Algorithmic Privacy
What does algorithmic data privacy mean?
What does algo/data privacy mean?

- Control of access
- Control of use
- Knowledge of access/use
- Freedom from surveillance
- Ownership of data
- Opt in/out
- Anonymity
Privacy vs. Security

- Security: control access to "raw" data
  - Locks, keys, crypto

- Privacy: allow use of data but control inferences/exfiltration
  - Anonymization, differential privacy
**Privacy vs. Security**

<table>
<thead>
<tr>
<th>Name</th>
<th>Age</th>
<th>SSN</th>
<th>Major</th>
<th>GPA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Joe</td>
<td>19</td>
<td>547...</td>
<td>CIS</td>
<td>2.7</td>
</tr>
<tr>
<td>Mary</td>
<td>18</td>
<td>233...</td>
<td>Math</td>
<td>3.5</td>
</tr>
<tr>
<td>Bill</td>
<td>20</td>
<td>713...</td>
<td>English</td>
<td>3.2</td>
</tr>
<tr>
<td>Eve</td>
<td>17</td>
<td>492...</td>
<td>History</td>
<td>3.4</td>
</tr>
</tbody>
</table>

- **Security**: Keep locked, control access
- **Privacy**: Model should not reveal Joe's GPA, Eve's SSN
A Bit of (Old School) Crypto

1. Suppose I want to send you a message \( a \in \{0, 1\} \).
2. We first meet and choose a random \( b \in \{0, 1\} \).
3. Later, I send you:

   \[
   c = a \oplus b = \begin{cases} 
   0 & \text{if } a \neq b \\
   1 & \text{if } a = b 
   \end{cases}
   \]

4. You decrypt:

   \[
   c \oplus b = (a \oplus b) \oplus b = a \oplus (b \oplus b) = a
   \]
A Bit of (Old School) Crypto

- Eavesdropper sees only $c$ which is random
- "One-time pad"
- Why? $(a \oplus b) \oplus (a' \oplus b) = a \oplus a'$
- Longer pads for longer messages/files
Public-Key Crypto

- OTP security is absolute
- But:
  - Have to meet privately
  - Keys are long
  - Every exchange needs new/different keys

- Public-key crypto:
  - Separates encryption/decryption
  - Encryption keys public
  - Security based on computational hardness
  - Underlies modern Internet (e.g. HTTPS)
Goldilocks & the 3 Privacy

Anonymization: Not private enough
(too useful)

Differential Privacy: Strong privacy & utility

"No Harm whatsoever": Too private (not useful)
“Anonymization”

- Basic idea: start with some sensitive dataset $D$, transform to anonymized version $D'$
- $D \rightarrow D'$ by two operations:
  - Redaction: remove entire fields/columns
  - Coarsening: reduce resolution of field
### Anonymized Data Isn’t

<table>
<thead>
<tr>
<th>Name</th>
<th>Age</th>
<th>Gender</th>
<th>Zip Code</th>
<th>Smoker</th>
<th>Diagnosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>*</td>
<td>60-70</td>
<td>Male</td>
<td>191**</td>
<td>Y</td>
<td>Heart disease</td>
</tr>
<tr>
<td>*</td>
<td>60-70</td>
<td>Female</td>
<td>191**</td>
<td>N</td>
<td>Arthritis</td>
</tr>
<tr>
<td>*</td>
<td>60-70</td>
<td>Male</td>
<td>191**</td>
<td>Y</td>
<td>Lung cancer</td>
</tr>
<tr>
<td>*</td>
<td>60-70</td>
<td>Female</td>
<td>191**</td>
<td>N</td>
<td>Crohn’s Disease</td>
</tr>
<tr>
<td>*</td>
<td>60-70</td>
<td>Male</td>
<td>191**</td>
<td>Y</td>
<td>Lung Cancer</td>
</tr>
<tr>
<td>*</td>
<td>50-60</td>
<td>Female</td>
<td>191**</td>
<td>N</td>
<td>HIV</td>
</tr>
<tr>
<td>*</td>
<td>50-60</td>
<td>Male</td>
<td>191**</td>
<td>Y</td>
<td>Lyme Disease</td>
</tr>
<tr>
<td>*</td>
<td>50-60</td>
<td>Male</td>
<td>191**</td>
<td>Y</td>
<td>Seasonal Allergies</td>
</tr>
<tr>
<td>*</td>
<td>50-60</td>
<td>Female</td>
<td>191**</td>
<td>N</td>
<td>Ulcerative Colitis</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Name</th>
<th>Age</th>
<th>Gender</th>
<th>Zip Code</th>
<th>Diagnosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>*</td>
<td>50-60</td>
<td>Female</td>
<td>191**</td>
<td>HIV</td>
</tr>
<tr>
<td>*</td>
<td>50-60</td>
<td>Female</td>
<td>191**</td>
<td>Lupus</td>
</tr>
<tr>
<td>*</td>
<td>50-60</td>
<td>Female</td>
<td>191**</td>
<td>Hip Fracture</td>
</tr>
<tr>
<td>*</td>
<td>60-70</td>
<td>Male</td>
<td>191**</td>
<td>Pancreatic Cancer</td>
</tr>
<tr>
<td>*</td>
<td>60-70</td>
<td>Male</td>
<td>191**</td>
<td>Ulcerative Colitis</td>
</tr>
<tr>
<td>*</td>
<td>60-70</td>
<td>Male</td>
<td>191**</td>
<td>Flu Like Symptoms</td>
</tr>
</tbody>
</table>
A Precise Definition

- Say $D'$ is $K$-anonymous (w.r.t. certain columns) if for any tuple of values in $D'$, there are at least $K$ copies.

- Privacy by confusion?

- What guarantees does this give you?
Anonymization...

- Is brittle
- Pretends D is the only dataset in the world
- Has no meaningful semantics
- Is demonstrably vulnerable to reidentification attacks
Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization

Paul Ohm

Computer scientists have recently undermined our faith in the privacy-protecting power of anonymization, the name for techniques that protect the privacy of individuals in large databases by deleting information like names and social security numbers. These scientists have demonstrated that they can often "reidentify" or "deanonymize" individuals hidden in anonymized data with astonishing ease. By understanding this research, we realize we have made a mistake, labored beneath a fundamental misunderstanding, which has assured us much less privacy than we have assumed. This mistake pervades nearly every information privacy law, regulation, and debate, yet regulators and legal scholars have paid it scant attention. We must respond to the surprising failure of anonymization, and this Article provides the tools to do so.

INTRODUCTION..................................................................................................................1703
I. ANONYMIZATION AND REIDENTIFICATION ......................................................1706
   A. The Past: Robust Anonymization ..................................................................1706
       1. Ubiquitous Anonymization ......................................................................1707
          a. The Anonymization/Reidentification Model ....................................1707

* Associate Professor, University of Colorado Law School. This Article was presented at the Privacy Law Scholars Conference and at conferences and faculty workshops at Harvard's Center for Research and Computer Science and Berkman Center, Princeton's Center for Information Technology Policy, Fordham University Center for Law and Information Policy, University of Washington School of Law, University of Washington's Computer Science & Engineering Department, NYU Information Law Institute, DePaul Center for IP Law and Information Technology, International Association of Privacy Professionals Global Privacy Summit, and the University of Colorado Law School. I thank all participants for their comments.

Thanks in particular to Caspar Bowden, Ramon Caceres, Ryan Calo, Deborah Cantrell, Danielle Citron, Nestor Davidson, Pierre de Vries, Vasant Dhar, Cynthia Dwork, Jed Ela, Ed Felten, Victor Fleischer, Susan Freiwald, Brett Frischmann, Michael Froomkin, Simson Garfinkel, Lauren Gelman, Eric Goldman, James Grimmelmann, Mike Hintze, Chris Hoofnagle, Clare Huntington, Jeff Jonas, Jerry Kang, Nancy Kim, Jon Kleinberg, Sarah Krakoff, Tim Lee, William McGeveran, Deven McGraw, Viva Moffat, Tyler Moore, Arvind Narayanan, Helen Nissenbaum, Scott Peppett, Jules Polonetsky, Foster Provost, Joel Reidenberg, Ira Rubinstein, Andrew Schwartz, Ari Schwartz, Vitaly Shmatikov, Chris Sohoian, Dan Solove, Latanya Sweeney, Peter Swire, Salil Vadhan, Michael Waggoner, Phil Weiser, Rebecca Wright, Felix Wu, and Michael Zimmer for their comments. This research was supported by a pre-tenure research leave grant by the University of Colorado Law School, and for this I thank Dean David Getches and Associate Dean Dayna Matthew. Finally, I thank my research assistant, Jerry Green.
Another Try...

- Anonymization: no actual privacy guarantee
- "No harm whatsoever" privacy
- Compare:

```
D
  ↓
A  ↓
   ↓
you
  ↓
result
```

```
D
  ↓
A  ↓
   ↓
you
  ↓
result
```
• Chance of (any) harm to you in A & B should be identical.

• Good definition?

• Let's consider smoking & lung cancer...
So... what's a better idea?
A More Refined Comparison

- Compare:

D

\[ \text{you} \]

A

r

D'

\[ \text{you} \]

A

r'

r & r' should be "indistinguishable"
Indistinguishability

• Intuition: Observer seeing only output "can't tell" if input was D or D'
  
• But if $r = 1.0$ and $r' = 0.999$, can tell!

• Example: average salary

• Going to need randomized algorithms & probabilistic indistinguishability
Randomized Response (1965)

Have you deliberately violated social distancing?

- Privacy: Plausible deniability
- Utility:
  \[ Pr[\text{"yes"}|\text{yes}] = \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} = \frac{3}{4} \]
  \[ Pr[\text{"yes"}|\text{no}] = 0 + \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4} \]
  \[ Pr[\text{"yes"}] = p\left(\frac{3}{4}\right) + (1-p)\left(\frac{1}{4}\right) \]
RR as an "algorithm"

output distribution

Idea: Output distribution should be almost the same if you are y or n.
Differential Privacy

Defn (Randomized) algo \( A \) is \( \epsilon \)-DP if for any neighboring \( D \) & \( D' \) if:

\[
D \xrightarrow{A} \text{output} \quad \epsilon\text{-close}
\]

\[
D' \xrightarrow{A} \text{output} \quad \epsilon\text{-close}
\]

Output space
Formally: For any subset $S \subseteq \Omega$:

$$\Pr[A(D) \in S] \leq e^\varepsilon \Pr[A(D' \in S]$$

$$\Pr[A(D) \in S] \geq \frac{1}{e^\varepsilon} \Pr[A(D') \in S]$$

- $\varepsilon \to 0$: perfect privacy
- $\varepsilon \to \infty$: no privacy
- $S$ as your "privacy fear"

Important: Does not promise $A(D), A(D')$ small, just that they are close

Property of algo, not data!

Contrast anonymization
DP analysis of RR

• Let \( \text{Oi} \in \{\text{"yes"}, \text{"no"}\} \) be output of participant \( i \)

• Note: for any inputs \( (y_1^n) \),

\[
\Pr[0,0,0,\ldots,0_n] = \Pr[0_1]\Pr[0_2]\ldots\Pr[0_n]
\]

(independence)

so we can just analyze \( y_1^n \).

• Need to relate \( \Pr[\text{"yes"}|\text{yes}] \) & \( \Pr[\text{"yes"}|\text{no}] \)

• But \( \frac{\Pr[\text{"yes"}|\text{yes}]}{\Pr[\text{"yes"}|\text{no}]} = \frac{3}{4} \)

\[
\frac{\Pr[\text{"yes"}|\text{yes}]}{\Pr[\text{"yes"}|\text{no}]} = \frac{3}{4}
\]

• \( e^c = 3, \ c = \ln(3) \approx 1.1 \)

= 3.
Then:

\[ P[r \mid y] = q + (1 - q)/2 \]
\[ = (1 + q)/2 \]
\[ P[r \mid n] = 0 + (1 - q)/2 \]
\[ = (1 - q)/2 \]

Ratio \[= \frac{1 + q}{1 - q} = e^\Delta \]

\[ \Delta = \ln \left( \frac{1 + q}{1 - q} \right) = 3 \]
General Tools for Differential Privacy: The Laplace Mechanism
The Laplace Mechanism

• Consider data \( x \in \{0, 1\}^n \)
• So i's data is number \( x_i \)
• Want to (DP)compute some function:
  \[ f(x) = f(x_1, x_2, \ldots, x_n) \in \mathbb{R} \]
  • E.g. average, median, std, max, \( x_1x_2^3 + 10x_3^2 + \ldots \)
• Want a general way of DP-computing \( f(x) \) accurately
The Laplace Mechanism

• Compute $f(x)$ exactly then “add noise”
  • What kind? How much?
  • How much? Define sensitivity of $f(x)$:

$$\Delta f = \max_{\text{neighboring } x, x' \in [0,1]^n} \left\{ \frac{1}{2} |f(x) - f(x')| \right\}$$
The Laplace Mechanism

Examples:

- $f = \text{average}, \quad \Delta f = \frac{1}{n} \cdot (111 \ldots 1 \rightarrow 111 \ldots 10)$
- $f = \text{std} = \sqrt{\frac{1}{n} \sum (x_i - \mu)^2}, \quad \Delta f = \frac{1}{\sqrt{n}}$
- $f(x) = \max(x_1, \ldots, x_n), \quad \Delta f = 1 \cdot (00 \ldots 0 \rightarrow 00 \ldots 01)$
- $f(x) = x_1 \cdot x_2 \cdots x_n, \quad \Delta f = 1$
- $f(x) = \text{median}(x_1, \ldots, x_n) = 1$
The Laplace Mechanism

• Laplace distribution:
  - Randomly choose value \( v \) with probability:
    \[
    \frac{1}{2b} e^{-|v|/b}
    \]
  - \( b \) is a parameter
  - Larger values of \( |v| \) are exponentially less probable
The Laplace Mechanism

- mean = 0, variance = 2b
- larger b: more noise
- smaller b: less noise
The Laplace Mechanism

• Finally:
  - Compute $f(x)$
  - Output $f(x) + v$

where $v$ is Laplace with $b = \Delta f / \epsilon$

Sensitivity of $f$ for $\epsilon$-DP
Theorem: Laplace mech. satisfies $\varepsilon$-DP.

Proof: Let $x, x' \in [0,1]^n$ be neighbors with LM distributions $P_x, P_{x'}$.

For any output value $o$:

$$P_x(o) = \frac{\left(\frac{1}{2b} e^{-|f(x)-o|/b}\right)}{\left(\frac{1}{2b} e^{-|f(x')-o|/b}\right)} = e^{\left(\frac{|f(x')-o| - |f(x)-o|}{b}\right)/b}$$

$$\leq |f(x') - f(x)|$$

$f(x)$ $\circ$ $f(x')$ $\circ$ $f(x)$ $\circ$ $f(x')$
\[ \frac{f(x') - f(x)}{b} \leq e \Delta f/b \]
\[ \leq e^{\Delta f/\varepsilon} = e^{\varepsilon}. \checkmark \]

\[ \text{If } S = \{o_1, o_2, \ldots, o_k\} \text{ is a set:} \]
\[ \frac{p_x(s)}{p_x(s')} = \frac{\sum_{i=1}^{k} p_x(o_i)}{\sum_{i=1}^{k} p_x(o_i)} \leq \frac{\sum_{i=1}^{k} e^{\varepsilon} p_x(o_i)}{\sum_{i=1}^{k} p_x(o_i)} \leq e^{\varepsilon}. \]

LM is also useful

- E.g. if $\Delta f = \frac{1}{\sqrt{n}}$ (average) then $b = \frac{\Delta f / \varepsilon}{\frac{1}{\sqrt{n}}} = \frac{1}{\sqrt{n}\varepsilon}$.
- $b \to 0$ as $n \to \infty$ for fixed $\varepsilon$.

- $\Delta f = \frac{1}{\sqrt{n}}$, $b = \frac{1}{\varepsilon \sqrt{n}}$

- General: any $f$ s.t. $\Delta f \to 0$ as $n \to \infty$
General Tools for DP II:
The Exponential Mechanism
Complex Inputs/Outputs

- Laplace: numbers → number
- What about:
  - Dataset → decision tree, neural net, ...
  - Social network → clustering
  - Votes → chosen/ideal outcome
  - Auction bids → winners & prices

Can't just "add noise" to output!
The Setting

- Input space $I$
- Output space $O$

- For $x \in I$, $o \in O$, notion of quality of output $o$: $U(x, o) \in \mathbb{R}$

Examples:
- $x$ = dataset, $o$ = neural net, $U = \text{error of } o \text{ on } x$
- $x$ = votes, $o$ = outcome, $U = \text{agreement of } o \text{ on } x$
- $x$ = bids, $o$ = winners, $U = \text{social welfare}$
The Setting

- Absent privacy goal is $D^* = \arg\max_{\delta \in \mathcal{D}} \mathbb{E} u(x, 0, \delta)$
  
i.e. pick best output.
- But this won’t be DP!
- Example: exfiltrating training data in ML.
- What can we do?
Generalized Sensitivity

- Let's define

$$
\Delta u = \max_{x,x' \in \Omega} \max_{O \in \Omega} \left| u(x,O) - u(x',O) \right|
$$

- How much can changing a single input change the quality of some output?

- E.g. ML: \( \Delta u = \frac{1}{n} \)
- E.g. average of \( x_1, \ldots, x_n \): \( \Delta u = \frac{1}{n} \)
- E.g. winning auction price: \( \Delta u \) large
The Exponential Mechanism

- On input $x$, output each $o \in O$ with probability:
  \[ p_x(o) = \frac{e^{\frac{E_u(x,o) / 2\Delta u}{\mathcal{Z}(x)}}}{\mathcal{Z}(x)} \]

  where $\mathcal{Z}(x) = \sum_{o \in O} e^{\frac{E_u(x,o) / 2\Delta u}{\mathcal{Z}(x)}}$

- Every $o \in O$ might be output, but more likely
Exp. Mechanism is $\varepsilon$-DP

Proof: $\forall$ nbrs $x, x' \in I, 0 \in O,$

\[
\begin{align*}
P_{x\mid o} &= e^{\varepsilon \cdot u(x, o)/2\Delta u}/z(x) \\
P_{x'\mid o} &= e^{\varepsilon \cdot u(x', o)/2\Delta u}/z(x') \\
&= e^{\varepsilon(u(x', o) - u(x, o))/2\Delta u} \\
&\leq e^{\varepsilon\Delta u/2\Delta u} \\
&= e^{\varepsilon/2}.
\end{align*}
\]
Now: \[
\frac{Z(x)}{Z(x')} = \sum_0 \frac{e^{\epsilon \eta(x,0)/2\Delta u}}{\sum_0 e^{\epsilon \eta(x,0)/2\Delta u}} \leq \sum_0 e^{\epsilon \eta(x,0)/2\Delta u} \leq \sum_0 e^{\epsilon \frac{\partial u}{2\Delta u}} \leq \sum_0 e^{\epsilon \eta(x,0)/2\Delta u} \]

\[
e^{\epsilon \Delta u/2\Delta u} \leq e^{\epsilon \eta(x,0)/2\Delta u} \]

\[
e^{\epsilon \Delta u/2\Delta u} \leq e^{\epsilon \eta(x,0)/2\Delta u} \]

\[
= e^{\epsilon \Delta u/2\Delta u}, \quad \text{and} \quad e^{\epsilon \Delta u/2\Delta u} e^{\epsilon \Delta u/2} = e^{\epsilon \Delta u}.
\]
Utility of \( \text{Exp. Mech.} \)

- Special case: \( O \) finite

- Fix \( x \), let \( EM(x) \) denote \( \text{Exp. Mech.} \)

- Then with high probability:

\[
U(x, O^*) - U(x, EM(x)) \leq \frac{2d}{\ln |O|} \]

- E.g. ML: \( \Delta u = \frac{1}{n} \), \( |O| = 2^d \),

set \( \frac{2d}{\ln n} \ll 1 \Rightarrow n \gg \frac{2d}{\e} \).
So I, O and u(x,0) can be anything and we can be DP!

What's the catch?
DP Immunity to Post-Processing
Then $B(A(x))$ also $\varepsilon$-DP.

**Proof:**

For all $x, x', \forall o' \in O'$, let $T = \{ O \subseteq O : B(o) = o' \}$

Then:

$$\Pr[B(A(x)) = o'] = \Pr[A(x) \in T]$$

$$\leq e^{\varepsilon} \Pr[A(x')] \in T]$$

$$= e^{\varepsilon} \Pr[B(A(x')) = o']$$

\[\checkmark\]
Special case:

O' = "input was \( x \)"

Then for small \( \epsilon \), have:

\[
\Pr[\text{BLA}(x)] = "x" \] (right)
\[
\Pr[\text{BLA}(x')] = "x" \] (wrong)

\[
= 1 - \Pr[\text{BLA}(x')] = "x'"
\]

\[
\leq \Pr[\text{BLA}(x)] + \Pr[\text{BLA}(x')] = "x'"
\] \approx 1.

"Not (much) better than random guessing"
Composition
Properties of DP
Parallel DP Algos

• Suppose $A_1$ is $\epsilon_1$-DP and $A_2$ is $\epsilon_2$-DP

• $A(x) = (A_1(x), A_2(x))$

• Then $A$ is $(\epsilon_1 + \epsilon_2)$-DP

• Proof: $P_x^{A}(0_1, 0_2)$

  $= P_x^{A_1}(0_1) P_x^{A_2}(0_2)$ \textbf{indep.}

  $\leq e^{\epsilon_1} P_x^{A_1}(0_1) e^{\epsilon_2} P_x^{A_2}(0_2)$

  $= e^{\epsilon_1 + \epsilon_2} P_x^{A}(0_1, 0_2)$. \checkmark
Suppose your data is in many datasets $x_1, x_2, x_3, \ldots$
"Theorem." If there are \( k \) datasets & algos, each \( \varepsilon\)-DP, then composition is \( k \cdot \varepsilon\)-DP.

Has led to development of an algo toolkit for DP.
Applied Case Study: DP Synthetic Data Generation
The Goal

\[ D \xrightarrow{\varepsilon} A \xrightarrow{\varepsilon} D' \]

- \( D' \) is a dataset that is DP but still "looks like" \( D \)
- Like anonymization but better
"Looks like D"

- \( D' \) approx. preserves statistical props of \( D' \)
- Conditional queries:
  - e.g. fraction of rows s.t. \( \text{age} \geq 25 \& \text{gender} = F \& \text{job} = \text{VP} \)
  - e.g. avg. salary of same
- Generally low sensitivity \( (\sim \frac{1}{n}) \)
- If query \( q \), want \( |q(D) - q(D')| \) small
How?

- Generally hand
- Could use Exp. Mechanism, but...

Instead:

- Add Laplace noise to the $g(D) \rightarrow g'(D)$
- Treat entries of $D'$ as variables
- Use gradient descent to reduce $\max_{D'} |g'(D) - g(D')|$
Algorithm 2 Relaxed Adaptive Projection (RAP)

Input: A dataset $D$, a collection of $m$ statistical queries $Q$, a ‘‘queries per round’’ parameter $K \leq m$, a ‘‘number of iterations’’ parameter $T \leq m/K$, a synthetic dataset size $n'$, and differential privacy parameters $\epsilon, \delta$.

Let $\rho$ be such that:

$$\epsilon = \rho + 2\sqrt{\rho \log(1/\delta)}$$

if $T = 1$ then
  for $i = 1$ to $m$ do
    Let $\hat{a}_i = G(D, q_i, \rho/m)$.
  end for
  Randomly initialize $D' \in (\mathcal{X}^r)^{n'}$.
  Output $D' = RP(q, \hat{a}, D')$.
else
  Let $Q_S = \emptyset$ and $D'_0 \in (\mathcal{X}^r)^{n'}$ be an arbitrary initialization.
  for $t = 1$ to $T$ do
    for $k = 1$ to $K$ do
      Define $\hat{q}^{Q \setminus Q_S}(x) = (\hat{q}_i(x) : q_i \in Q \setminus Q_S)$ where $\hat{q}_i$ is an equivalent extended differentiable query for $q_i$.
      Let $q_i = RNM(D, q^{Q \setminus Q_S}, q^{Q \setminus Q_S}(D'_t-1), \frac{\rho}{2T \cdot K})$.
      Let $Q_S = Q_S \cup \{q_i\}$.
      Let $\hat{a}_i = G(D, q_i, \frac{\rho}{2T \cdot K})$.
    end for
    Define $q^{Q_S}(x) = (q_i(x) : q_i \in Q_S)$ and $\hat{a} = \{\hat{a}_i : q_i \in Q_S\}$ where $\hat{q}_i$ is an equivalent extended differentiable query for $q_i$. Let $D'_t = RP(q^{Q_S}, \hat{a}, D'_t-1)$.
  end for
  Output $D'_T$.
end if

import jax.numpy as np
def threeway_marginals(D):
  return np.einsum('ij, ik, il->jkl', D, D, D)/D.shape[0]

Figure 1: Python function used to compute 3-way product queries
<table>
<thead>
<tr>
<th>Dataset</th>
<th>Records</th>
<th>Features</th>
<th>Transformed Binary Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADULT</td>
<td>48842</td>
<td>15</td>
<td>588</td>
</tr>
<tr>
<td>LOANS</td>
<td>42535</td>
<td>48</td>
<td>4427</td>
</tr>
</tbody>
</table>

Table 1: Datasets. Each dataset starts with the given number of original (categorical and real valued) features. After our transformation, it is encoded as a dataset with a larger number of binary features.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K$</td>
<td>Queries per round</td>
<td>5 10 25 50 100</td>
</tr>
<tr>
<td>$T$</td>
<td>Number of iterations</td>
<td>2 5 10 25 50</td>
</tr>
</tbody>
</table>

Table 2: RAP hyperparameters tested in our experiments
Figure 2: Max-error for 3 and 5-way marginal queries on different privacy levels. The number of marginals is fixed at 64.
Figure 3: Max error for increasing number of 3 and 5-way marginal queries with $\epsilon = 0.1$