galois

JANA - PRACTICAL PDAAS

"BENE VIXIT, BENE QVI LATVIT." - OVID

galois

University of BRISTOL

RUTGERS

GEORGE MASON UNIVERSITY

Rebecca Wright
Differential Privacy
Applied Cryptography

David Cash
Public Key
Cryptography

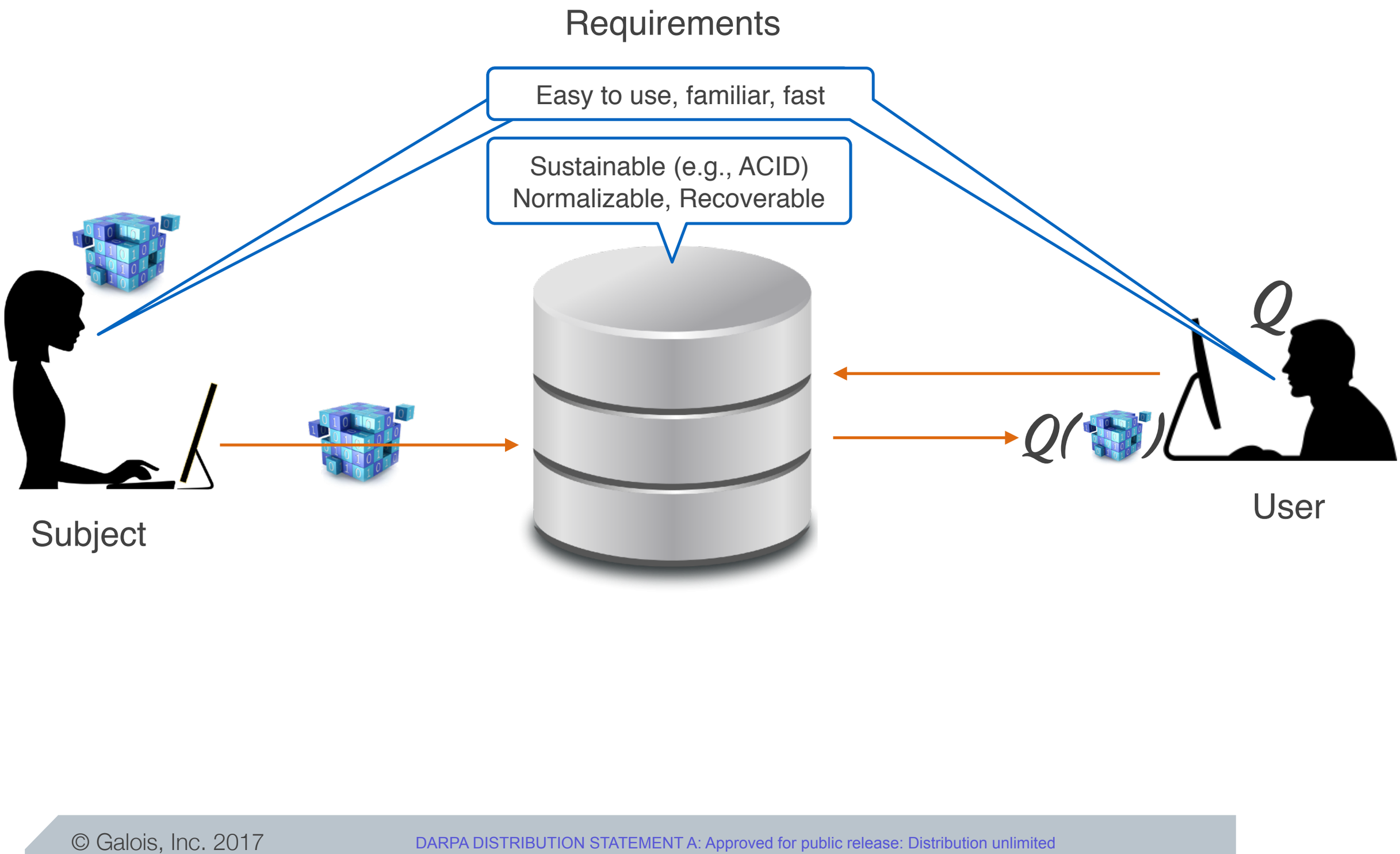
Dov Gordon
Scalable Secure
Computation

Anand Sarwate
Differential Privacy
Machine Learning

Dave Archer
Data-intensive Systems
Secure Computation

Nigel Smart
Cryptography
Secure Computation

Data as a Service



Second Thought: Xenosecurity

…network traces
…personnel data
…

Integrity, Confidentiality, Availability, Non-repudiation
Curious, Covert, Malicious
Collusion

Second Thought: Xenosecurity

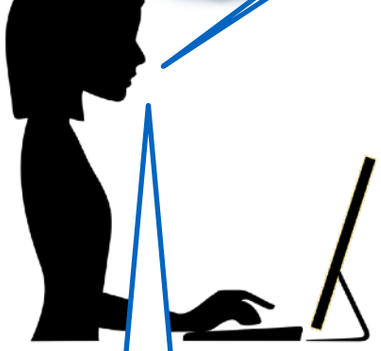
Requirements

Easy to Verify Security

Still fast

Easy to Configure Security

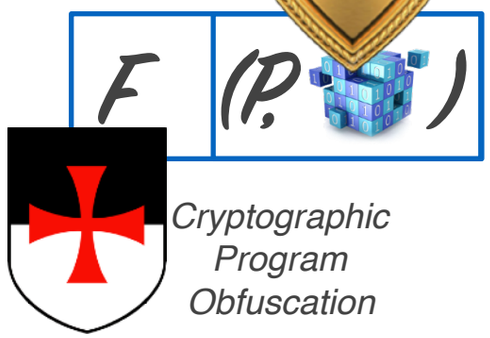
…network traces
…personnel data
…



CommSec
TLS

Ciphers
AES

Secure
Comp.
SMC,…



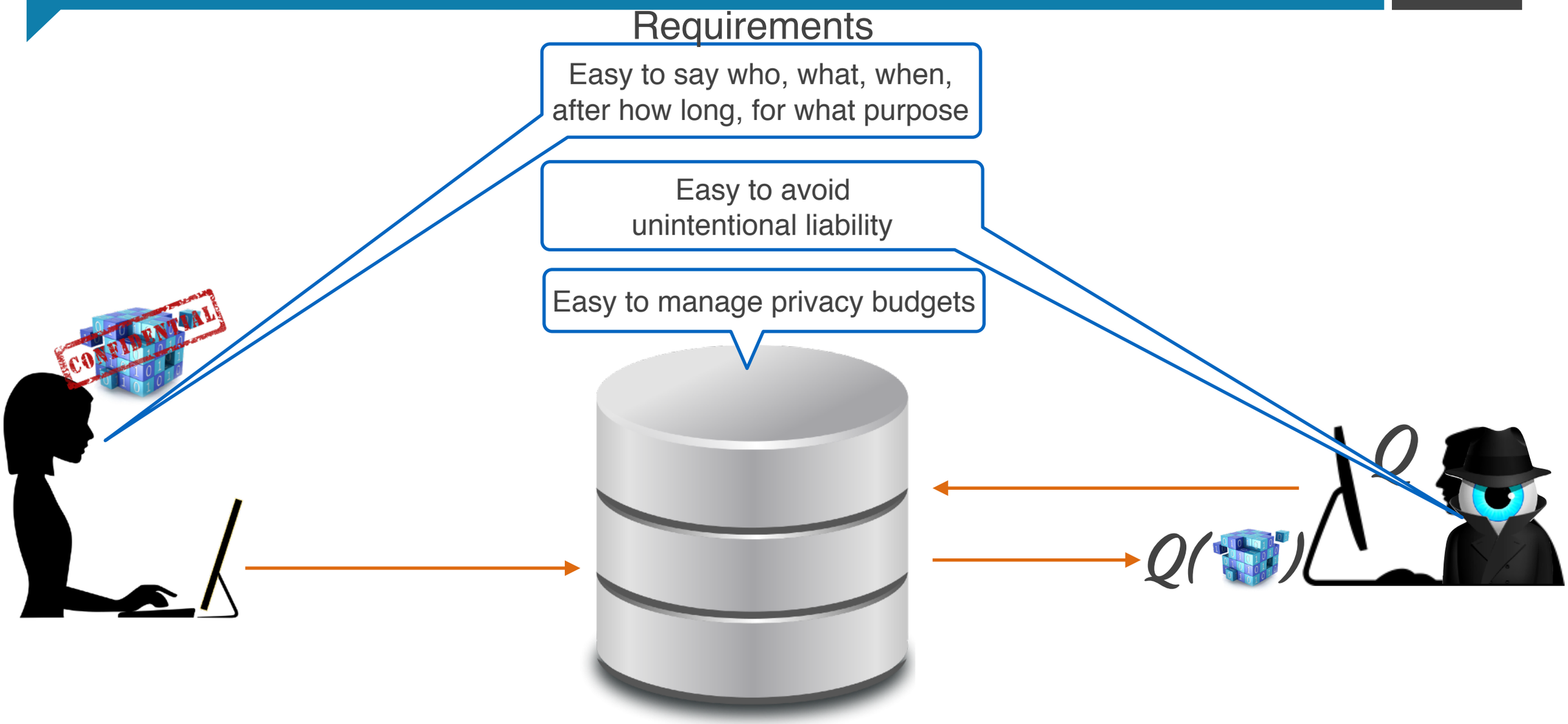
Cryptographic Program Obfuscation

Blinded from every vantage

Integrity, Confidentiality, Availability, Non-repudiation
Curious, Covert, Malicious
Collusion

P.S. You don't really believe it's ok to ever have data in the clear, do you?

Third Thought: “GeitoSecurity”



Privacy Policy Options

Policy approach	Impact on Privacy	Impact on Utility	Challenges
Collection policy	Poor	Nearly ideal	Policy-ese, Dancing pigs
Use policy	Good	Unacceptable	Liability, human negotiation
Storage policy Anonymization, pseudonymization, hashing, blinding	Poor	Poor	Re-identification, limited query capability
Human-in-loop sharing policy	Good	Good	Qualitative, capacity limited

Idea: Jana - PDaaS Research Vehicle

Requirements	Jana Principles
Easy to use, familiar, fast	Relational, allow for trade-offs
Sustainable (e.g., ACID) Normalizable, Recoverable	Based on COTS RDBMS
Easy to Configure Security	Characterize multiple security options
Easy to Verify Security	Remote attestation
Easy to state permissions	Controls by user, time, maturity, accuracy (DP), or function (FE) Deny, allow, aggregate only, DP only
Easy to avoid unintentional liability	Only query results seen by users
Easy privacy budgeting	Differential privacy budgeting included

Project Goal: study trade-offs
privacy vs. performance
policy complexity vs. privacy vs. utility

Jana Primer - Private Data as a Service

End to end+

Not pre-processed functions

Familiar, expressive: SQL + RDBMS

Easy to use: standard web service

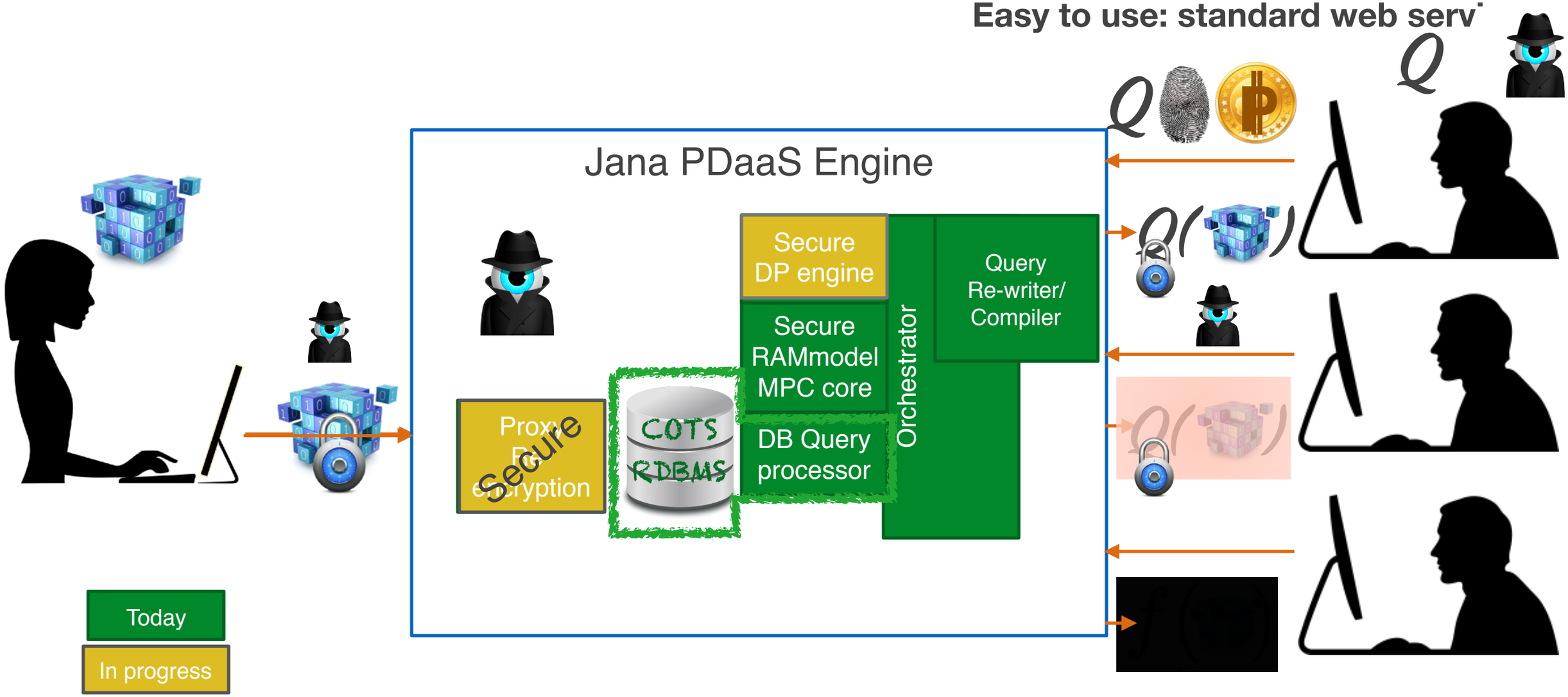
Jana Primer - Private Data as a Service

End to end+

Not pre-processed functions

Familiar, expressive: SQL + RDBMS

Easy to use: standard web serv˙



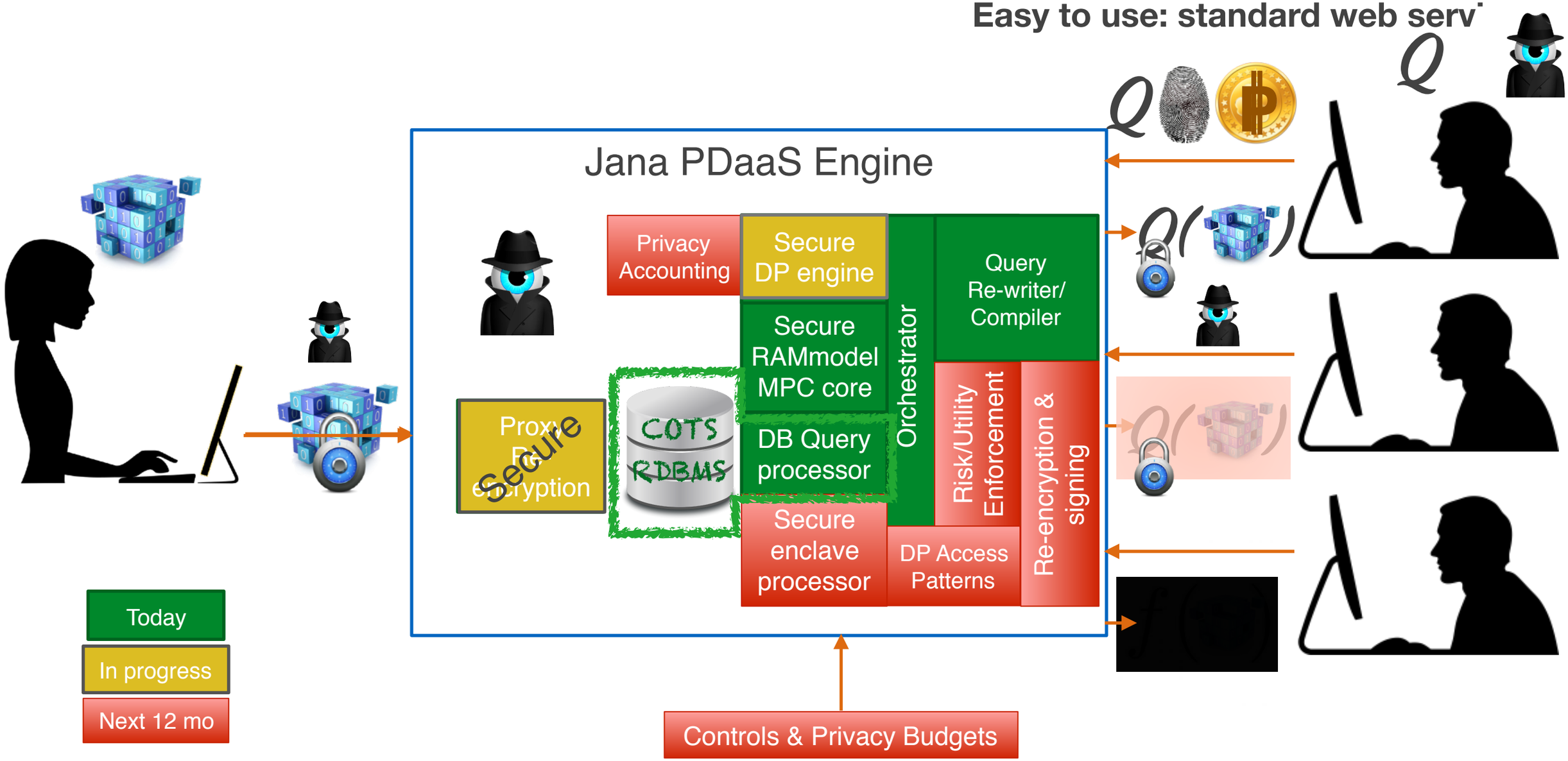
Jana Primer - Private Data as a Service

End to end+

Not pre-processed functions

Familiar, expressive: SQL + RDBMS

Easy to use: standard web serv˙



Jana Capabilities

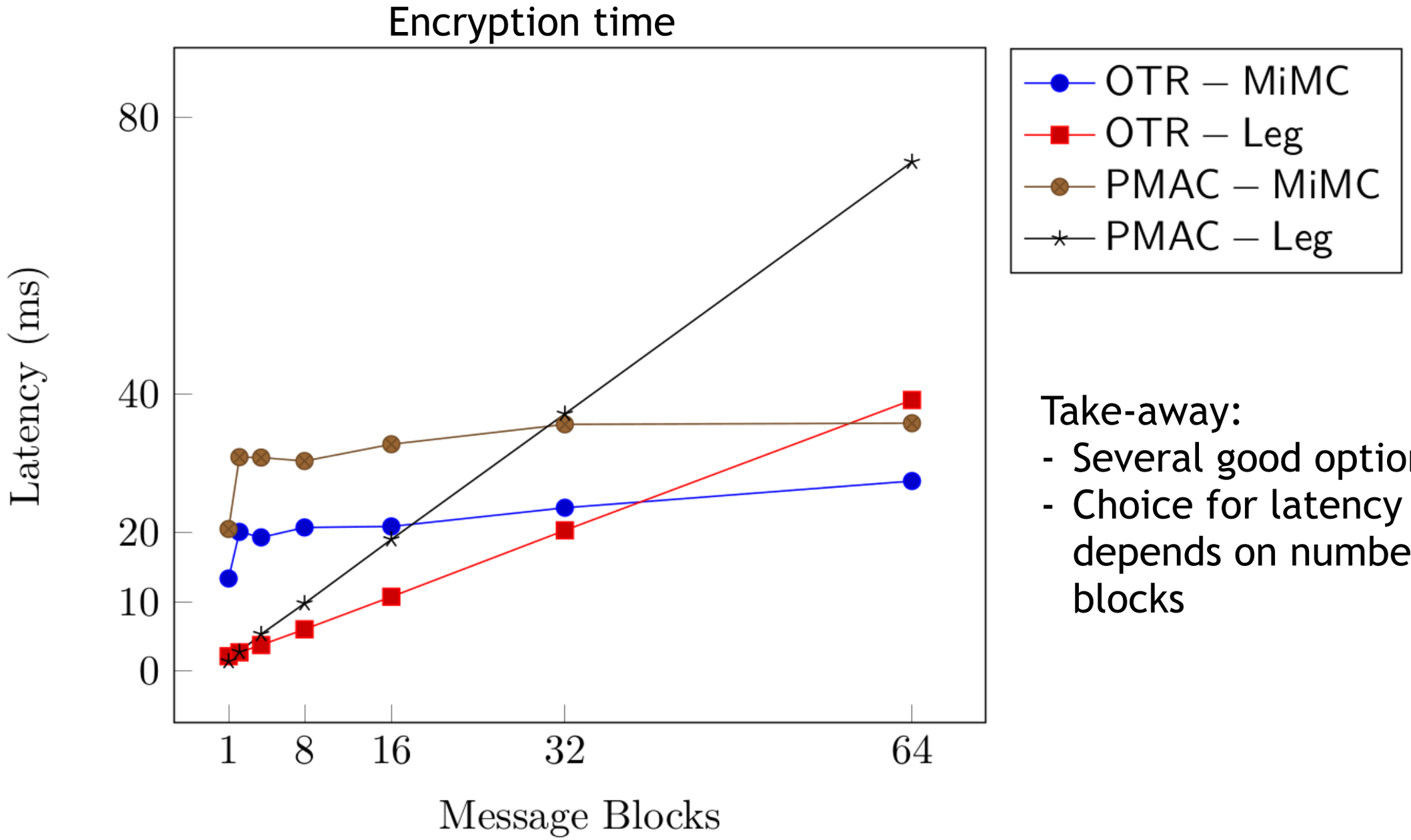
- Functionality
  - Generous subset of SQL
  - RDBMS ACID properties
- Privacy
  - DIT - Public key crypto + proxy re-encryption
  - DAR - Deterministic, random, searchable
  - Computation - in RDBMS using DET & searchable, in SPDZ MPC, soon in SGX
  - Results - Differential privacy applied (if needed) while in MPC
- Performance
  - 10Ks of records moving to 100Ks, queries in seconds, but YMMV!
- Deployment
  - Web service with RESTful API
  - Docker or similar appliance soon

By Generous Subset, We Mean…

- SPJ, UNION, INTERSECT, EXCEPT
- Integer, String, Boolean, Enum, Fixed-Point, Date
- Nested query support

```
SELECT person.person_id, lastname, firstname, diseasestate, gender, birthdate
FROM person
  JOIN community ON community.community_id = person.residence
  JOIN person2diseasestate ON person2diseasestate.person_id = person.person_id
  JOIN policyauthority2community ON policyauthority2community.community_id = community.community_id
  JOIN policyauthority ON policyauthority.authority_id = policyauthority2community.authority_id
WHERE person2diseasestate.transitiondate < '04-20-2017'
  AND person2diseasestate.diseasestate IN ('I')
  AND policyauthority.authority = 'CebuCityCommunityPA'
  AND person.person_id NOT IN
    (SELECT person.person_id
      FROM person
        JOIN community ON community.community_id = person.residence
        JOIN person2diseasestate ON person2diseasestate.person_id = person.person_id
        JOIN policyauthority2community ON policyauthority2community.community_id = community.community_id
        JOIN policyauthority ON policyauthority.authority_id = policyauthority2community.authority_id
      WHERE person2diseasestate.transitiondate < '04-20-2017'
        AND person2diseasestate.diseasestate IN ('R', 'D')
        AND policyauthority.authority = 'CebuCityCommunityPA');
```

Results: MPC-Friendly PRFs and Modes



Take-away:

- Several good options
- Choice for latency depends on number of blocks

Results: Searchable Encryption, Characterized

- **Result**: Privacy vs. Performance for 10 distinct range query approaches