Cryptography and the Data Industrial Complex

Alexander C. Mueller PhD
CEO and Founder of Capnion

April 9, 2019

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

Anomalous Goatees in Media

Visitors from a Parallel Dimension

Someone like you from another dimension, the same but different, arrives to tell you something about your world by describing their own.

Agenda

PHIL 100 for A.I.

- Who are you?
- Why are you here?

POLISCI 100 for A.I.

- Tim Cook's Data Industrial Complex
- Iron Triangles
- Bureaucracies / Legislatures / Interest Groups

Transformative Technology

- Idiosyncrasies of data ownership
- What? and why? of homomorphic encryption
- Transformation of business models

Who *are* you?

What would be left of your life if all the **information about you** held by business and government vanished overnight?

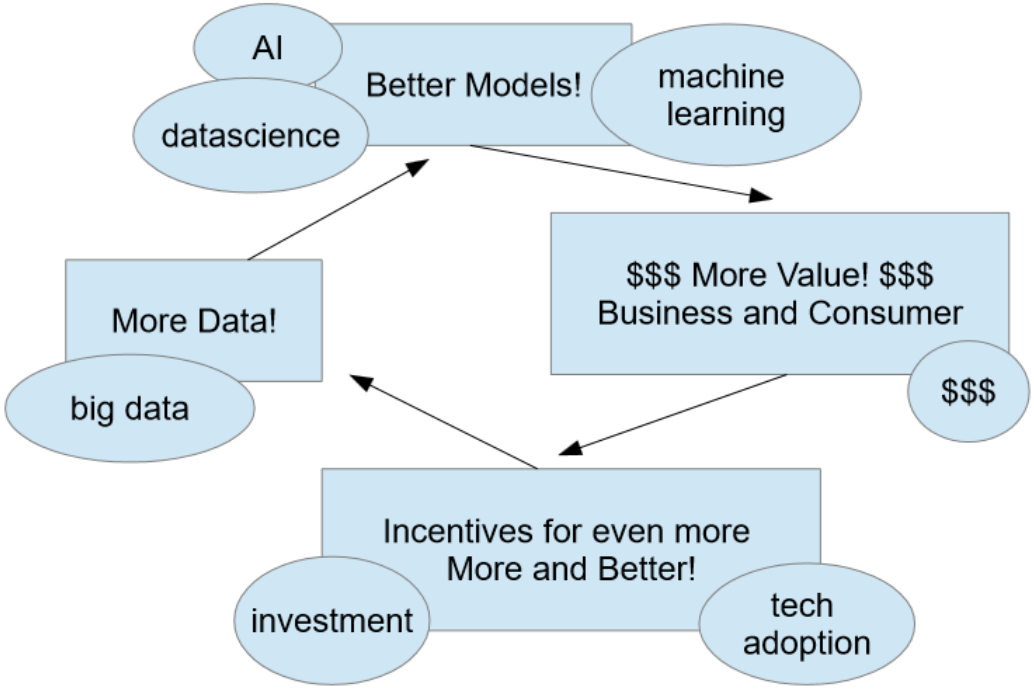
You would retain your human relationships, but you might struggle to access and maintain them. You would **have a rough time** at the grocery store, the bank, the social security office, ...

The functional member-of-society “you” is dangerously close to being the **collection of your P.I.I.** such as your name, SSN, driver’s license number, mother’s maiden name...

How much of “you” have you **handed over to robots** recently?

Why are you here?

Our buzzwords are really all about the possibility of **runaway, reinforcing data throughput**...



We're all here because we're interested in watching the merry-go-round spin faster and faster, but **we are also riding it**.

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

What iron triangles, if any, exist around the data economy?

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 (notably e2e encrypted chat 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: You!

Hopefully, at least one interest group springs to mind...

What sorts of regulations on data appeal to you...

- as a citizen and consumer?
- as a data professional?

Are the answers to these two questions the same?

Prospects for Future Regulation

"If I’m emailing within WhatsApp ... does that inform your advertisers?" - a United States Senator

Who will (realistically) provide the details of future regulations?

Some Radical Ideas

Data Ownership 1

Would you like to “own” your personal data?

What does it mean to own information?

Data Ownership 2

What sorts of activities typify our idea of ownership?

- Physical posession
- Employment in useful work at will (or not)
- Destruction or denial of use to others

Data is unusual for its easy reproducibility.

- Useful work (typically) requires possesion, but ...
- Possession is not easily revokable, and ...
- The power to deny use is quickly lost forever.

How can we separate work from possession?

Back to Basics

If only I could find someone with some **basic decency** and respect for other **people's privacy** to help me out.

I am not asking for anything too crazy, just someone competent in **basic arithmetic** to help with some **bookkeeping**.

Is that person you? Would you **care to volunteer**?

Assignment: Basic Bookkeeping

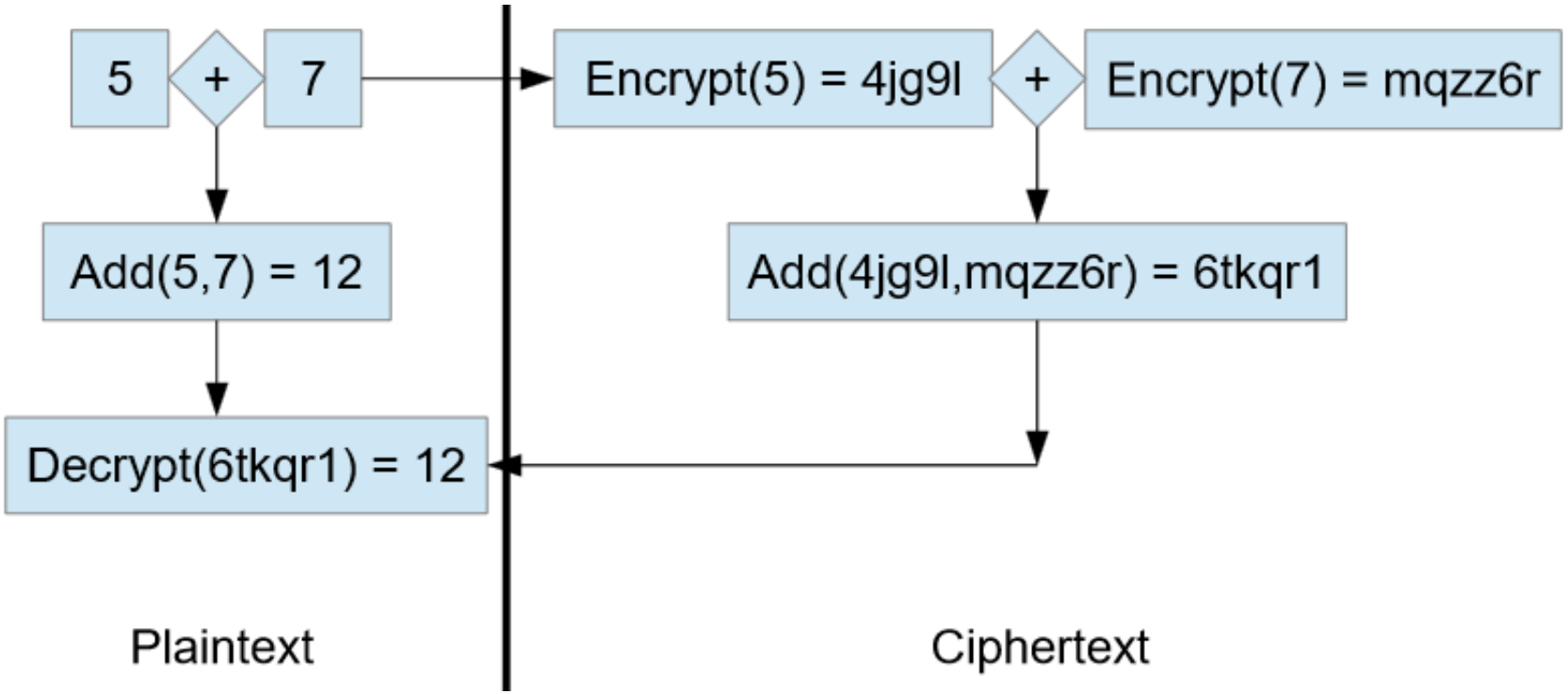
Please go ahead and total the expenses by category.

Expense	Category	Amount
Southern Illinois Networking	fdg4shN8rgj!	hw54hW&s
Jet Charter	4fh5$7Dd27	bZs74jj5v
Elite Executive Pens Dot Com	lyBe9zy&b2	64nwHsn
...	...	...

Easy, right? Thanks for handling this. It's so hard to find decent help these days. I need that jet deduction!

What's the problem here? Who owns this data?

Homomorphic Encryption



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

Buzzword Salad

If you want in on some of this action
but you have no expertise in your organization...
What do you do?

Analytics Partners

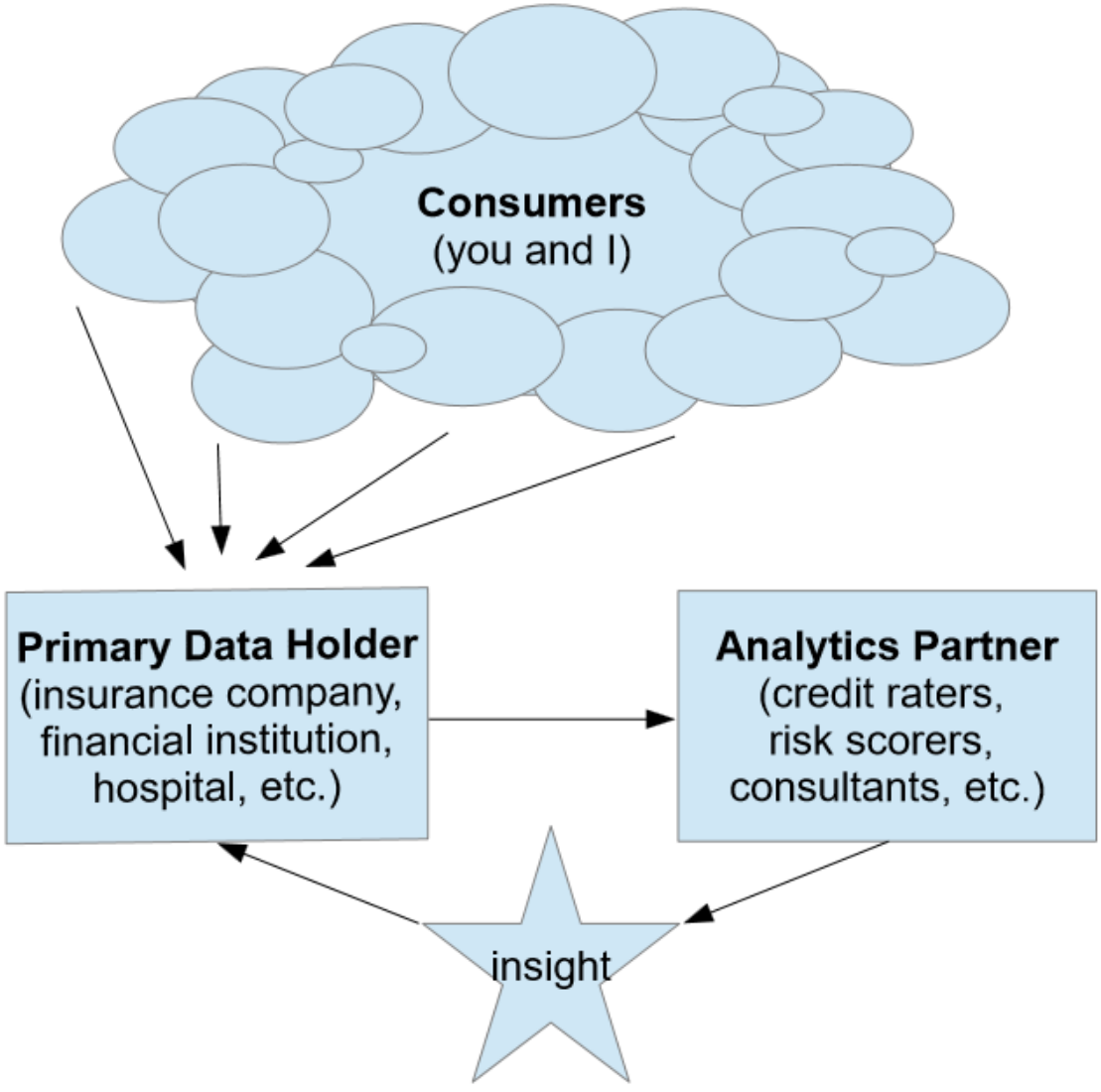
Personal data you give to a business often passes through a number of hands in its lifetime.

Partners that primary data holders (example: a hospital) work with to better analyze your data (example: **A.I. companies**) are an interesting example we'll examine in detail.

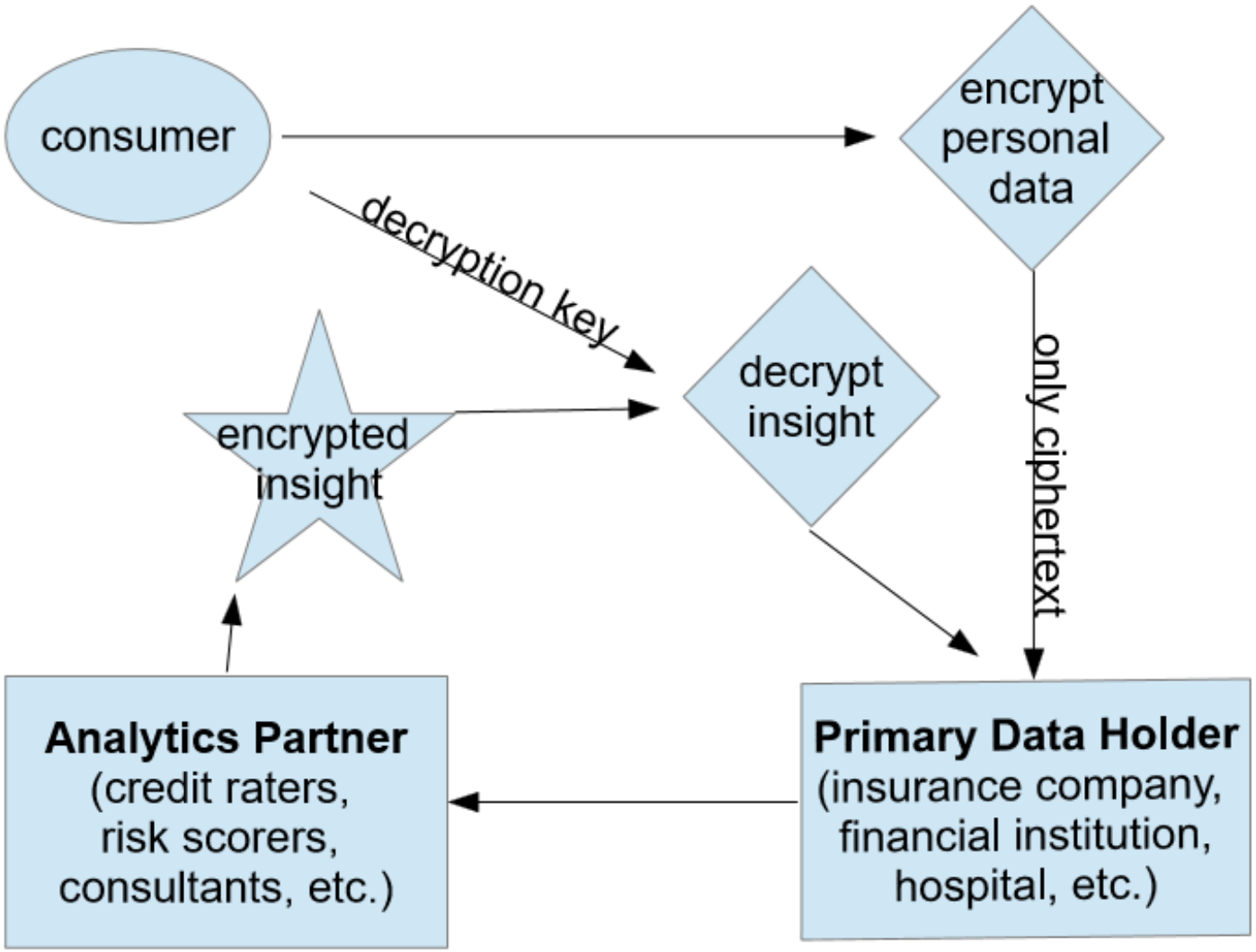
There are many businesses in this rough category...

- Raters of consumer credit
- Clearing houses for financial transaction
- Risk scoring in insurance
- **Probably your employer!?!?**

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 sensitive plaintext.
- analytics insights are the **only plaintext in circulation**.

When data is decrypted...

- it requires the **approval of the consumer**.
- it occurs late and the data passes through fewer hands.
- it reveals only **practically necessary information**.

The consumer has expanded powers of ownership as they can

- tactically **allow or prevent use** of data
- all but **destroy the data** by destroying keys

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/