Observable…Yet Still Private? An Offensive Privacy Perspective on Observability

PRESENTER

Cat Easdon
Dynatrace Research

PRESENTER

Patrick Berchtold
Dynatrace & ISEC, TU Graz

Intro to Observability

The Three Pillars

Logs

Timestamped structured or unstructured text, e.g. error logs

Metrics

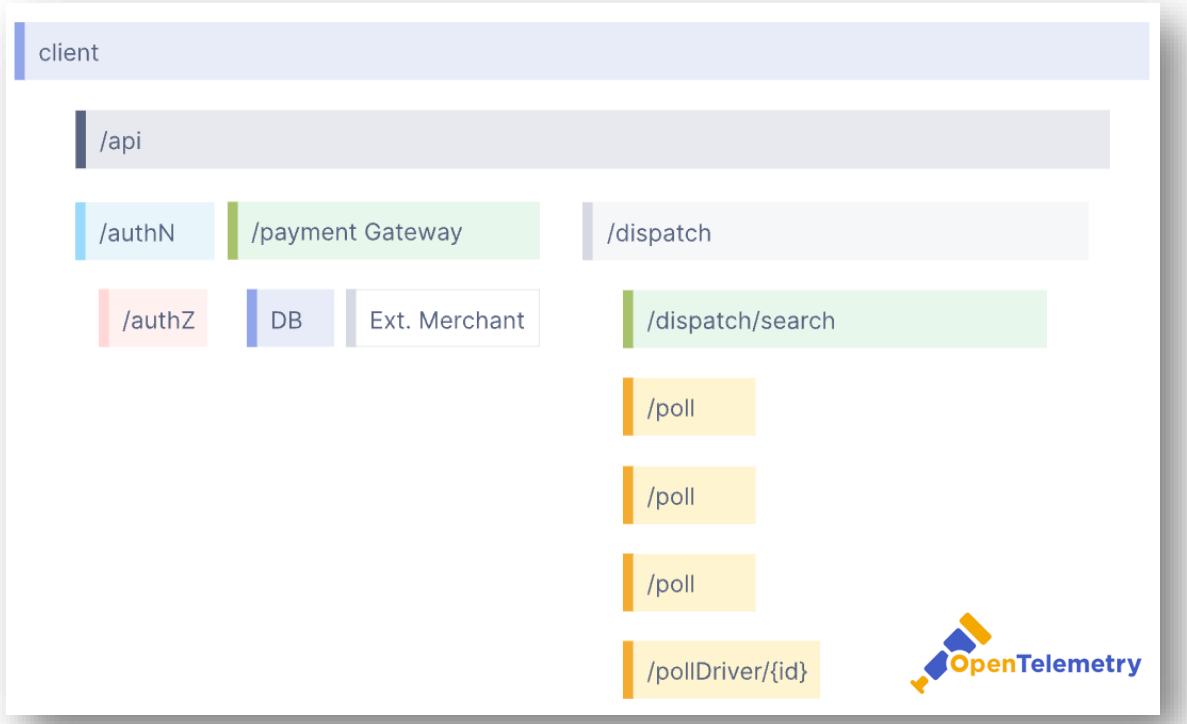
Numeric data, typically tracking performance and resource usage over time

Traces

Trace the request flow end-to-end (through frontend, backend microservices, DB...)

Frontend Observability

- Client-side instrumentation links client-side events to server-side events for tracing
- Real user monitoring (RUM) adds detail about the user's experience and interactions
  - Useful for troubleshooting and business decision-making
- Offered by many vendors; experimental support in OpenTelemetry



Observability and Privacy

- Privacy features and developer education are essential
  - Sensitive data masking at each ingest stage, plus sensitive data detection as a last line of defense
  - Privacy-preserving API design

- Even if you don't use observability tooling, consider that someone else in the request chain (ISP, CDN, …) probably does

Reconstruction Attacks

Linear Reconstruction

Fundamental Law of Information Recovery: *Overly accurate answers to too many questions will destroy privacy in a spectacular way (Dwork & Roth, 2014)*

Provided enough statistics, the underlying dataset can be fully recovered.

- **Linear reconstruction:** solve a linear program formed by the statistics
- Provides exact solution, but time- and memory-intensive for high dimensional datasets

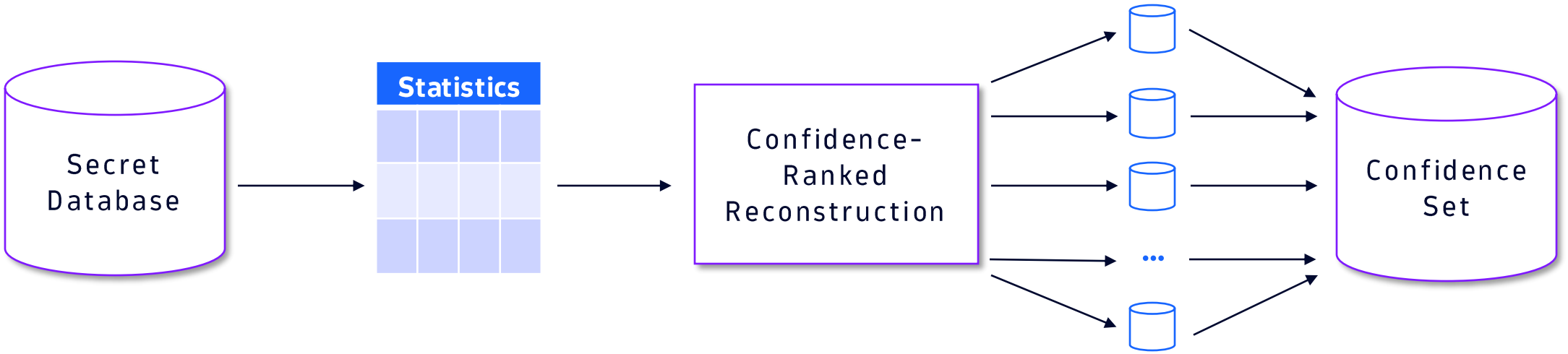
Age	Count	Median	Average
Total	3	30	44
Men	2	28	28

Reconstruction

	Gender	Age
A	Male	26
B	Male	30
C	Female	76

Confidence-Ranked Reconstruction

- Exploits privacy leakage in synthetic data generation
  - Create synthetic datasets from the same distribution as the original dataset
  - Convert to a non-convex optimization problem and use the Relaxed Adaptive Projection optimization heuristic
  - Output a list (ordered by confidence) of candidate rows in the dataset
- Only partial reconstruction possible



Dick, Travis, et al. "Confidence-ranked reconstruction of census microdata from published statistics." Proceedings of the National Academy of Sciences 120.8 (2023): e2218605120.

Reconstruction Attack Settings

Public Dataset

Offline, non-interactive,
e.g. census statistics

Malicious Provider

Online, interactive

Untrusted system

Malicious User

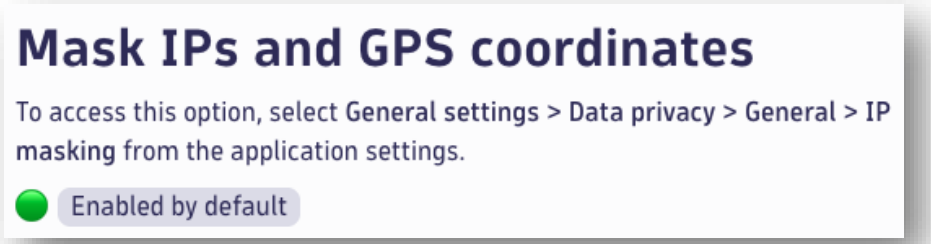
Online, interactive

Trusted system

Our Scenario

Context: B2B

- B2B SaaS observability provider
- Customer confidentiality is essential, especially in highly-regulated sectors
- Protect privacy via ‘paved roads’
  - Privacy nudges in UI and docs
  - Privacy features enabled by default where feasible
- Privacy threat modeling and offensive privacy research inform privacy feature investments



Context: Proposed Privacy Feature

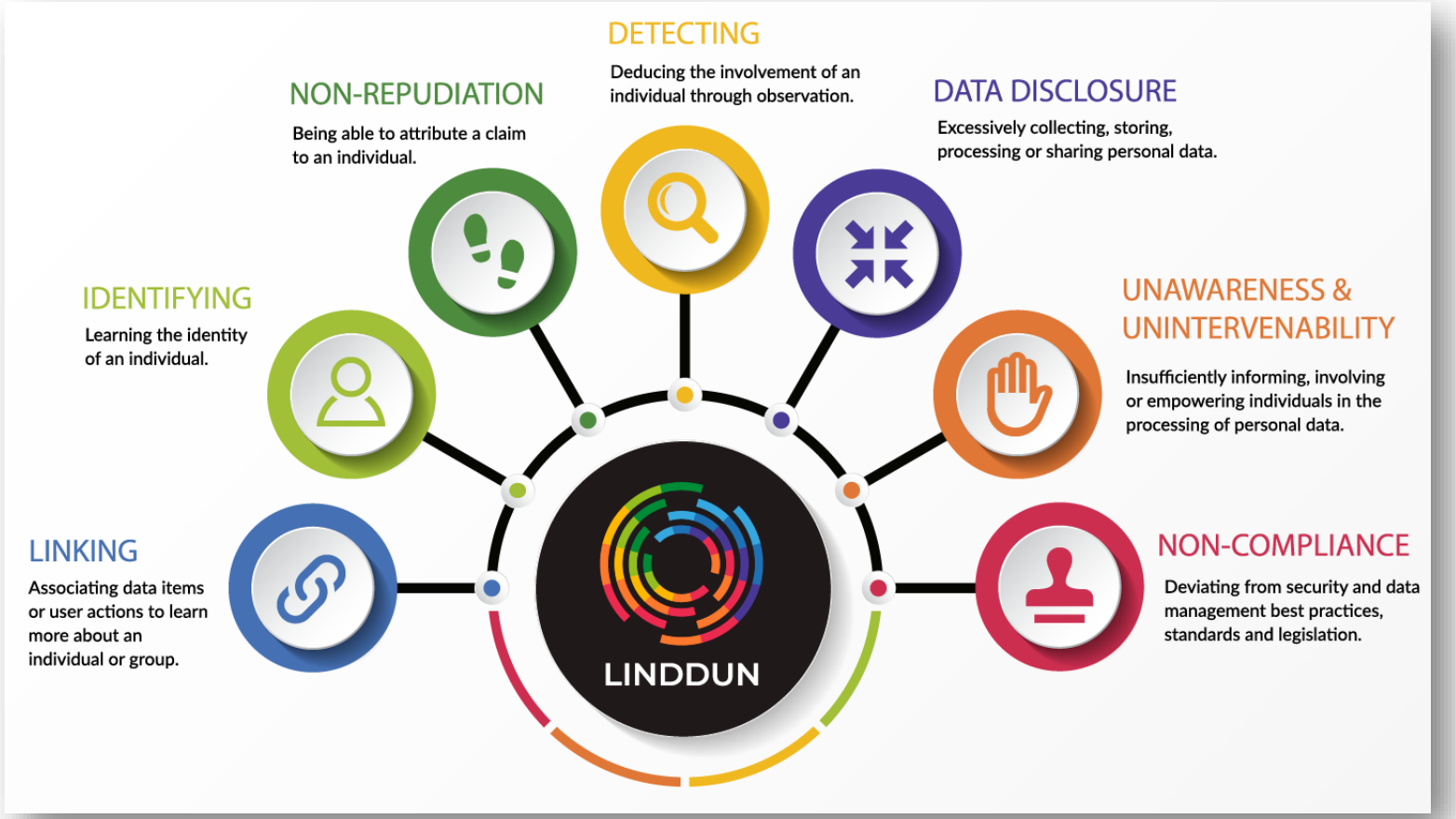
- Currently only two access levels for RUM data per application: view all, or none
- Idea: let's add a new feature, a **statistics-only** API
- But estimated effort is high
  - Especially with differential privacy
  - Could we consider noise-free statistics?
  - To find out and strengthen the business case, let's threat model and attack 😈

This will help protect users' privacy! Let's build it!

Are you sure it'll provide meaningful privacy guarantees?

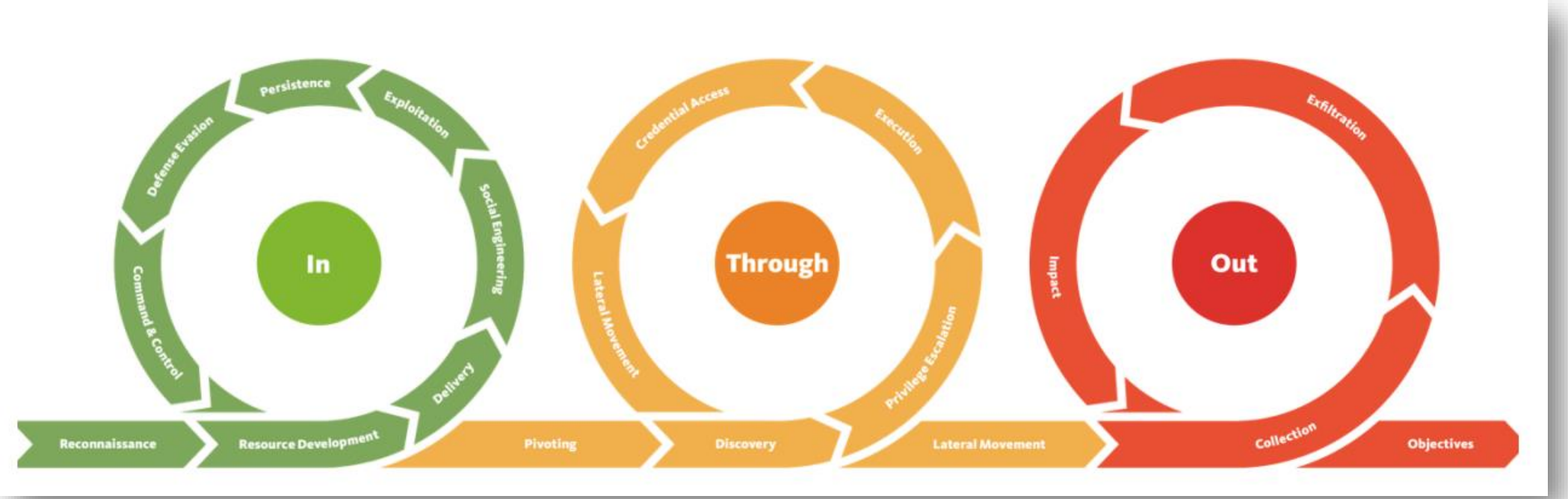
Threat Modeling RUM Statistics

LINDDUN



Every single one of these might or might not be a threat to monitored users, depending on how our customers use RUM. We need a more tactical perspective.

The Unified Kill Chain (Security)



A Privacy Kill Chain for Reconstruction Attacks

Data Collection

How will they get access to RUM statistics?

Data Reconstruction

How will they try to recover the underlying dataset, and how successful will they be?

Data Exploitation

What are their objectives, and can they meet them with the recovered data?

The Actors (With Statistics-Only Feature)

Developer

Needs (audited) access to RUM data for troubleshooting

Business Analyst

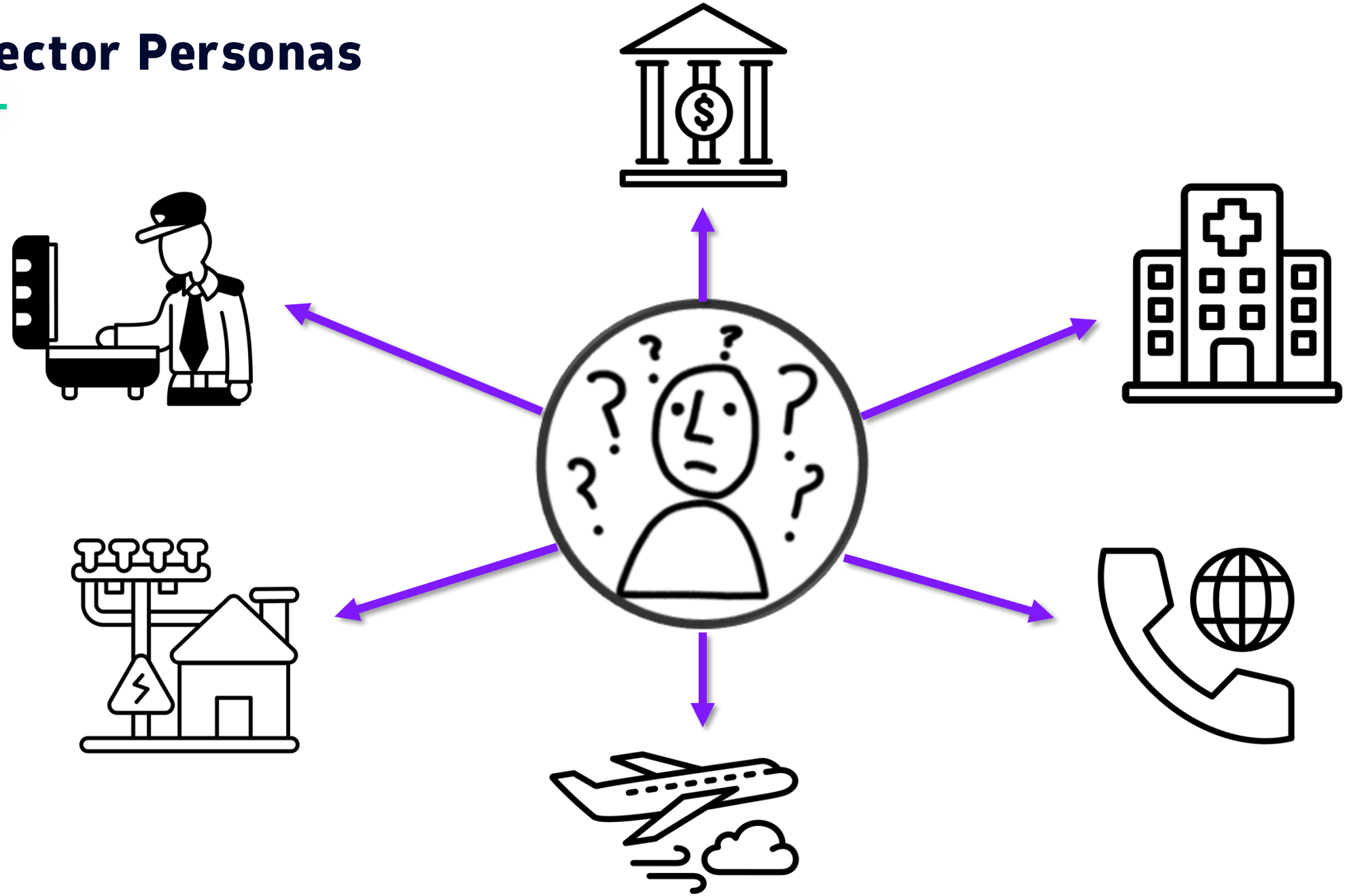
Only has (audited) access to statistics to identify trends

Insider Threat

Also has (audited) access to statistics

Not a developer

Per-Sector Personas



Experimental Setup and Results

Experimental Setup and Challenges

Attacker setup

- Low computational budget: Intel® Core™ i7-10875H (up to 5.1 GHz), 32GB RAM
- Confidence-ranked attack code in Python, adapted from Relaxed Adaptive Projection paper with statistical noise removed

Dataset

- Subset of real data model; demo data generated with open-source application EasyTravel
- 5000 rows (user sessions), dimensionality: 1.71×10^{17}

Feature	# Unique Values
browserMajorVersion	67
country	101
region	335
city	549
displayResolution	28
screenHeight	23
screenWidth	25
osVersion	17
isp	503

Aydore et al. "Differentially Private Query Release Through Adaptive Projection." ICML (2021).

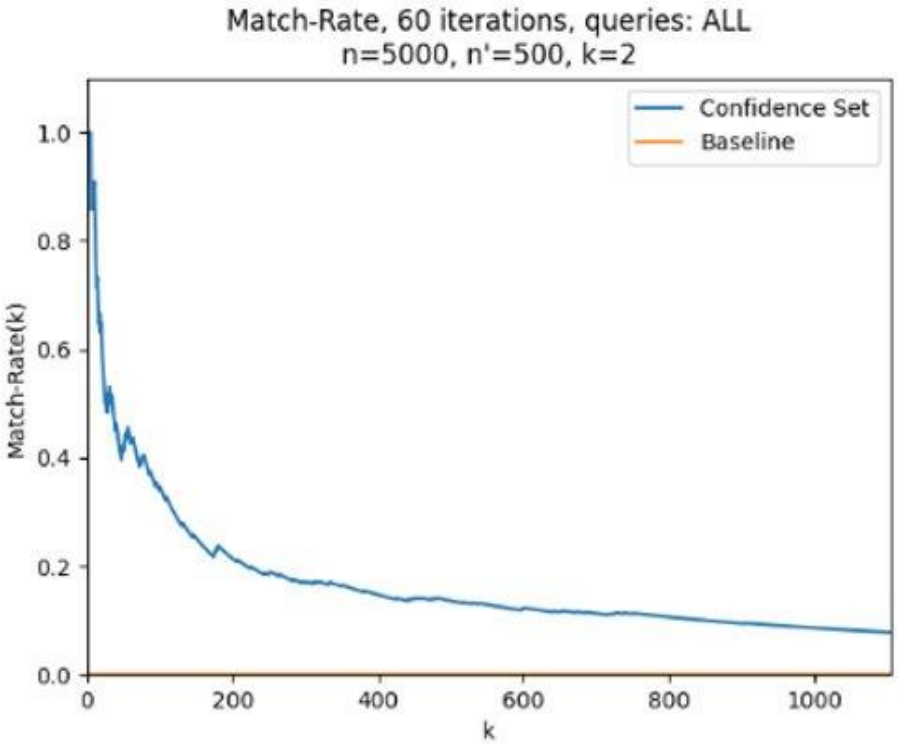
Results: Linear Reconstruction

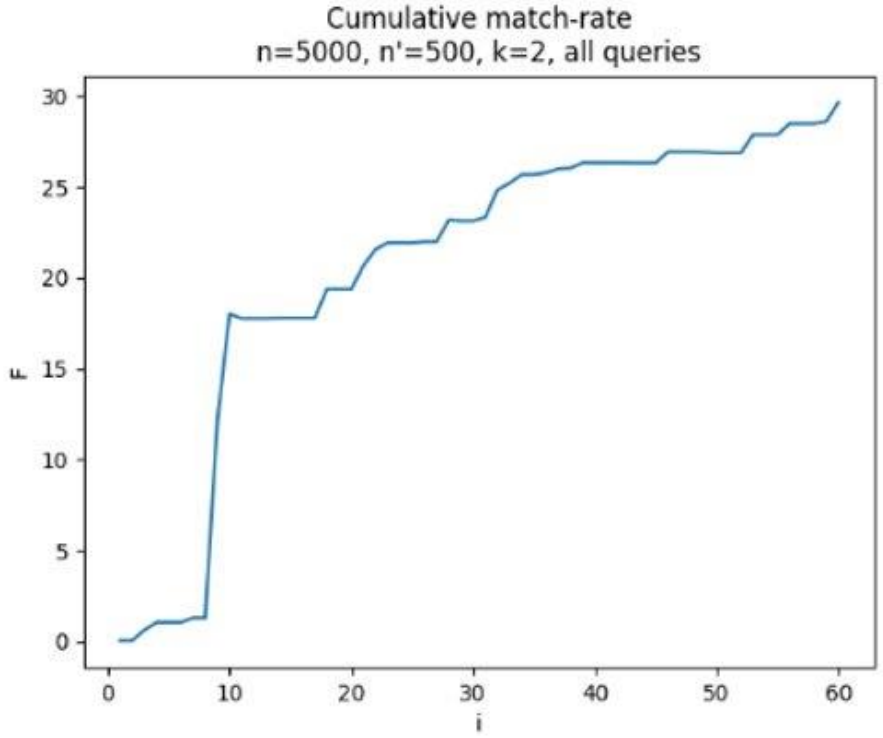
- Attack code: in R using lp_solve library
- Assumption: constant sparsity
- Effective for small datasets
- Infeasible for our dataset even with significantly more computational capacity
  - High memory usage
  - Runtime increases exponentially

Rows	Dim.	# Statistics	Runtime (s)
20	420	160	0.187
40	840	320	0.835
80	1680	640	4.673
160	3360	1280	41.5
320	6720	2560	635
640	13440	5120	53873

Results: Confidence-Ranked Reconstruction

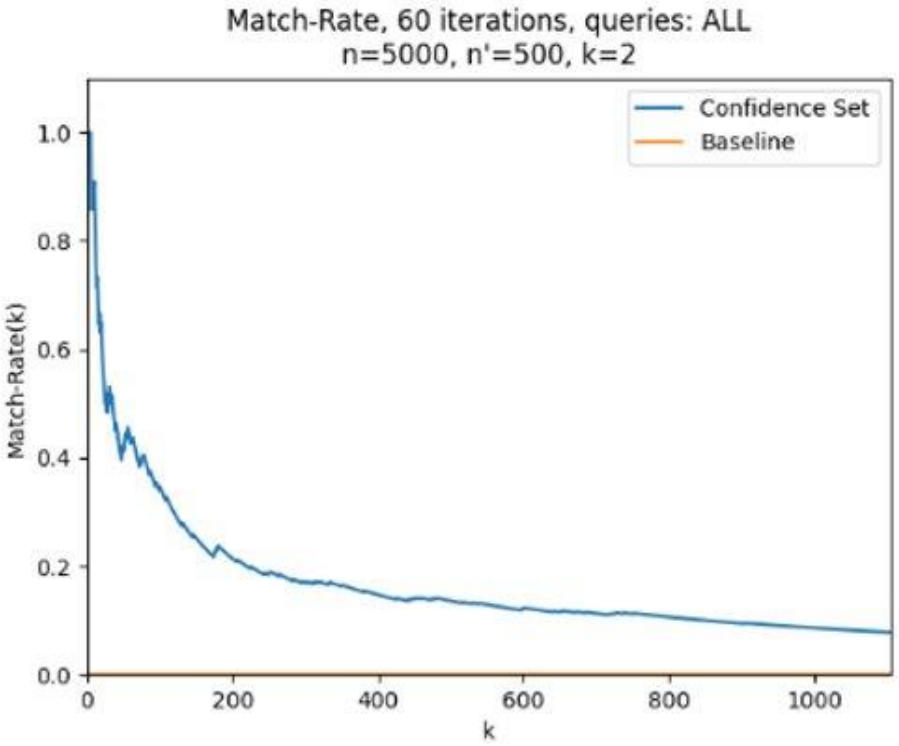
- No direct comparison possible with linear reconstruction (different setting & output)
- Instead analyze **(cumulative) match-rate**
- 72hrs computation for <2% reconstructed with high confidence (all 2-way queries)

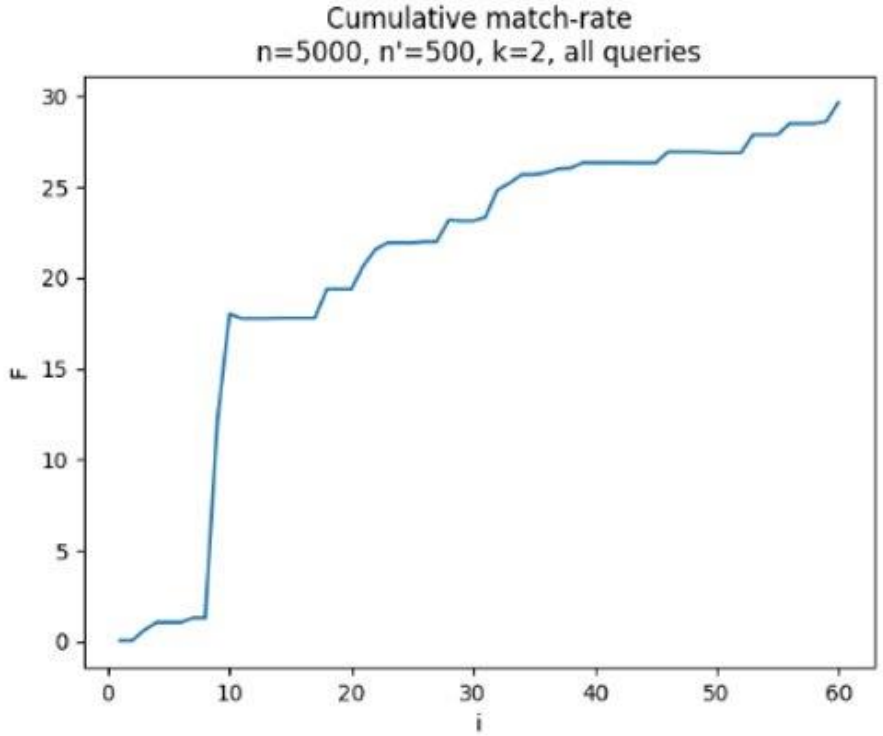




Results: Confidence-Ranked Reconstruction

- Time-consuming but partial reconstruction (<2%) is feasible
- Targeting a specific individual is hard due to the low proportion of data recovered
- Additional challenge in practice: rolling dataset that grows during attack





Mitigations

Mitigating The Reconstruction Attack Kill Chain

Data Collection

Mitigate with API rate limiting and anomalous behavior detection

Data Reconstruction

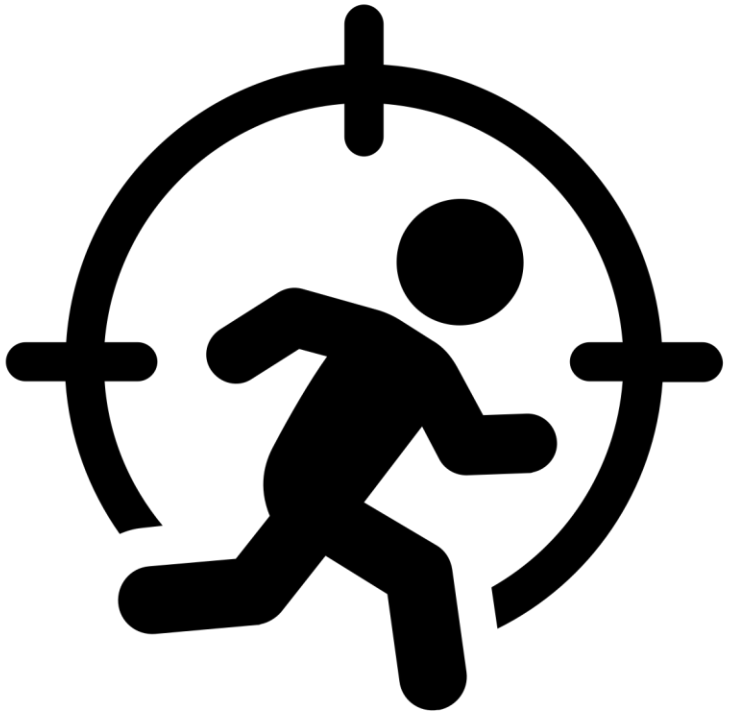
Mitigate with differential privacy

Data Exploitation

Mitigate with data minimization and segregation

Mitigations for Data Collection

- Feasible for the malicious user attack setting
- Confidence-ranked reconstruction requires 1 million API queries!
  - Data collection takes 2 weeks (rate limiting)
  - Anomalous request traffic easily detectable
  - Audit logging enables attacker identification
- Bonus 1: no assumptions necessary about attacker's computational resources
- Bonus 2: low implementation effort in our context



Conclusion

- Linear reconstruction: computationally infeasible for datasets of realistic size
- Confidence-ranked reconstruction: easily detectable during data collection
- -> Meaningful privacy guarantees without differential privacy
- Offensive privacy analysis can help support the business case for privacy features and prove privacy claims to customers

We tested that the privacy guarantees are robust!

Thanks, this helps me decide which features we should invest in.

Thanks for listening!
Any questions?

Copyright Notice

- **Dynatrace content and branding:** © 2025 Dynatrace LLC
- **Third-party images, text, and videos:** see links for attribution
- **Unattributed images:** used under license from the Noun Project
- **All other content:** original work by the authors, may be reused with attribution

CLOUD DONE RIGHT