Learn Differential Privacy!
A 4-Part Webinar Series

Michael Hay

CTO
Tumult Labs

Damien Desfontaines

Senior Scientist
Tumult Labs

Hal Triedman

Privacy Engineer
Wikimedia Foundation

Goals of the series

1. Learn privacy risks in analyzing and sharing sensitive data about individuals.

2. Understand what is differential privacy (DP), what use cases are good fits for DP.

3. Learn how to design and tune differentially private data releases with Tumult Analytics.

4. Understand the DP deployment process and what it could look like at Wikimedia.

Outline of webinar series

1. **A Non-Technical Introduction to Differential Privacy (DP)**
   Today!

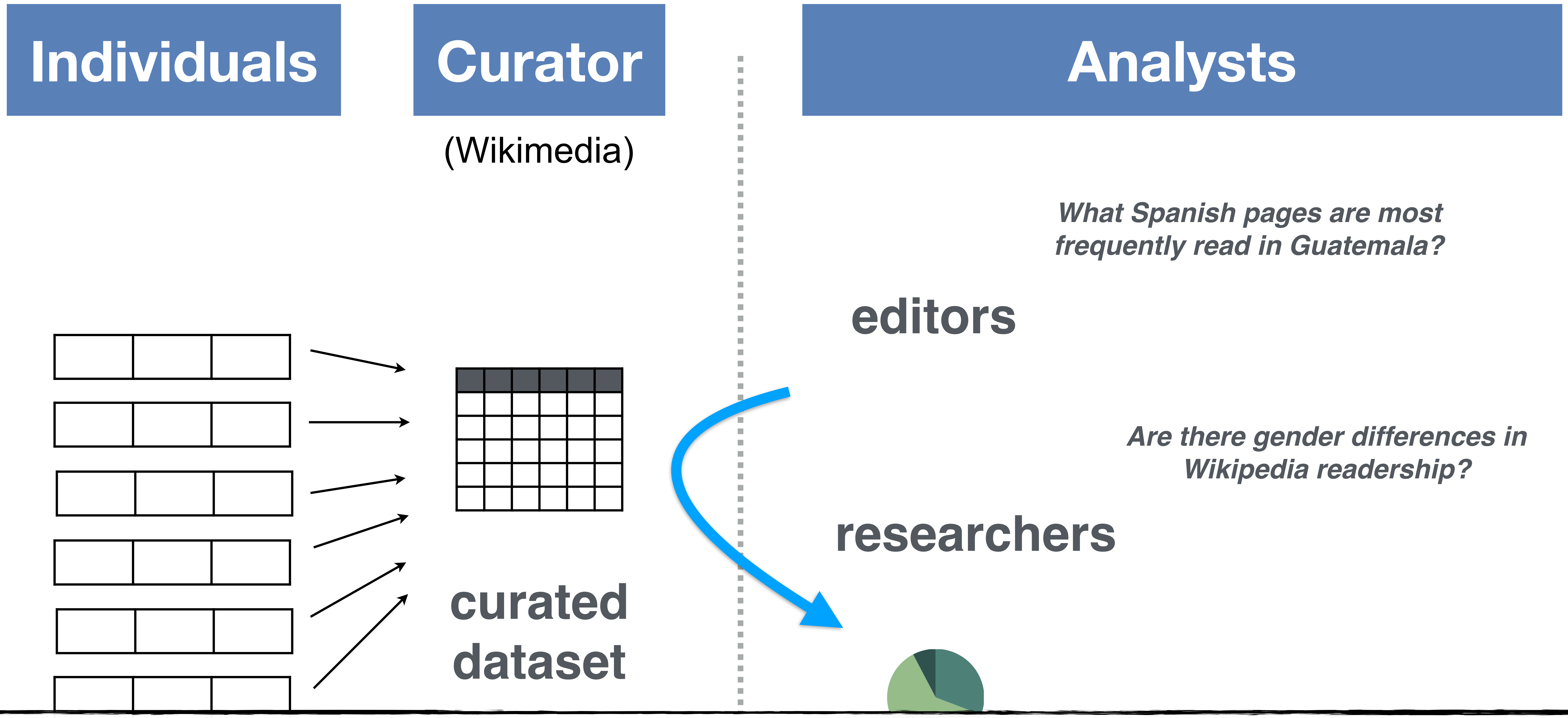
2. **Differential privacy fundamentals, first steps using Tumult Analytics**
   Wednesday, July 13th, 3pm GMT

3. **Boosting the utility of differentially private mechanisms**
   Monday, July 25th, 3pm GMT

4. **Deploying differential privacy at Wikimedia**
   Wednesday, July 27th, 3pm GMT

Outline for Module 1: A Non-Technical Introduction to Differential Privacy

- Privacy in data sharing and why this is a hard problem
- Differential Privacy (DP)
- DP Deployment
- What DP deployment at WMF might look like



The Problem

Finding a method for sharing useful data without compromising privacy

Data sharing use cases

WMF wants to give editors an idea of which pages are most visited

Hospitals want to share healthcare data with researchers/students

Government agencies release data to researchers and policy makers

Ride-hailing companies want/need to share data with cities to optimize traffic and make policies on infrastructure use.

Common theme: Need to learn aggregate trends in the data

But sharing raw data is often infeasible

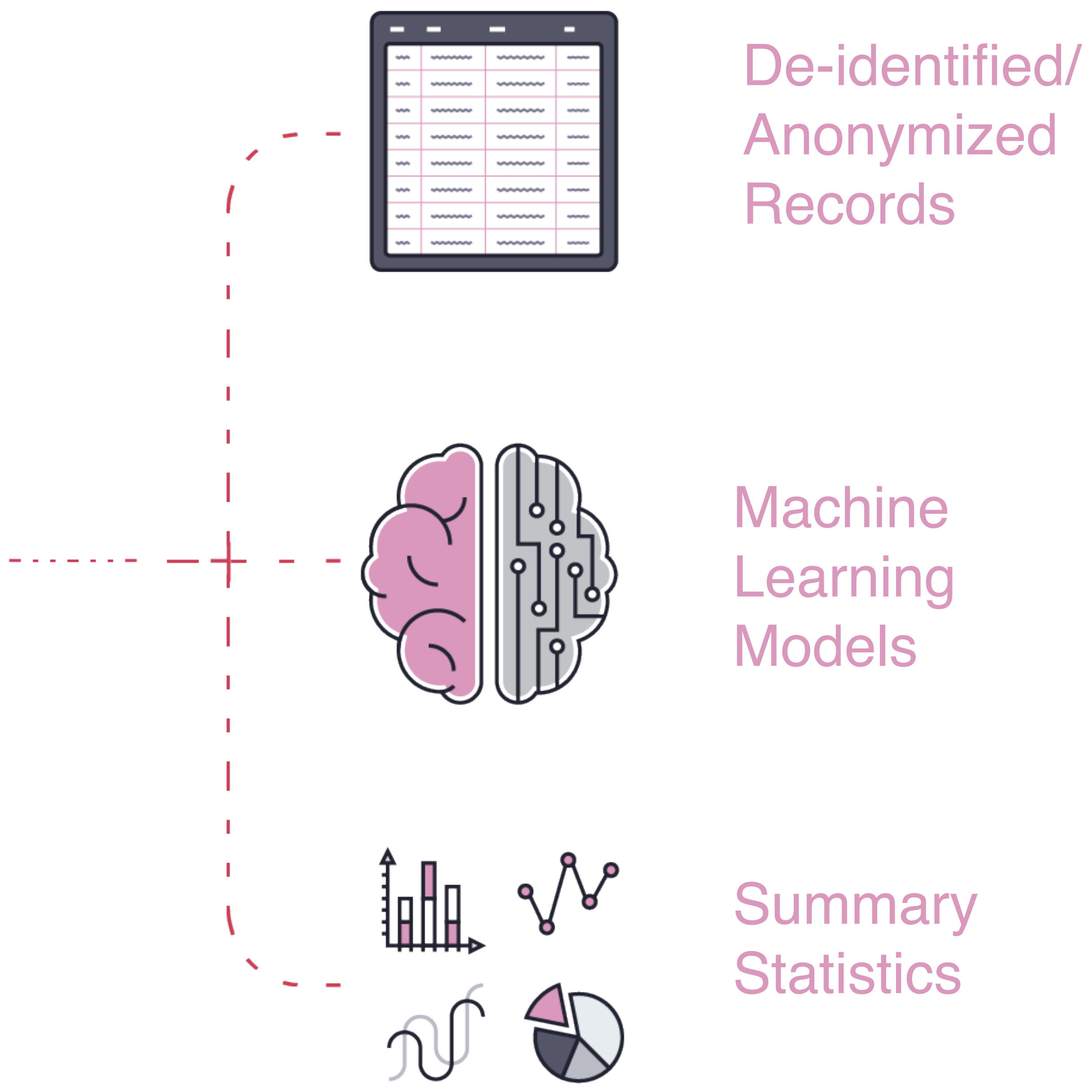
- Data is often very sensitive
- Data sharing, especially at the level of individuals, is often heavily regulated
- Approval procedures introduce delays in sharing
- May be too risky, or violate trust of data participants

"Google and the University of Chicago are Sued Over Data Sharing" NYTimes June 26, 2019

Organizations share anonymized/aggregated data

Sensitive individual-level data

FIRST	LAST	ZIP	EMAIL	SSN
Justin	Roberson	51507	R.Justin@abc.com	104--16-7991
Regina	Hendrix	72010	H.Richar@def.com	506-14-0301
Driscoll	Gonzales	81234	G.Driscoll@ghi.com	904-07-2168
Elliot	Cruz	71804	E.Cruz@jkl.com	206-18-7246
Sylvia	Hayes	20765	H.Sylvia@mno.com	001-29-2511
Nathaniel	Hendricks	18564	H.Nathaniel@pqr.com	103-13-5147
Erica	Solomon	64435	S.Erica@stu.com	003-08-6941
Julian	Preston	09659	P.Julian@vwx.com	307-11-6329
Keely	Bond	43891	B.Keely@yand.com	602-12-3128
Patricia	Williamson	07696	W.Patricia@zee.com	901-10-5103
Denton	Lawson	72344	L.Denton@abc.com	512-18-6293
Branden	Stanton	73389	S.Branden@def.com	091-05-3705
Bethany	Mullins	25279	M.Bethany@ghi.com	164-50-2124
Xena	Humphrey	56418	H.Xena@jkl.com	550-11-8312
Althea	York	40877	Y.Althea@mno.com	604-10-9665
Martin	Arnold	12503	A.Martin@pqr.com	224-54-8270



Protecting privacy is hard even when the data is anonymized or aggregated.

Attacks show data sharing methods fail to protect privacy

Sensitive individual-level data

FIRST	LAST	ZIP	EMAIL	SSN
Justin	Roberson	51507	R.Justin@abc.com	104--16-7991
Regina	Hendrix	72010	H.Richar@def.com	506-14-0301
Driscoll	Gonzales	81234	G.Driscoll@ghi.com	904-07-2168
Elliot	Cruz	71804	E.Cruz@jkl.com	206-18-7246
Sylvia	Hayes	20765	H.Sylvia@mno.com	001-29-2511
Nathaniel	Hendricks	18564	H.Nathaniel@pqr.com	103-13-5147
Erica	Solomon	64435	S.Erica@stu.com	003-08-6941
Julian	Preston	09659	P.Julian@vwx.com	307-11-6329
Keely	Bond	43891	B.Keely@yand.com	602-12-3128
Patricia	Williamson	07696	W.Patricia@zee.com	901-10-5103
Denton	Lawson	72344	L.Denton@abc.com	512-18-6293
Branden	Stanton	73389	S.Branden@def.com	091-05-3705
Bethany	Mullins	25279	M.Bethany@ghi.com	164-50-2124
Xena	Humphrey	56418	H.Xena@jkl.com	550-11-8312
Althea	York	40877	Y.Althea@mno.com	604-10-9665
Martin	Arnold	12503	A.Martin@pqr.com	224-54-8270

Anonymized data are susceptible to re-identification attacks

"A Face is Exposed for AOL Searcher No. 4417749"
New York Times, Aug. 9, 2006

"Why 'Anonymous' Data Sometimes Isn't"
Wired Magazine, Dec. 12, 2007

"Public NYC Taxicab Database Lets You See How Celebrities Tip"
Gawker, Oct. 23, 2014

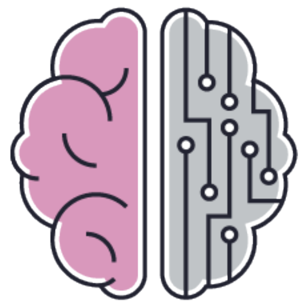
"Credit Card Study Blows Holes in Anonymity"
Science, Jan. 30, 2015

"'Anonymous' Genomes Identified"
The Scientist, May 3, 2013

Attacks show data sharing methods fail to protect privacy

Sensitive individual-level data

FIRST	LAST	ZIP	EMAIL	SSN
Justin	Roberson	51507	R.Justin@abc.com	104--16-7991
Regina	Hendrix	72010	H.Richar@def.com	506-14-0301
Driscoll	Gonzales	81234	G.Driscoll@ghi.com	904-07-2168
Elliot	Cruz	71804	E.Cruz@jkl.com	206-18-7246
Sylvia	Hayes	20765	H.Sylvia@mno.com	001-29-2511
Nathaniel	Hendricks	18564	H.Nathaniel@pqr.com	103-13-5147
Erica	Solomon	64435	S.Erica@stu.com	003-08-6941
Julian	Preston	09659	P.Julian@vwx.com	307-11-6329
Keely	Bond	43891	B.Keely@yand.com	602-12-3128
Patricia	Williamson	07696	W.Patricia@zee.com	901-10-5103
Denton	Lawson	72344	L.Denton@abc.com	512-18-6293
Branden	Stanton	73389	S.Branden@def.com	091-05-3705
Bethany	Mullins	25279	M.Bethany@ghi.com	164-50-2124
Xena	Humphrey	56418	H.Xena@jkl.com	550-11-8312
Althea	York	40877	Y.Althea@mno.com	604-10-9665
Martin	Arnold	12503	A.Martin@pqr.com	224-54-8270



Machine learning is susceptible to inference attacks

"Google, Apple, and others show large language models trained on public data expose personal information"
Venture Beat, Dec. 16, 2020

"Membership inference attacks detect data used to train machine learning models"
Venture Beat, Apr. 28, 2021

Attacks show data sharing methods fail to protect privacy

Sensitive individual-level data

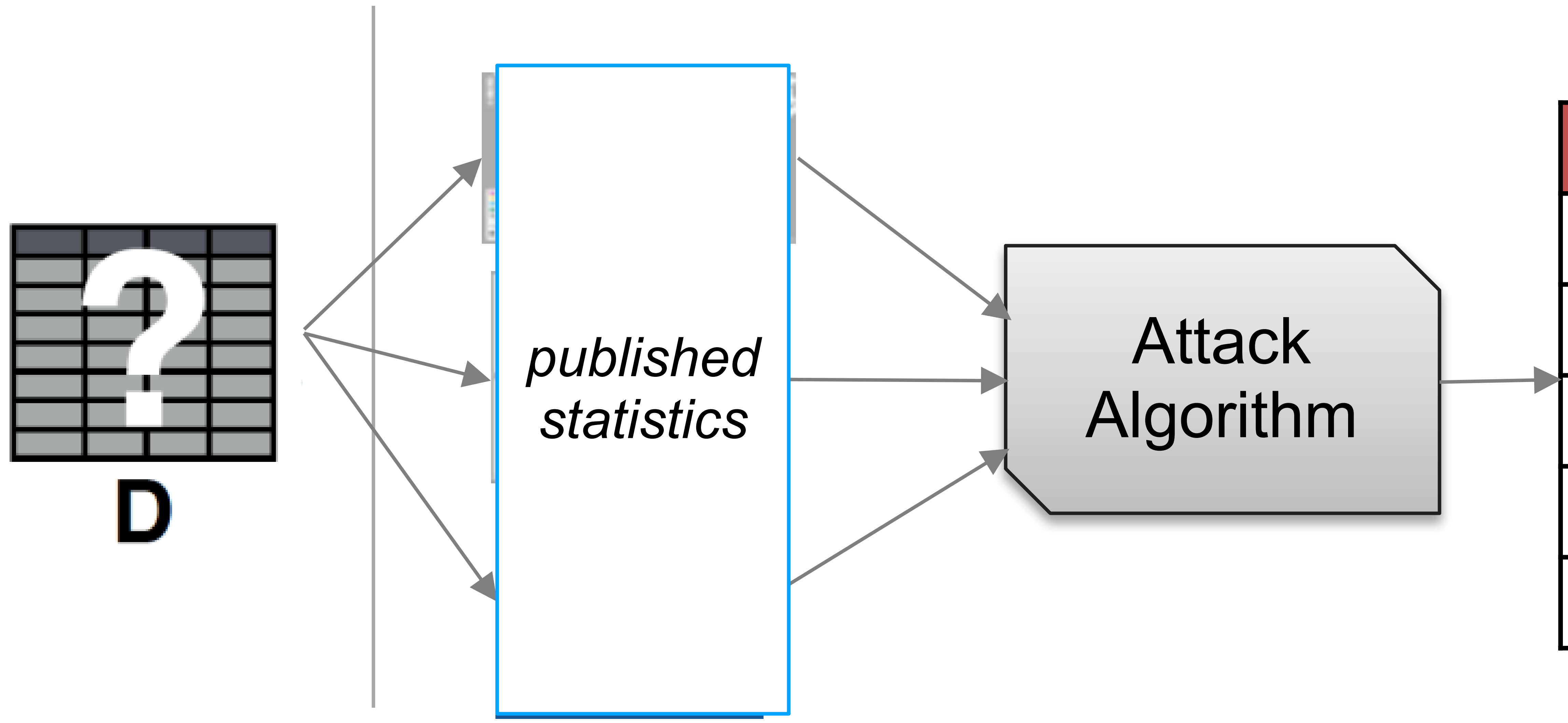
FIRST	LAST	ZIP	EMAIL	SSN
Justin	Roberson	51507	R.Justin@abc.com	104--16-7991
Regina	Hendrix	72010	H.Richar@def.com	506-14-0301
Driscoll	Gonzales	81234	G.Driscoll@ghi.com	904-07-2168
Elliot	Cruz	71804	E.Cruz@jkl.com	206-18-7246
Sylvia	Hayes	20765	H.Sylvia@mno.com	001-29-2511
Nathaniel	Hendricks	18564	H.Nathaniel@pqr.com	103-13-5147
Erica	Solomon	64435	S.Erica@stu.com	003-08-6941
Julian	Preston	09659	P.Julian@vwx.com	307-11-6329
Keely	Bond	43891	B.Keely@yand.com	602-12-3128
Patricia	Williamson	07696	W.Patricia@zee.com	901-10-5103
Denton	Lawson	72344	L.Denton@abc.com	512-18-6293
Branden	Stanton	73389	S.Branden@def.com	091-05-3705
Bethany	Mullins	25279	M.Bethany@ghi.com	164-50-2124
Xena	Humphrey	56418	H.Xena@jkl.com	550-11-8312
Althea	York	40877	Y.Althea@mno.com	604-10-9665
Martin	Arnold	12503	A.Martin@pqr.com	224-54-8270

Statistical data products are susceptible to reconstruction attacks

"Privacy researchers device a noise-exploitation attack that defeats dynamic anonymity"
Tech Crunch, Aug. 17, 2019

"Potential privacy lapse found in Americans' 2010 census data"
AP Feb 16, 2019

Database reconstruction demonstrated by the US Census Bureau



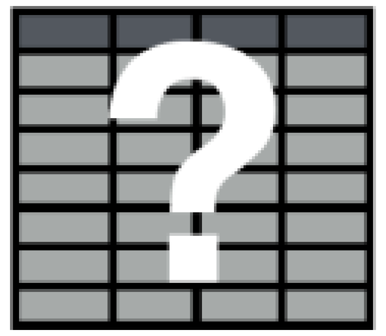
ID	Sex	...	HID
122	M	...	H6
123	F	...	H6
124	M	...	H7
125	M	...	H8
126	F	...	H8

46% of the records (142 million) had exact match on block, sex, age, race categories

2010 Decennial Census

2019 Database Reconstruction attack

Database reconstruction demonstrated by the US Census Bureau



D

published statistics

Attack Algorithm

External Commercial Dataset with identifiers

nam	gen.	age	HR	...
Alice	F	83	65	...
Bob	M	50	85	...
Carl	M	23	61	...
...	...		...	

ID	Sex	...	HID	Name	SSN
122	M	...	H6	Smith	111
123	F	...	H6	Ravi	222
124	M	...	H7	Susan	333
125	M	...	H8	Bob	444
126	F	...	H8	Jay	555

17% of the records (52 million) exact re-identification

2010 Decennial Census

2019 Database Reconstruction attack

Database reconstruction demonstrated by the US Census Bureau

Death by a thousand cuts!

Each release is “safe” (satisfied Census statistical disclosure limitation process)

But in combination, they can reveal the entire underlying database

Fundamental Law of Information Reconstruction
[Dinur-Nissim PODS 2003]

Takeaways from attacks

- Aggregated data can contain facts about individuals that can be extracted through attacks
- Privacy attacks are becoming increasingly sophisticated
- Attacks show that conventional data sharing methods fail to protect privacy, especially when multiple releases are made independently.
- Organizations are turning to mathematically rigorous privacy standards like differential privacy.

The Problem

Finding a method for sharing useful data without compromising privacy

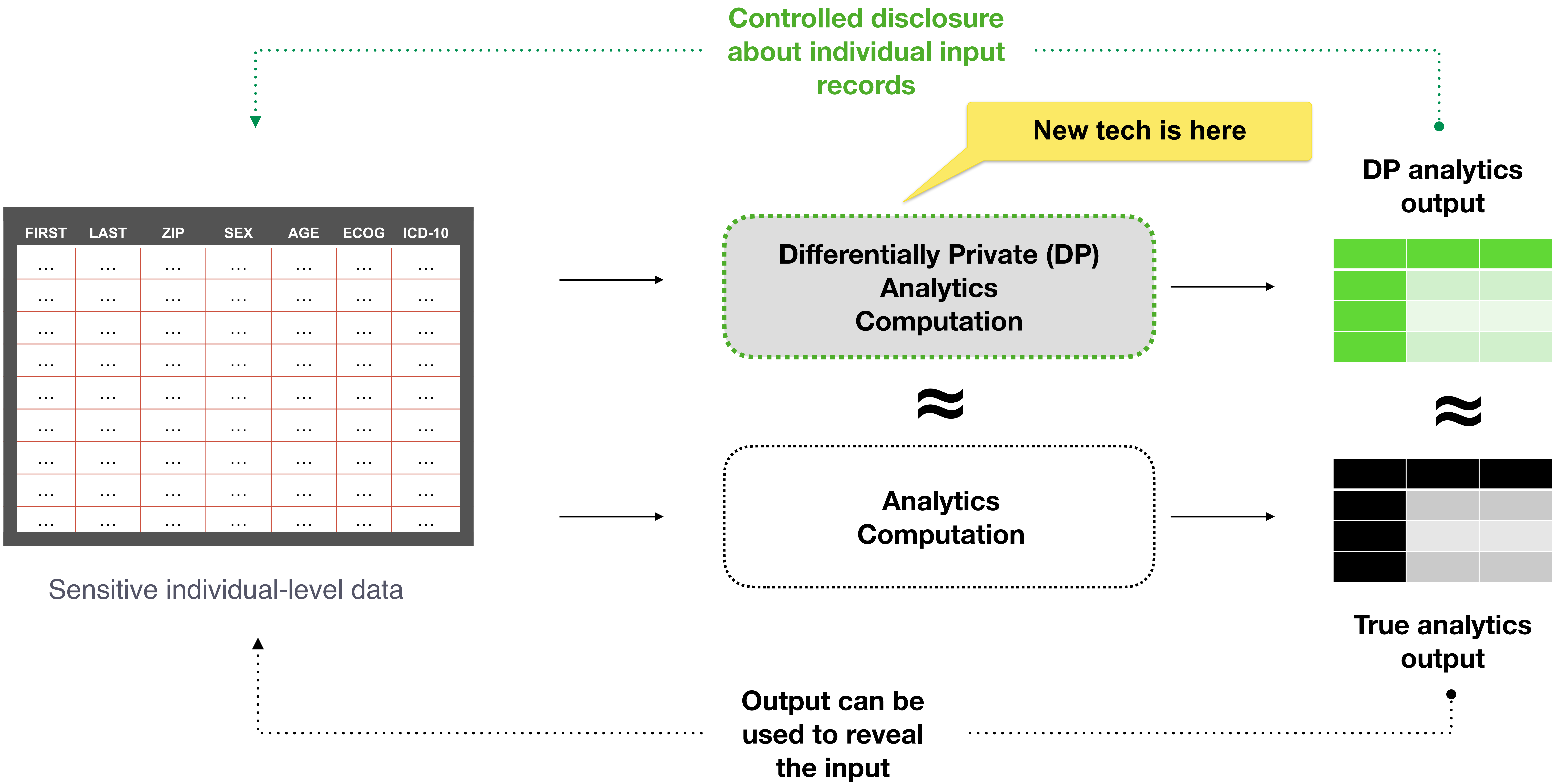
What would a good solution look like?

1. Clear guarantee that quantifies the privacy loss
2. Resilience to attack.
3. Multiple data releases from the same source should NOT lead to total privacy failure
4. Be able to share useful, trust worthy data

Outline for Module 1: A Non-Technical Introduction to Differential Privacy

- Privacy in data sharing and why this is a hard problem
- Differential Privacy (DP)
- DP Deployment
- What DP deployment at WMF might look like

Differential privacy
a standard for computations on data
that limits the personal information that could be revealed by the output.



Two key differences…

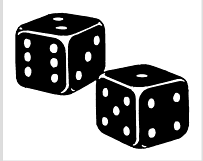
First key difference: randomness

Some “noise” in output

FIRST	LAST	ZIP	SEX	AGE	ECOG	ICD-10
…	…	…	…	…	…	…
…	…	…	…	…	…	…
…	…	…	…	…	…	…
…	…	…	…	…	…	…
…	…	…	…	…	…	…
…	…	…	…	…	…	…
…	…	…	…	…	…	…
…	…	…	…	…	…	…
…	…	…	…	…	…	…

Sensitive individual-level data

Differentially Private (DP) Analytics Computation



≈

Analytics Computation

DP analytics output

	Male	Female
Group=0	322	1034
Group=1	198	2501
Group=2	167	1624

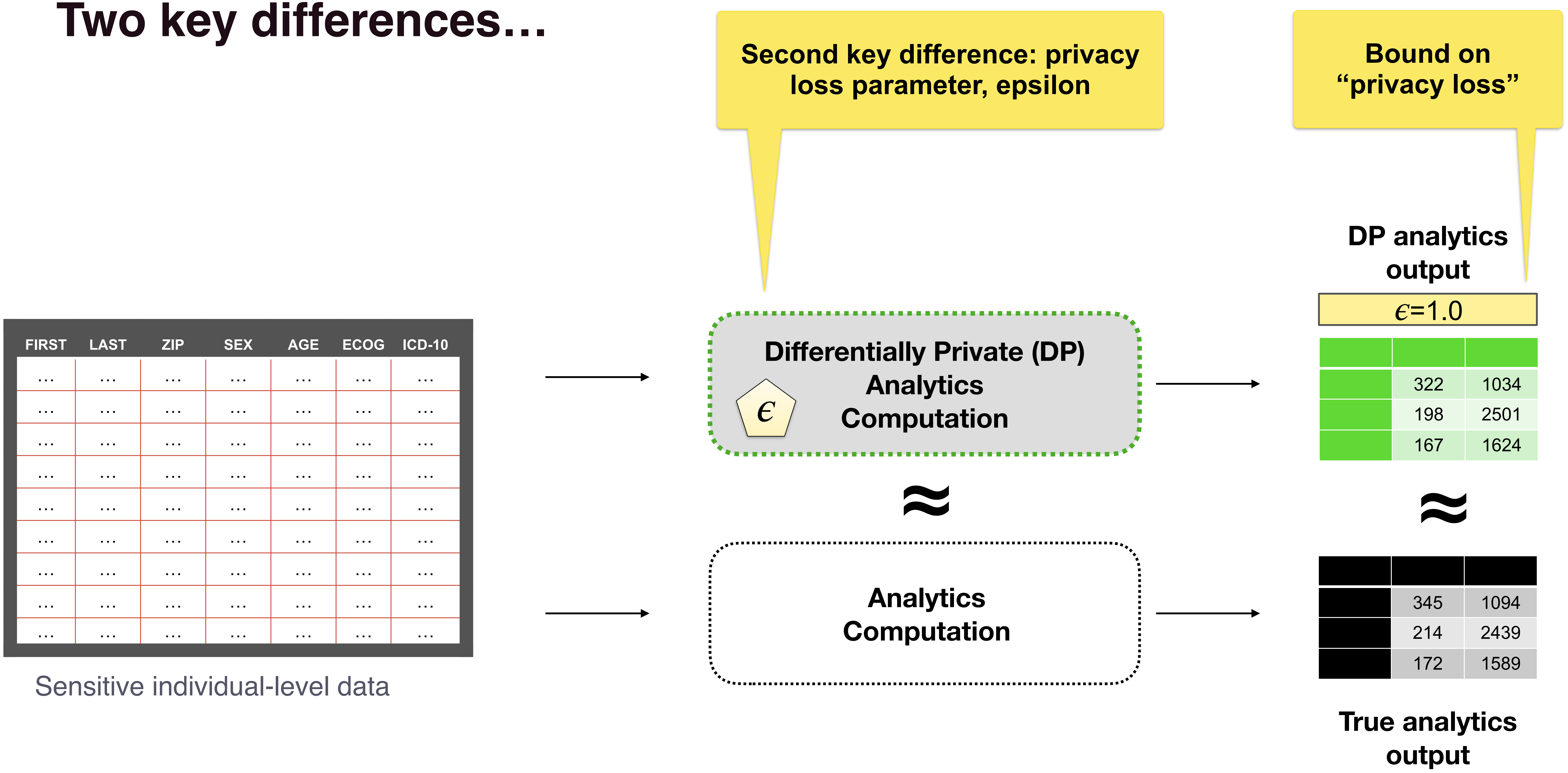
≈

	Male	Female
Group=0	345	1094
Group=1	214	2439
Group=2	172	1589

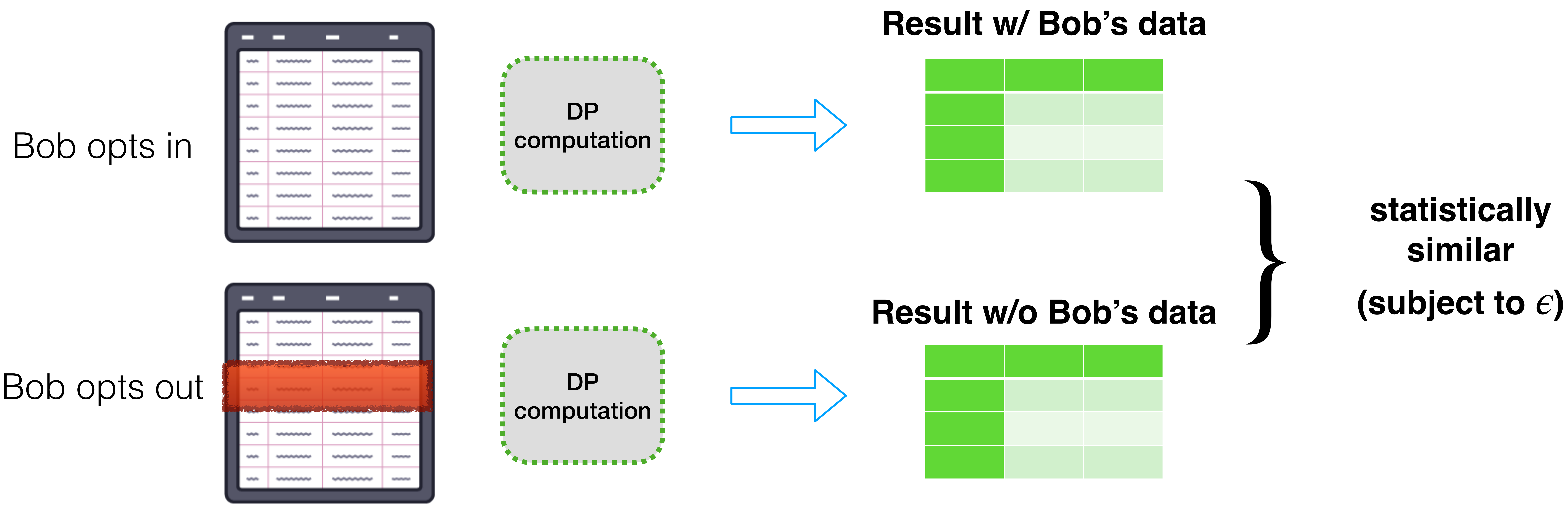
True analytics output

Dice By Pearson Scott Foresman - This file has been extracted from another file, Public Domain, https://commons.wikimedia.org/w/index.php?curid=5553369

Two key differences…



The differential privacy standard (informally)



The computation must be insensitive to a "small" change in the input (adding or removing the data of ***any*** single person*)*

Interpreting differential privacy (DP)

- **DP ensures Plausible deniability:**
  - Suppose attacker uses result to make an inference about Bob's record.
  - Bob can *plausibly deny* the attacker's claim by arguing the attacker would have made the *same inference* even if Bob's record were not in the data.
- **DP is a relative guarantee:**
  - Differential privacy does not mean there is no risk to any individuals.
  - Differential privacy bounds the *additional risk* to an individual due to their data being included in the computation.
  - Parameter epsilon quantifies the additional risk of any "worst-case" scenario

Differential privacy

a ^parameterized^ standard for computations on data
that limits the personal information that could be revealed by the output.

Guarantee of limited disclosure about input

FIRST	LAST	ZIP	SEX	AGE	ECOG	ICD-10
…	…	…	…	…	…	…
…	…	…	…	…	…	…
…	…	…	…	…	…	…
…	…	…	…	…	…	…
…	…	…	…	…	…	…
…	…	…	…	…	…	…
…	…	…	…	…	…	…
…	…	…	…	…	…	…
…	…	…	…	…	…	…

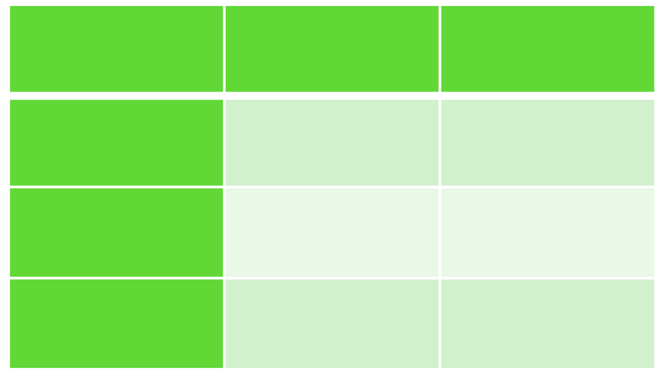
Sensitive individual-level data

Differentially Private (DP) Analytics Computation

ϵ

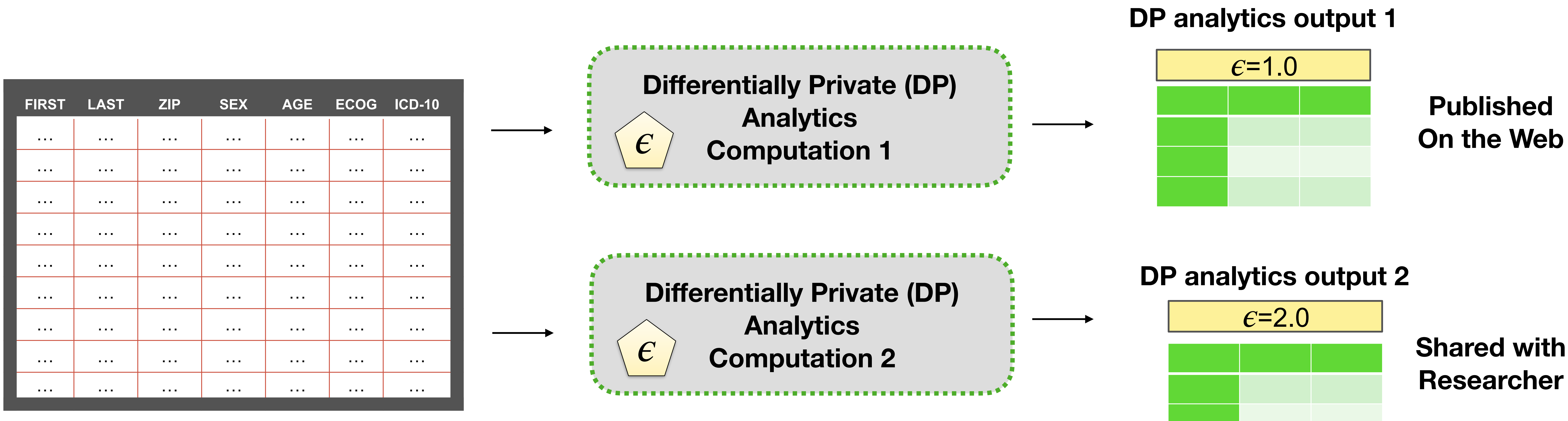
DP analytics output

ϵ=1.0



- Every individual protected.
- Every attribute protected.
- The guarantee holds, regardless of compute power or knowledge of potential attacker.
- Resists current and future attacks

Managing cumulative privacy loss



Sensitive individual-level data

Privacy Accounting

Release	Recipient	Privacy Loss
Analytics 1	Web Pub	ϵ=1.0
Analytics 2	Researcher	ϵ=2.0
TOTAL PRIVACY LOSS		ϵ=3.0

Dice By Pearson Scott Foresman - This file has been extracted from another file, Public Domain, https://commons.wikimedia.org/w/index.php?curid=5553369

Ahead of Regulatory Requirements

The privacy protection provided by differential privacy is stronger than most regulatory requirements.

- Differentially-private outputs are no longer “personal data”.
- Differential privacy has been adopted by the US Census Bureau, and its protection deemed sufficient to meet Title 13.
- Differential privacy has been adopted by the IRS deemed sufficient to meet Title 26
- DP outputs are widely considered to satisfy GPDR’s anonymization standard (prohibiting singling-out of individuals)

Returning to our desiderata…

	Differential privacy
1. Clear guarantee that quantifies the privacy loss	✔
2. Resilience to attack	✔
3. Multiple data releases from the same source should NOT lead to total privacy failure	✔
4. Be able to share useful, trust worthy data	✔

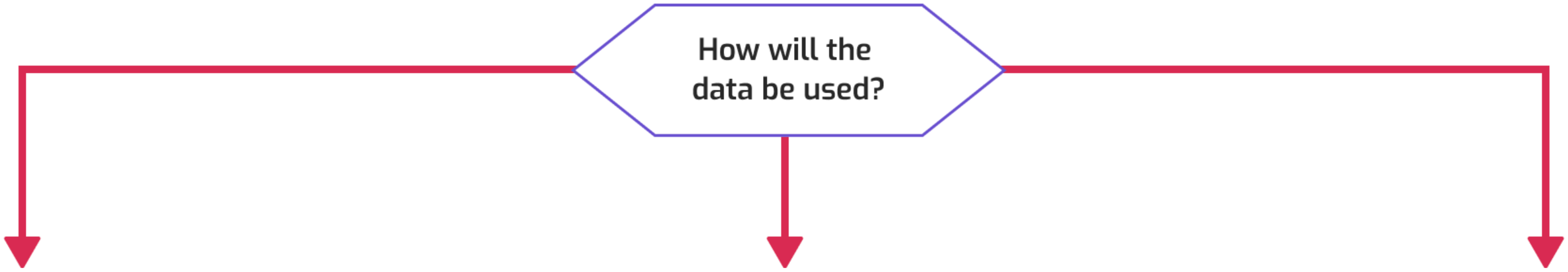
Check out module 2 to find out how!

Outline for Module 1: A Non-Technical Introduction to Differential Privacy

- Privacy in data sharing and why this is a hard problem
- Differential Privacy (DP)
- DP Deployment
- What DP deployment at WMF might look like

https://www.tmlt.io/research/is-differential-privacy-the-right-fit-for-your-problem

What problems are well-suited to DP?



What problems are well-suited to DP?

How will the data be used?

For robust analyses that don't depend too much on individual data points

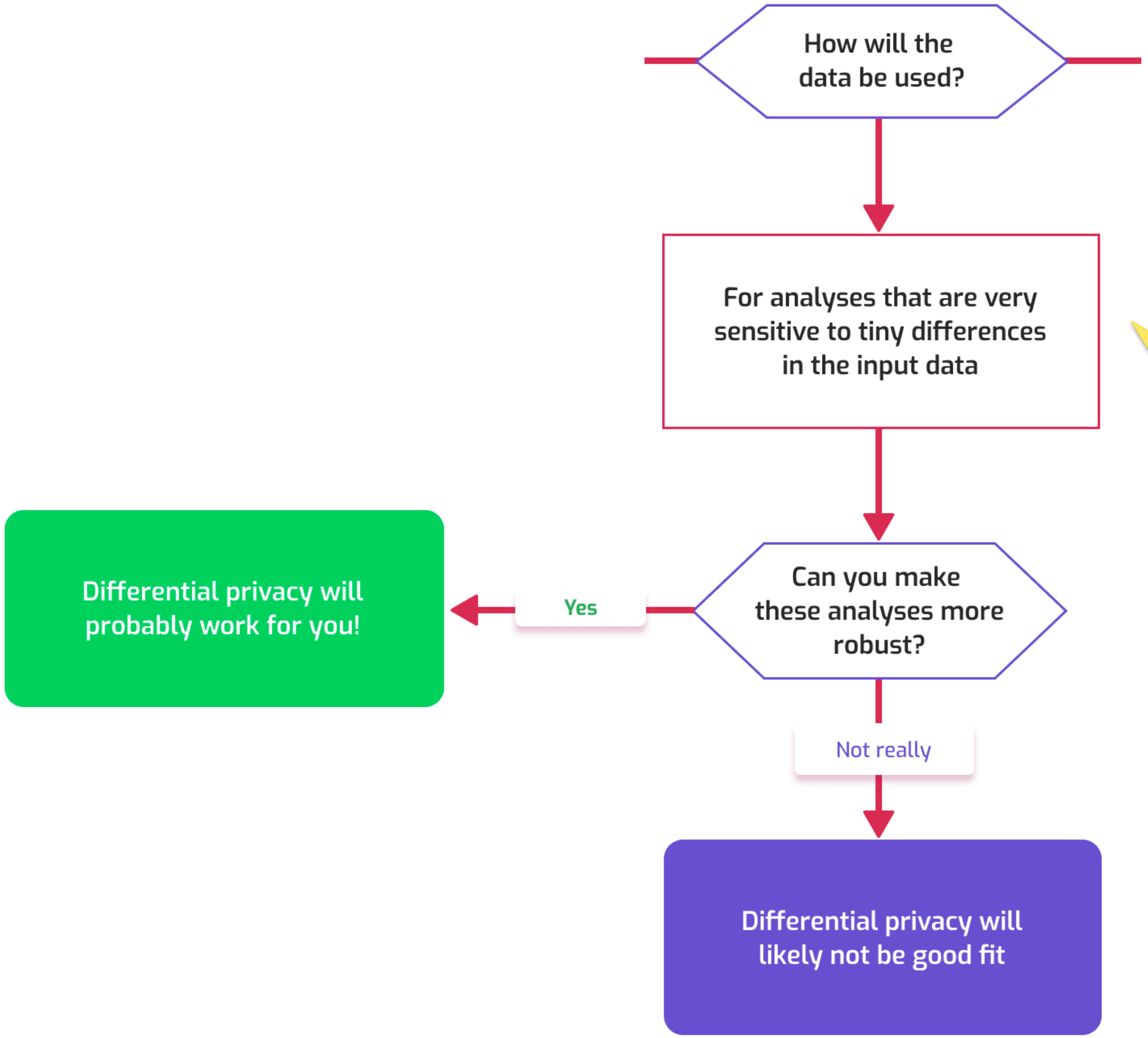
Differential privacy will probably work for you!

- **Statistics and trends about "large" groups**
  Averages, Sums, Counts, (mid range) Quantiles, Histograms, Correlations, Regression, Clustering, Temporal trends, etc.

- **Even if preceded by complicated analysis**
  Joins, Filters, Maps, FlatMaps, Splitting/Grouping

https://www.tmlt.io/research/is-differential-privacy-the-right-fit-for-your-problem

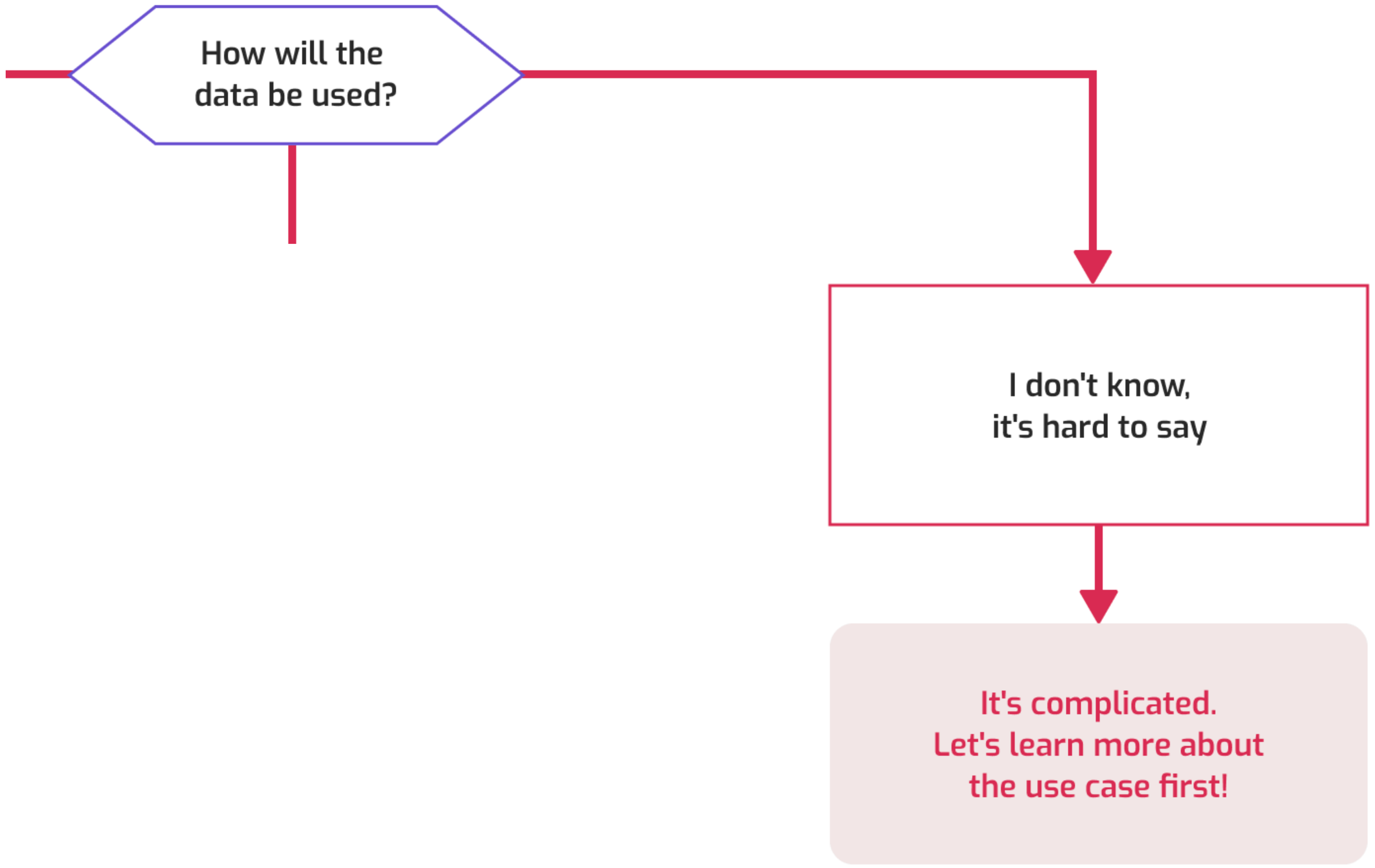
What problems are well-suited to DP?



- **Outlier detection**
- **Small Populations**
- **Analyses that preserve linkability with external data**

https://www.tmlt.io/research/is-differential-privacy-the-right-fit-for-your-problem

What problems are well-suited to DP?



DP Deployment is becoming easier

Differential Privacy Tools

Analytics

Tumult Analytics

Diffprivlib

PipelineDP

OpenDP

SQL

GoogleDP

SmartNoise

Machine Learning

Tensorflow Privacy

Opacus

Synthetic Data

Gretel Synthetics

Sarus

DP Deployment is a process

Improve the strategy

Collect requirements → Build a prototype algorithm → Measure and communicate the error → Deploy 🚀 → Monitor

Outline for Module 1: A Non-Technical Introduction to Differential Privacy

- Privacy in data sharing and why this is a hard problem
- Differential Privacy (DP)
- DP Deployment
- What DP deployment at WMF might look like

Conclusions

DP offers…

- a ***reliable guarantee*** of privacy for all individuals in the source data
- accurate privacy loss “accounting” to facilitate risk management across multiple releases
- ability to perform a variety of statistical computations accurately

Deploying DP…

- is made easier by the availability of consulting services and open-source software
- is a ***process*** that involves design, tuning, monitoring
- is best suited for use cases with clear goals around analyses that are robust to small changes in input

Outline of webinar series

1. **A Non-Technical Introduction to Differential Privacy (DP)**
   Today!

2. **Differential privacy fundamentals, first steps using Tumult Analytics**
   Wednesday, July 13th, 3pm GMT

3. **Boosting the utility of differentially private mechanisms**
   Monday, July 25th, 3pm GMT

4. **Deploying differential privacy at Wikimedia**
   Wednesday, July 27th, 3pm GMT

Thank you! Questions?

Michael Hay,
Tumult Labs
michael@tmlt.io