Distributed Private Machine Learning

Abhradeep Guha Thakurta

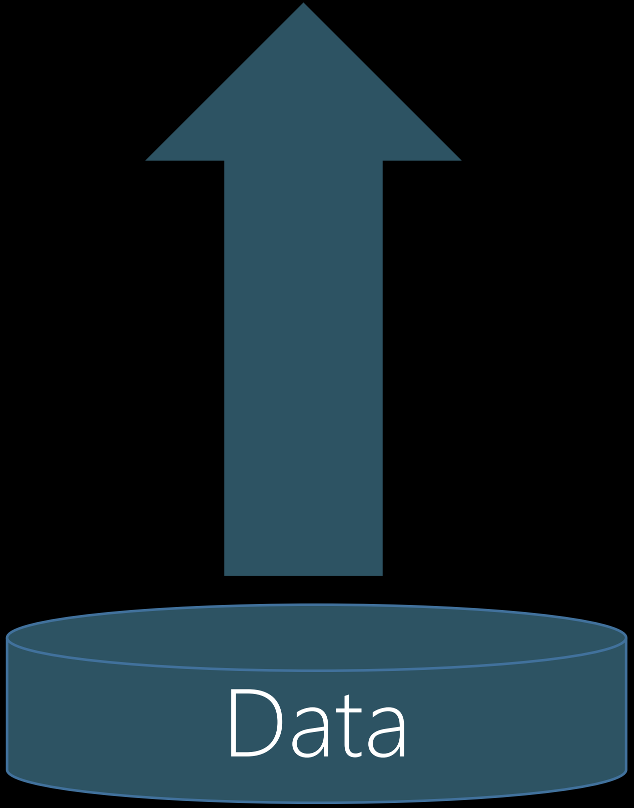
University of California Santa Cruz

Distributed learning from private data

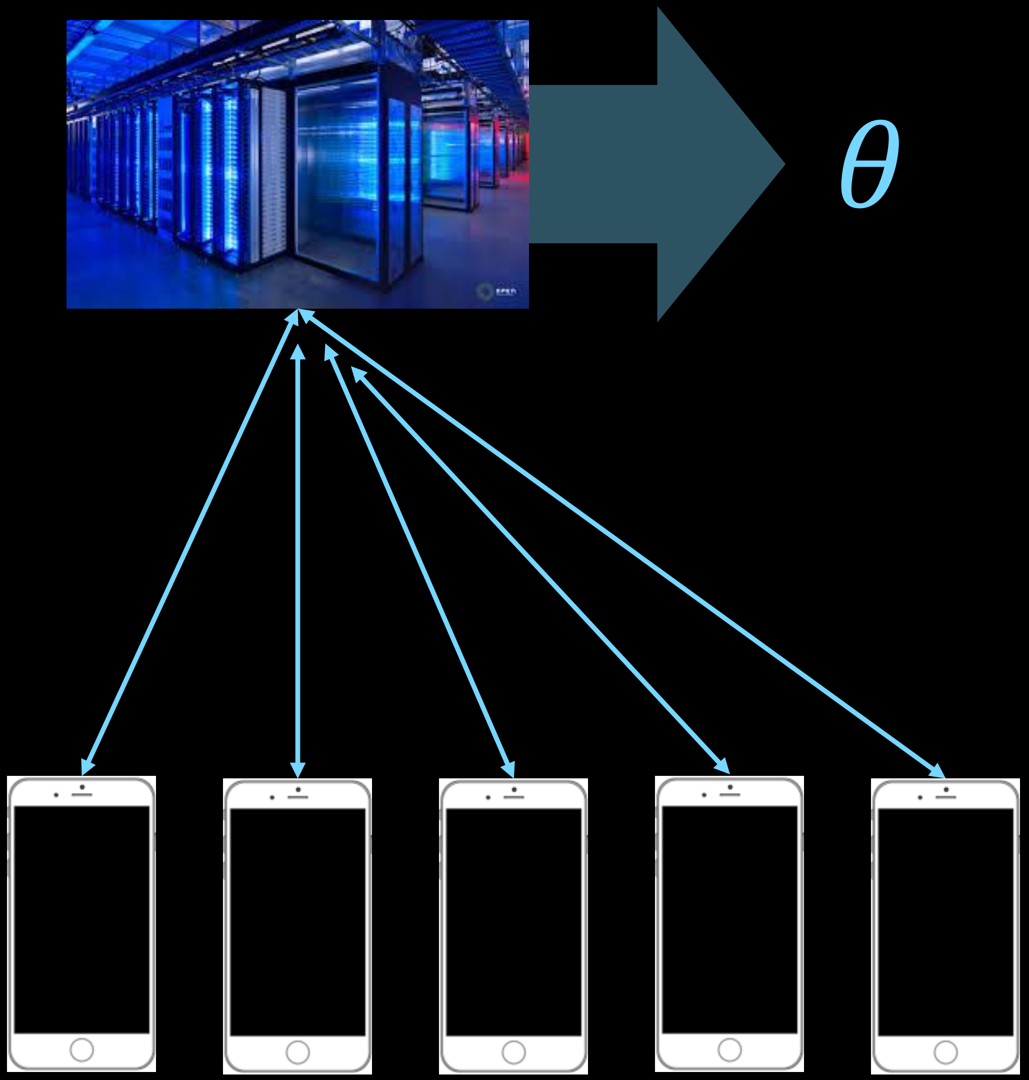
Distributed machine learning - Setup

Traditional machine learning

θ : Model / statistic

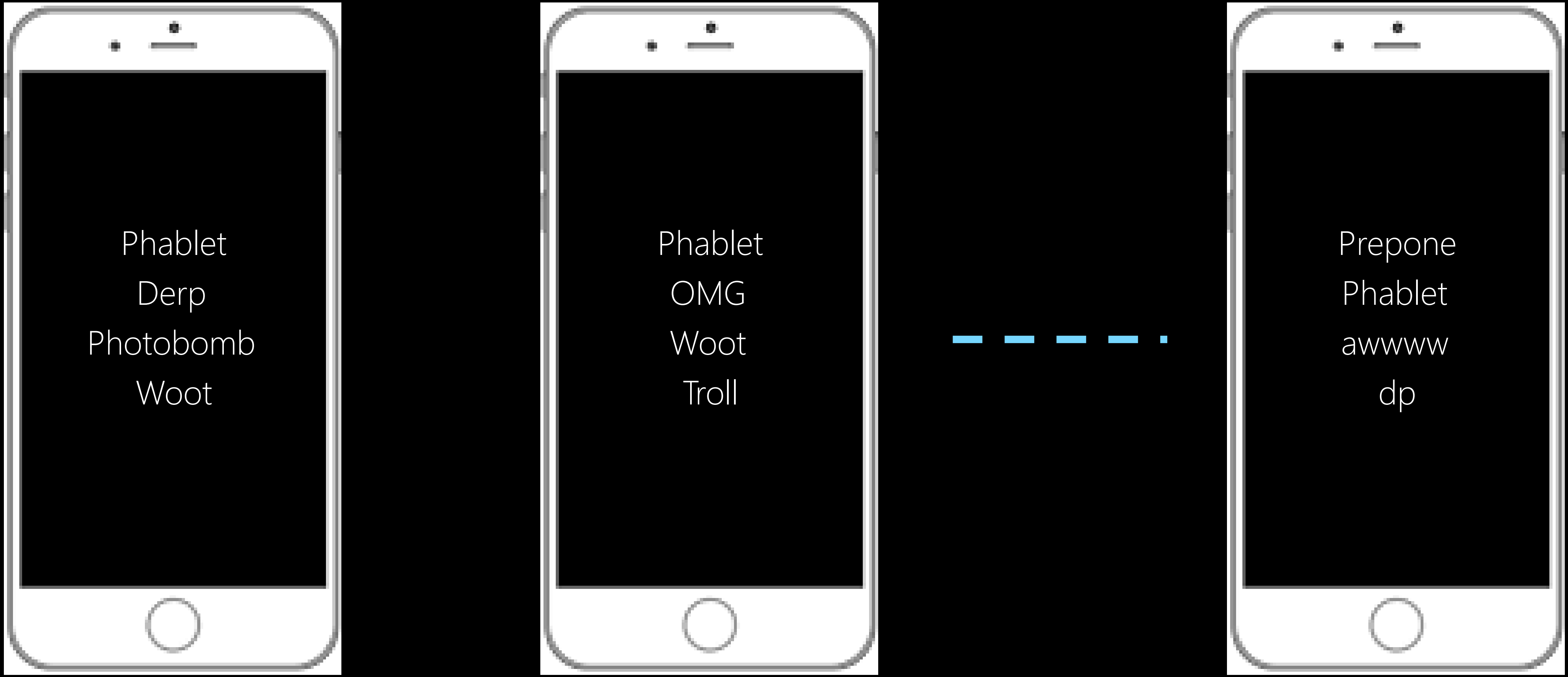


Distributed machine learning



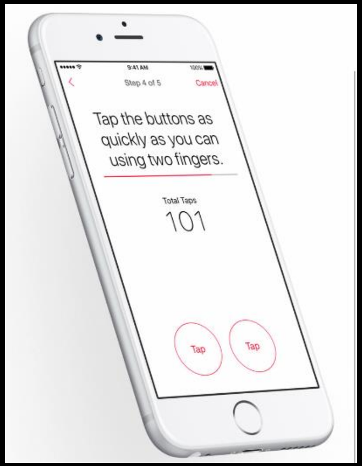
Learning from private data

Learn new (and frequent) words typed



Learning from private data

Predict if a person has Parkinson’s disease



Get measurements from
gyroscope, display screen etc.

YES / NO

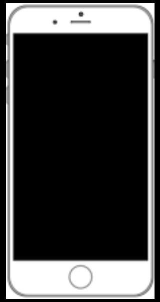
Model / classifier

θ

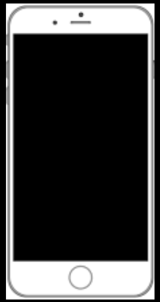
Image courtesy: Research kit (Apple)

Learning from private data

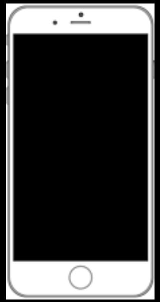
Collaborative filtering



1	?	3	?	?	2



2	?	?	?	?	4

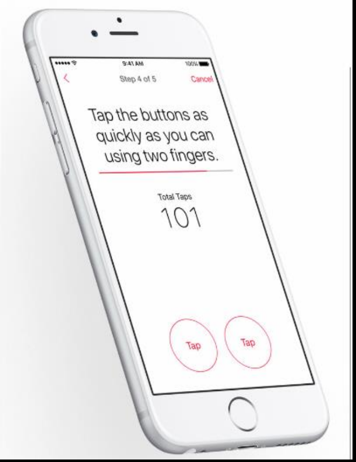


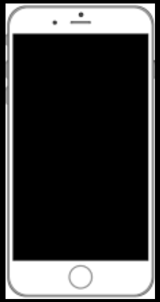
?	3	9	?	3	6

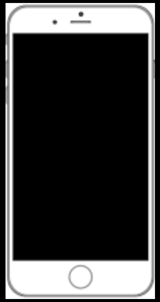
1	1	3	1	1	2
2	2	6	2	2	4
3	3	9	3	3	6

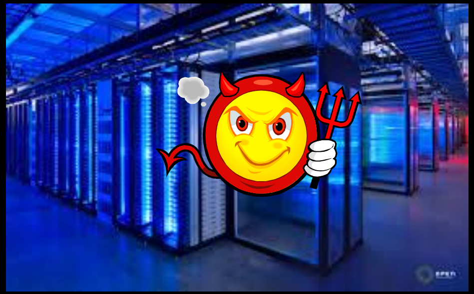
Assumption: Hidden matrix has some structure (e.g., low-rank)

Need for privacy







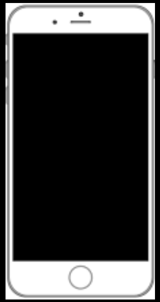


Trust boundary

Model / statistic

Local differential privacy [Warner65,EGS03,DMNS06]

Data sample: d



$\mathcal{A}(d)$

Data sample: d'

$\mathcal{A}(d')$

Requirement: $\mathcal{A}(d)$ and $\mathcal{A}(d')$ should be close in distribution

Local differential privacy [Warner65,EGS03,DMNS06]

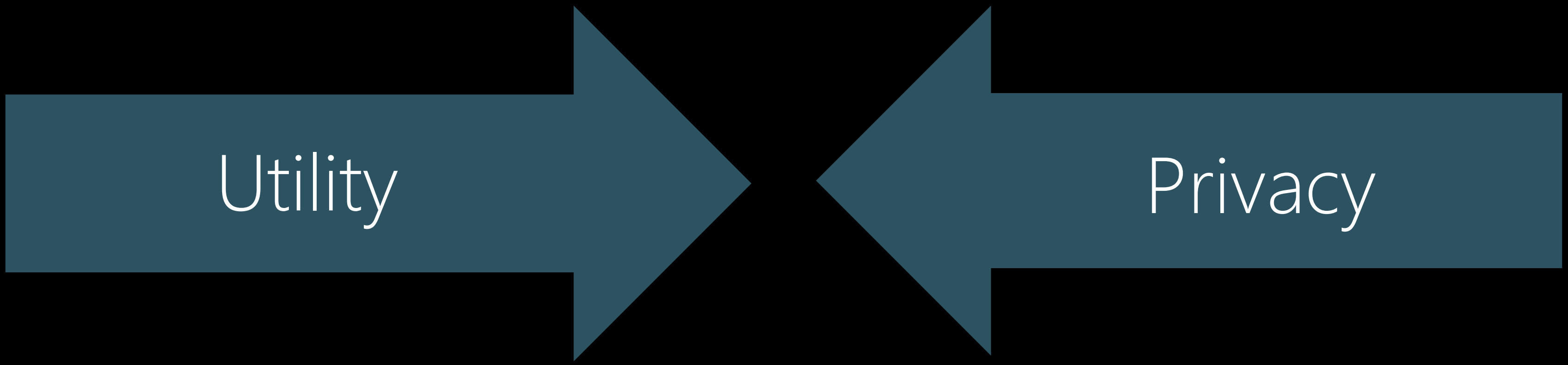
ϵ: Privacy parameter (smaller value implies stronger privacy)

Resilient against arbitrary side information

Provably protects against membership attacks

Challenge: Balancing trade-offs

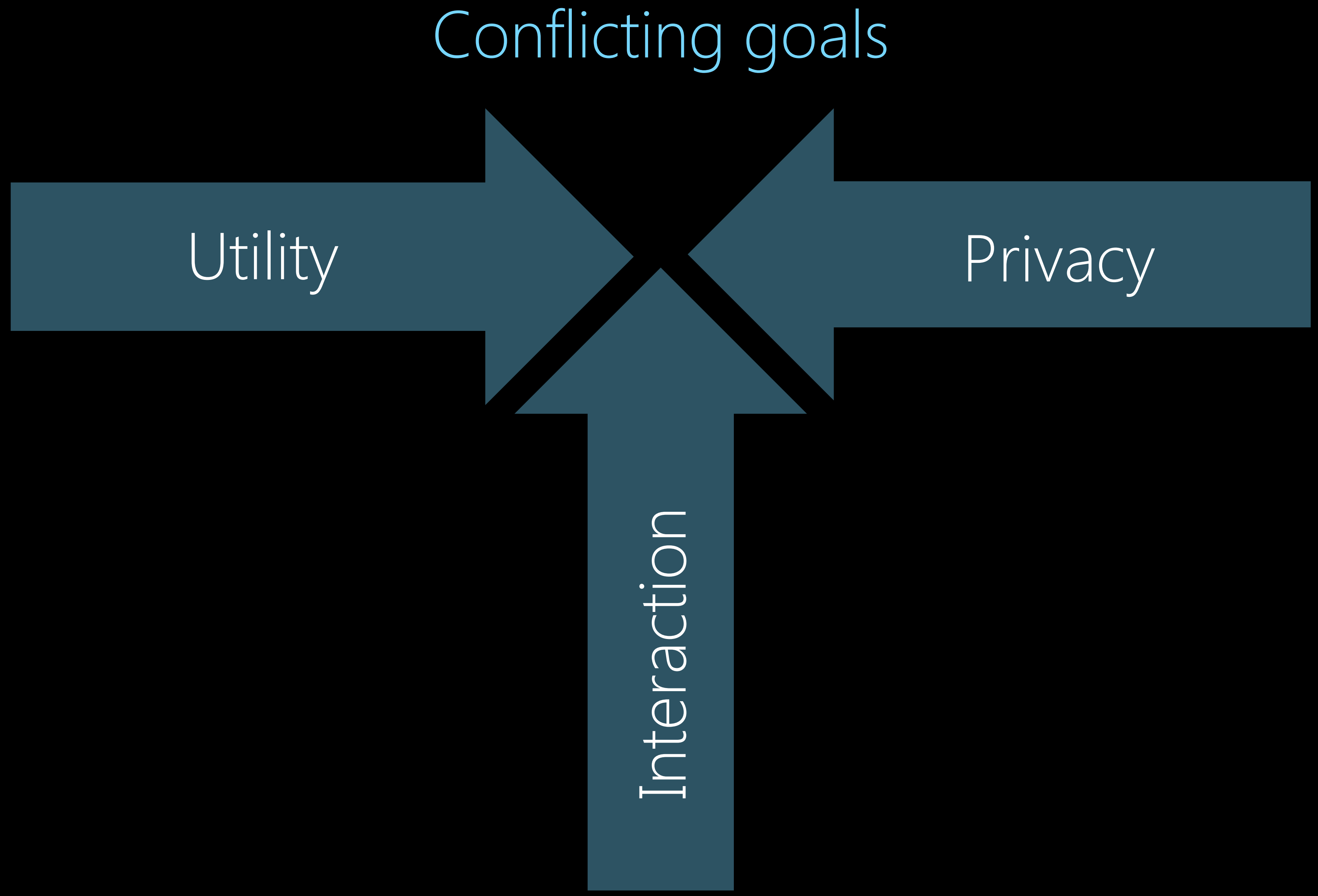
Conflicting goals



Balancing the tradeoff is hard:

- AOL fiasco: *CNBC 101 dumbest moments in business*
- Netflix attack [NS08], Facebook attack [Korolova11], …

This talk



Distributed Private Machine Learning

1. Learning from private data

2. Private distributed model selection

3. Private on-device learning

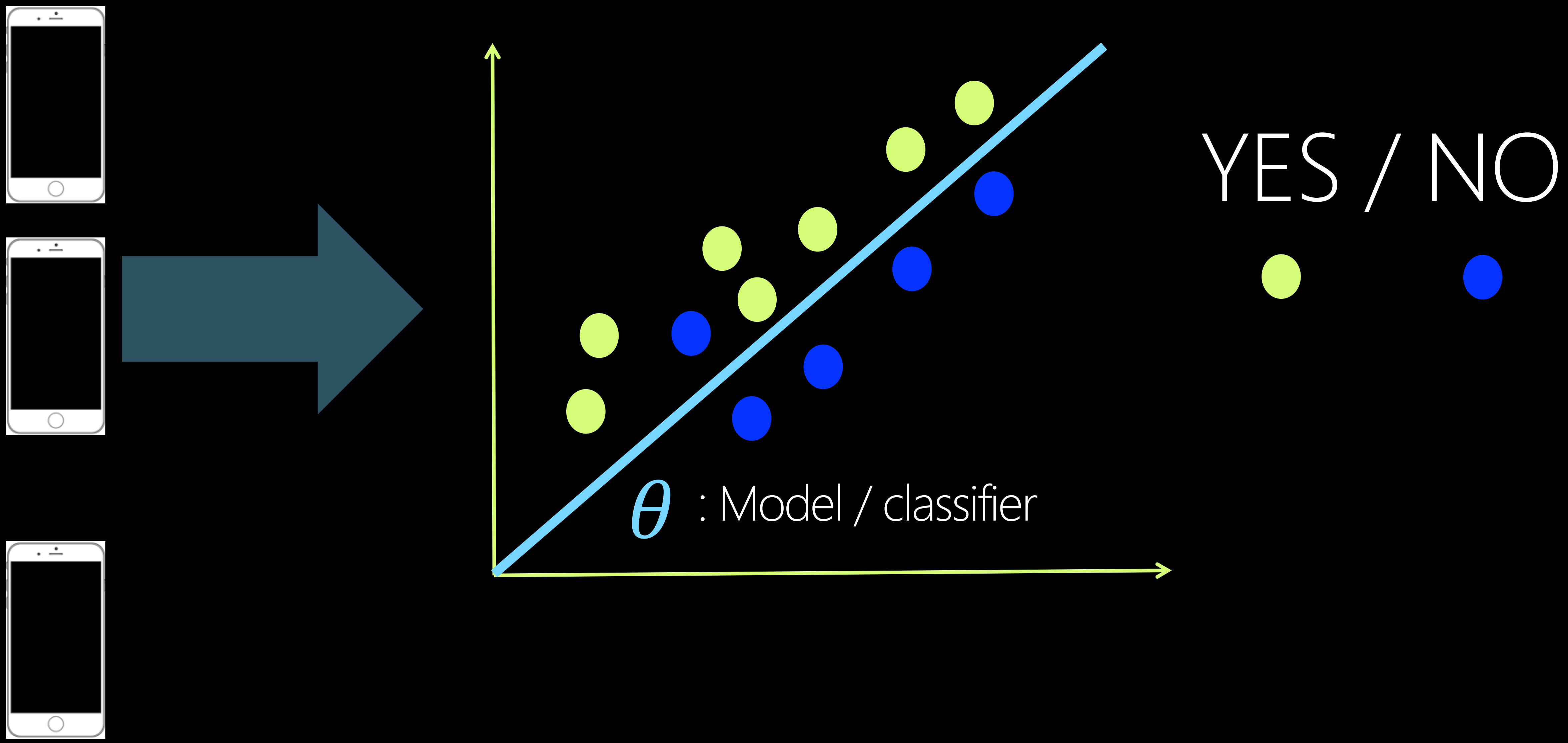
Distributed Private Machine Learning

1. Learning from private data
2. Private distributed model selection
3. Private on-device learning

Private distributed model seleciton

Learning from private data

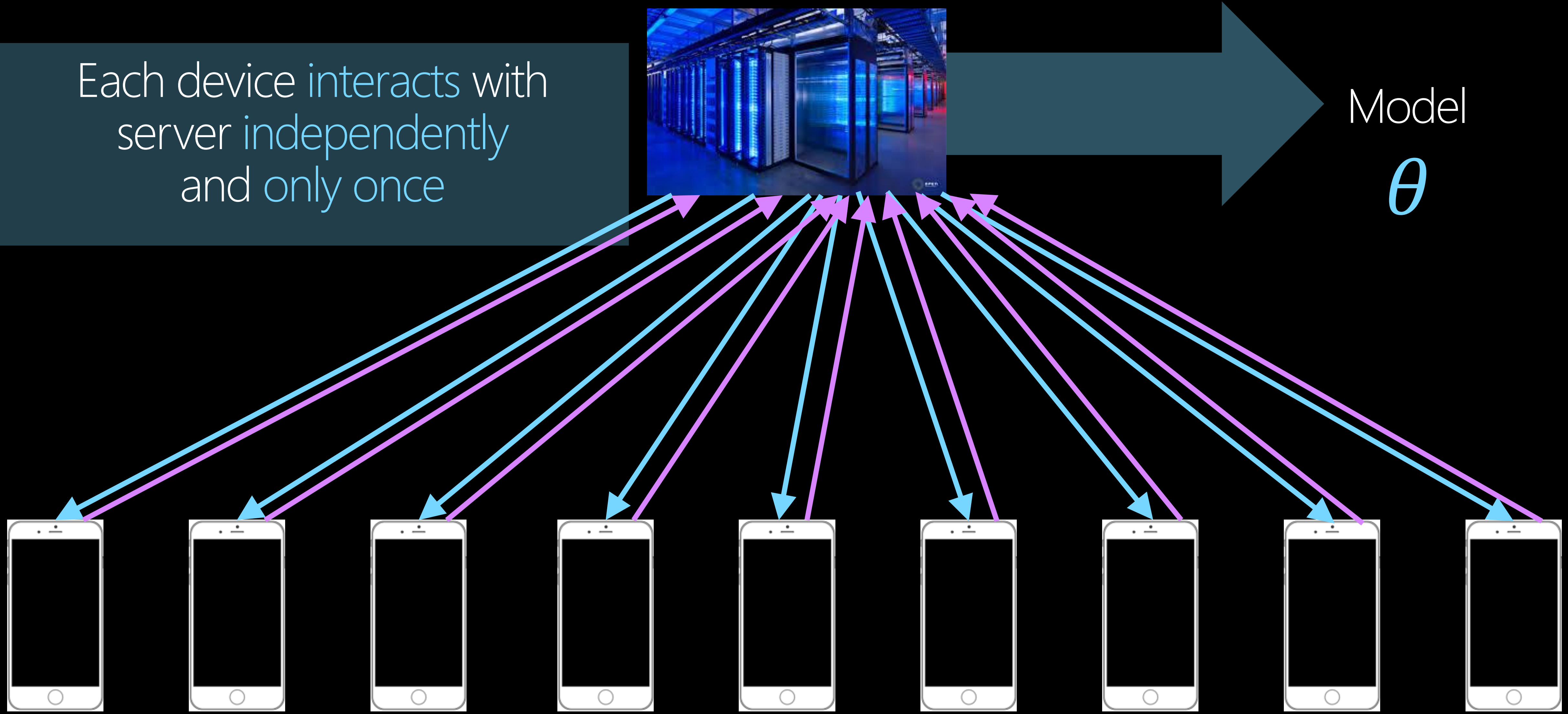
Predict if a person has Parkinson's disease



Towards engineering distributed learning systems

Ideal scenario: Complete parallelism

Each device interacts with server independently and only once



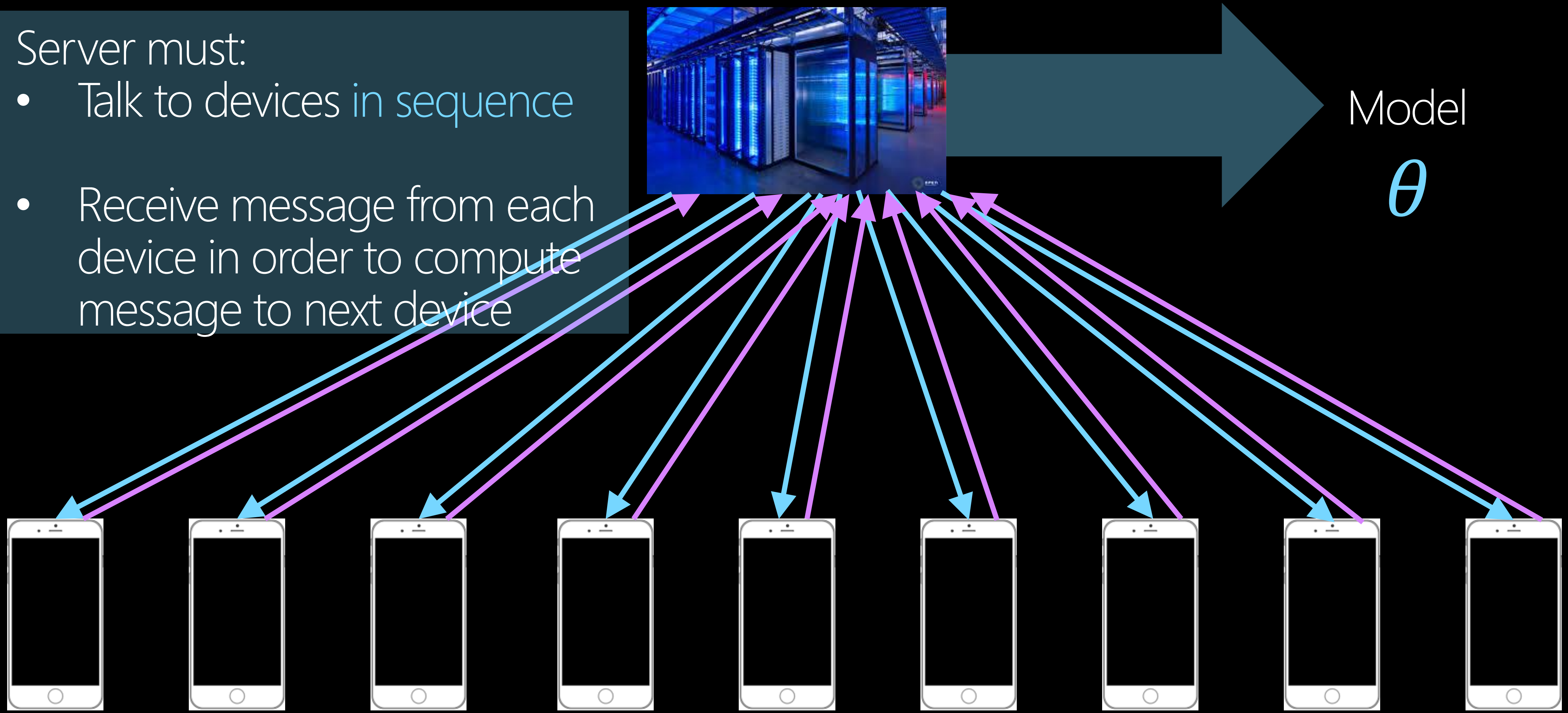
Model θ

Towards engineering distributed learning systems

State of the art [DJW'13]: Completely adaptive interaction

Server must:

- Talk to devices in sequence
- Receive message from each device in order to compute message to next device



This talk [Smith T. Upadhyay' 17]

Distributed private learning with local differential privacy

New algorithms that use little or no adaptive interaction

Lower bound: Cannot get accurate, general algorithms that use no adaptive interaction

Previous work

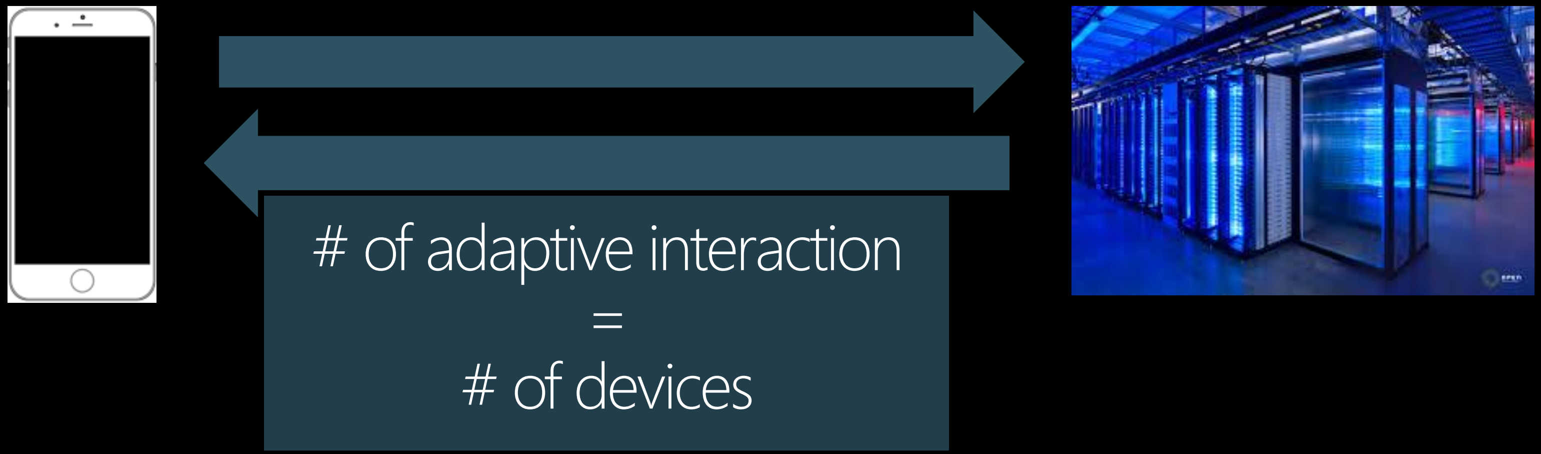
Distributed private learning with local differential privacy

Kasivishwanathan *et al. 2008*

Introduced the problem of local private learning

Duchi *et al. 2013*

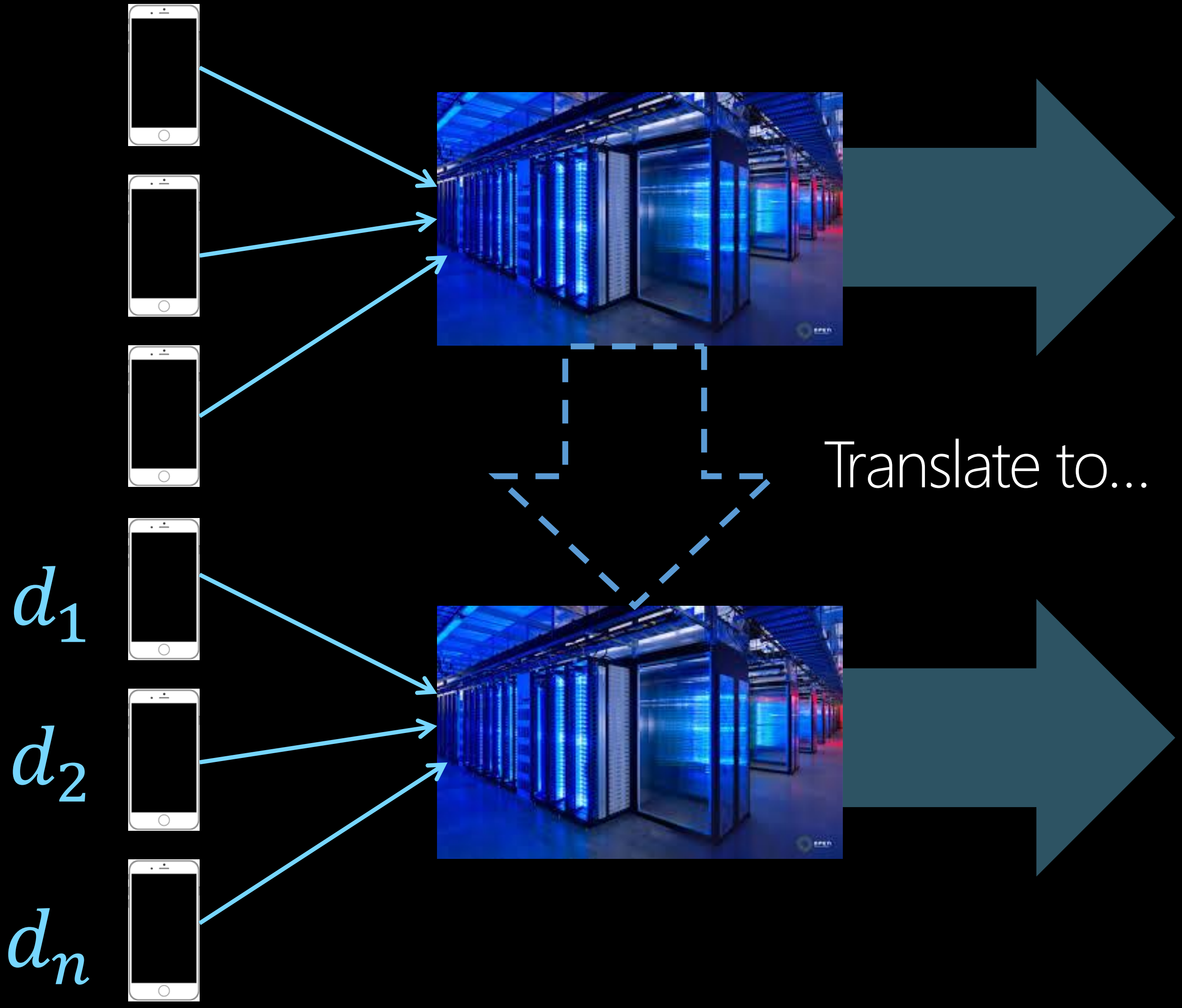
Tight upper and lower bounds on accuracy



Talks to each device only once

Key New Results

Single parameter learning: Minimal error with full parallelism



Diff. private model

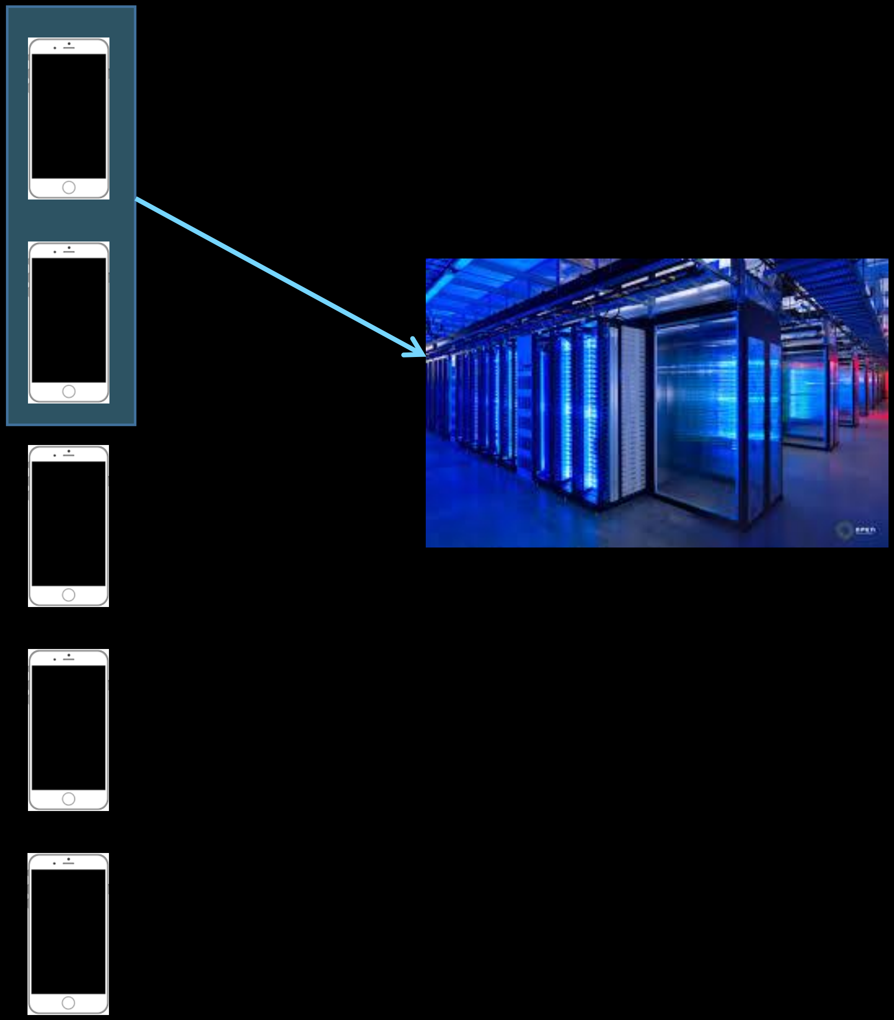
$\theta_{priv} \in \mathbb{R}$

Diff. private estimate of

Median $\{d_1, d_2, \ldots, d_n\}$

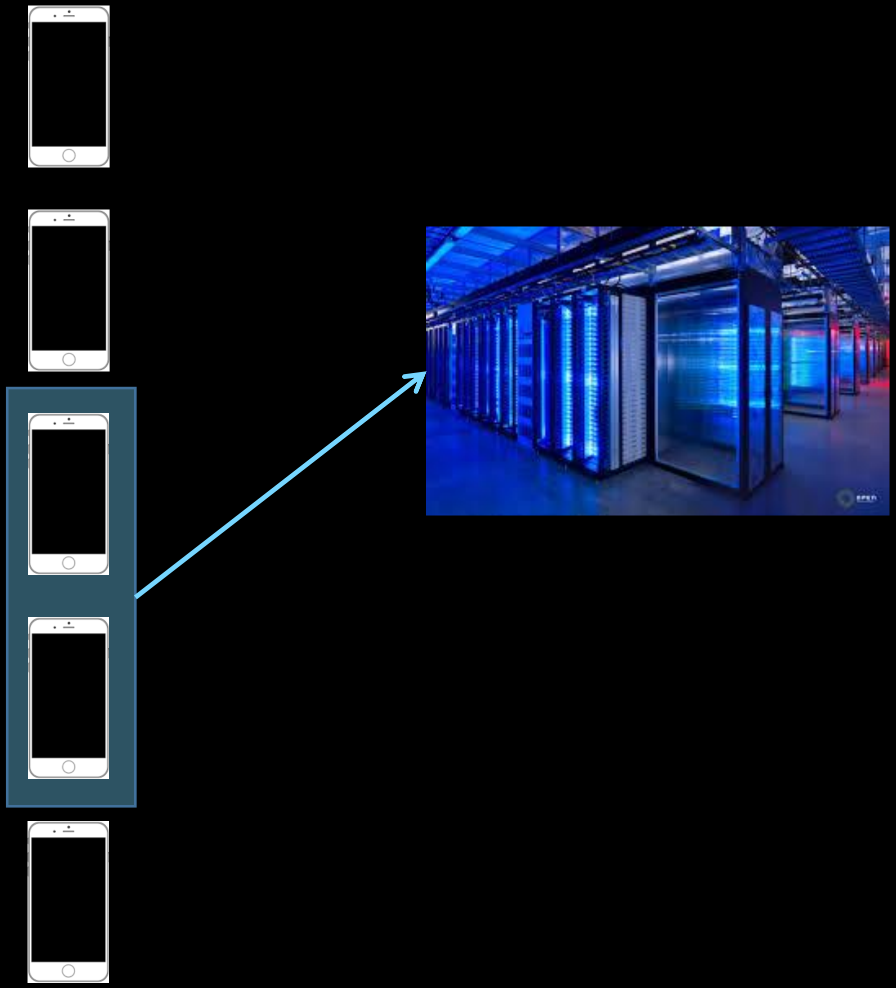
Key New Results

Multi parameter learning: Minimal error with few rounds of adaptivity



Key New Results

Multi parameter learning: Minimal error with few rounds of adaptivity



Key New Results

Multi parameter learning: Minimal error with few rounds of adaptivity

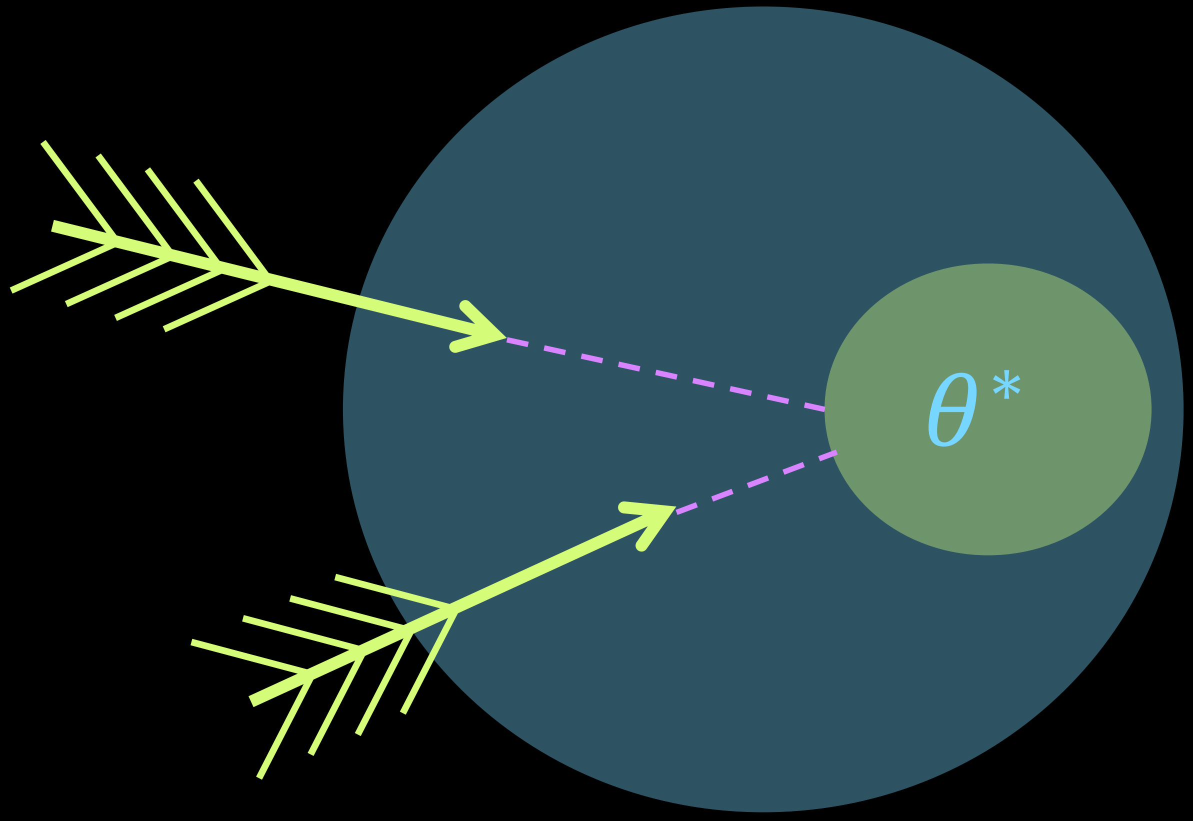
Diff. private model
$\theta_{priv} \in \mathbb{R}^p$

Exponential improvement in the rounds of adaptivity

Still interact with one device only once

Key New Results

Lower bound: Minimal error needs few rounds of adaptivity



θ^*: Best model

Next Steps

Implement the algorithms and evaluate empirically

Deploy the project in practice

Current lower bounds are only for gradient based methods

- Obtaining non-adaptive algorithms will analyzing non-gradient based methods

Distributed Private Machine Learning

1. Learning from private data

2. Private distributed model selection

3. Private on-device learning

Distributed Private Machine Learning

1. Learning from private data
2. Private distributed model selection
3. Private on-device learning

Private on-device learning

On-device learning with sensitive data

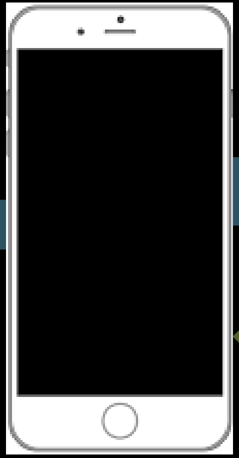
Privacy preserving personalization

Health analytics

Language models

Collaborative filtering

Locally learned models



Model estimates via
differentially private channel

Receive model updates

Local learning

On-device learning with sensitive data

Privacy preserving personalization

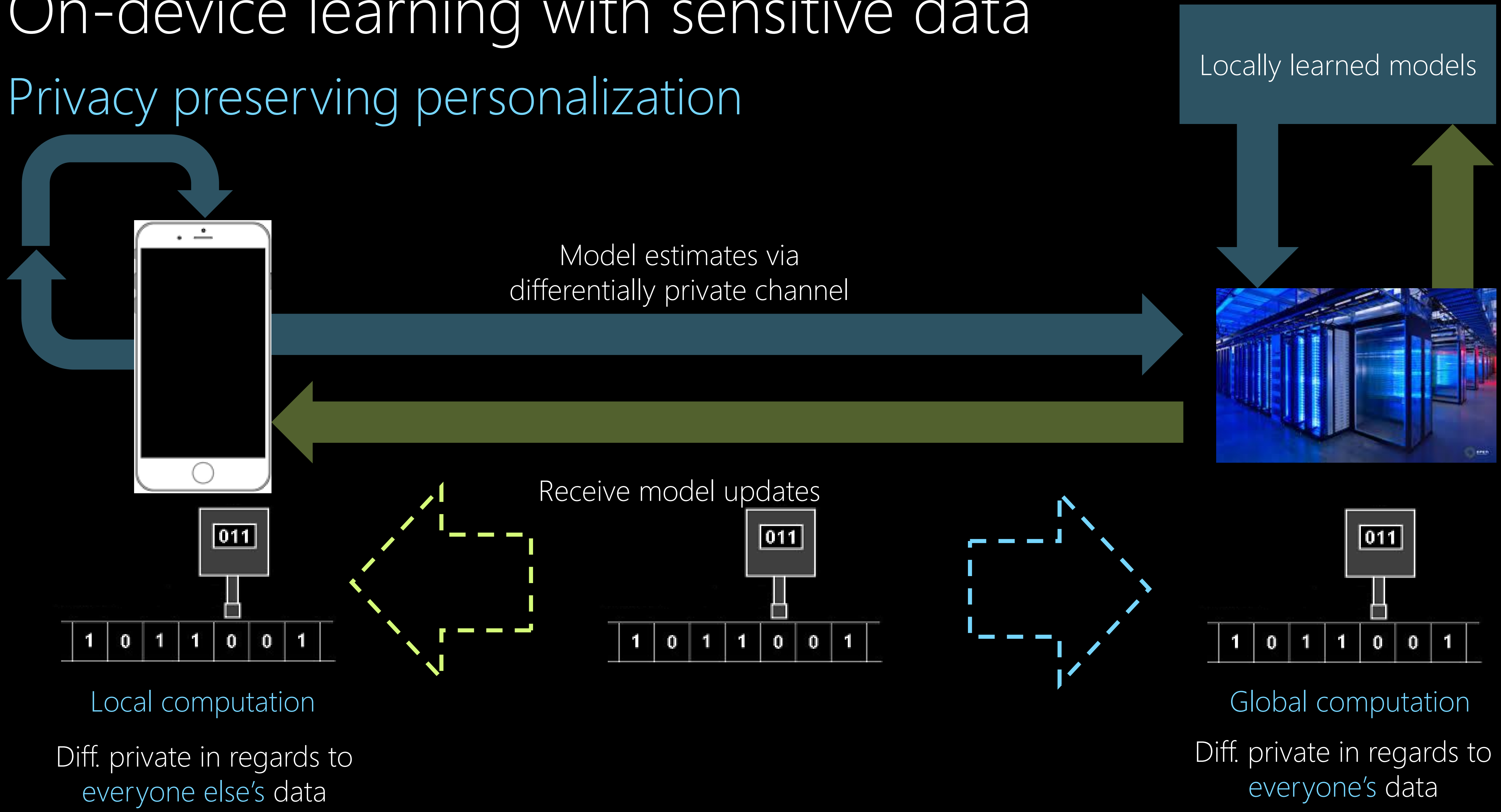
Health analytics

Language models

Collaborative filtering

On-device learning with sensitive data

Privacy preserving personalization



Differentially private on-device learning

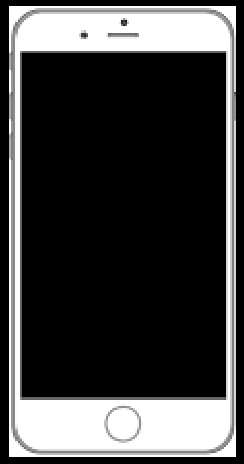
New results and future direction

Collaborative filtering [Jain T. Thakkar]

First algorithm with formal error guarantee

- Global component: Error covariance

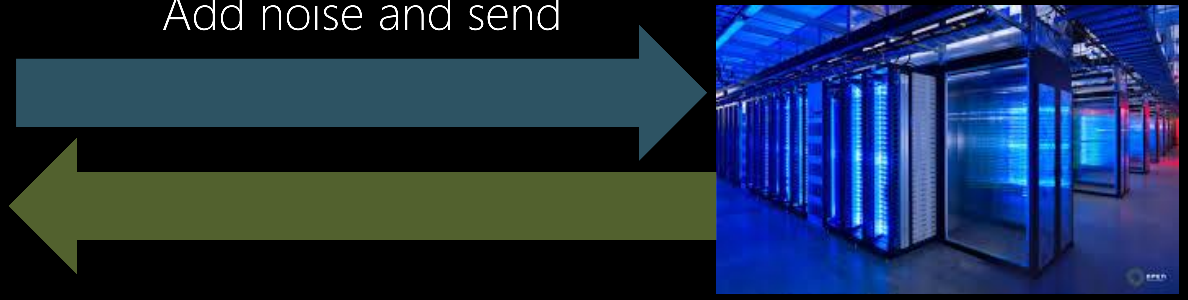
1	1	3	1	1	2
2	2	6	2	2	4
3	3	9	3	3	6



0
0
4
0
0
-1

0	0	4	0	0	-1

Error covariance



Differentially private on-device learning

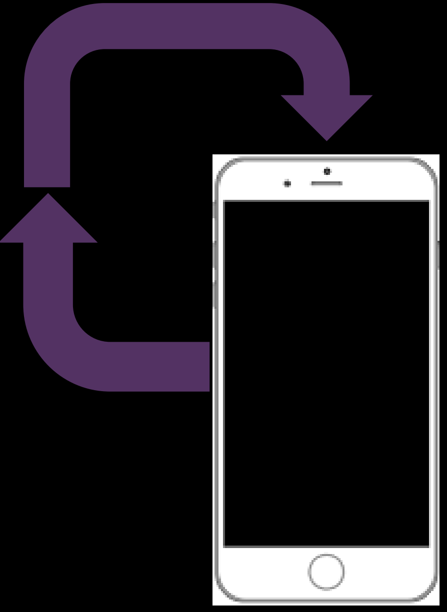
New results and future direction

Collaborative filtering [Jain T. Thakkar]

First algorithm with formal error guarantee

- Global component: Error covariance
- Local component: Compute the prediction

1	1	3	1	1	2
2	2	6	2	2	4
3	3	9	3	3	6



[HR12] Hints at trivial error if predictions are public

Next step

Improve on-device machine learning by harnessing global computation

Distributed Private Machine Learning

1. Learning from private data
2. Private distributed model selection
3. Private on-device learning

Conflicting goals