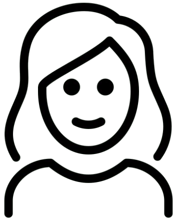
QuerySnout: A tool to Automate the Discovery of Privacy Vulnerabilities in Query-Based Systems

Ana-Maria Crețu, in collaboration with Florimond Houssiau (Alan Turing Institute, UK), Antoine Cully and Yves-Alexandre de Montjoye (Imperial College London, UK)

CNIL Privacy Research Day, 14/06/2023, Paris, France

An interface to compute answers about a private dataset

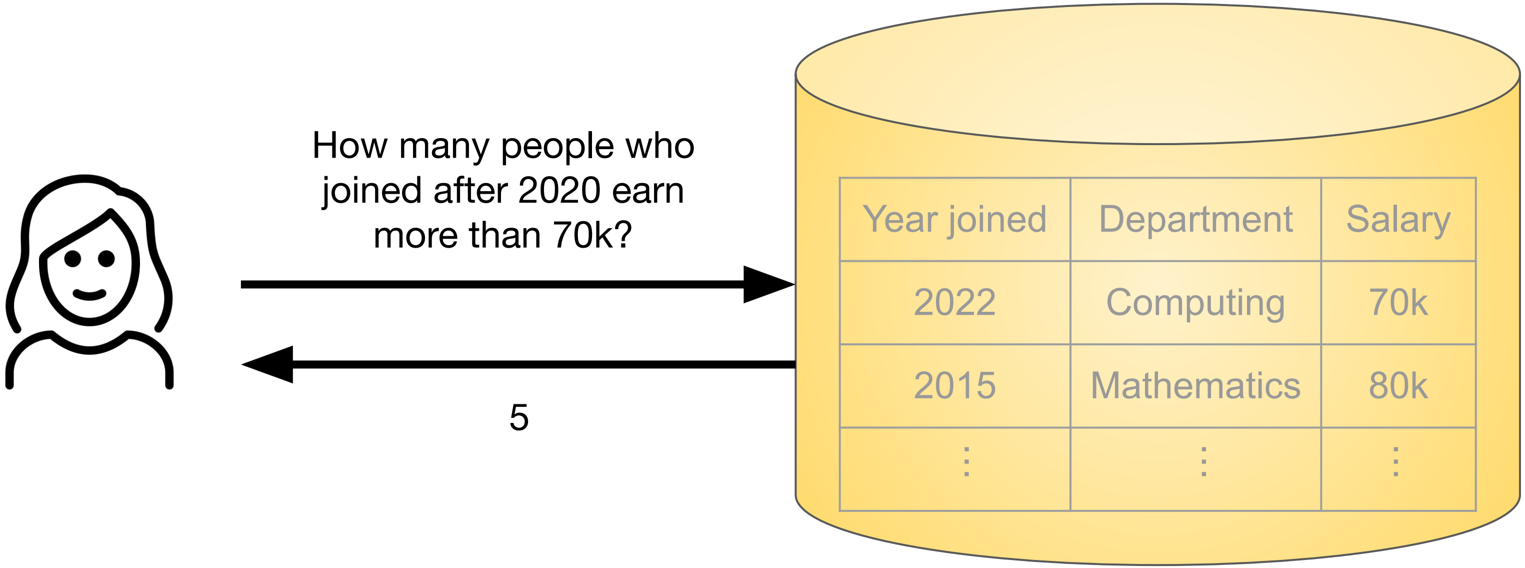


How many people who joined after 2020 earn more than 70k?

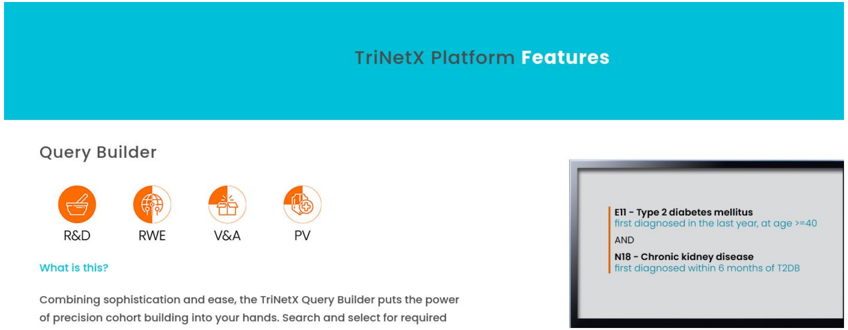
5

Year joined	Department	Salary
2022	Computing	70k
2015	Mathematics	80k
⋮	⋮	⋮

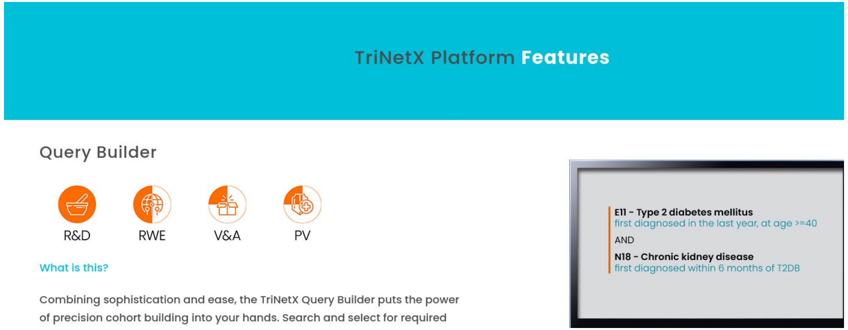
An interface to compute answers about a private dataset

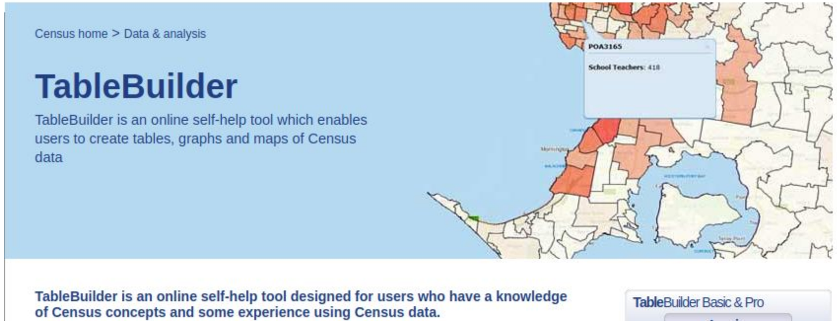


Query-based systems are a promising way to share data

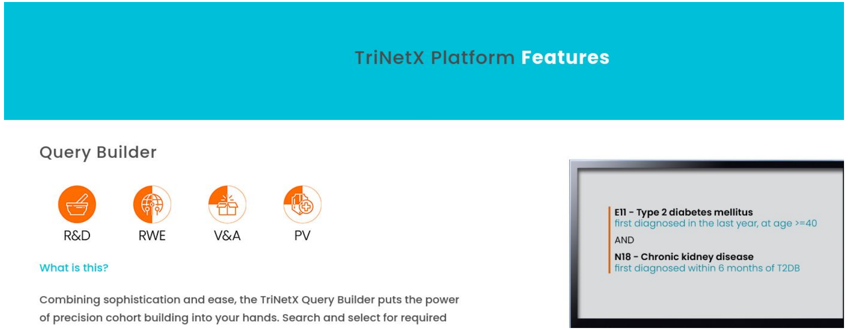


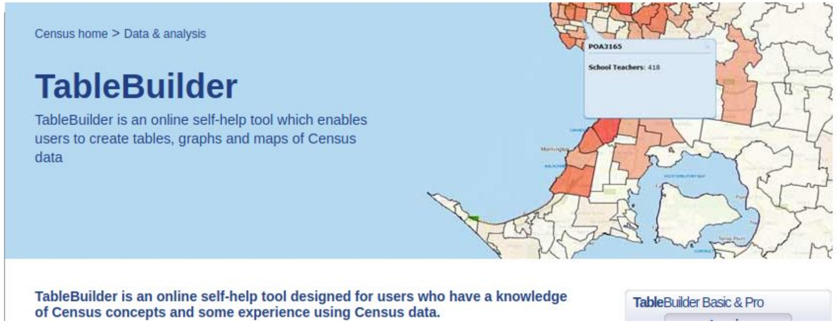
Query-based systems are a promising way to share data

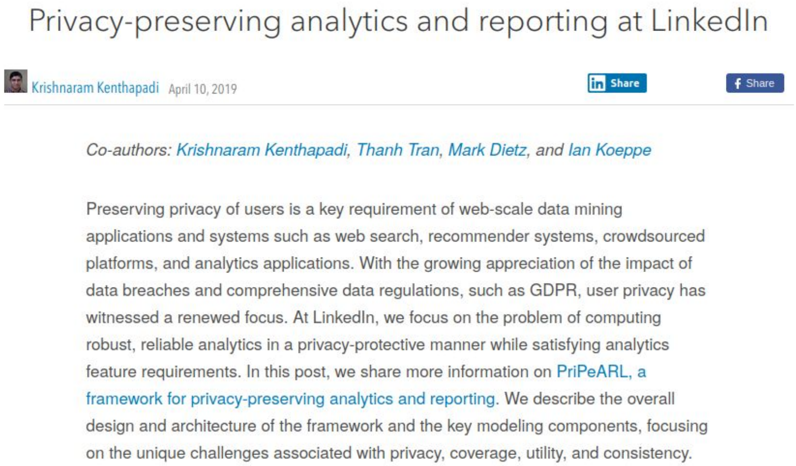




Query-based systems are a promising way to share data







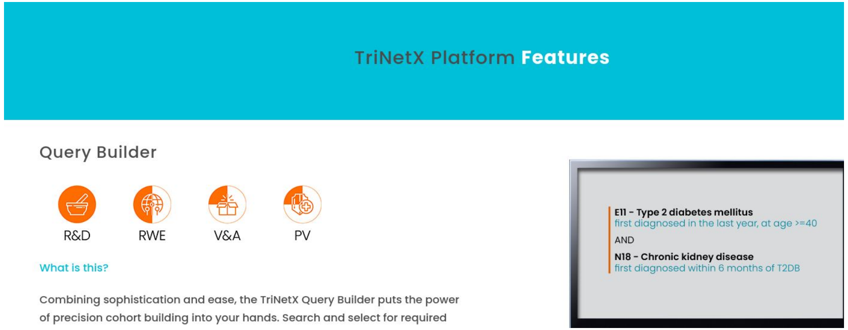
Privacy-preserving analytics and reporting at LinkedIn

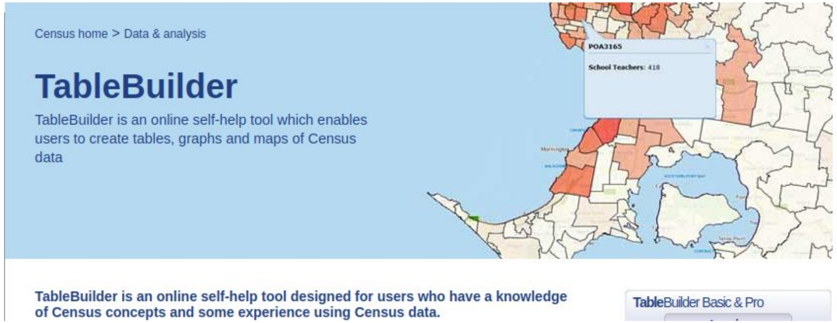
Krishnaram Kenthapadi April 10, 2019 Share Share

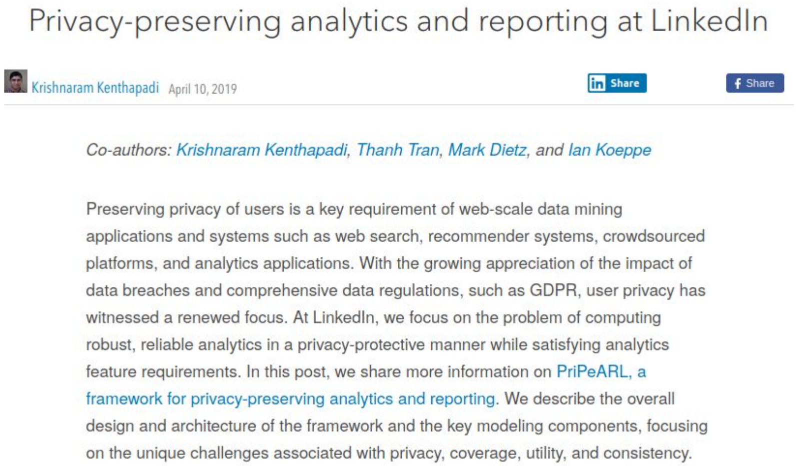
Co-authors: Krishnaram Kenthapadi, Thanh Tran, Mark Dietz, and Ian Koeppe

Preserving privacy of users is a key requirement of web-scale data mining applications and systems such as web search, recommender systems, crowdsourced platforms, and analytics applications. With the growing appreciation of the impact of data breaches and comprehensive data regulations, such as GDPR, user privacy has witnessed a renewed focus. At LinkedIn, we focus on the problem of computing robust, reliable analytics in a privacy-protective manner while satisfying analytics feature requirements. In this post, we share more information on PriPeARL, a framework for privacy-preserving analytics and reporting. We describe the overall design and architecture of the framework and the key modeling components, focusing on the unique challenges associated with privacy, coverage, utility, and consistency.

Query-based systems are a promising way to share data





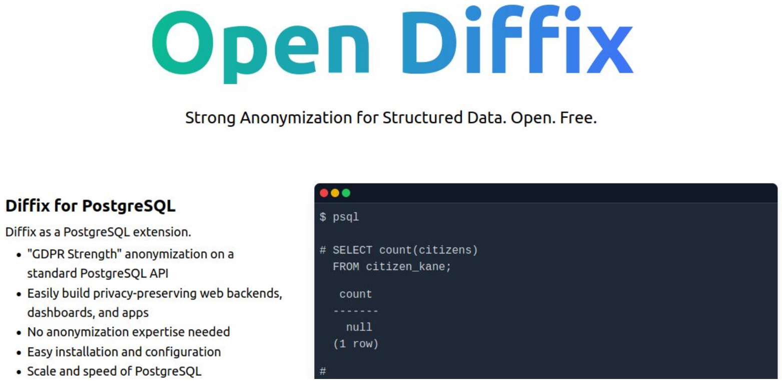


Privacy-preserving analytics and reporting at LinkedIn

Krishnaram Kenthapadi April 10, 2019 Share Share

Co-authors: Krishnaram Kenthapadi, Thanh Tran, Mark Dietz, and Ian Koeppe

Preserving privacy of users is a key requirement of web-scale data mining applications and systems such as web search, recommender systems, crowdsourced platforms, and analytics applications. With the growing appreciation of the impact of data breaches and comprehensive data regulations, such as GDPR, user privacy has witnessed a renewed focus. At LinkedIn, we focus on the problem of computing robust, reliable analytics in a privacy-protective manner while satisfying analytics feature requirements. In this post, we share more information on PriPeARL, a framework for privacy-preserving analytics and reporting. We describe the overall design and architecture of the framework and the key modeling components, focusing on the unique challenges associated with privacy, coverage, utility, and consistency.



Query-based systems can leak sensitive information

Assume I know Alice is the only person hired in the Department of Computing in 2022.

Year joined	Department	Salary
2022	Computing	70k
2015	Mathematics	80k
⋮	⋮	⋮

Query-based systems can leak sensitive information

Assume I know Alice is the only person hired in the Department of Computing in 2022.

Q1: How many people in the Department of Computing hired in 2022 earn a salary of 70k?

Year joined	Department	Salary
2022	Computing	70k
2015	Mathematics	80k
⋮	⋮	⋮

Query-based systems can leak sensitive information

Assume I know Alice is the only person hired in the Department of Computing in 2022.

Q1: How many people in the Department of Computing hired in 2022 earn a salary of 70k? 0 or 1

Year joined	Department	Salary
2022	Computing	70k
2015	Mathematics	80k
⋮	⋮	⋮

Query-based systems can leak sensitive information

Assume I know Alice is the only person hired in the Department of Computing in 2022.

Q1: How many people in the Department of Computing hired in 2022 earn a salary of 70k? 0 or 1

Always 0

\+ **Query-set size restriction**

Year joined	Department	Salary
2022	Computing	70k
2015	Mathematics	80k
⋮	⋮	⋮

Query-based systems can leak sensitive information

Assume I know Alice is the only person hired in the Department of Computing in 2022.

+ **Query-set size restriction**

Year joined	Department	Salary
2022	Computing	70k
2015	Mathematics	80k
⋮	⋮	⋮

Query-based systems can leak sensitive information

Assume I know Alice is the only person hired in the Department of Computing in 2022.

Q'$_1$: How many people in the Department of Computing earn a salary of 70k?

Q'$_2$: How many people in the Department of Computing **who did not join in 2022 earn a salary of 70k?**

+ Query-set size restriction

Year joined	Department	Salary
2022	Computing	70k
2015	Mathematics	80k
⋮	⋮	⋮

Query-based systems can leak sensitive information

Assume I know Alice is the only person hired in the Department of Computing in 2022.

Q'$_1$: How many people in the Department of Computing earn a salary of 70k? → x

Q'$_2$: How many people in the Department of Computing **who did not join in 2022 earn a salary of 70k?** → x or x-1

\+ **Query-set size restriction**

Year joined	Department	Salary
2022	Computing	70k
2015	Mathematics	80k
⋮	⋮	⋮

Query-based systems can leak sensitive information

Assume I know Alice is the only person hired in the Department of Computing in 2022.

Q'$_1$: How many people in the Department of Computing earn a salary of 70k? → x

Q'$_2$: How many people in the Department of Computing **who did not join in 2022 earn a salary of 70k?** → x or x-1

+ Query-set size restriction

Year joined	Department	Salary
2022	Computing	70k
2015	Mathematics	80k
⋮	⋮	⋮

Attack: Q'$_1$- Q'$_2$ > 0 implies that Alice's salary is 70k

[1] Denning D. E. and Denning P. J. The tracker: A threat to statistical database security. TODS (1979).

Query-based systems can leak sensitive information

Assume I know Alice is the only person hired in the Department of Computing in 2022.

Q'_1: How many people in the Department of Computing earn a salary of 70k? → x

→ x + 

Q'_2: How many people in the Department of Computing **who did not join in 2022 earn a salary of 70k?** → x or x-1

→ x or x-1 + 

+ **Noise addition**
+ **Query-set size restriction**

Year joined	Department	Salary
2022	Computing	70k
2015	Mathematics	80k
⋮	⋮	⋮

Attack: $Q'_1 - Q'_2 > 0$ implies that Alice's salary is 70k

[1] Denning D. E. and Denning P. J. The tracker: A threat to statistical database security. TODS (1979).

Query-based systems can leak sensitive information

Assume I know Alice is the only person hired in the Department of Computing in 2022.

Q'_1: How many people in the Department of Computing earn a salary of 70k? → x +

Q'_2: How many people in the Department of Computing **who did not join in 2022 earn a salary of 70k?** → x or x-1 +

\+ **Noise addition**

\+ **Query-set size restriction**

Year joined	Department	Salary
2022	Computing	70k
2015	Mathematics	80k
⋮	⋮	⋮

Query-based systems can leak sensitive information

Assume I know Alice is the only person hired in the Department of Computing in 2022.

Q'_1: How many people in the Department of Computing earn a salary of 70k? → x +

Q'_2:

not

Q'_1: How many people in the Department of Computing earn a salary of 70k? → x +

Q'_2: How many people in the Department of Computing **who did not join in 2022 earn a salary of 70k?** → x or x-1 +

+ **Noise addition**
+ **Query-set size restriction**

Year joined	Department	Salary
2022	Computing	70k
2015	Mathematics	80k
⋮	⋮	⋮

Query-based systems can leak sensitive information

Assume I know Alice is the only person hired in the Department of Computing in 2022.

Q'$_1$: How many people in the Department of Computing earn a salary of 70k? → x +

Q'$_2$:

not

Q'$_1$: How many people in the Department of Computing earn a salary of 70k? → x +

Q'$_2$:

not j

Q'$_1$: How many people in the Department of Computing earn a salary of 70k? → x +

Q'$_2$: How many people in the Department of Computing **who did not join in 2022 earn a salary of 70k?** → x or x-1 +

\+ **Noise addition**

\+ **Query-set size restriction**

Year joined	Department	Salary
2022	Computing	70k
2015	Mathematics	80k
⋮	⋮	⋮

Query-based systems can leak sensitive information

Assume I know Alice is the only person hired in the Department of Computing in 2022.

Q'$_1$: How many people in the Department of Computing earn a salary of 70k? → x +

Q'$_2$:

not

Q'$_1$: How many people in the Department of Computing earn a salary of 70k? → x +

Q'$_2$:

not j

Q'$_1$: How many people in the Department of Computing earn a salary of 70k? → x +

Q'$_2$: How many people in the Department of Computing **who did not join in 2022 earn a salary of 70k?** → x or x-1 +

- **Deterministic noise addition**
- **Query-set size restriction**

Year joined	Department	Salary
2022	Computing	70k
2015	Mathematics	80k
⋮	⋮	⋮

[2] O'Keefe, C.M., Haslett, S., Steel, D., and Chambers, R. Table builder problem-confidentiality for linked tables. (2008)

[3] Diffix-Birch: Extending Diffix-Aspen. Francis, P. et al. https://arxiv.org/abs/1806.02075 (2018)

[4] Felix Bauer. 2017. Announcing the first ever bug bounty program for a privacy protection solution.

[5] Gadotti, A., Houssiau, F., Rocher, L., Livshits, B., and de Montjoye Y.-A. When the signal is in the noise: Exploiting Diffix's Sticky Noise. USENIX Security, (2019).

[6] Pyrgelis, A. On Location, Time, and Membership: Studying How Aggregate Location Data Can Harm Users' Privacy. (2018)

Can we automate the discovery of attacks against query-based systems?

Attribute inference threat model

Consider a database D protected by a query-based system (QBS).

Attribute inference threat model

Consider a database D protected by a query-based system (QBS).

Year joined	Department	Salary
2022	Computing	?

Alice’s partial record

Alice’s unknown sensitive attribute

Attribute inference threat model

Consider a database D protected by a query-based system (QBS).

Year joined	Department	Salary
2022	Computing	?

Alice’s partial record

Alice’s unknown sensitive attribute

⚠️ A successful attack leads to the disclosure of Alice’s private information.

Search space of attacks

The attacks we search for consist of two components:

1. A set of queries $q_1, \ldots, q_m$

2. A mathematical (arithmetic) function G to combine their answers in order to infer the sensitive attribute s.

Search space of attacks

The attacks we search for consist of two components:

1. A set of queries $q_1, \ldots, q_m$

2. A mathematical (arithmetic) function G to combine their answers in order to infer the sensitive attribute s.

Our goal is to find the best queries and the best function G allowing to accurately infer s.

QuerySnout

1. Given any set of queries, we use machine learning to learn a function G that infers the unknown attribute based on their answers.

QuerySnout

1. Given any set of queries, we use machine learning to learn a function G that infers the unknown attribute based on their answers.
2. We explore the space of queries to find the best attacks using evolutionary search techniques.

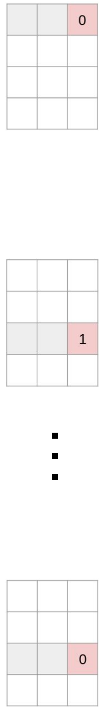
QuerySnout

1. Given any set of queries, we use machine learning to learn a function G that infers the unknown attribute based on their answers.
2. We explore the space of queries to find the best attacks using evolutionary search techniques.

Learning how to combine answers to queries $q_1, \ldots, q_m$

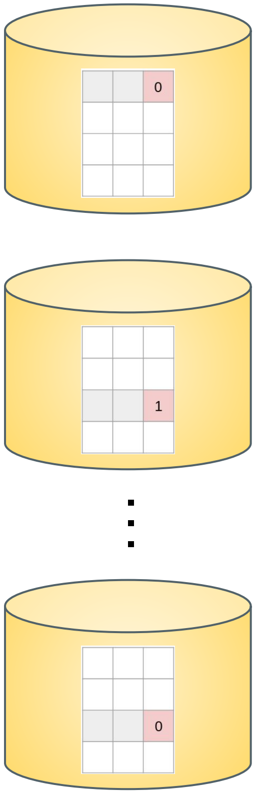
Learning how to combine answers to queries $q_1, \ldots, q_m$

Step 1: Sample "proxy" datasets



Learning how to combine answers to queries $q_1, \ldots, q_m$

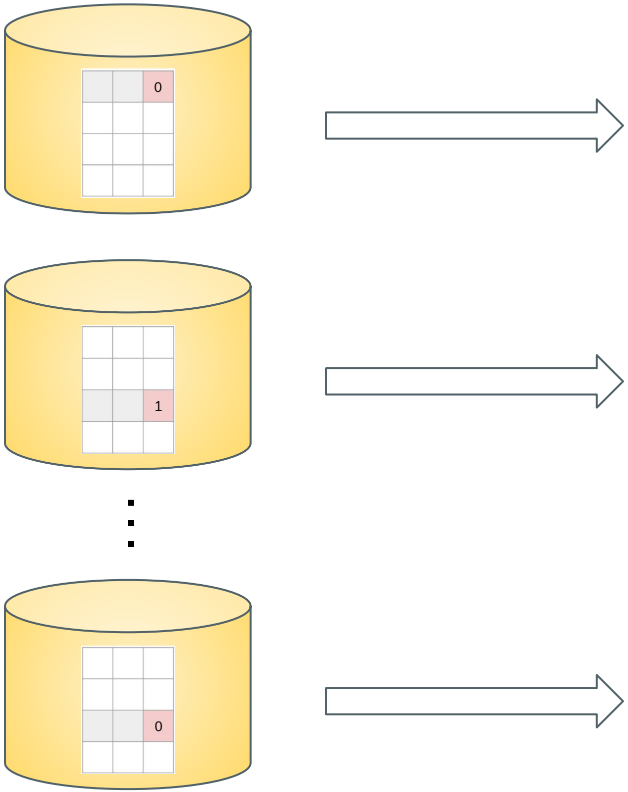
Step 1: Sample "proxy" datasets



Learning how to combine answers to queries $q_1, \ldots, q_m$

Step 1: Sample "proxy" datasets

Step 2: Retrieve answers



$\mathbf{q_1}$ $\mathbf{q_2}$ $\mathbf{q_3}$ $\mathbf{q_m}$

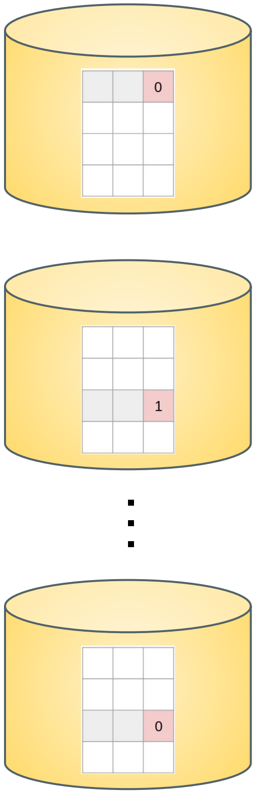
20, 21, 5, …, 50

19, 19, 3, … , 54

24, 25, 7, …., 44

Learning how to combine answers to queries $q_1, \ldots, q_m$

Step 1: Sample "proxy" datasets

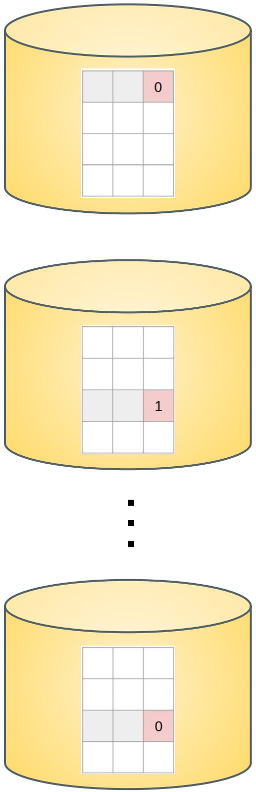


Step 2: Retrieve answers

X_{train} (q_1, q_2, q_3, …, q_m)	Y_{train}
20, 21, 5, …, 50	0
19, 19, 3, … , 54	1
24, 25, 7, …., 44	0

Learning how to combine answers to queries $q_1, \ldots, q_m$

Step 1: Sample “proxy” datasets



Step 2: Retrieve answers

X_{train}: q_1 q_2 q_3 … q_m	Y_{train}
20, 21, 5, …, 50	0
19, 19, 3, … , 54	1
24, 25, 7, …., 44	0

Step 3: Learn a rule to combine answers to queries

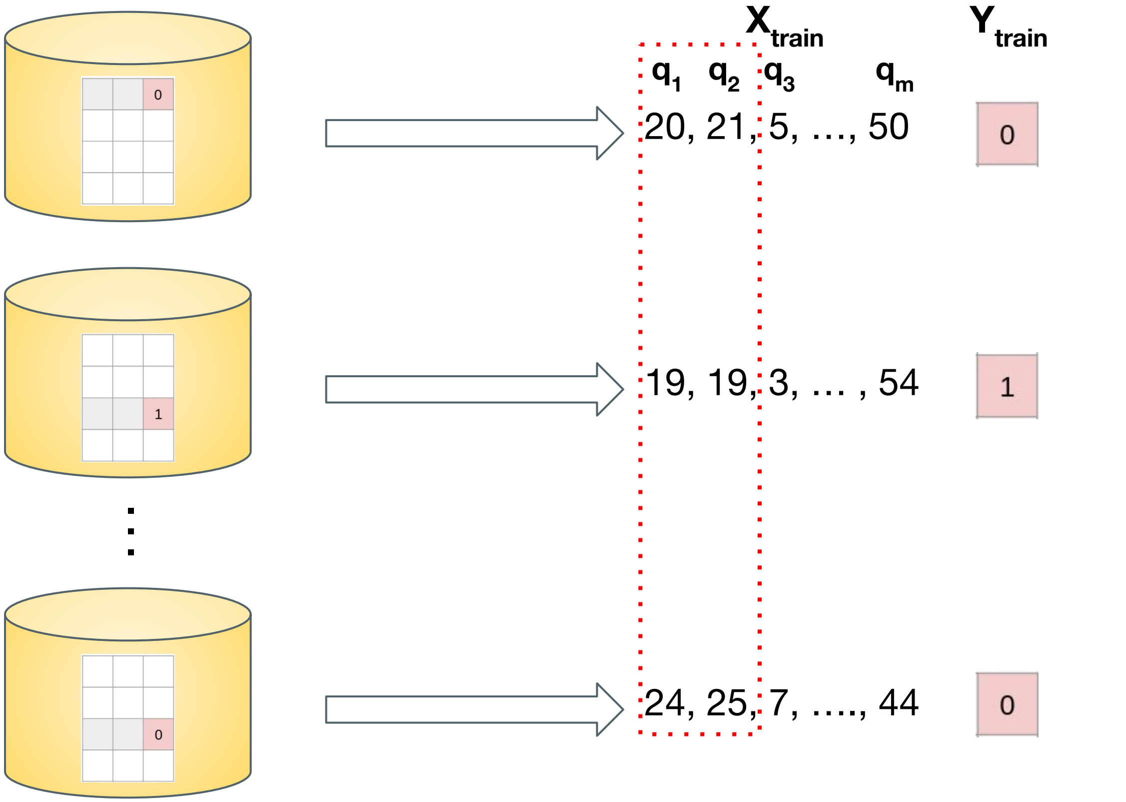
Machine learning classifier

X_{train} Y_{train}

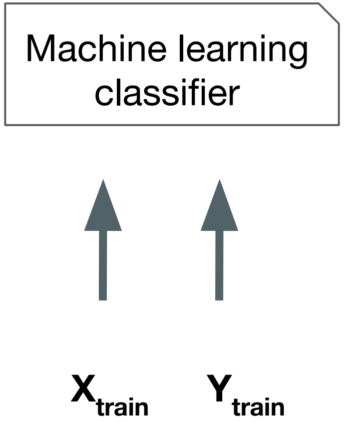
Learning how to combine answers to queries $q_1, \ldots, q_m$

Step 1: Sample "proxy" datasets

Step 2: Retrieve answers



Step 3: Learn a rule to combine answers to queries



Results

Empirical validation

- We apply QuerySnout to two real-world mechanisms implementing deterministic noise addition.

Empirical validation

- We apply QuerySnout to two real-world mechanisms implementing deterministic noise addition.
- We compare the performance of QuerySnout with manual attacks from prior works.

Empirical validation

- We apply QuerySnout to two real-world mechanisms implementing deterministic noise addition.
- We compare the performance of QuerySnout with manual attacks from prior works.

Diffix mechanism

(a) **AUXILIARY**	Adult	Census	Insurance
QuerySnout (automated)	**77.8** (0.5)	**78.3** (1.4)	**80.1** (0.6)
Gadotti et al. [30] (manual)	76.3 (0.8)	76.9 (1.4)	73.0 (1.2)

The attacks found by QuerySnout correctly guess Alice's secret 4 times out of 5

Table Builder mechanism

(a) **AUXILIARY**	Adult	Census	Insurance
QuerySnout (automated)	**84.5** (0.6)	**85.5** (1.4)	**85.4** (0.6)
Rinott et al.[56] (manual)	76.1 (7.5)	78.1 (7.0)	56.9 (4.6)
Chip. et al.[11] (manual)	61.2 (3.5)	62.4 (3.1)	52.8 (1.8)

Empirical validation

- We apply QuerySnout to two real-world mechanisms implementing deterministic noise addition.
- We compare the performance of QuerySnout with manual attacks from prior works.

Diffix mechanism

(a) AUXILIARY	Adult	Census	Insurance
QuerySnout (automated)	**77.8** (0.5)	**78.3** (1.4)	**80.1** (0.6)
Gadotti et al. [30] (manual)	76.3 (0.8)	76.9 (1.4)	73.0 (1.2)

(b) EXACT-BUT-ONE	Adult	Census	Insurance
QuerySnout (automated)	**90.2** (0.6)	**88.3** (0.9)	**91.6** (1.2)
Gadotti et al. [30] (manual)	77.1 (0.9)	77.5 (2.0)	74.4 (0.7)

Table Builder mechanism

(a) AUXILIARY	Adult	Census	Insurance
QuerySnout (automated)	**84.5** (0.6)	**85.5** (1.4)	**85.4** (0.6)
Rinott et al.[56] (manual)	76.1 (7.5)	78.1 (7.0)	56.9 (4.6)
Chip. et al.[11] (manual)	61.2 (3.5)	62.4 (3.1)	52.8 (1.8)

(b) EXACT-BUT-ONE	Adults	Census	Insurance
QuerySnout (automated)	**98.1** (0.7)	**96.6** (0.9)	**98.8** (0.7)
Rinott et al.[56] (manual)	83.1 (8.7)	72.1 (13.4)	76.5 (2.3)
Chip. et al.[11] (manual)	72.3 (6.4)	64.4 (7.2)	67.2 (1.4)

Empirical validation

- We apply QuerySnout to two real-world mechanisms implementing deterministic noise addition.
- We compare the performance of QuerySnout with manual attacks from prior works.

Diffix mechanism

(a) AUXILIARY	Adult	Census	Insurance
QuerySnout (automated)	**77.8** (0.5)	**78.3** (1.4)	**80.1** (0.6)
Gadotti et al. [30] (manual)	76.3 (0.8)	76.9 (1.4)	73.0 (1.2)

(b) EXACT-BUT-ONE	Adult	Census	Insurance
QuerySnout (automated)	**90.2** (0.6)	**88.3** (0.9)	**91.6** (1.2)
Gadotti et al. [30] (manual)	77.1 (0.9)	77.5 (2.0)	74.4 (0.7)

Table Builder mechanism

(a) AUXILIARY	Adult	Census	Insurance
QuerySnout (automated)	**84.5** (0.6)	**85.5** (1.4)	**85.4** (0.6)
Rinott et al.[56] (manual)	76.1 (7.5)	78.1 (7.0)	56.9 (4.6)
Chip. et al.[11] (manual)	61.2 (3.5)	62.4 (3.1)	52.8 (1.8)

(b) EXACT-BUT-ONE	Adults	Census	Insurance
QuerySnout (automated)	**98.1** (0.7)	**96.6** (0.9)	**98.8** (0.7)
Rinott et al.[56] (manual)	83.1 (8.7)	72.1 (13.4)	76.5 (2.3)
Chip. et al.[11] (manual)	72.3 (6.4)	64.4 (7.2)	67.2 (1.4)

👉 We obtain similar results on other systems, including non-deterministic ones and a budget-based mechanism implementing ε-differential privacy.

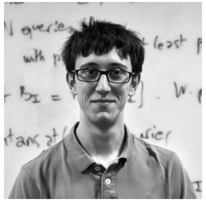
Takeaways

👉 Query-based systems (QBS) are one of the main ways to share data anonymously.

👉 QuerySnout can automatically find privacy vulnerabilities in QBSs, at the “pressing of a button” □.

👉 Our work opens the door to automatically auditing the privacy of QBSs in a context-dependent way.

👉 Our code is available at https://github.com/computationalprivacy/querysnout.

@AnaMariaCretu5 @fhoussiau @CULLYAntoine @yvesalexandre

Thank you!

Can automated attacks be mitigated by reducing the number of queries?

Limiting the number of queries is a popular defense.

👉 The performance of QuerySnout increases with the number of queries.

👉 It's still very good even with 10 queries.

👉 It consistently outperforms the random search and the random solution.