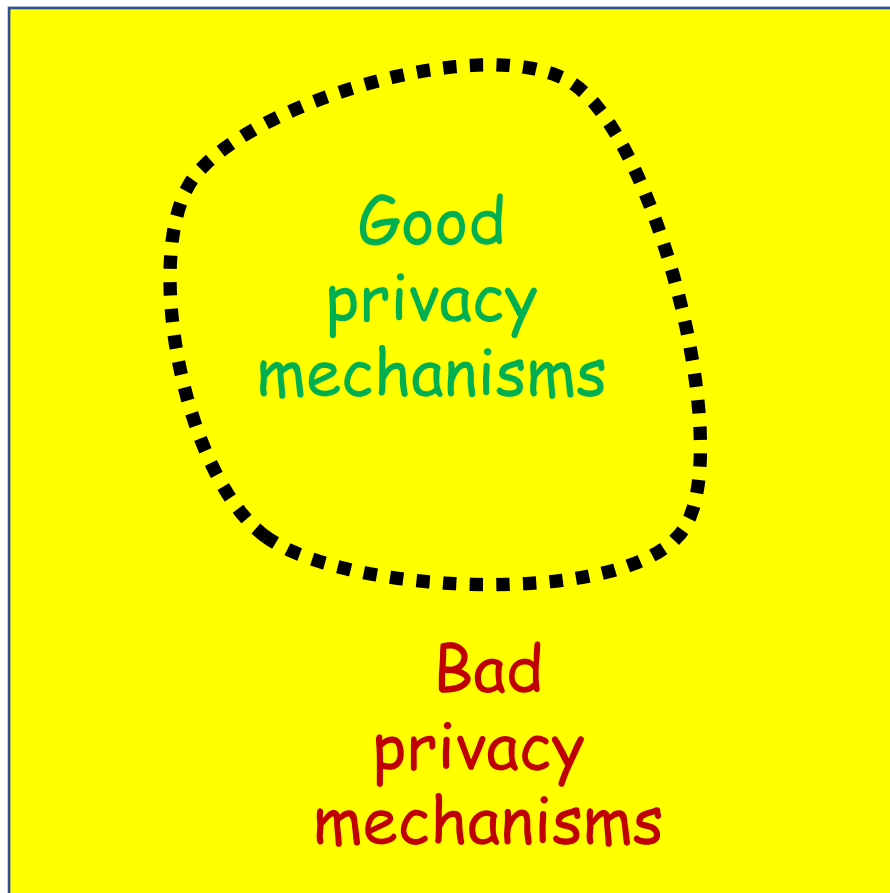


$$\forall P \in R \exists \Delta I. V^A \cap C \wedge Y$$
$$\Pr \left[\frac{d\theta}{dx} R(x) \right] \leq \vec{\nabla} \times e^\epsilon \cdot \Pr_A [M(x'') | z^3]$$

Spoiler: this will not happen

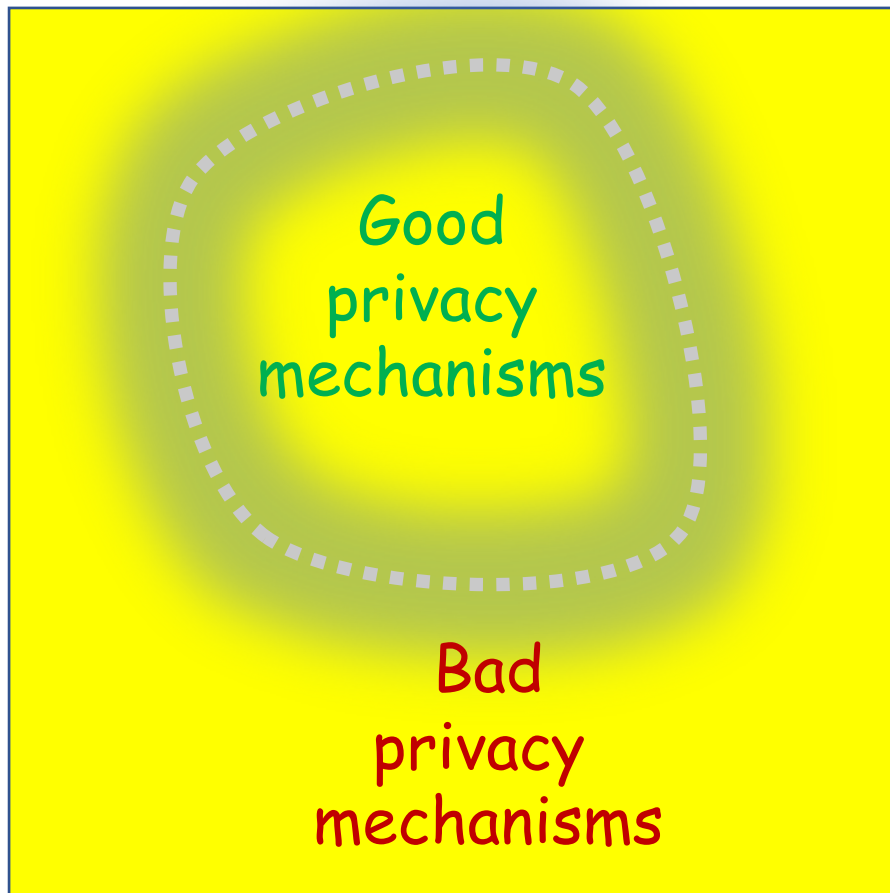
Legal privacy standards under a computational lens

- Privacy regulations have many components: admin requirements, enforcement mechanisms, exceptions for various purposes, remediation, ...
- We focus on restrictions on the set of mechanisms that are considered to "preserve privacy"
- Definitions of legal standards are not formal from a technical point of view



Legal privacy standards under a computational lens

- Privacy regulations have many components: admin requirements, enforcement mechanisms, exceptions for various purposes, remediation, ...
- We focus on restrictions on the set of mechanisms that are considered to "preserve privacy"
- Definitions of legal standards are not formal from a technical point of view



Legal privacy standards under a computational lens

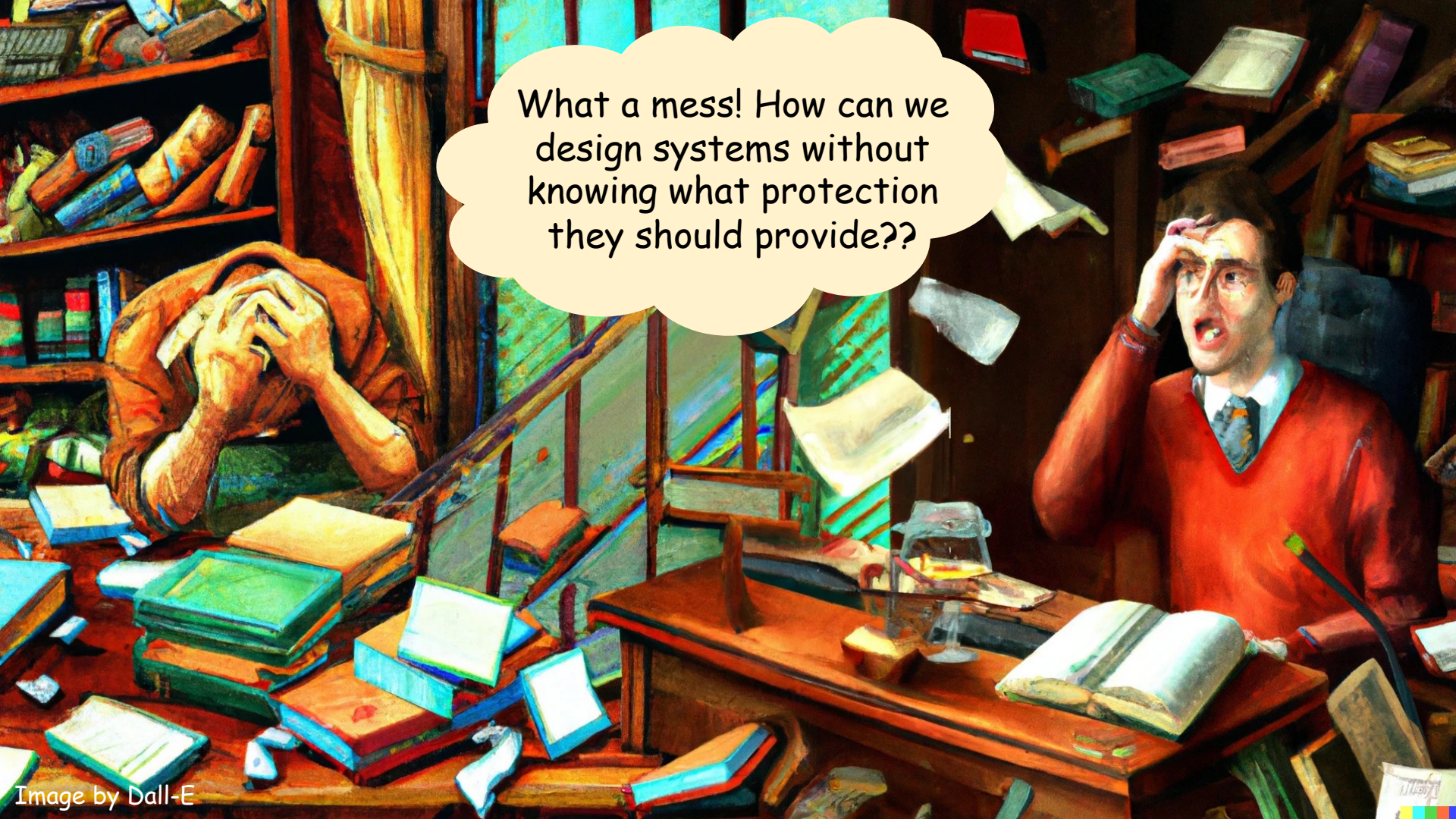
- Privacy regulations have many components: admin requirements, enforcement mechanisms, exceptions for various purposes, remediation, ...
- We focus on restrictions on the set of mechanisms that are considered to "preserve privacy"
- Definitions of legal standards are not formal from a technical point of view



Legal privacy standards under a computational lens

- Privacy regulations have many components: admin requirements, enforcement mechanisms, exceptions for various purposes, remediation, ...
- We focus on restrictions on the set of mechanisms that are considered to "preserve privacy"
- Definitions of legal standards are not formal from a technical point of view

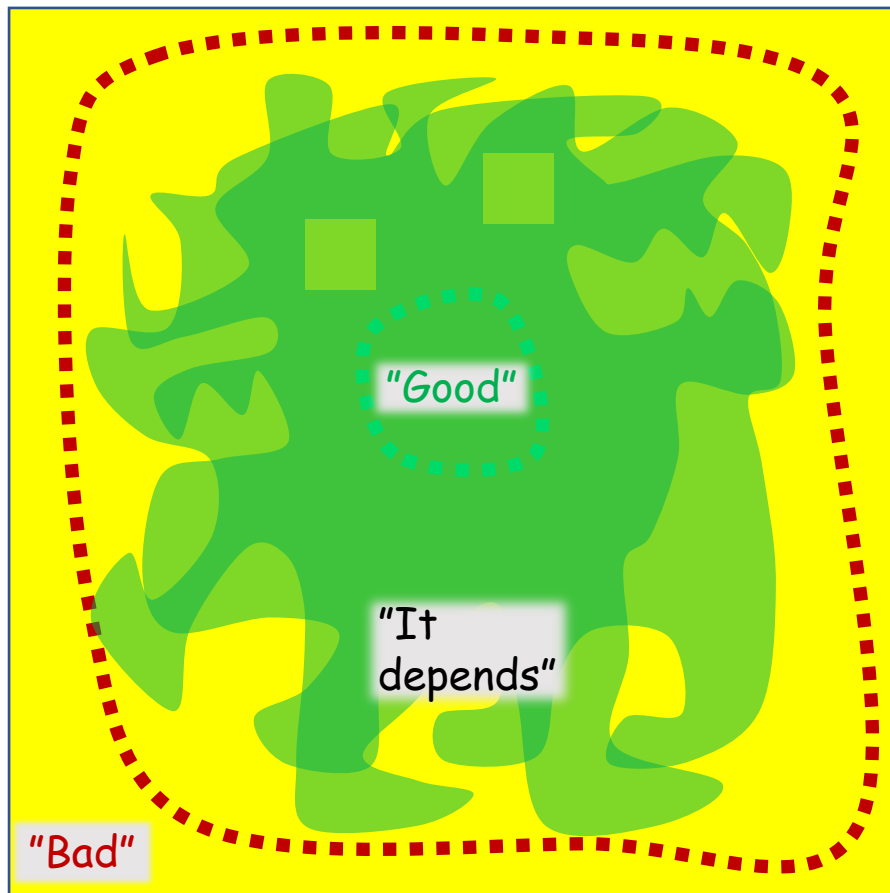




What a mess! How can we design systems without knowing what protection they should provide??

Legal privacy standards under a computational lens

- What if we consider all possible interpretations?
- Well defined boundaries are helpful!

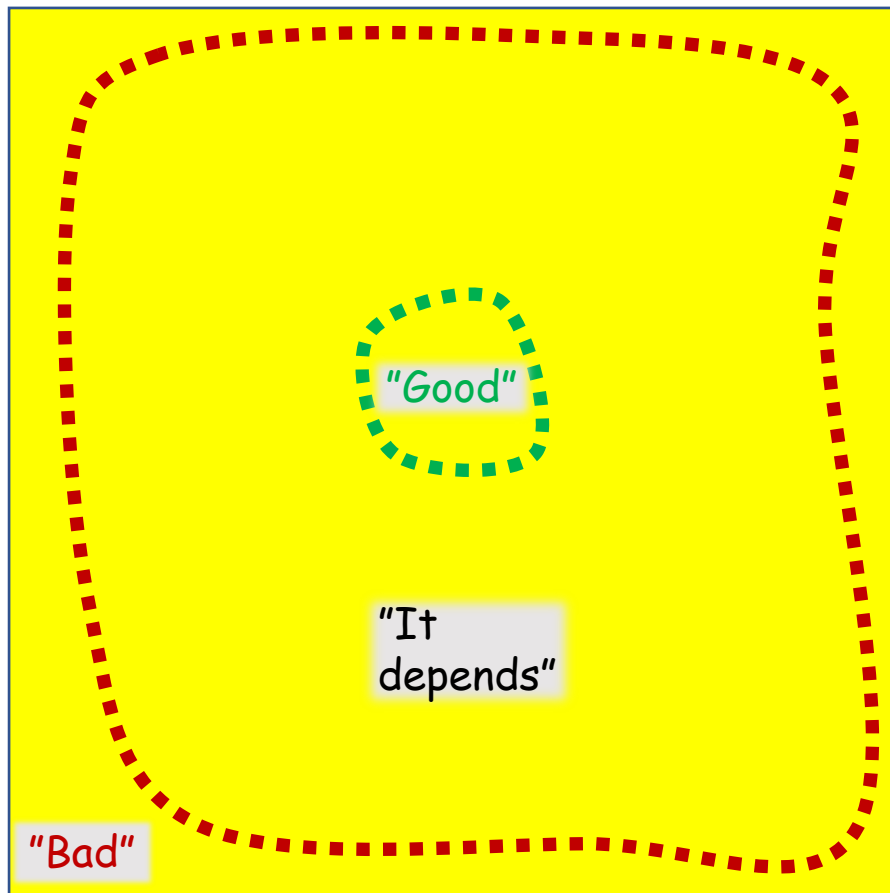


Legal privacy standards under a computational lens

- What if we consider all possible interpretations?

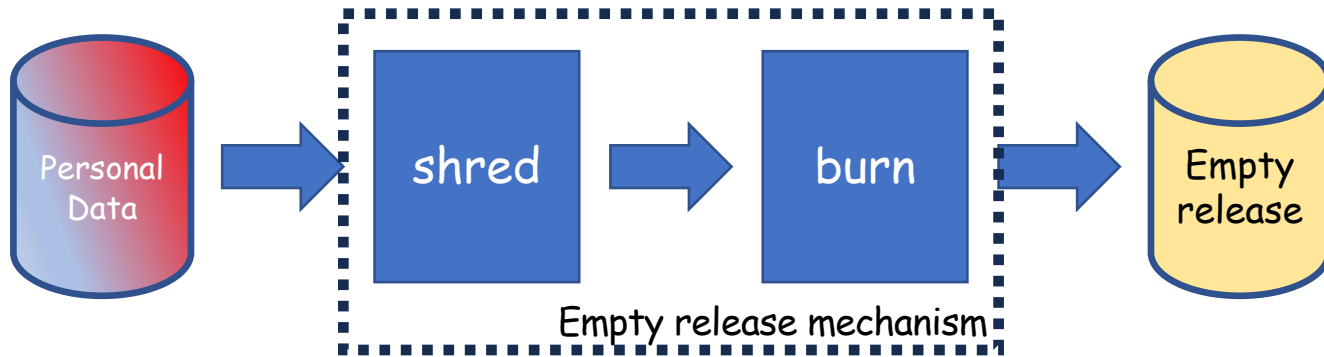


- Well defined boundaries are helpful!
- Boundaries ~~may~~ will become tighter with improved analysis and communication with regulators



Putting simple mechanisms on the map

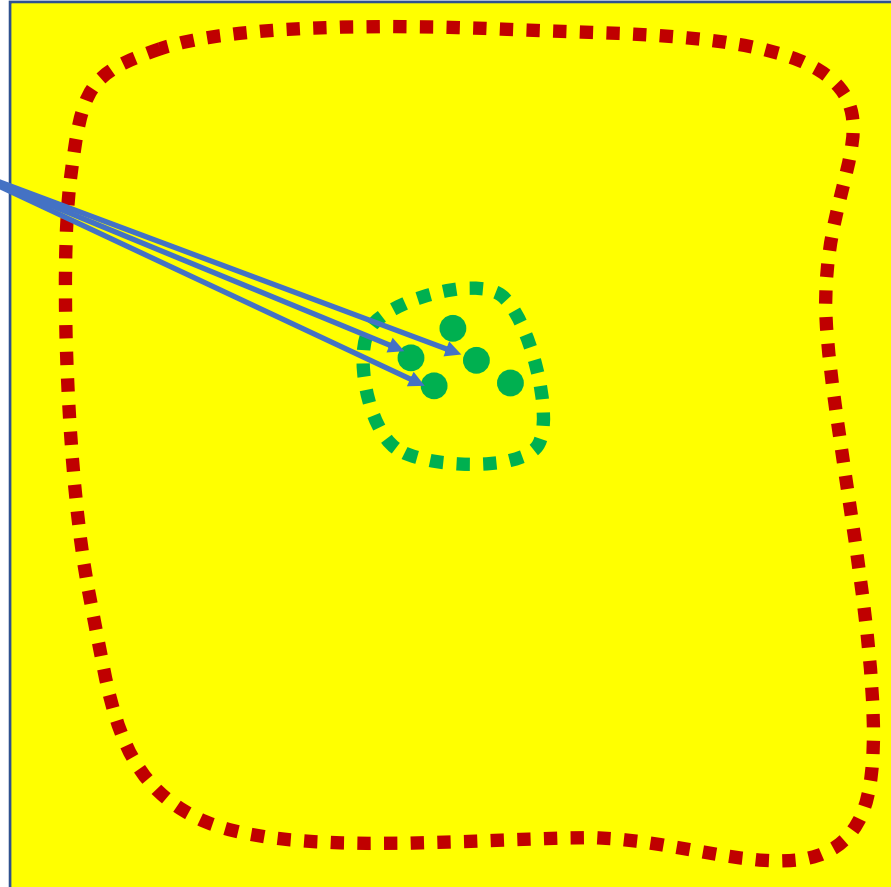
Empty release mechanisms



- May not be a good use of taxpayer money ...
 - ... but **always protects privacy!**
- **More mechanisms of this family:** all mechanisms that ignore their input data
 - E.g. the mechanism that outputs "Abracadabra" on all inputs

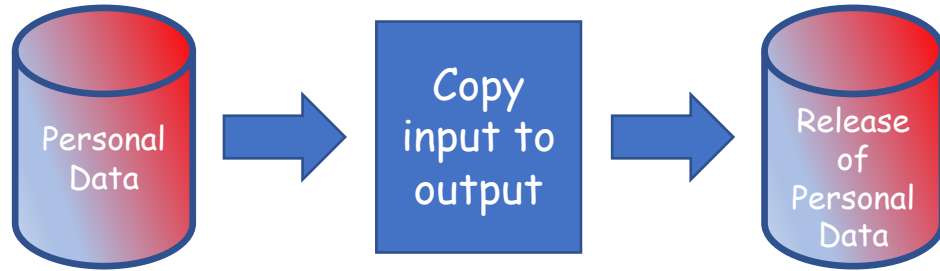
Principles for Privacy Regulation [Altman, Cohen, N]

- At Least The Empty Release (ALTER): Any privacy regulation should deem empty release mechanisms as providing privacy



mechanisms

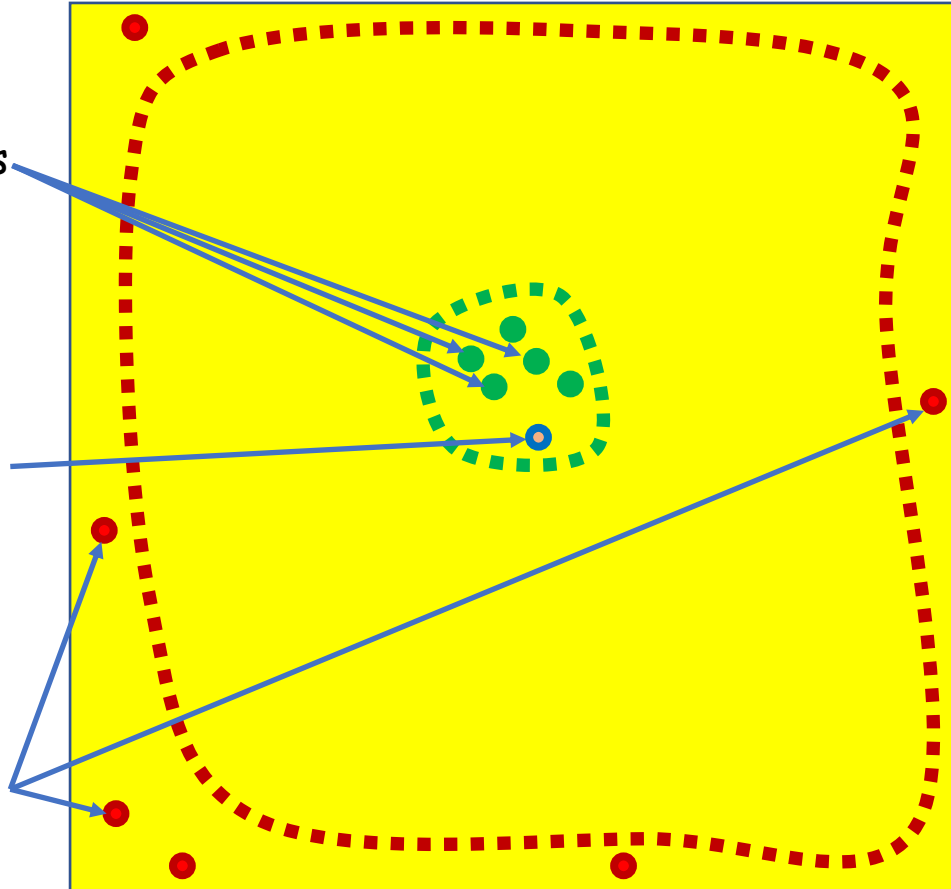
Identity mechanism(s)



- Never protects privacy!
- **More mechanisms in this family:** any mechanism whose output can be post-processed to result in identity
 - Aka reconstruction attacks [Dinur N 03]

Principles for Privacy Regulation [Altman, Cohen, N]

- **At Least The Empty Release (ALTER):** Any privacy regulation should deem empty release mechanisms as providing privacy
- **Not Just the Empty Release (NJER):** Any privacy regulation should deem at least some non-empty release mechanisms as providing privacy
- **Identity:** Any privacy regulation should deem identity mechanisms as not providing privacy



mechanisms

Principles? This is ~~trivial~~ idiotic!



Answer, part I*

- General Principles [Altman+ 2022] [Altman Cohen N]
 - Process protection
 - ALTER, NJER, identity, indistinguishability
 - Inclusion-based protection
 - Individual relatedness, individual unrelatedness
 - Format neutrality
 - Post processing, indistinguishability
 - Composition awareness
- Protective assumptions
- Transparency
 - Transparency complement

Answer, part II

These ~~idiotic~~ simple principles are useful for analyzing a major regulation

Data anonymization

- Many privacy and data protection laws around the globe conceive of some **anonymization** or **de-identification** process



- **Health Insurance Portability and Accountability Act (HIPAA 1996) Privacy Rule:** governs the use of 'protected health information' but not '[h]ealth information that does not identify an individual and with respect to which there is no reasonable basis to believe that the information can be used to identify an individual'

Data anonymization

- Many privacy and data protection laws around the globe conceive of some **anonymization** or **de-identification** process

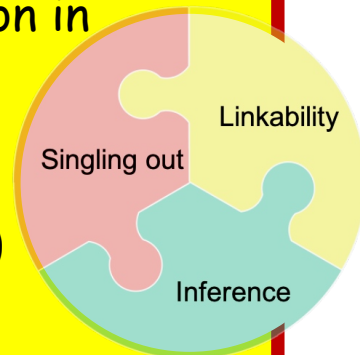


- **General Data Protection Regulation (GDPR 2016)**: governs the processing of 'personal data' but not 'personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable'

What do regulations mean by 'anonymization'?

- Often left undefined 😞
- Sometimes understood as the scraping of certain identifying attributes
 - E.g., HIPAA safe-harbor method specifies a list of identifiers to be removed
 - Such techniques have been repeatedly demonstrated insufficient for guaranteeing reasonable privacy 😞 [Sweeney 2000, Narayanan Shmatikov 2006,...]

- Most well-developed treatment of the concept of anonymization in regulatory guidance available today is from opinions of EU's Article 29 Data Protection Working Party
- A 2014 WP opinion breaks down anonymization into protection from three types of attacks on unregulated (publicly released) data: **linkability**, **singling out**, and **inference**



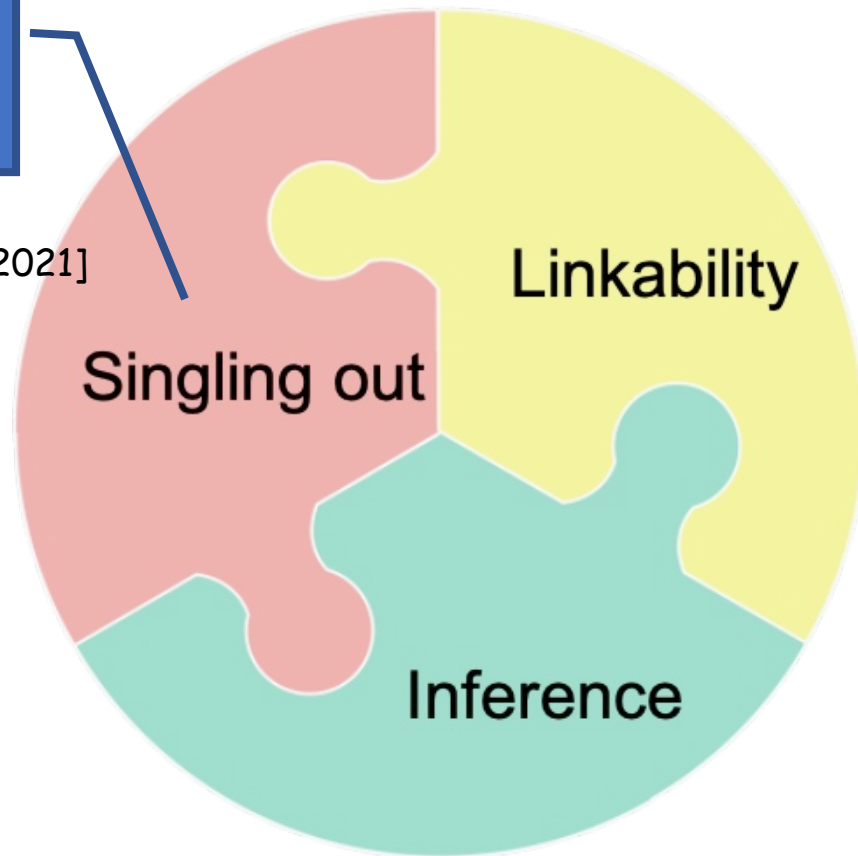
Article 29 WP assessment of privacy technologies

	Is Singling out still a risk?	Is Linkability still a risk?	Is Inference still a risk?
Pseudonymisation	Yes	Yes	Yes
Noise addition	Yes	May not	May not
Substitution	Yes	Yes	May not
Aggregation or K-anonymity	No	Yes	Yes
L-diversity	No	Yes	May not
Differential privacy	May not	May not	May not
Hashing/Tokenization	Yes	Yes	May not

Table 6. Strengths and Weaknesses of the Techniques Considered

What is singling out?

Based on [Cohen, N 2020]
[Altman, Cohen, N, Wood 2021]



Art. 29 WP general notions of attacks on released data

Context: the GDPR notion of *Singling out*

GDPR, Article 1:

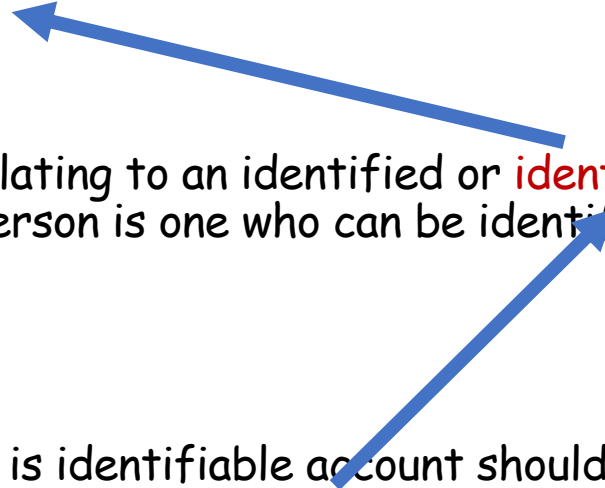
"This Regulation lays down rules relating to the protection of natural persons with regard to the processing of **personal data** . . ."

GDPR, Article 4:

"Personal data means any information relating to an identified or **identifiable** natural person; an identifiable natural person is one who can be identified, directly or indirectly . . ."

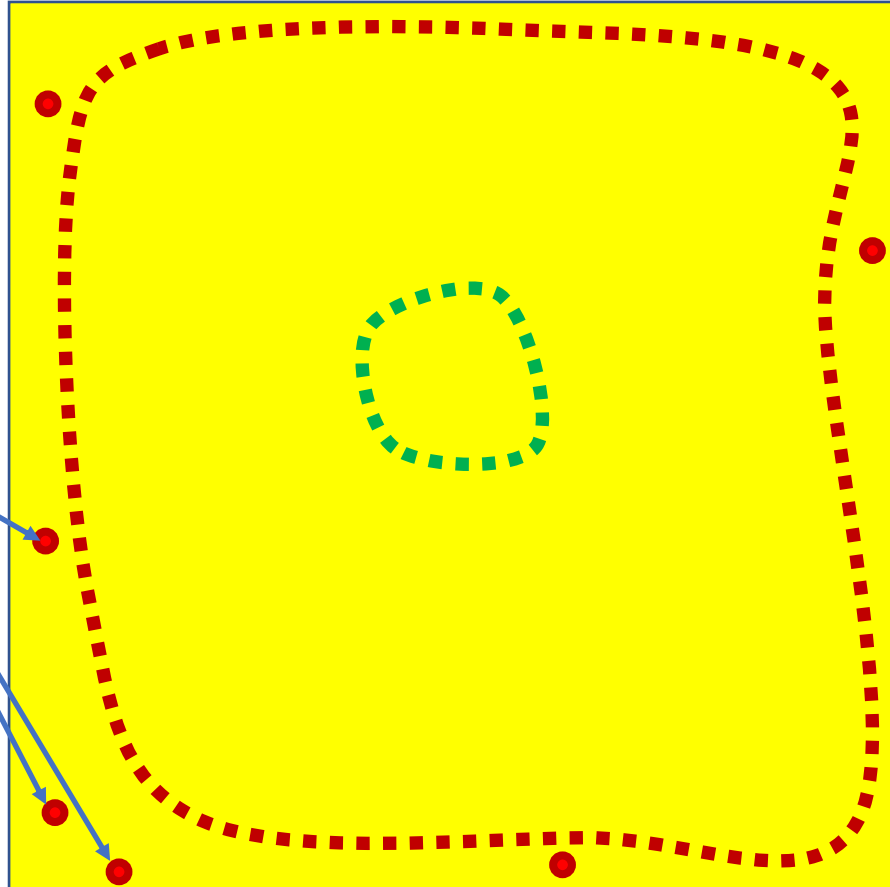
GDPR, Recital 26:

"To determine whether a natural person is identifiable account should be taken of all the means reasonably likely to be used, such as **singling out** . . . to identify the natural person directly or indirectly."



Putting singling out on or map

- Mechanisms demonstrated to allow singling out



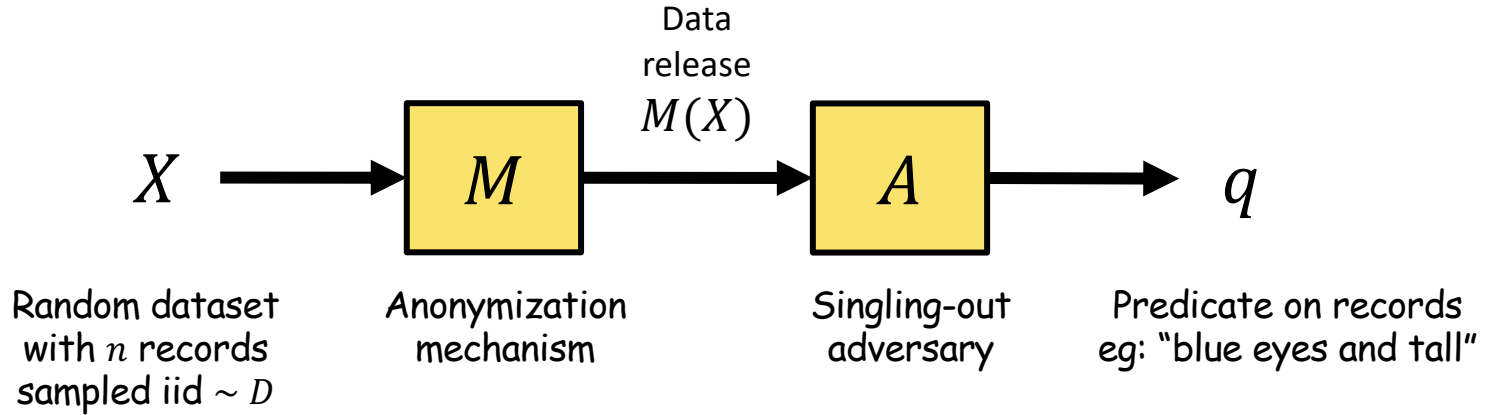
But - what is singling out?

- The A29WP 2014 guidance understands singling out as the ability to 'isolate':
 - **`To isolate`**: to identify a combination of attributes that distinguishes an individual from all other individuals in the data underlying a given data release

ID	Movie	Date	Rating	Movie	Date	Rating	Movie	Date	Rating
1	Fargo	Jan 1	5	Mulan	Feb 2	5	Crash	Mar 3	5
2	Fargo	Jan 11	5	Mulan	Feb 29	5	Crash	Mar 13	5
3	The Sting	Jan 1	5	Mulan	Feb 2	5	Mad Max	Mar 3	5

- **Isolation examples**: there is **exactly one row** in the **underlying data** of a person that
 1. ... watched "The Sting"
 2. ... watched "Mulan" between Feb 19 and Mar 10
 3. ... doesn't satisfy 1 or 2:
 - i.e. did not watch "The Sting" nor watched "Mulan" between Feb 19 and Mar 10

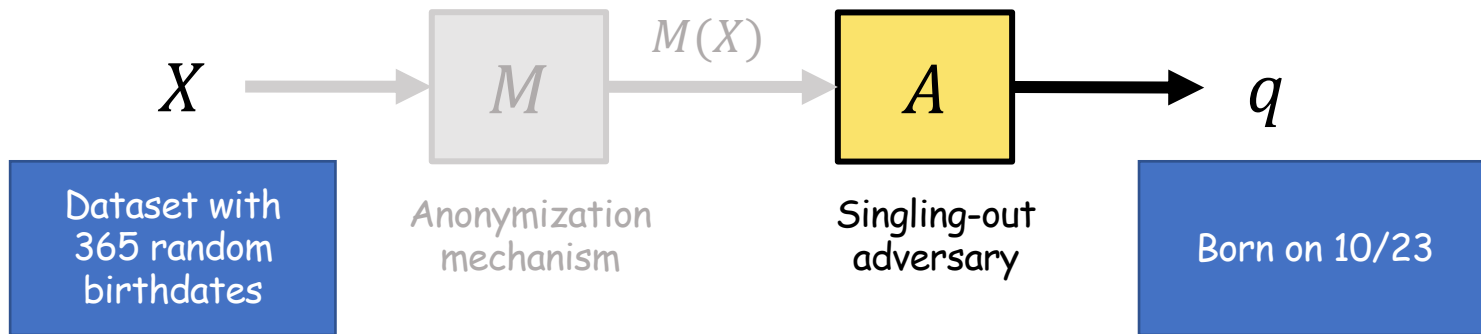
Singling out = Isolation ?



Adversary's goal: Given $M(X)$ output predicate q matching exactly 1 row in X

Definition attempt: M is secure against singling out if no adversary can isolate a row except with negligible probability (over coins of X, M, A)

Isolation with empty release

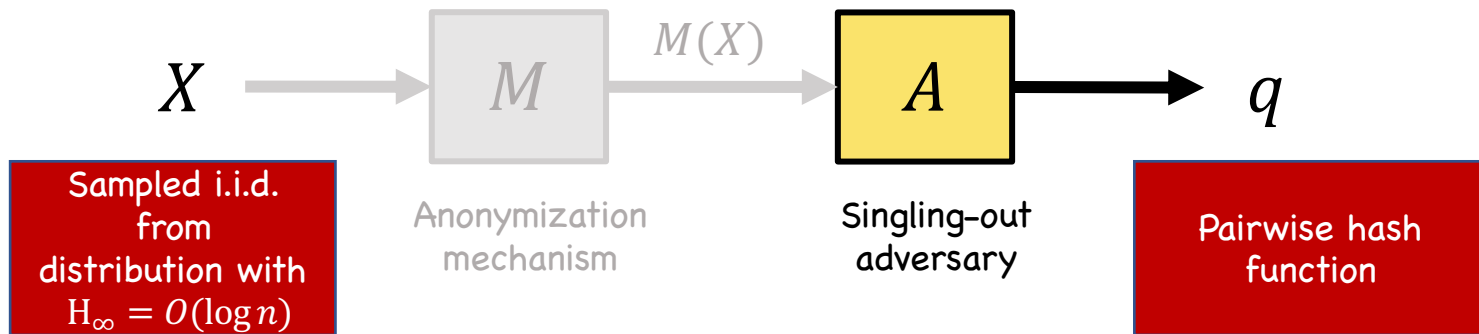


- q matches a $1/365$ fraction of the universe

$$\Pr[q^* \text{ isolates a row}] = \sum_{i=1}^n \Pr[q^* \text{ isolates row } i] = 365 \times \left(\frac{1}{365}\right) \left(1 - \frac{1}{365}\right)^{365-1} \approx 0.37$$

- Can isolate with empty release and succeed with prob. $\approx 37\%$

Isolation with empty release



- q matches a $1/n$ fraction of the universe (via application of the leftover hash lemma)

$$\Pr[q^* \text{ isolates a row}] = n \times \left(\frac{1}{n}\right) \left(1 - \frac{1}{n}\right)^{n-1} \approx \frac{1}{e} \approx 0.37$$

- Can isolate given an empty release and succeed with prob. $\approx 37\%$



Violates ALTER!

By the A29WG interpretation of singling out even empty release mechanisms do not anonymize data

Fixing the isolation criteria

- **Predicate singling out** happens when an attacker manages to isolate much better than with the empty release [Cohen N 20]
 - In particular, when the attacker isolates with a predicate of negligible probability
- E.g., "Vegetarian, Columbian, female theoretical computer scientist, plays the violin, races cars, and fluent in Hebrew, Italian, and Japanese"
 - A priori, it is likely that no person with this description exists
 - → with empty release any attacker has only a negligible chance of isolation
 - If given a release an attacker succeeds with even 1% success probability with such predicates - significant isolation
- Such a definition immediately satisfies **ALTER (at least the empty release)**
- Can be shown to satisfy **NJER (not just the empty release)**

Article 29 WP assessment of k -anonymity

	Is Singling out still a risk?	Is Linkability still a risk?	Is Inference still a risk?
Pseudonymisation	Yes	Yes	Yes
Noise addition	Yes	May not	May not
Substitution	Yes	Yes	May not
Aggregation or K -anonymity	No	Yes	Yes
L-diversity	No	Yes	May not
Differential privacy	May not	May not	May not
Hashing/Tokenization	Yes	Yes	May not

Table 6. Strengths and Weaknesses of the Techniques Considered

k-anonymity [Samarati Sweeney '98, Sweeney '02]

A *k*-anonymizer is a process that suppresses information in a dataset to make every combination of potentially identifying attributes appear at least *k* times

potentially identifying

ZIP	Age	sex	Disease
23456	55	Female	Heart
12345	30	Male	Heart
12346	33	Male	Heart
13144	45	Female	Cancer
13155	42	Male	Hepatitis
23456	42	Male	Viral

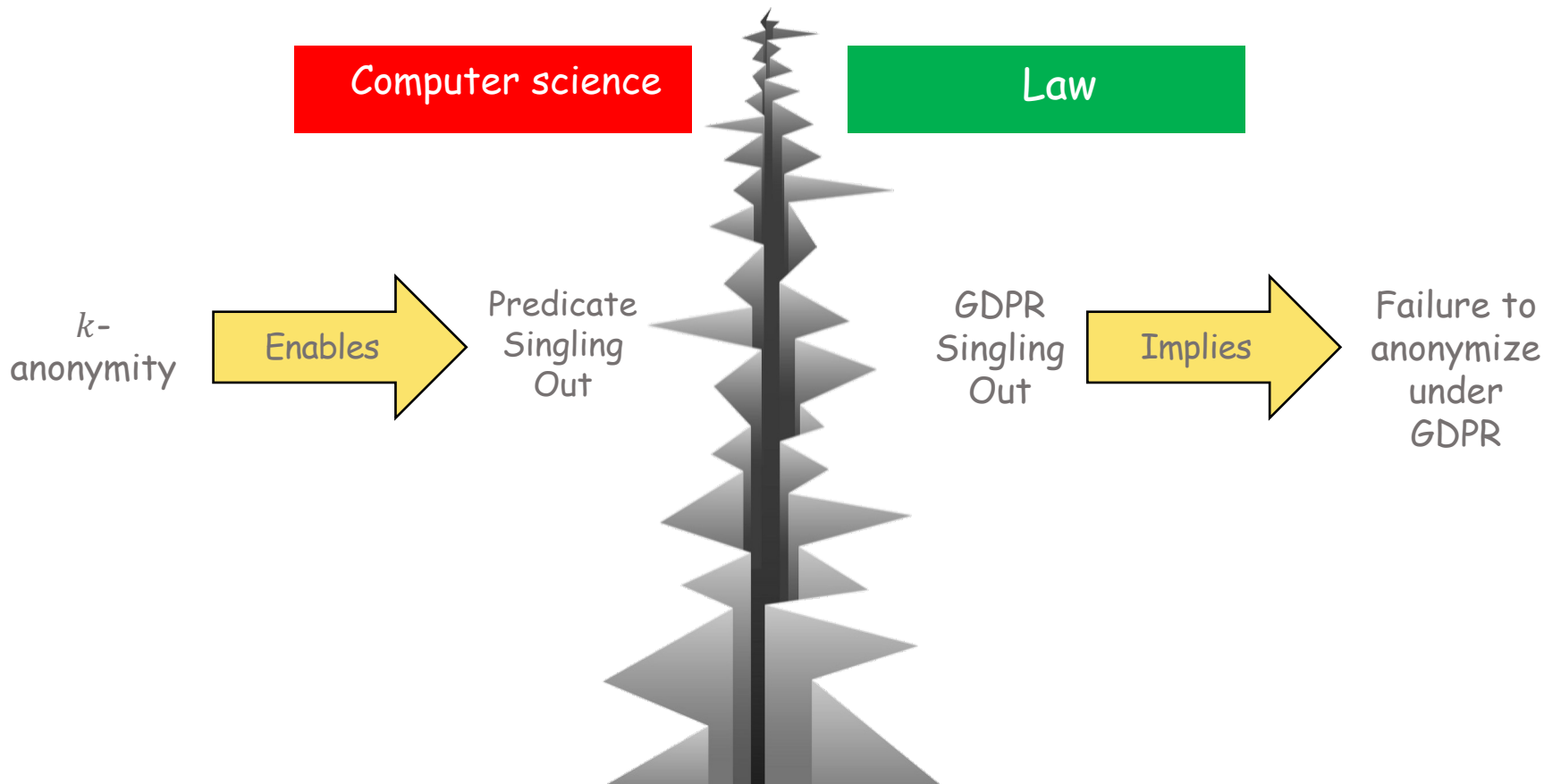
2-anonymous

ZIP	Age	sex	Disease
23456	**	*	Heart
1234*	3*	Male	Heart
1234*	3*	Male	Heart
131**	4*	*	Cancer
131**	4*	*	Hepatitis
23456	**	*	Viral

Does k-anonymity provide security against predicate singling out?

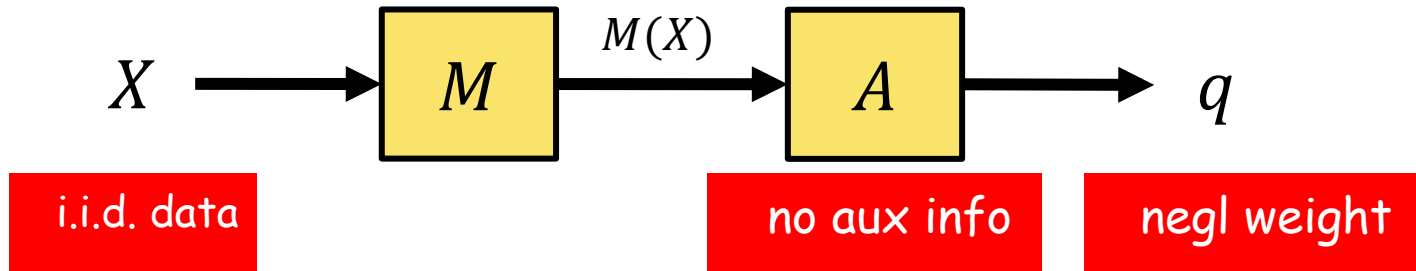
- **Theorem (informal) [Cohen, N 20]:** k-anonymity typically enables predicate singling out
- **Proof:** demonstrates that typically the k-anonymizer would do the hard work for the attacker, needs to be complemented with a trivial attacker (using leftover hash lemma)
- [Cohen 22] introduced downcoding attacks and proved that a large class of k-anonymizers is vulnerable to downcoding attacks.
- [Cohen 22] used LinkedIn.com to reidentify 3 students in a k-anonymized dataset published by EdX
- But, does k-anonymity satisfy the **GDPR anonymization standard?**

Does k -anonymity satisfy the GDPR anonymization standard?



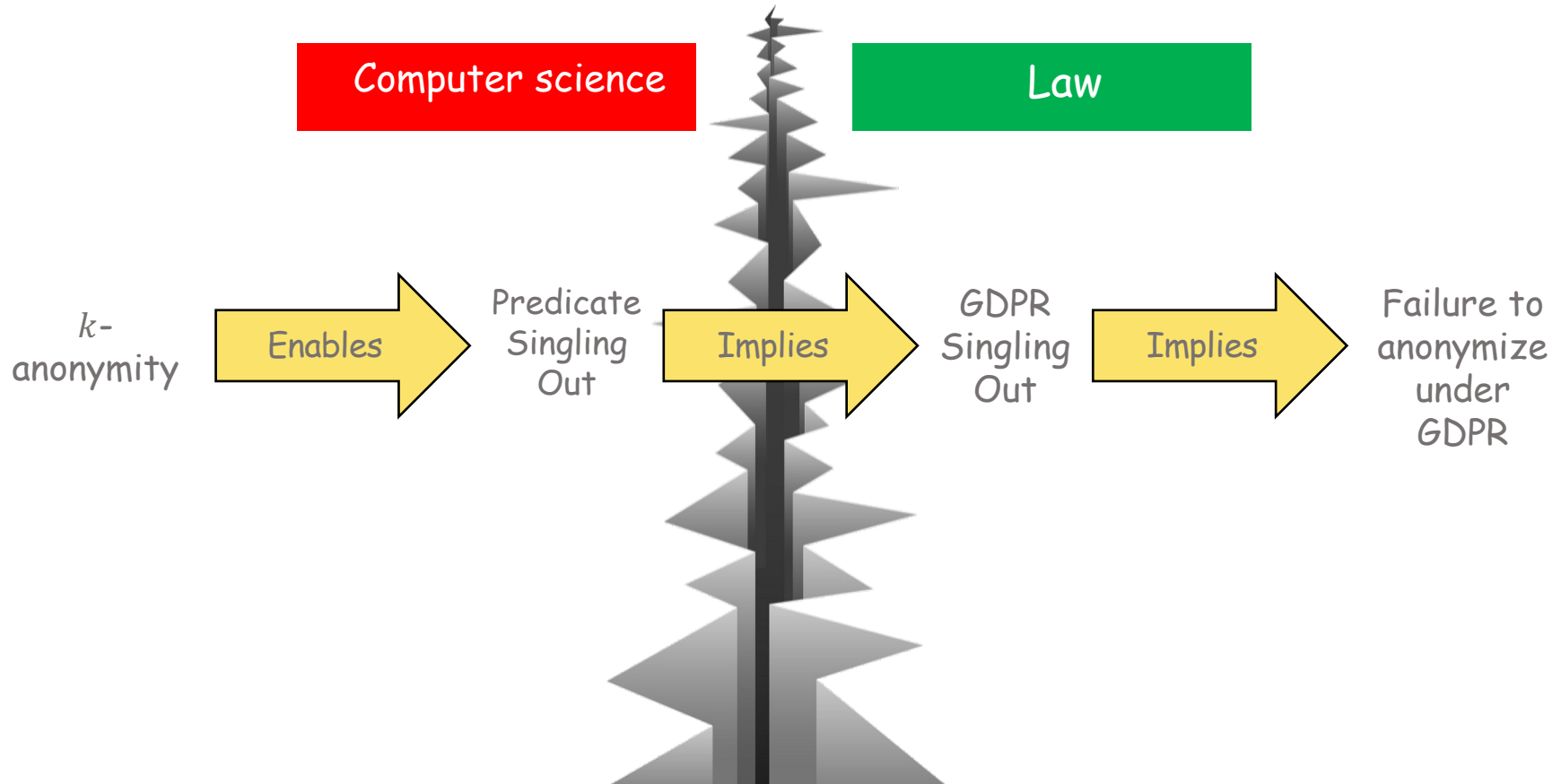
Let's review our modeling assumptions

- Design choices for security against predicate singling out:



- Likely, more restricted than what GDPR regulators had in mind for singling out
- Failure to protect against predicate singling out likely implies failure to protect against GDPR singling out

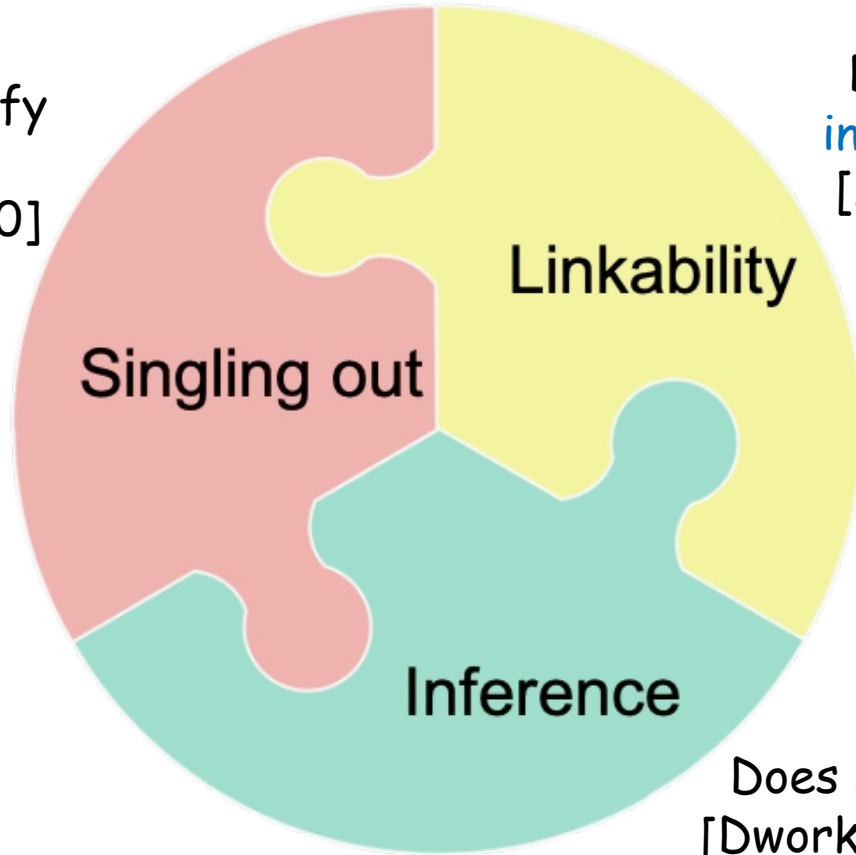
A "legal theorem" for singling out



Does not satisfy
ALTER
[Cohen N 2020]



Predicate
singling out



Does not satisfy
indistinguishability
[Altman Cohen N]

Does not satisfy **NJER**
[Dwork Naor 2006, 2010]



Summary: elements of a methodology

- **Significant gaps** between regulatory and technical conceptions of privacy cause uncertainty, exploitable loopholes, risks to individual privacy
 - Need for a coordination of the two disciplines in the area of privacy [N Wood 2001]
- **A major obstacle: type mismatch**
 - Legal concepts are inherently fuzzy and CS concepts are (and need to be) crisply defined
 - To achieve certainty - find safe zones 'far enough' from the area of legal uncertainty

Summary: elements of a methodology

- Principles for privacy regulation:
 - Reveals weaknesses in existing regulation
 - Provide guidance for future regulation
- Work needed beyond anonymization concepts:
 - Concepts from regulation needing technical treatment: data deletion, statistical purposes, opt out, consent, ...
 - Concepts from technical literature that need to be embedded in regulation: composition, privacy budget, ...

Thank you!