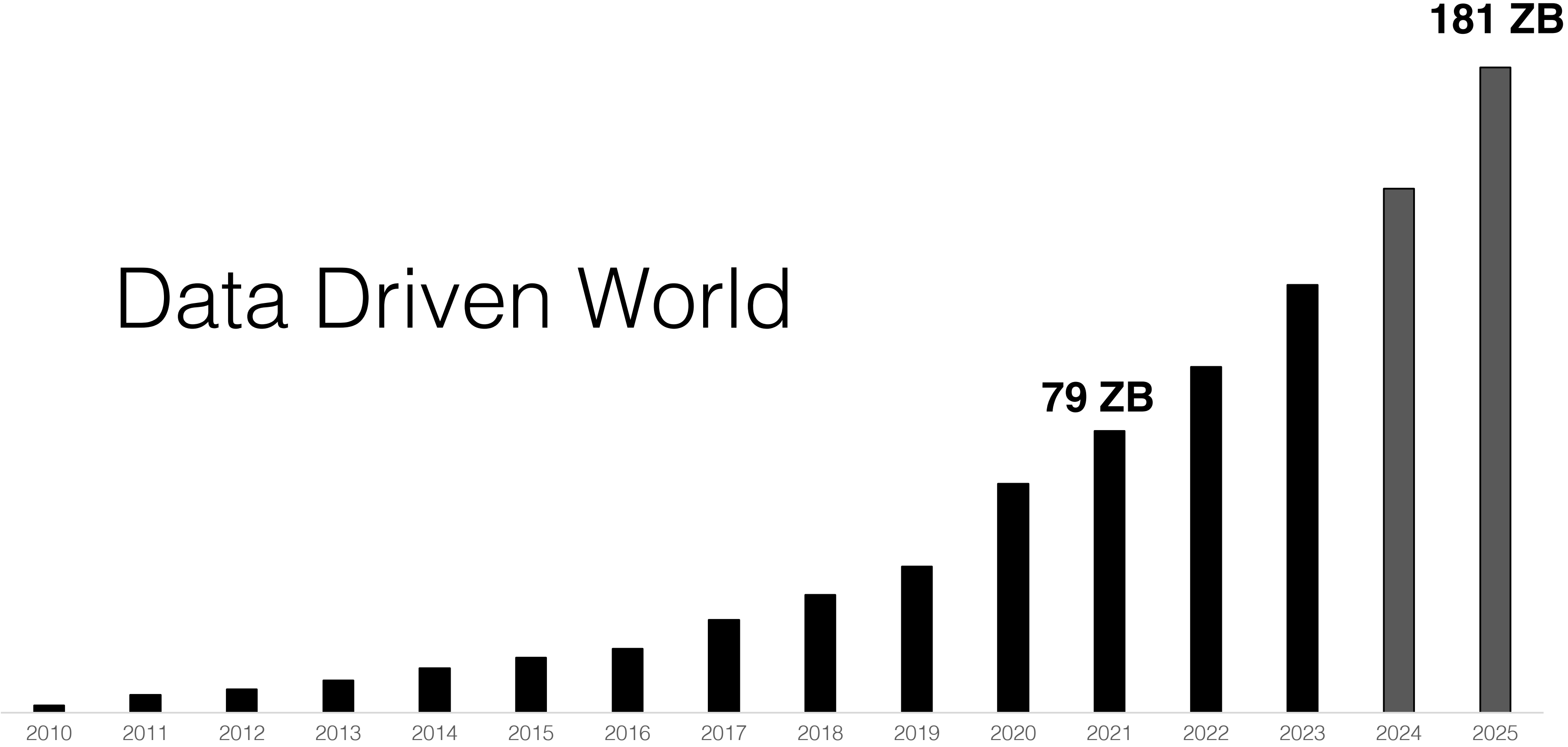
Privacy-Preserving Systems for a Data-Driven World

Anwar Hithnawi

ETH zürich

Data Driven World



Sensitive Data

Smart Homes

Genetics

Dating

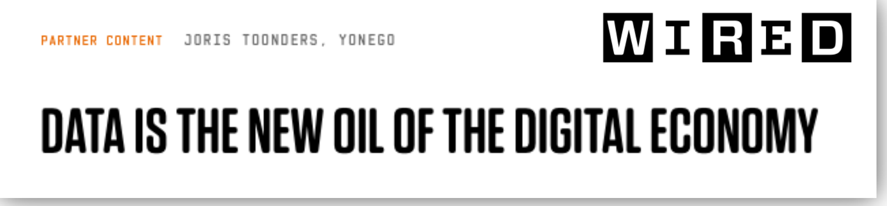
Geolocation

Finance

Health

Government

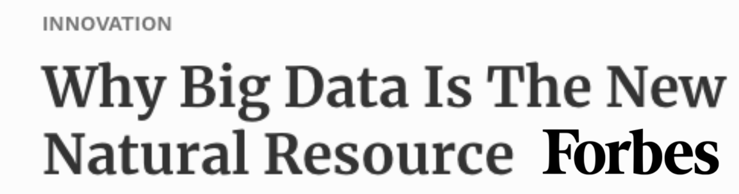
Personal



PARTNER CONTENT JORIS TOONDERS, YONEGO

WIRED

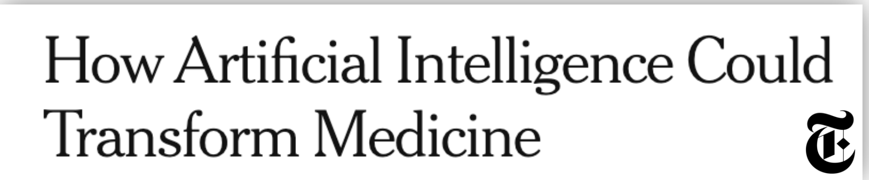
DATA IS THE NEW OIL OF THE DIGITAL ECONOMY



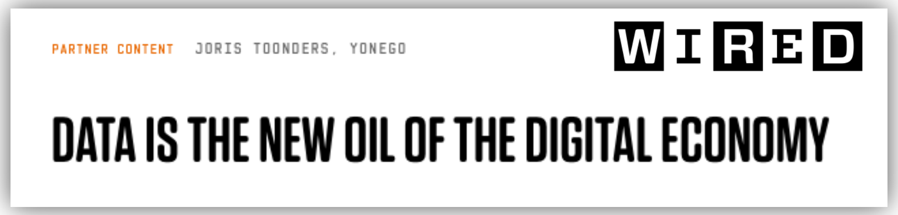
INNOVATION

Why Big Data Is The New Natural Resource

Forbes



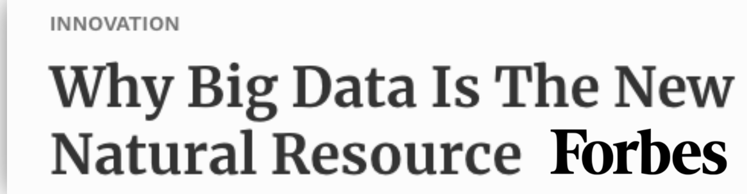
How Artificial Intelligence Could Transform Medicine



PARTNER CONTENT JORIS TOONDERS, YONEGO

WIRED

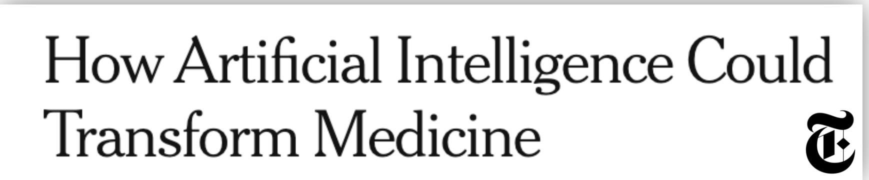
DATA IS THE NEW OIL OF THE DIGITAL ECONOMY



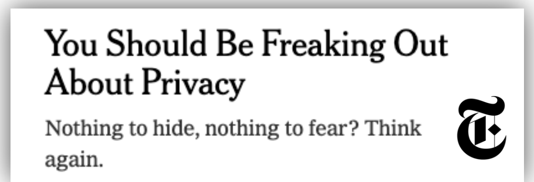
INNOVATION

Why Big Data Is The New Natural Resource

Forbes

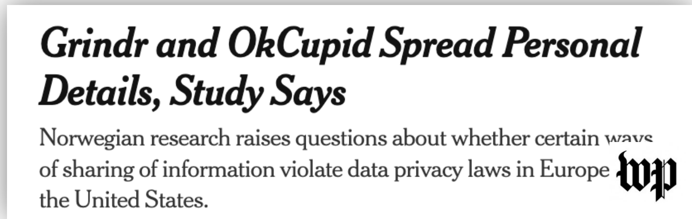


How Artificial Intelligence Could Transform Medicine



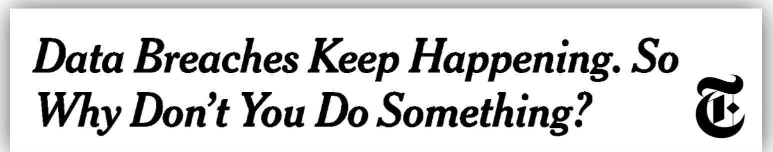
You Should Be Freaking Out About Privacy

Nothing to hide, nothing to fear? Think again.

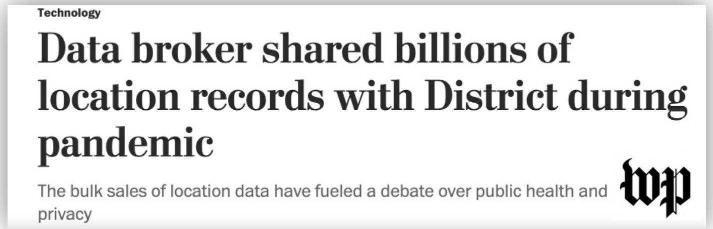


Grindr and OkCupid Spread Personal Details, Study Says

Norwegian research raises questions about whether certain ways of sharing of information violate data privacy laws in Europe the United States.



Data Breaches Keep Happening. So Why Don't You Do Something?

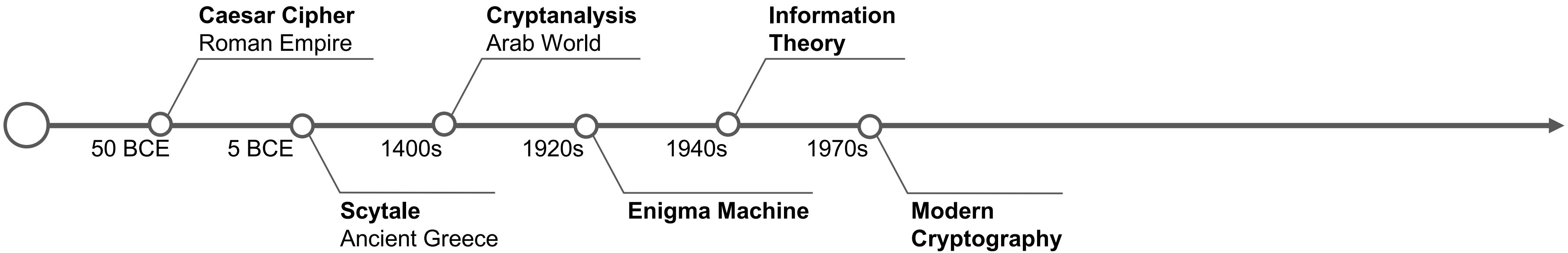


Technology

Data broker shared billions of location records with District during pandemic

The bulk sales of location data have fueled a debate over public health and privacy

Data Protection: An Age-Old Concern



Caesar Cipher
Roman Empire

Scytale
Ancient Greece

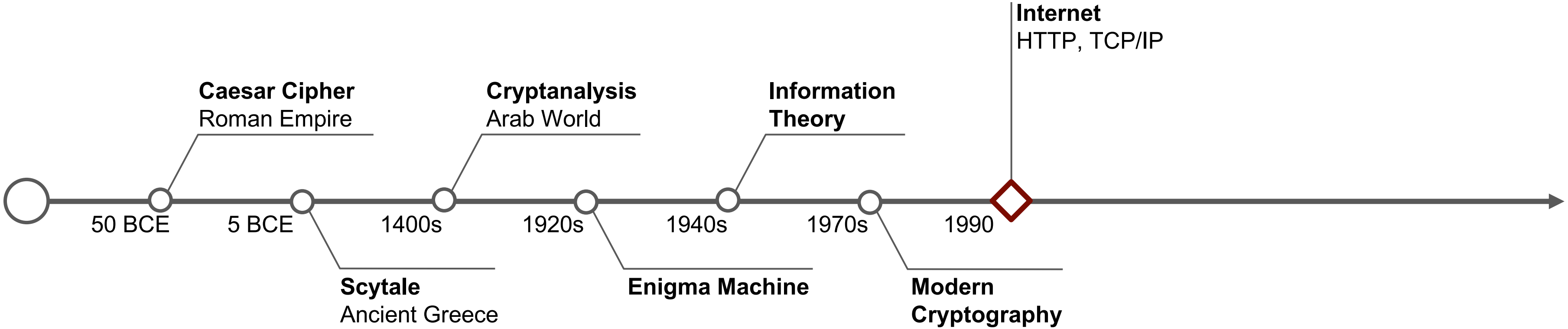
Cryptanalysis
Arab World

Enigma Machine

Information Theory

Modern Cryptography

Data Protection: An Age-Old Concern



Caesar Cipher
Roman Empire

50 BCE

Scytale
Ancient Greece

5 BCE

Cryptanalysis
Arab World

1400s

Enigma Machine

1920s

Information Theory

1940s

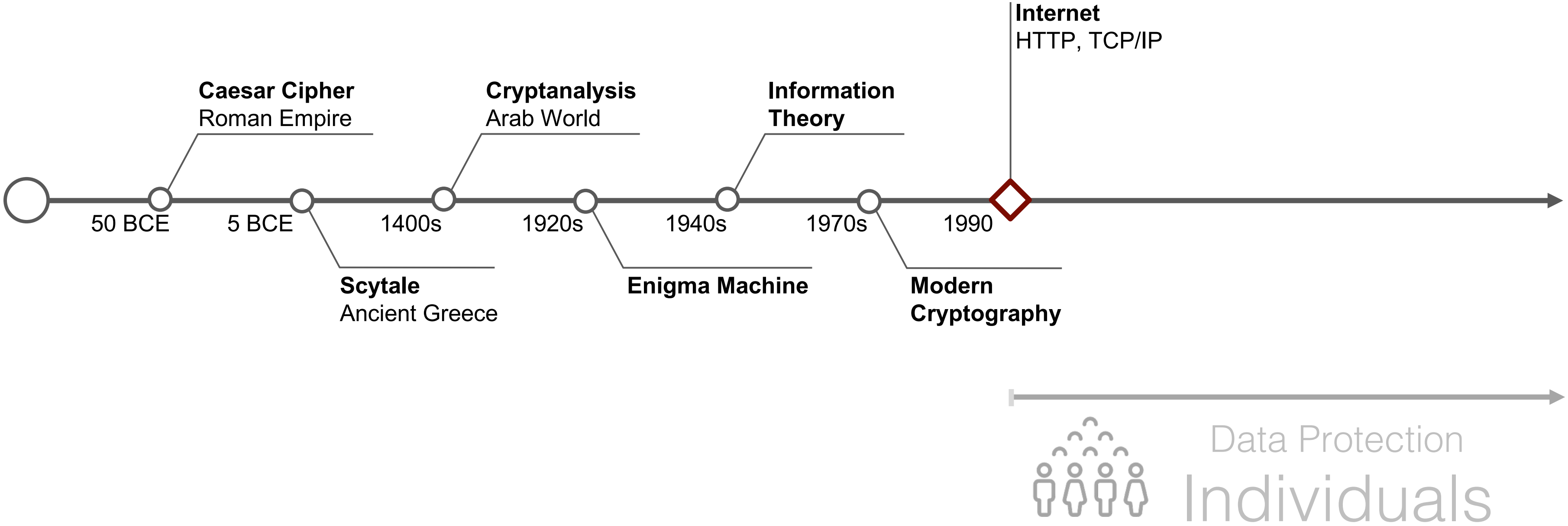
Modern Cryptography

1970s

Internet
HTTP, TCP/IP

1990

Data Protection: An Age-Old Concern



50 BCE

Caesar Cipher
Roman Empire

5 BCE

Scytale
Ancient Greece

1400s

Cryptanalysis
Arab World

1920s

Enigma Machine

1940s

Information Theory

1970s

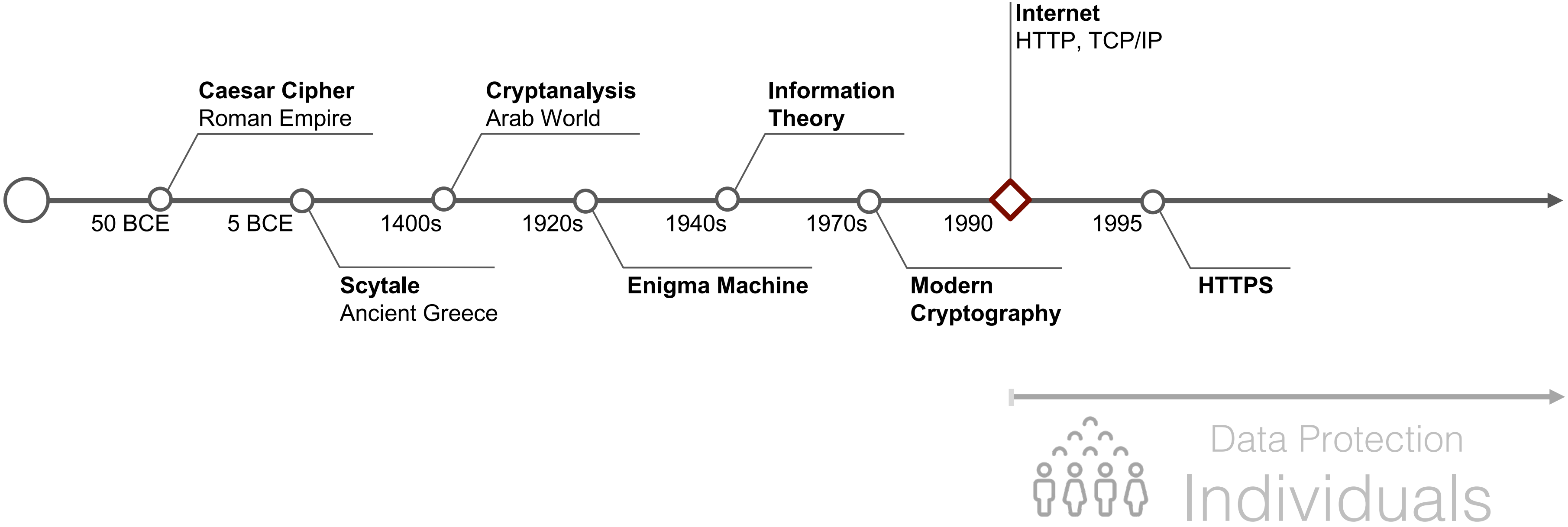
Modern Cryptography

1990

Internet
HTTP, TCP/IP

Data Protection
Individuals

Data Protection: An Age-Old Concern



50 BCE

Caesar Cipher
Roman Empire

5 BCE

Scytale
Ancient Greece

1400s

Cryptanalysis
Arab World

1920s

Enigma Machine

1940s

Information Theory

1970s

Modern Cryptography

1990

Internet
HTTP, TCP/IP

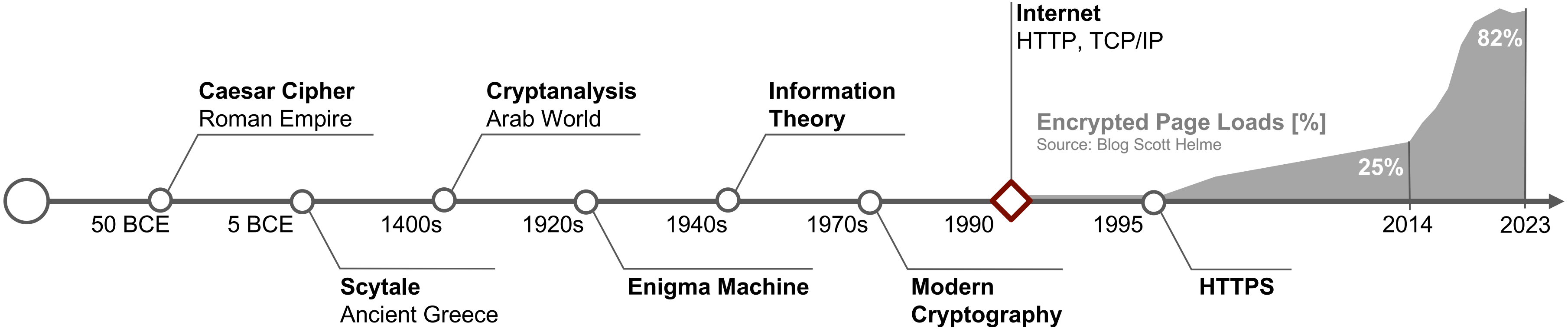
1995

HTTPS

Data Protection

Individuals

Data Protection: An Age-Old Concern



Caesar Cipher
Roman Empire

Scytale
Ancient Greece

Cryptanalysis
Arab World

Enigma Machine

Information Theory

Modern Cryptography

Internet
HTTP, TCP/IP

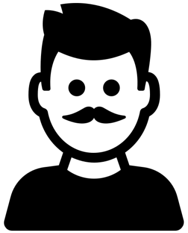
HTTPS

Encrypted Page Loads [%]
Source: Blog Scott Helme

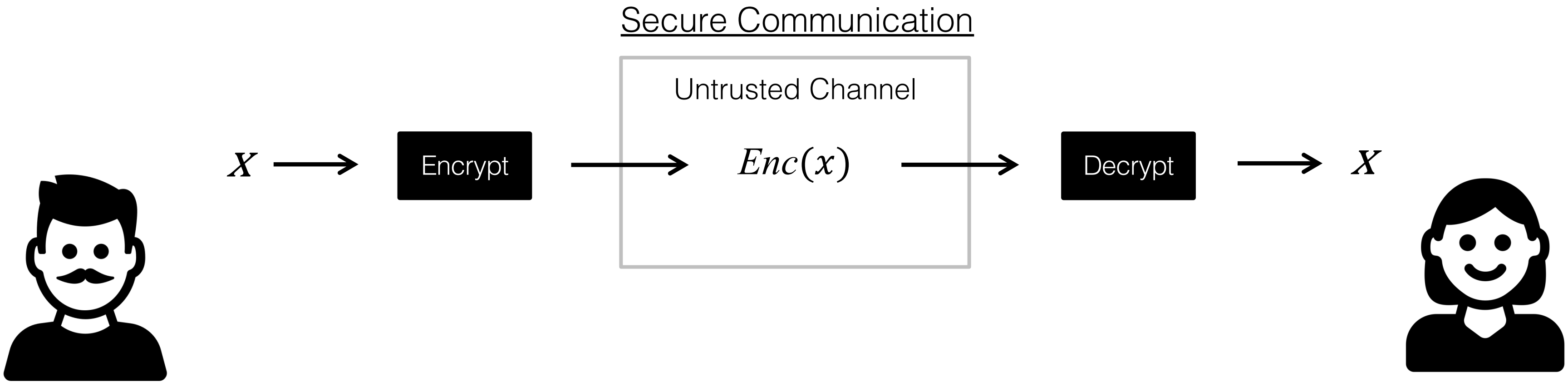
Data Protection

Individuals

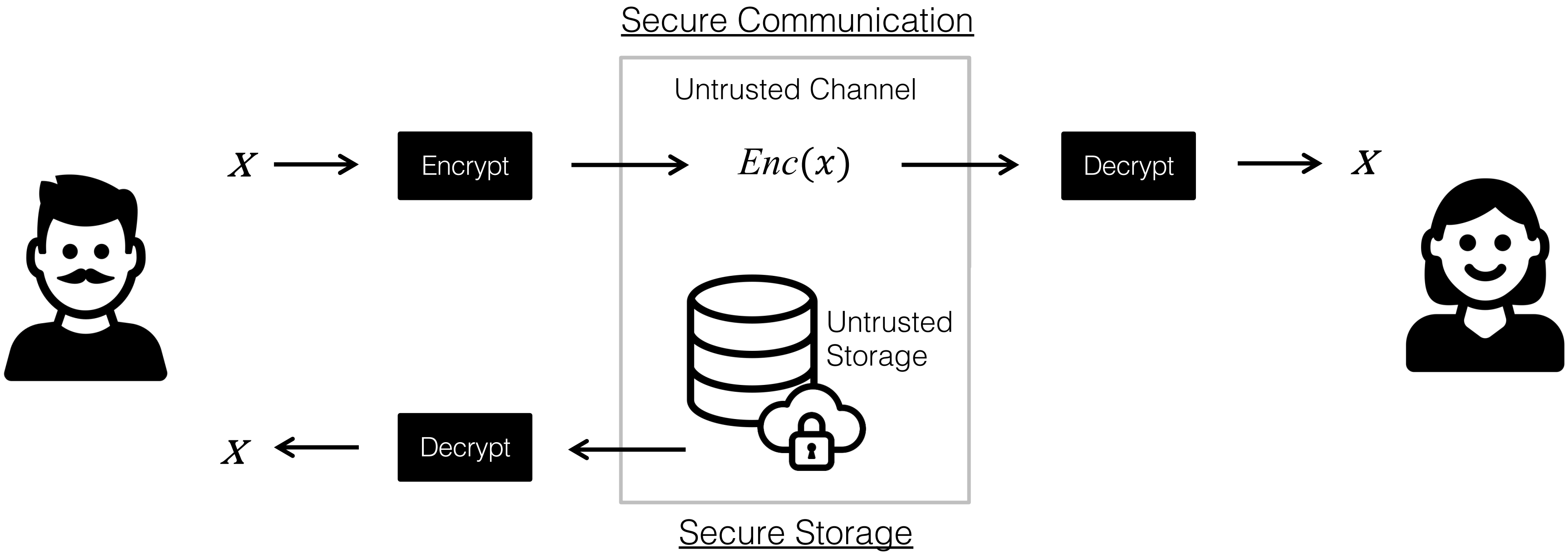
Securing Data: Building Blocks



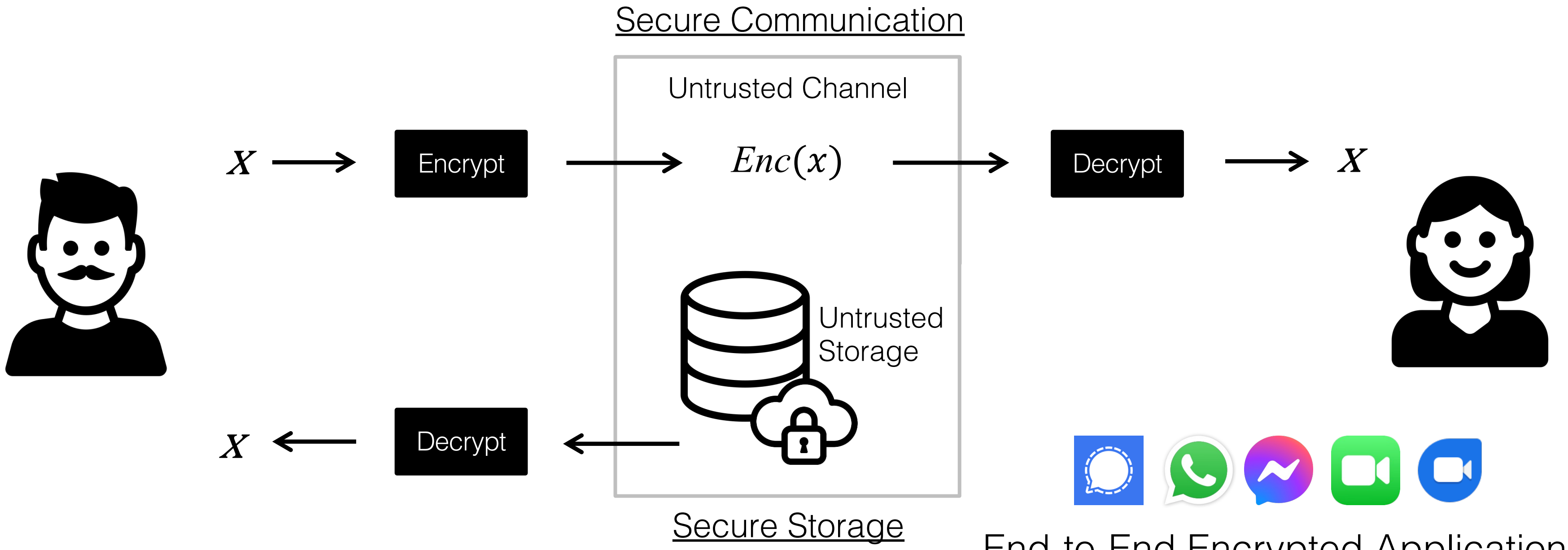
Securing Data: Building Blocks



Securing Data: Building Blocks

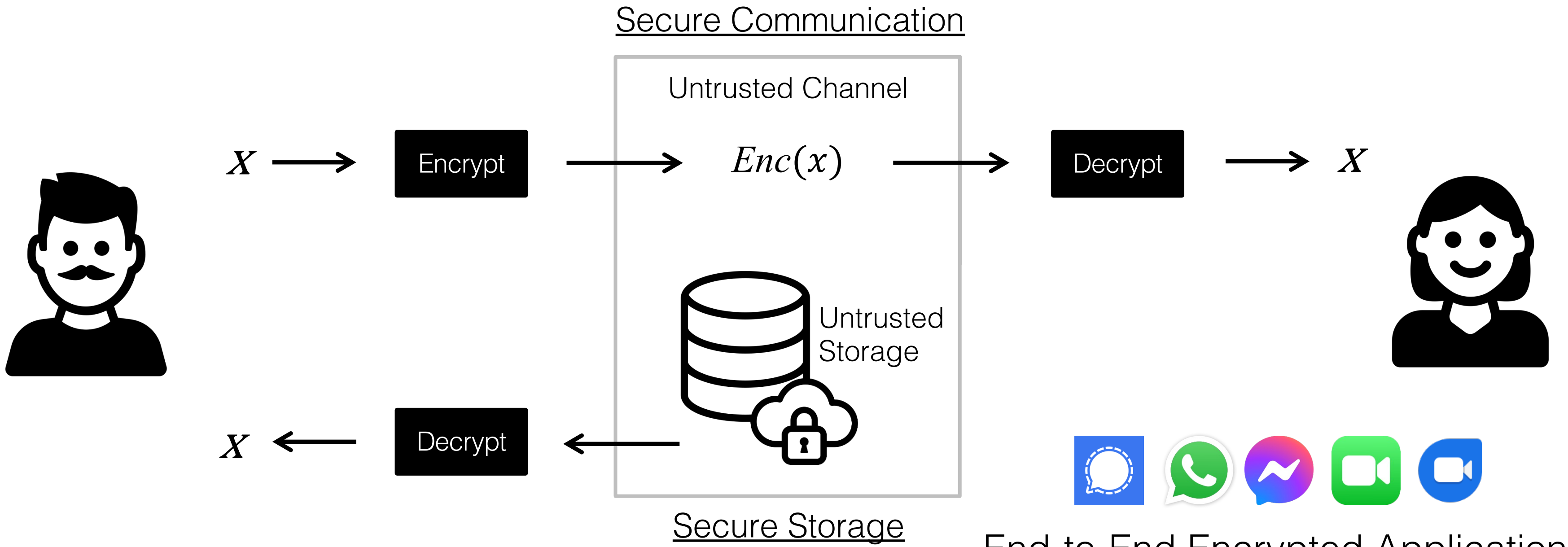


Securing Data: Building Blocks



End-to-End Encrypted Applications

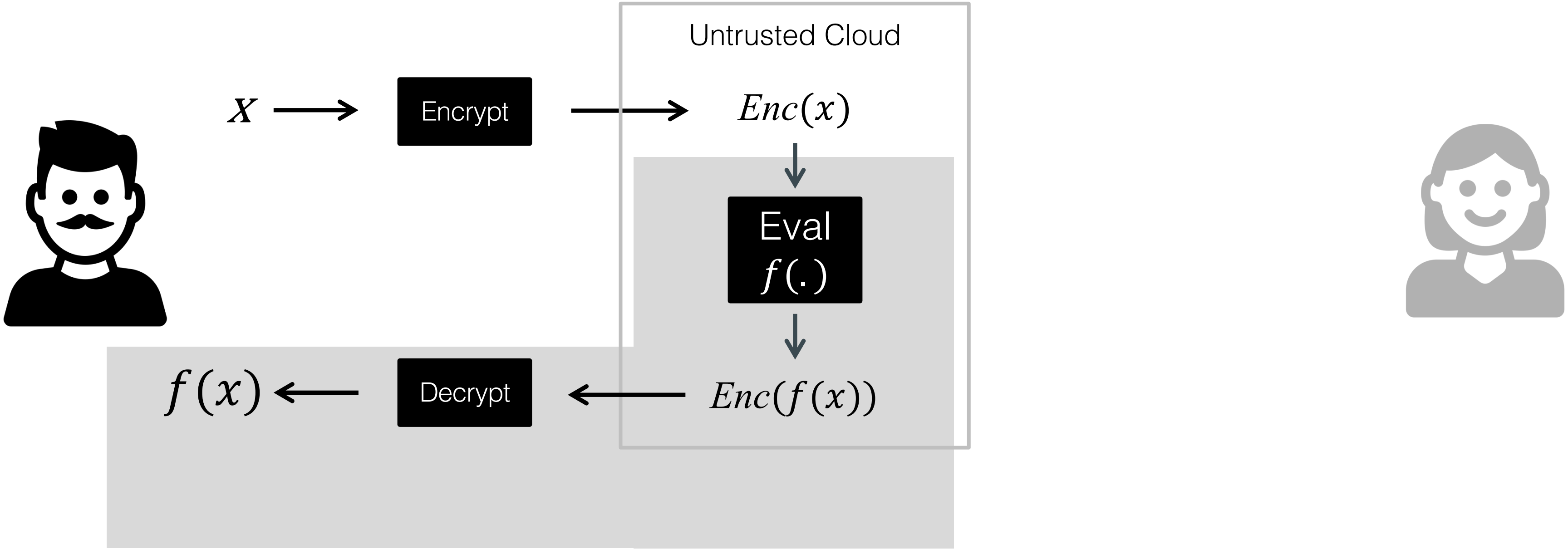
Securing Data: Building Blocks



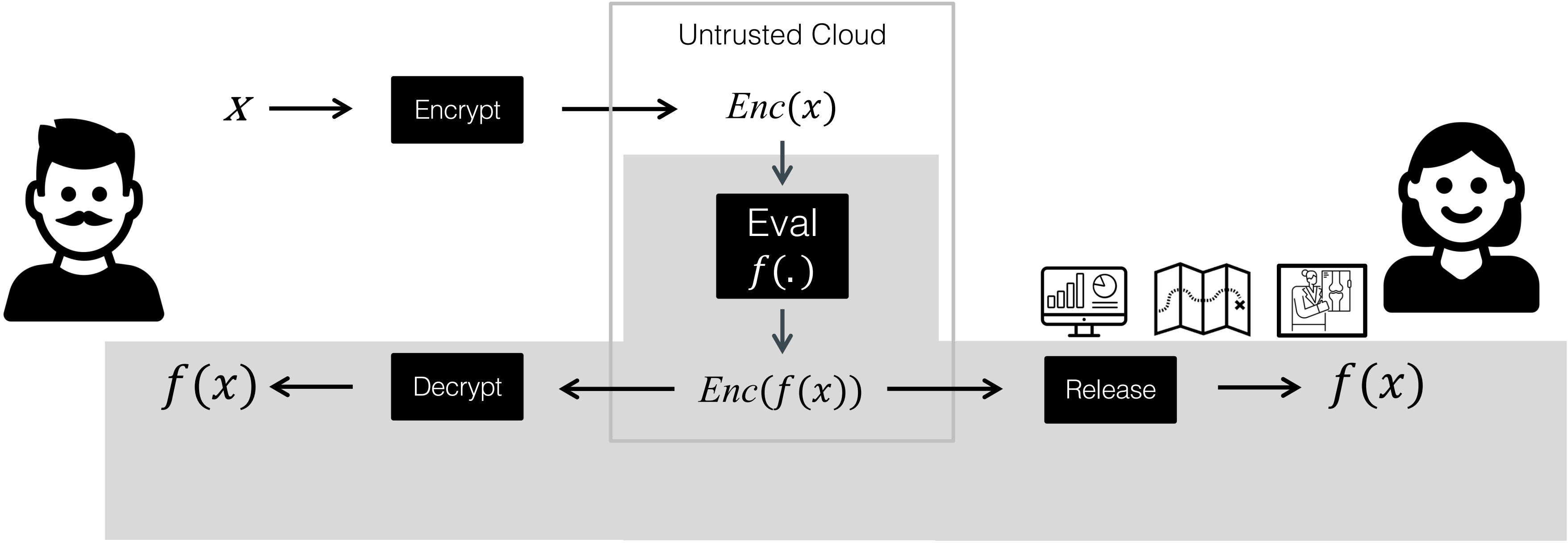
End-to-End Encrypted Applications

More Applications?

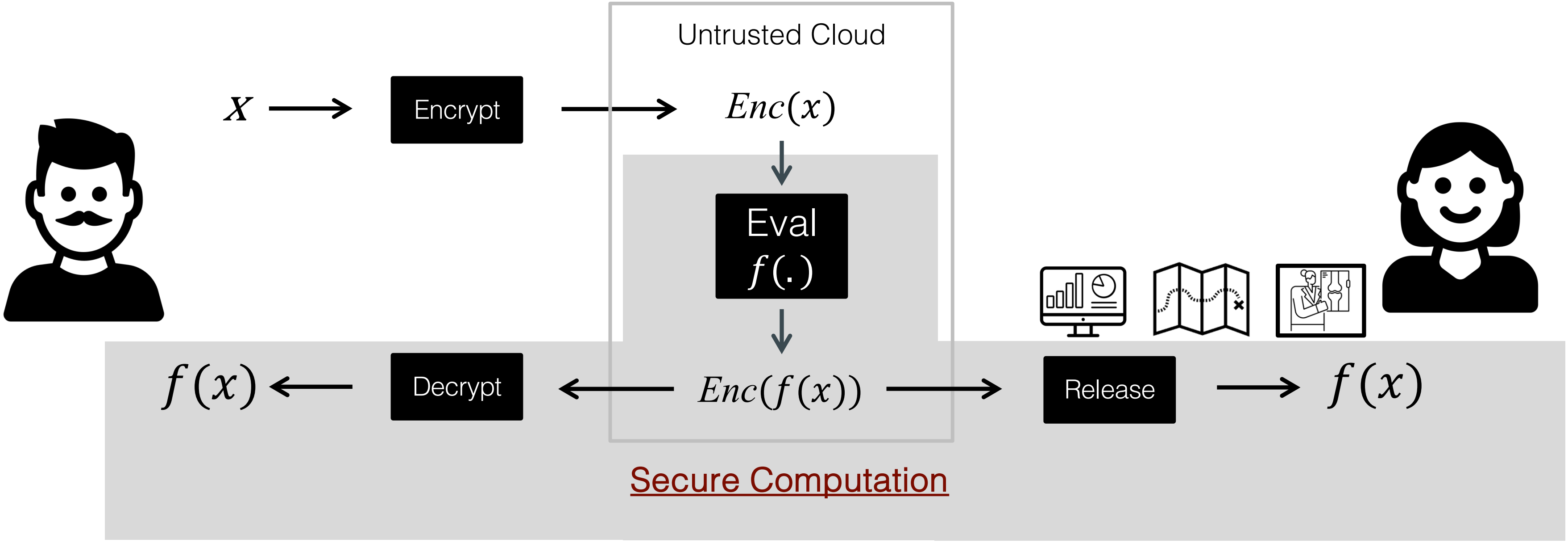
Securing Data in Use: Modern Applications



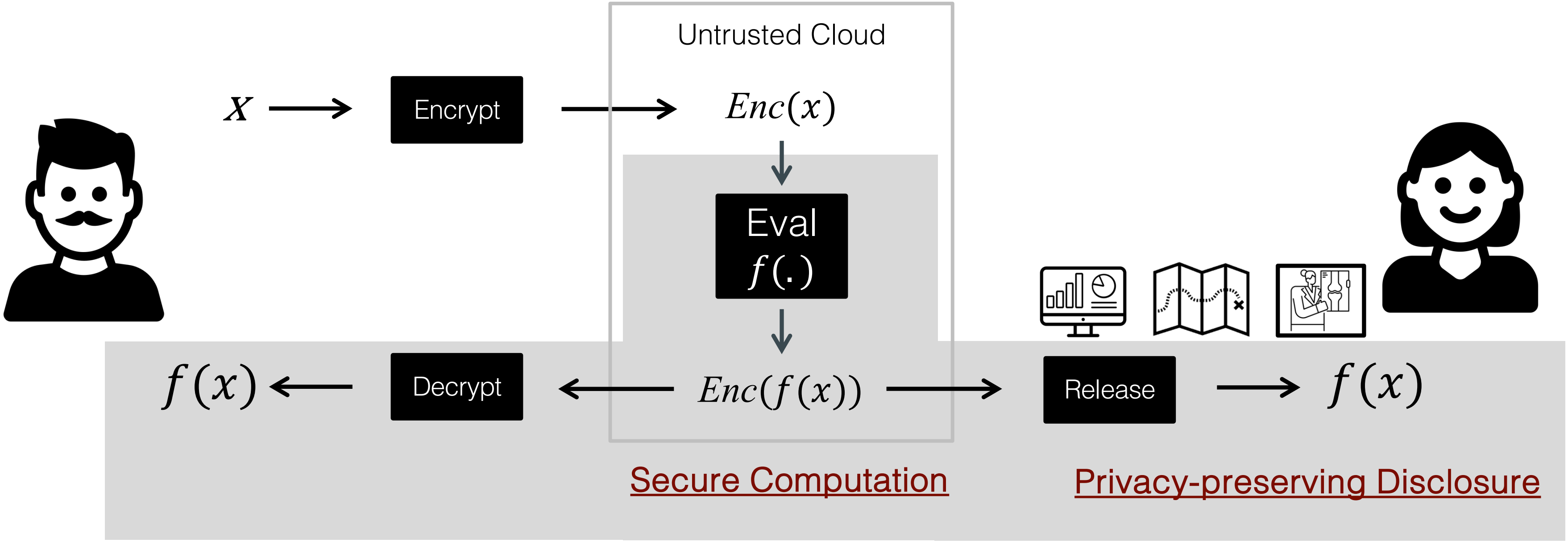
Securing Data in Use: Modern Applications



Securing Data in Use: Modern Applications



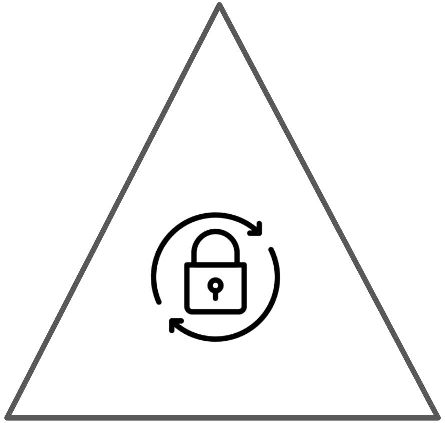
Securing Data in Use: Modern Applications



Secure Computation

Privacy-preserving Disclosure

End-to-End Security

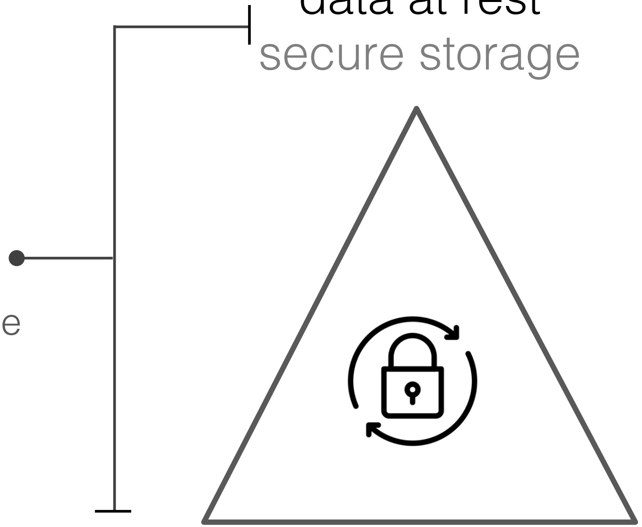


data at rest
secure storage

data in transit
secure communication

data in use
secure computation

End-to-End Security



data at rest

secure storage

Ubiquitous Adoption

Conventional Crypto

Encryption & Digital Signature

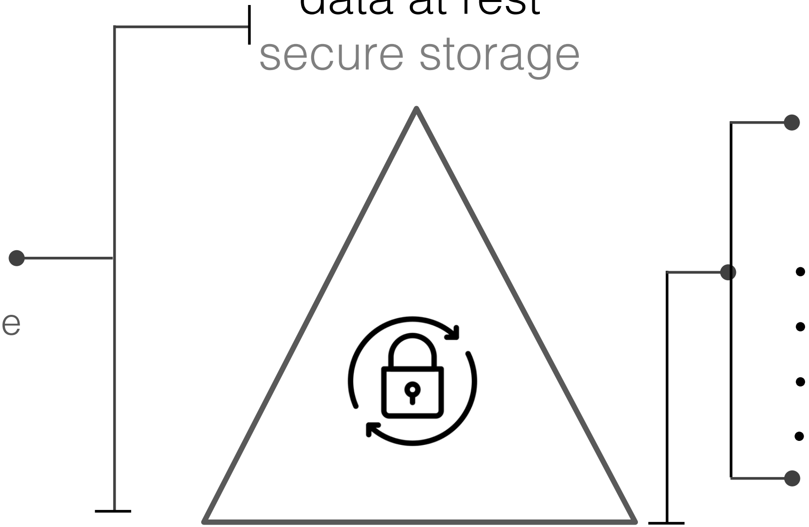
data in transit

secure communication

data in use

secure computation

End-to-End Security



data at rest

secure storage

Ubiquitous Adoption

Conventional Crypto

Encryption & Digital Signature

Just Starting

Privacy - Enhancing Technologies (PETs)

- Homomorphic Encryption
- Secure Multi-party Computation
- Zero Knowledge Proofs
- Differential Privacy

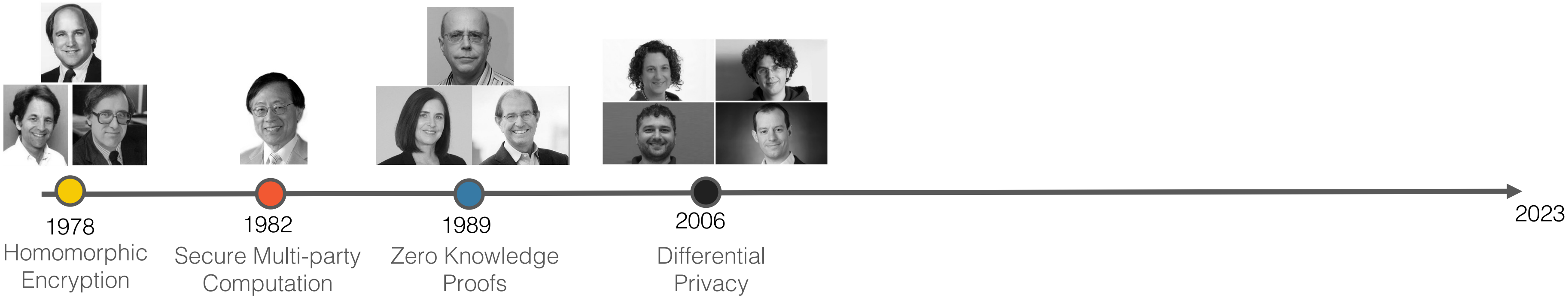
data in transit

secure communication

data in use

secure computation

~ 40 Years of History



1978
Homomorphic Encryption

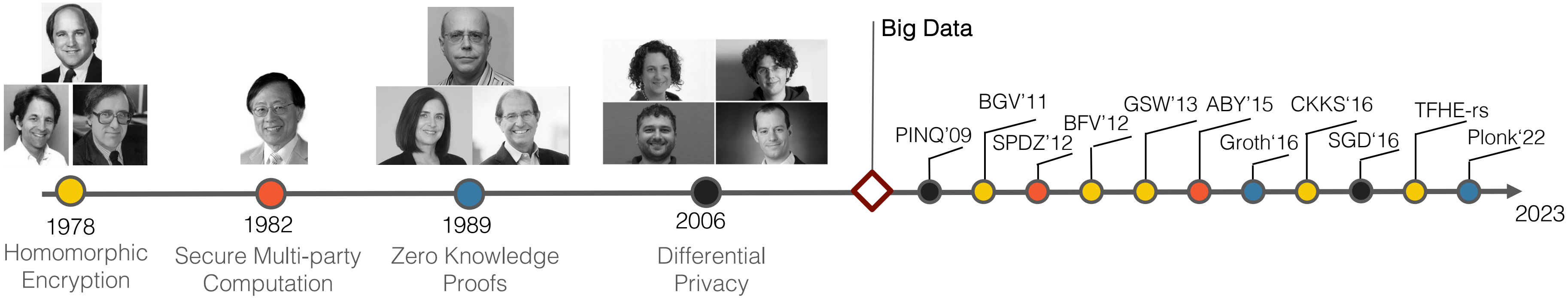
1982
Secure Multi-party Computation

1989
Zero Knowledge Proofs

2006
Differential Privacy

2023

~ 40 Years of History



1978
Homomorphic
Encryption

1982
Secure Multi-party
Computation

1989
Zero Knowledge
Proofs

2006
Differential
Privacy

Big Data

PINQ’09

BGV’11

SPDZ’12

BFV’12

GSW’13

ABY’15

Groth‘16

CKKS‘16

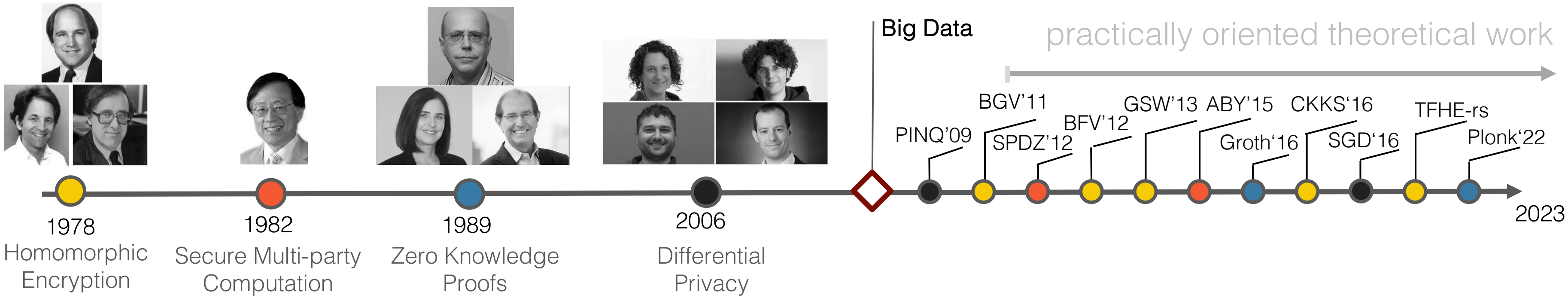
SGD‘16

TFHE-rs

Plonk‘22

2023

~ 40 Years of History



1978
Homomorphic
Encryption

1982
Secure Multi-party
Computation

1989
Zero Knowledge
Proofs

2006
Differential
Privacy

Big Data

practically oriented theoretical work

PINQ’09

BGV’11

SPDZ’12

BFV’12

GSW’13

ABY’15

Groth‘16

CKKS‘16

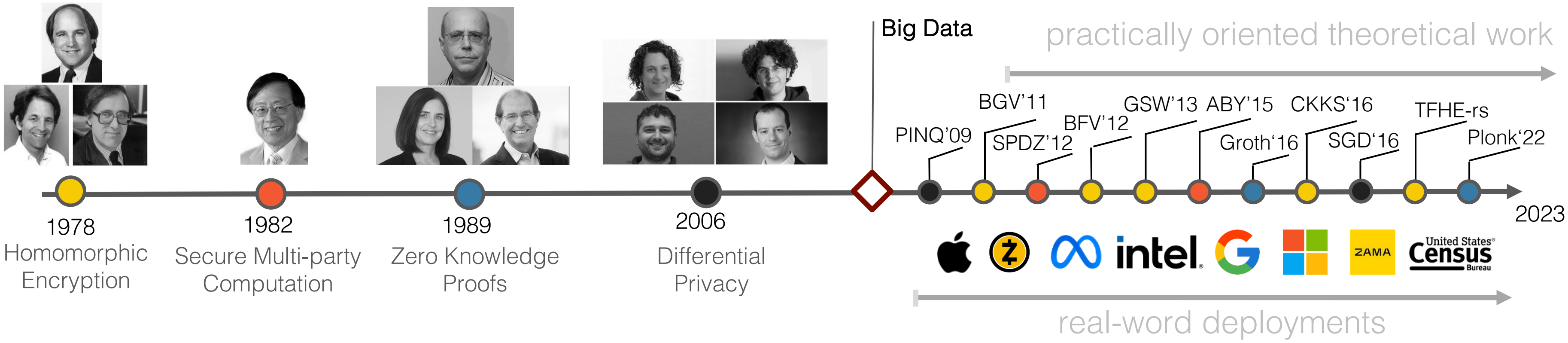
SGD‘16

TFHE-rs

Plonk‘22

2023

~ 40 Years of History



1978
Homomorphic Encryption

1982
Secure Multi-party Computation

1989
Zero Knowledge Proofs

2006
Differential Privacy

Big Data

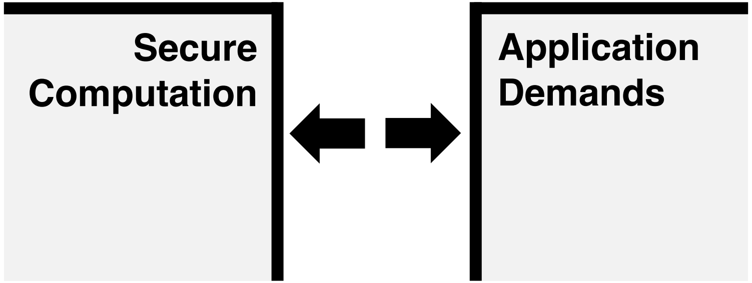
practically oriented theoretical work

PINQ'09
BGV'11
SPDZ'12
BFV'12
GSW'13
ABY'15
Groth'16
CKKS'16
SGD'16
TFHE-rs
Plonk'22

2023

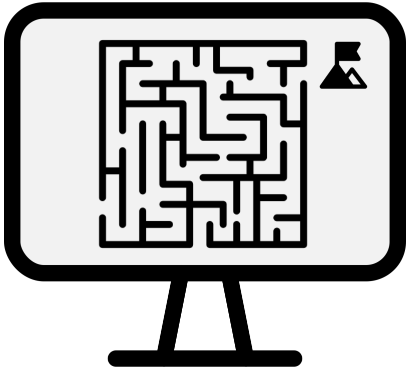
real-word deployments

Theory to Practice: Barriers to Broad Adoption



Secure Computation

Application Demands



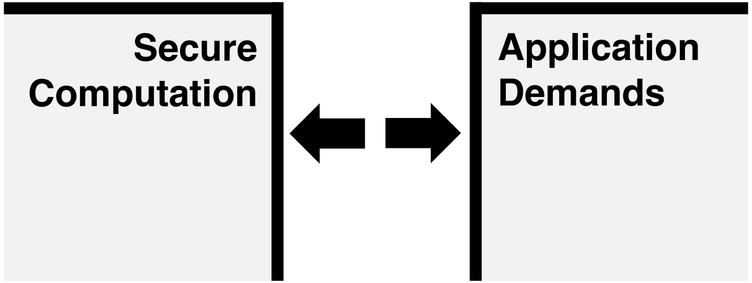
Performance Gap

Practical for numerous applications but remains beyond reach for constrained use cases.

Complexity

There's a gap between the capabilities of PETs today and organizations' ability to incorporate them into applications.

Theory to Practice: Barriers to Broad Adoption

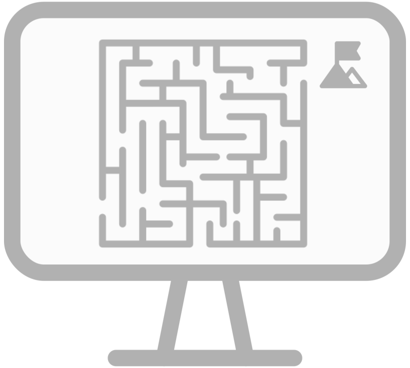


Secure Computation

Application Demands

Performance Gap

Practical for numerous applications but remains beyond reach for constrained use cases.

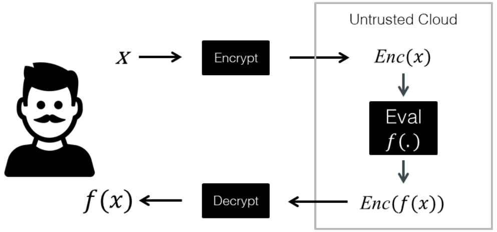


Complexity

There's a gap between the capabilities of PETs today and organizations' ability to incorporate them into applications.

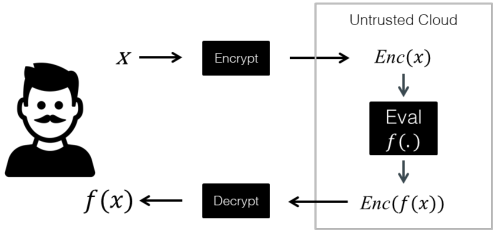
Performance Gap

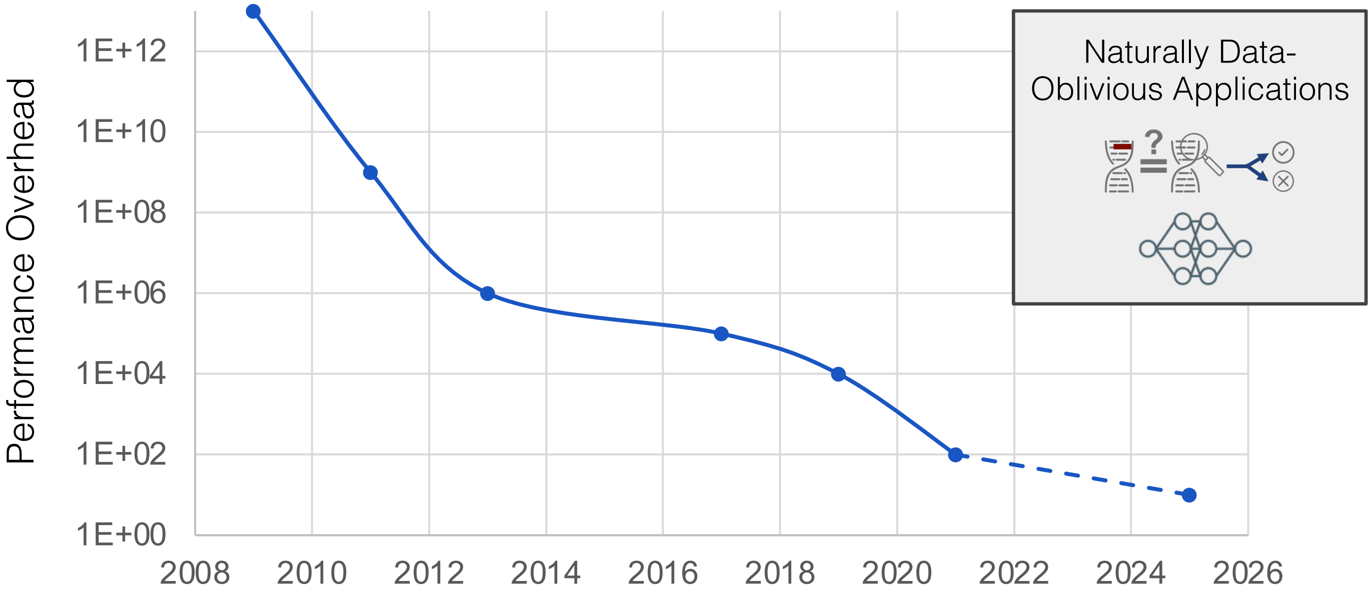
Fully Homomorphic Encryption



Performance Gap

Fully Homomorphic Encryption



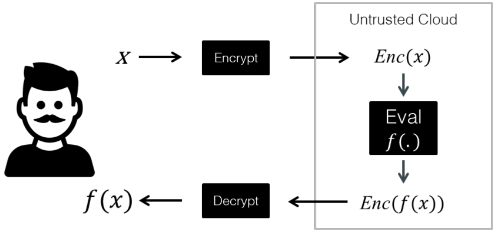


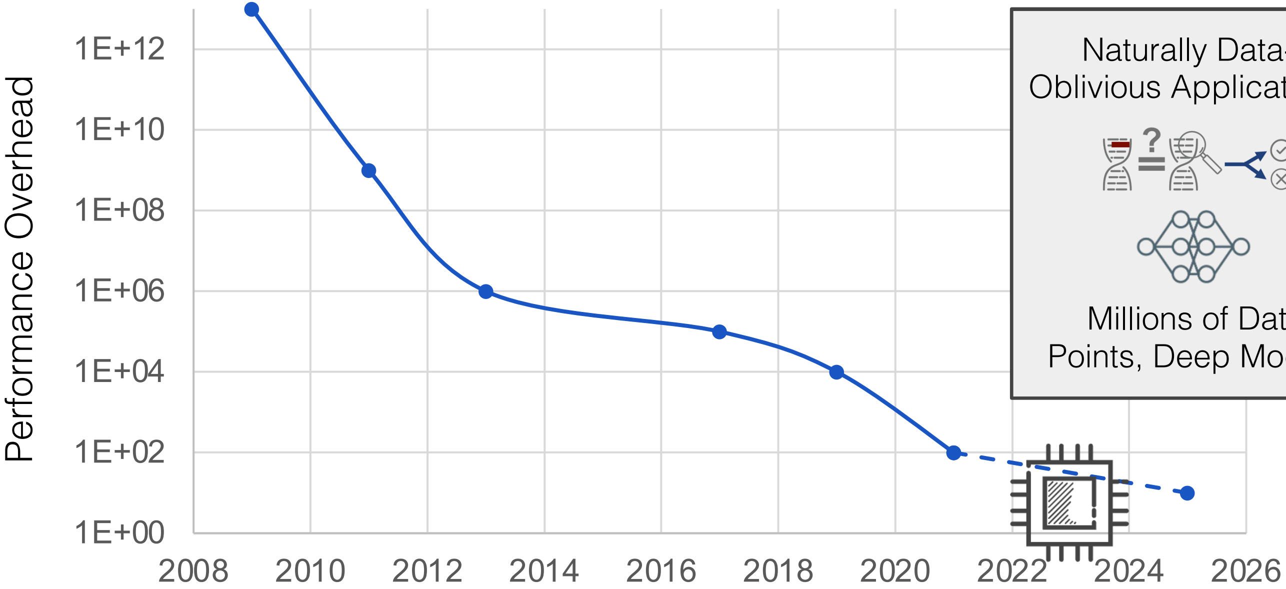
Naturally Data-Oblivious Applications

Graph Adapted from Kristin Lauter, Talk @PriCon2020

Performance Gap

Fully Homomorphic Encryption





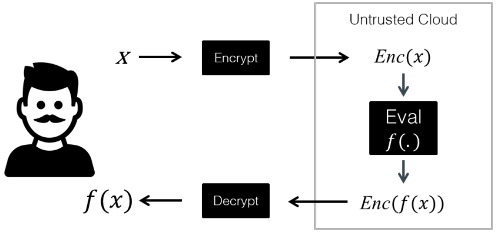
Naturally Data-Oblivious Applications

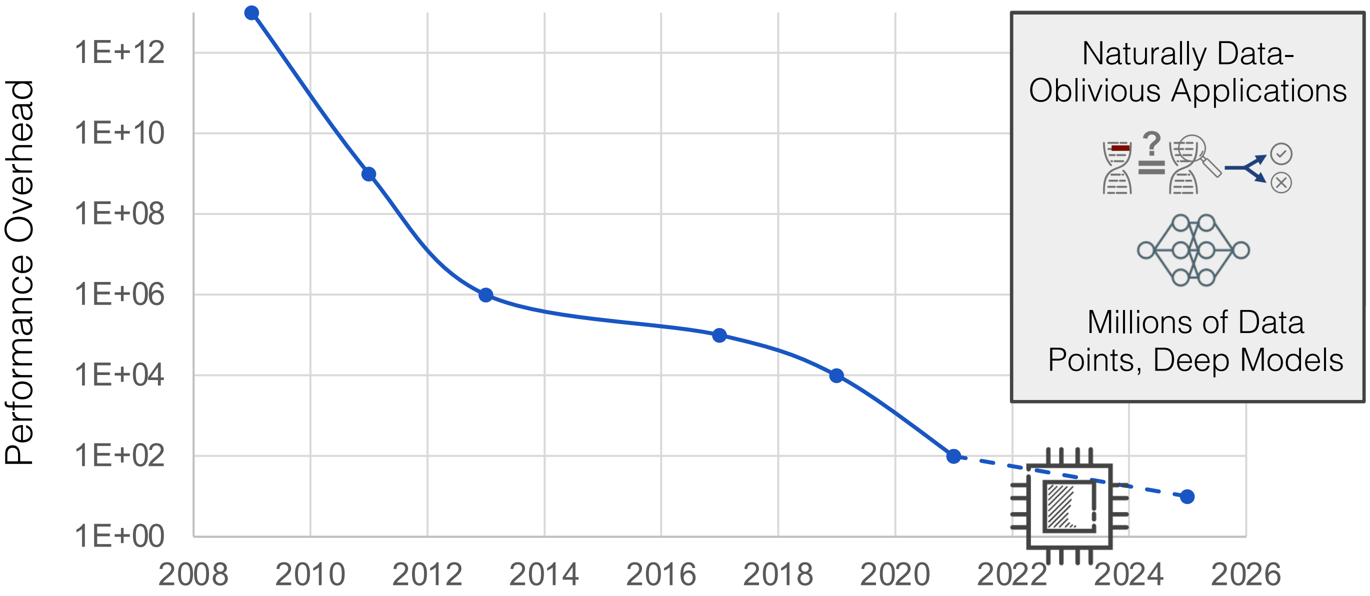
Millions of Data Points, Deep Models

Graph Adapted from Kristin Lauter, Talk @PriCon2020

Performance Gap

Fully Homomorphic Encryption





Naturally Data-Oblivious Applications

Millions of Data Points, Deep Models

Highly Interactive Applications

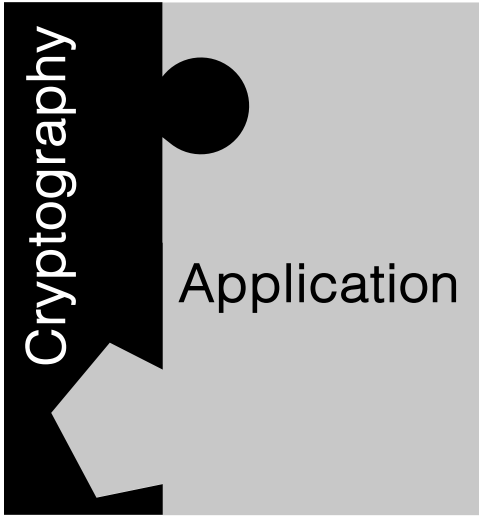
Constrained Environments

Graph Adapted from Kristin Lauter, Talk @PriCon2020

Approach to Efficiency

Empower Constrained Environments with Encrypted Data Processing.

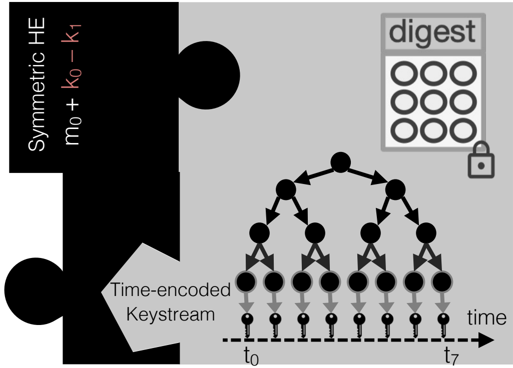
co-design



Encrypted Data Stream Processing at Scale

[Constrained Data Sources, Large Scale, Low-Latency]

[TimeCrypt - USENIX NSDI’20]

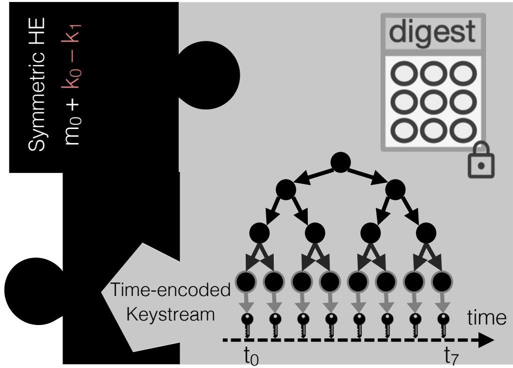


co-design

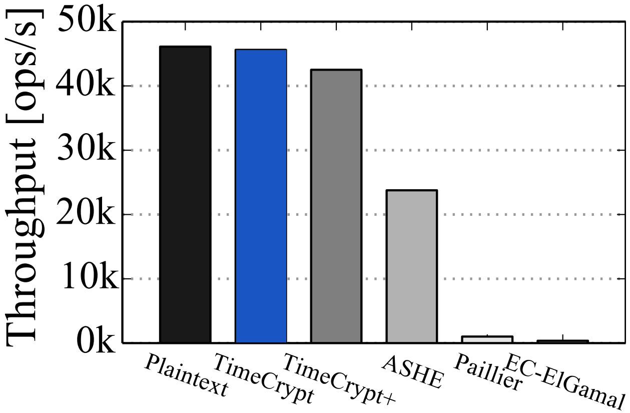
Encrypted Data Stream Processing at Scale

[Constrained Data Sources, Large Scale, Low-Latency]

[TimeCrypt - USENIX NSDI'20]



co-design

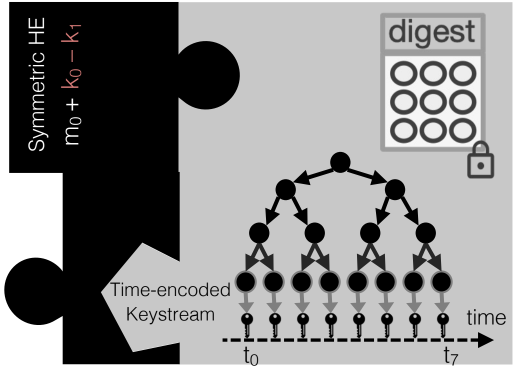


System Performance

Encrypted Data Stream Processing at Scale

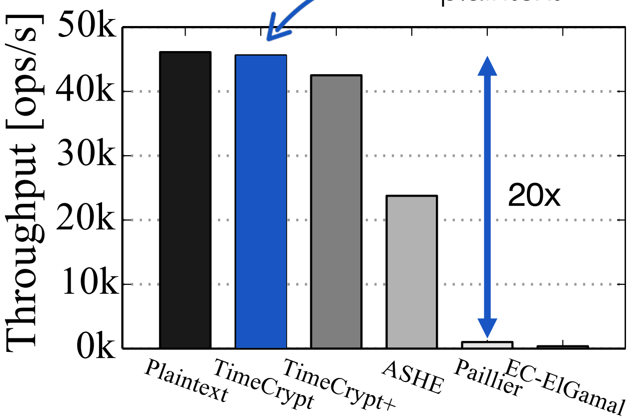
[Constrained Data Sources, Large Scale, Low-Latency]

[TimeCrypt - USENIX NSDI’20]



co-design

2% slowdown compared to plaintext

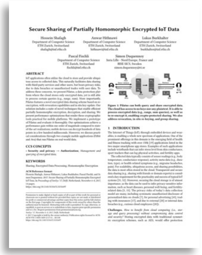


System Performance

Privacy-preserving, functional, and performant systems

My work aims to **build** practical systems that use
cryptography to empower users and preserve their privacy.

Talos: Encrypted Query Processing for the Internet of Things



Secure Sharing of Partially Homomorphic Encrypted IoT Data

TimeCrypt: Encrypted Data Stream Processing at Scale with Cryptographic Access Control

Droplet: Decentralized Authorization and Access Control for Encrypted Data Streams

Zeph: Cryptographic Enforcement of End-to-End Data Privacy

VF-PS: How to Select Important Participants in Vertical Federated Learning, Efficiently and Securely?

RoFL: Attestable Robustness for Secure Federated Learning

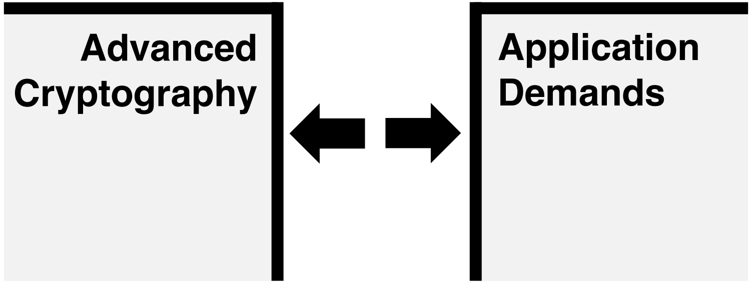
Talos	Pilatus	TimeCrypt	Droplet	Zeph	VF-PS	RoFL
ACM SenSys	ACM SenSys	USENIX NSDI	USENIX Security	USENIX OSDI	NeurIPS	IEEE S&P

Internet of Things

Streaming

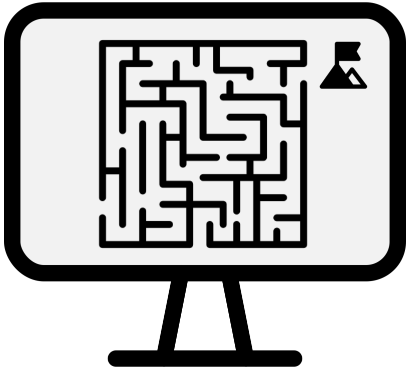
Collaborative ML

Theory to Practice: Barriers to Broad Adoption

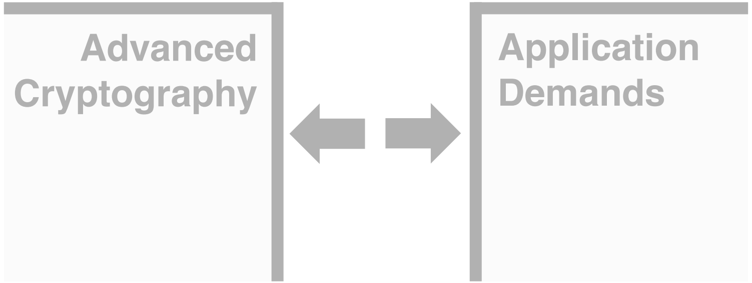

Advanced Cryptography

Application Demands

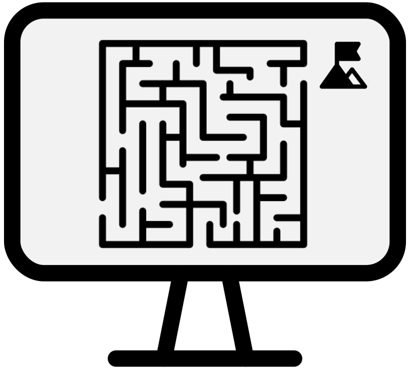
Performance Gap


Complexity

Theory to Practice: Barriers to Broad Adoption



Performance Gap



Complexity

Democratize Privacy-Preserving Computation

My work aims to **democratize** access to privacy-preserving computation with new tools, systems, and abstractions.

Secure Computation

SoK: Fully Homomorphic Encryption Compilers

Alexander Viand, ETH Zurich, alexander.viand@inf.ethz.ch
Patrick Jattke, ETH Zurich, pjattke@ethz.ch
Anwar Hithnawi, ETH Zurich, anwar.hithnawi@inf.ethz.ch

FHE Compilers

IEEE S&P

HECO: Fully Homomorphic Encryption Compiler

Alexander Viand, Patrick Jattke, Miro Haller, Anwar Hithnawi

ETH Zurich

HECO

USENIX Security

Differential Privacy

Cohere: Managing Differential Privacy in Large Scale Systems

Nicolas Küchler[1], Emanuel Opel[1], Hidde Lycklama[1], Alexander Viand[1], Anwar Hithnawi[1]

[1]ETH Zurich [2]Intel Labs

Cohere

IEEE S&P

Programmability

Deployments

Democratize Privacy-Preserving Computation

My work aims to **democratize** access to privacy-preserving computation with new tools, systems, and abstractions.

Secure Computation

SoK: Fully Homomorphic Encryption Compilers

Alexander Viand, ETH Zurich, alexander.viand@inf.ethz.ch
Patrick Jattke, ETH Zurich, pjattke@ethz.ch
Anwar Hithnawi, ETH Zurich, anwar.hithnawi@inf.ethz.ch

FHE Compilers

IEEE S&P

HECO: Fully Homomorphic Encryption Compiler

Alexander Viand, Patrick Jattke, Miro Haller, Anwar Hithnawi

ETH Zurich

HECO

USENIX Security

Differential Privacy

Cohere: Managing Differential Privacy in Large Scale Systems

Nicolas Küchler[1], Emanuel Opel[1], Hidde Lycklama[1], Alexander Viand[2], Anwar Hithnawi[1]

[1]ETH Zurich [2]Intel Labs

Cohere

IEEE S&P

Programmability

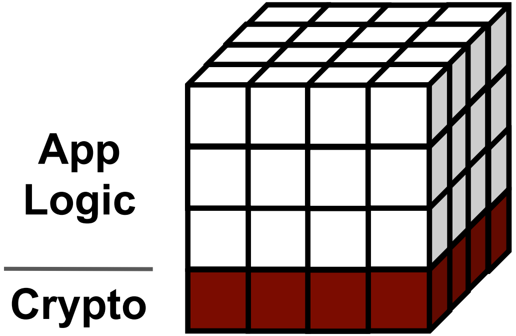
Deployments

Developing and Deploying Privacy-preserving Applications is

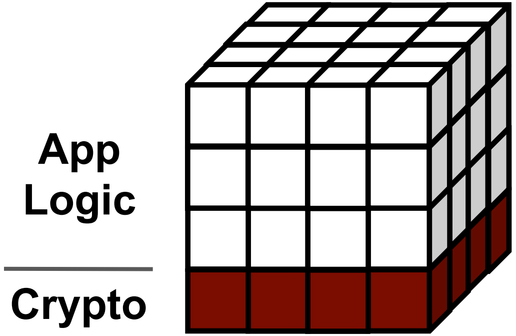
Notoriously Hard

What does “developing these applications” entail?

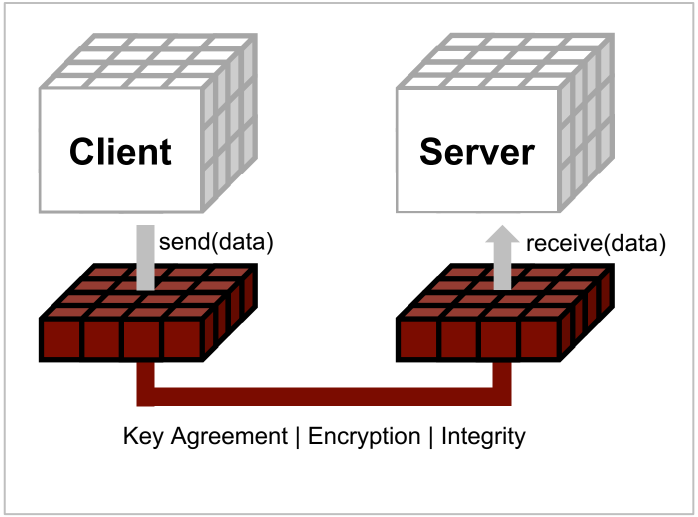
Conventional Cryptography



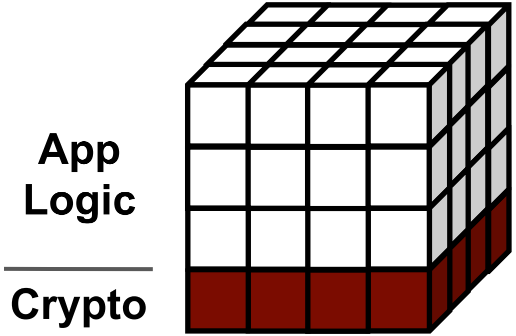
Conventional Cryptography



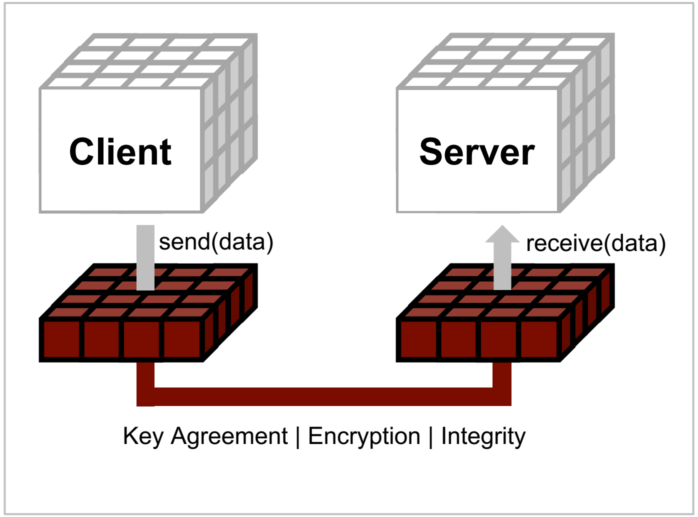
Secure Communication



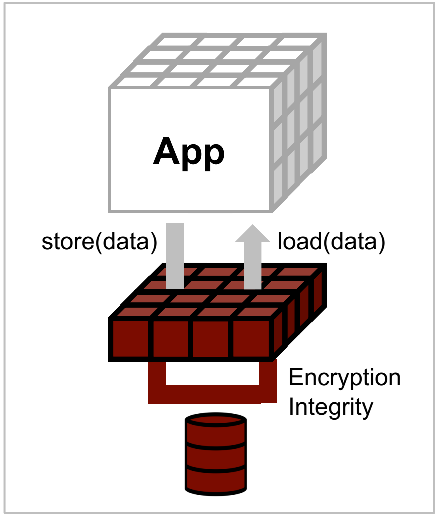
Conventional Cryptography



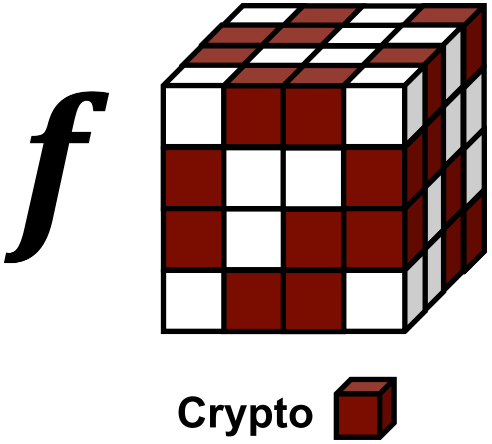
Secure Communication

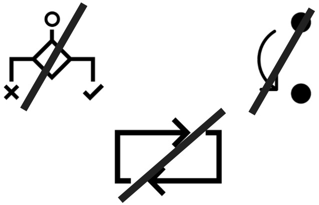


Secure Storage

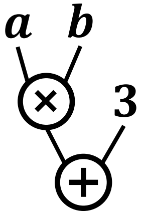


Advanced Cryptography: Secure Computation

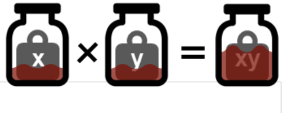




Data Oblivious

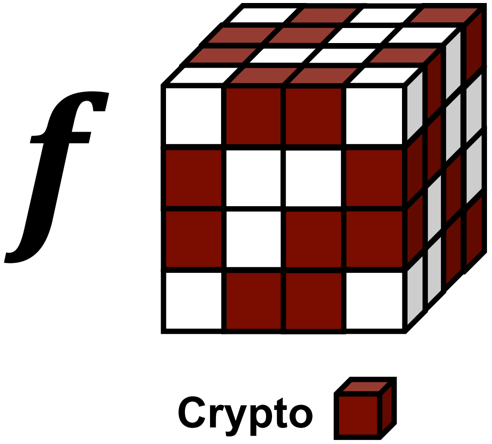


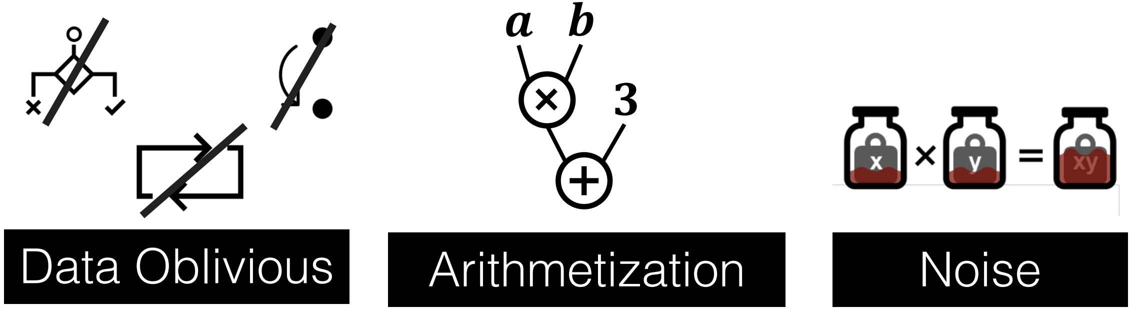
Arithmetization



Noise

Advanced Cryptography: Secure Computation





Data Oblivious

Arithmetization

Noise

Functionality and performance depend on f's representation:

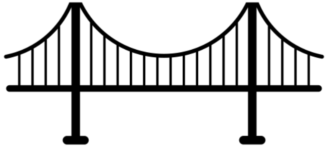
- How do we express f
- How do we optimize f

Usable Fully Homomorphic Encryption

(IEEE S&P’21, USENIX Security‘23)

Usable FHE

Advanced Cryptography



Programming Languages

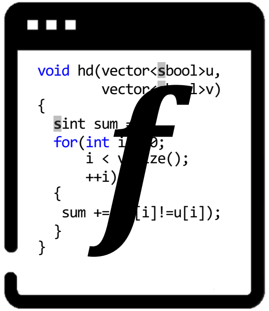
1 What makes developing FHE applications hard?

[IEEE S&P'21]

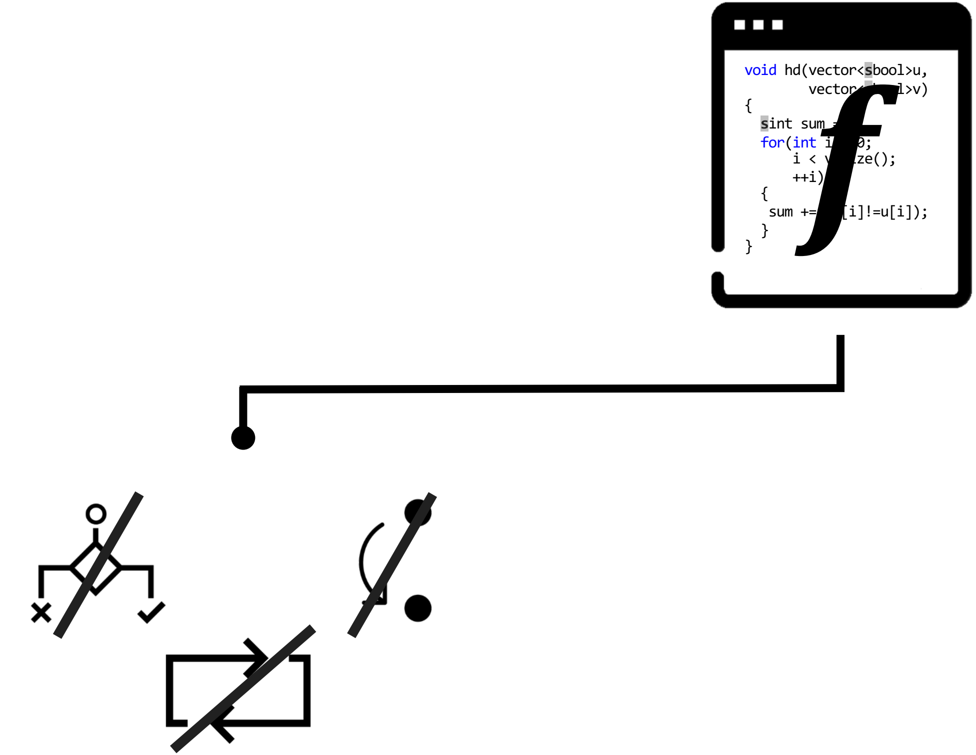
2 How can compilers address these complexities?

[USENIX Security'23]

Fully Homomorphic Encryption Programming Paradigm

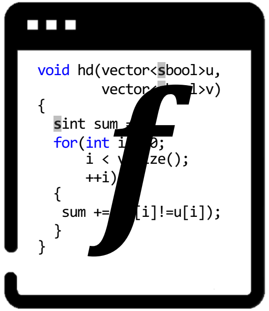


```
void hd(vector<sbool>u,
        vector<   l>v)
{
  sint sum =
  for(int i   0;
      i < v  ze();
      ++i)
  {
   sum +=   [i]!=u[i]);
  }
}
```

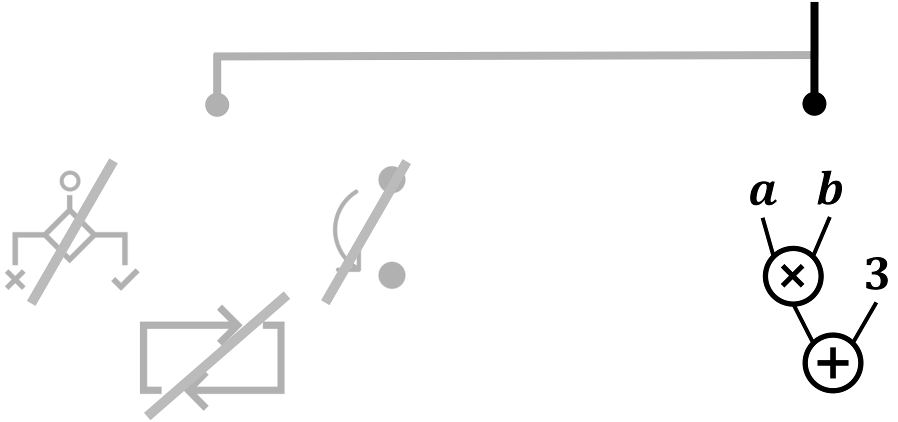


```
void hd(vector<sbool>u,
           vector<    l>v)
{
  sint sum =
  for(int i   0;
       i < v   ze();
       ++i)
  {
   sum +=    [i]!=u[i]);
  }
}
```

Data Oblivious

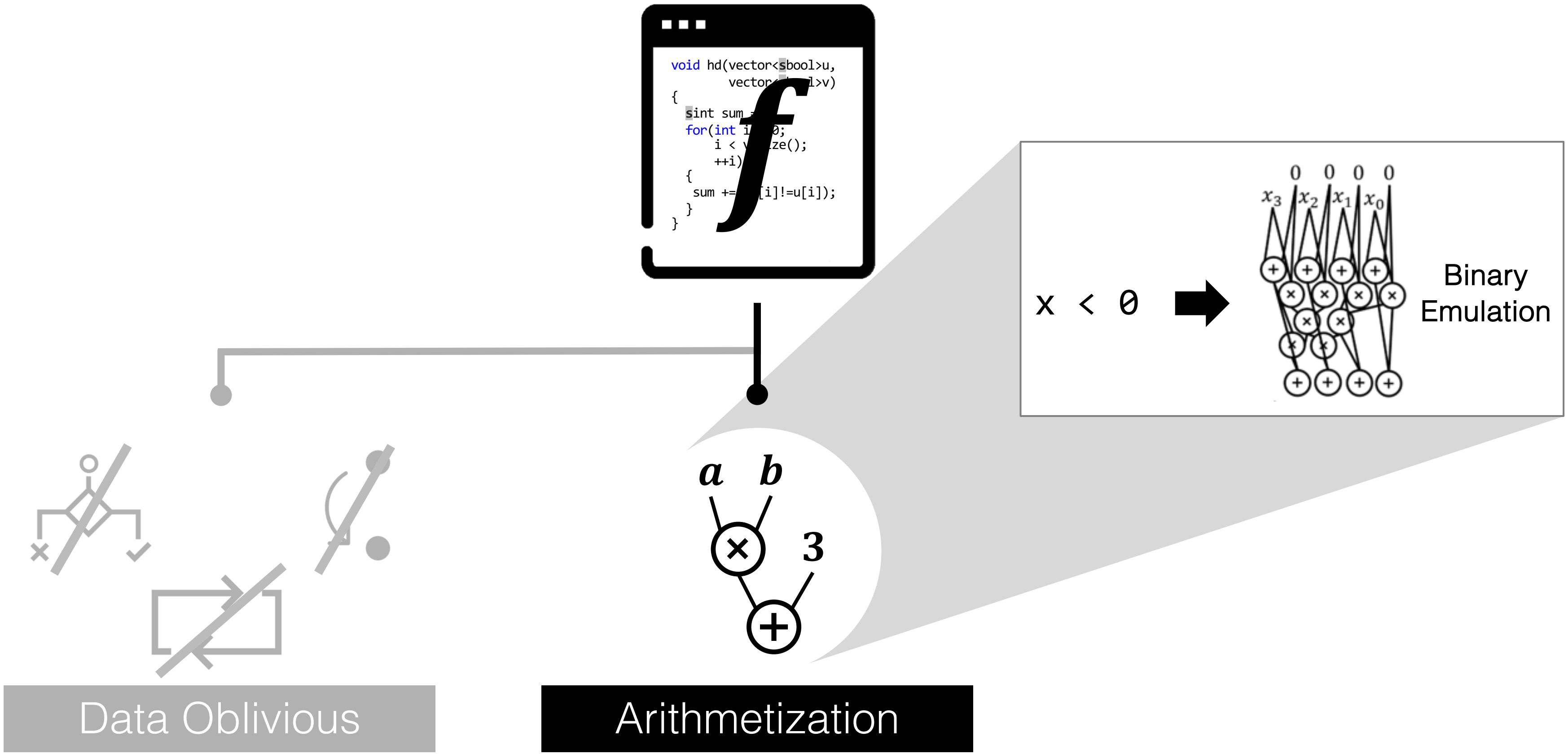


```
void hd(vector<sbool>u,
        vector<sbool>v)
{
  sint sum = 0;
  for(int i = 0;
      i < v.size();
      ++i)
  {
   sum += (v[i]!=u[i]);
  }
}
```



Data Oblivious

Arithmetization



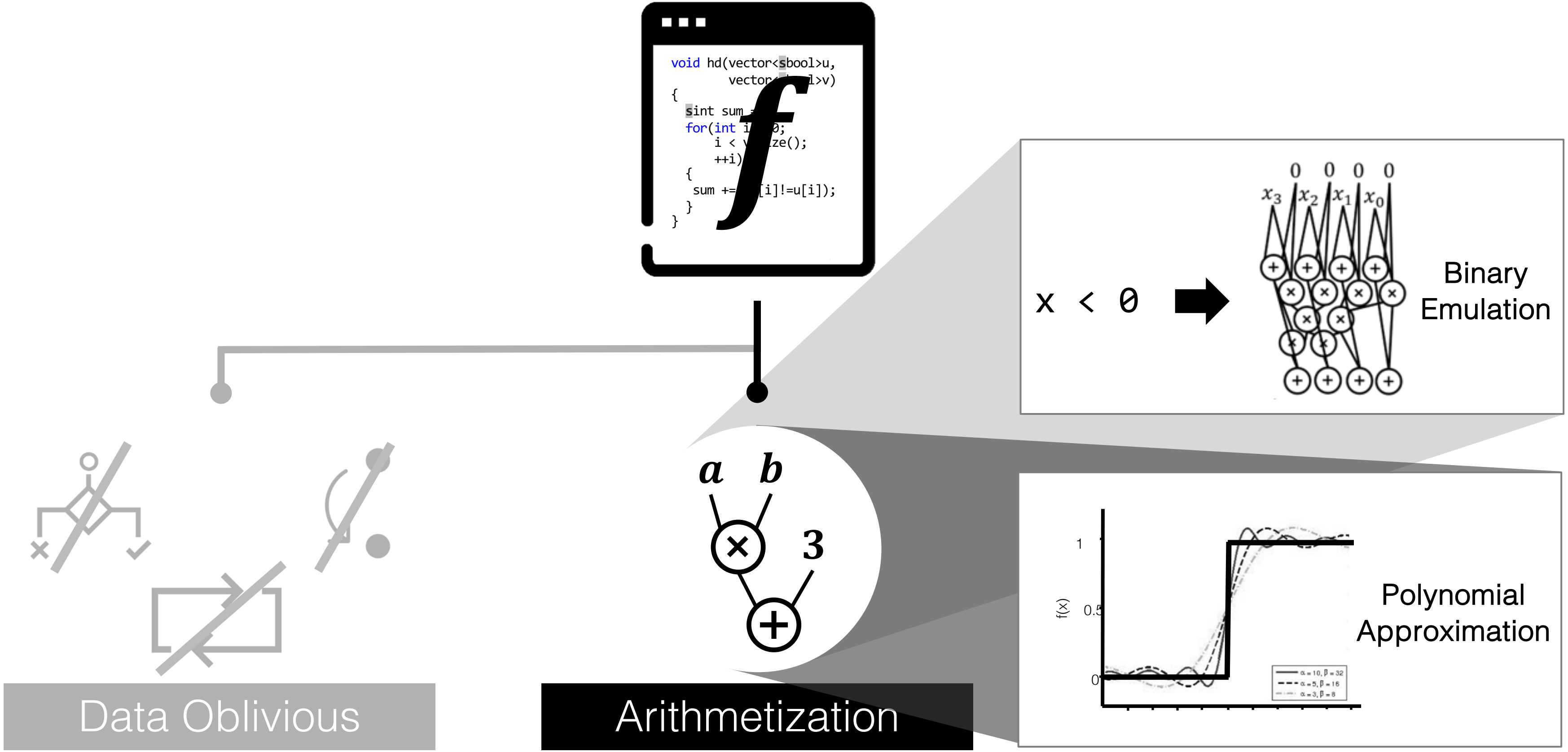
```
void hd(vector<sbool>u,
        vector<    l>v)
{
  sint sum =
  for(int i   0;
      i < v   ze();
      ++i)
  {
   sum +=   [i]!=u[i]);
  }
}
```

x < 0

Binary Emulation

Data Oblivious

Arithmetization



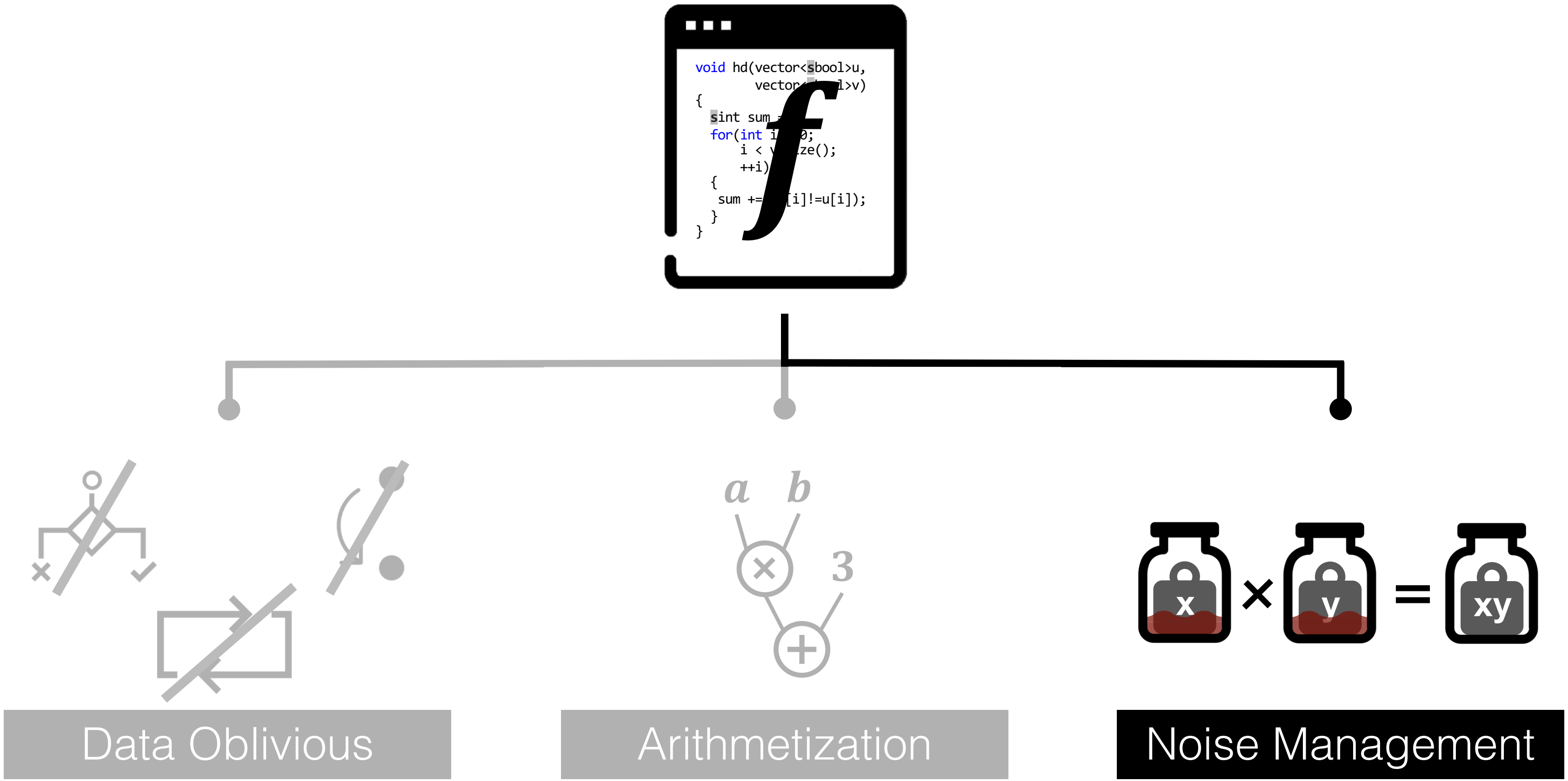
```
void hd(vector<sbool>u,
        vector<    l>v)
{
  sint sum
  for(int i    0;
      i < v    ze();
      ++i)
  {
   sum +=    [i]!=u[i]);
  }
}
```

Binary Emulation

Polynomial Approximation

Data Oblivious

Arithmetization

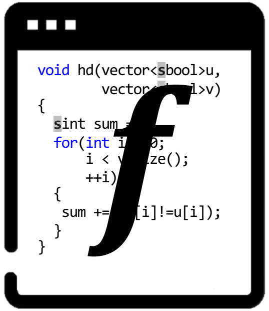


```
void hd(vector<sbool>u,
        vector<    l>v)
{
  sint sum =
  for(int i   0;
      i < v   ze();
      ++i)
  {
   sum +=    [i]!=u[i]);
  }
}
```

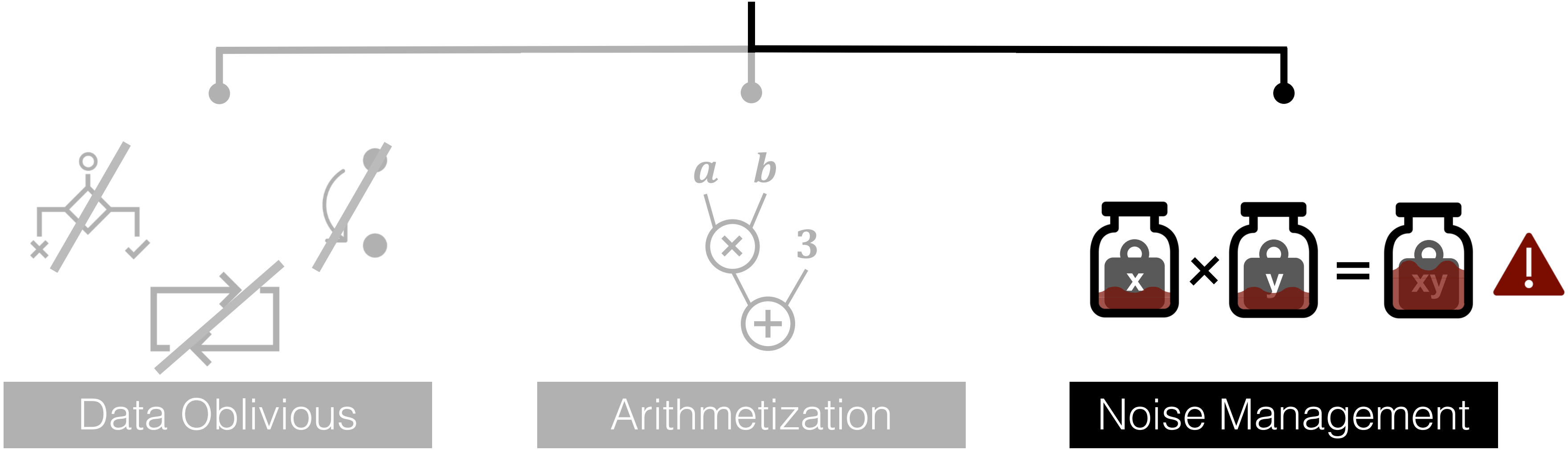
Data Oblivious

Arithmetization

Noise Management



```
void hd(vector<sbool>u,
        vector<     l>v)
{
  sint sum
  for(int i   0;
      i < v   ze();
      ++i)
  {
   sum +=   [i]!=u[i]);
  }
}
```

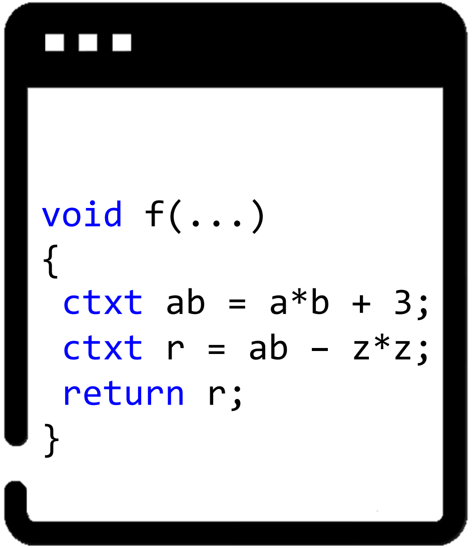


Data Oblivious

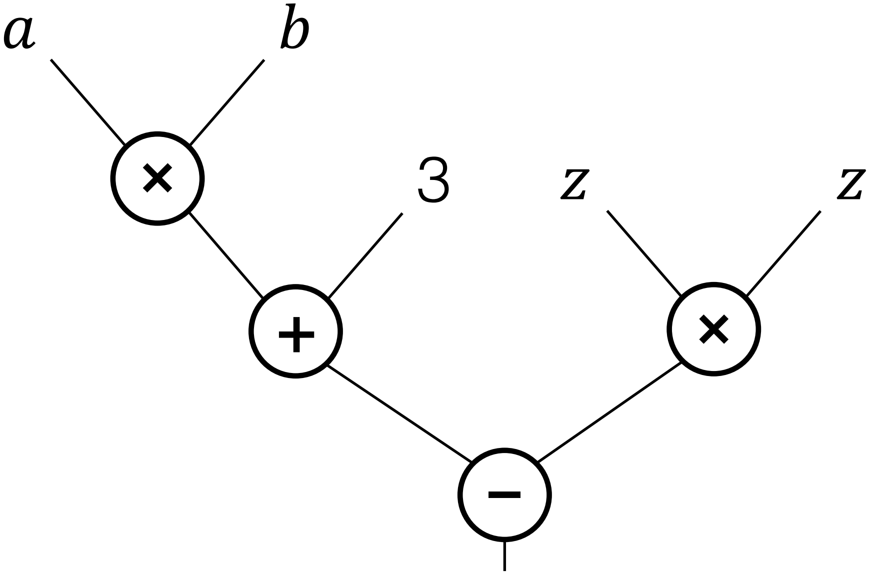
Arithmetization

Noise Management

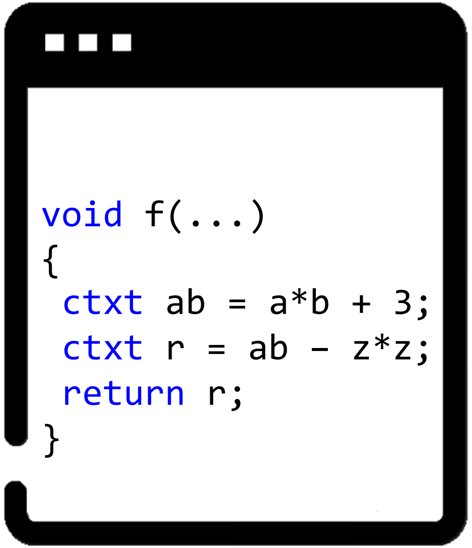
FHE Noise Management



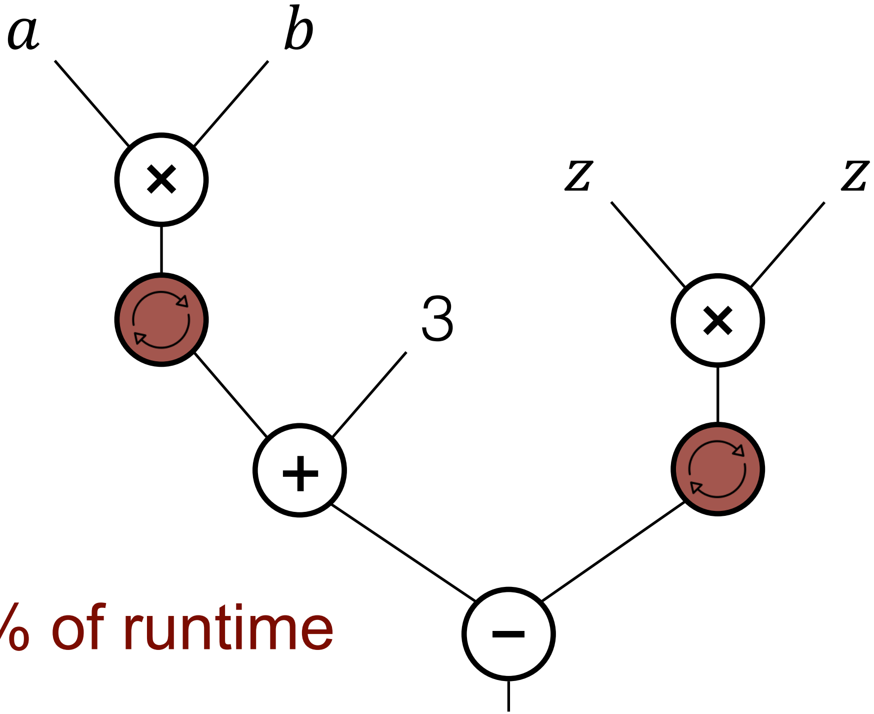
```
void f(...)
{
 ctxt ab = a*b + 3;
 ctxt r = ab – z*z;
 return r;
}
```



FHE Noise Management

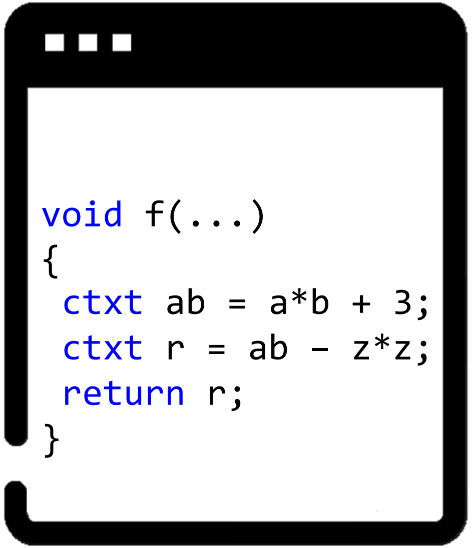


```
void f(...)
{
 ctxt ab = a*b + 3;
 ctxt r = ab - z*z;
 return r;
}
```



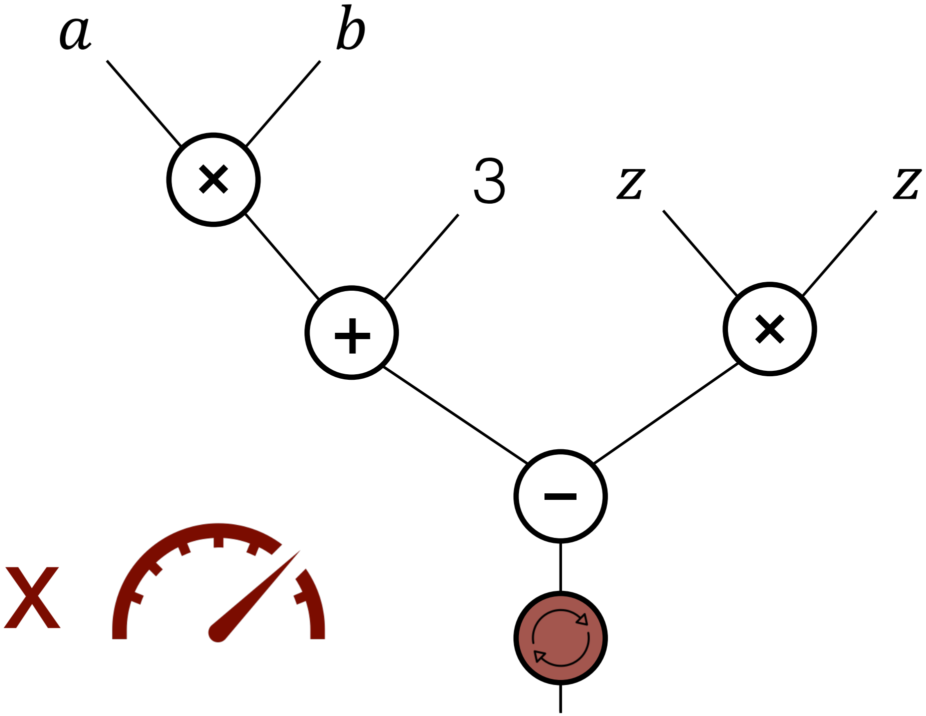
● 90% of runtime

FHE Noise Management

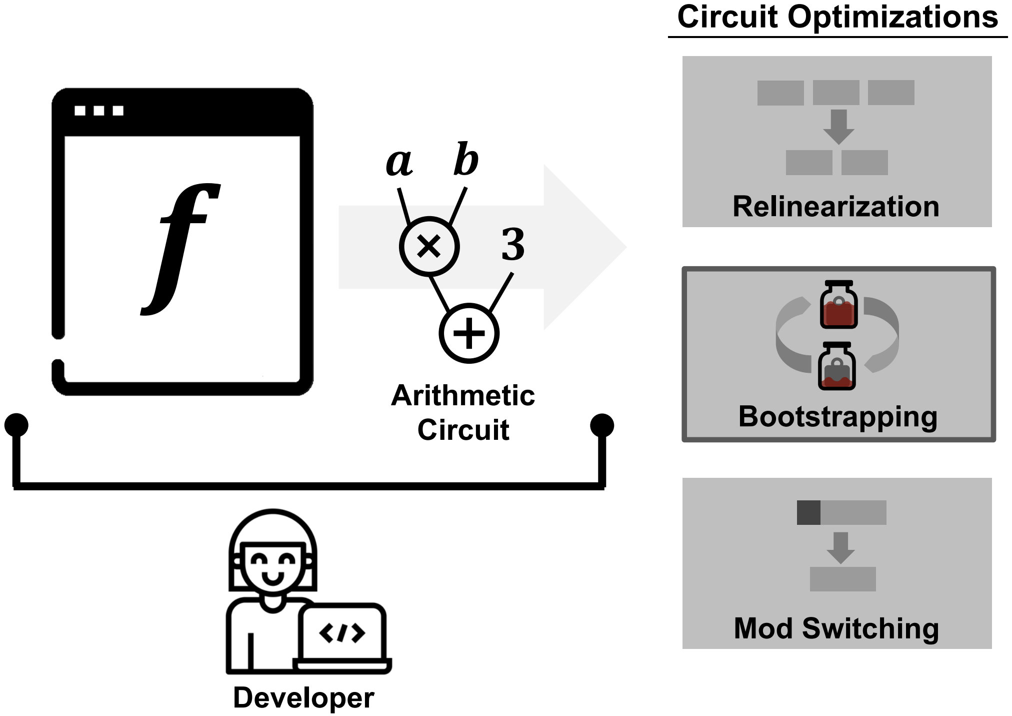


```
void f(...)
{
 ctxt ab = a*b + 3;
 ctxt r = ab – z*z;
 return r;
}
```

2x



Developing FHE Applications



Arithmetic Circuit

Developer

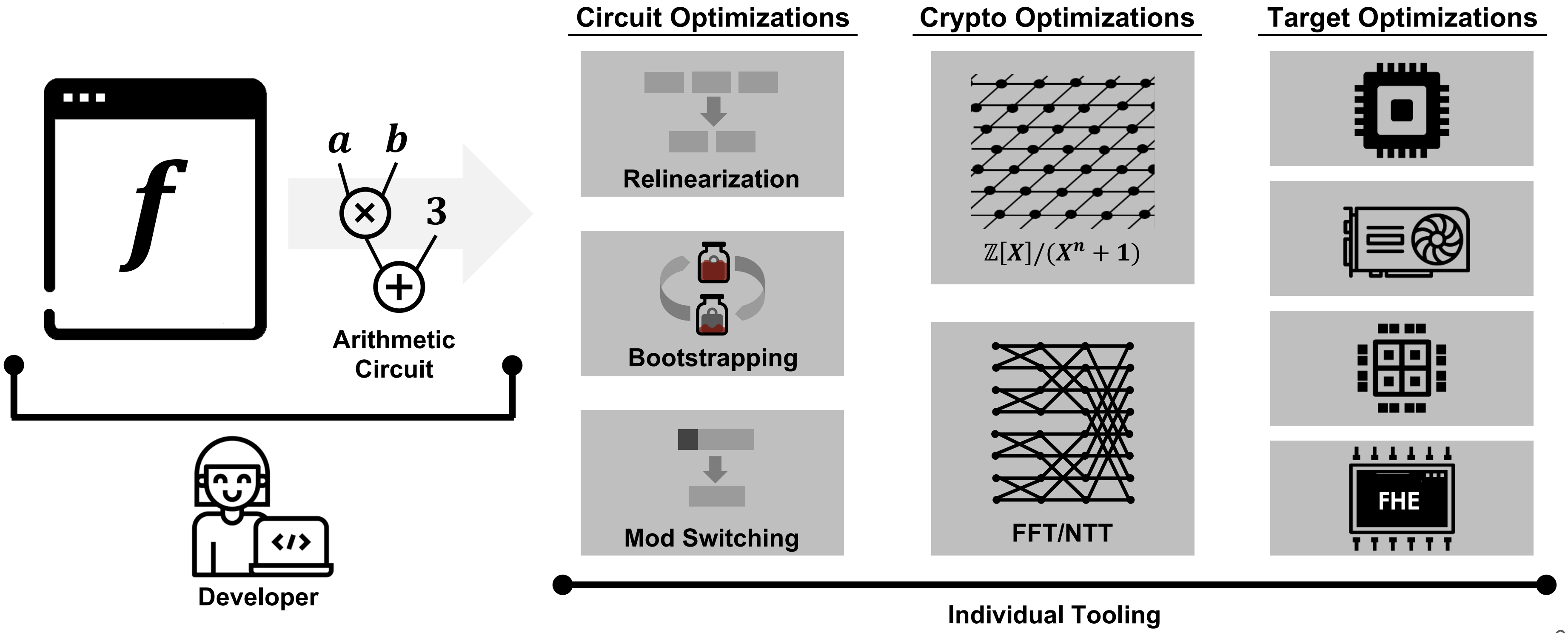
Circuit Optimizations

Relinearization

Bootstrapping

Mod Switching

Developing FHE Applications



f

a b 3

Arithmetic Circuit

Developer

Circuit Optimizations

Relinearization

Bootstrapping

Mod Switching

Crypto Optimizations

$\mathbb{Z}[X]/(X^n + 1)$

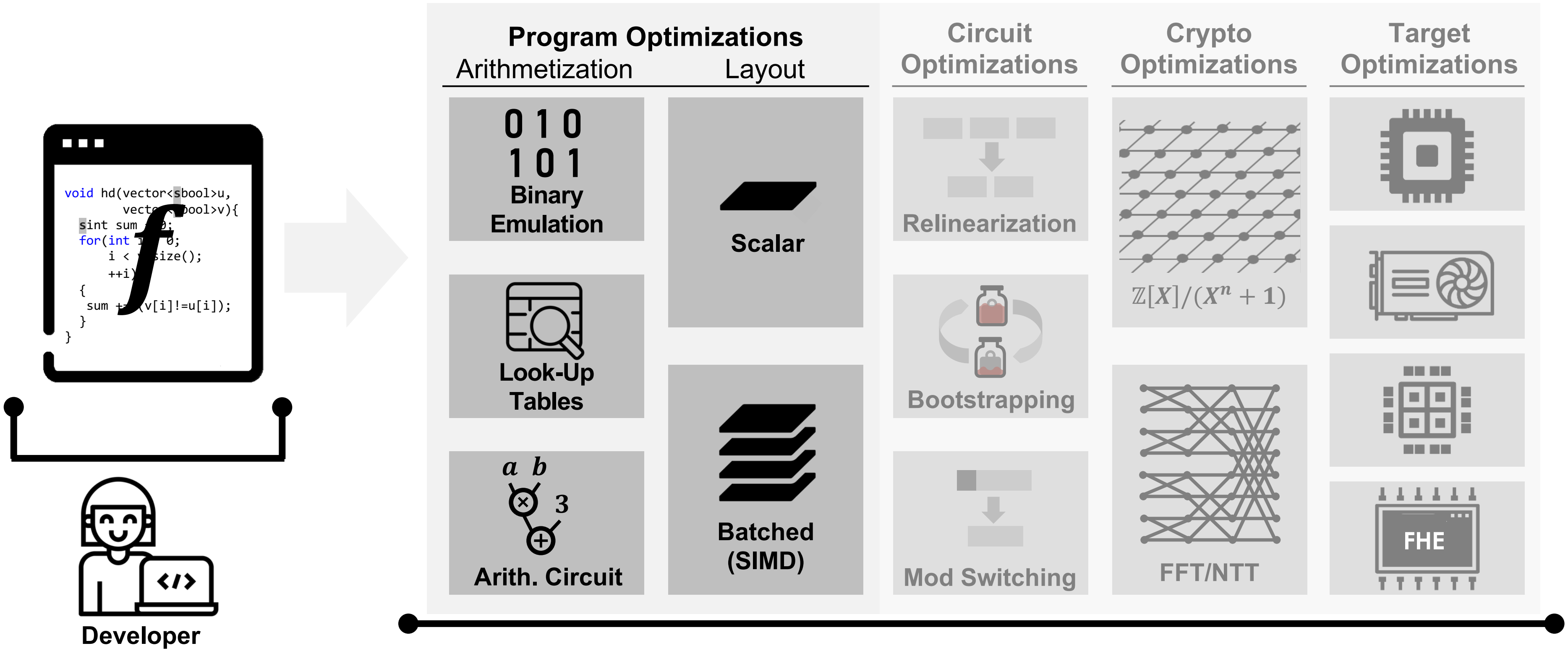
FFT/NTT

Target Optimizations

FHE

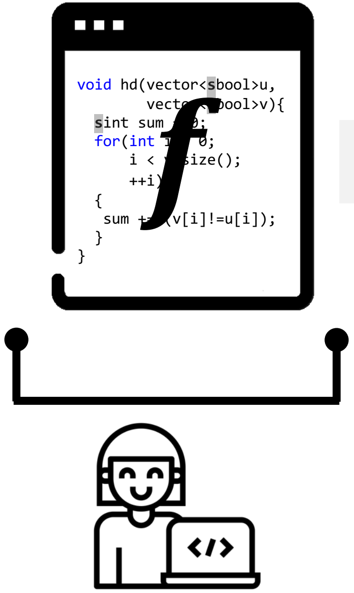
Individual Tooling

HECO



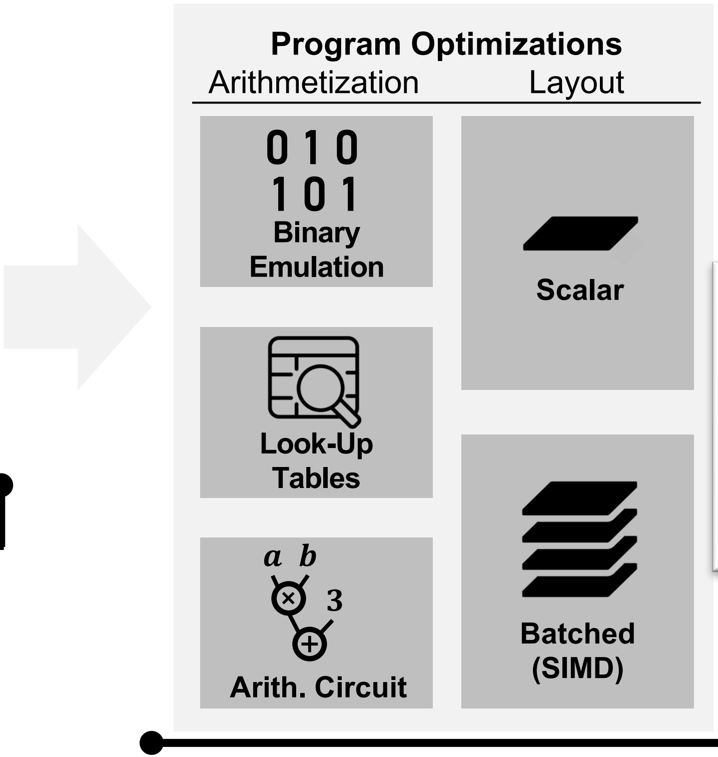
```
void hd(vector<sbool>u,
        vector<sbool>v){
  sint sum = 0;
  for(int i = 0;
      i < v.size();
      ++i)
  {
   sum += (v[i]!=u[i]);
  }
}
```

HECO: Transform High-level Programs to Efficient FHE Solutions



```
void hd(vector<sbool>u,
        vector<sbool>v){
  sint sum = 0;
  for(int i = 0;
      i < v.size();
      ++i)
  {
   sum += (v[i]!=u[i]);
  }
}
```

Developer



Program Optimizations

Circuit Optimizations

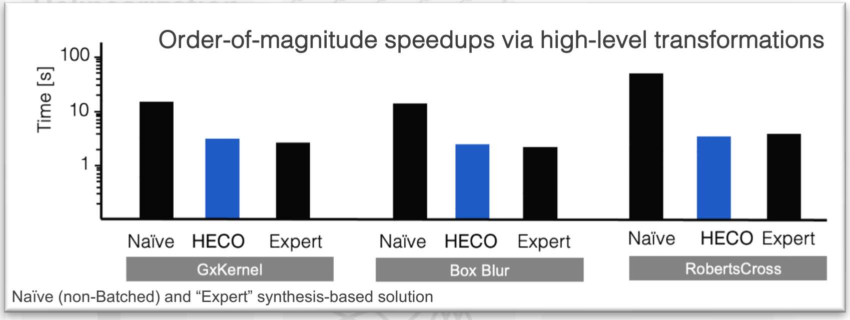
Crypto Optimizations

Target Optimizations

Mod Switching

FFT/NTT

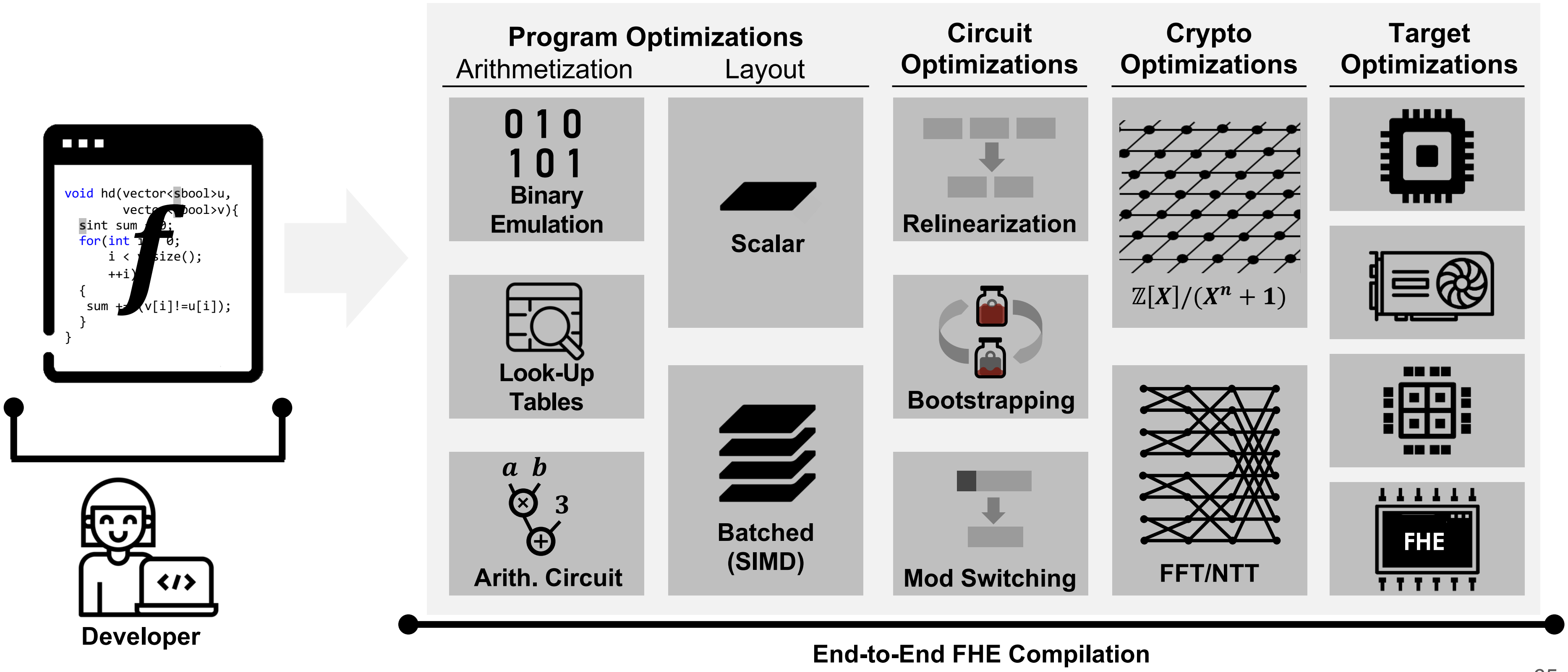
FHE



Order-of-magnitude speedups via high-level transformations

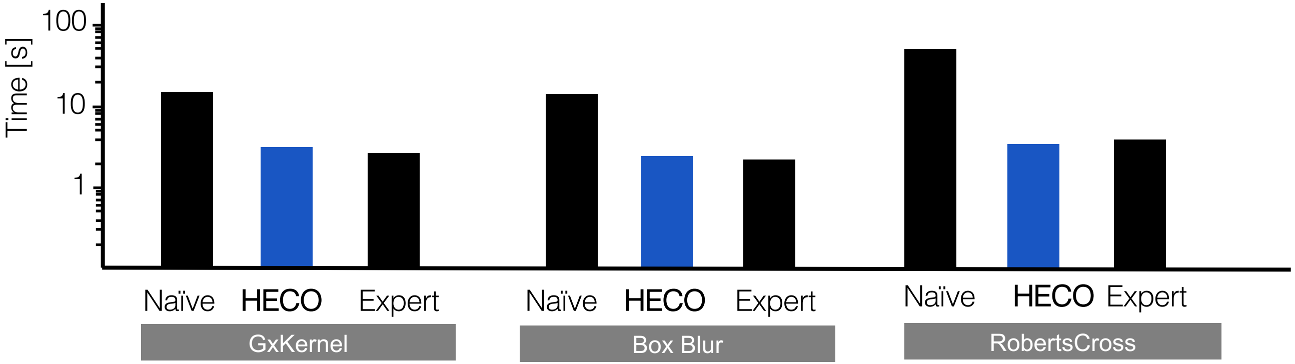
Naïve (non-Batched) and “Expert” synthesis-based solution

HECO: End-to-End FHE Compilation



```
void hd(vector<sbool>u,
        vector<sbool>v){
  sint sum = 0;
  for(int i = 0;
      i < v.size();
      ++i)
  {
   sum += (v[i]!=u[i]);
  }
}
```

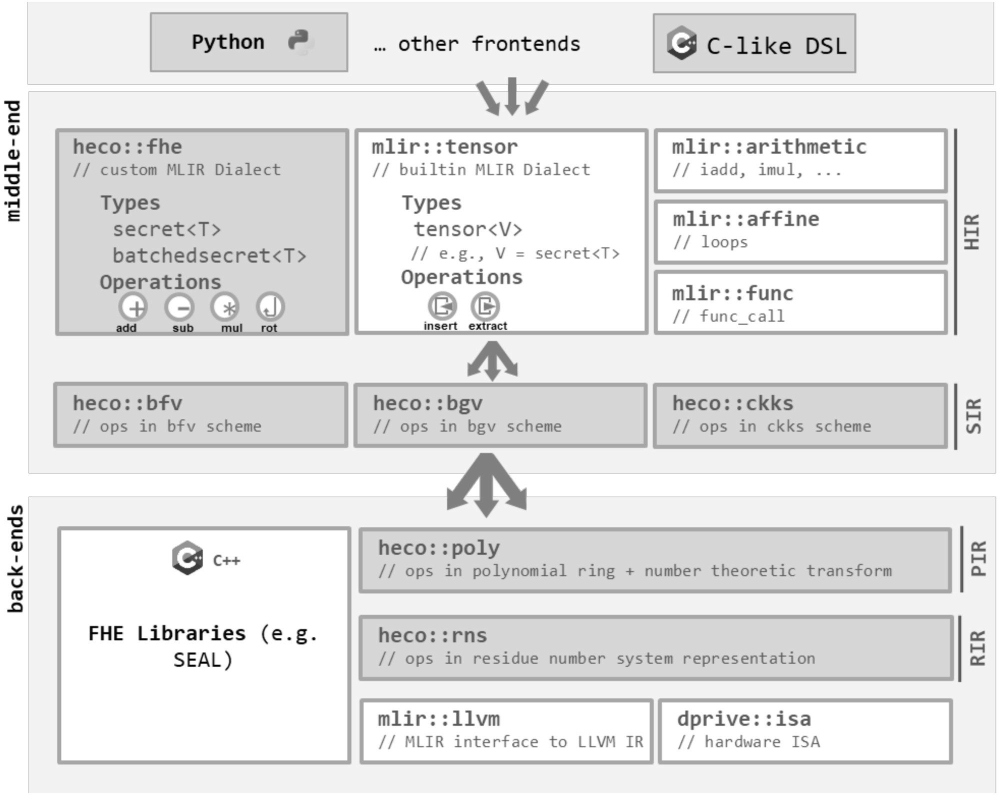
Evaluation: Effect of Batching Opimizations



[CD+21] Cowan, M. et al. 2021. Porcupine: A Synthesizing Compilerfor Vectorized Homomorphic Encryption – PLDI 2021, 375–389.

HECO: Compiler for FHE

[USENIX Security'23]



open source, **automated**
end-to-end optimization for FHE

Democratize Privacy-Preserving Computation

My work aims to **democratize** access to privacy-preserving computation with new tools, systems, and abstractions.

Secure Computation

SoK: Fully Homomorphic Encryption Compilers

Alexander Viand, ETH Zurich · Patrick Jattke, ETH Zurich · Anwar Hithnawi, ETH Zurich

FHE Compilers

IEEE S&P

HECO: Fully Homomorphic Encryption Compiler

Alexander Viand, Patrick Jattke, Miro Haller, Anwar Hithnawi

ETH Zurich

HECO

USENIX Security

Differential Privacy

Cohere: Managing Differential Privacy in Large Scale Systems

Nicolas Küchler[1], Emanuel Opel[1], Hidde Lycklama[1], Alexander Viand[1], Anwar Hithnawi[1]

[1]ETH Zurich [2]Intel Labs

Cohere

IEEE S&P

Programmability

Deployments

Democratize Privacy-Preserving Computation

My work aims to **democratize** access to privacy-preserving computation with new tools, systems, and abstractions.

Secure Computation

SoK: Fully Homomorphic Encryption Compilers

HECO: Fully Homomorphic Encryption Compiler

FHE Compilers
IEEE S&P

HECO
USENIX Security

Differential Privacy

Cohere: Managing Differential Privacy in Large Scale Systems

Cohere
IEEE S&P

Programmability

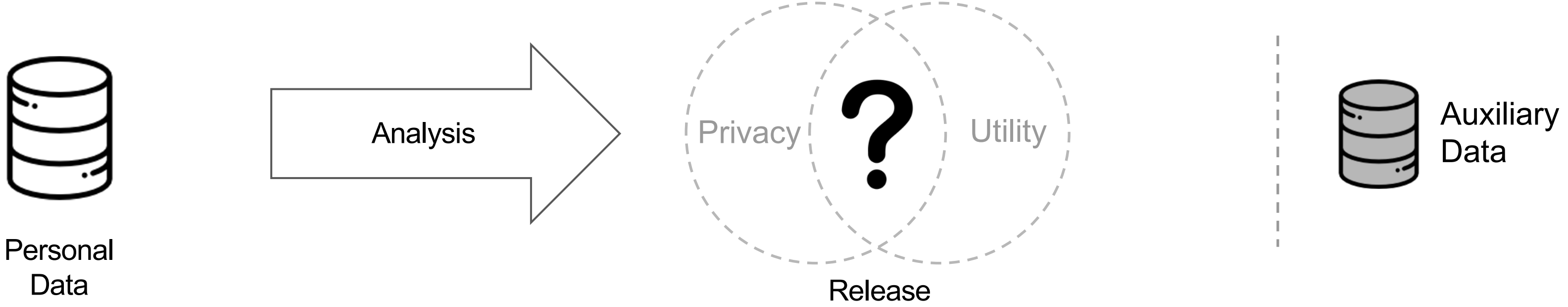
Deployments

Differential Privacy in Large-Scale Systems

(IEEE S&P‘24)

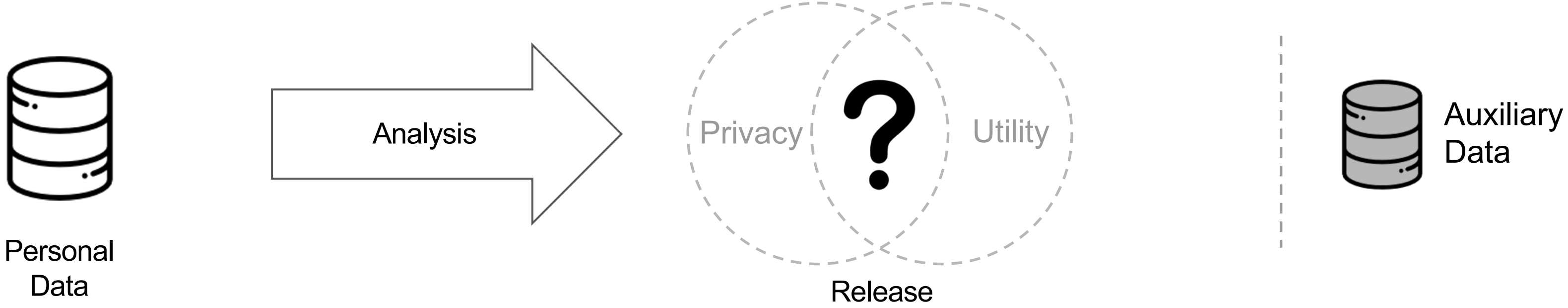
Statistical Release

How can we release useful information without compromising privacy?



Statistical Release

How can we release useful information without compromising privacy?



Personal Data

Analysis

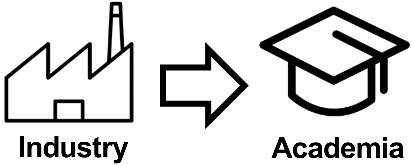
Privacy

?

Utility

Release

Auxiliary Data

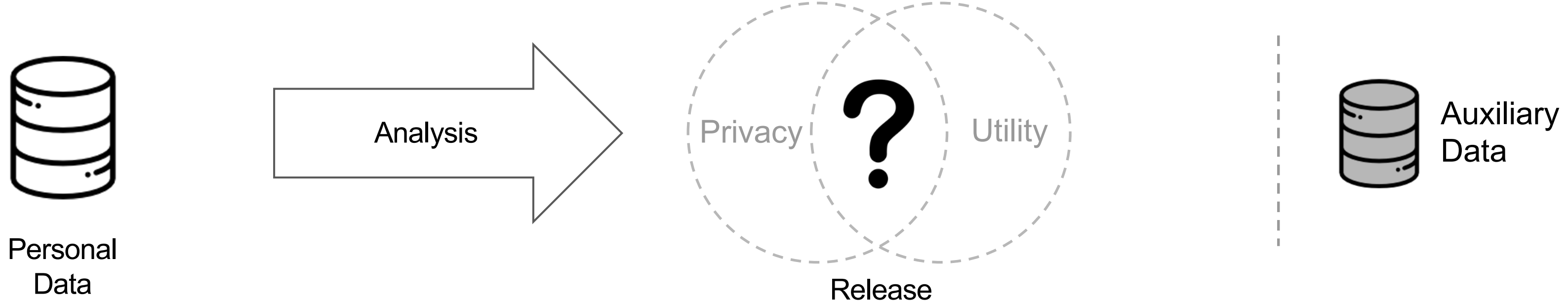


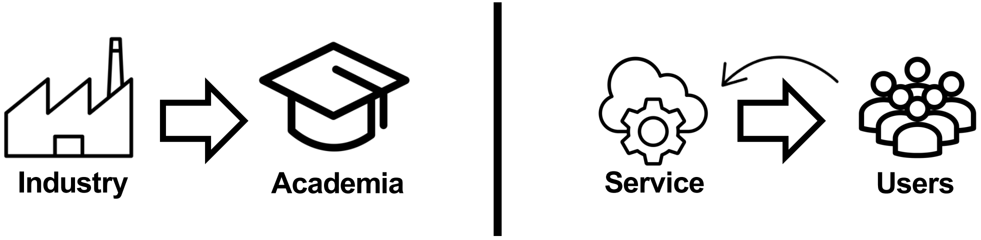
Industry

Academia

Statistical Release

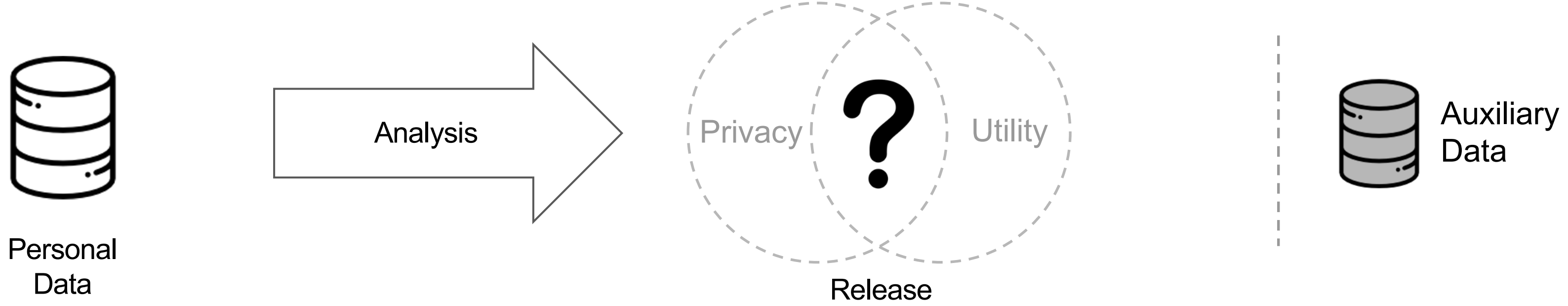
How can we release useful information without compromising privacy?





Statistical Release

How can we release useful information without compromising privacy?



Personal Data

Analysis

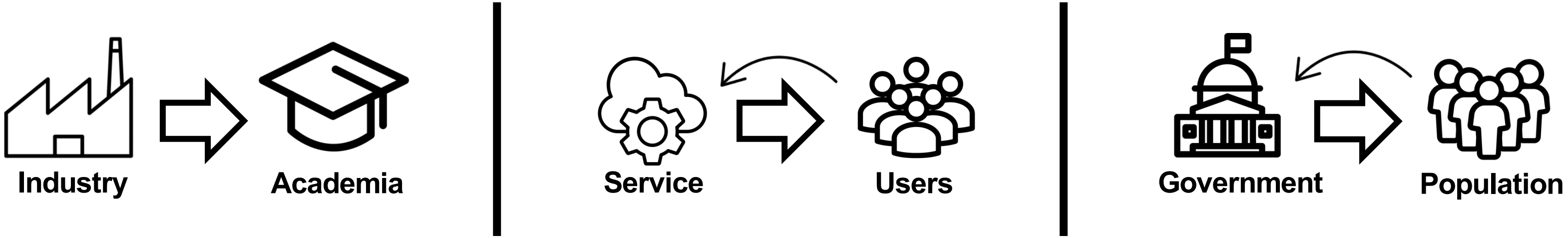
Privacy

?

Utility

Release

Auxiliary Data



Industry

Academia

Service

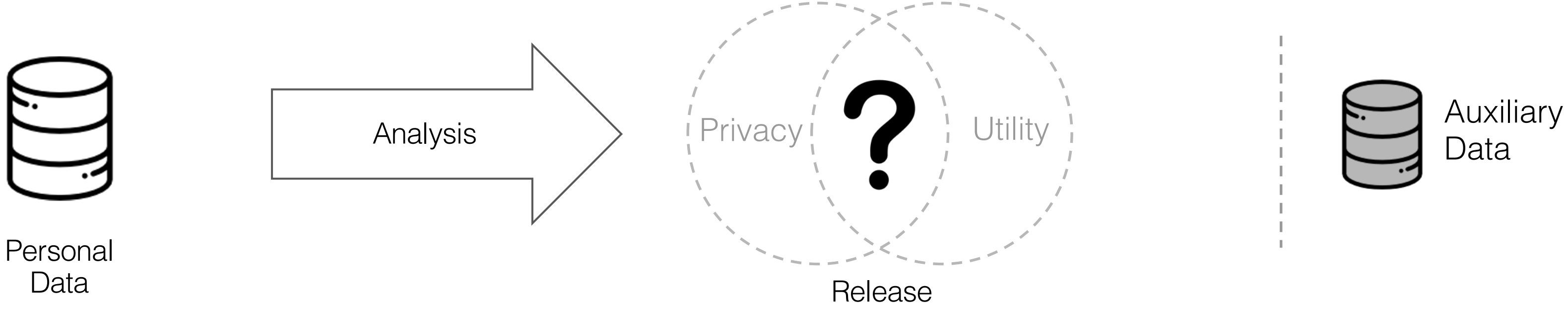
Users

Government

Population

Statistical Release

How can we release useful information without compromising privacy?



Personal Data

Analysis

Privacy

?

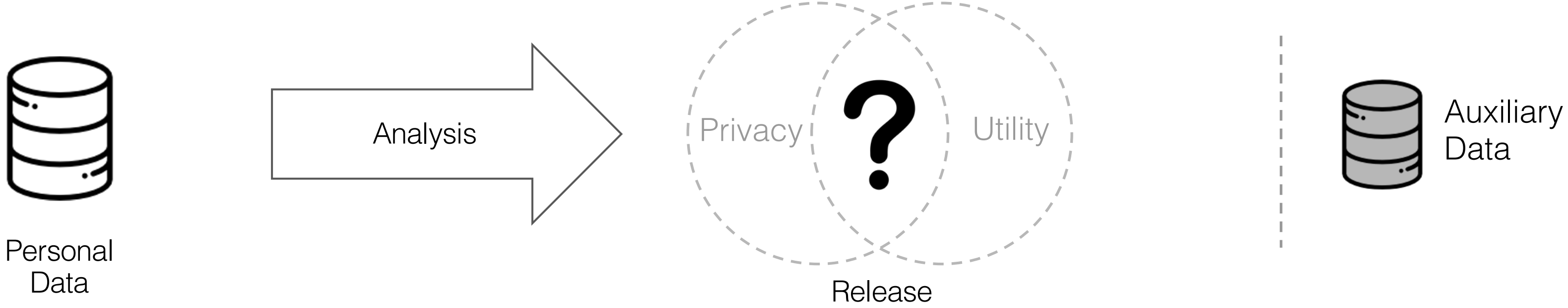
Utility

Release

Auxiliary Data

Statistical Release

How can we release useful information without compromising privacy?



Personal Data

Analysis

Privacy

?

Utility

Release

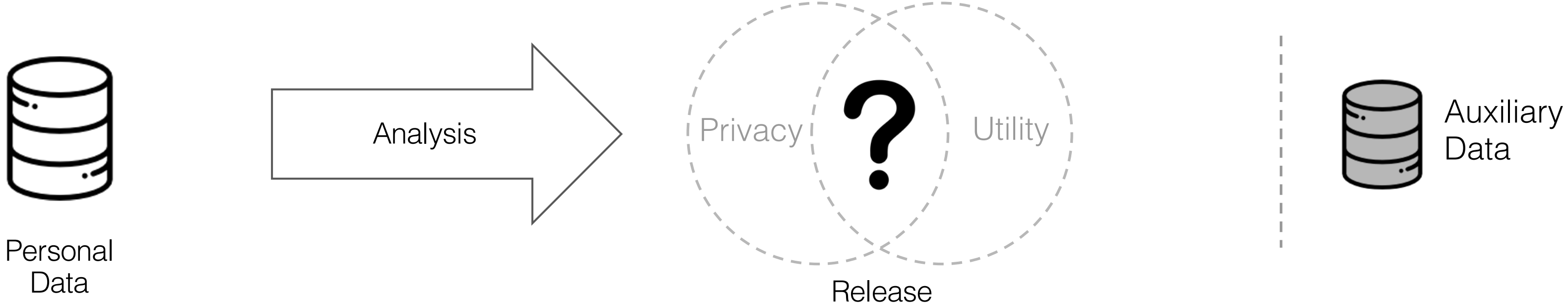
Auxiliary Data

- Anonymization
  Redact Personally Identifiable Information

Name	Region	...	Value
[redacted]	CH		100
[redacted]	DE		237

Statistical Release

How can we release useful information without compromising privacy?



Personal Data

Analysis

Privacy

?

Utility

Release

Auxiliary Data

- Anonymization
  Redact Personally Identifiable Information

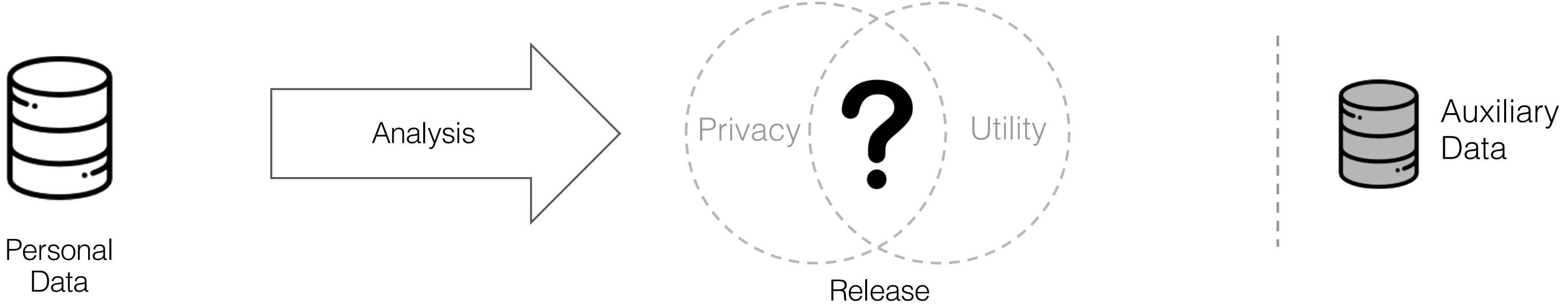
Name	Region	...	Value
[redacted]	CH		100
[redacted]	DE		237

- Release Aggregates

Descriptive Statistics / ML Model

Statistical Release

How can we release useful information without compromising privacy?



Personal Data

Analysis

Privacy ? Utility

Release

Auxiliary Data

- ⚠ Anonymization
  Redact Personally Identifiable Information

Name	Region	...	Value
[redacted]	CH		100
[redacted]	DE		237

- ⚠ Release Aggregates

Descriptive Statistics / ML Model

Privacy Attacks

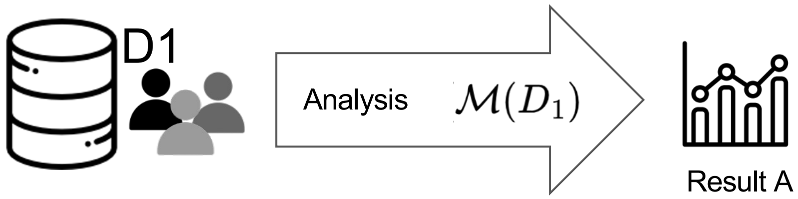
- Re-Identification (NYC TAXI)
- Database Reconstruction (United States Census 2010)
- Membership Inference (LLM)

Differential Privacy

Mathematical definition of privacy in the context of statistical releases

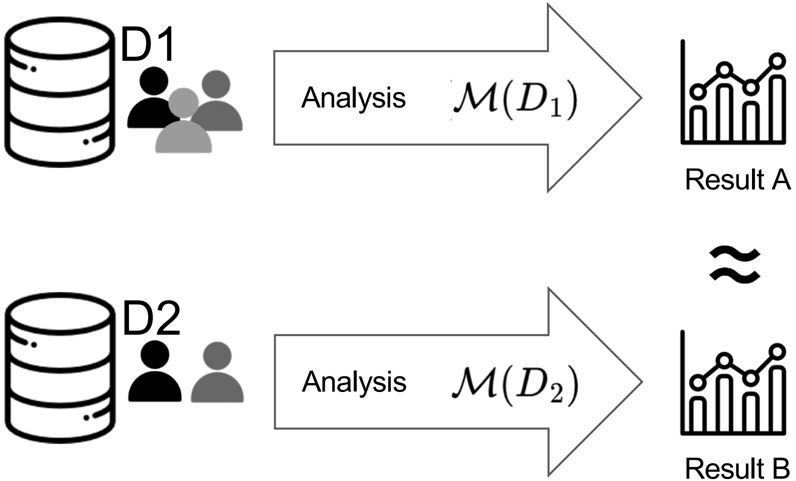
Differential Privacy

Mathematical definition of privacy in the context of statistical releases



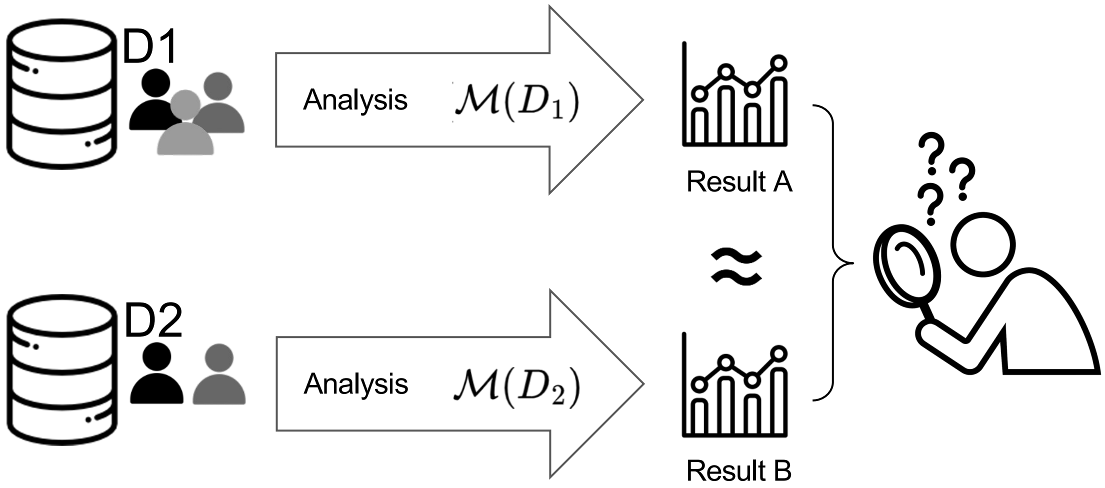
Differential Privacy

Mathematical definition of privacy in the context of statistical releases



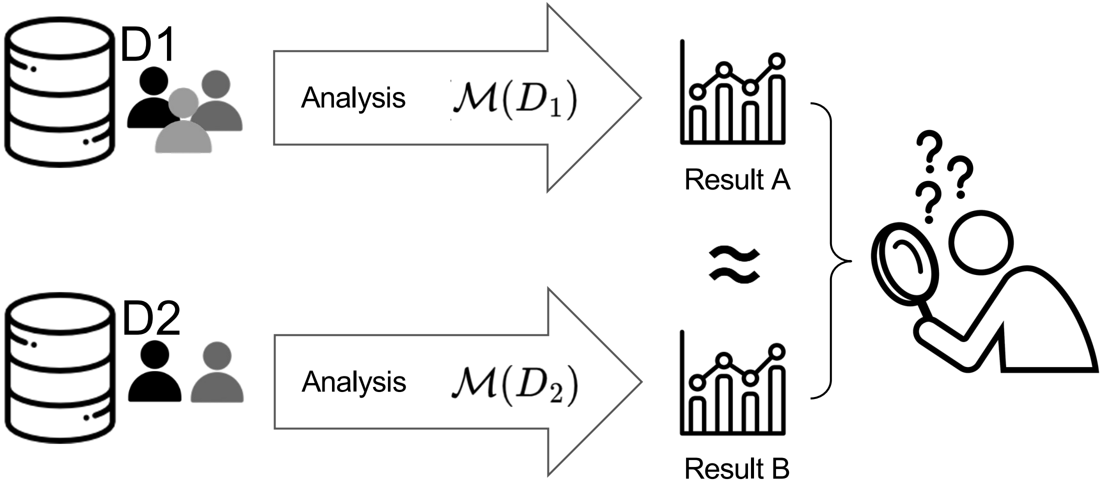
Differential Privacy

Mathematical definition of privacy in the context of statistical releases



Differential Privacy

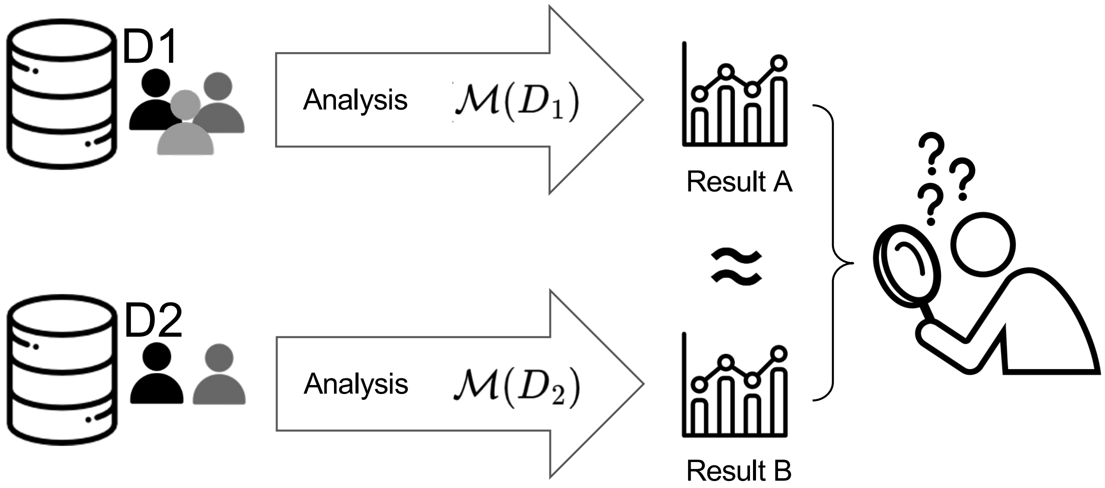
Mathematical definition of privacy in the context of statistical releases



$$Pr\left[\mathcal{M}(D_1) \in \mathcal{S}\right] \leq e^{\epsilon} \cdot Pr\left[\mathcal{M}(D_2) \in \mathcal{S}\right] + \delta$$

Differential Privacy

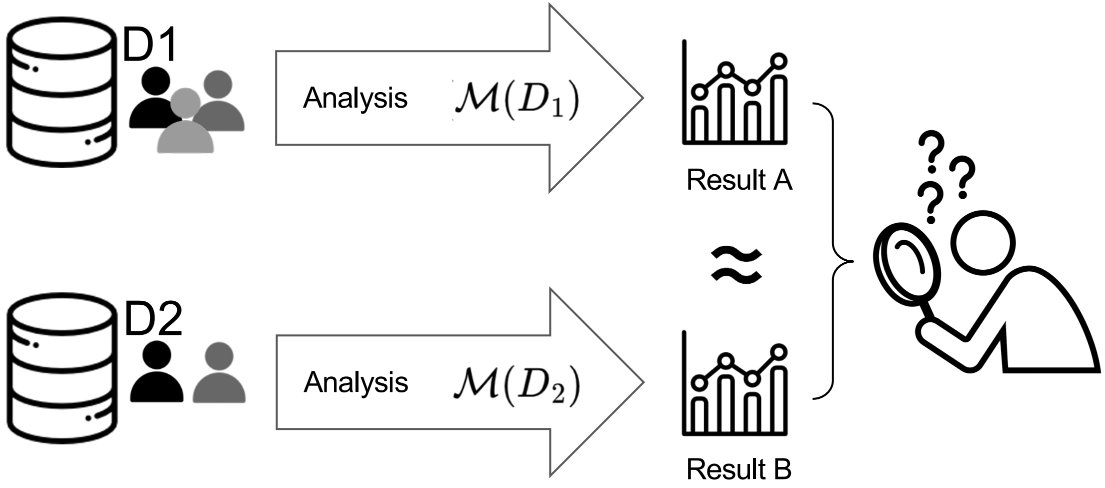
Mathematical definition of privacy in the context of statistical releases

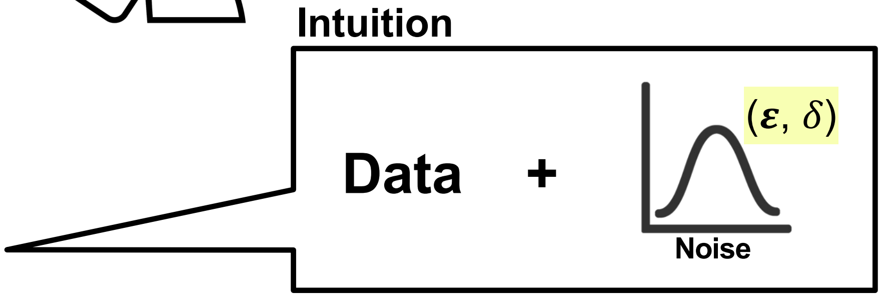


$$Pr\left[\mathcal{M}(D_1) \in \mathcal{S}\right] \leq e^{\epsilon} \cdot Pr\left[\mathcal{M}(D_2) \in \mathcal{S}\right] + \delta$$

Differential Privacy

Mathematical definition of privacy in the context of statistical releases

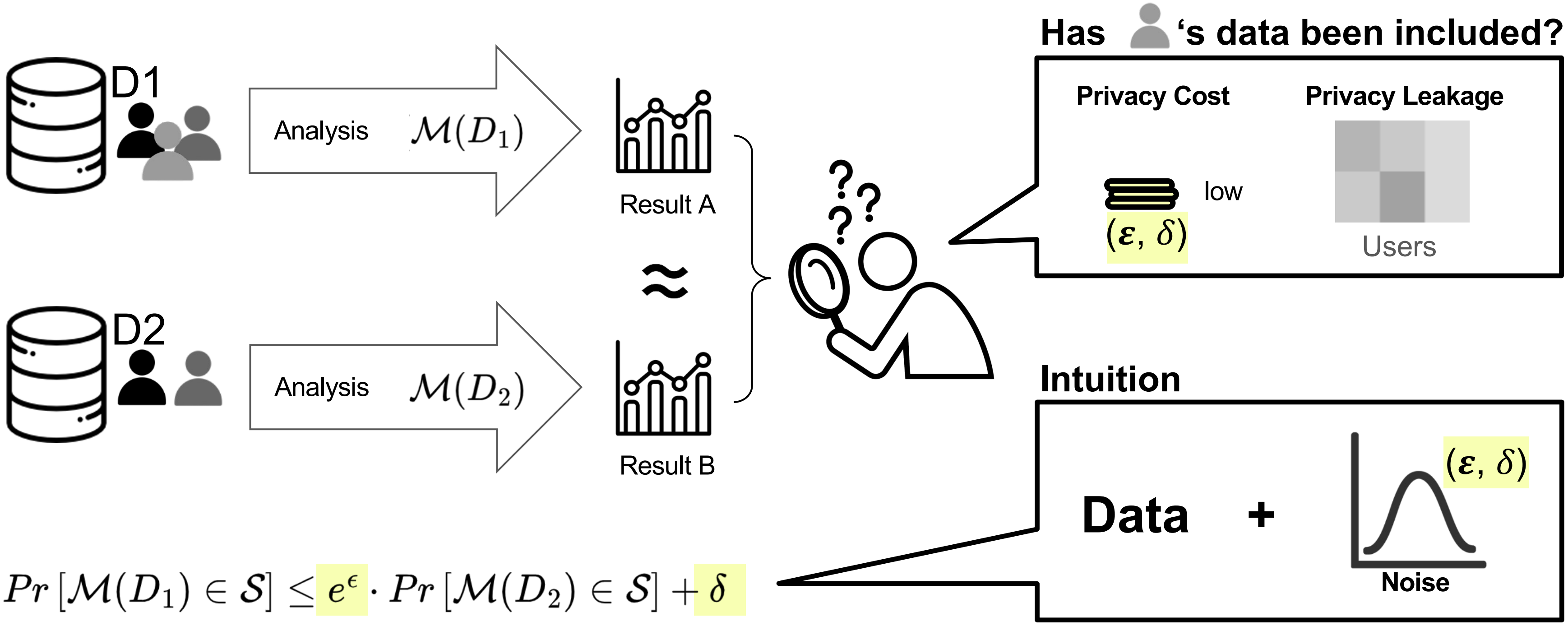




$$Pr\left[\mathcal{M}(D_1) \in \mathcal{S}\right] \leq e^{\epsilon} \cdot Pr\left[\mathcal{M}(D_2) \in \mathcal{S}\right] + \delta$$

Differential Privacy

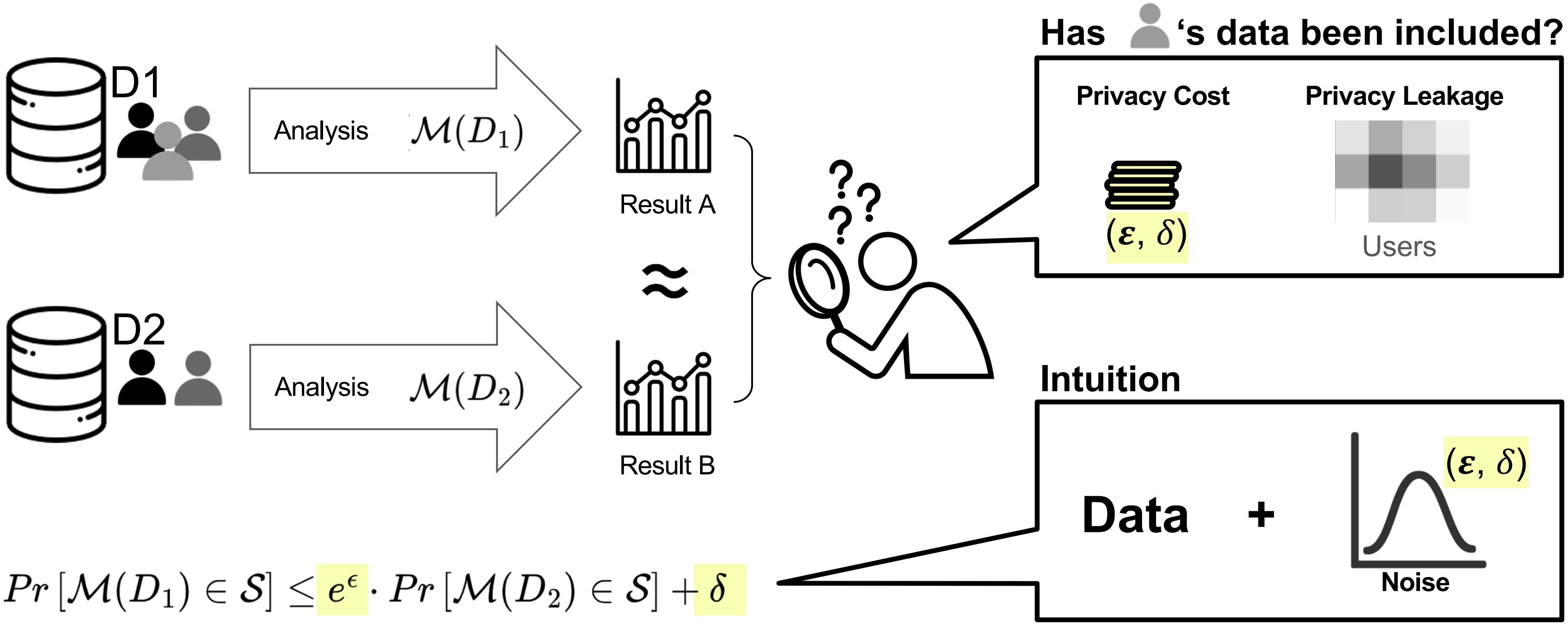
Mathematical definition of privacy in the context of statistical releases



$$Pr\left[\mathcal{M}(D_1) \in \mathcal{S}\right] \leq e^{\epsilon} \cdot Pr\left[\mathcal{M}(D_2) \in \mathcal{S}\right] + \delta$$

Differential Privacy

Mathematical definition of privacy in the context of statistical releases



D1

Analysis $\mathcal{M}(D_1)$

Result A

$\approx$

D2

Analysis $\mathcal{M}(D_2)$

Result B

Has ‘s data been included?

Privacy Cost

(ε, δ)

Privacy Leakage

Users

Intuition

Data +

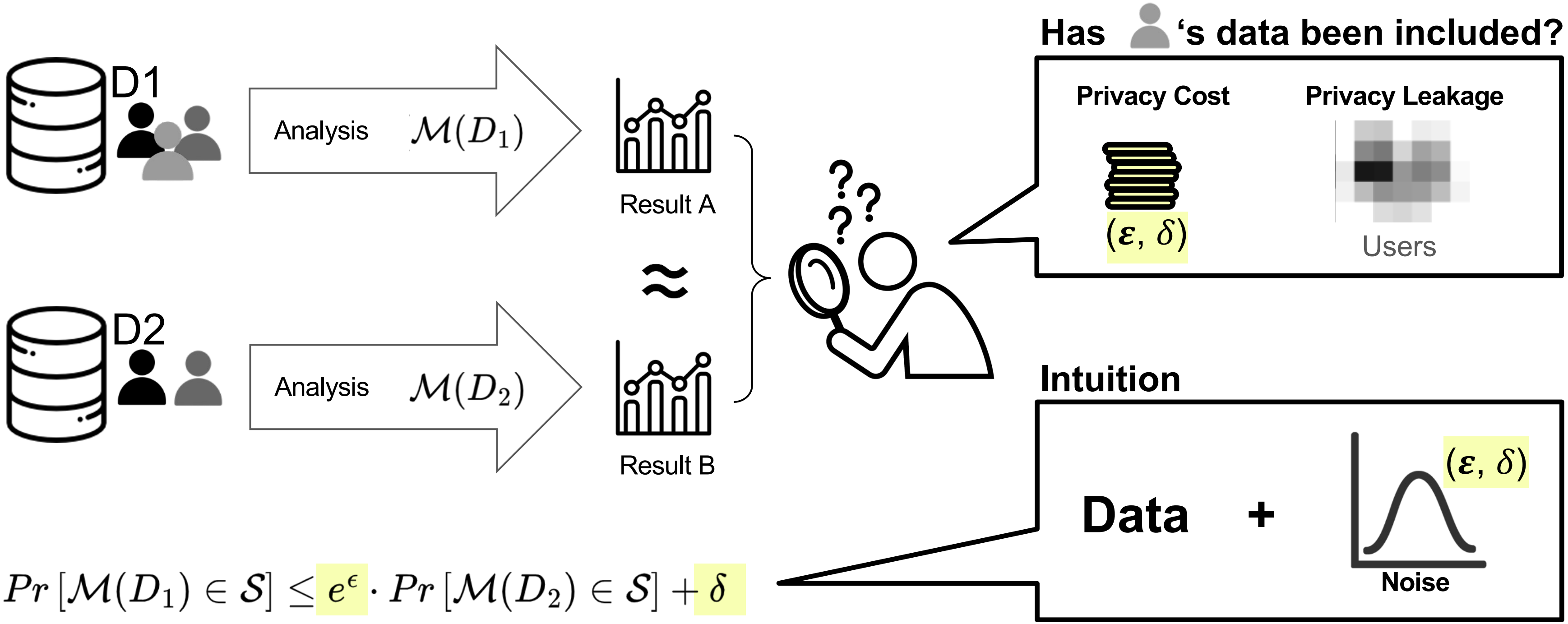
Noise

(ε, δ)

$$Pr\left[\mathcal{M}(D_1) \in \mathcal{S}\right] \leq e^{\epsilon} \cdot Pr\left[\mathcal{M}(D_2) \in \mathcal{S}\right] + \delta$$

Differential Privacy

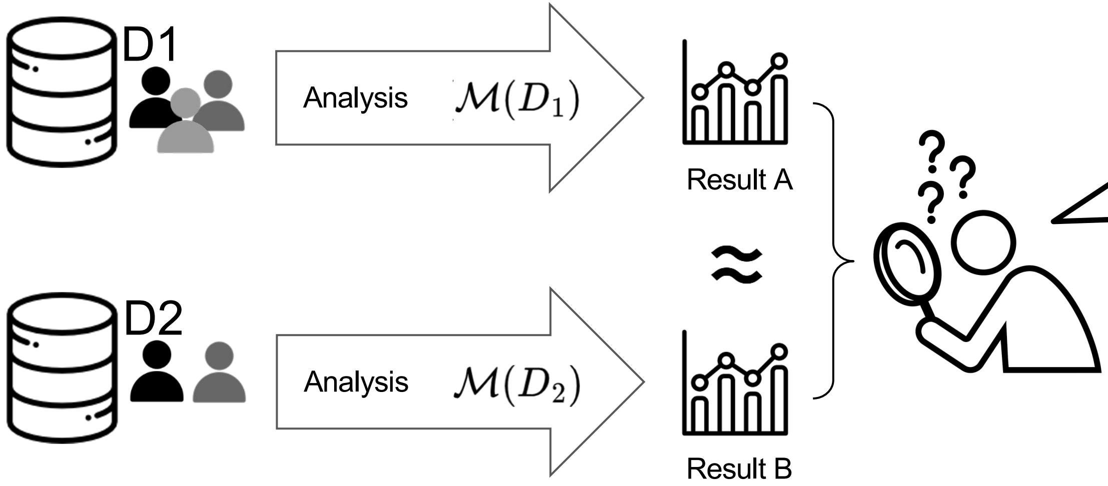
Mathematical definition of privacy in the context of statistical releases



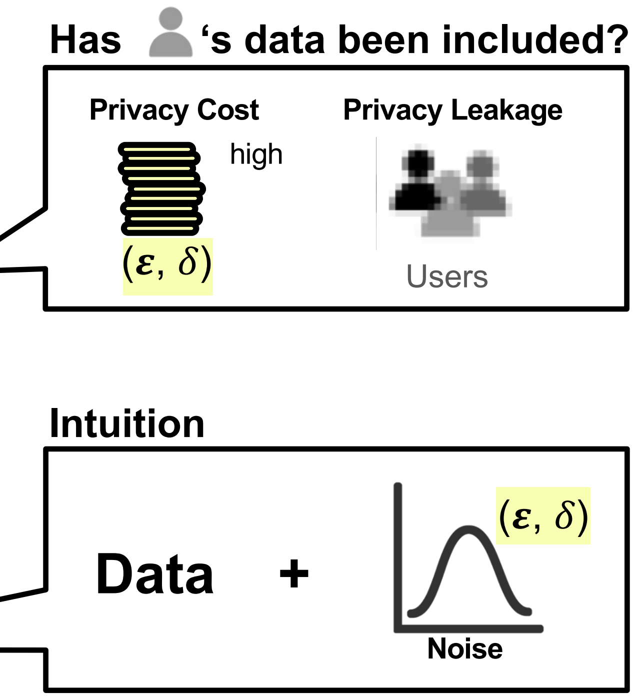
$$Pr\left[\mathcal{M}(D_1) \in \mathcal{S}\right] \leq e^{\epsilon} \cdot Pr\left[\mathcal{M}(D_2) \in \mathcal{S}\right] + \delta$$

Differential Privacy

Mathematical definition of privacy in the context of statistical releases



$$Pr\left[\mathcal{M}(D_1) \in \mathcal{S}\right] \leq e^{\epsilon} \cdot Pr\left[\mathcal{M}(D_2) \in \mathcal{S}\right] + \delta$$


Has 👤's data been included?

Privacy Cost

high

$(\boldsymbol{\varepsilon}, \delta)$

Privacy Leakage

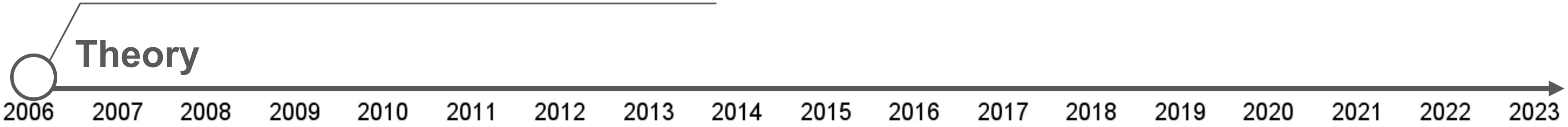
Users

Intuition

Data + Noise

$(\boldsymbol{\varepsilon}, \delta)$

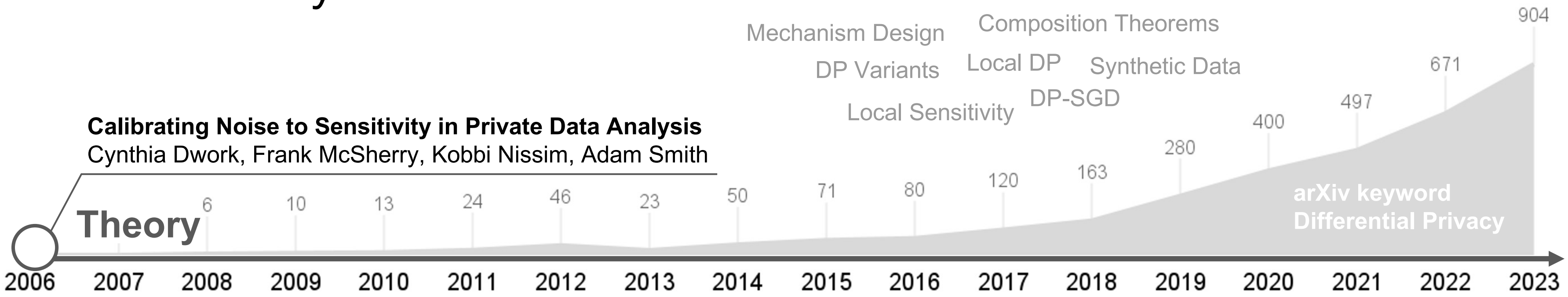
From Theory to Practice



Calibrating Noise to Sensitivity in Private Data Analysis
Cynthia Dwork, Frank McSherry, Kobbi Nissim, Adam Smith

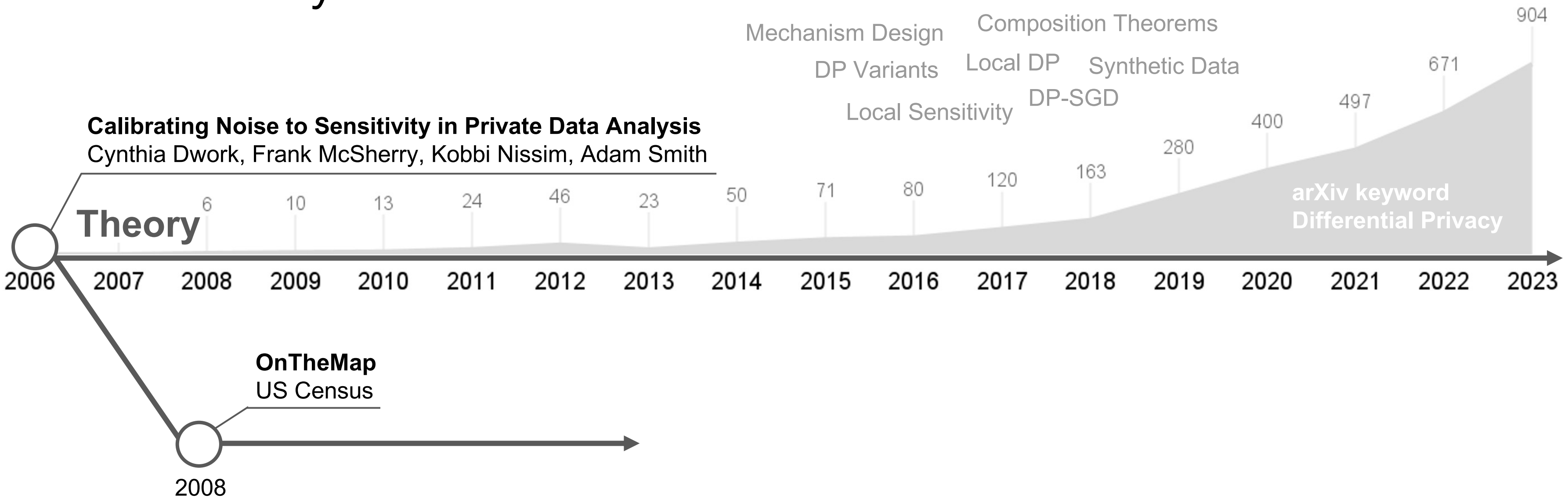
Theory

From Theory to Practice



Calibrating Noise to Sensitivity in Private Data Analysis
Cynthia Dwork, Frank McSherry, Kobbi Nissim, Adam Smith

From Theory to Practice



Calibrating Noise to Sensitivity in Private Data Analysis
Cynthia Dwork, Frank McSherry, Kobbi Nissim, Adam Smith

Theory

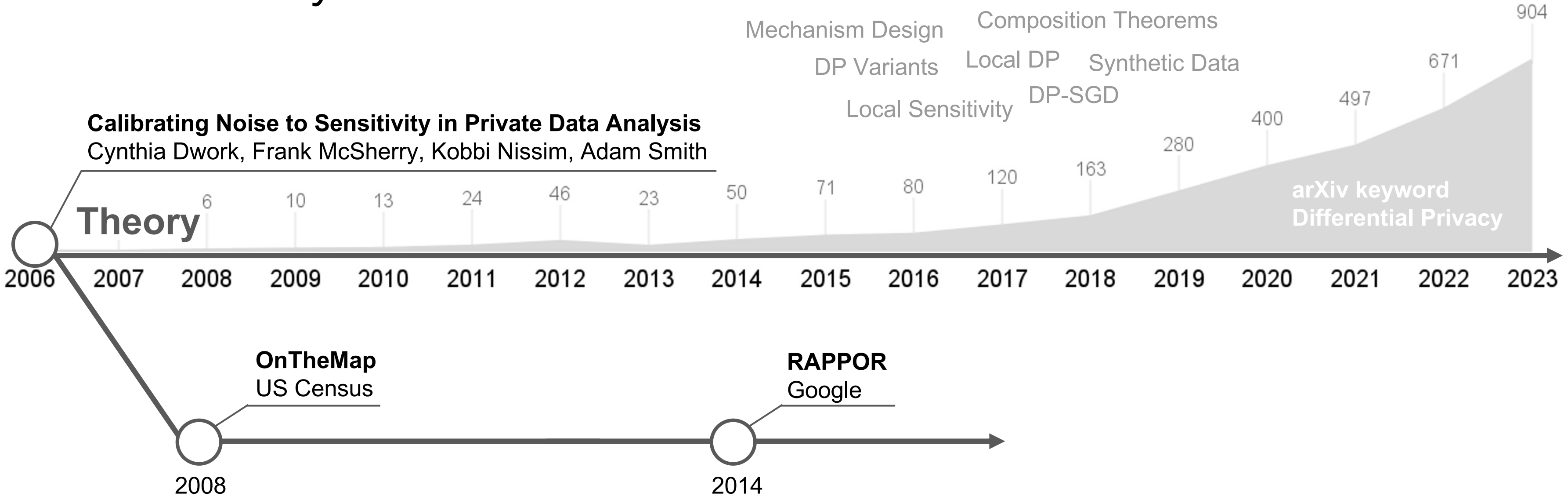
OnTheMap
US Census

2008

Real-World Applications

[Desfontaines Blog, 2021]

From Theory to Practice



Calibrating Noise to Sensitivity in Private Data Analysis
Cynthia Dwork, Frank McSherry, Kobbi Nissim, Adam Smith

Theory

Mechanism Design

DP Variants

Local Sensitivity

Composition Theorems

Local DP

DP-SGD

Synthetic Data

arXiv keyword
Differential Privacy

OnTheMap
US Census

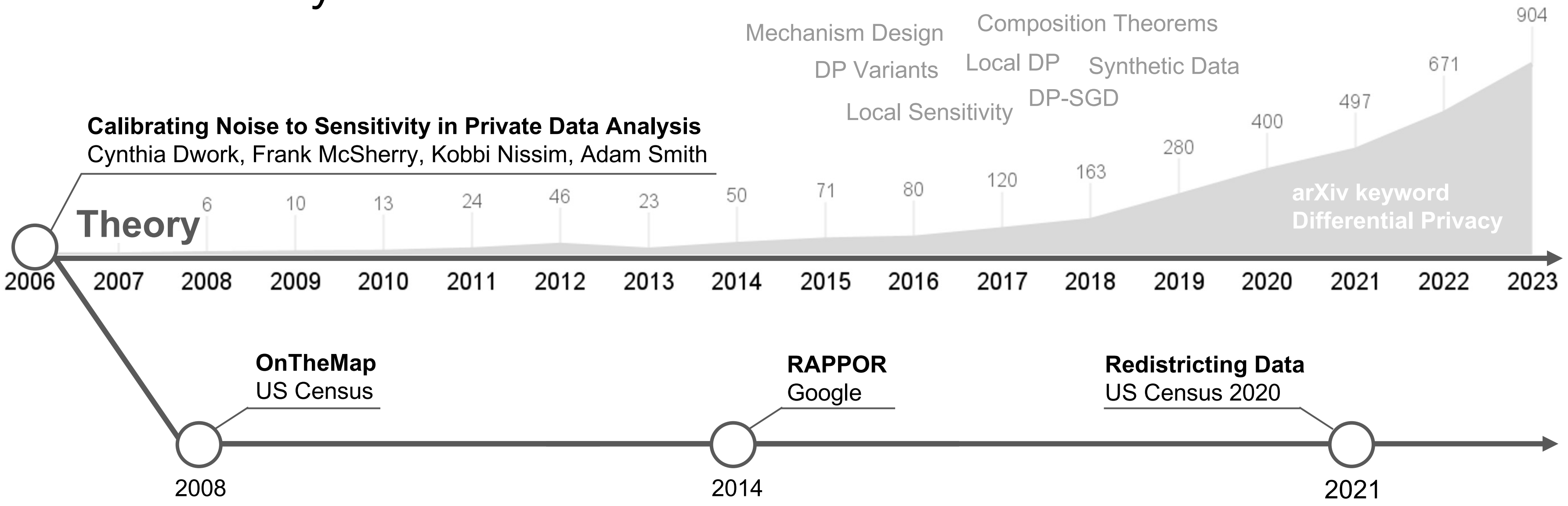
2008

RAPPOR
Google

2014

Real-World Applications
[Desfontaines Blog, 2021]

From Theory to Practice



Mechanism Design

Composition Theorems

DP Variants

Local DP

Synthetic Data

DP-SGD

Local Sensitivity

Calibrating Noise to Sensitivity in Private Data Analysis
Cynthia Dwork, Frank McSherry, Kobbi Nissim, Adam Smith

Theory

arXiv keyword Differential Privacy

OnTheMap
US Census
2008

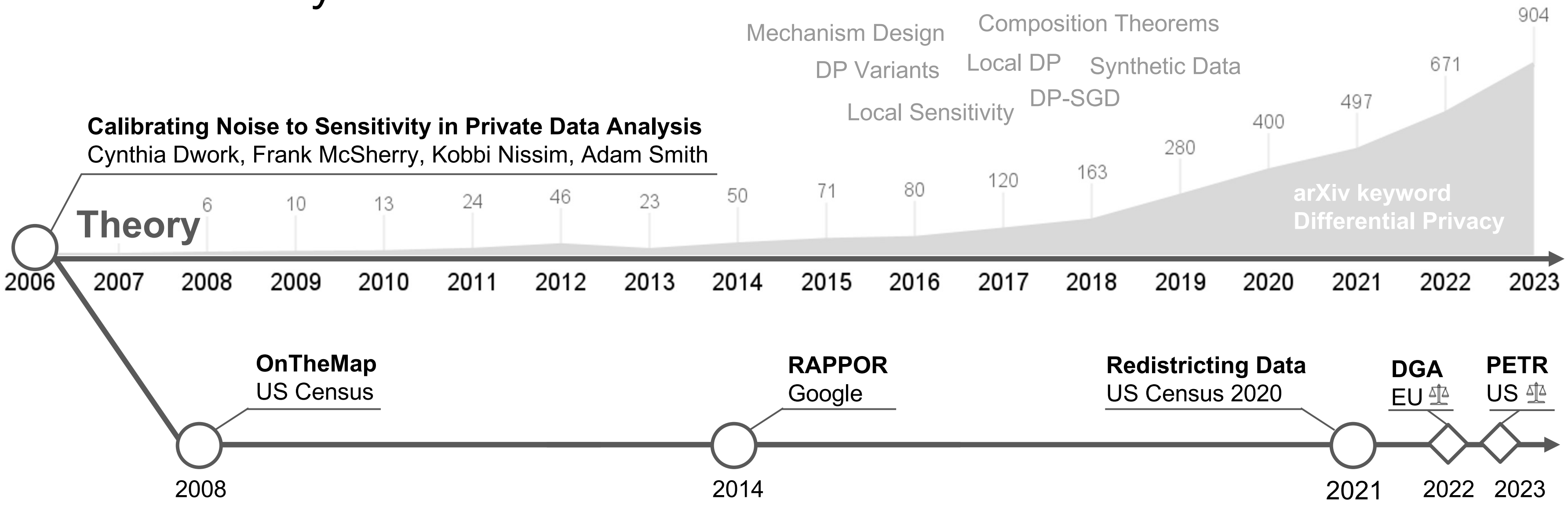
RAPPOR
Google
2014

Redistricting Data
US Census 2020
2021

Real-World Applications

[Desfontaines Blog, 2021]

From Theory to Practice



Calibrating Noise to Sensitivity in Private Data Analysis
Cynthia Dwork, Frank McSherry, Kobbi Nissim, Adam Smith

Mechanism Design
DP Variants
Local Sensitivity
Composition Theorems
Local DP
DP-SGD
Synthetic Data

OnTheMap
US Census
2008

RAPPOR
Google
2014

Redistricting Data
US Census 2020
2021

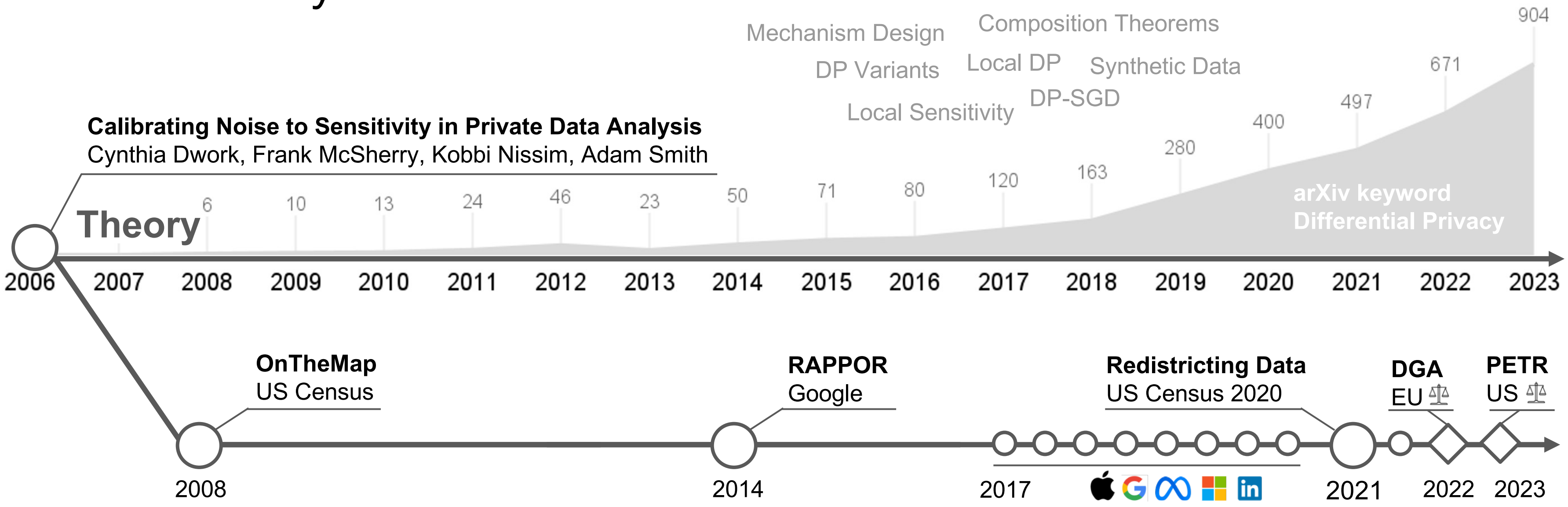
DGA
EU ⚖
2022

PETR
US ⚖
2023

Real-World Applications

[Desfontaines Blog, 2021]

From Theory to Practice



Calibrating Noise to Sensitivity in Private Data Analysis
Cynthia Dwork, Frank McSherry, Kobbi Nissim, Adam Smith

Theory

Mechanism Design
DP Variants
Local Sensitivity
Composition Theorems
Local DP
DP-SGD
Synthetic Data

arXiv keyword Differential Privacy

OnTheMap
US Census

RAPPOR
Google

Redistricting Data
US Census 2020

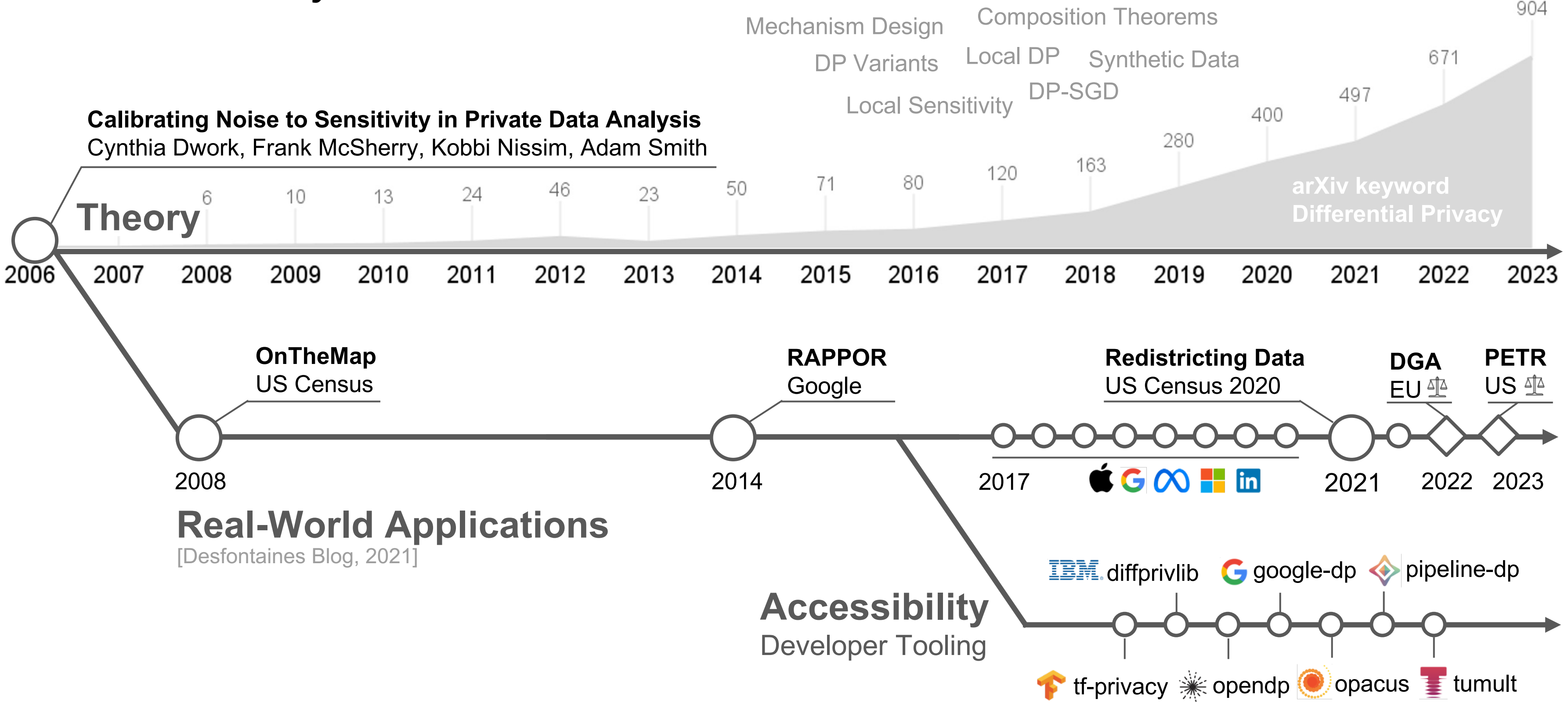
DGA
EU ⚖

PETR
US ⚖

Real-World Applications

[Desfontaines Blog, 2021]

From Theory to Practice



Calibrating Noise to Sensitivity in Private Data Analysis
Cynthia Dwork, Frank McSherry, Kobbi Nissim, Adam Smith

Theory

Mechanism Design
DP Variants
Local Sensitivity
Composition Theorems
Local DP
DP-SGD
Synthetic Data

arXiv keyword Differential Privacy

OnTheMap
US Census

RAPPOR
Google

Redistricting Data
US Census 2020

DGA
EU ⚖

PETR
US ⚖

Real-World Applications
[Desfontaines Blog, 2021]

Accessibility
Developer Tooling

IBM diffprivlib
google-dp
pipeline-dp
tf-privacy
opendp
opacus
tumult

Deploying DP Applications

Deploying DP Applications

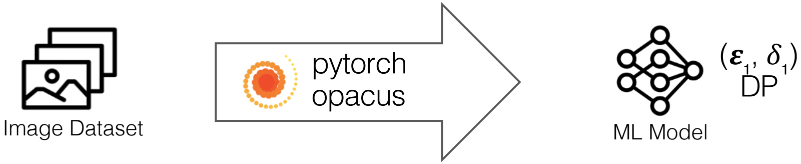


Image Dataset

pytorch
opacus

ML Model

$(\varepsilon_1, \delta_1)$
DP

Deploying DP Applications

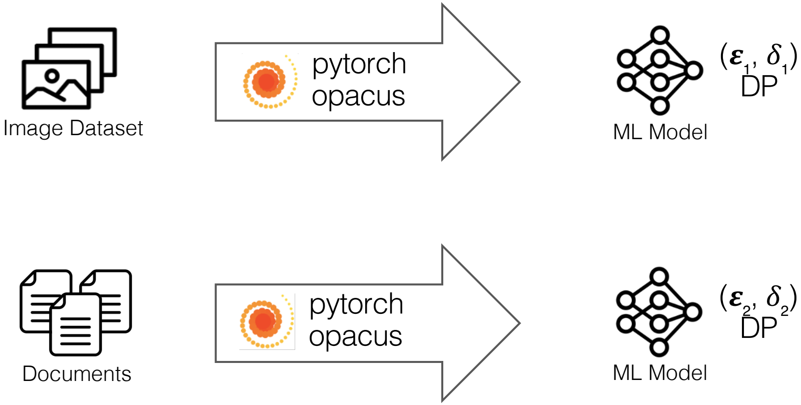


Image Dataset

pytorch opacus

ML Model

$(\boldsymbol{\varepsilon}_1, \delta_1)$ DP

Documents

pytorch opacus

ML Model

$(\boldsymbol{\varepsilon}_2, \delta_2)$ DP

Deploying DP Applications

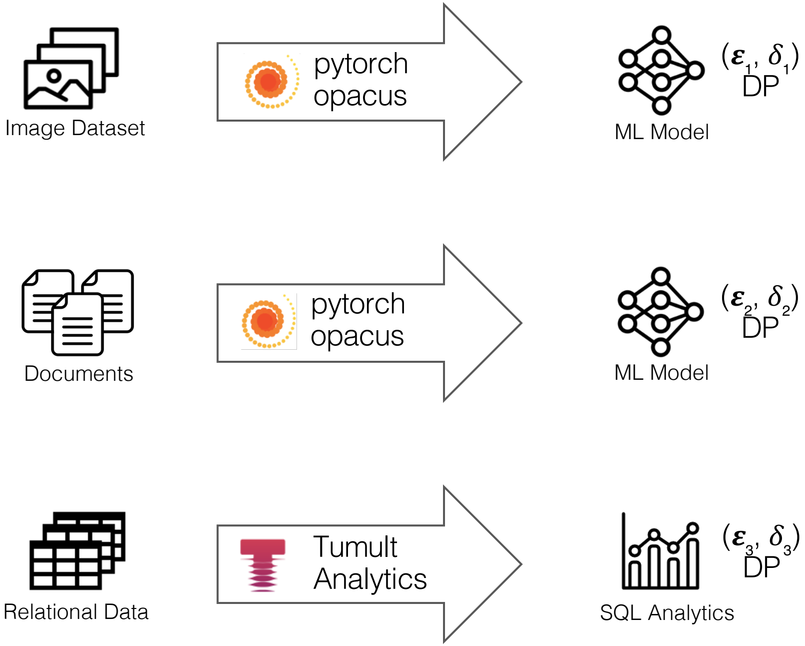


Image Dataset

pytorch opacus

ML Model

$(\varepsilon_1, \delta_1)$ DP

Documents

pytorch opacus

ML Model

$(\varepsilon_2, \delta_2)$ DP

Relational Data

Tumult Analytics

SQL Analytics

$(\varepsilon_3, \delta_3)$ DP

Deploying DP Applications

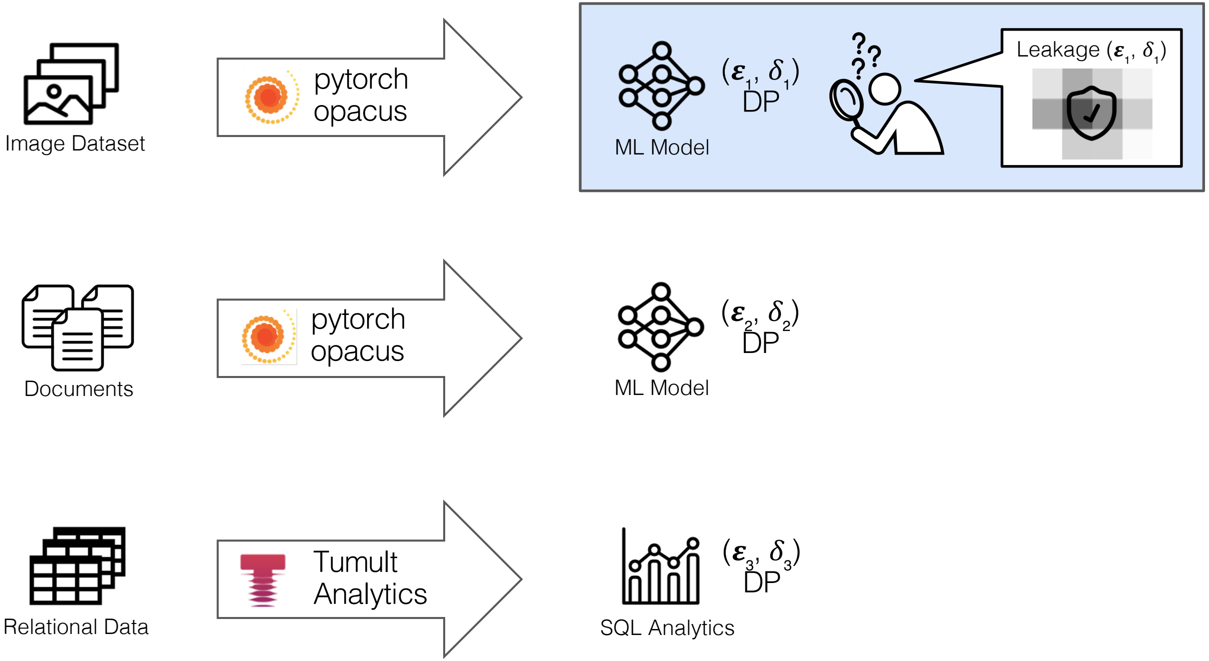


Image Dataset

pytorch opacus

ML Model

$(\varepsilon_1, \delta_1)$ DP

Leakage $(\varepsilon_1, \delta_1)$

Documents

pytorch opacus

ML Model

$(\varepsilon_2, \delta_2)$ DP

Relational Data

Tumult Analytics

SQL Analytics

$(\varepsilon_3, \delta_3)$ DP

Deploying DP Applications

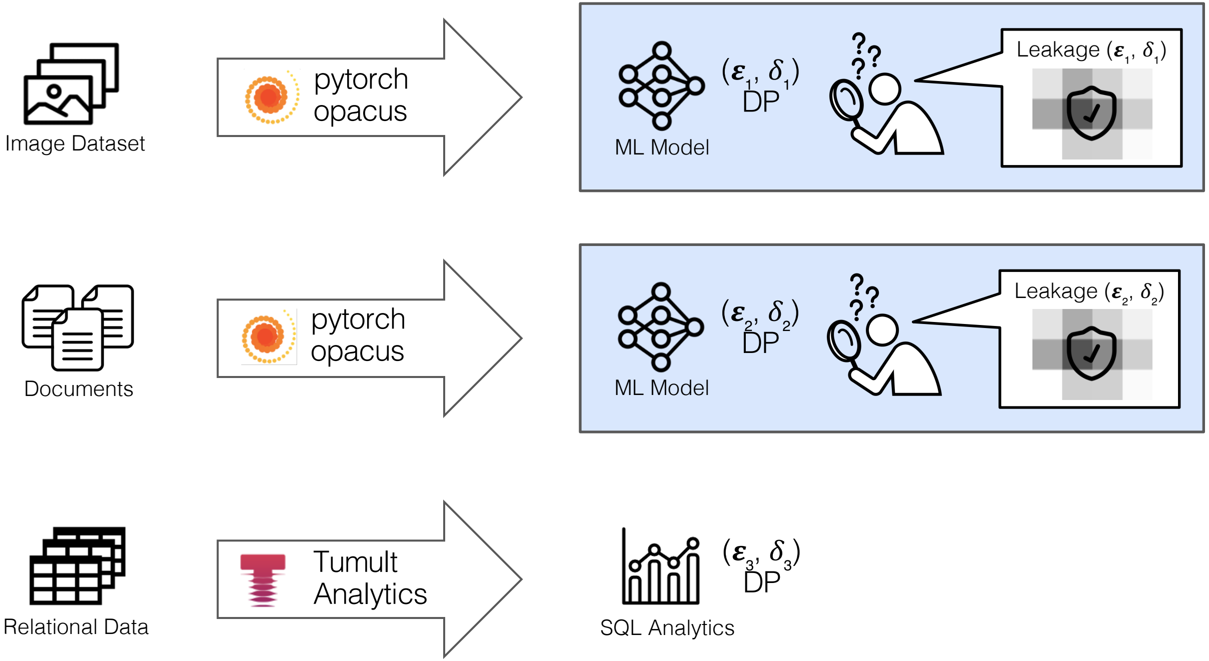


Image Dataset

pytorch opacus

ML Model

$(\varepsilon_1, \delta_1)$ DP

Leakage $(\varepsilon_1, \delta_1)$

Documents

pytorch opacus

ML Model

$(\varepsilon_2, \delta_2)$ DP

Leakage $(\varepsilon_2, \delta_2)$

Relational Data

Tumult Analytics

SQL Analytics

$(\varepsilon_3, \delta_3)$ DP

Deploying DP Applications

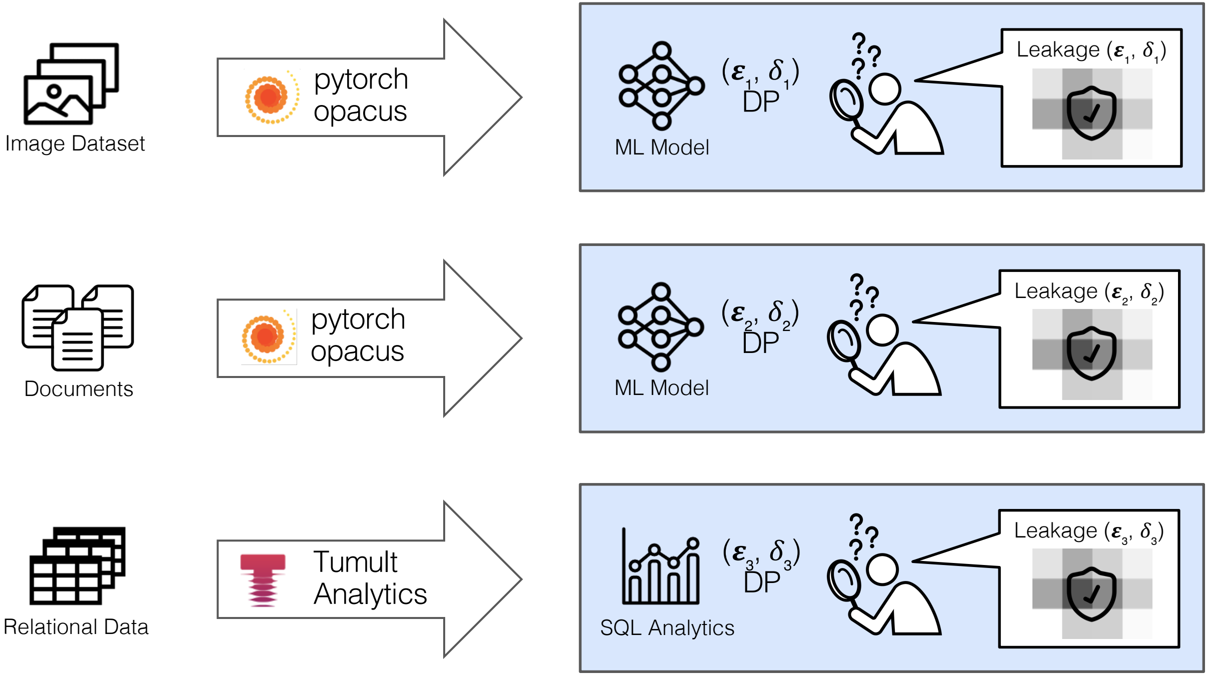


Image Dataset
pytorch opacus
ML Model
$(\varepsilon_1, \delta_1)$ DP
Leakage $(\varepsilon_1, \delta_1)$
Documents
pytorch opacus
ML Model
$(\varepsilon_2, \delta_2)$ DP
Leakage $(\varepsilon_2, \delta_2)$
Relational Data
Tumult Analytics
SQL Analytics
$(\varepsilon_3, \delta_3)$ DP
Leakage $(\varepsilon_3, \delta_3)$

Deploying DP Applications

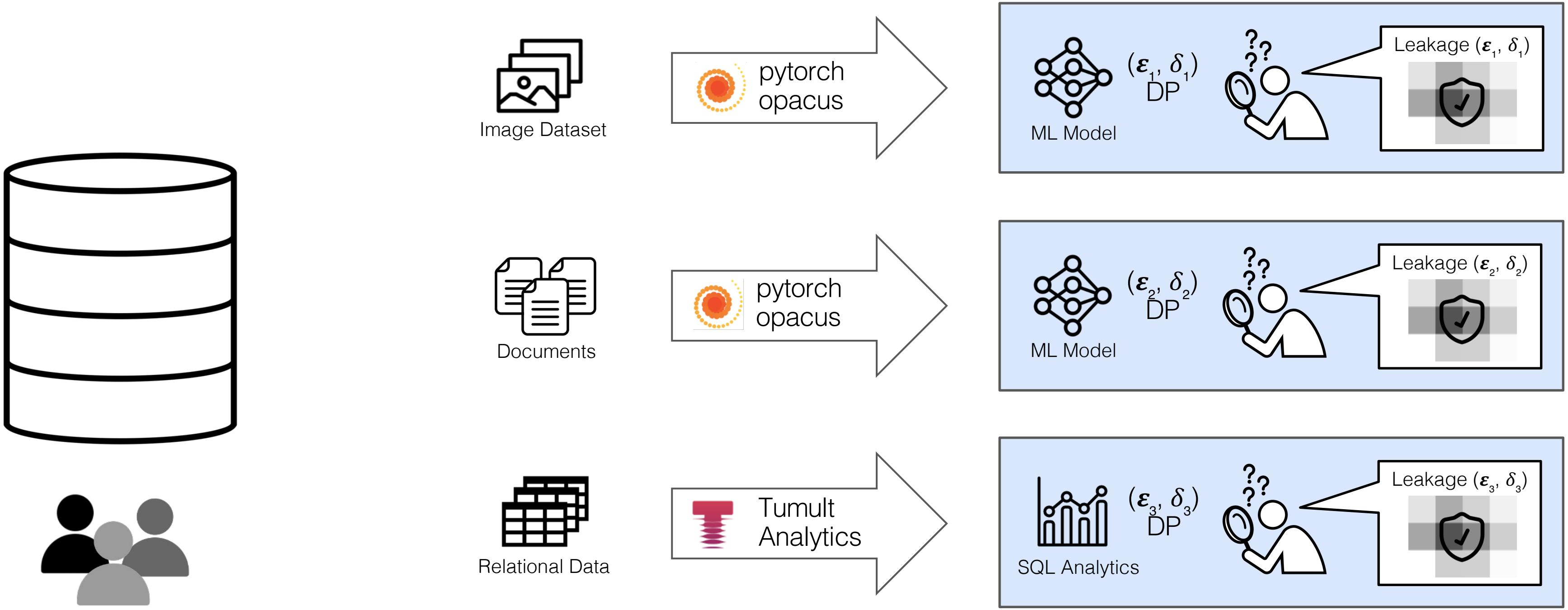


Image Dataset
pytorch opacus
ML Model
$(\varepsilon_1, \delta_1)$ DP
Leakage $(\varepsilon_1, \delta_1)$
Documents
pytorch opacus
ML Model
$(\varepsilon_2, \delta_2)$ DP
Leakage $(\varepsilon_2, \delta_2)$
Relational Data
Tumult Analytics
SQL Analytics
$(\varepsilon_3, \delta_3)$ DP
Leakage $(\varepsilon_3, \delta_3)$

Deploying DP Applications

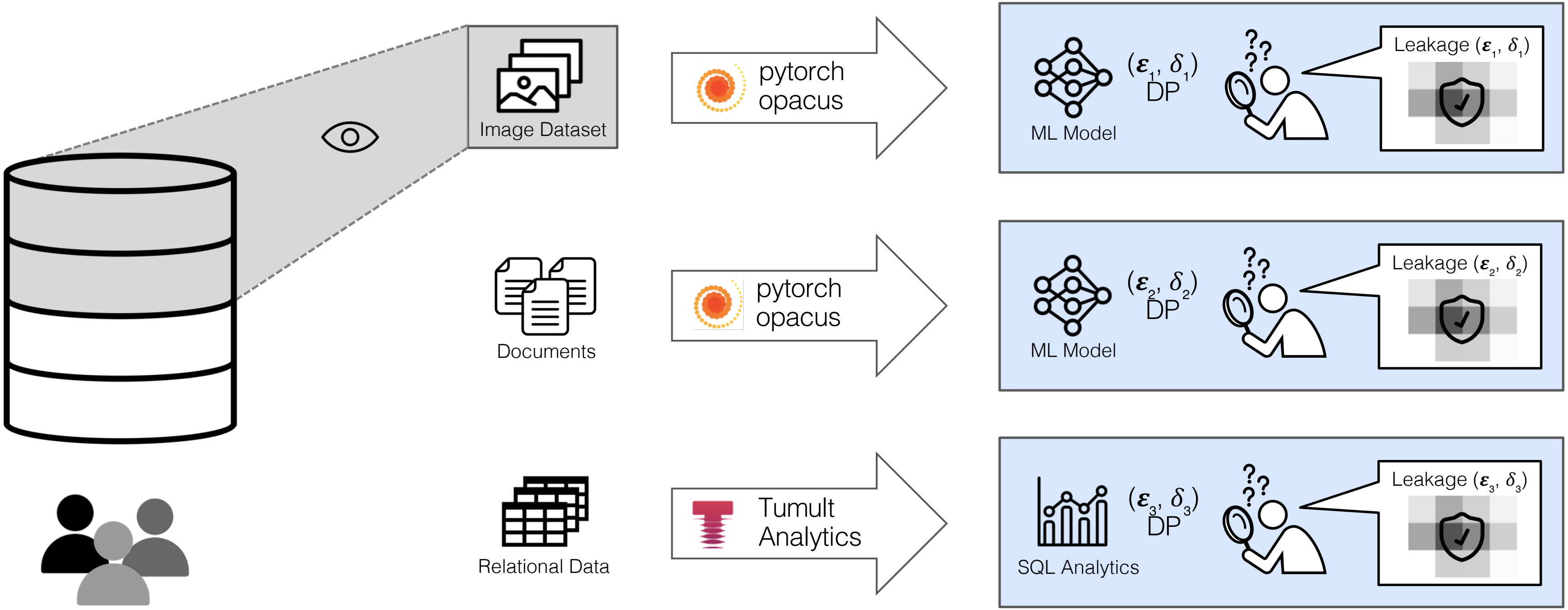


Image Dataset

pytorch opacus

ML Model

$(\varepsilon_1, \delta_1)$ DP

Leakage $(\varepsilon_1, \delta_1)$

Documents

pytorch opacus

ML Model

$(\varepsilon_2, \delta_2)$ DP

Leakage $(\varepsilon_2, \delta_2)$

Relational Data

Tumult Analytics

SQL Analytics

$(\varepsilon_3, \delta_3)$ DP

Leakage $(\varepsilon_3, \delta_3)$

Deploying DP Applications

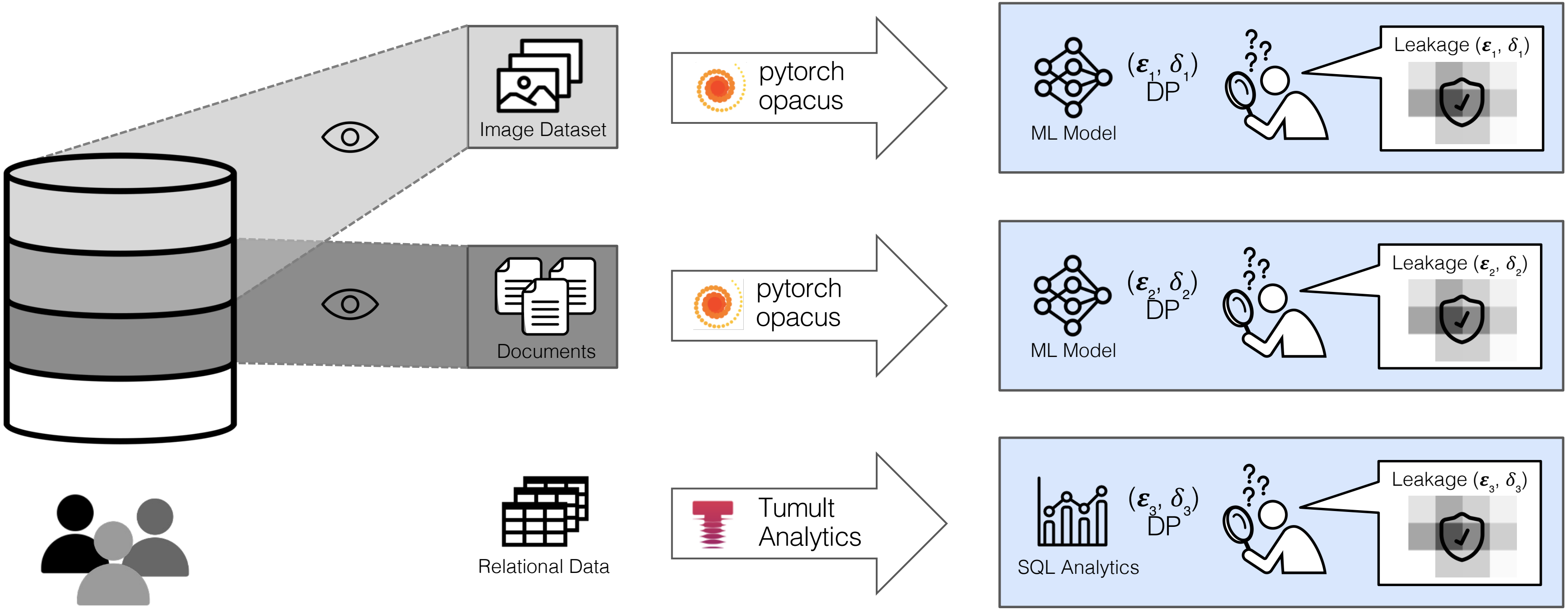


Image Dataset

Documents

Relational Data

pytorch opacus

pytorch opacus

Tumult Analytics

ML Model

$(\varepsilon_1, \delta_1)$ DP

Leakage $(\varepsilon_1, \delta_1)$

ML Model

$(\varepsilon_2, \delta_2)$ DP

Leakage $(\varepsilon_2, \delta_2)$

SQL Analytics

$(\varepsilon_3, \delta_3)$ DP

Leakage $(\varepsilon_3, \delta_3)$

Deploying DP Applications

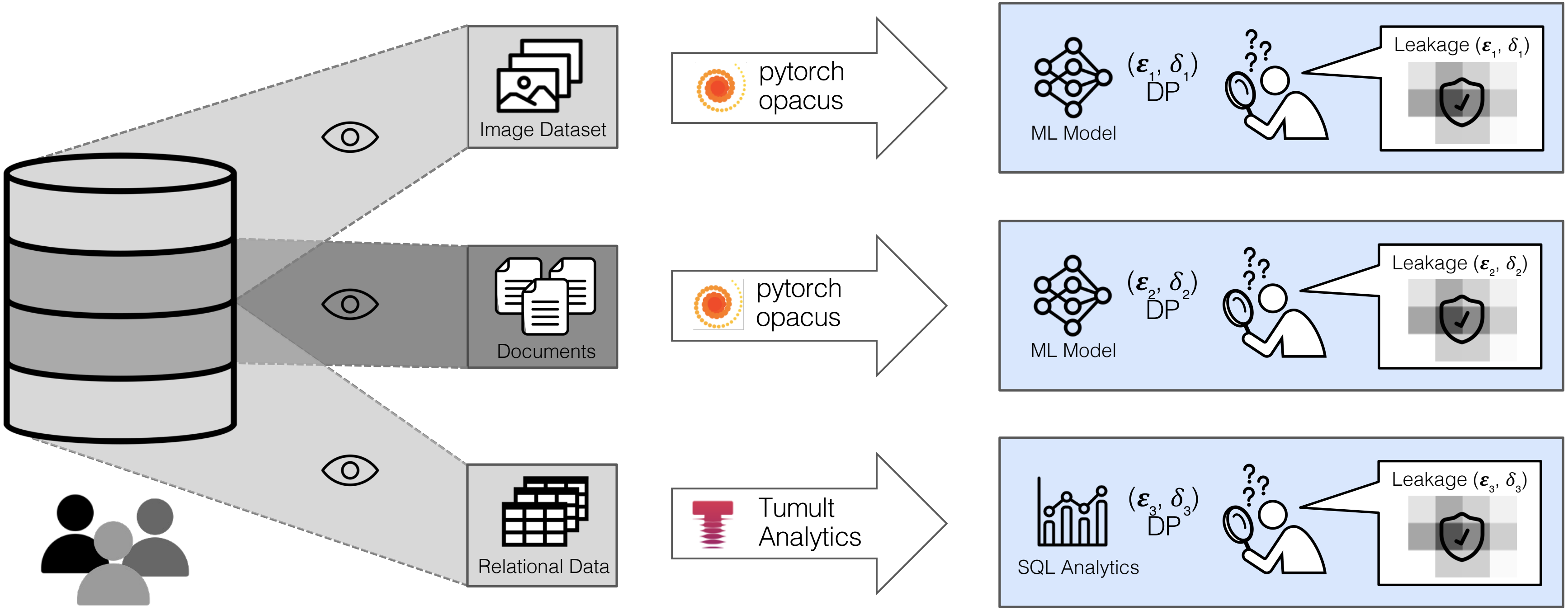


Image Dataset
Documents
Relational Data
pytorch opacus
pytorch opacus
Tumult Analytics
ML Model
$(\varepsilon_1, \delta_1)$ DP
Leakage $(\varepsilon_1, \delta_1)$
ML Model
$(\varepsilon_2, \delta_2)$ DP
Leakage $(\varepsilon_2, \delta_2)$
SQL Analytics
$(\varepsilon_3, \delta_3)$ DP
Leakage $(\varepsilon_3, \delta_3)$
??
??
??

Deploying DP Applications

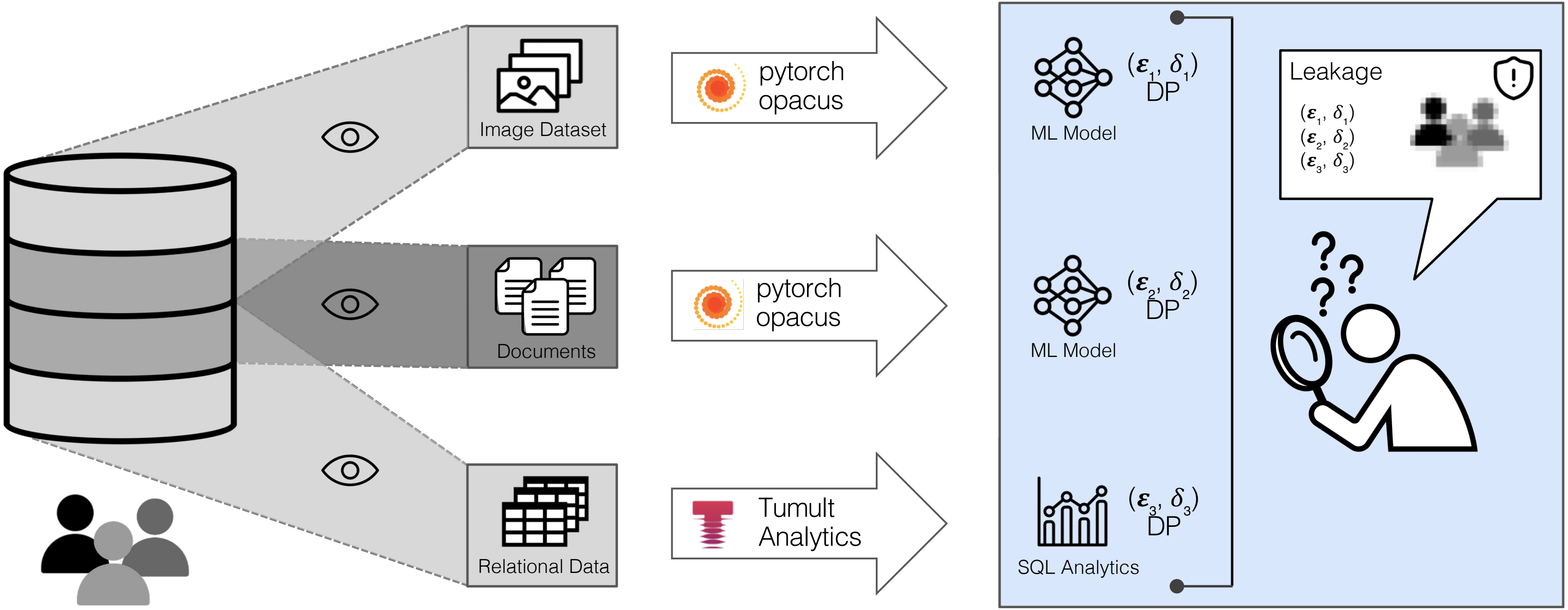


Image Dataset
Documents
Relational Data
pytorch opacus
pytorch opacus
Tumult Analytics
ML Model
$(\varepsilon_1, \delta_1)$ DP
ML Model
$(\varepsilon_2, \delta_2)$ DP
SQL Analytics
$(\varepsilon_3, \delta_3)$ DP
Leakage
$(\varepsilon_1, \delta_1)$
$(\varepsilon_2, \delta_2)$
$(\varepsilon_3, \delta_3)$
? ? ?

Deploying DP Applications

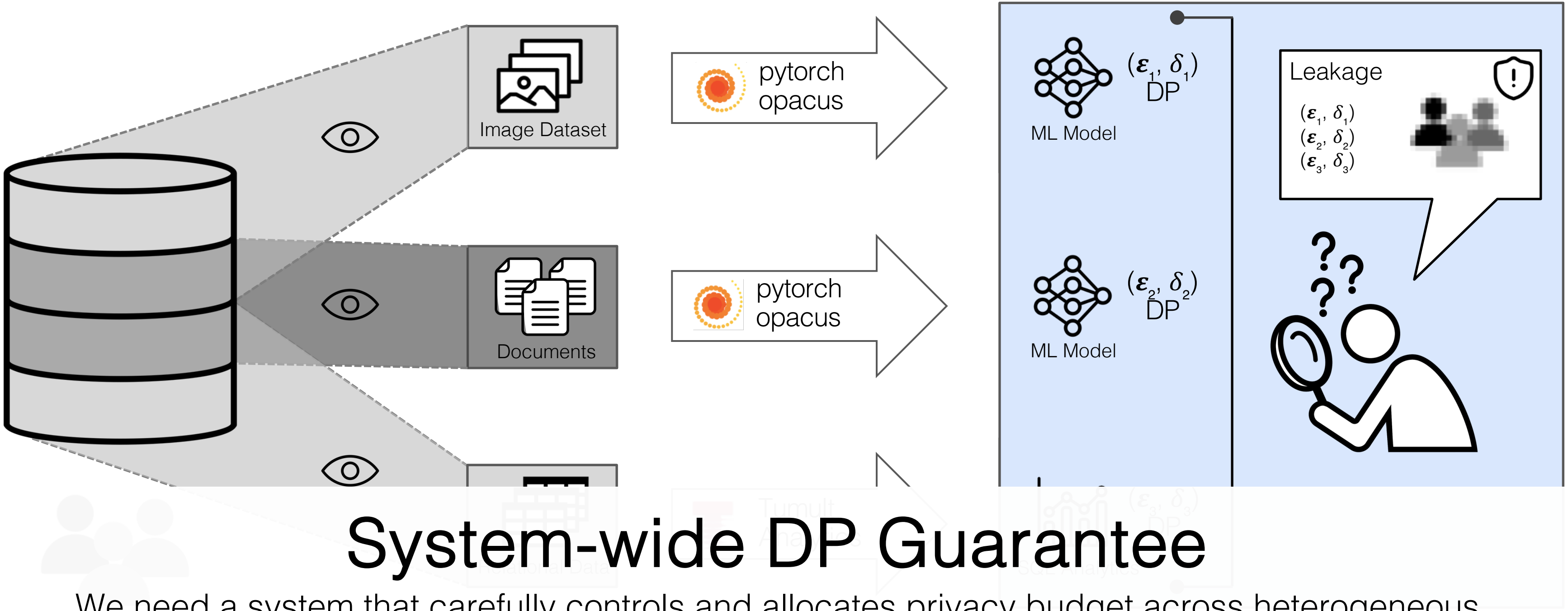
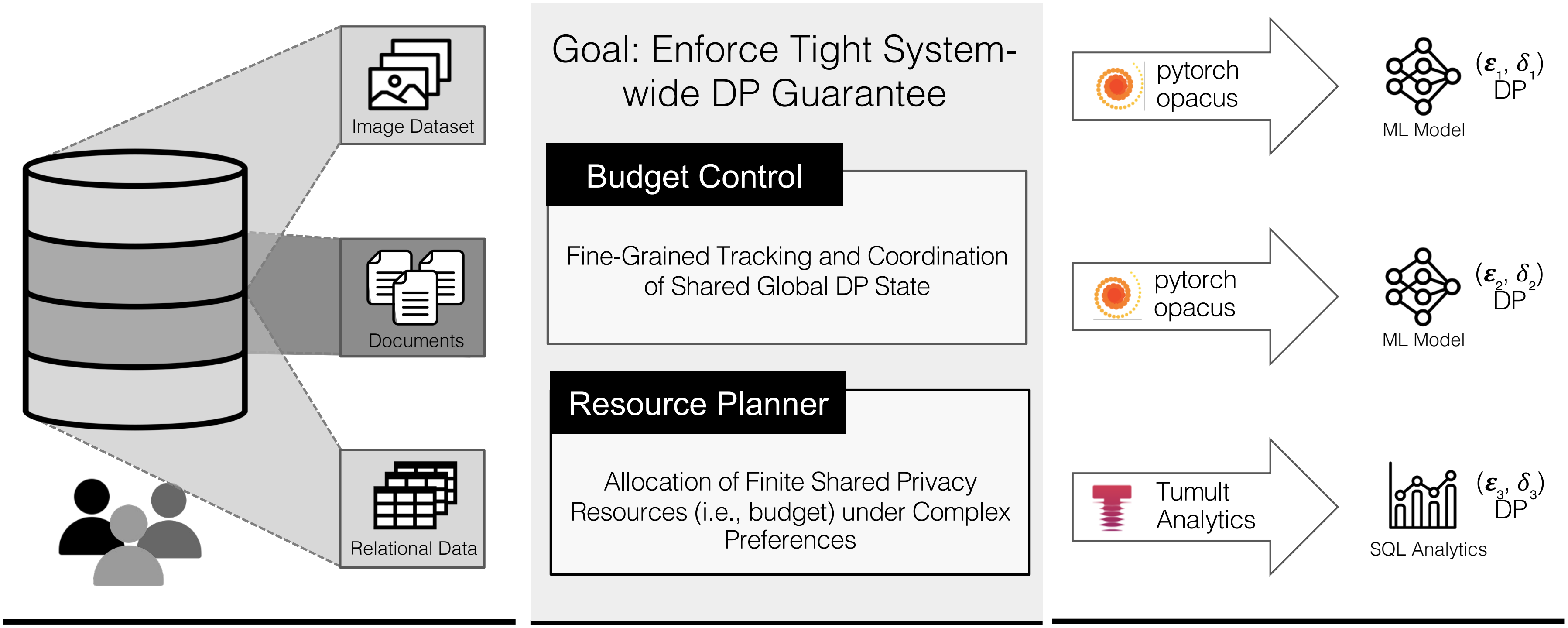


Image Dataset
Documents
pytorch
opacus
pytorch
opacus
ML Model
$(\varepsilon_1, \delta_1)$ DP
ML Model
$(\varepsilon_2, \delta_2)$ DP
Leakage
$(\varepsilon_1, \delta_1)$
$(\varepsilon_2, \delta_2)$
$(\varepsilon_3, \delta_3)$
?
?
?

System-wide DP Guarantee

We need a system that carefully controls and allocates privacy budget across heterogeneous applications and data systems over time.

Cohere: Unified System Architecture for DP



Goal: Enforce Tight System-wide DP Guarantee

Budget Control

Fine-Grained Tracking and Coordination of Shared Global DP State

Resource Planner

Allocation of Finite Shared Privacy Resources (i.e., budget) under Complex Preferences

Data Layer

DP Management Layer

Application Layer

Challenges: System-wide Privacy Guarantee

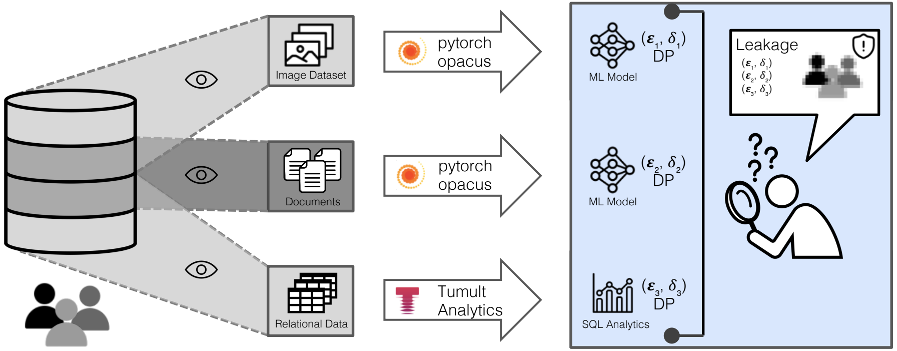


Image Dataset
Documents
Relational Data
pytorch opacus
pytorch opacus
Tumult Analytics
$(\varepsilon_1, \delta_1)$ DP
ML Model
$(\varepsilon_2, \delta_2)$ DP
ML Model
$(\varepsilon_3, \delta_3)$ DP
SQL Analytics
Leakage
$(\varepsilon_1, \delta_1)$
$(\varepsilon_2, \delta_2)$
$(\varepsilon_3, \delta_3)$
???

Challenges: System-wide Privacy Guarantee

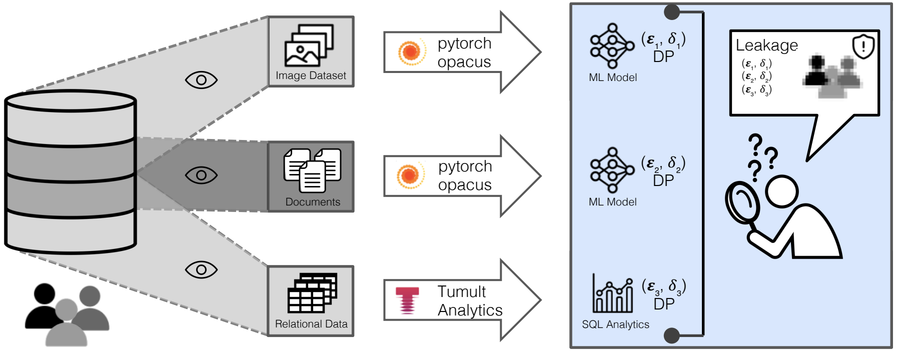
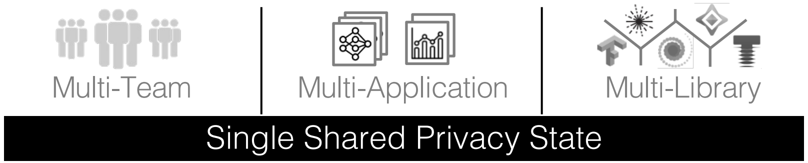


Image Dataset
Documents
Relational Data
pytorch opacus
pytorch opacus
Tumult Analytics
$(\varepsilon_1, \delta_1)$ DP
ML Model
$(\varepsilon_2, \delta_2)$ DP
ML Model
$(\varepsilon_3, \delta_3)$ DP
SQL Analytics
Leakage
$(\varepsilon_1, \delta_1)$
$(\varepsilon_2, \delta_2)$
$(\varepsilon_3, \delta_3)$
??

1. Coordination Problem



Challenges: System-wide Privacy Guarantee

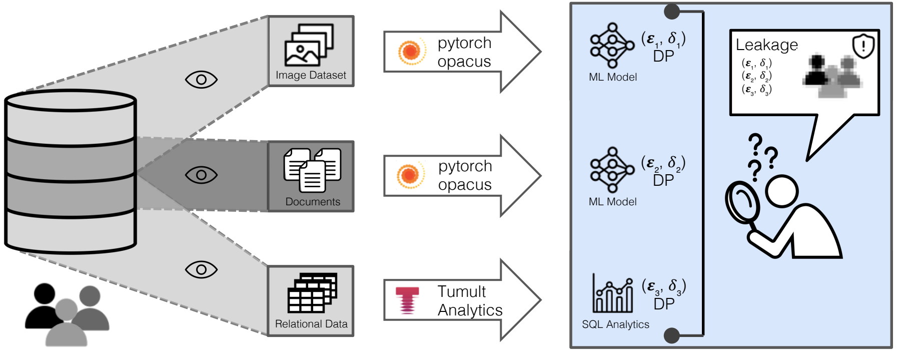
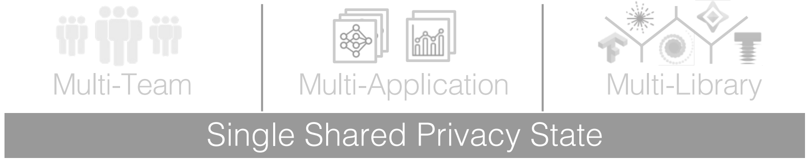


Image Dataset
Documents
Relational Data
pytorch opacus
pytorch opacus
Tumult Analytics
ML Model
$(\varepsilon_1, \delta_1)$ DP
ML Model
$(\varepsilon_2, \delta_2)$ DP
SQL Analytics
$(\varepsilon_3, \delta_3)$ DP
Leakage
$(\varepsilon_1, \delta_1)$
$(\varepsilon_2, \delta_2)$
$(\varepsilon_3, \delta_3)$
???

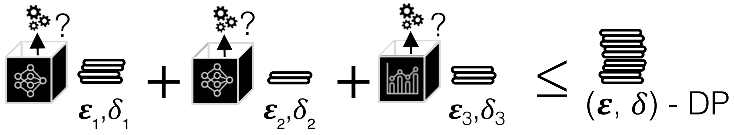
1. Coordination Problem



Multi-Team | Multi-Application | Multi-Library

Single Shared Privacy State

2. Composition Complexity



ε_1,δ_1 + ε_2,δ_2 + ε_3,δ_3 $\leq$ (ε, δ) - DP

Challenges: System-wide Privacy Guarantee

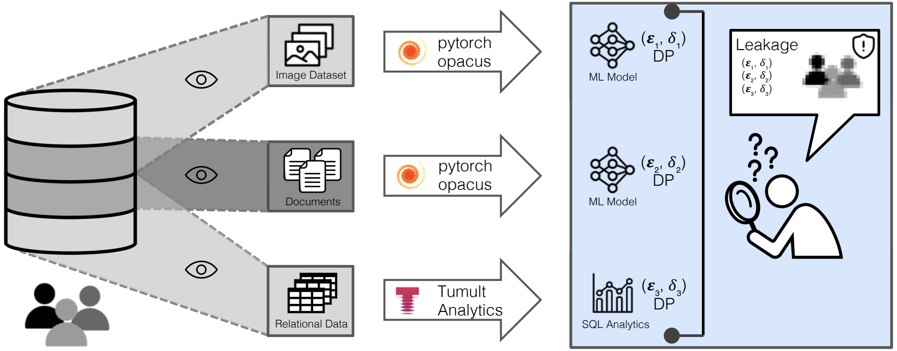
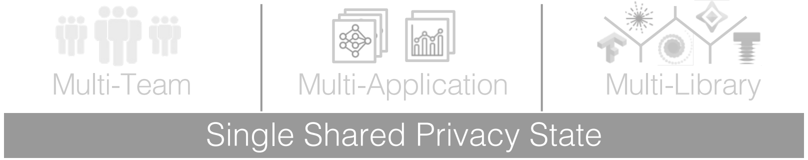
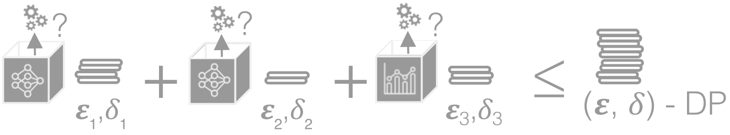


Image Dataset
Documents
Relational Data
pytorch opacus
pytorch opacus
Tumult Analytics
$(\varepsilon_1, \delta_1)$ DP
ML Model
$(\varepsilon_2, \delta_2)$ DP
ML Model
$(\varepsilon_3, \delta_3)$ DP
SQL Analytics
Leakage
$(\varepsilon_1, \delta_1)$
$(\varepsilon_2, \delta_2)$
$(\varepsilon_3, \delta_3)$

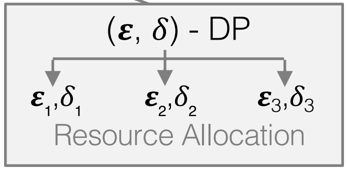
1. Coordination Problem



2. Composition Complexity



3. Scarce and Finite Resource



Challenges: System-wide Privacy Guarantee

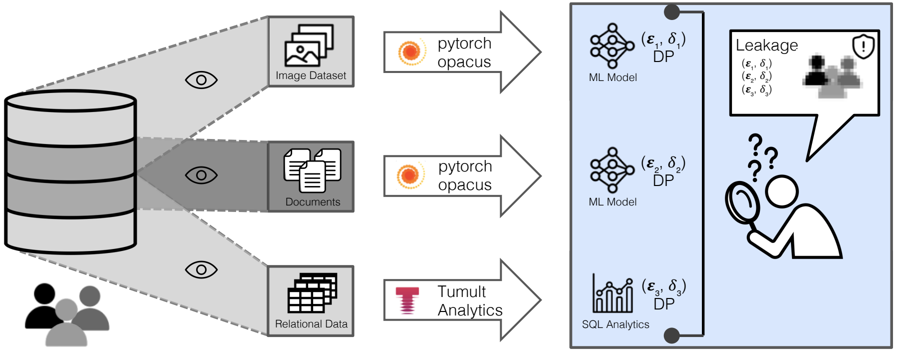
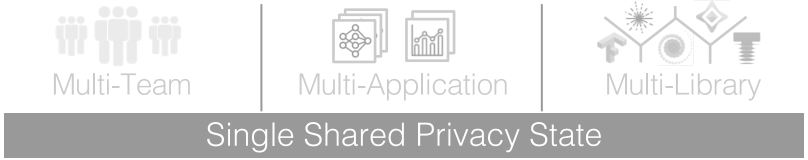


Image Dataset
Documents
Relational Data
pytorch opacus
pytorch opacus
Tumult Analytics
ML Model
$(\varepsilon_1, \delta_1)$ DP
ML Model
$(\varepsilon_2, \delta_2)$ DP
SQL Analytics
$(\varepsilon_3, \delta_3)$ DP
Leakage
$(\varepsilon_1, \delta_1)$
$(\varepsilon_2, \delta_2)$
$(\varepsilon_3, \delta_3)$
??

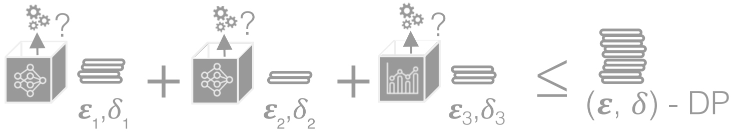
1. Coordination Problem



Multi-Team | Multi-Application | Multi-Library

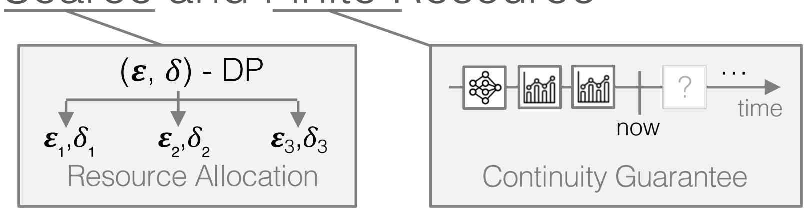
Single Shared Privacy State

2. Composition Complexity



ε_1,δ_1 + ε_2,δ_2 + ε_3,δ_3 ≤ (ε, δ) - DP

3. Scarce and Finite Resource



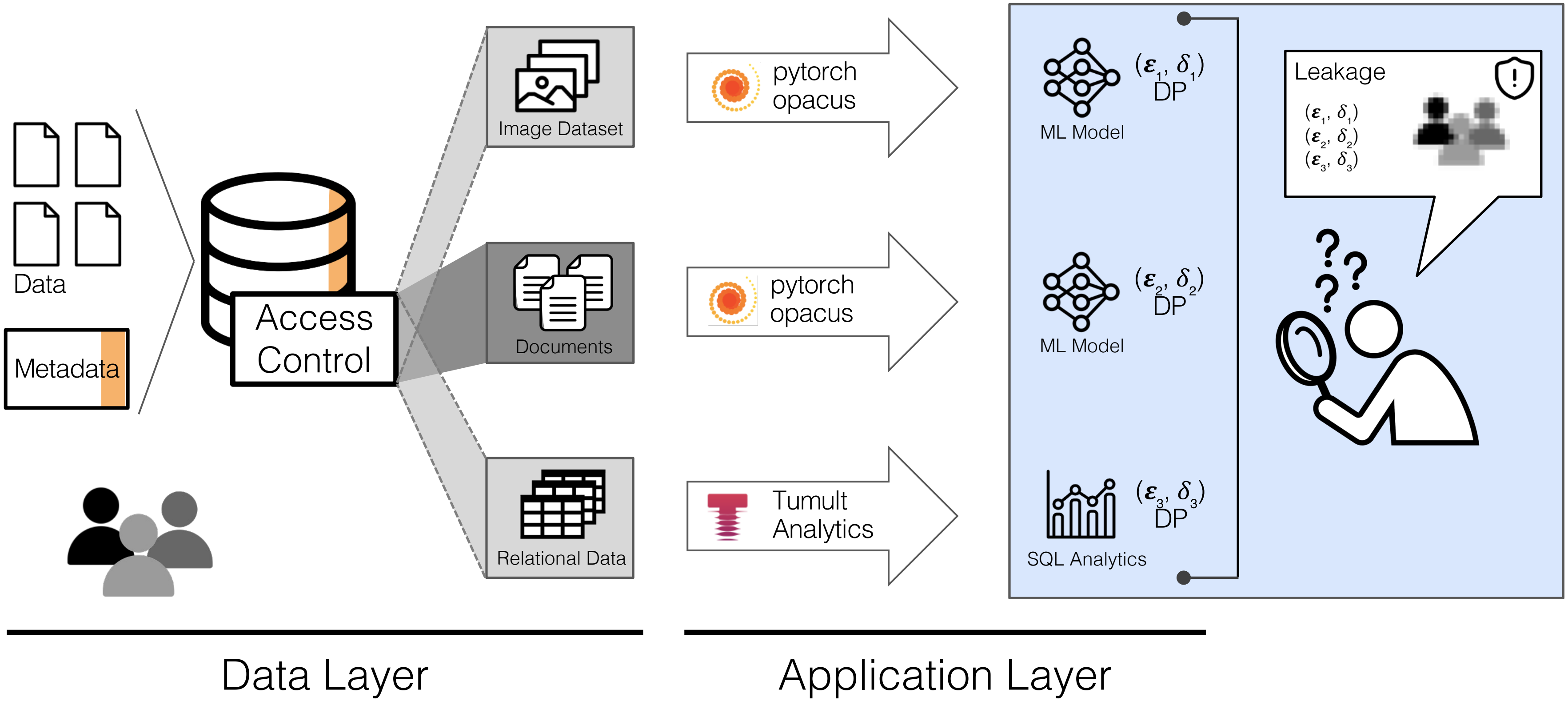
(ε, δ) - DP

ε_1,δ_1 ε_2,δ_2 ε_3,δ_3

Resource Allocation

Continuity Guarantee

Unified System Architecture for DP



Unifying the Application Layer

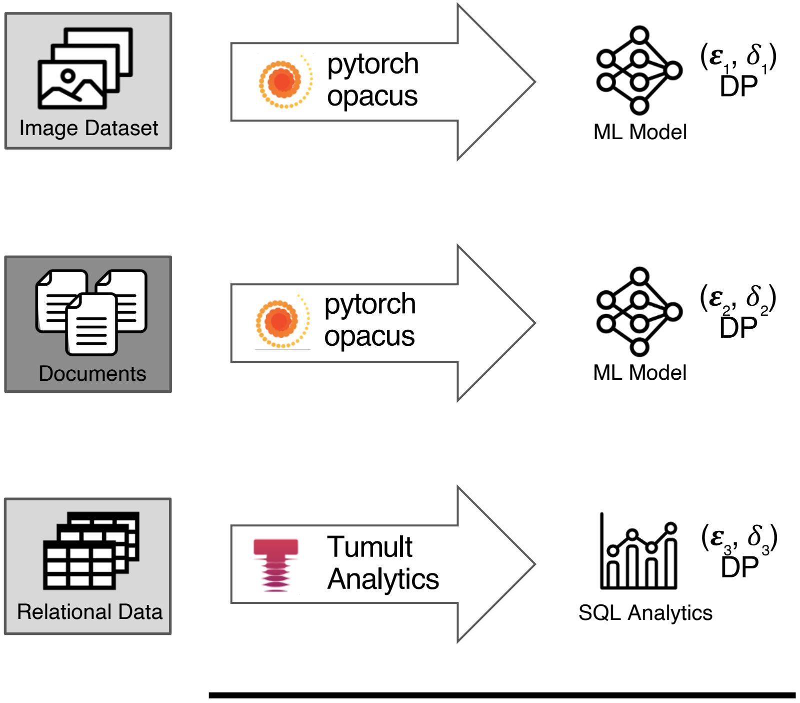


Image Dataset

pytorch opacus

ML Model

$(\boldsymbol{\varepsilon}_1, \delta_1)$ DP

Documents

pytorch opacus

ML Model

$(\boldsymbol{\varepsilon}_2, \delta_2)$ DP

Relational Data

Tumult Analytics

SQL Analytics

$(\boldsymbol{\varepsilon}_3, \delta_3)$ DP

Application Layer

Unifying the Application Layer

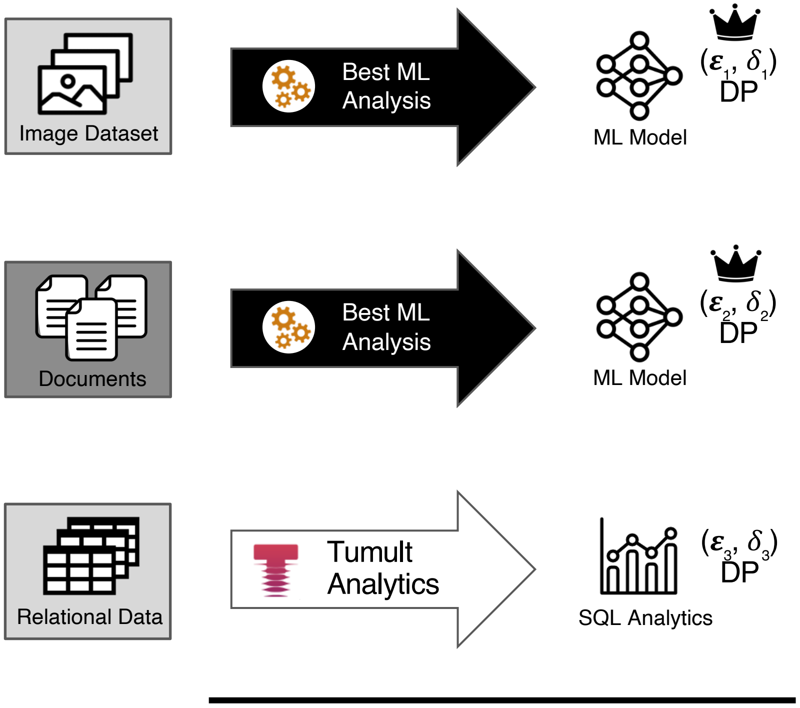


Image Dataset

Best ML Analysis

ML Model

$(\boldsymbol{\varepsilon}_1, \delta_1)$ DP

Documents

Best ML Analysis

ML Model

$(\boldsymbol{\varepsilon}_2, \delta_2)$ DP

Relational Data

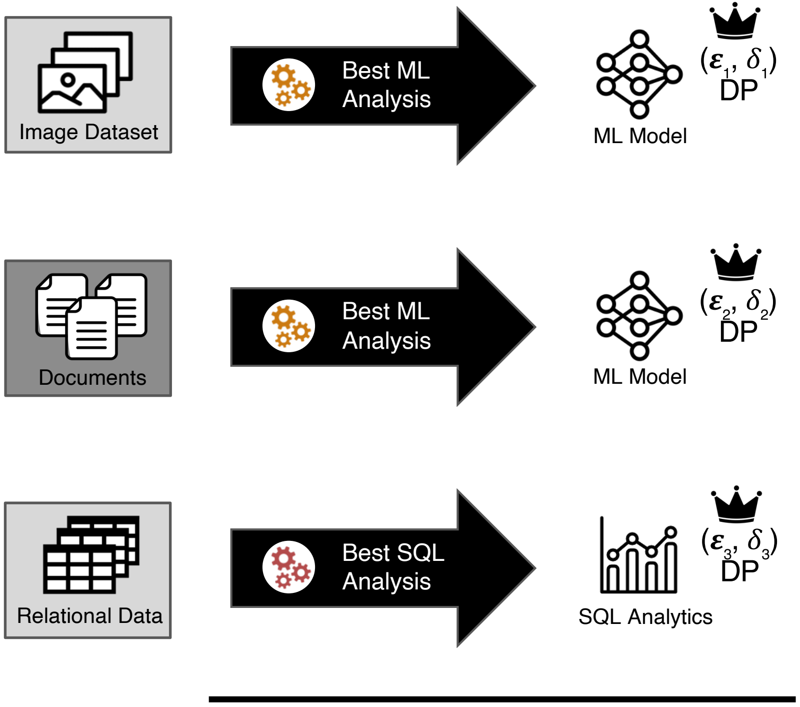
Tumult Analytics

SQL Analytics

$(\boldsymbol{\varepsilon}_3, \delta_3)$ DP

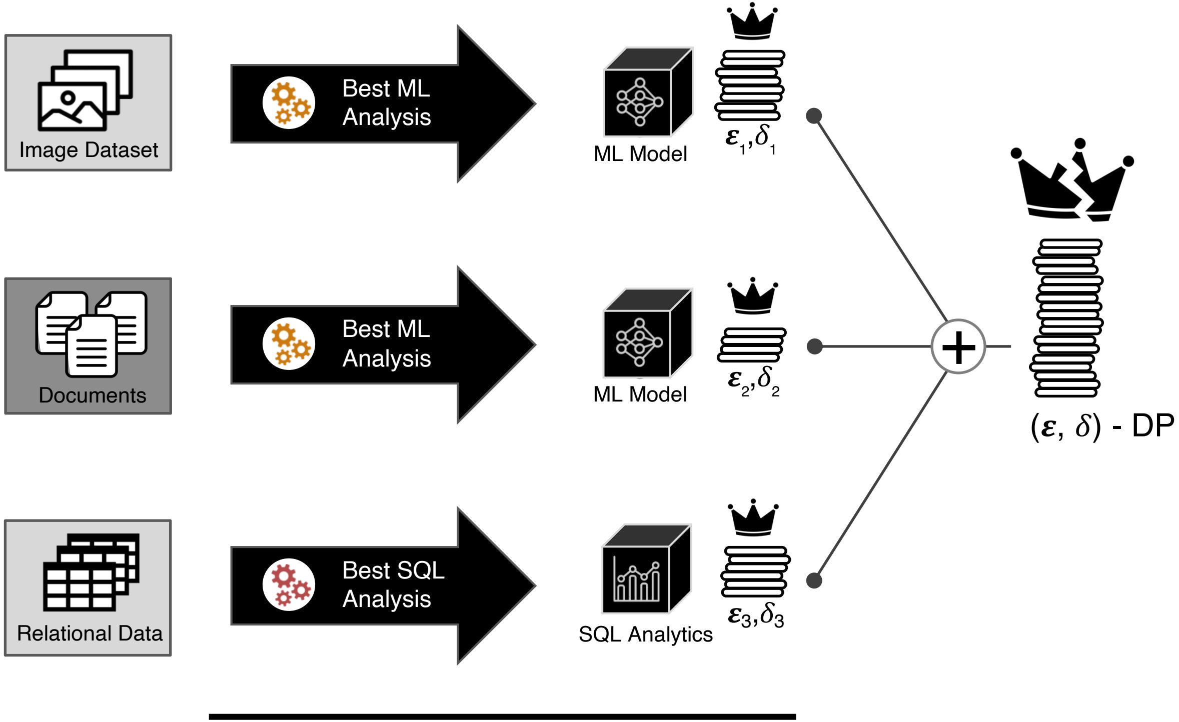
Application Layer

Unifying the Application Layer



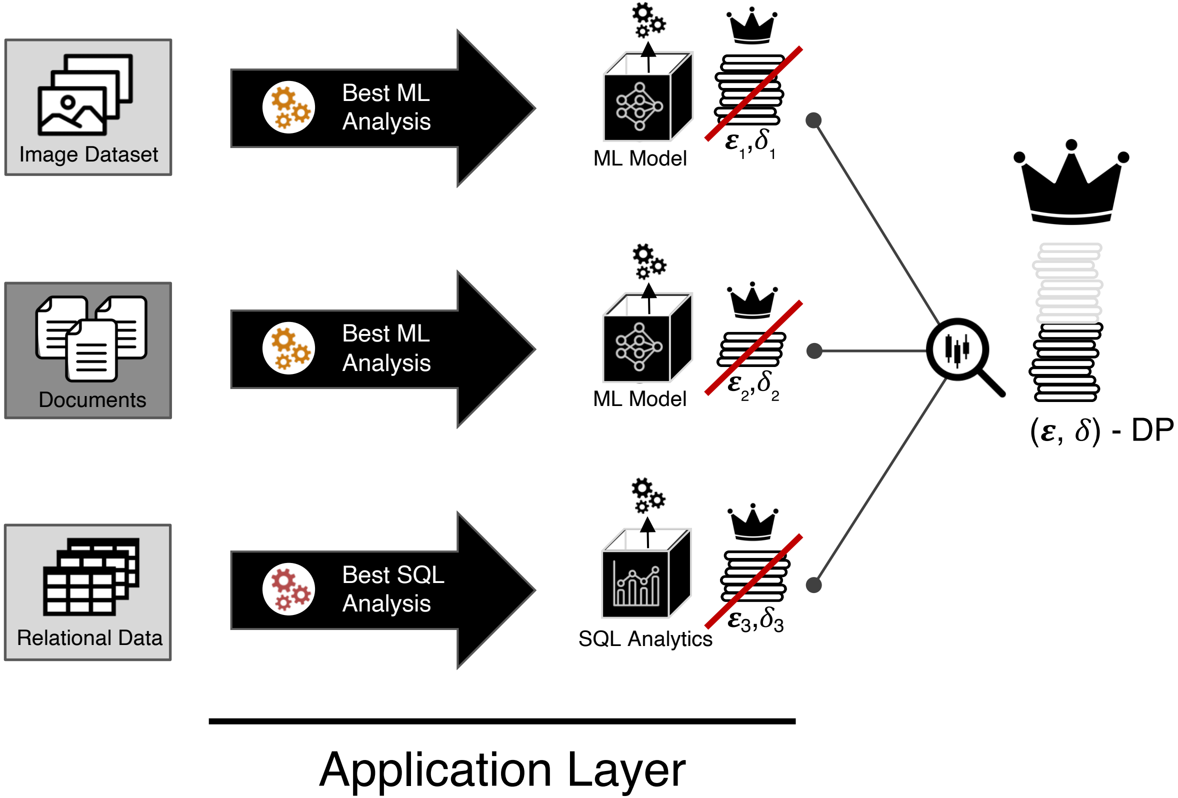
Application Layer

Unifying the Application Layer



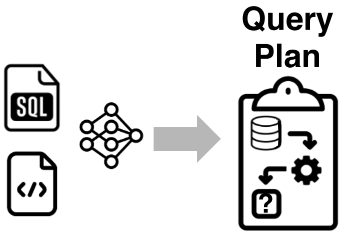
Application Layer

Unifying the Application Layer

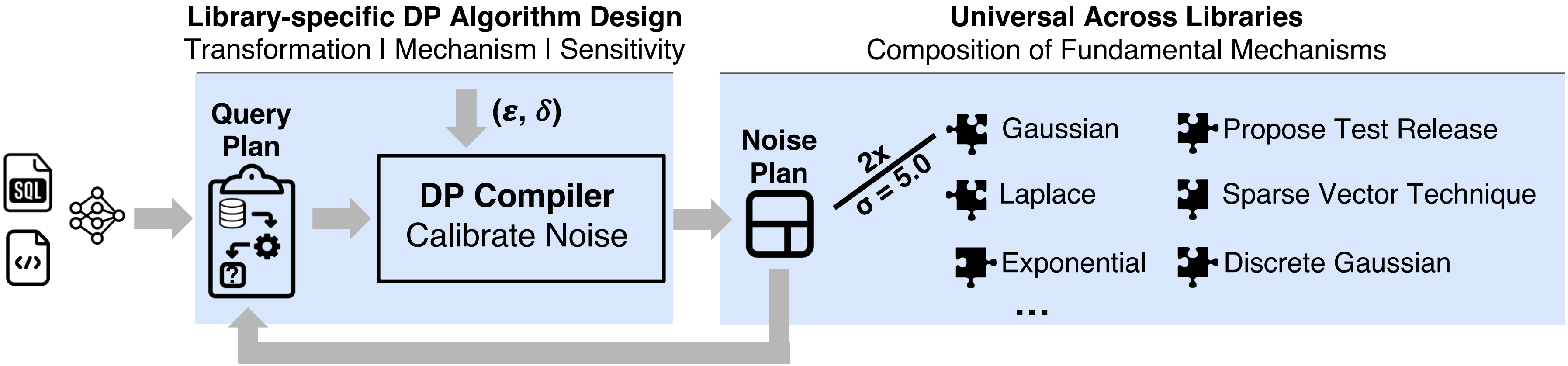


Moving Beyond Local Optima

DP Libraries: In a Nutshell



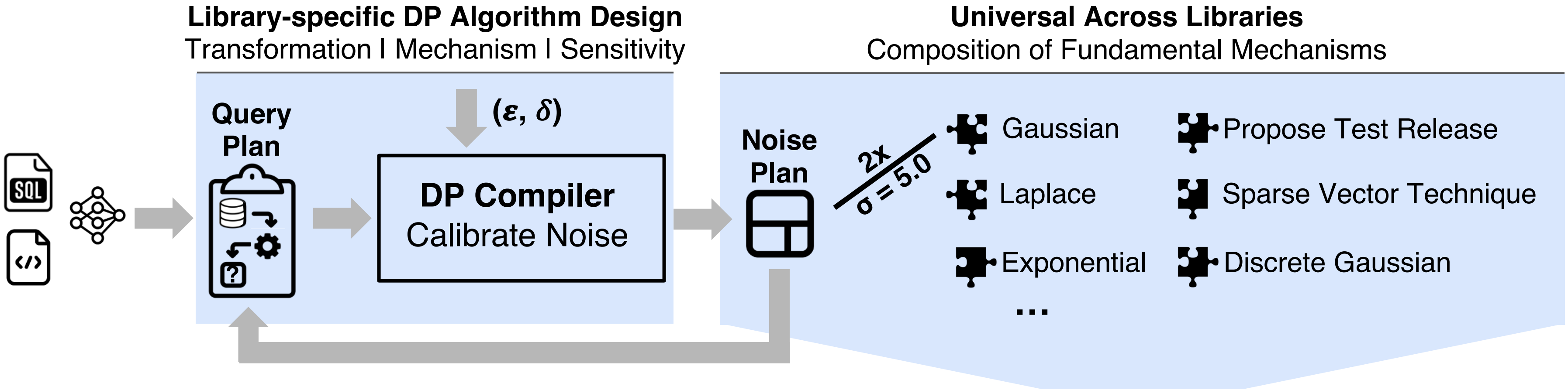
DP Libraries: In a Nutshell



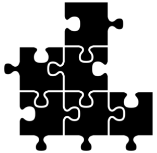
Library-specific DP Algorithm Design
Transformation I Mechanism I Sensitivity

Universal Across Libraries
Composition of Fundamental Mechanisms

DP Libraries: In a Nutshell



If we can compose all fundamental mechanisms, we can support a variety of heterogeneous libraries through a unified noise plan.



Unifying the Application Layer

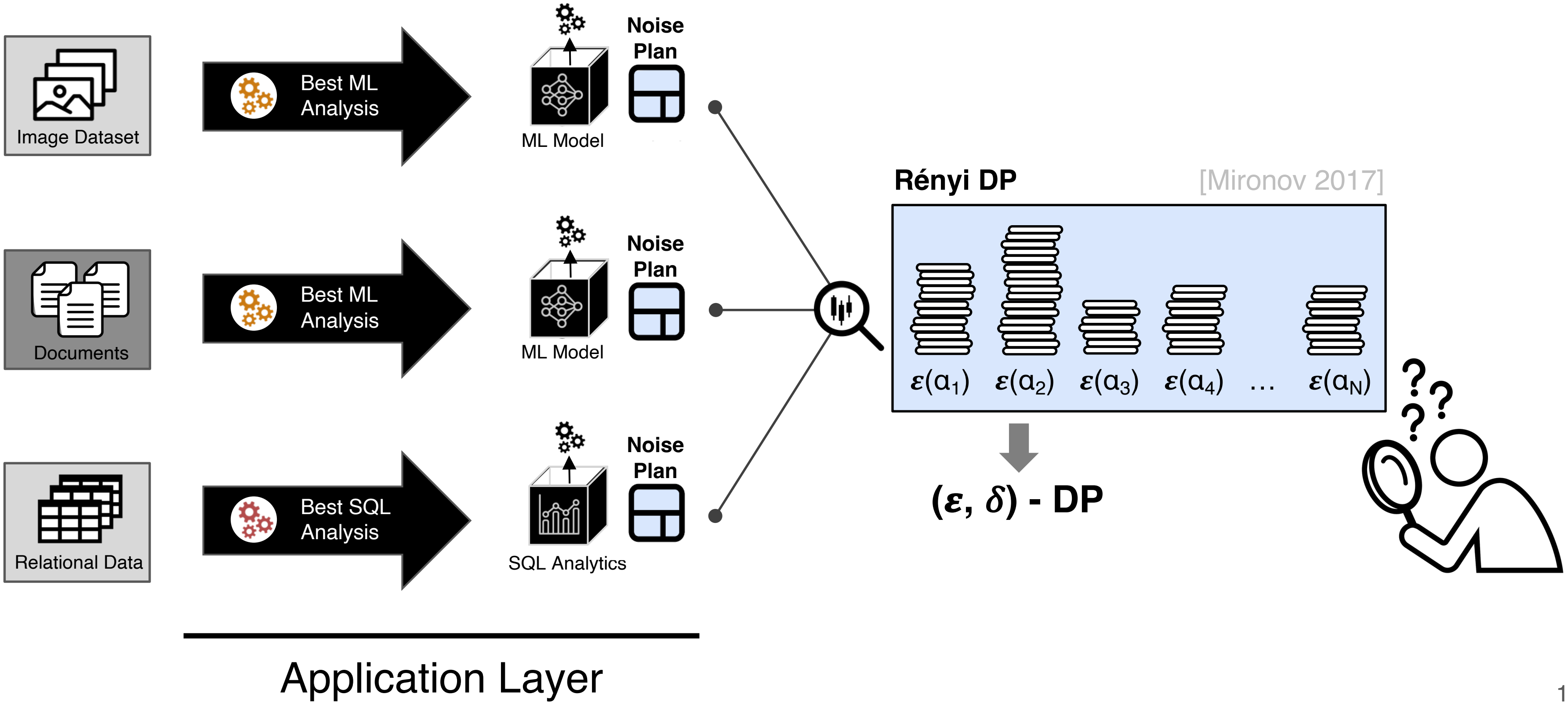


Image Dataset

Best ML Analysis

ML Model

Noise Plan

Documents

Best ML Analysis

ML Model

Noise Plan

Relational Data

Best SQL Analysis

SQL Analytics

Noise Plan

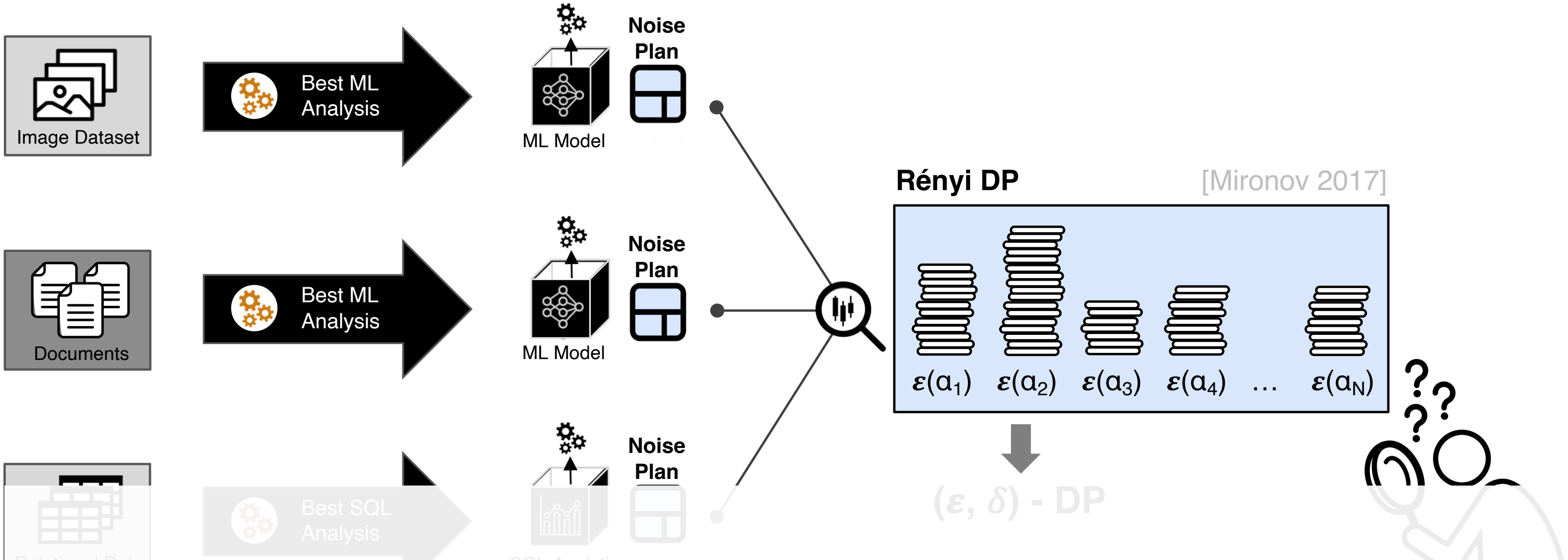
Application Layer

Rényi DP [Mironov 2017]

$\varepsilon(\alpha_1)$ $\varepsilon(\alpha_2)$ $\varepsilon(\alpha_3)$ $\varepsilon(\alpha_4)$ … $\varepsilon(\alpha_N)$

(ε, δ) - DP

Unifying the Application Layer



Is this the best we can do?

Assumptions: All applications are presumed to access every user.

Application Layer

Fine-grained Privacy Analysis

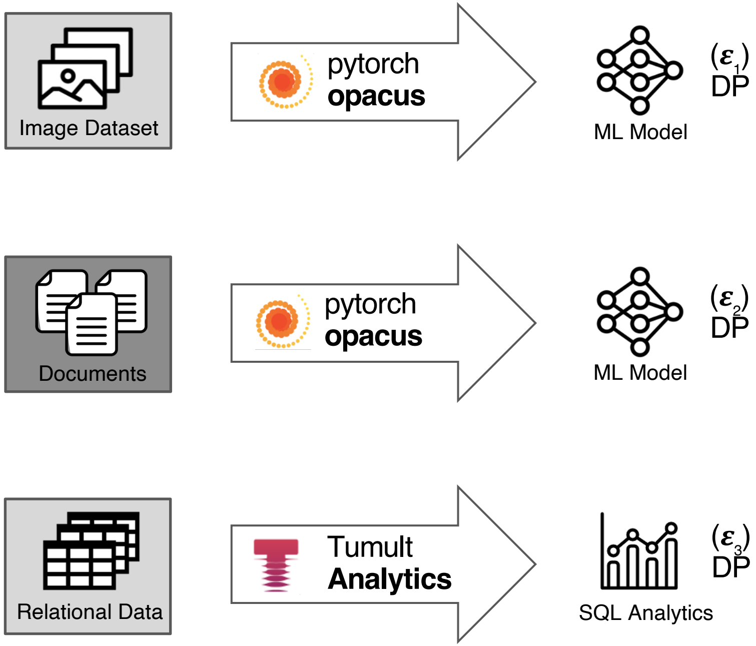
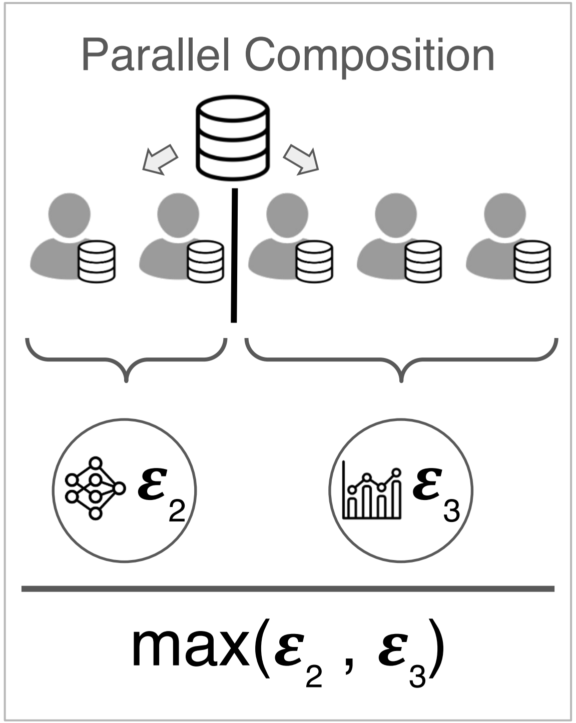


Image Dataset → pytorch **opacus** → ML Model (ε_1) DP

Documents → pytorch **opacus** → ML Model (ε_2) DP

Relational Data → Tumult **Analytics** → SQL Analytics (ε_3) DP



Parallel Composition

ε_2 ε_3

$$\max(\varepsilon_2\ ,\ \varepsilon_3)$$

[McSherry 2009]

Fine-grained Privacy Analysis

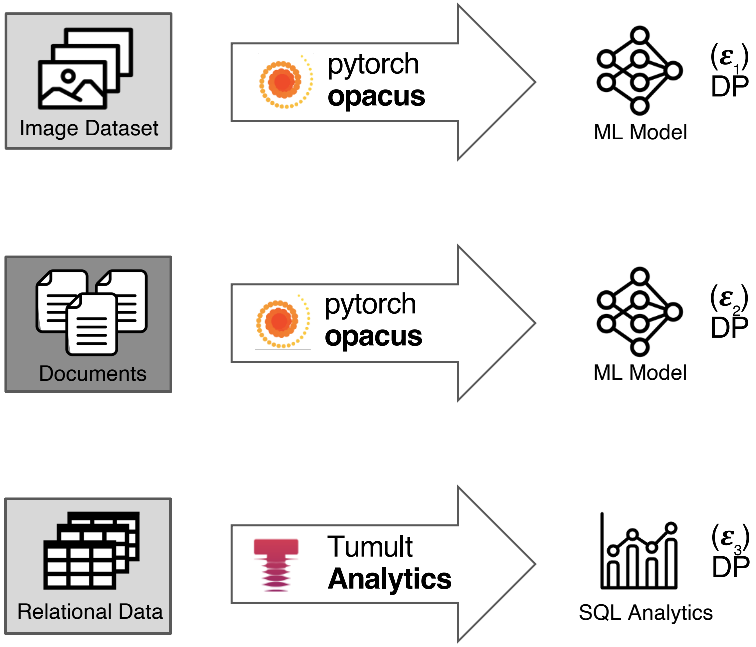
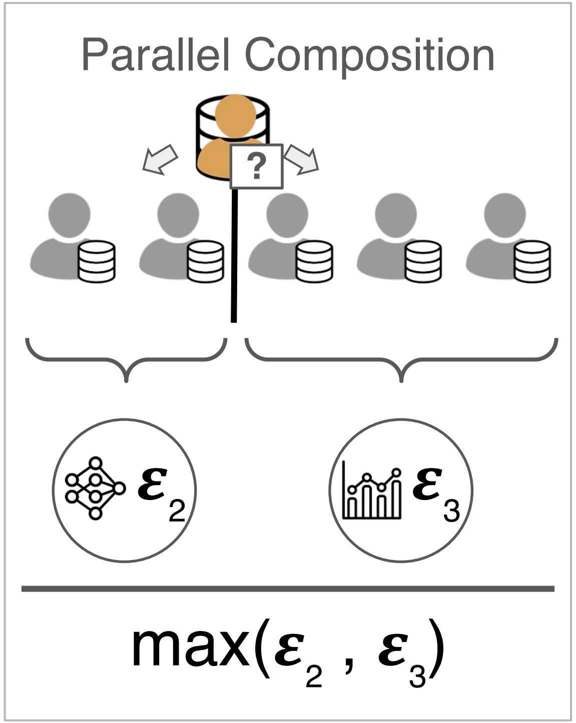


Image Dataset → pytorch **opacus** → ML Model (ε_1) DP

Documents → pytorch **opacus** → ML Model (ε_2) DP

Relational Data → Tumult **Analytics** → SQL Analytics (ε_3) DP



Parallel Composition

$$\max(\varepsilon_2 , \varepsilon_3)$$

[McSherry 2009]

Fine-grained Privacy Analysis

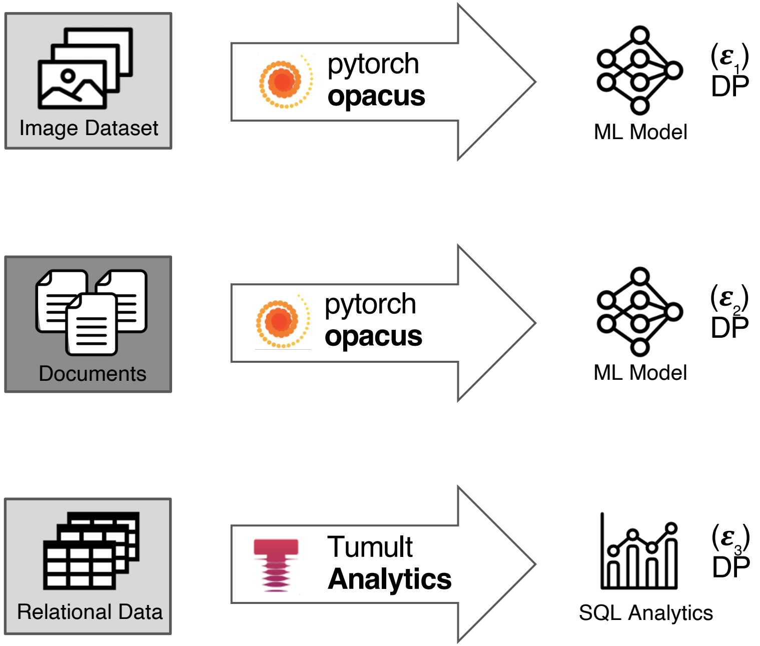
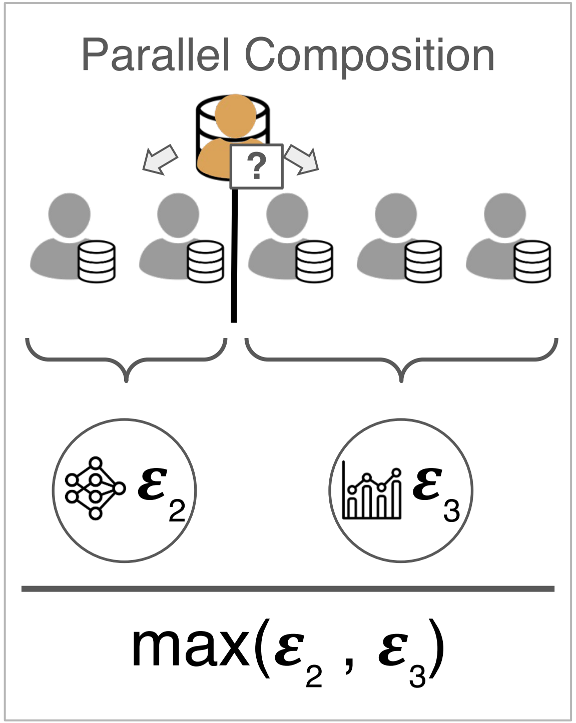


Image Dataset → pytorch **opacus** → ML Model (ε_1) DP

Documents → pytorch **opacus** → ML Model (ε_2) DP

Relational Data → Tumult **Analytics** → SQL Analytics (ε_3) DP



Parallel Composition

?

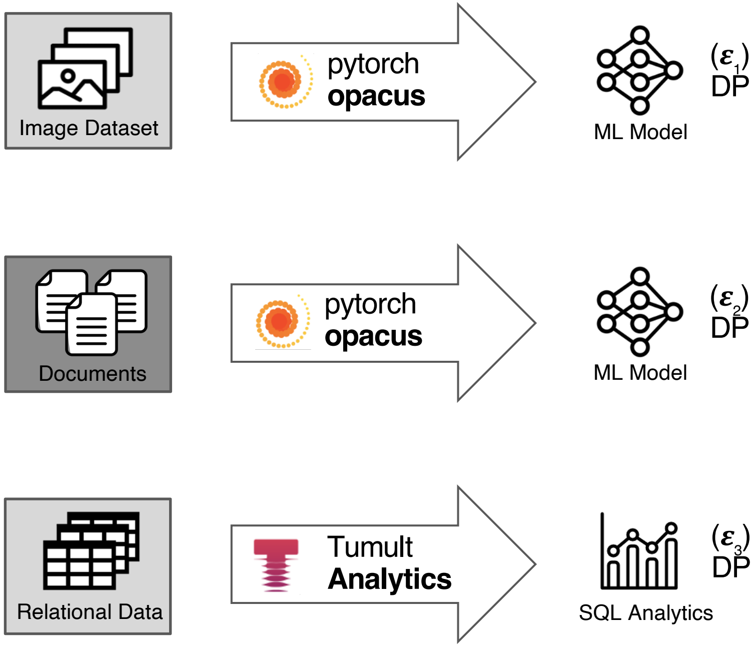
ε_2

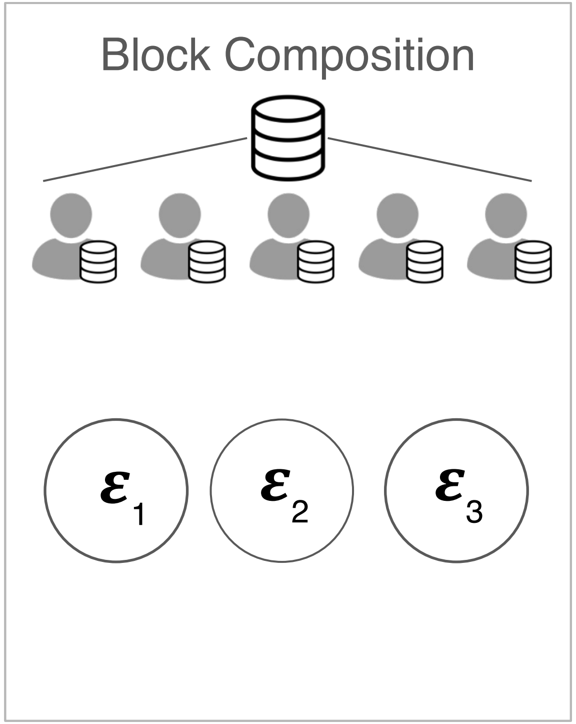
ε_3

$\max(\varepsilon_2 , \varepsilon_3)$

[McSherry 2009]

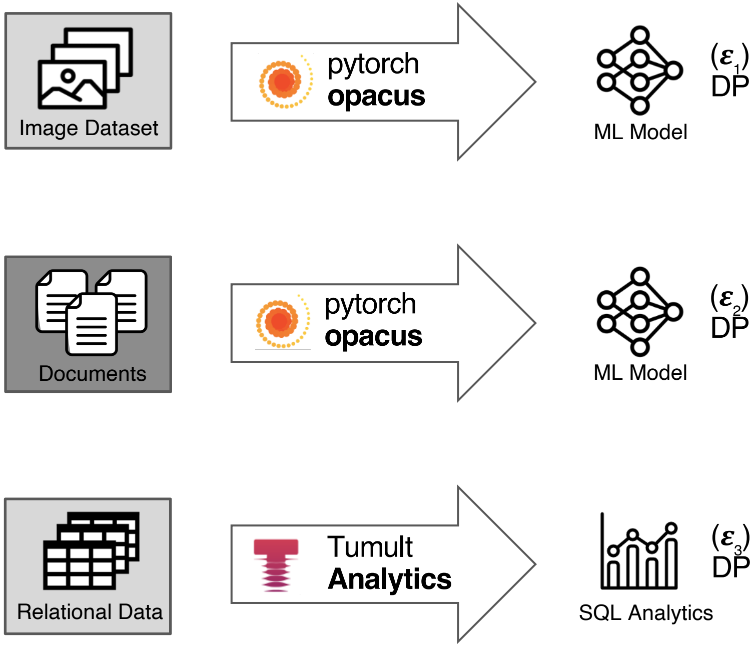
Fine-grained Privacy Analysis

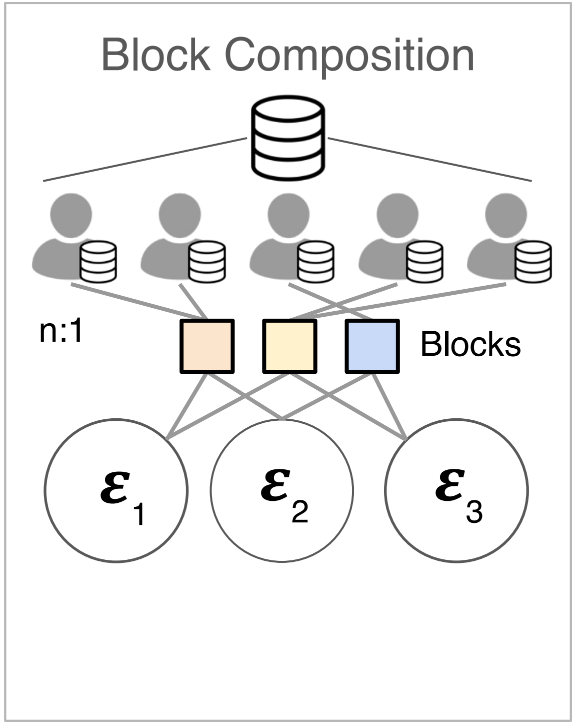




[Lécuyer SOSP'19]

Fine-grained Privacy Analysis





[Lécuyer SOSP’19]

Fine-grained Privacy Analysis

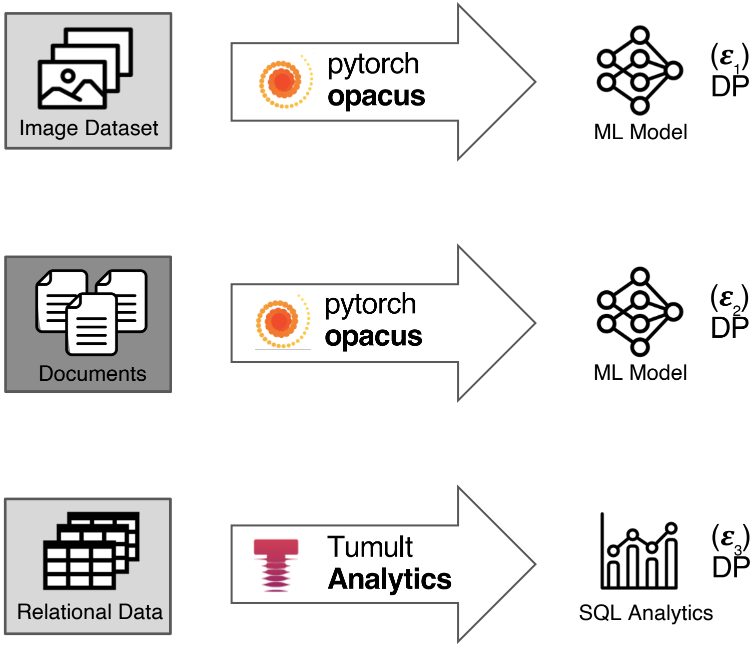


Image Dataset

pytorch **opacus**

ML Model

(ε_1) DP

Documents

pytorch **opacus**

ML Model

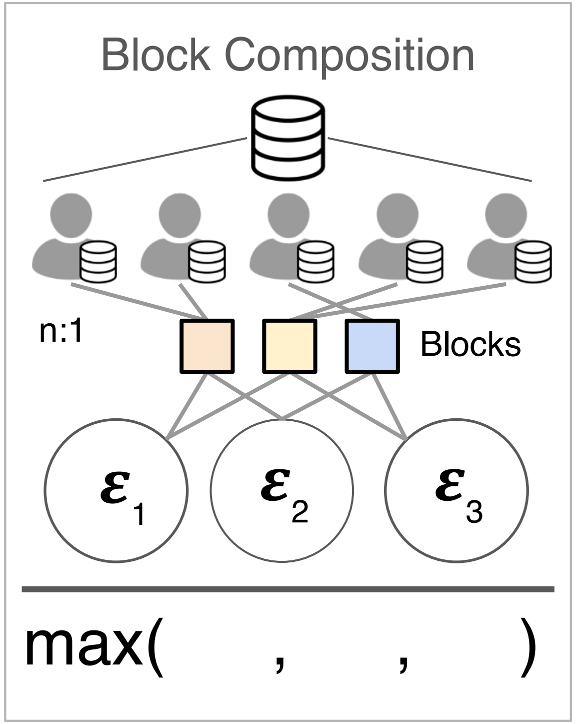
(ε_2) DP

Relational Data

Tumult **Analytics**

SQL Analytics

(ε_3) DP



Block Composition

n:1

Blocks

ε_1 ε_2 ε_3

max(, ,)

[Lécuyer SOSP'19]

Fine-grained Privacy Analysis

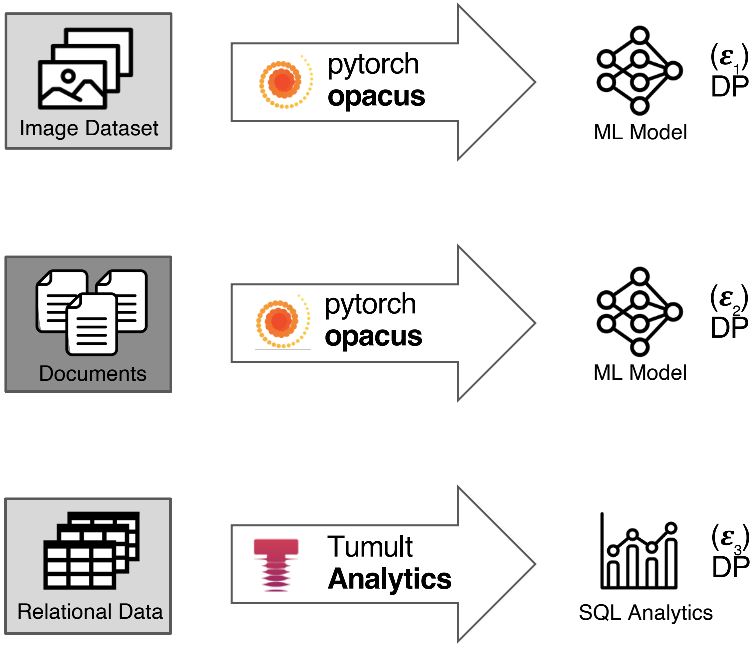


Image Dataset

pytorch **opacus**

ML Model

(ε_1) DP

Documents

pytorch **opacus**

ML Model

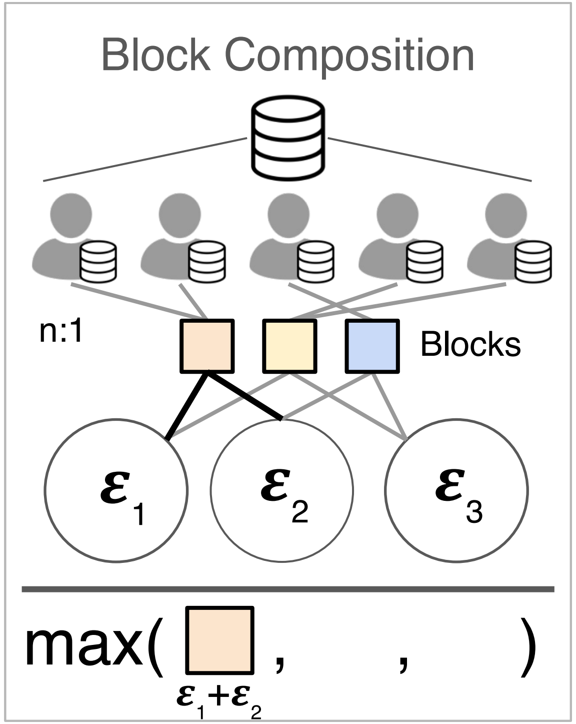
(ε_2) DP

Relational Data

Tumult **Analytics**

SQL Analytics

(ε_3) DP



Block Composition

n:1

Blocks

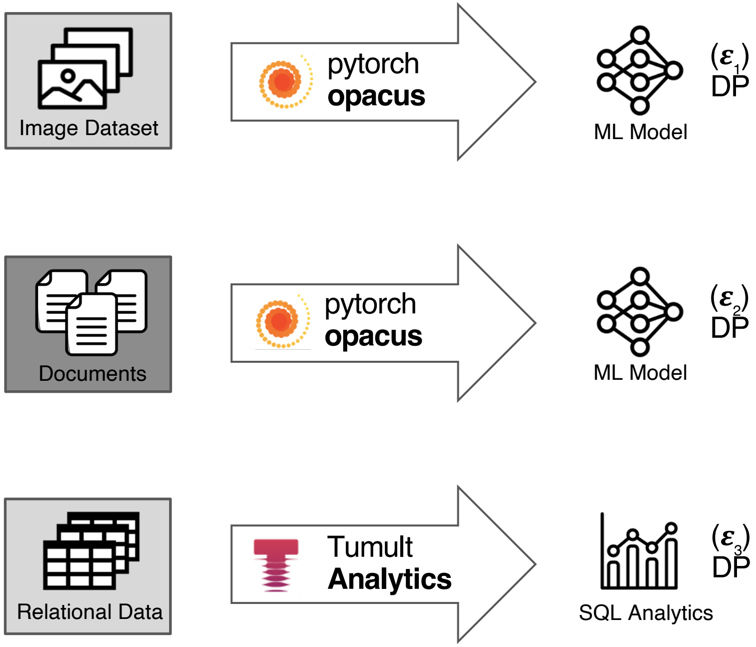
ε_1 ε_2 ε_3

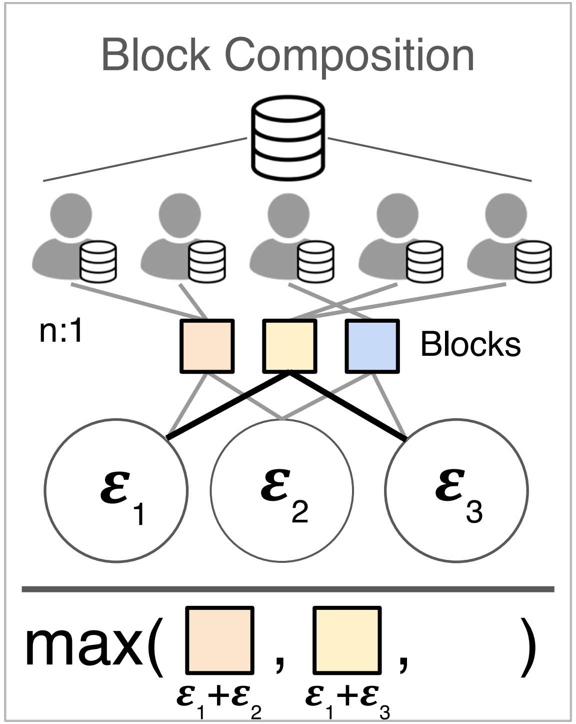
max(□ , ,)

$\varepsilon_1+\varepsilon_2$

[Lécuyer SOSP’19]

Fine-grained Privacy Analysis





[Lécuyer SOSP'19]

Fine-grained Privacy Analysis

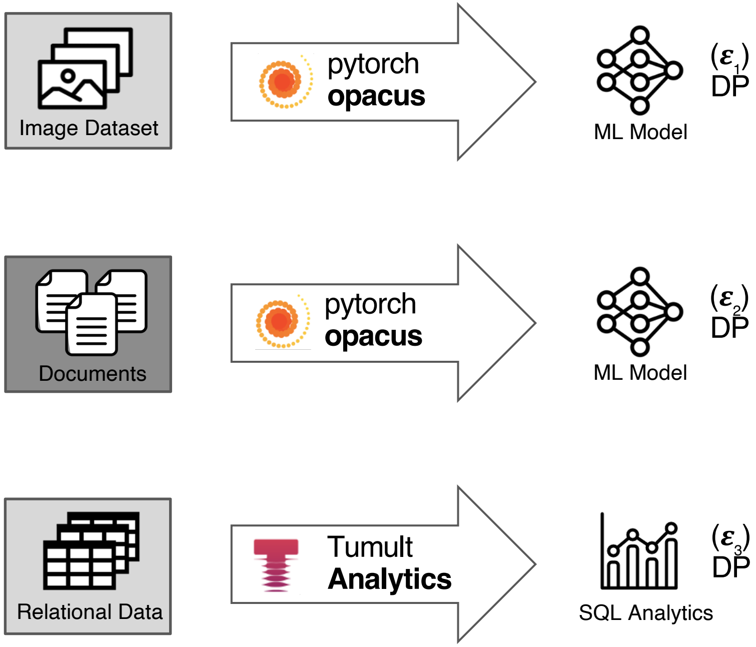


Image Dataset

pytorch **opacus**

ML Model

(ε_1) DP

Documents

pytorch **opacus**

ML Model

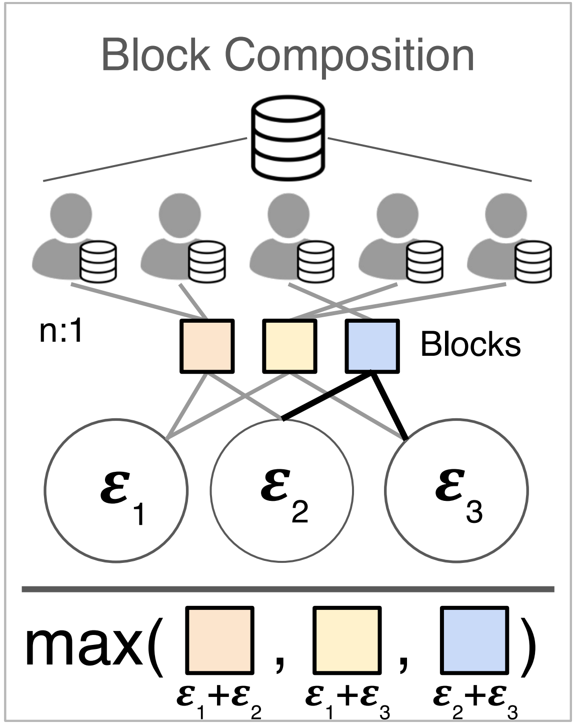
(ε_2) DP

Relational Data

Tumult **Analytics**

SQL Analytics

(ε_3) DP



Block Composition

n:1

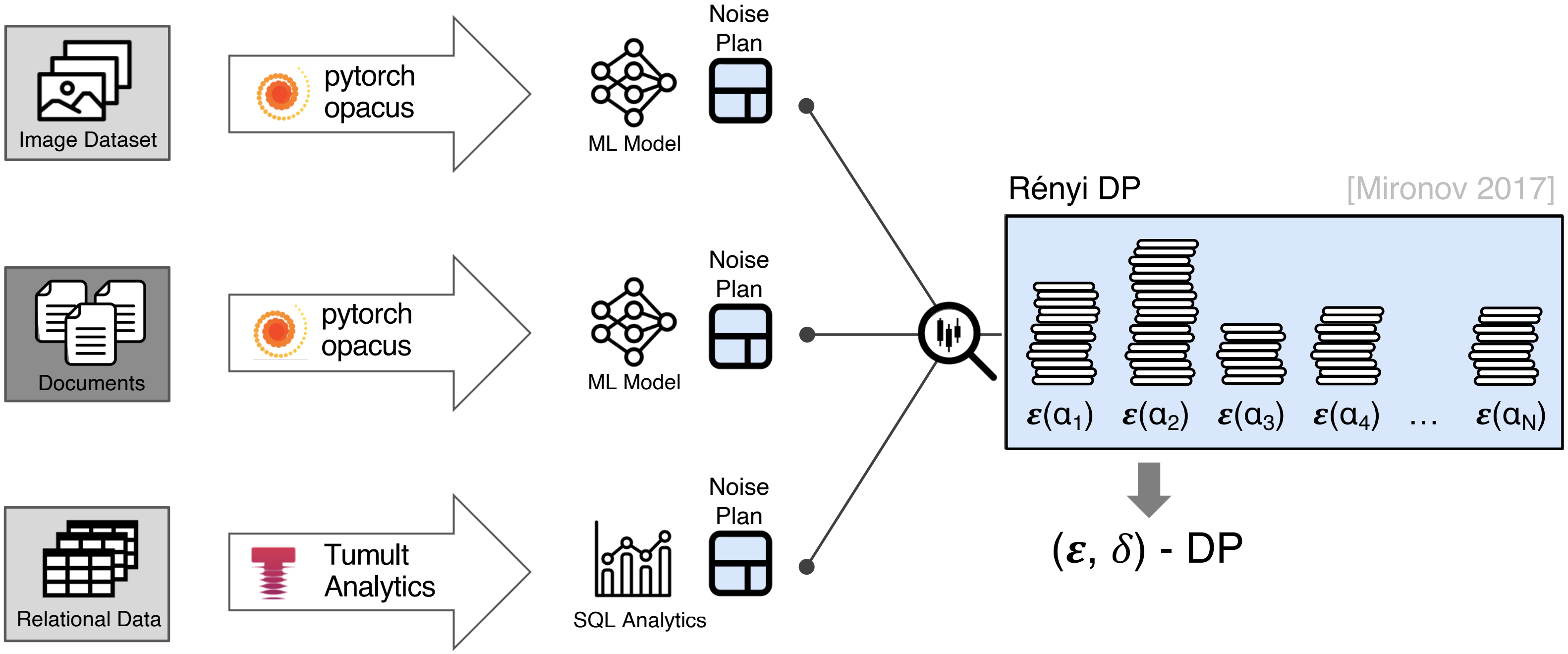
Blocks

ε_1 ε_2 ε_3

$\max(\underset{\varepsilon_1+\varepsilon_2}{\square}, \underset{\varepsilon_1+\varepsilon_3}{\square}, \underset{\varepsilon_2+\varepsilon_3}{\square})$

[Lécuyer SOSP'19]

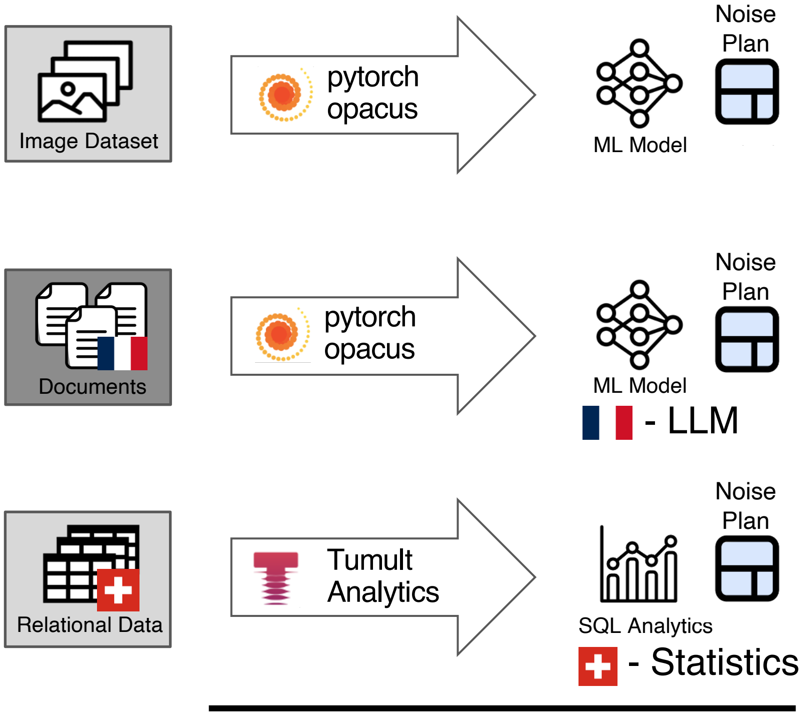
Fine-grained Privacy Analysis



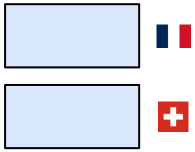
(ε, δ) - DP

Application Layer

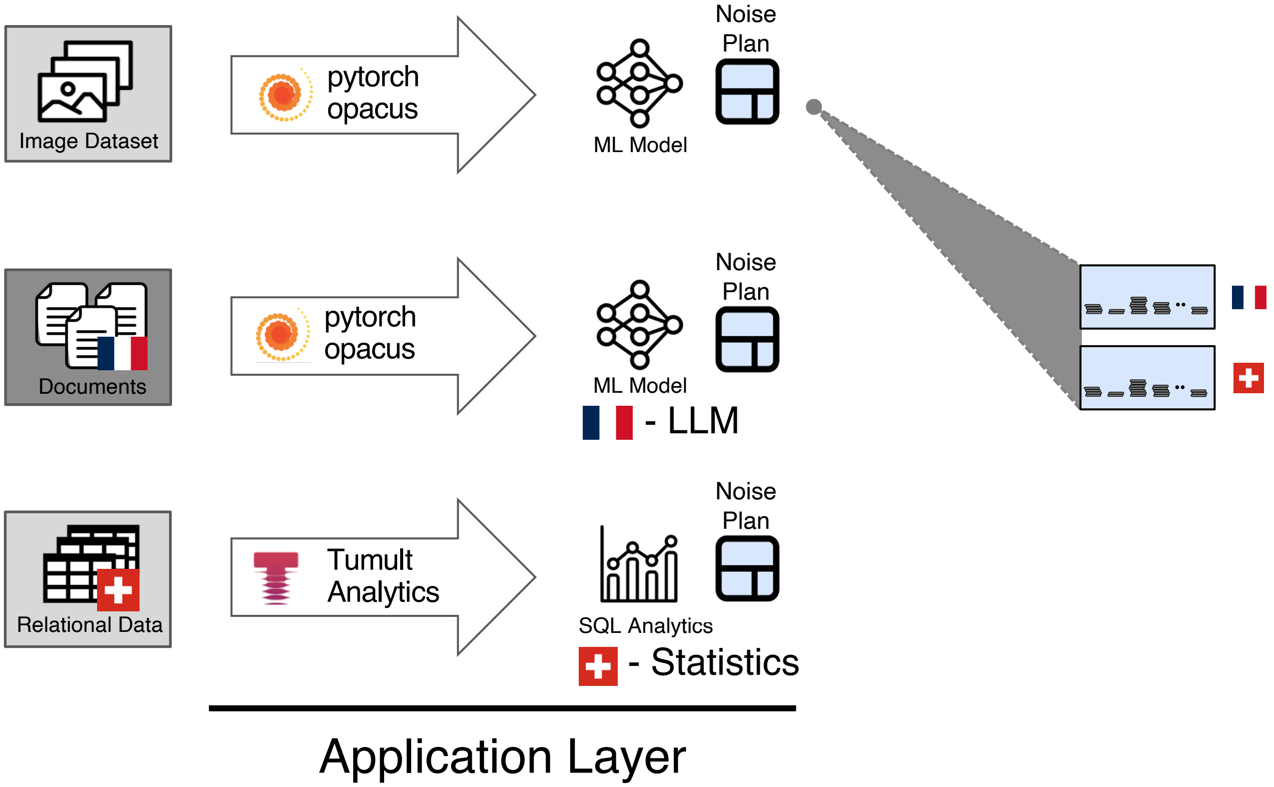
Fine-grained Privacy Analysis



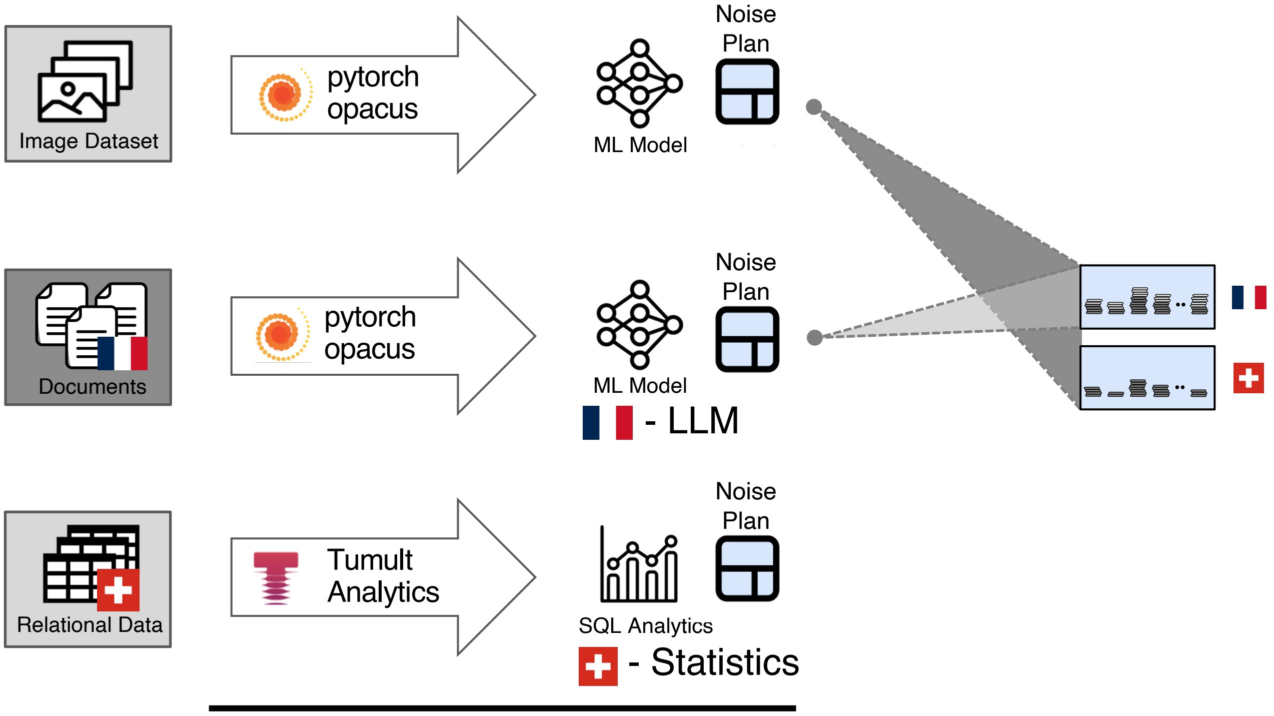
Application Layer



Fine-grained Privacy Analysis

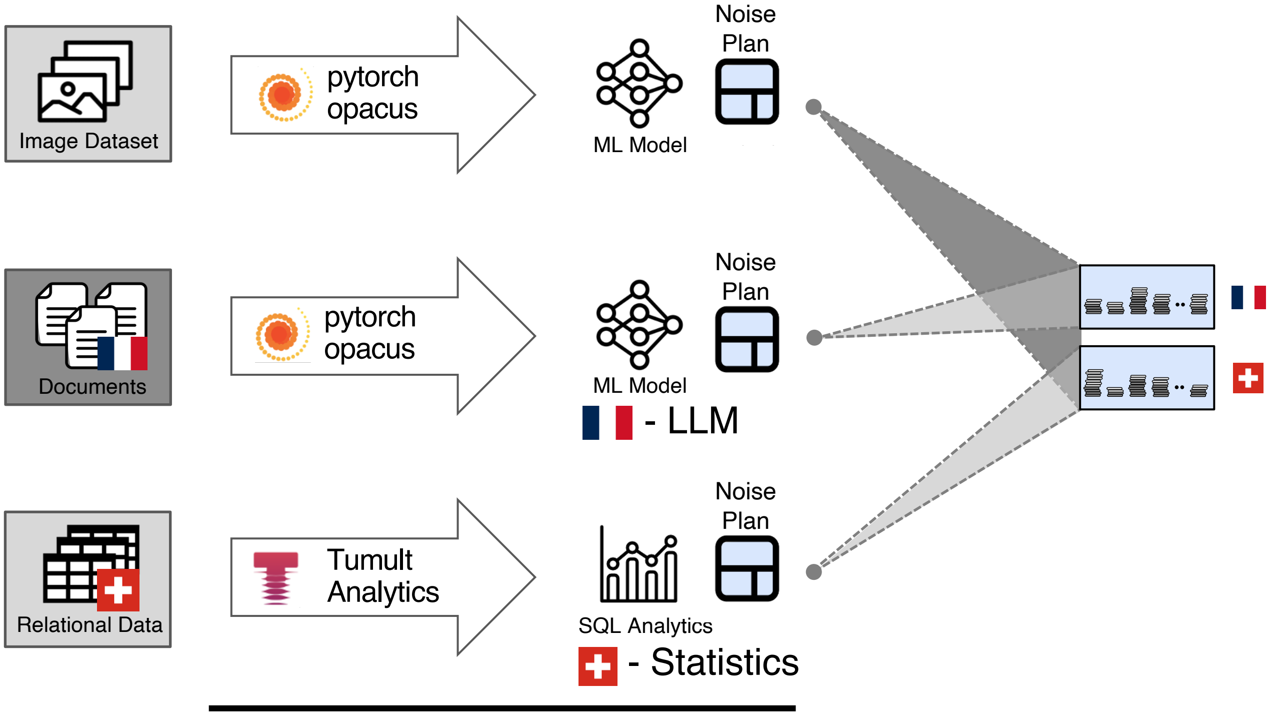


Fine-grained Privacy Analysis



Application Layer

Fine-grained Privacy Analysis



Application Layer

Fine-grained Privacy Analysis

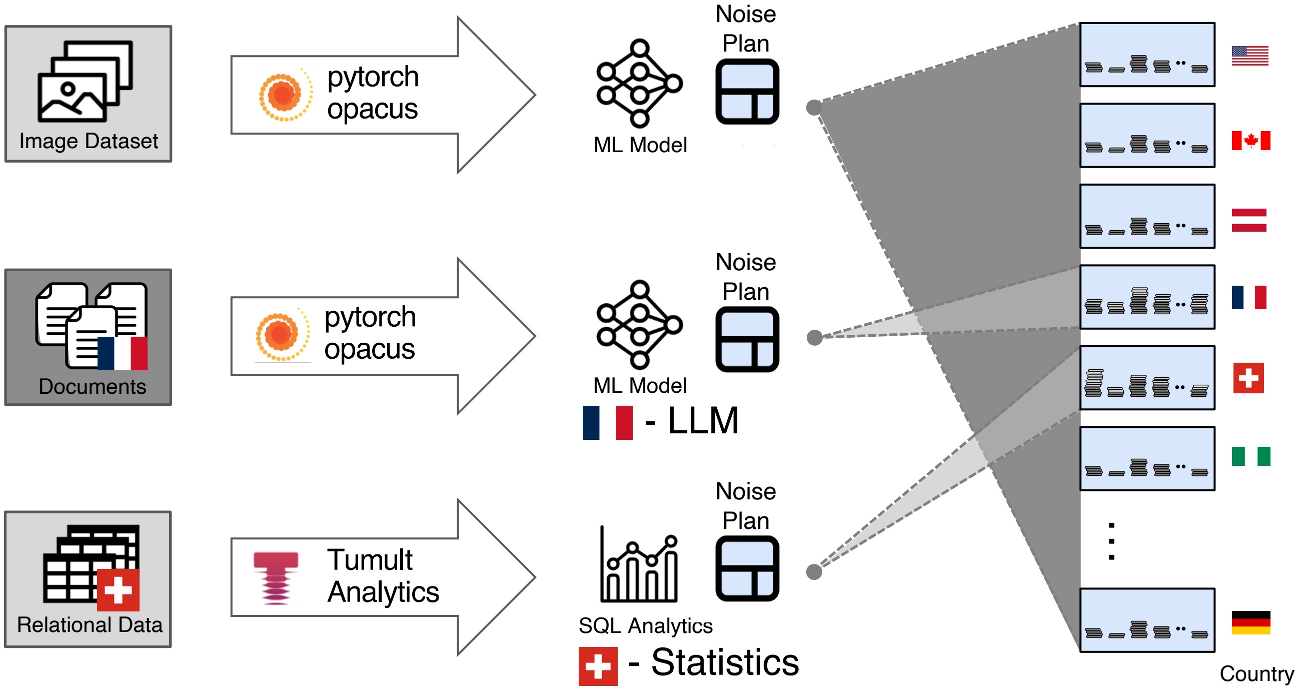


Image Dataset

pytorch opacus

ML Model

Noise Plan

Documents

pytorch opacus

ML Model

Noise Plan

- LLM

Relational Data

Tumult Analytics

SQL Analytics

Noise Plan

- Statistics

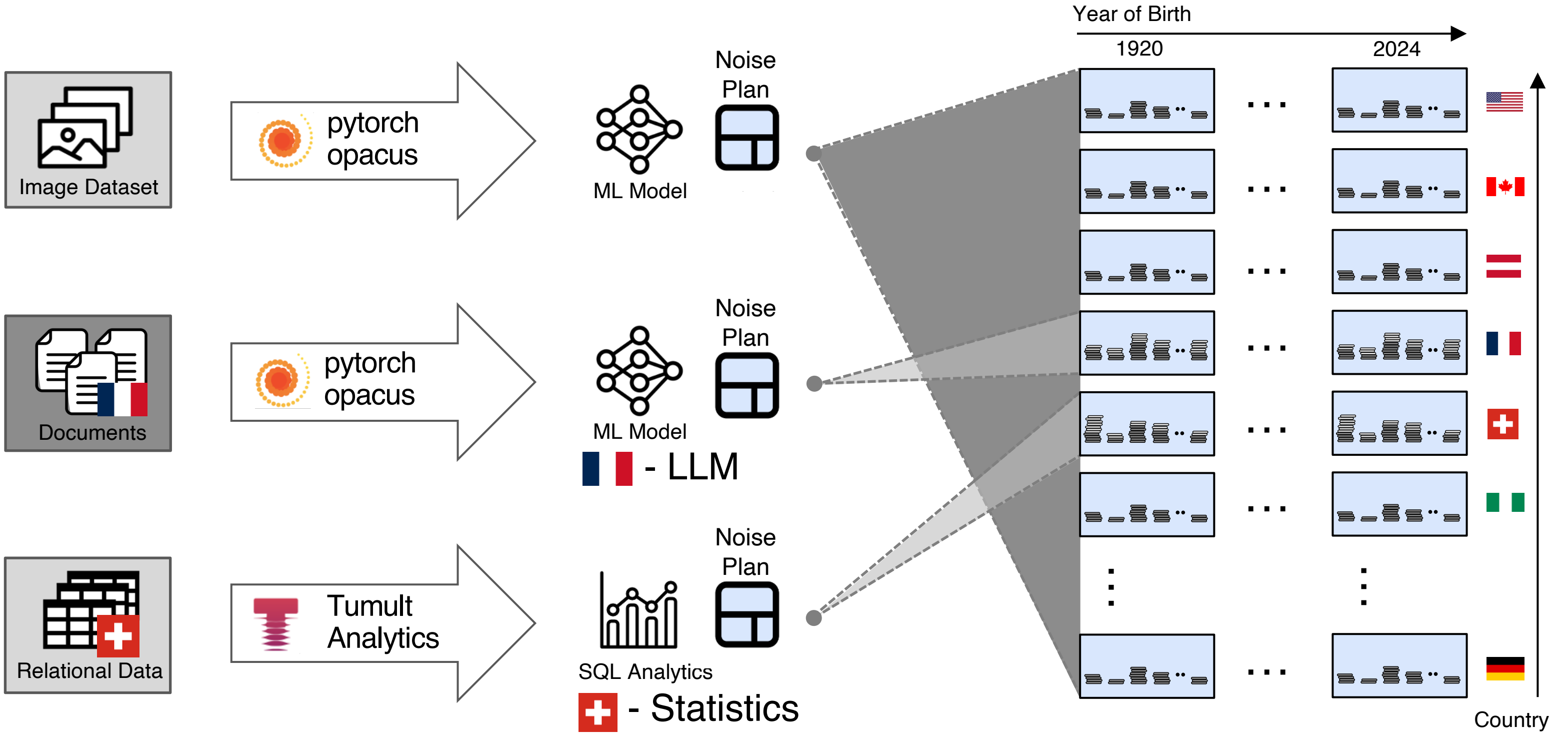
Country

Application Layer

Partitioning Attributes

Schema must be known in advance

Fine-grained Privacy Analysis

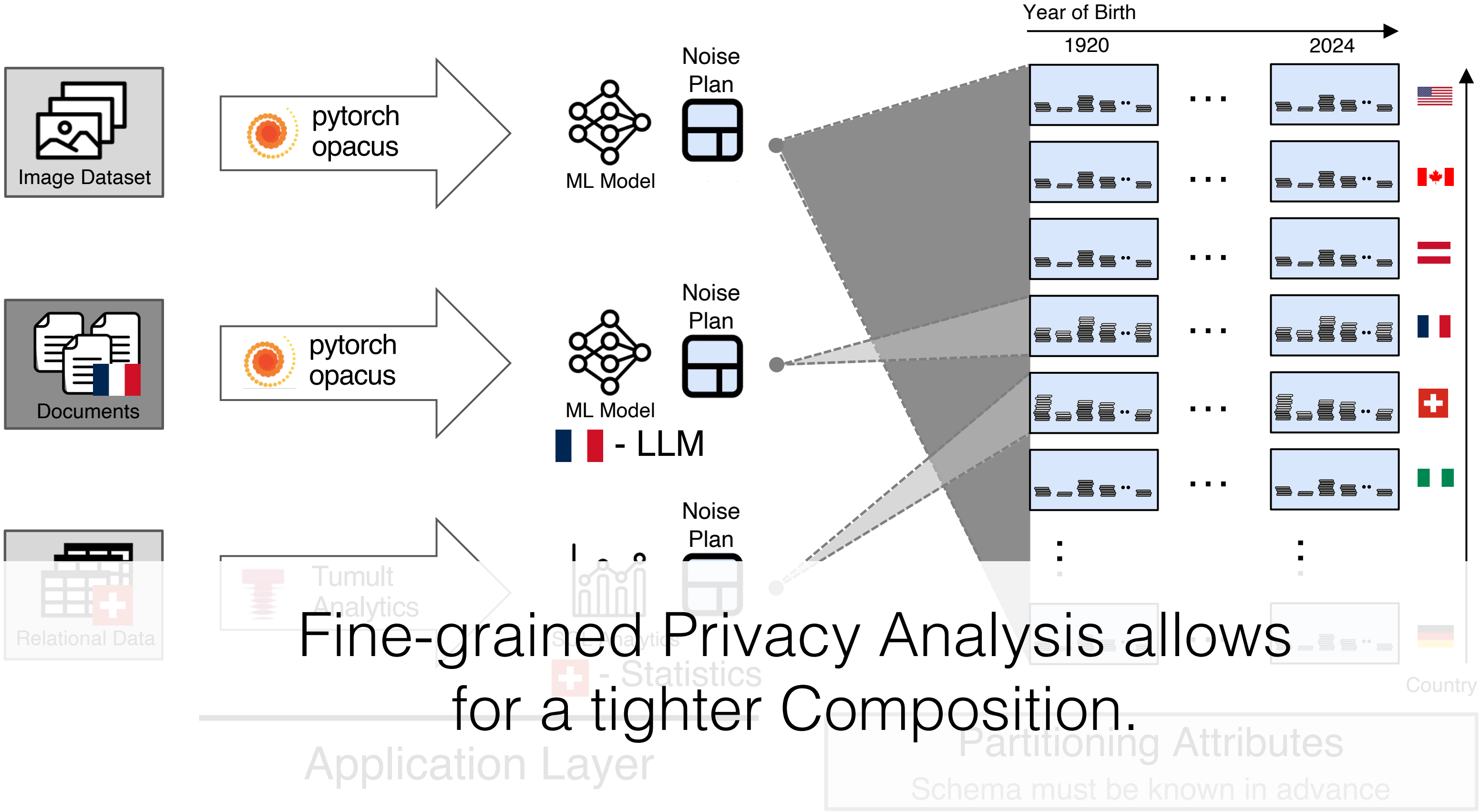


Application Layer

Partitioning Attributes

Schema must be known in advance

Fine-grained Privacy Analysis



Fine-grained Privacy Analysis allows for a tighter Composition.

Sampling: Random Subset Selection

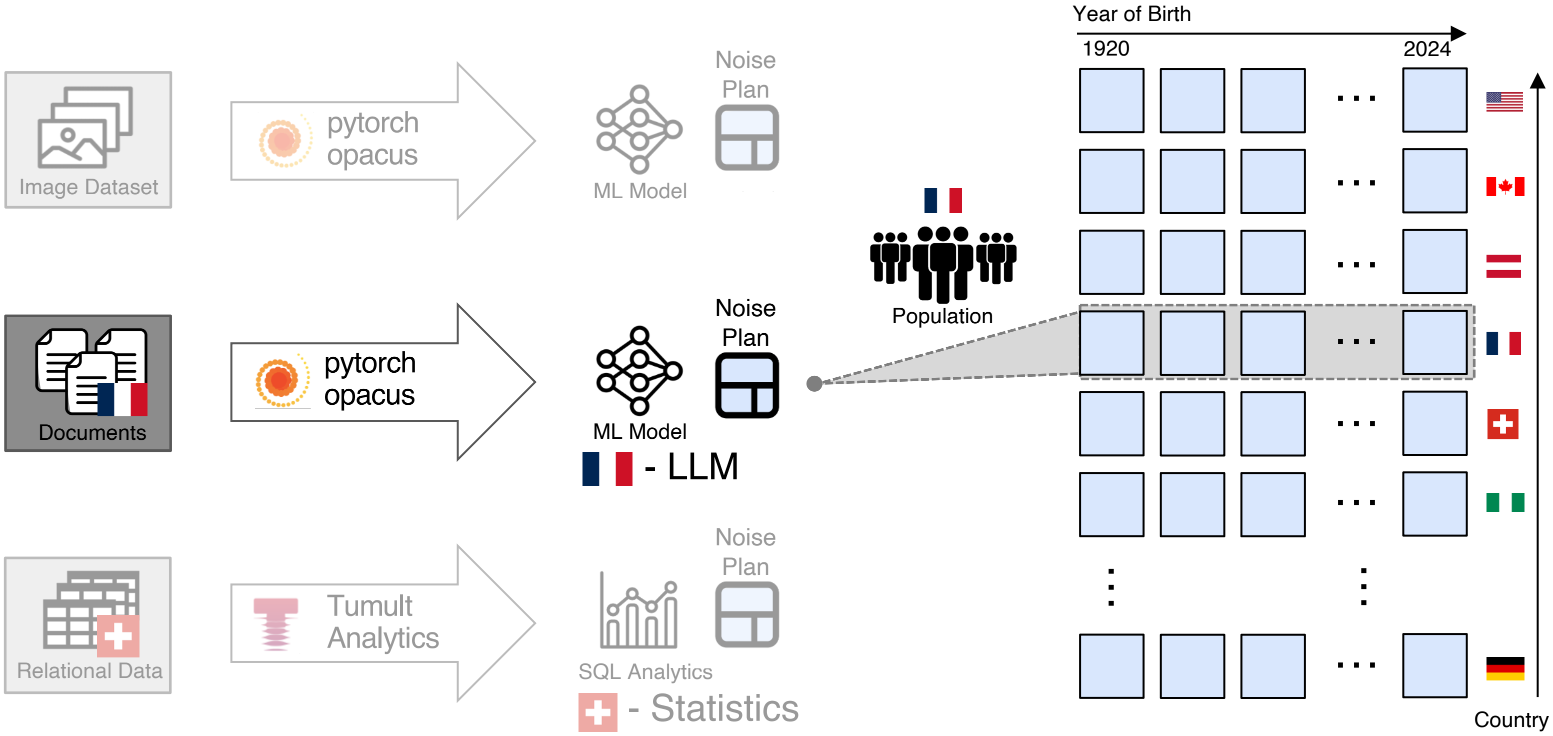


Image Dataset

pytorch opacus

ML Model

Noise Plan

Documents

pytorch opacus

ML Model

Noise Plan

🇫🇷 - LLM

Relational Data

Tumult Analytics

SQL Analytics

Noise Plan

🇨🇭 - Statistics

Population

Year of Birth

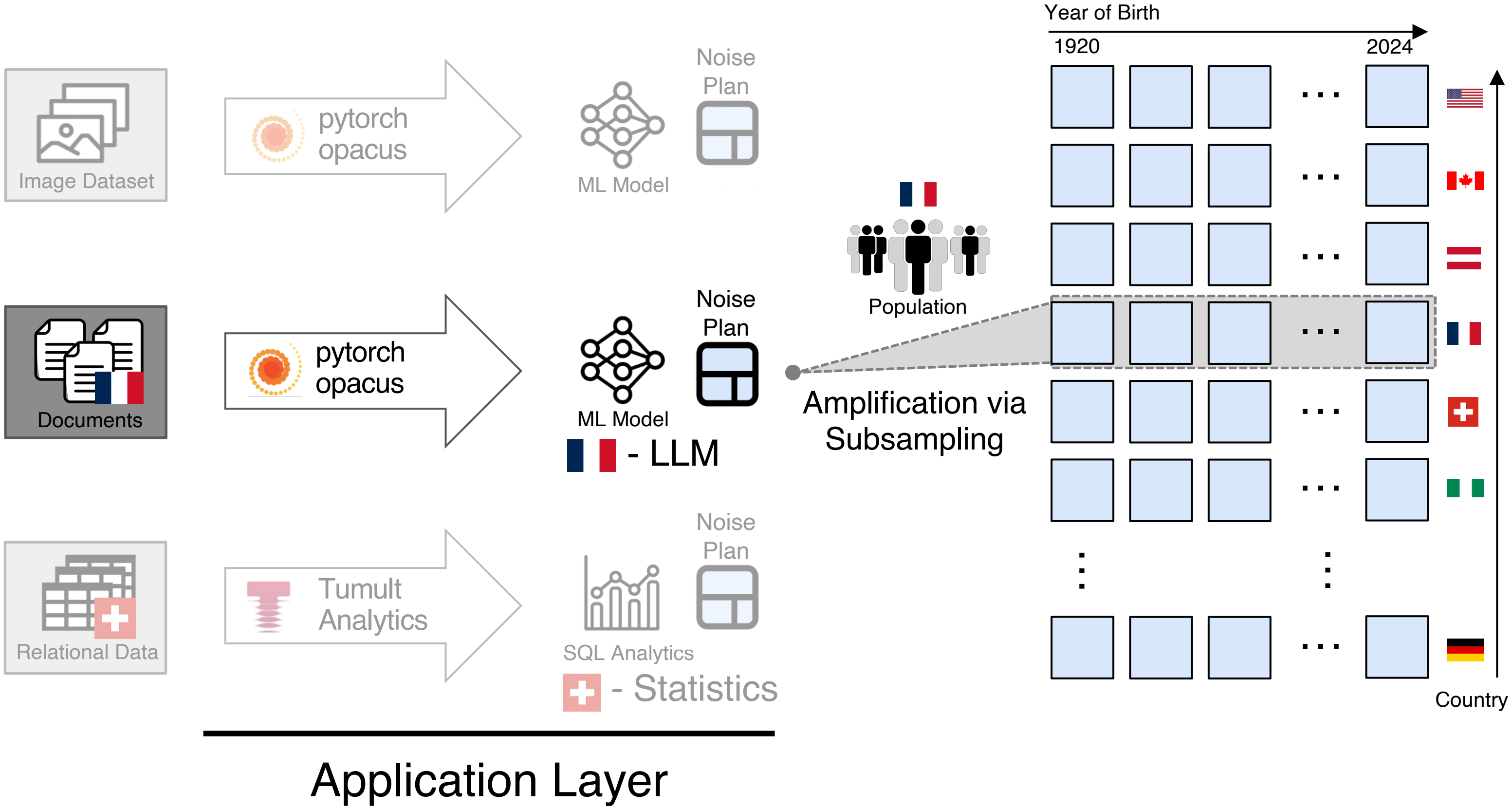
1920

2024

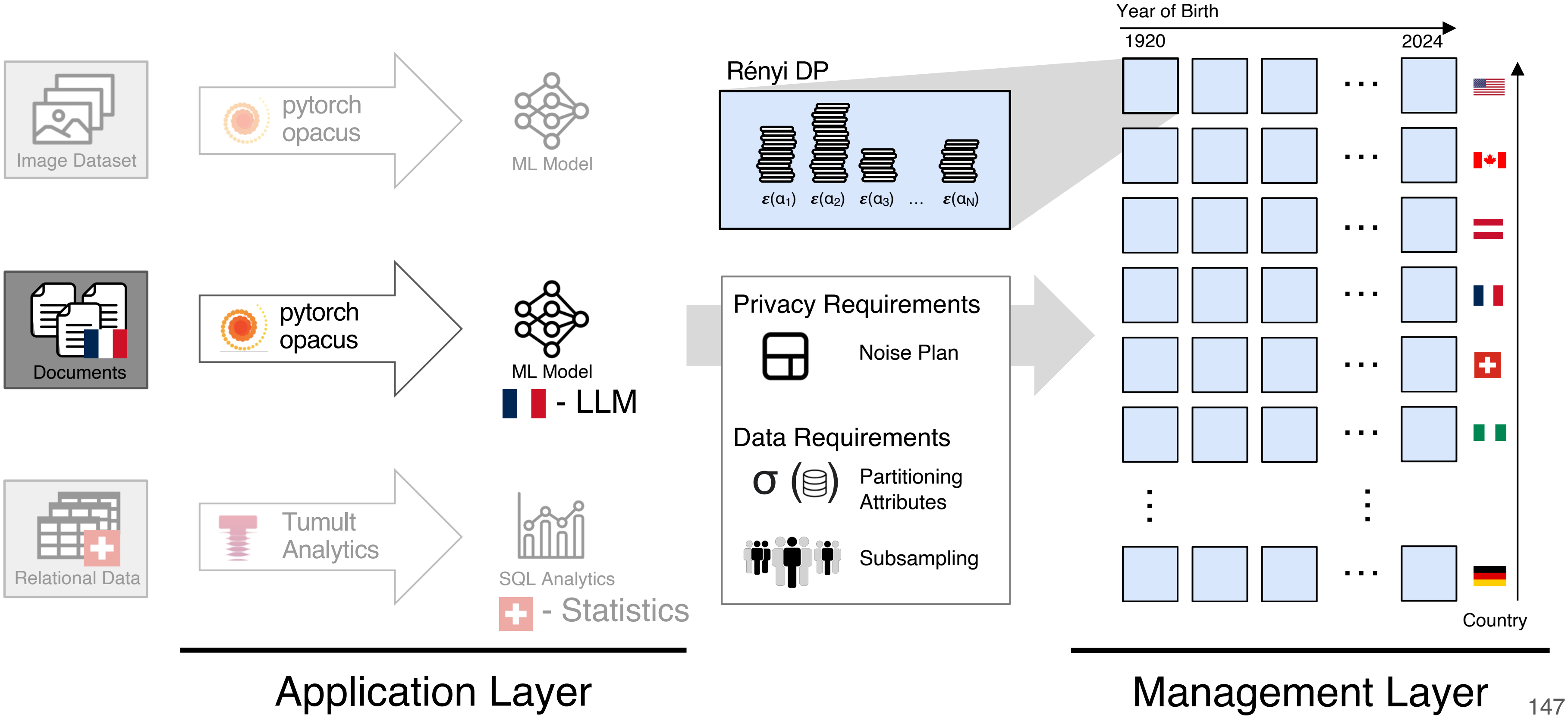
Country

Application Layer

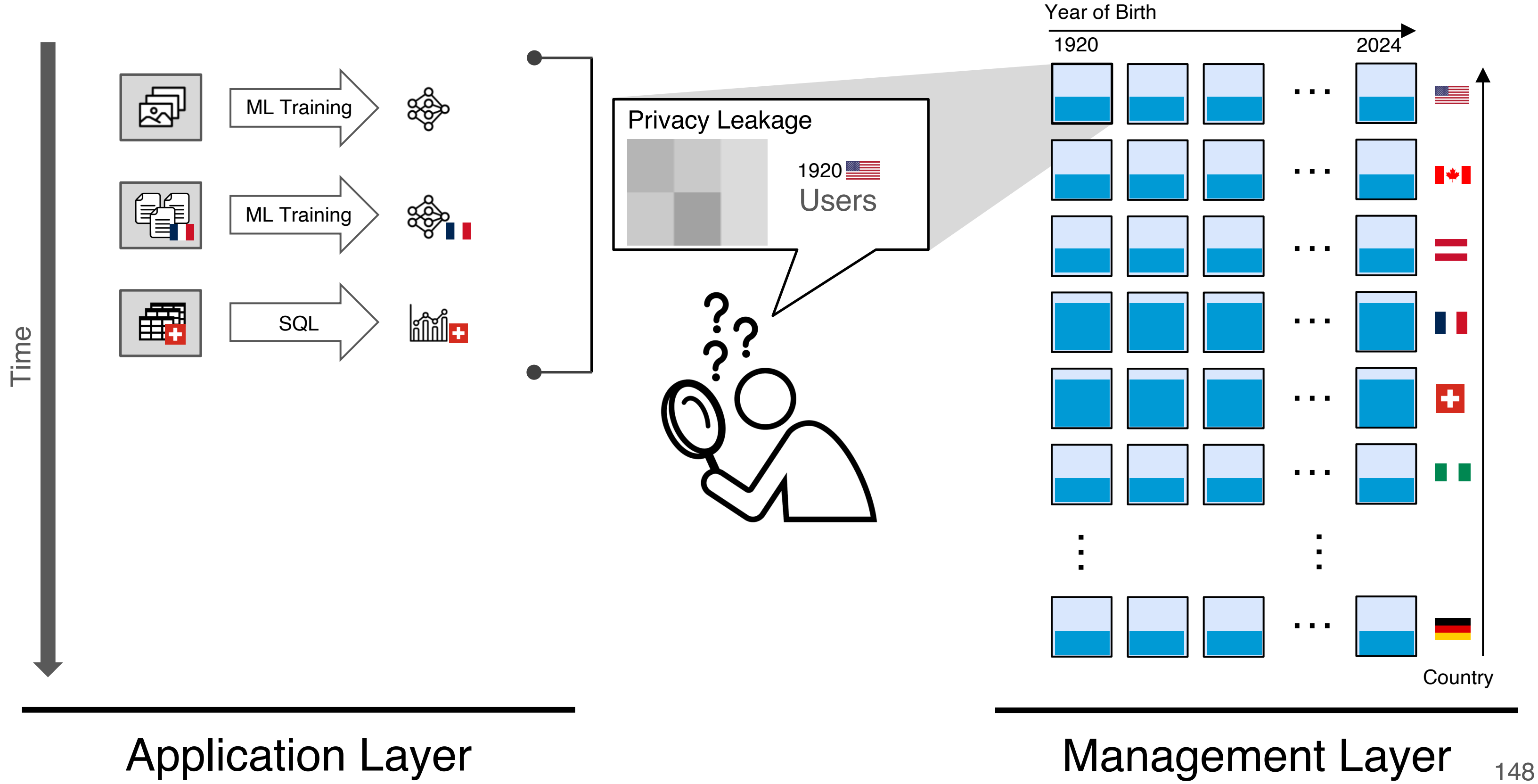
Sampling: Random Subset Selection



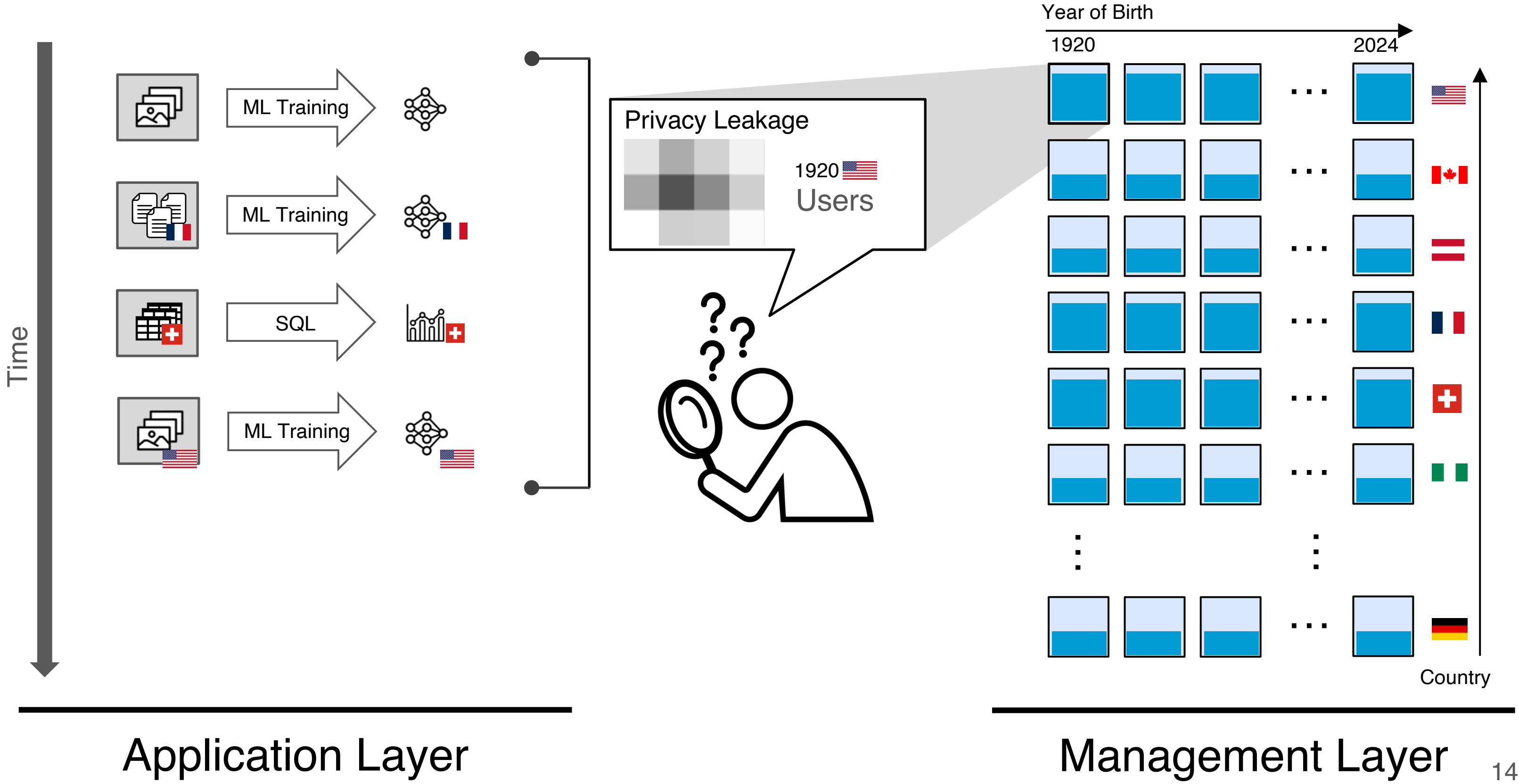
Scarce and Finite Resource



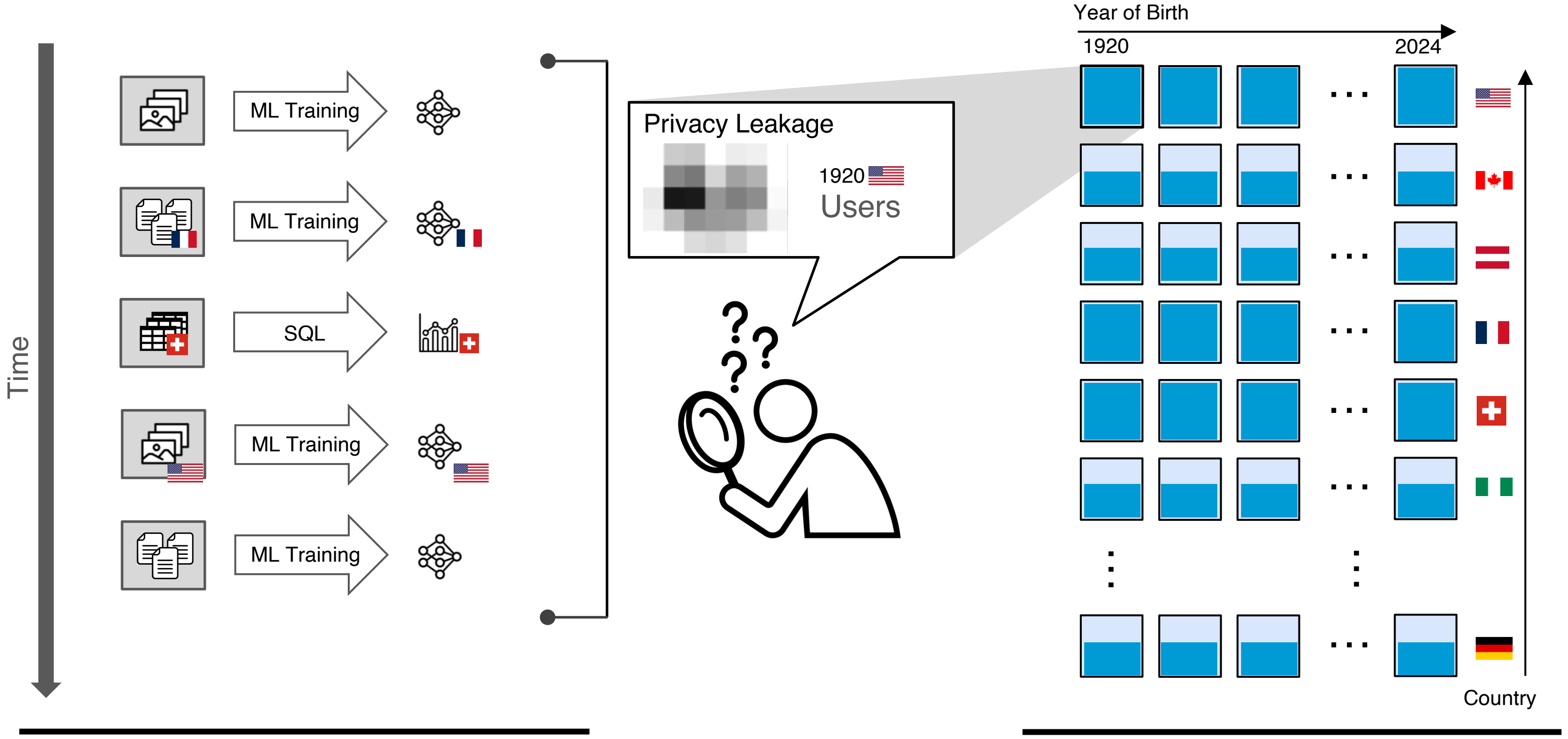
Scarce and Finite Resource



Scarce and Finite Resource



Scarce and Finite Resource



Time

ML Training

ML Training

SQL

ML Training

ML Training

Privacy Leakage

1920 🇺🇸

Users

Year of Birth

1920

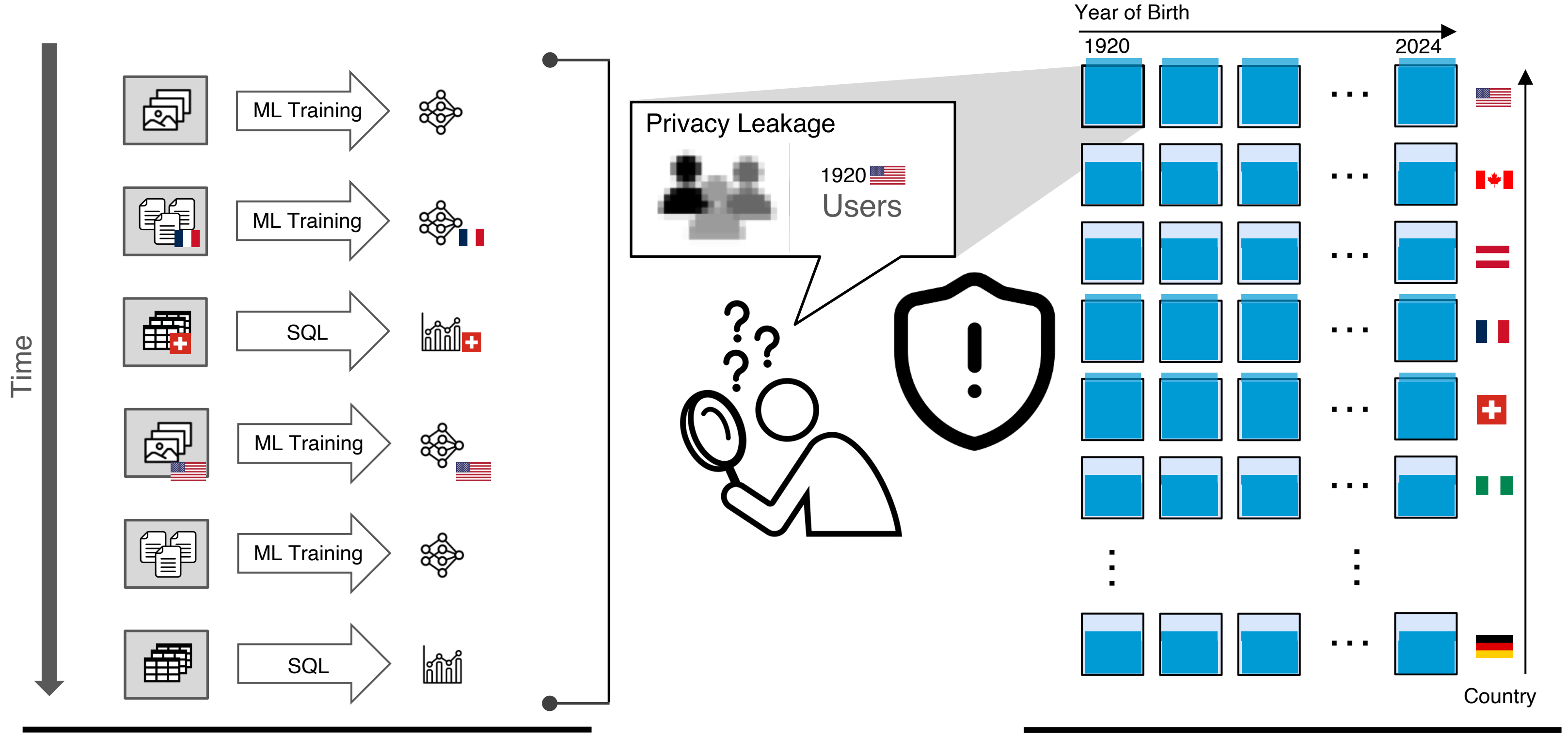
2024

Country

Application Layer

Management Layer

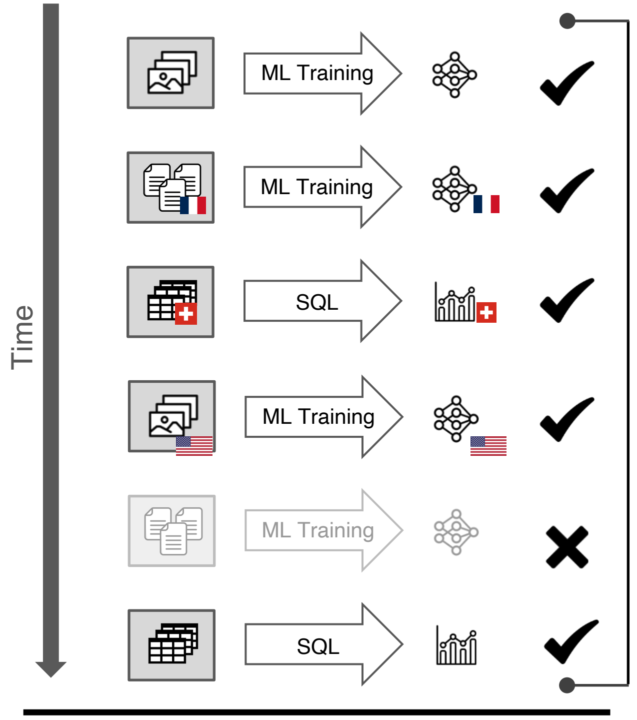
Scarce and Finite Resource



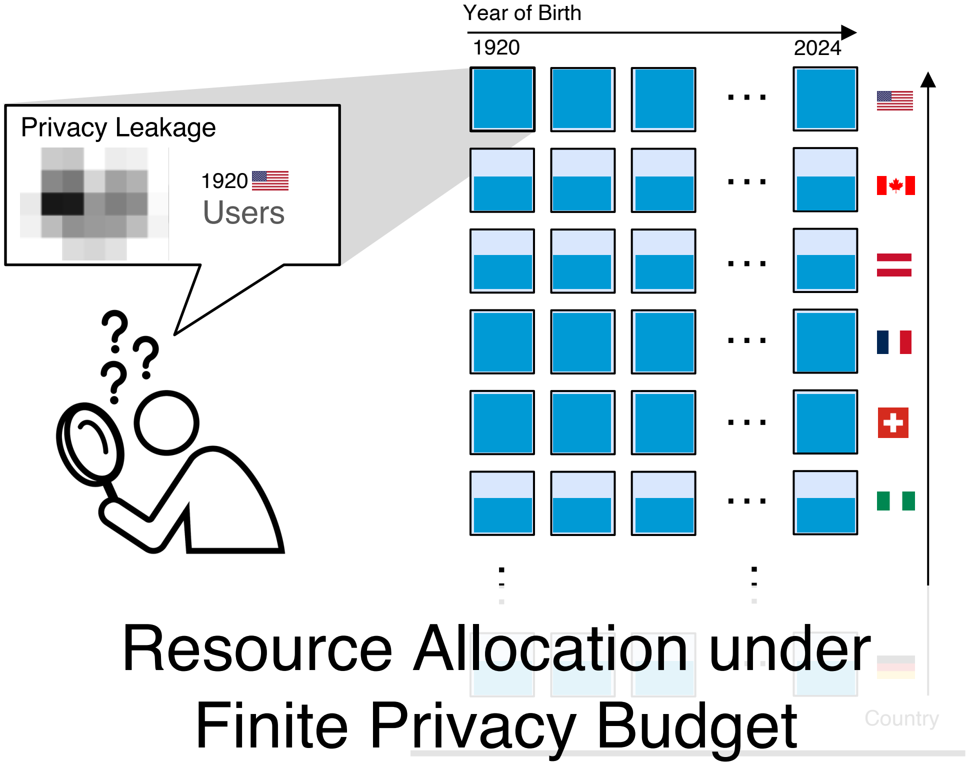
Application Layer

Management Layer

Scarce and Finite Resource



Application Layer



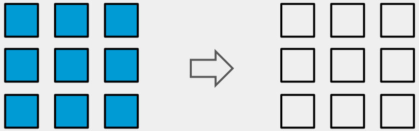
Resource Allocation under Finite Privacy Budget

Management Layer

Continuity under a Finite Budget

Ensuring Sustained Budget Allocation Over Time

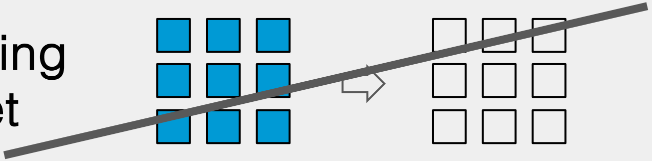
Resetting Budget



Continuity under a Finite Budget

Ensuring Sustained Budget Allocation Over Time

Resetting Budget

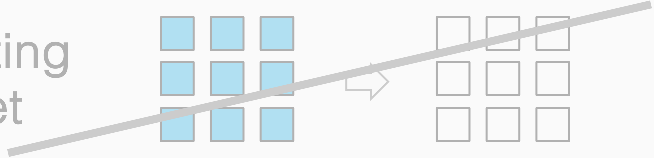


DP Violation

Continuity under a Finite Budget

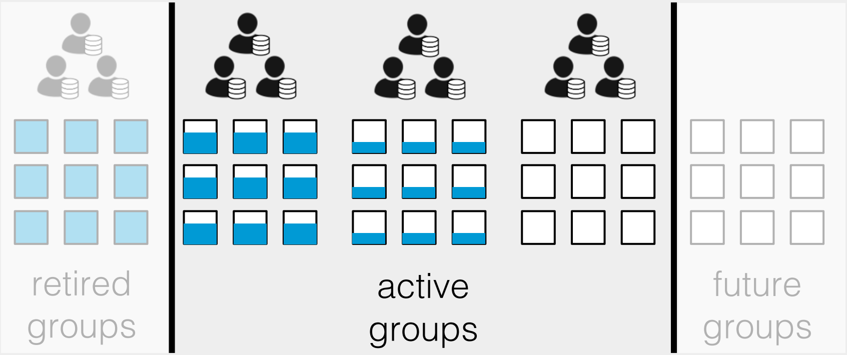
Ensuring Sustained Budget Allocation Over Time

Resetting Budget



DP Violation

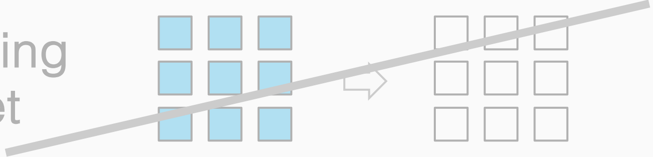
User Rotation



Continuity under a Finite Budget

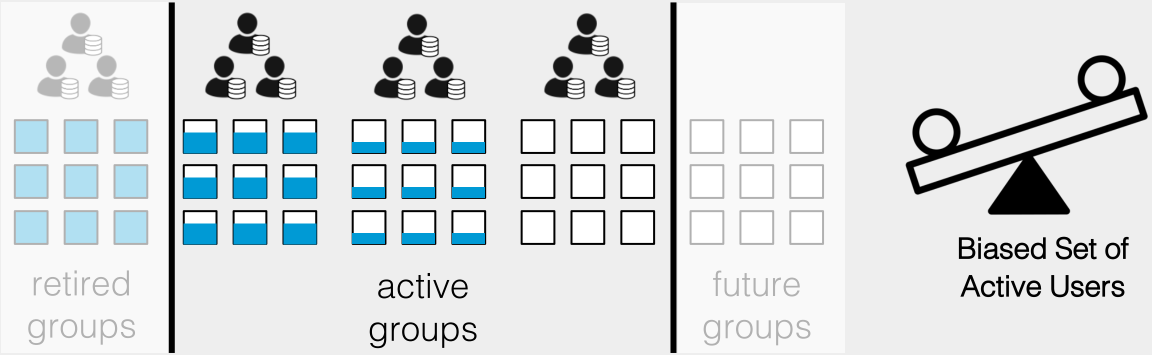
Ensuring Sustained Budget Allocation Over Time

~~Resetting Budget~~



DP Violation

User Rotation

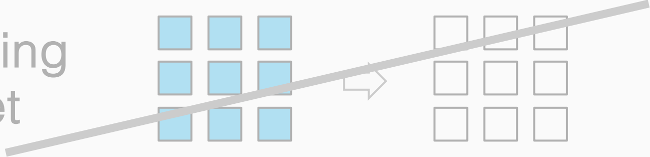


Biased Set of Active Users

Continuity under a Finite Budget

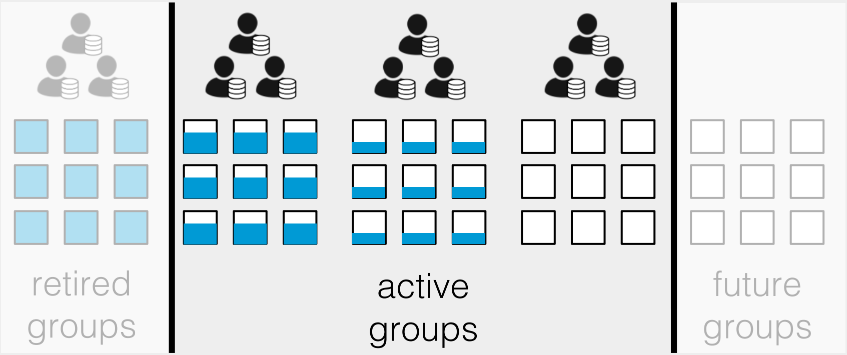
Ensuring Sustained Budget Allocation Over Time

Resetting Budget



DP Violation

User Rotation



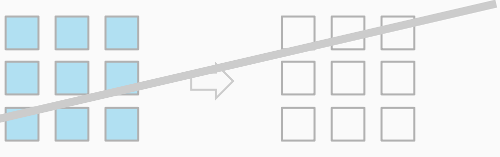
Biased Set of Active Users

Budget Guarantees with Unlocking

Continuity under a Finite Budget

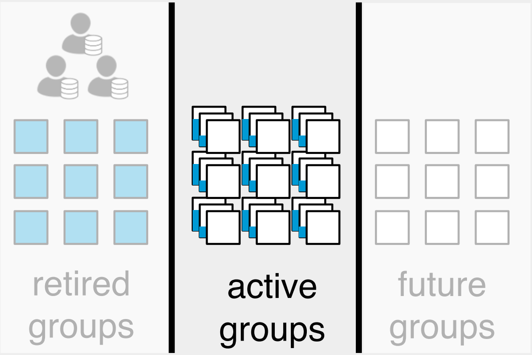
Ensuring Sustained Budget Allocation Over Time

Resetting Budget



DP Violation

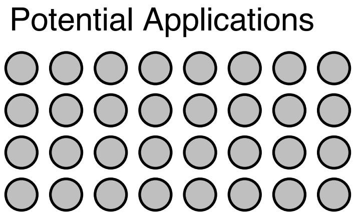
User Rotation



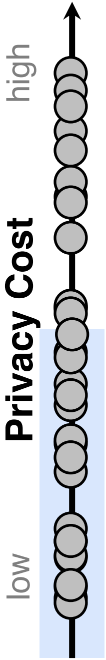
Biased Set of Active Users

Budget Guarantees with Unlocking

Privacy Resource Allocation

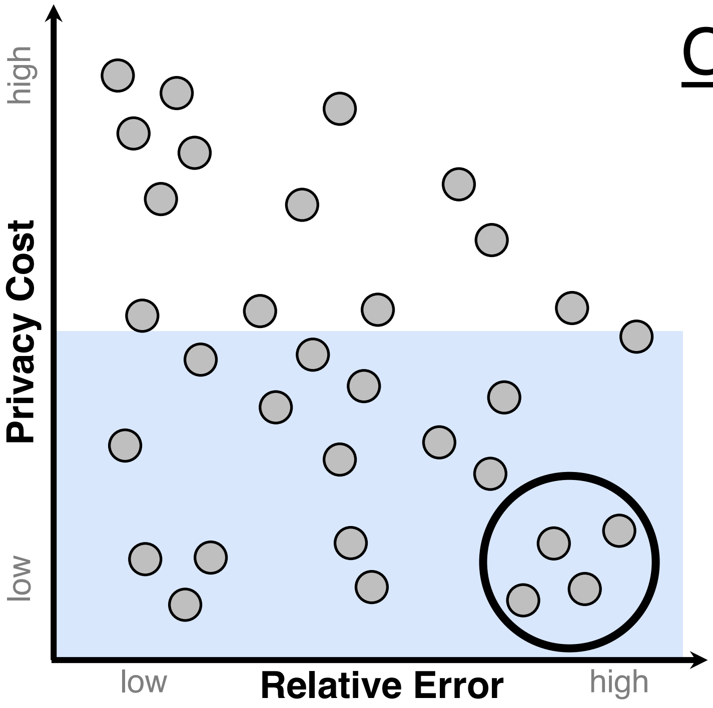


Privacy Resource Allocation



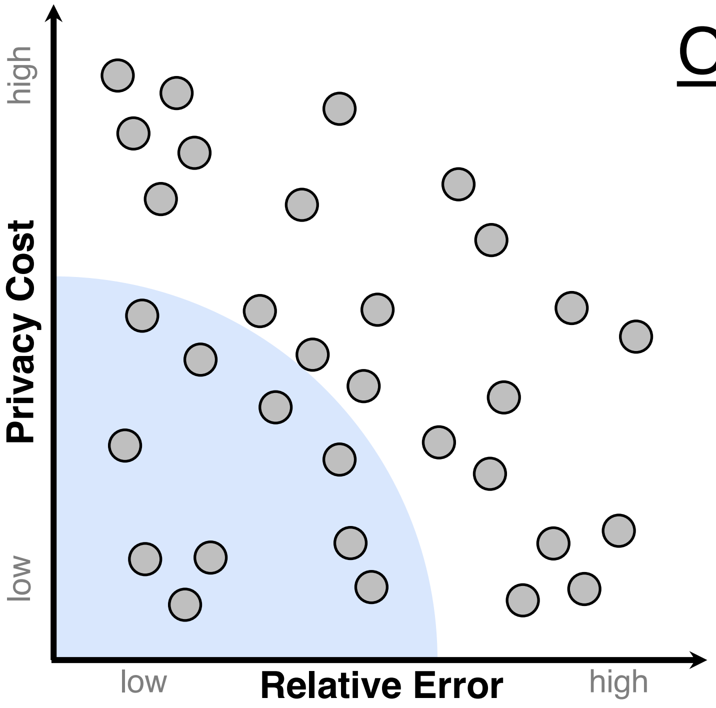
Optimize the number of Applications

Privacy Resource Allocation



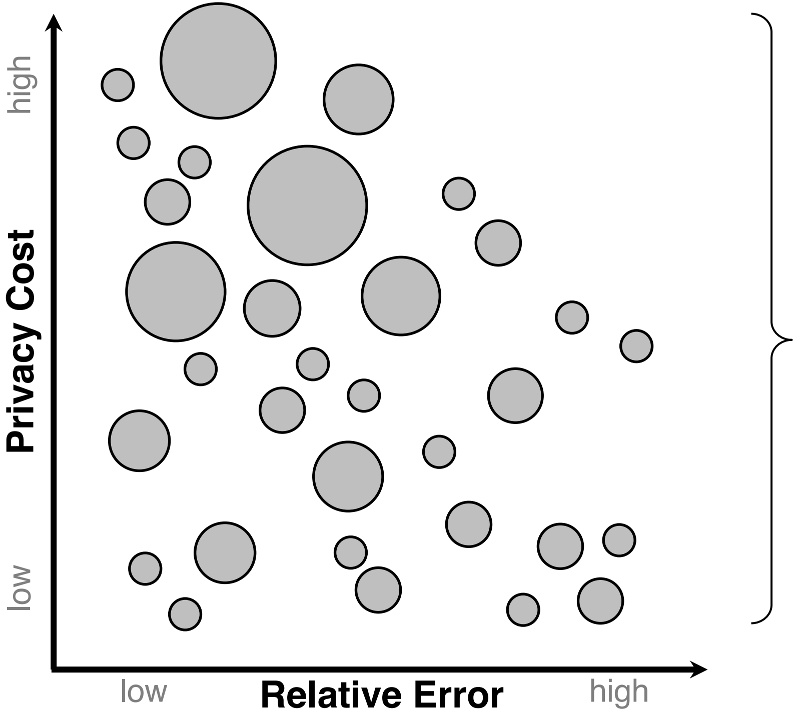
Optimize the number of Applications

Privacy Resource Allocation



Optimize Privacy Cost Relative to Error

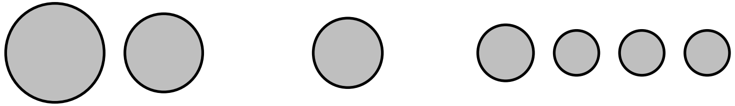
Privacy Resource Allocation



Optimize for Utility

Privacy Resource Allocation

Potential Applications



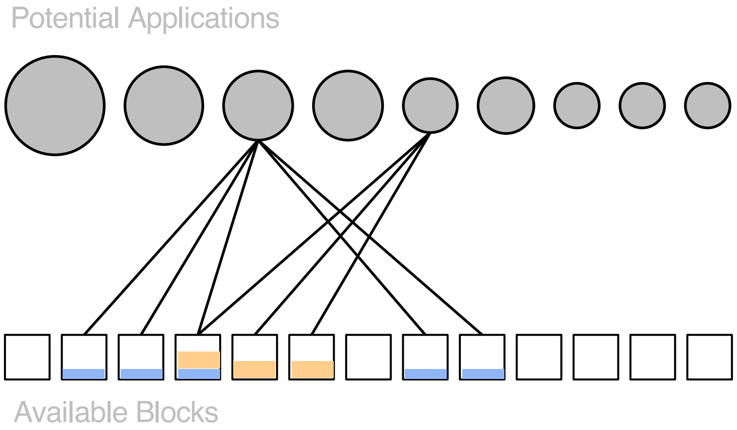
Multidimensional Knapsack Problem

Objective:

$$\max \sum_{i \in Apps} Utility_i * y_i$$

y_i = 1 if application i is allocated, else 0

Privacy Resource Allocation



Multidimensional Knapsack Problem

Objective:

$$\max \sum_{i \in Apps} Utility_i * y_i$$

y_i = 1 if application i is allocated, else 0

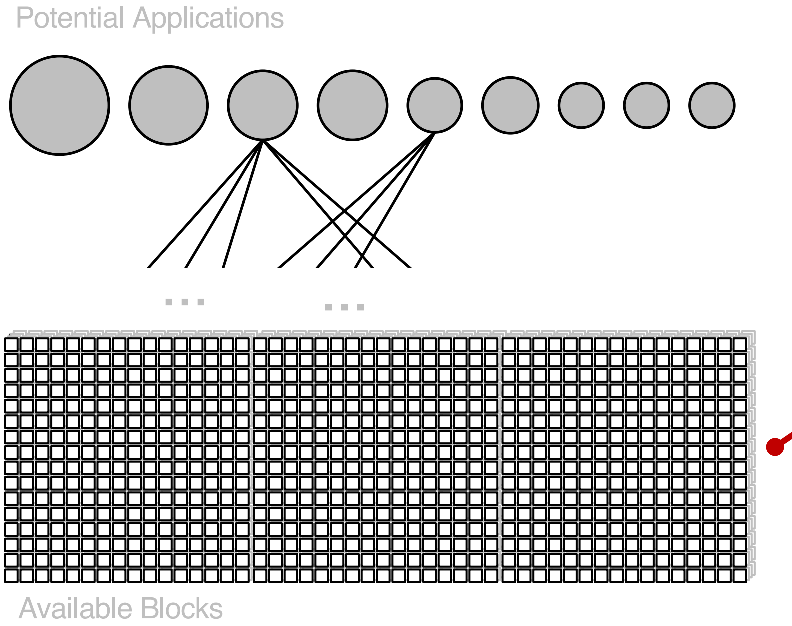
Budget Constraints:

$$s.t. \sum_{i \in Apps} \boldsymbol{\varepsilon}_{ij} * y_i \leq Budget_j \qquad \forall\, j \in Blocks$$

Privacy cost of application i for block j

* for simplicity we show the cost in $\boldsymbol{\varepsilon}$- DP rather than RDP

Privacy Resource Allocation



Multidimensional Knapsack Problem

Objective:

$$\max \sum_{i\,\in\,Apps} Utility_i * y_i$$

$y_i = 1$ if application i is allocated, else 0

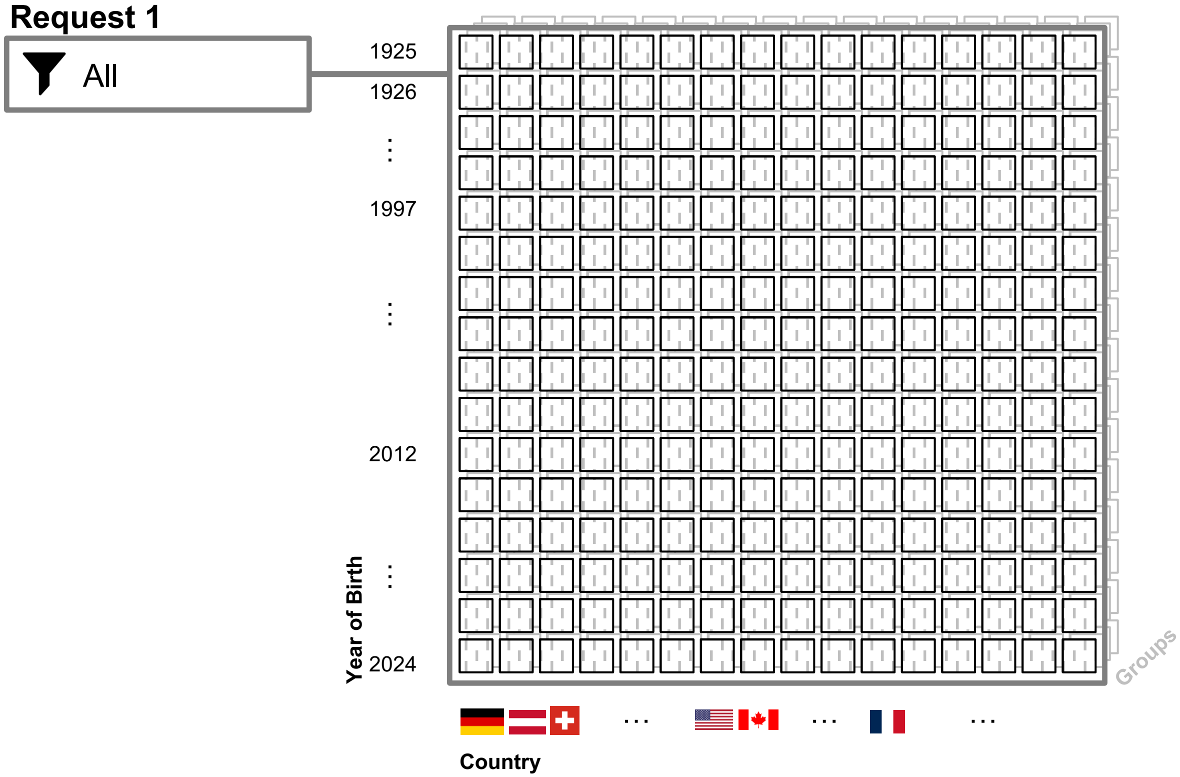
Budget Constraints:

$$s.t. \quad \sum_{i\,\in\,Apps} \boldsymbol{\varepsilon}_{ij} * y_i \le Budget_j \qquad \forall\, j \in Blocks$$

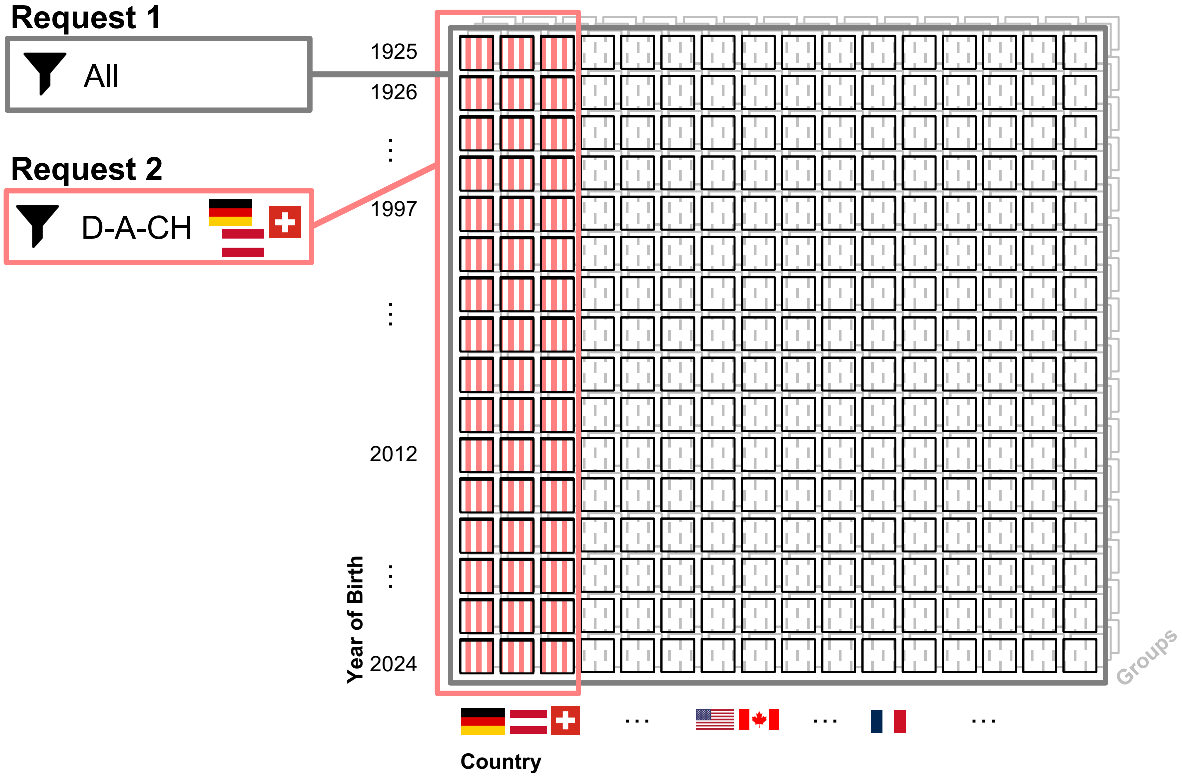
Privacy cost of application i for block j

* for simplicity we show the cost in ε- DP rather than RDP

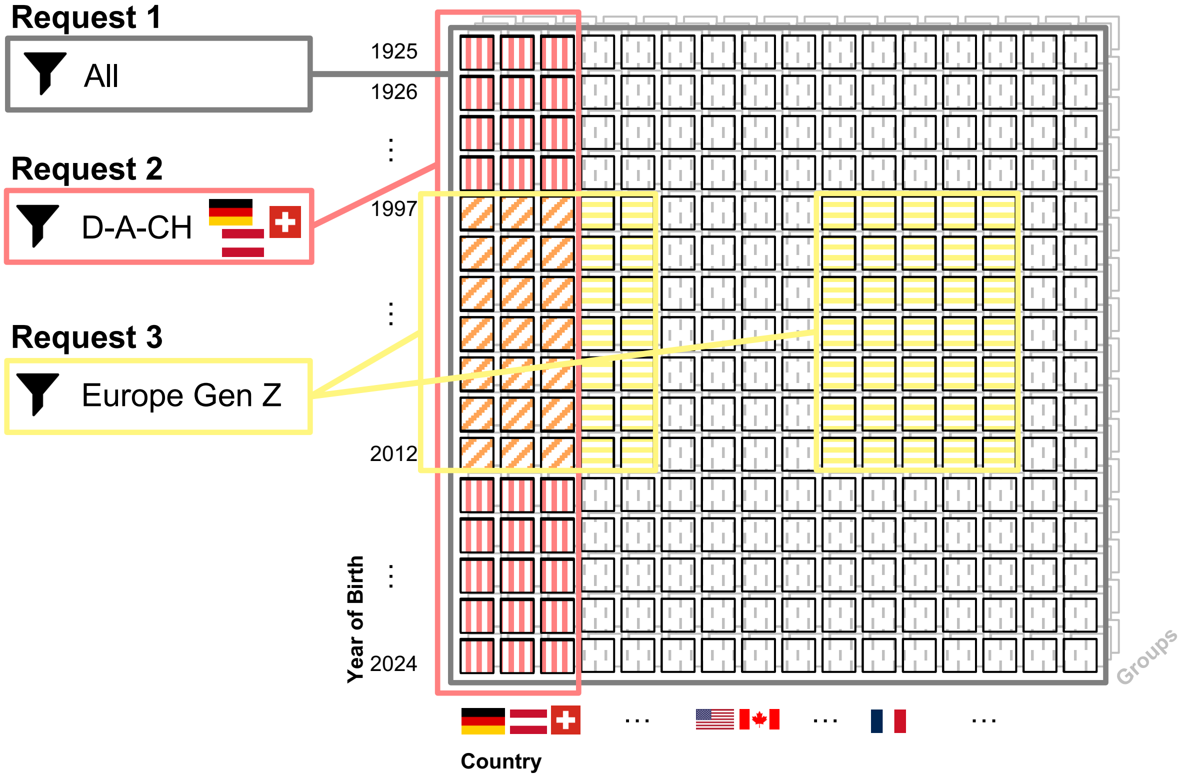
Resource Allocation: Taming the Complexity



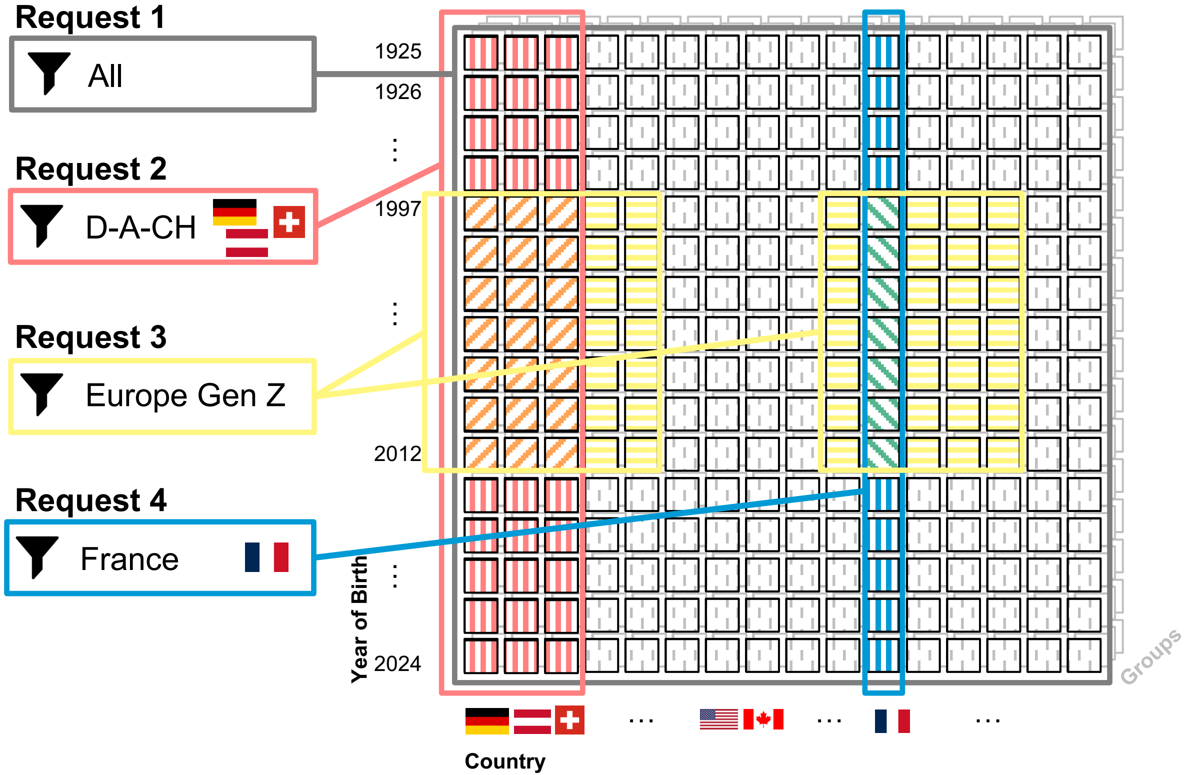
Resource Allocation: Taming the Complexity



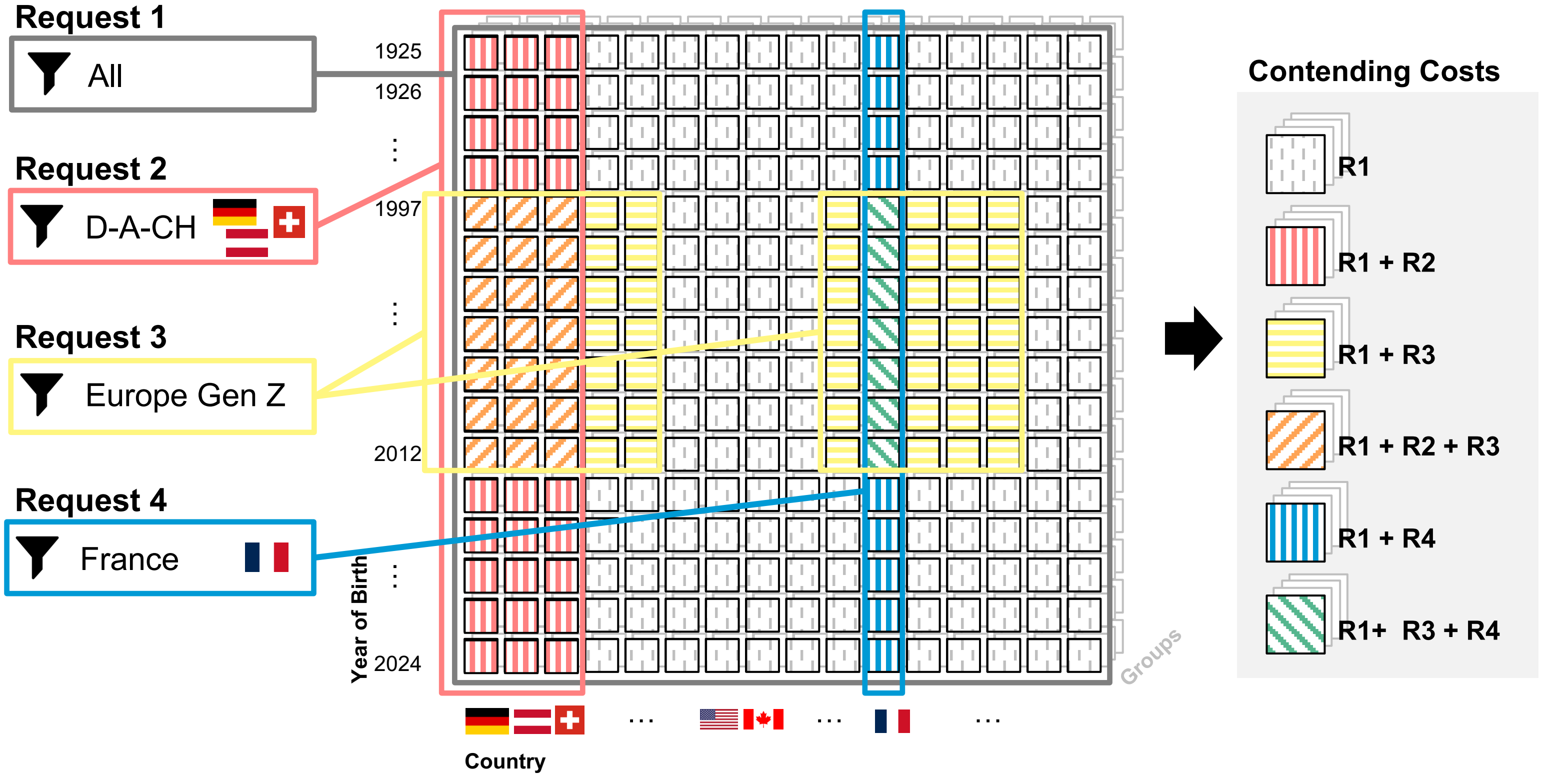
Resource Allocation: Taming the Complexity



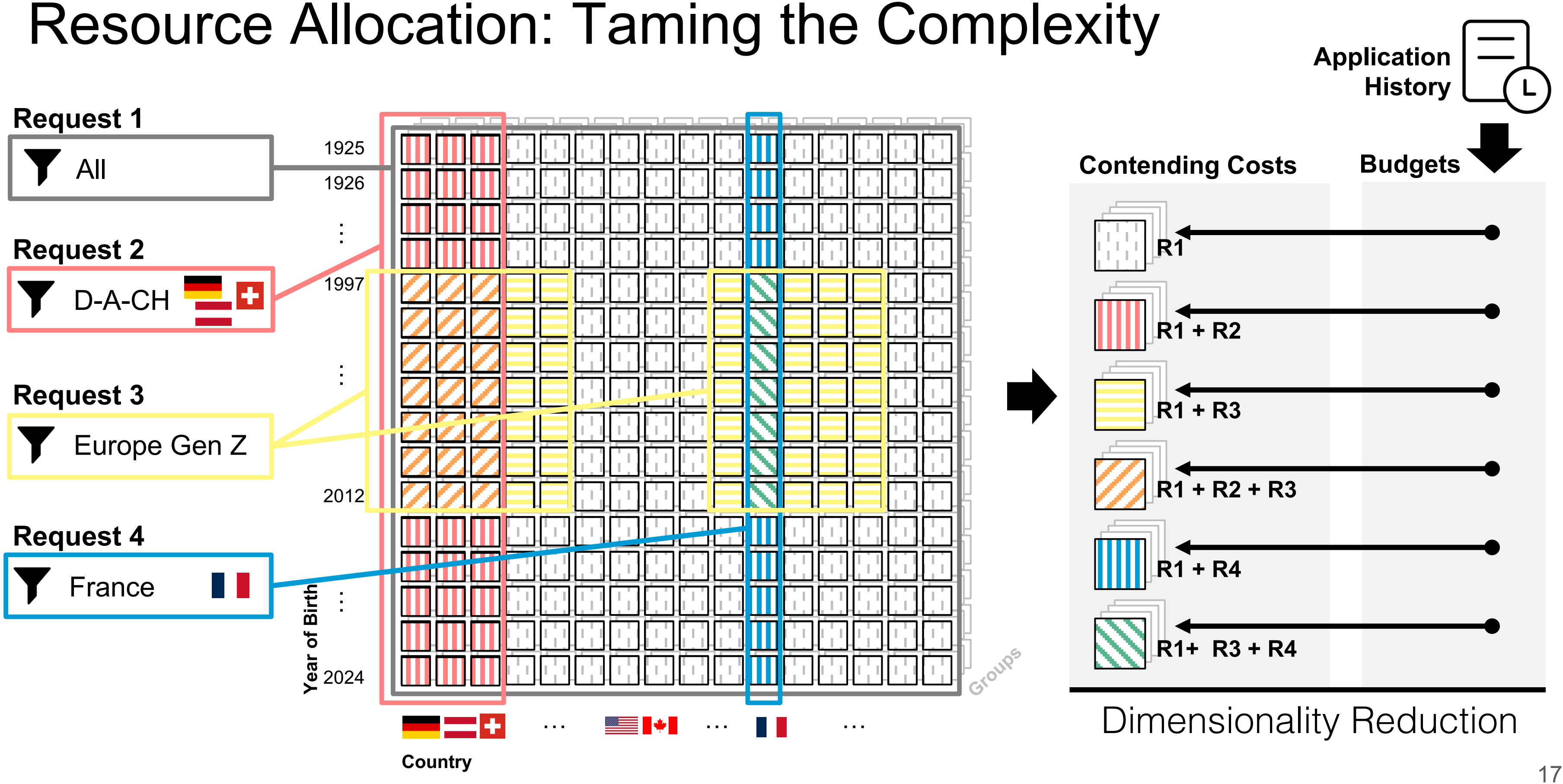
Resource Allocation: Taming the Complexity



Resource Allocation: Taming the Complexity

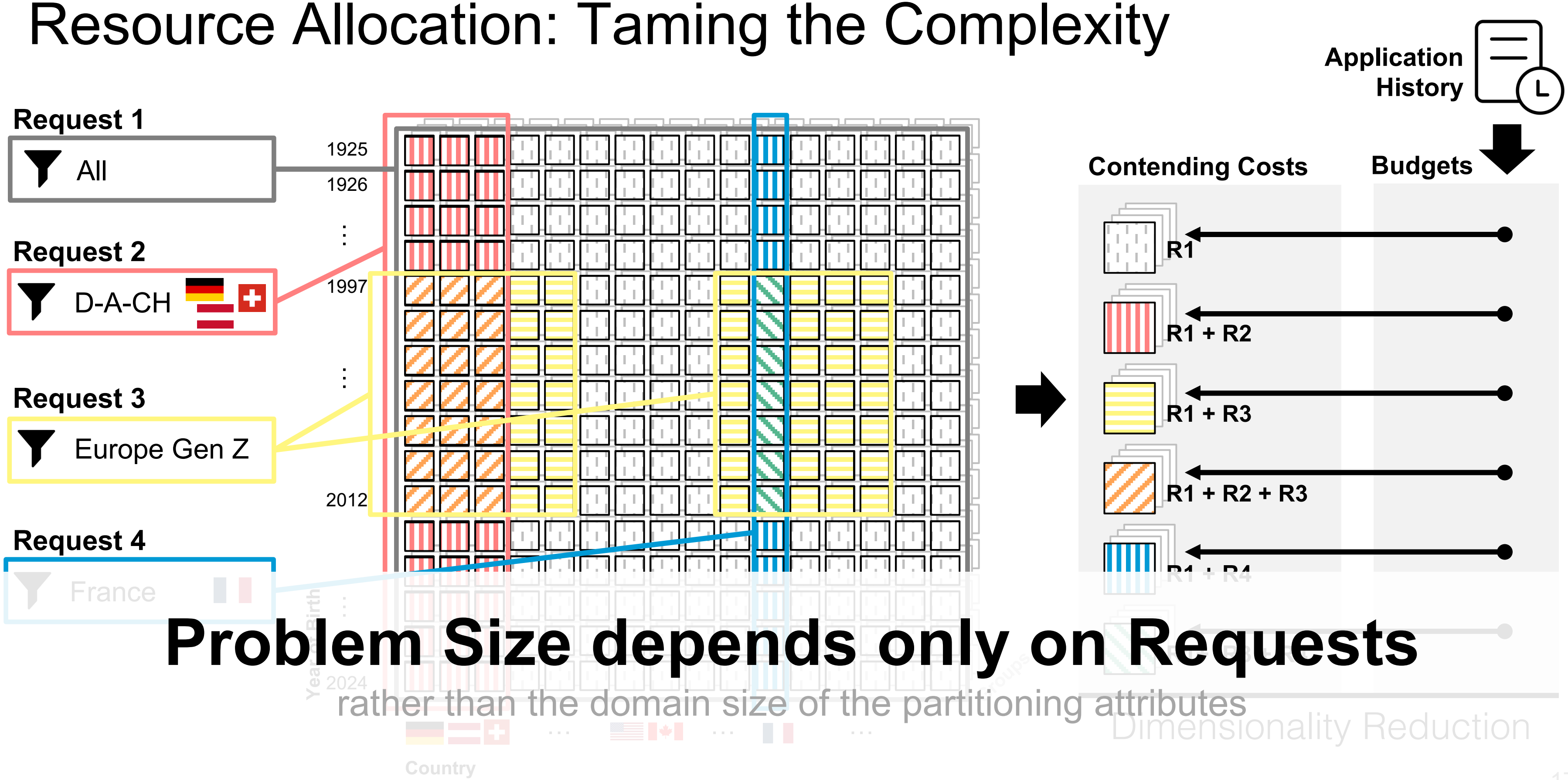


Resource Allocation: Taming the Complexity



Dimensionality Reduction

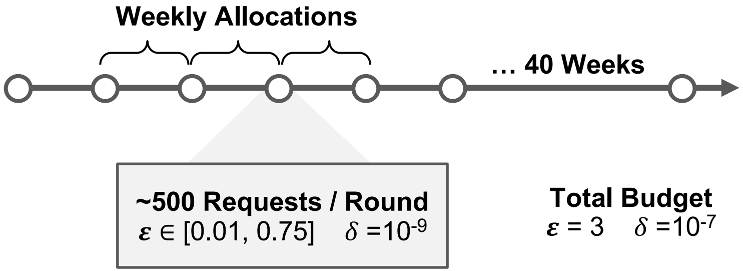
Resource Allocation: Taming the Complexity



Problem Size depends only on Requests

rather than the domain size of the partitioning attributes

Evaluation Scenario



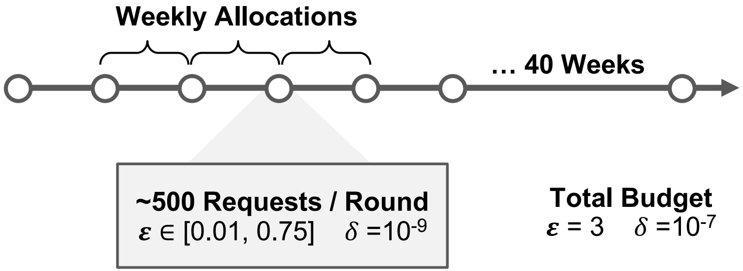
Weekly Allocations

… 40 Weeks

~500 Requests / Round
$\boldsymbol{\varepsilon} \in [0.01, 0.75]$ $\delta = 10^{-9}$

Total Budget
$\boldsymbol{\varepsilon} = 3$ $\delta = 10^{-7}$

Evaluation Scenario

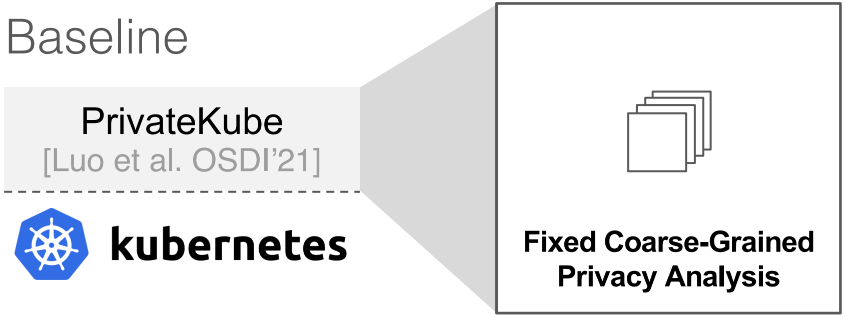


Weekly Allocations

… 40 Weeks

~500 Requests / Round
$\varepsilon \in [0.01, 0.75]$ $\delta = 10^{-9}$

Total Budget
$\varepsilon = 3$ $\delta = 10^{-7}$



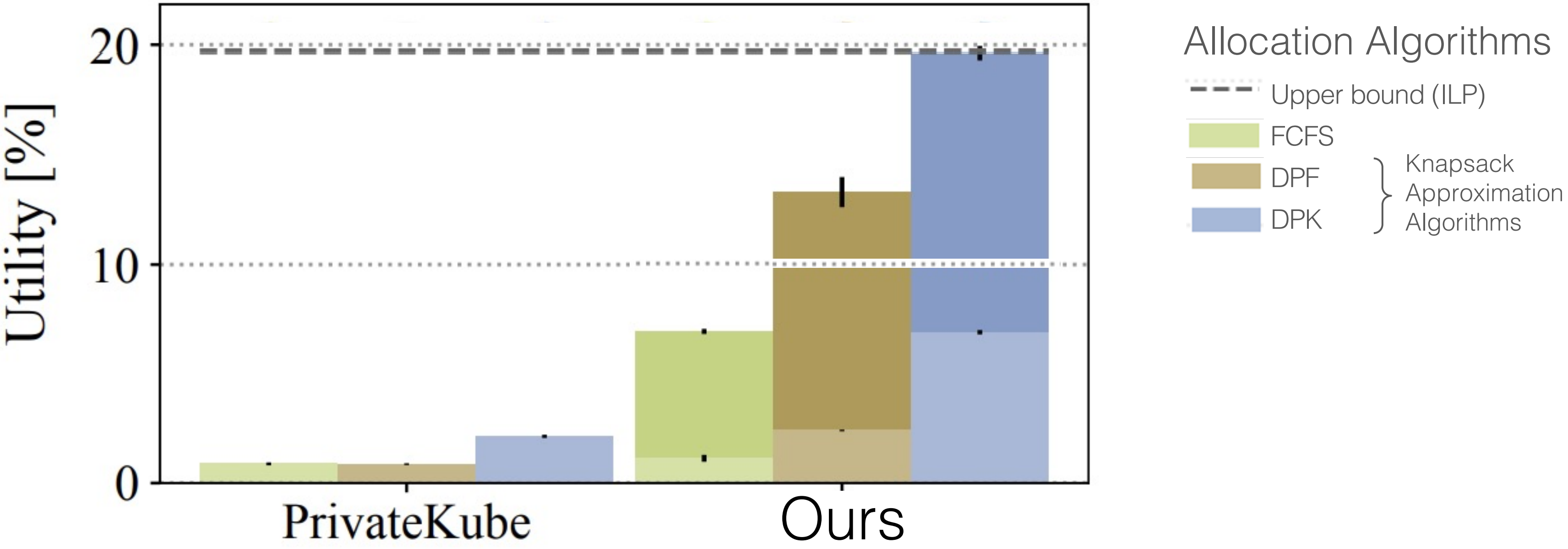
Baseline

PrivateKube
[Luo et al. OSDI'21]

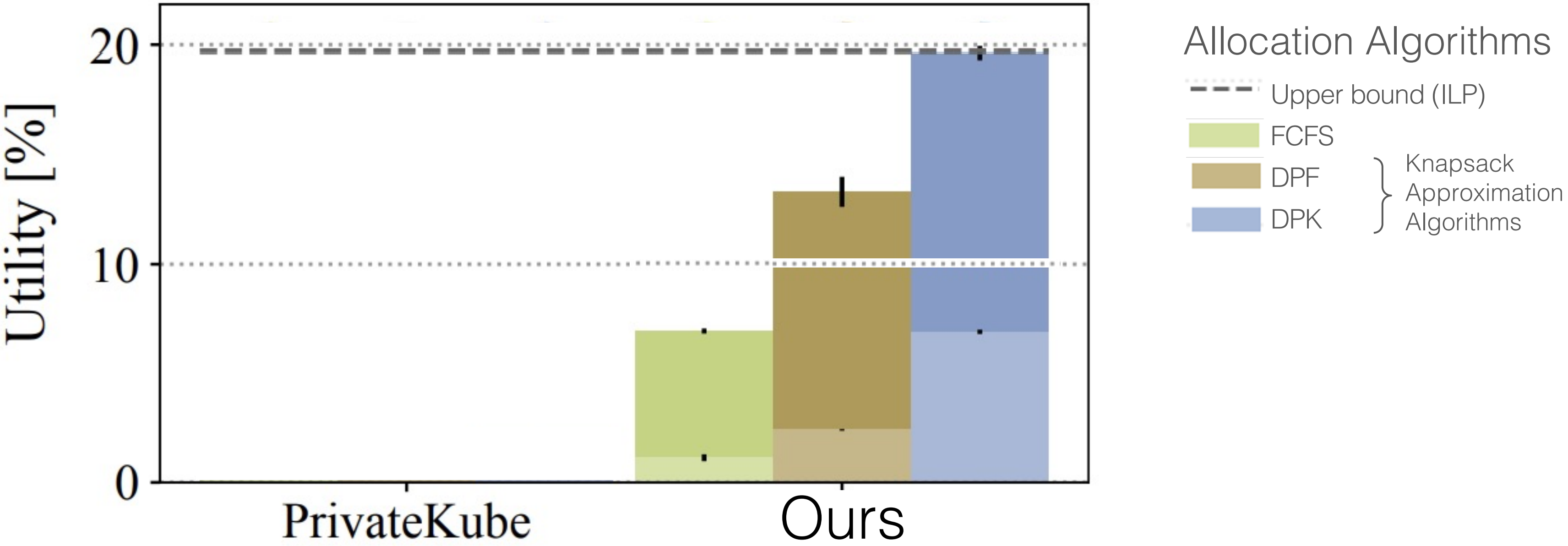
kubernetes

Fixed Coarse-Grained Privacy Analysis

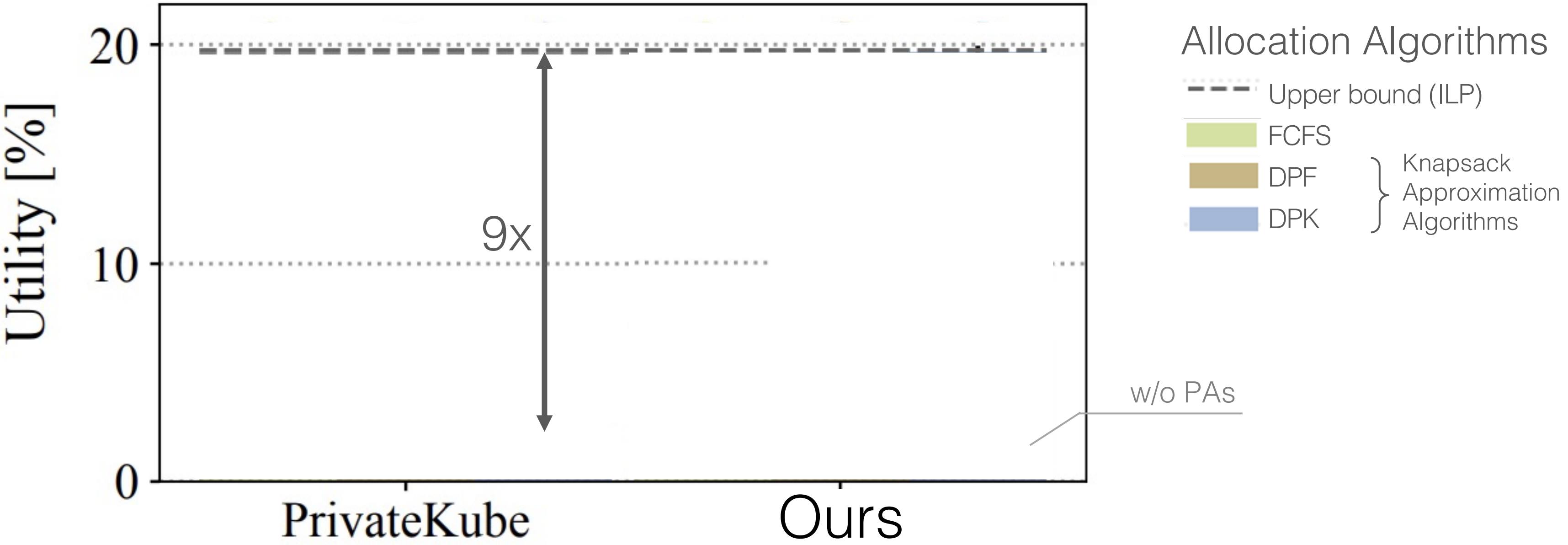
Workload: Mixture of Analytics and ML Tasks



Workload: Mixture of Analytics and ML Tasks

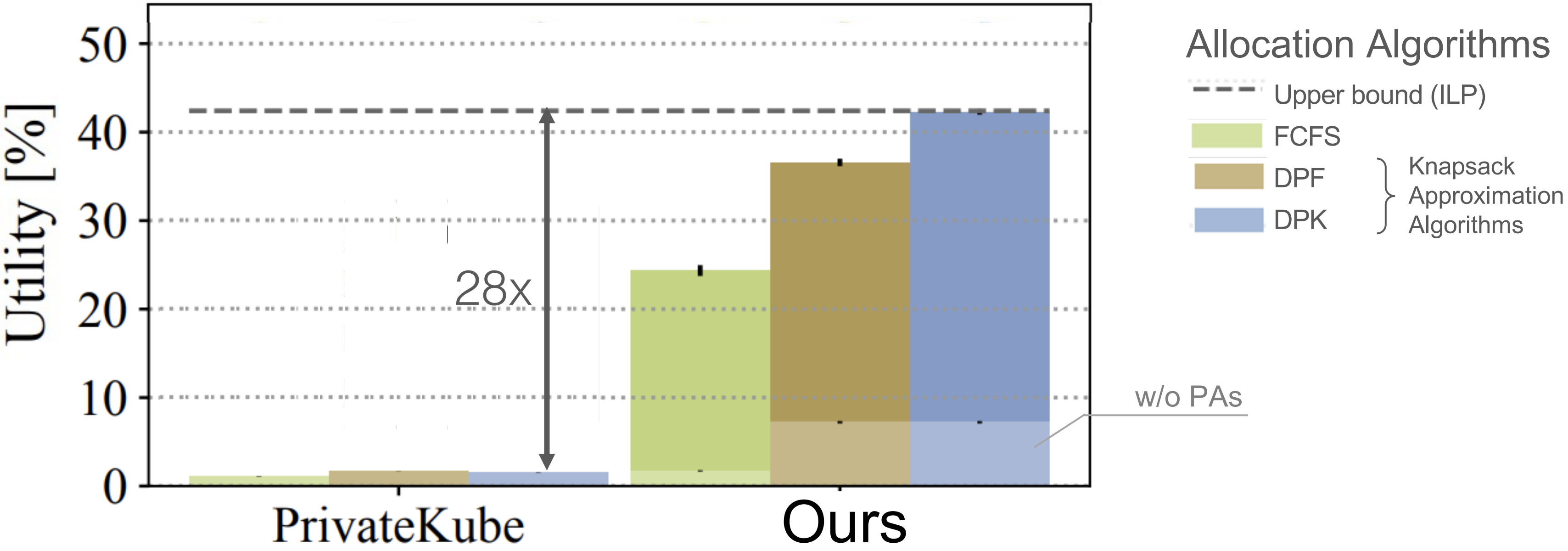


Workload: Mixture of Analytics and ML Tasks

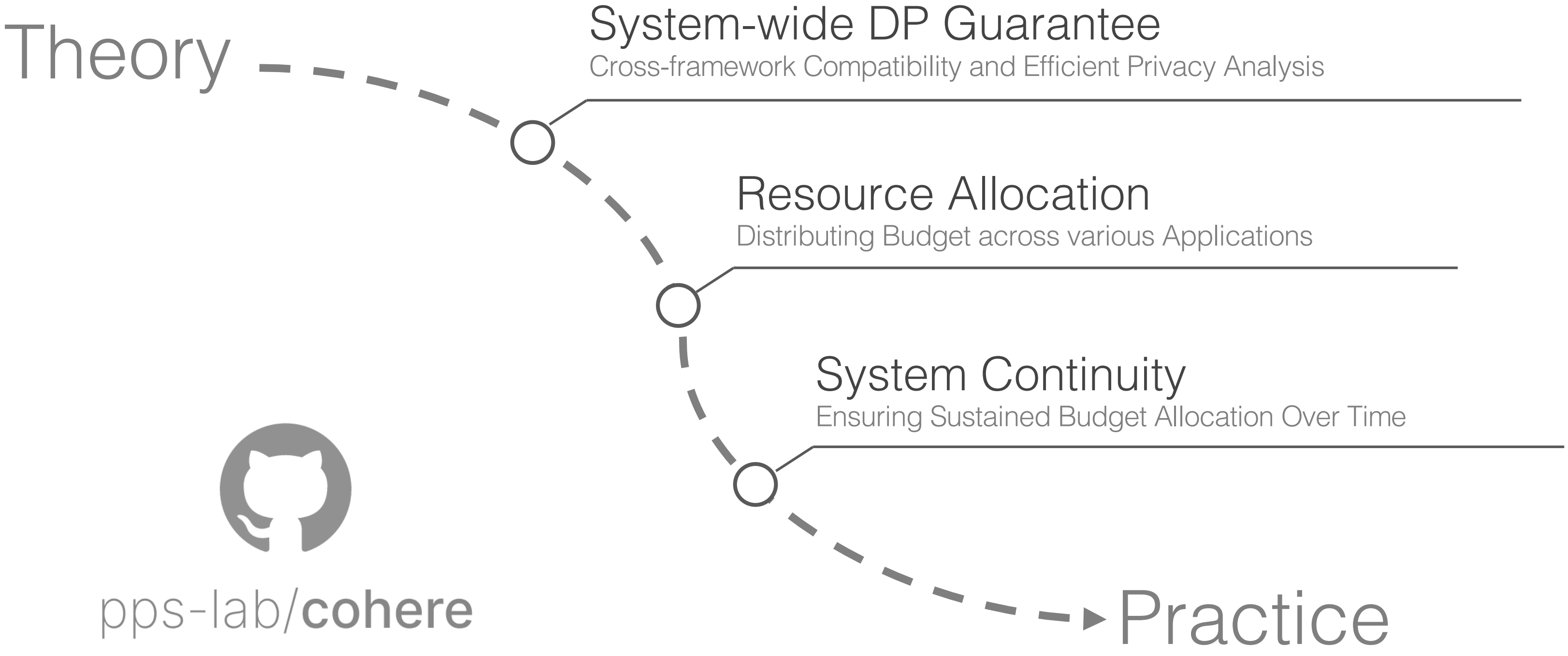


Workload: Predicate Counting Queries

SELECT Count(*) **FROM** x **WHERE Φ** (Only Gaussian Mechanism)



Differential Privacy



Theory

System-wide DP Guarantee

Cross-framework Compatibility and Efficient Privacy Analysis

Resource Allocation

Distributing Budget across various Applications

System Continuity

Ensuring Sustained Budget Allocation Over Time

Practice

pps-lab/**cohere**

Democratize Privacy-Preserving Computation

My work aims to **democratize access to privacy-preserving computation** with new tools, systems, and abstractions.

Secure Computation

SoK: Fully Homomorphic Encryption Compilers

Alexander Viand, ETH Zurich, alexander.viand@inf.ethz.ch
Patrick Jattke, ETH Zurich, pjattke@ethz.ch
Anwar Hithnawi, ETH Zurich, anwar.hithnawi@inf.ethz.ch

FHE Compilers
IEEE S&P

HECO: Fully Homomorphic Encryption Compiler

Alexander Viand, Patrick Jattke, Miro Haller, Anwar Hithnawi
ETH Zurich

HECO
USENIX Security

Differential Privacy

Cohere: Managing Differential Privacy in Large Scale Systems

Nicolas Küchler[1], Emanuel Opel[1], Hidde Lycklama[1], Alexander Viand[2], Anwar Hithnawi[1]
[1]ETH Zurich [2]Intel Labs

Cohere
IEEE S&P

Programmability

Deployments

Talos: Encrypted Query Processing for the Internet of Things
Secure Sharing of Partially Homomorphic Encrypted IoT Data
TimeCrypt: Encrypted Data Stream Processing at Scale with Cryptographic Access Control
Droplet: Decentralized Authorization and Access Control for Encrypted Data Streams
Zeph: Cryptographic Enforcement of End-to-End Data Privacy
VF-PS: How to Select Important Participants in Vertical Federated Learning, Efficiently and Securely?
RoFL: Attestable Robustness for Secure Federated Learning

Talos
ACM SenSys

Pilatus
ACM SenSys

TimeCrypt
USENIX NSDI

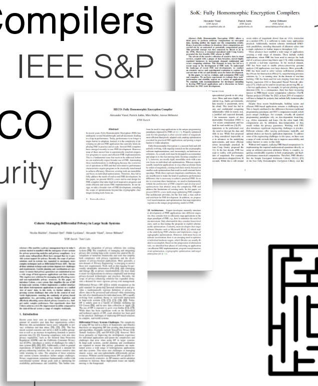
Droplet
USENIX Security

Zeph
USENIX OSDI

VF-PS
NeurIPS

RoFL
IEEE S&P

Privacy-Preserving
System Designs


SoK: Fully Homomorphic Encryption Compilers
HECO: Fully Homomorphic Encryption Compiler
Cohere: Managing Differential Privacy in Large Scale Systems

FHE Compilers
IEEE S&P

HECO
USENIX Security

Cohere
IEEE S&P

Democratize
Privacy-Preserving
Computation

My work aims to **build** practical systems that use **cryptography to empower users and preserve their privacy.**

Looking Forward

Democratize Privacy-Preserving Computation

Privacy-Preserving Systems Designs

Democratize Privacy-Preserving Computation

Hybrid Compilation

FHE ZKP

MPC

Secure Computation on Heterogeneous Hardware

Privacy-Preserving Systems Designs

Democratize Privacy-Preserving Computation

Hybrid Compilation

FHE ZKP

MPC

Secure Computation on Heterogeneous Hardware

Privacy-Preserving Systems Designs

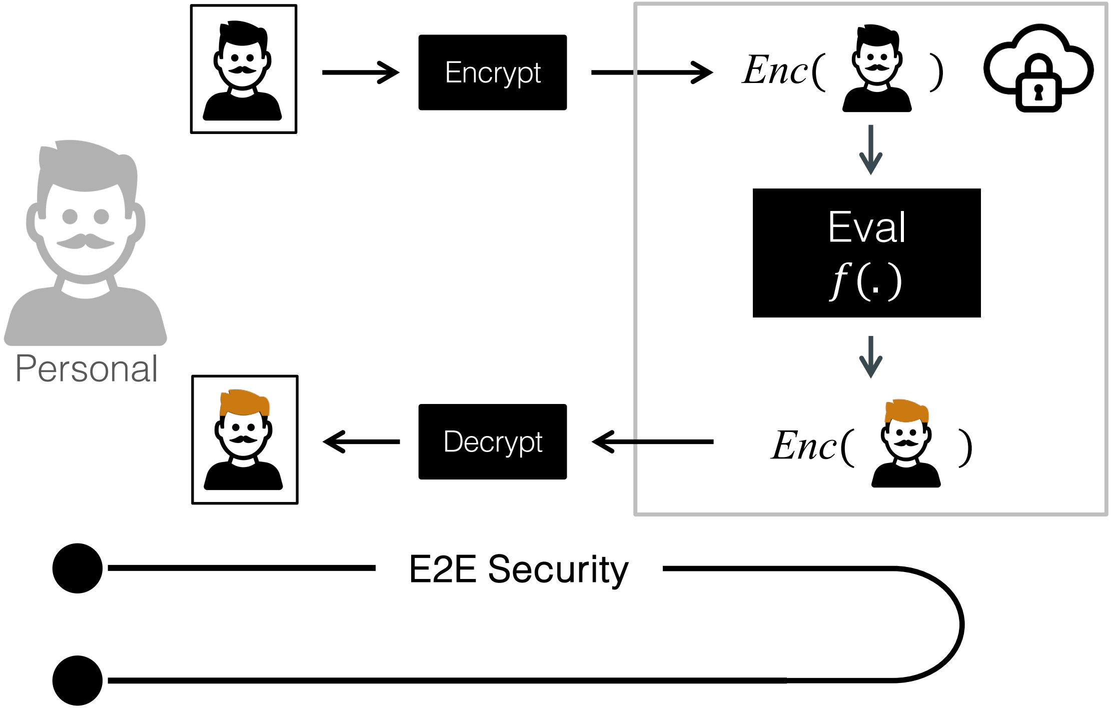
End-to-End Privacy

Privacy-Transparency Dichotomy

End-to-End Privacy

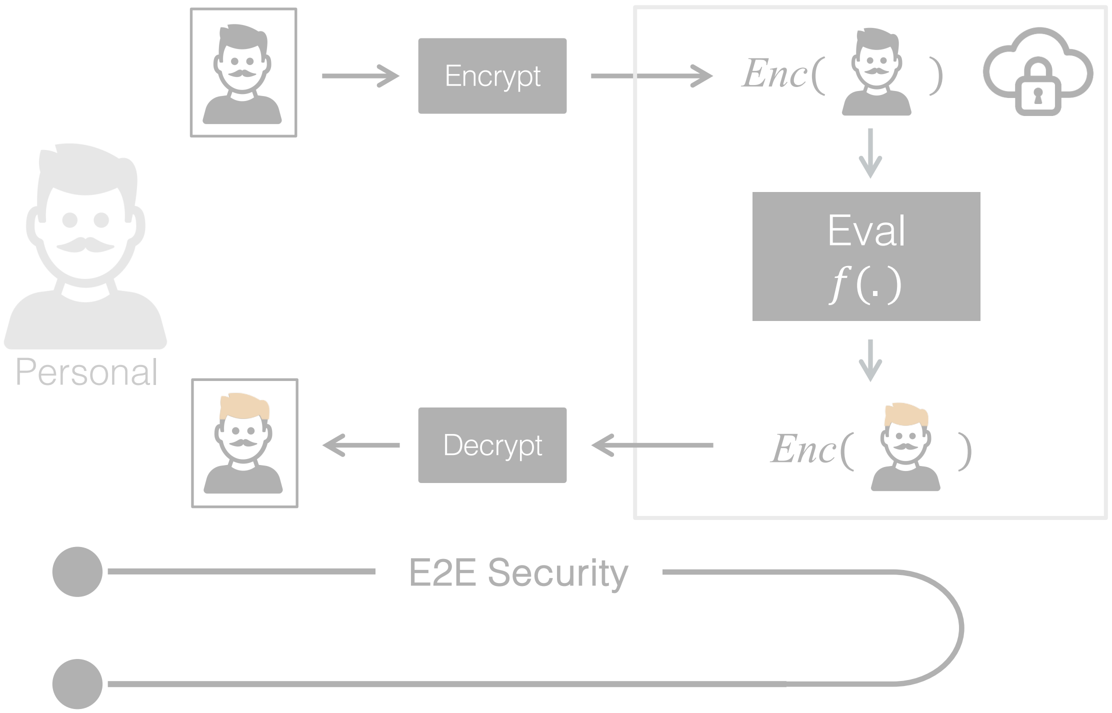
Secure Computation

Homomorphic Encryption | Secure Multi-party Computation



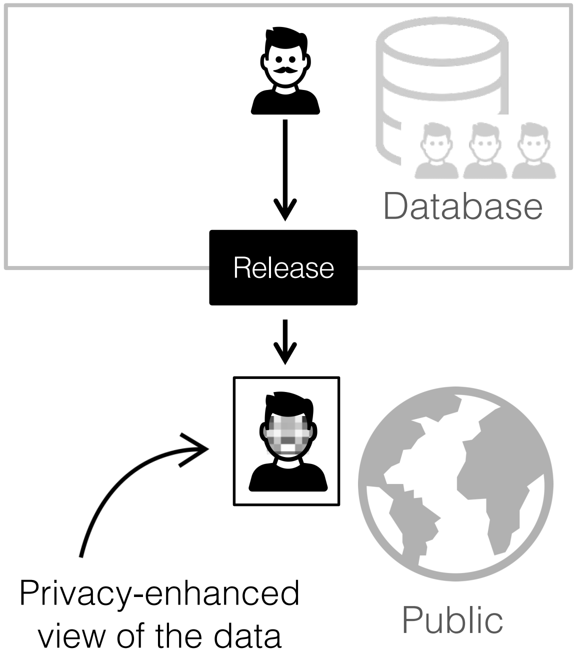
Secure Computation

Homomorphic Encryption | Secure Multi-party Computation



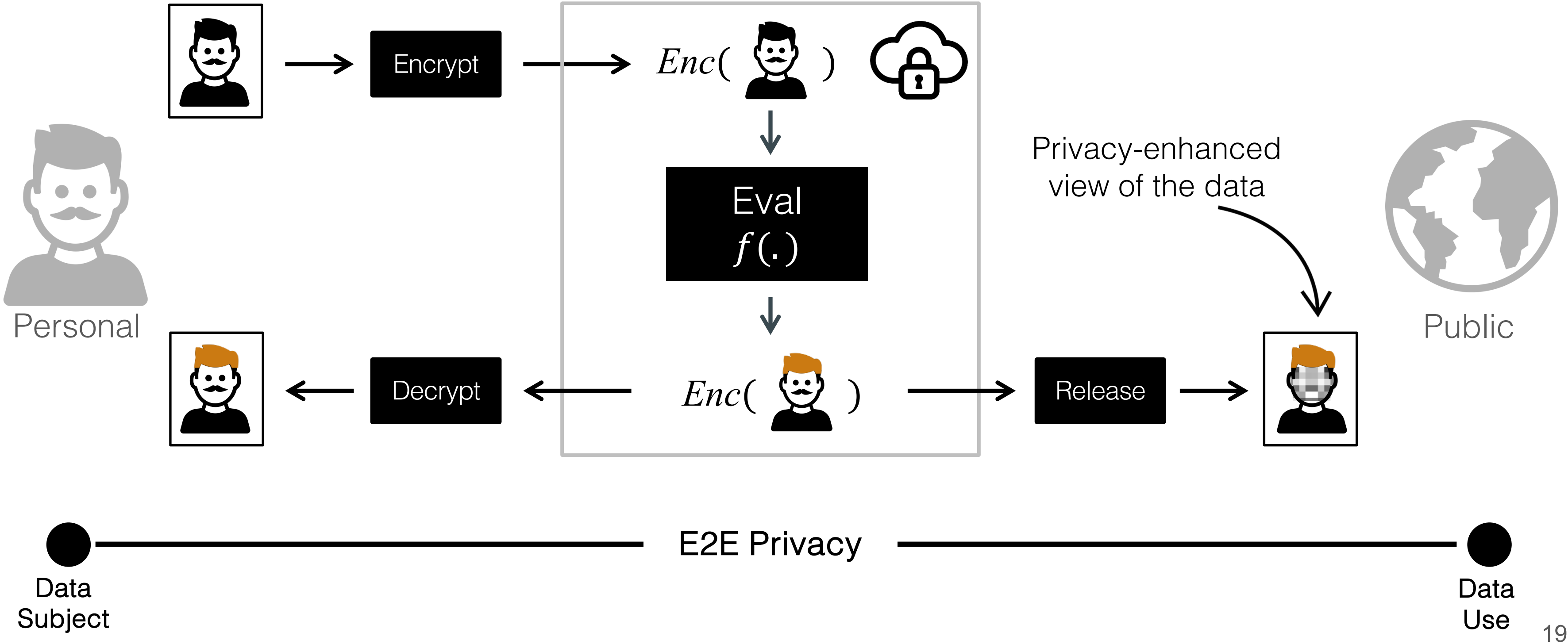
Releasing Data

Differential Privacy | Anonymization



End-to-End Privacy Platform

Homomorphic Encryption | Secure Multi-party Computation | Zero Knowledge Proofs | Differential Privacy

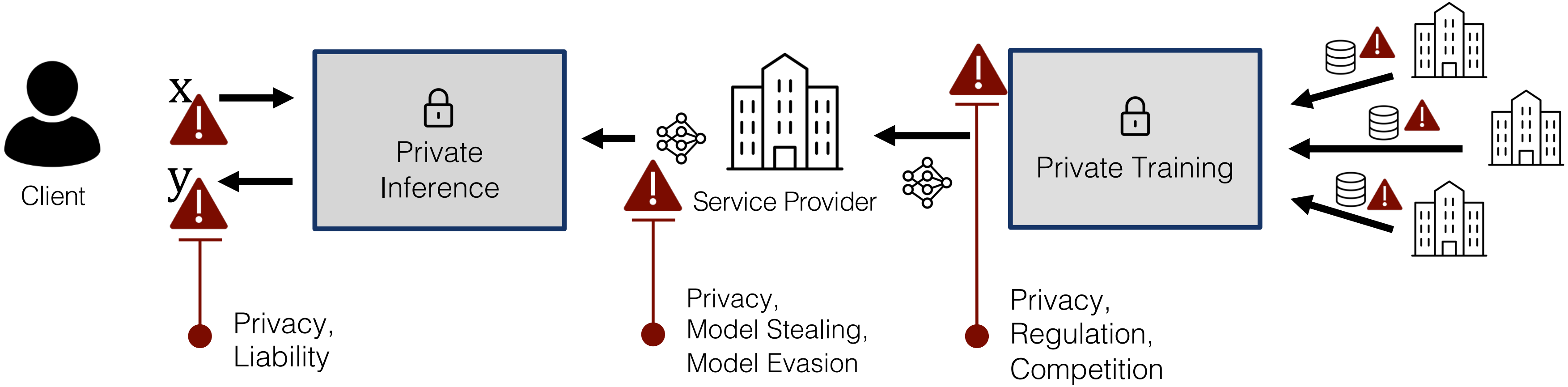


Privacy-Transparency Dichotomy

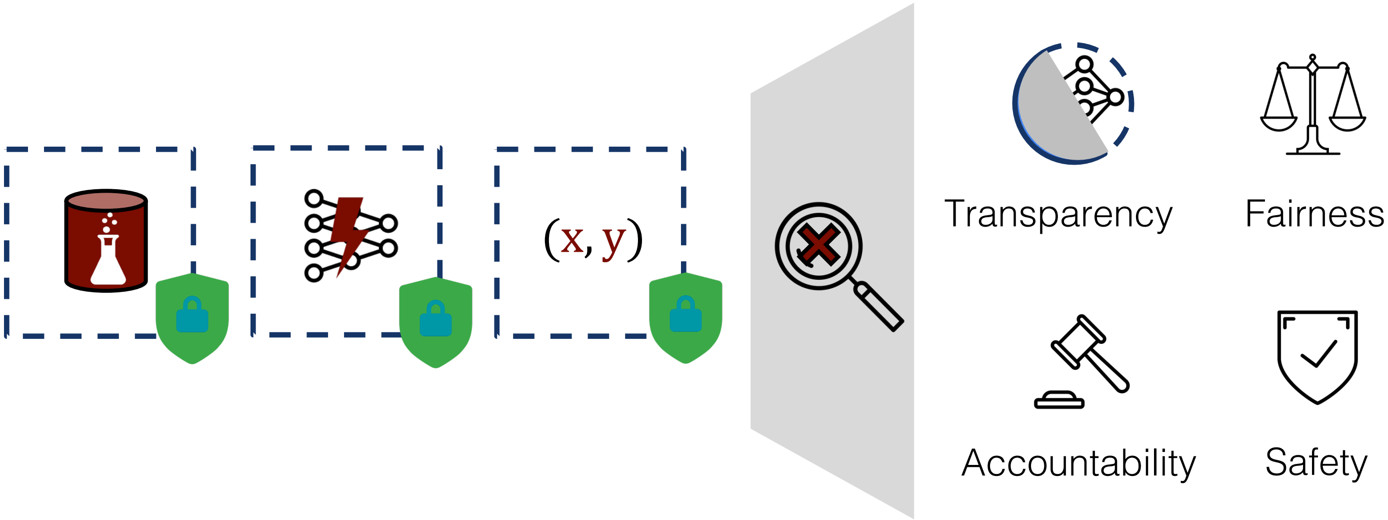
Privacy-Transparency Dichotomy

[Holding Secrets Accountable: Auditing Private ML Algorithms]

Privacy-Preserving Machine Learning



Verifiable Claims and Accountability in PPML



Acknowledgments

Students

Nicolas Küchler

Hidde Lycklama

Alexander Viand

Lukas Burkhalter

Miro Haller

Patrick Jattke

Christian Knabenhans

Emanuel Opel

Sponsors

Swiss National Science Foundation

intel

Meta

Talos: Encrypted Query Processing for the Internet of Things
Secure Sharing of Partially Homomorphic Encrypted IoT Data
TimeCrypt: Encrypted Data Stream Processing at Scale with Cryptographic Access Control
Droplet: Decentralized Authorization and Access Control for Encrypted Data Streams
Zeph: Cryptographic Enforcement of End-to-End Data Privacy
VF-PS: How to Select Important Participants in Vertical Federated Learning, Efficiently and Securely?
RoFL: Attestable Robustness for Secure Federated Learning

Privacy-Preserving
System Designs

Talos
ACM SenSys

Pilatus
ACM SenSys

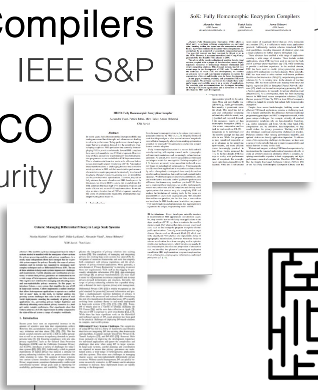
TimeCrypt
USENIX NSDI

Droplet
USENIX Security

Zeph
USENIX OSDI

VF-PS
NeurIPS

RoFL
IEEE S&P


SoK: Fully Homomorphic Encryption Compilers
HECO: Fully Homomorphic Encryption Compiler
Cohere: Managing Differential Privacy in Large Scale Systems

FHE Compilers
IEEE S&P

HECO
USENIX Security

Cohere
IEEE S&P

Democratize
Privacy-Preserving
Computation

My work aims to **build** practical systems that use
cryptography to empower users and preserve their privacy.