Now We Need To Encrypt Data WHILE We Use It? It will be more fun than you think.

Alexander C. Mueller PhD
CEO and Founder of Capnion

February 5, 2018

Who is the speaker?

30 second resume:

- an ancient metro Saint Louis townie
- grew up in University City
- University City High School
- B.A. Washington University (econ and math)
- Ph.D. University of Michigan (math)
- private sector data science
- founded data privacy company Capnion

The Counter-Earth

Orbiting exactly opposite the Earth you are familiar with is another Earth, alike in many ways but also different...

Agenda

Tim Cook's Data Industrial Complex

- Private information is bought and sold
- Iron Triangles
- Bureaucracies / Legislatures / Interest Groups
- Tim Cook's suggestions

Transformative Technology

- Homomorphic encryption and zero-knowledge proof
- Data ownership and transformed business models
- Tim Cook's suggestions revisited
- What's out there? Projects and Prophecy

Tim Cook's Data Industrial Complex

" the trade in digital data has exploded into a "data industrial complex" "

https://techcrunch.com/2018/10/24/apples-tim-cook-makes-blistering-attack-on-the-data-industrial-complex/

Data Bought and Sold, Lost All At Once

You were probably in the Exactis breach whether you've heard of the company or not.

- Exactis: a Florida based marketing and data aggregation company
- Reported a breach involving 218 million individuals and 110 million U.S. households this past June
- No financials but very personal information like smoker status, dog or cat person, religion

You gave someone your data and they sold it to Exactis without notifying you.

https://www.marketwatch.com/story/a-new-data-breach-may-have-exposed-personal-information-of-almost-every-american-adult-2018-06-27

POLSCI100: The Iron Triangle

Are there any iron triangles with negative effects on data privacy and cybersecurity?

Bureaucracy: NSA Backdoors

Eternal Blue is a Windows vulnerability and NSA exploit.

- Leaked to the public in April 2017 (oops!)
- The NSA knew of the vulnerability for years
- Patched in March 2018 after involvement in WannaCry
- Perennially popular as a route into unpatched systems
- (Some claims above are of course not acknowledged)

But surely the government would not actively, willfully introduce backdoors into widely used software?

https://www.wired.com/story/eternalblue-leaked-nsa-spy-tool-hacked-world/

Legislature: Required Backdoors

In December 2018 Australia passed Assistance and Access Bill

- Compels companies to provide access to encrypted data in apps
- One could say the problem is broad, vague language, but...
  - the law could be interpreted to require Australian nationals, even working abroad, to introduce backdoors into software *without notifying their employer*
- Lawmakers say they have no intention of demanding backdoors, engineers say there is no other way to comply with the law.

https://www.wired.com/story/australia-encryption-law-global-impact/

Interest Group: Adtech

It is maybe obvious that incentives here don't promote privacy in general, but it is also worth noting that these incentives align with law enforcement incentives.

It is pretty easy to find Facebook's guidelines (below) for providing information to law enforcement or even *civil litigants*, for example.

We'll talk more about encryption controlled by the user. This would make these data handoffs impossible.

Other things being equal, enabling law enforcement is great. Iron triangles often turn out not to be.

https://www.facebook.com/safety/groups/law/guidelines/

Tim Cook's Fixes

Tim Cook advocated four values for handling data.

- data minimization
  - collect less and de-identify more
- transparency
  - right to knowledge of how data is collected, shared, etc.
- the right to access
  - users should be able to view, amend, or delete their data
- the right to security
  - stop losing our data!

Do you think a data ecosystem that reflects these values is achievable? What are some obstacles?

Back to Basics: Simple Arithmetic

Who will volunteer to do some simple arithmetic for me?

Assignment: Basic Bookkeeping

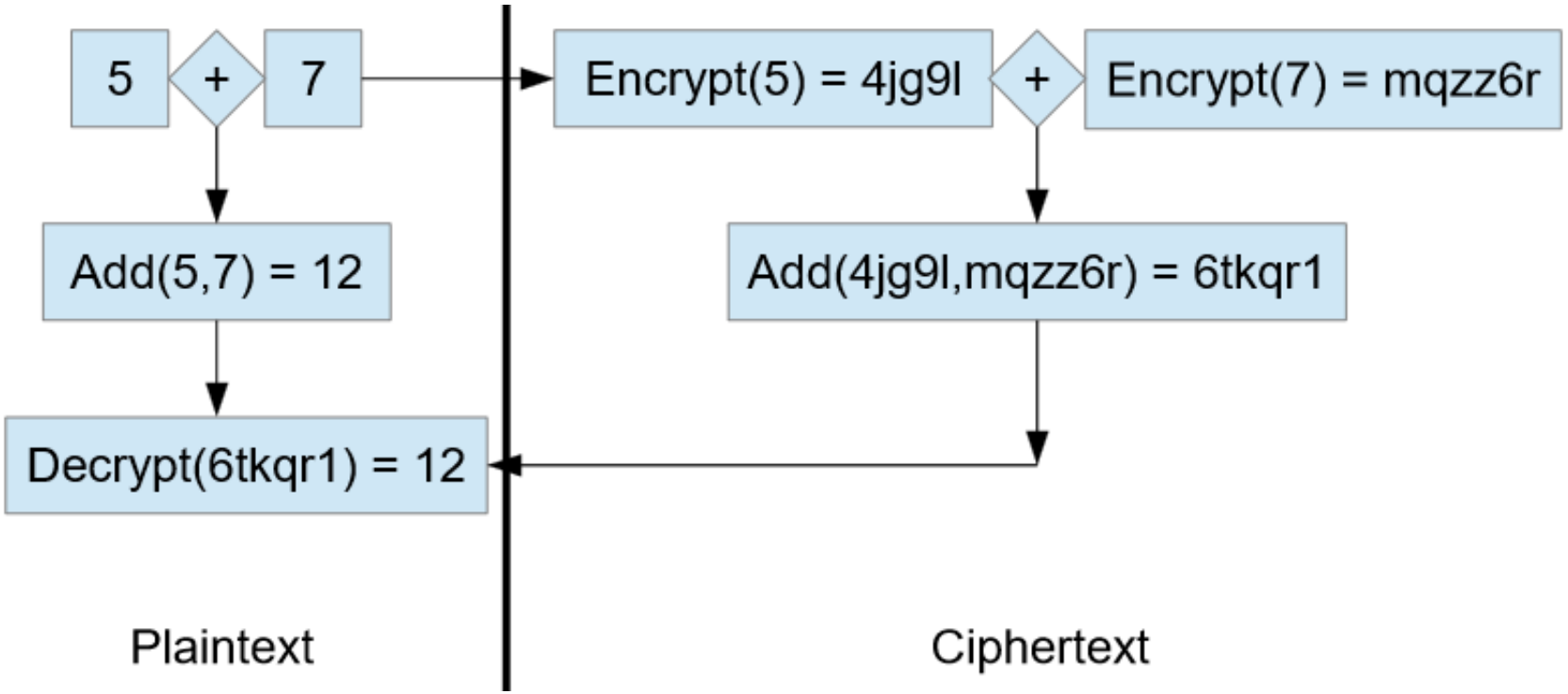
Please go ahead and...

- verify that the ending funds figure is correct, but
- maintain the secrecy of expense amounts and starting funds

Expense	Amount	Starting Funds
fancy pens	XXXXXXXXX	XXXXXXXXX
private jet	XXXXXXXXX	**Ending Funds**
"business entertainment"	XXXXXXXXX	$30173
attorneys	XXXXXXXXX	

Thanks! Easy as pie, right? RIGHT?!

Homomorphic Encryption



Homomorphic encryption allows one to do computations on encrypted data and get the “correct” answer after decryption.

What is a zero knowledge proof?

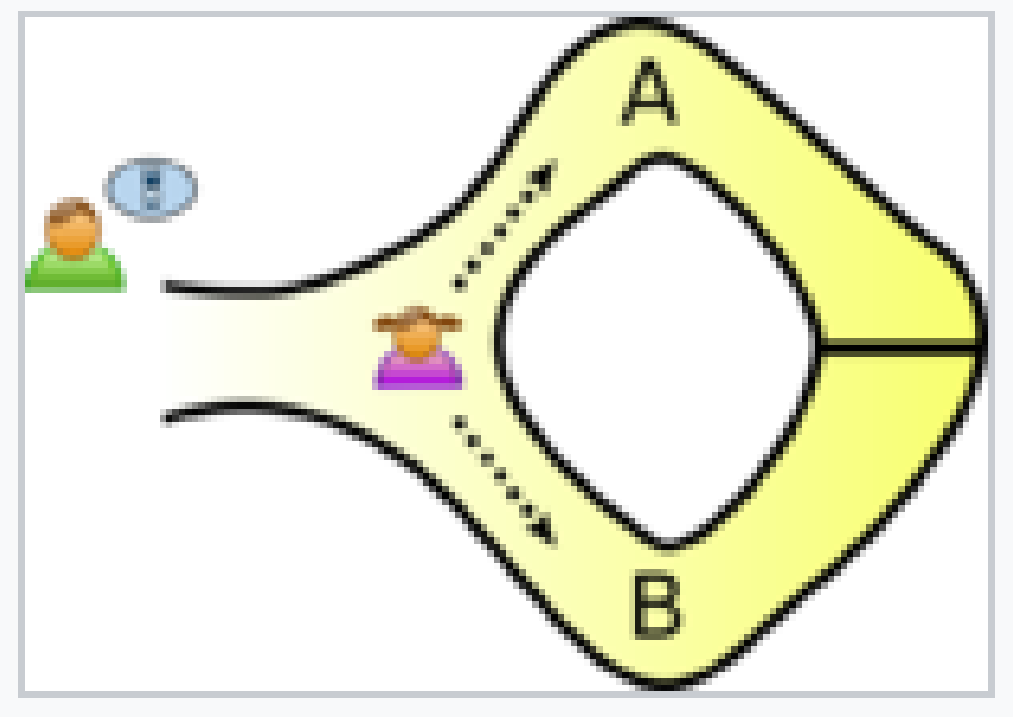
A **zero-knowledge proof** is a method by which a *prover*, traditionally called Peggy, proves to a *verifier*, traditionally called Victor, that

- Peggy possess a particular piece of information, and...
- Peggy does not reveal that information to Victor.

Example: Digital Signatures

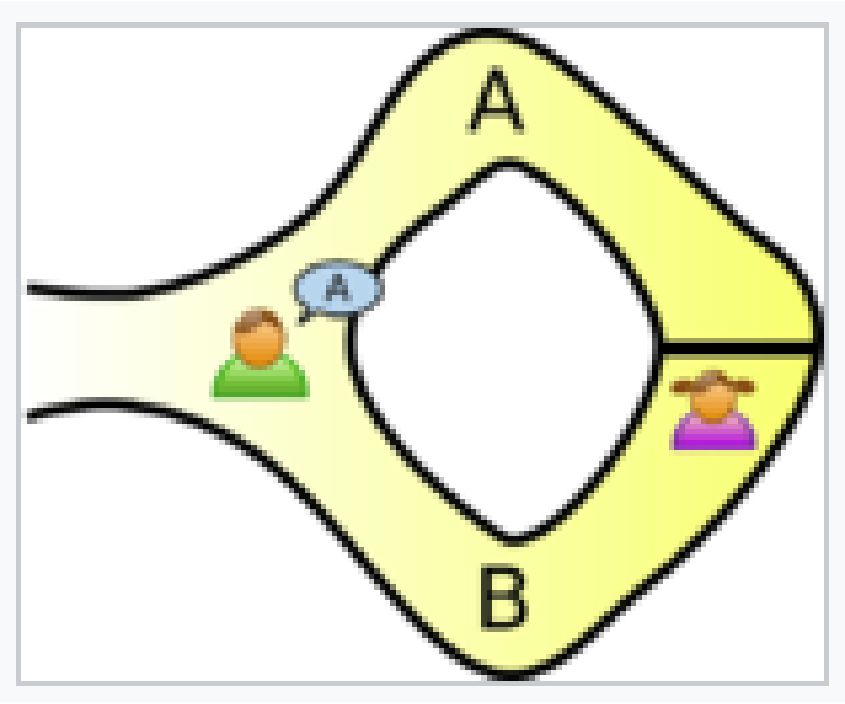
Peggy proves to Victor that she holds the private key corresponding to a particular public key but Victor does not receive any information about the private key itself.

Classic Cave Example 1



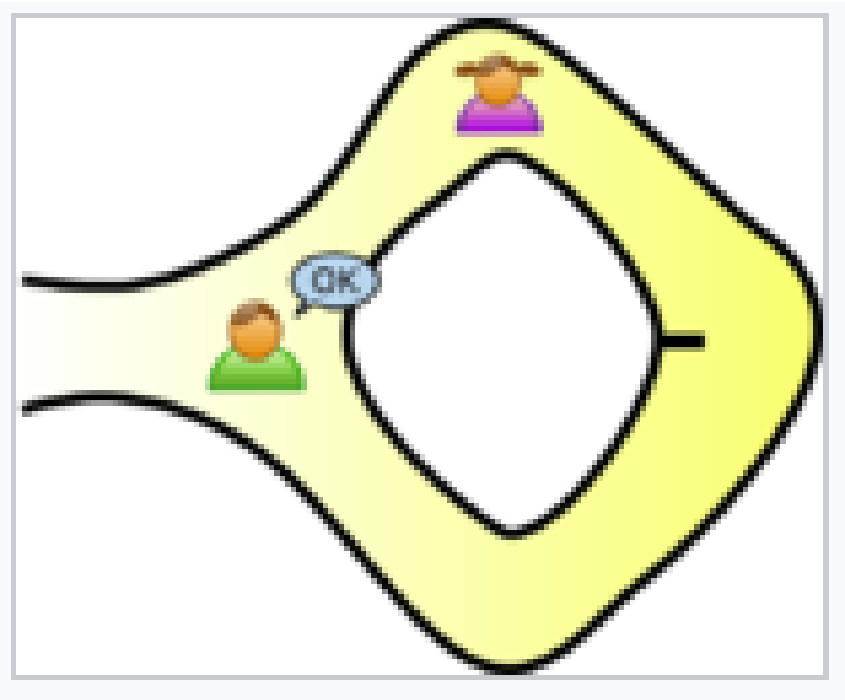
Peggy wishes to prove to Victor that she can **open the magic door** in the back of a donut-shaped cave, yet she wants to **keep her method for opening the door** a secret from Victor.

Classic Cave Example 2



After Peggy has entered the cave by one path or the other, Victor shouts into the cave whether Peggy should return by Path A or Path B.

Classic Cave Example 3



If Peggy can consistently return by the path Victor requests, she has demonstrated that she has a method for passing through the magic door, yet Victor has not seen Peggy open the door.

Basic Bookkeeping Revisited

Expense	Amount	Starting Funds
fancy pens	XXXXXXXXX	XXXXXXXXX
private jet	XXXXXXXXX	**Ending Funds**
"business entertainment"	XXXXXXXXX	$30173
attorneys	XXXXXXXXX	

Adding up the unknown, encrypted amounts to get an unknown and correct but still encrypted amount is **homomorphic encryption**.

Verifying that this encrypted amount is $30,173 without decrypting anything is a **zero-knowledge proof**.

Data Ownership

Would you like to “own” your personal data?

What does it mean to own information?

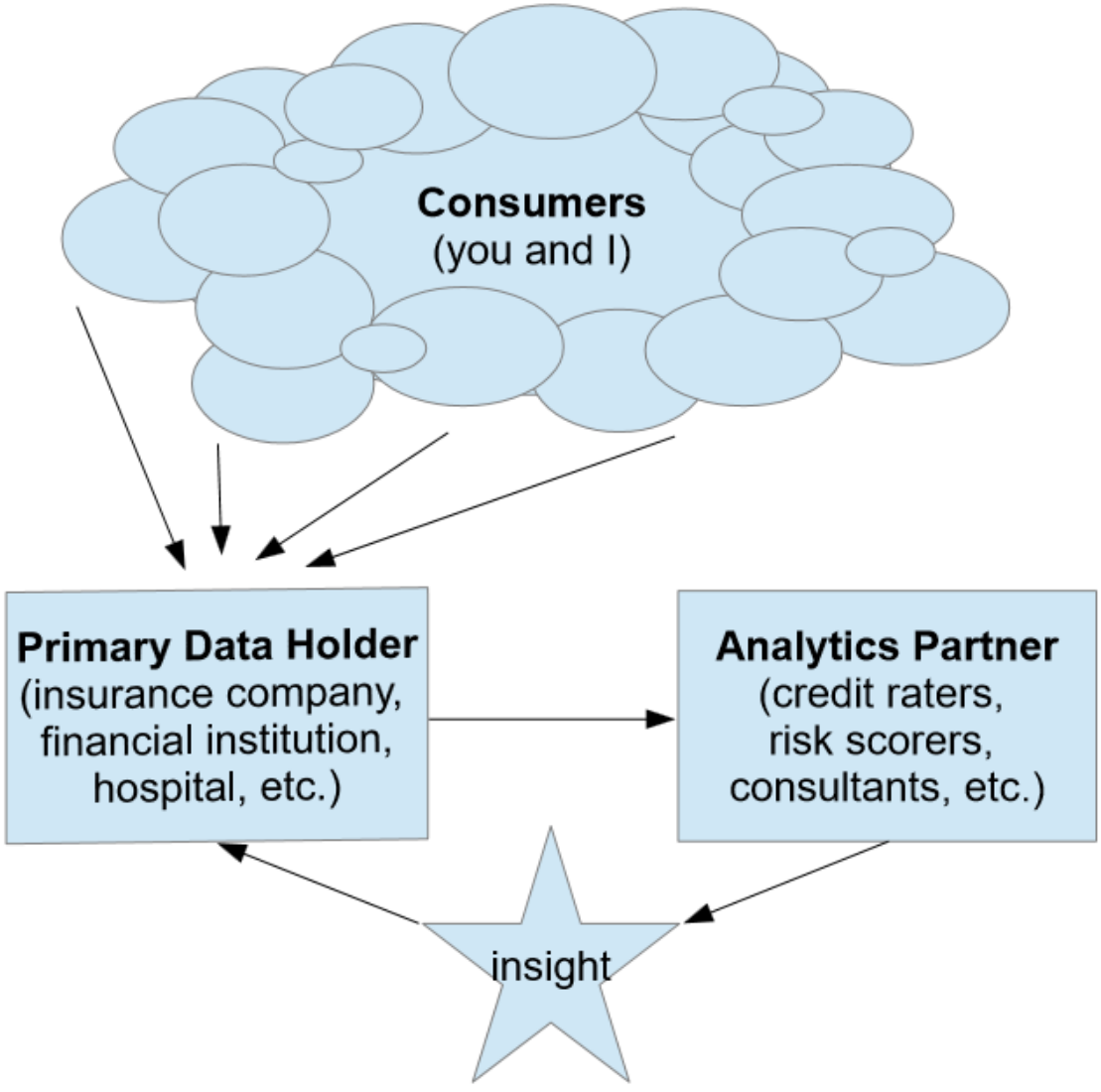
Analytics Partners

Personal data you give to a business often passes through a number of hands in its lifetime. Partners that primary data holders (example: a hospital) work with to better analyze your data (example: consultants) are an interesting example we'll examine in detail.

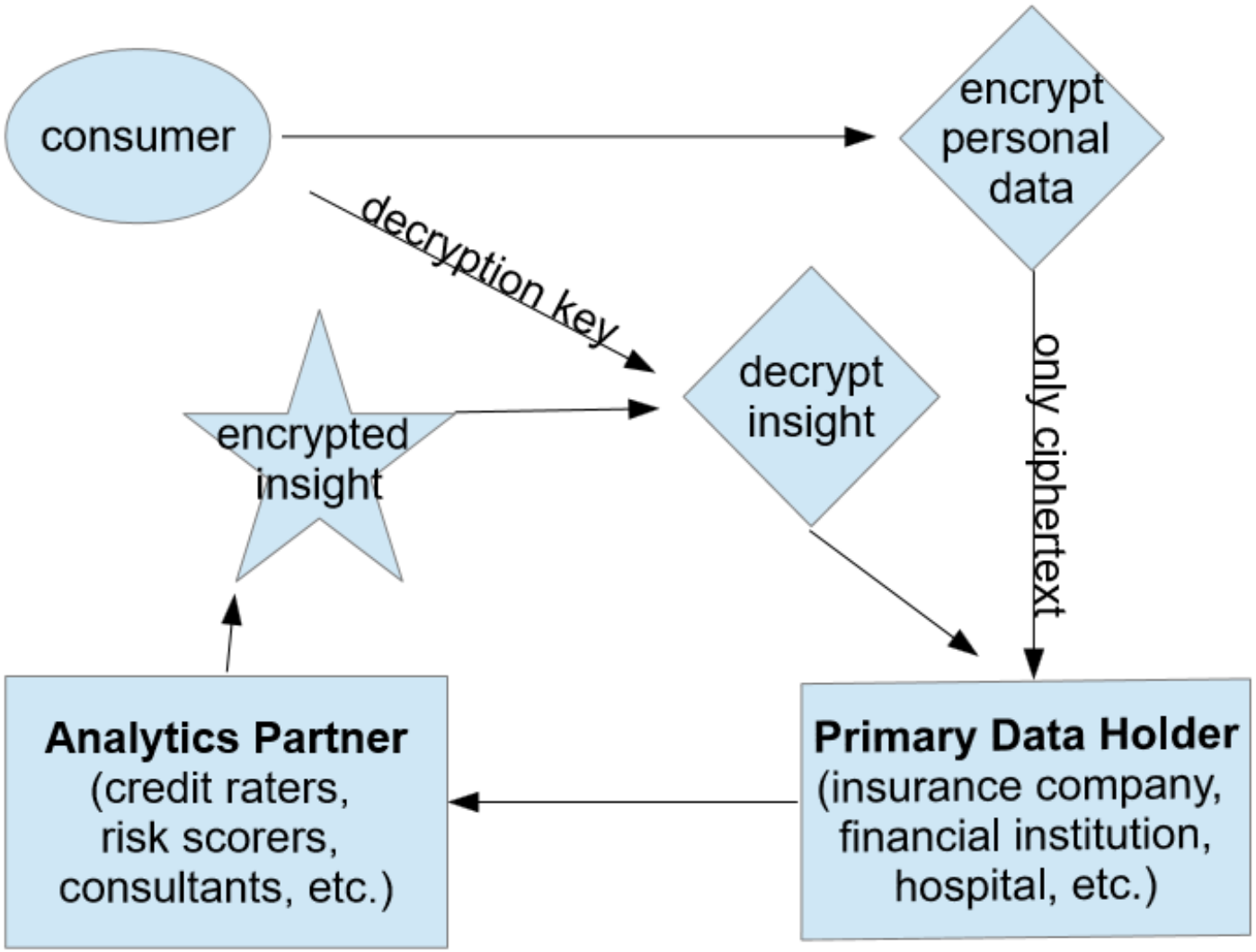
Depending on the industry, there are many businesses in this rough category...

- Raters of consumer credit
- Clearing houses for financial transactions
- Risk scoring in insurance
- Many forms of consulting

Typical Data Flow



Data Flow with Cryptographic Ownership



Contrast Between New and Old

In the newer cryptographic ownership model...

- no one but the consumer holds unencrypted personal information.
- analytics insights are the only plaintext in circulation.

When data is decrypted...

- it requires the approval of the consumer.
- it occurs late in the pipeline passes through fewer hands.
- it reveals only necessary, practical information.

The ability to tactically **allow or prevent utilization of a resource** agrees with our everyday intuition of what “ownership” should mean.

Tim Cook's Fixes

A data ecosystem with pervasive use of homomorphic encryption naturally enforces much of what Tim Cook suggested...

- data minimization
  - encryption is a form of de-identification
- transparency
  - data is useless to others until you give decrypt permission
- the right to access
  - decrypt permission is a natural opportunity of user access
- the right to security
  - in a sense, there is no longer anything to lose

What's Out There?

- HElib - fully homomorphic encryption software library
  - the algorithms, close to the metal in C++
  - https://github.com/shaih/HElib
- HomomorphicEncryption R Package
  - intuitive interface in R
  - http://www.louisaslett.com/HomomorphicEncryption/
- PySEAL Python Package
  - intuitive interface in Python
  - https://github.com/Lab41/PySEAL
- OpenMined - crowdsourced, private machine learning
  - https://github.com/OpenMined
- Zcash - cryptocurrency using zero-knowledge proofs
  - https://z.cash/

History and Prophecy

Some landmark papers and people...

- (Gentry 2009) “A Fully Homomorphic Encryption Scheme”
- (Gennaro, Gentry, Parno, Raykova 2013) “Quadratic Span Programs and Succinct NIZKs without PCPs”

Craig Gentry won a MacArthur Fellowship (the “genius grant”) for his 2009 PhD thesis work on fully homomorphic encryption (able to compute an arbitrary algorithm on ciphertext) and was widely interviewed afterwards.

He said at the time, almost 10 years ago, that it might be 10 years before the technology really started to catch on...

Questions

Any questions?

Feel free to contact me at acmueller@capnion.com

Slides are available at
https://github.com/capnion/ ...
random/blob/master/acm_capnion_feb19stlcyber.pdf