Welcome 2017 Faculty Summit Attendees

Faculty Summit 2017
microsoftfacultysummit.com

Microsoft Research
Microsoft.com/research

Facebook
@microsoftresearch

Twitter
@MSFTResearch
#FacSumm
#EdgeofAI

Microsoft Research
Faculty Summit **2017**

Microsoft Research
Faculty Summit
2017

Private AI

Kristin Lauter
Cryptography Research Group
Microsoft AI and Research

Cryptography Research Group

Kristin Lauter
Principal Researcher

Ran Gilad-Bachrach

Melissa Chase

Ranjit Kumaresan

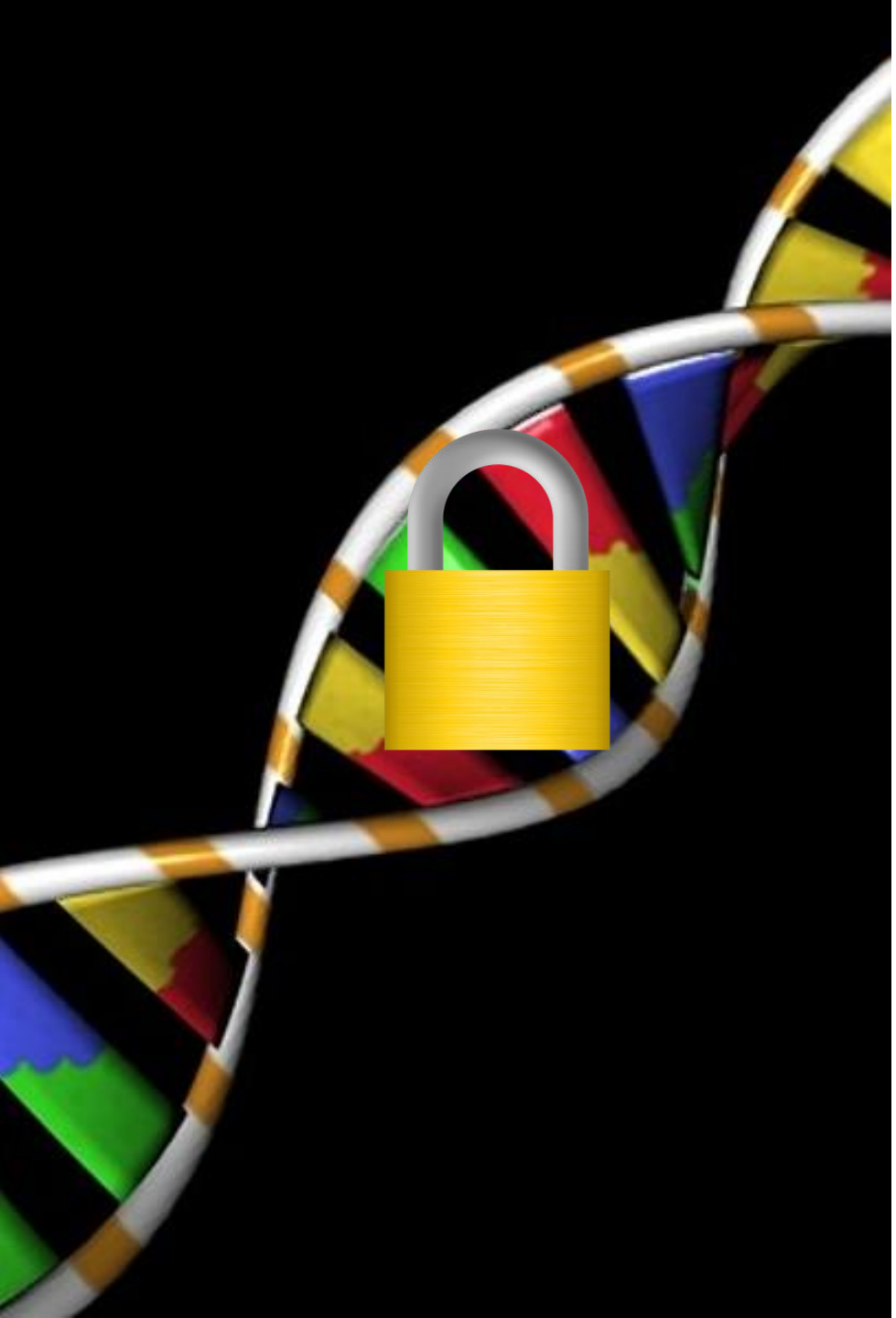
Hao Chen

Kim Laine

- Advances in AI allow us to use a trail of breadcrumbs of information to make complex predictions on people health, behavior and more.

- Are we at risk of losing our privacy?

- Will people share their data?



Private-AI

- Using advanced cryptology technologies (Homomorphic Encryption, Secure Multi-Party Computation, Differential Privacy) we develop AI tools that provide provable privacy guaranties.
- Private-AI tools allow the cloud to run AI tasks, such as making genomic predictions or training neural networks while being blind-folded.

Private AI

Add privacy to the entire machine learning process

Private AI: Challenges

Training

- Exposes training data to model builder

Prediction

- Exposes prediction data to model owner
- Leaks training data

Statistics

- Aggregation exposes data

Private AI: Our Technologies

Training

- Build models on distributed data without sharing the data

Prediction

- Make private predictions

Statistics

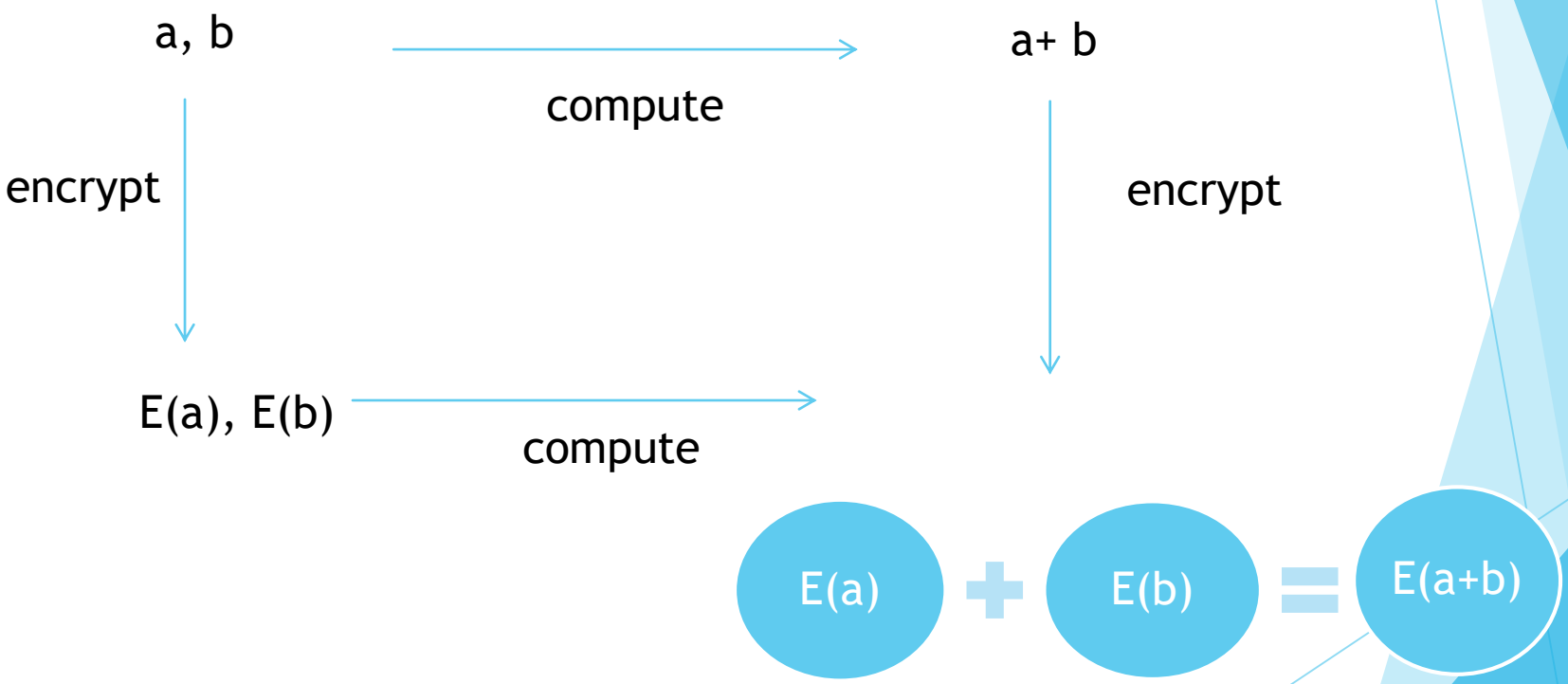
- Compute privately aggregated statistics

Private AI:
Privacy-Preserving Evaluation

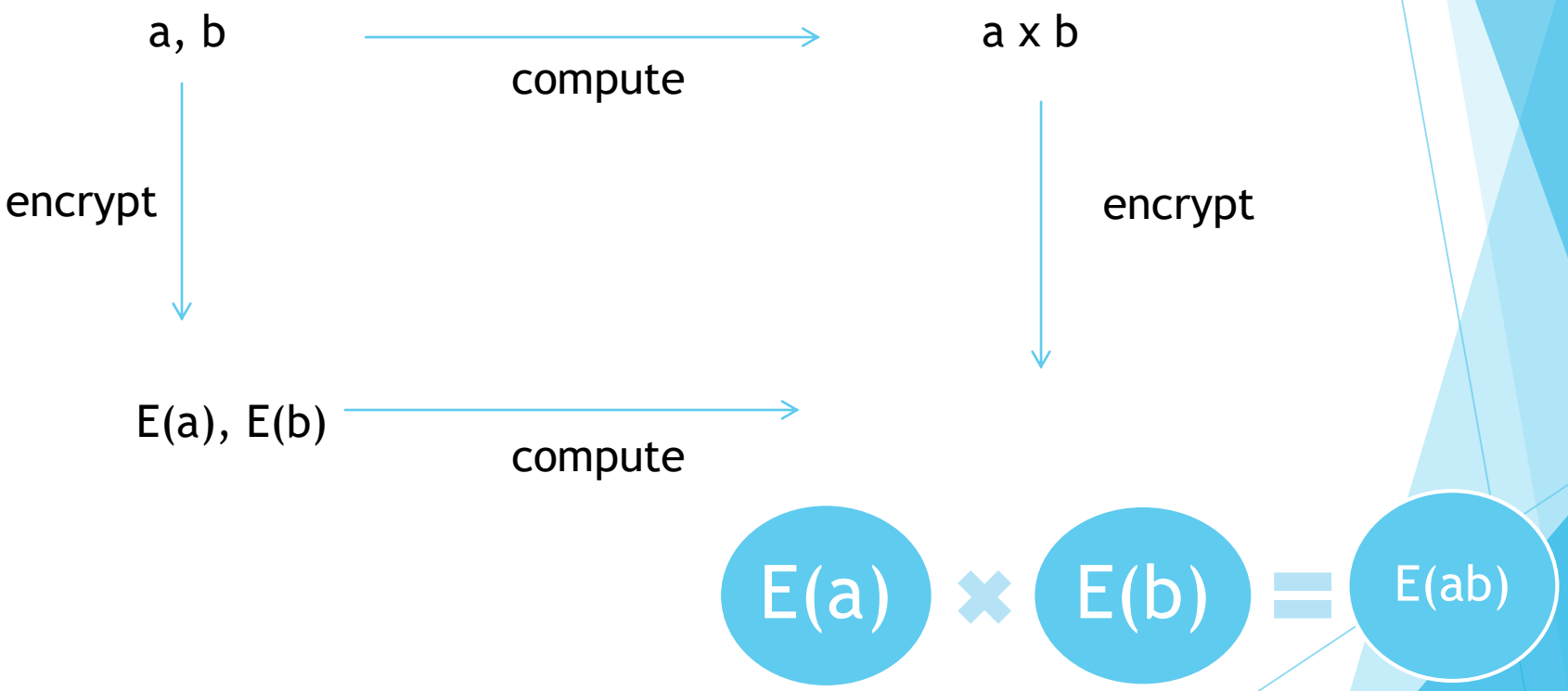
Technology: Homomorphic Encryption

- Encrypt input data
- Evaluate functions on encrypted data
- Result remains encrypted
- Data owner decrypts to obtain result

Homomorphic Encryption: addition



Homomorphic Encryption: multiplication



Protecting Data via Encryption:
Homomorphic encryption

1. Put your gold in a locked box.
2. Keep the key.
3. Let your jeweler work on it through a glove box.
4. Unlock the box when the jeweler is done!

Mathematics of Homomorphic Encryption

- New hard problems proposed (2009-2013), related to well-known hard lattice problems
  - Small Principal Ideal Problem, Approximate GCD, Learning With Errors (LWE), Ring-Learning With Errors
- Lattice-based Cryptography:
  - Compare to other public key systems: RSA (1975), ECC (1985), Pairings (2000)
  - Proposed by Hoffstein, Pipher, and Silverman in 1996 (NTRU), Aijtai-Dwork
- Hard Lattice Problems:
  - approximate Shortest Vector Problem, Bounded Distance Decoding
- SECURITY:
  - best attacks take exponential time
  - secure against quantum attacks (so far…)

Private AI:
Privacy-Preserving Evaluation

Technology: Homomorphic Encryption

- Simple Encrypted Arithmetic Library – SEAL
- Developed by our group since 2015
- Written in C++, includes .NET wrappers
- Freely available at http://sealcrypto.codeplex.com

The Register®
Biting the hand that feeds IT

Security

Microsoft boffins build better crypto for secure medical data crunching

Practical homomorphic encryption manual released

16 Nov 2015 at 03:57, Team Register

As genome research - and the genomes themselves - get passed around the scientific community, the world's woken up to the security and privacy risks this can involve. A Microsoft research quintet has therefore published ways to help scientists work on genomic data while reducing the risk of data theft.

The team published an informal manual to help scientists and other researchers to use the Simple Encrypted Arithmetic Library (SEAL).

Microsoft Helps Out Healthcare Sector with New Data Encryption Algorithm *UPDATE*

Microsoft reveals SEAL - Simple Encrypted Arithmetic Library

Previous reports have pointed the finger at the healthcare sector as being woefully unprepared for the modern age of Internet-enabled devices that are always online and present a constant danger to the patient, hospital, and insurer data.

This lack of security measures comes from the fact that, for many years, both the hardware and software part of healthcare applications couldn't handle the amount of data doctors and researchers needed, so no industry standards were put in place to protect sensitive data of any form.

Now, as technology has evolved, the healthcare sector is trying to catch up from behind with other industries, but the previous years, when it did not make a habit from protecting data, left a big hole to fill.

While many healthcare providers and medical research companies are putting more effort into catching up with modern-day security practices, there's still a lot of work to be done, which requires both time and financial resources to adapt various security tools to the medical industry.

Microsoft will provide a free tool to help with biomedical data processing and encryption

In a recent paper released on its research portal, Microsoft has announced a new encryption library that implements the theory of homomorphic encryption.

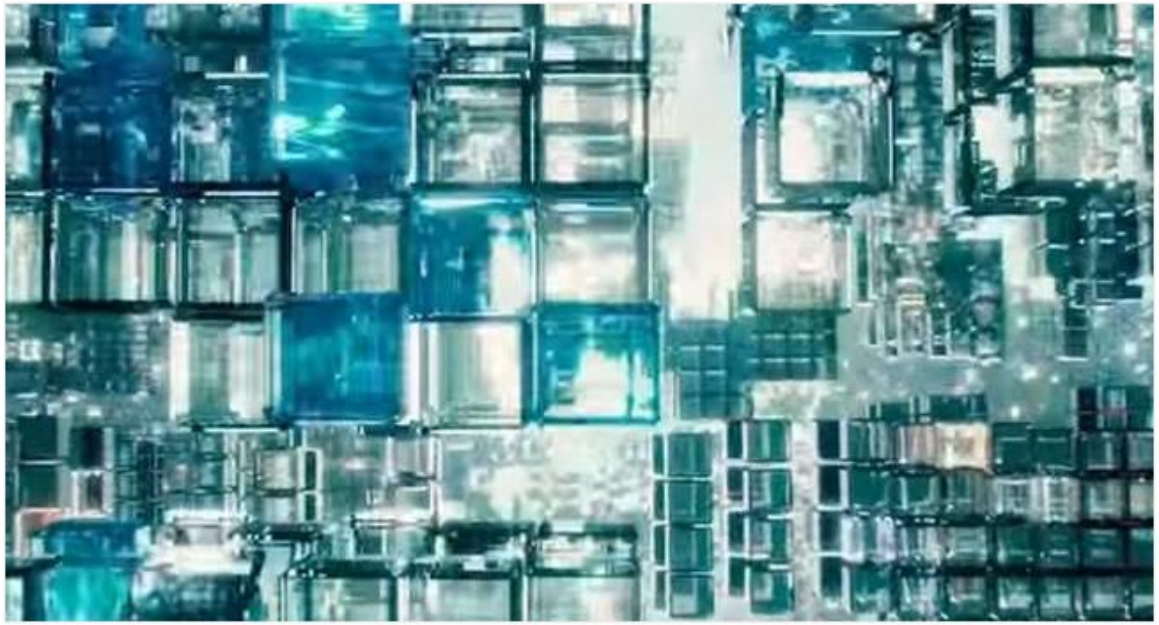
Homomorphic encryption is a method of encryption that encodes data in such a way that it allows developers to work with the encrypted data as if it were in unencrypted form.

The Register®
Biting the hand that feeds IT

Software

Microsoft researchers smash homomorphic encryption speed barrier

Artificial intelligence CryptoNets chew data fast but keep it safe



Ultron it isn't, thank goodness

9 Feb 2016 at 08:02, Iain Thomson

Exclusive Microsoft researchers, in partnership with academia, have published a paper detailing how they have dramatically increased the speed of homomorphic encryption systems.

With a standard encryption system, data is scrambled and then decrypted when it needs to be

Private AI:
Privacy-Preserving Evaluation

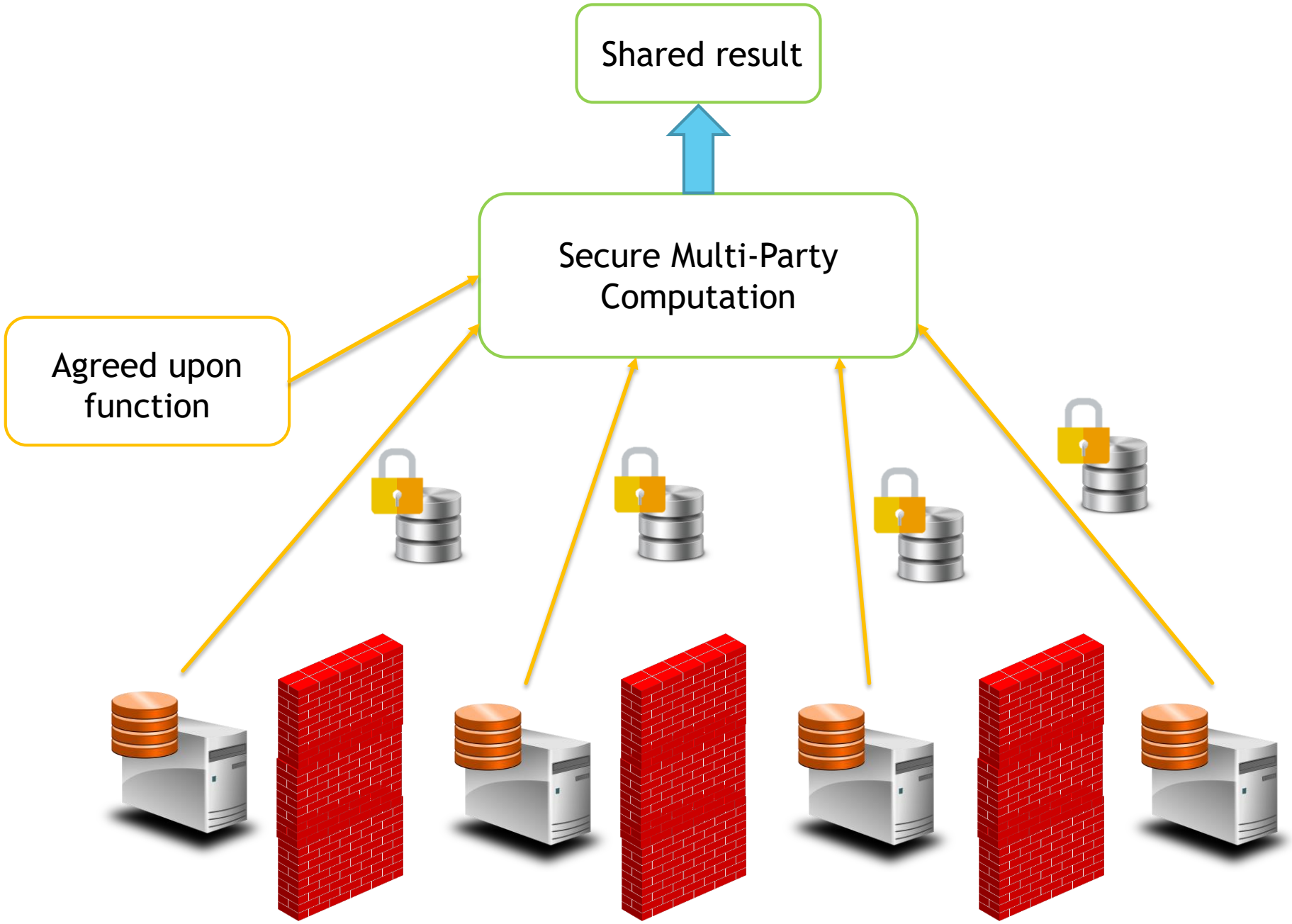
Technology: Homomorphic Encryption

- Demos: Using SEAL for private predictions

Private AI:
Computing on Distributed Data

Technology: Secure Multi-Party Computation (MPC)

- Privacy-preserving computation on distributed data


Shared result
Secure Multi-Party Computation
Agreed upon function

Private AI:
Training on Distributed Datasets

Technology: Secure Multi-Party Computation (MPC)

- Privacy-preserving computation on distributed data
- SMILY library (not currently released)

Private AI:
Training on Distributed Datasets

Technology: Secure Multi-Party Computation (MPC)

- Privacy-preserving computation on distributed data
- SMILY library (not currently released)
- MPC is a powerful tool
- Can be used to compute statistics on distributed datasets
- Expensive on large computations (communication bottleneck)
- Hard to apply *directly* to training (big input, big computation)

Private AI:
Training on Distributed Datasets

Technology: Secure Multi-Party Computation (MPC)

- Local computation to make secure step smaller (then iterate)
- Neural Networks, Boosted Trees, Random Forests, ...
- Many statistical models allow local computation (POC: ANOVA)
- Good for: Few parties with lot of data (lot of local computation)
- Challenge: Many parties with little data (little local computation)

Private AI: Privacy-Preserving Models

Technology: Differential Privacy

- Possible to extract sensitive training data from ML models
- Differential Privacy adds ”noise” during training
- *Prove* that resulting model does not leak training data
- Sometimes yields *better* models
- Combine with training on distributed data

Thank You!

Kristin Lauter (Principal Researcher)

klauter@microsoft.com

Ran Gilad-Bachrach

rang@microsoft.com

Hao Chen

haoche@microsoft.com

Melissa Chase

melissac@microsoft.com

Kim Laine

kim.laine@microsoft.com

Ranjit Kumaresan

ranjit.kumaresan@microsoft.com

SEAL library

http://sealcrypto.codeplex.com

Thank you


Microsoft Research
Faculty Summit
2017
The Edge of AI
Microsoft

Microsoft


Microsoft Research
Faculty Summit
2017
The Edge of AI
Microsoft