Practical privacy-preserving machine learning in Python

Catherine Nelson

Concur Labs @ SAP Concur

@DrCatNelson

#PyCon2020

Introducing myself

- Senior Data Scientist at Concur Labs
- Seattle PyLadies co-organizer
- O'Reilly Author
- Google Developer Expert in ML

Why privacy?

VB VentureBeat

Apple and Google weigh privacy concerns as EU demands coronavirus apps not track location data

Apps to help contain the spread of the novel coronavirus should not collect user location data, the European Commission said on Thursday.

3 days ago

WSJ 華爾街日報中文版

How Coronavirus Is Eroding Privacy

Governments are using technology to track and monitor individuals in an effort to slow the pandemic. Privacy advocates are wary, concerned ...

5 days ago

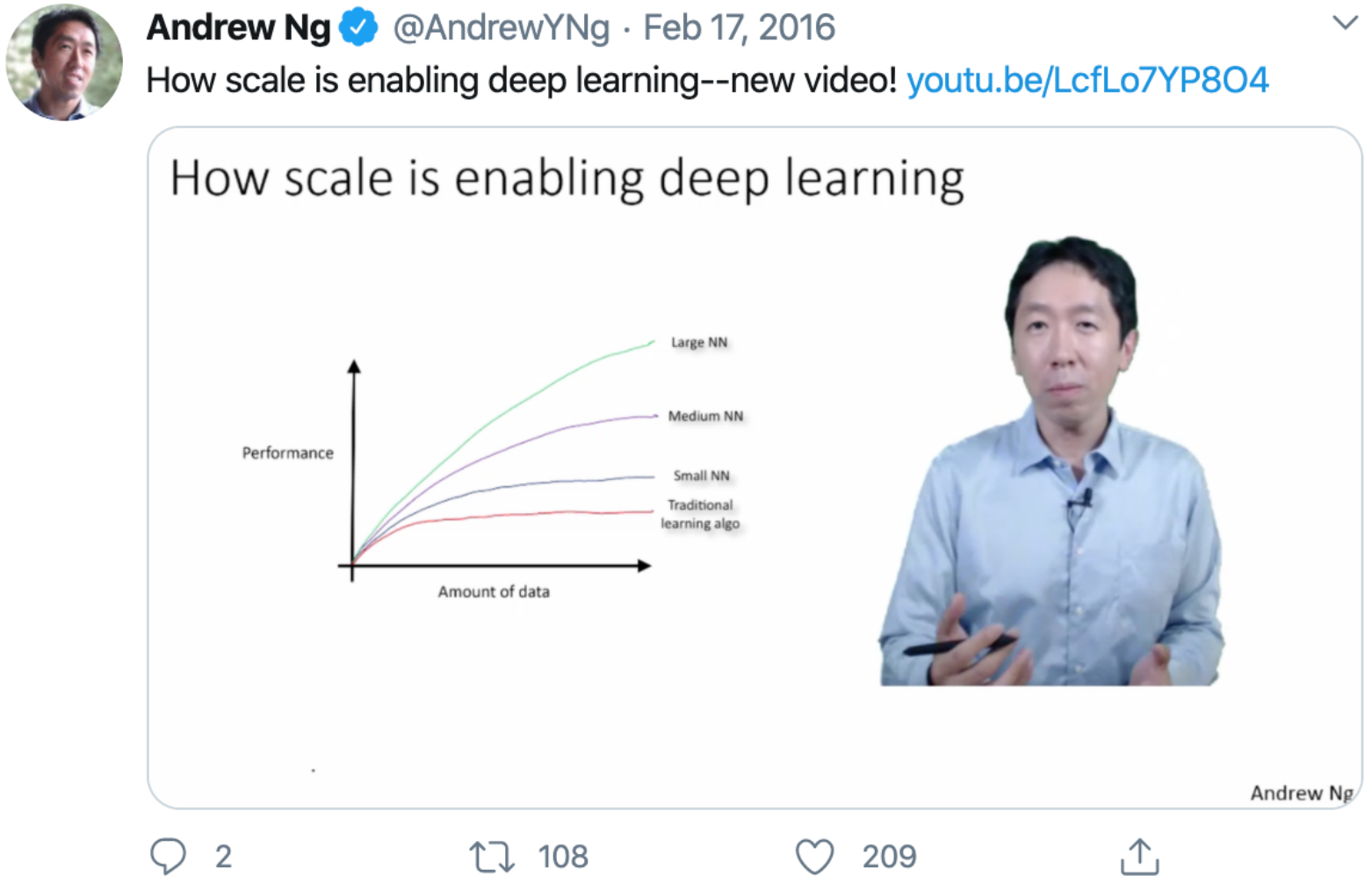
The Verge

Zoom faces a privacy and security backlash as it surges in popularity

Zoom risks becoming the victim of its own success as it faces a privacy and security backlash. A growing number of concerns have been raised ...

3 weeks ago

Machine learning is hungry for data



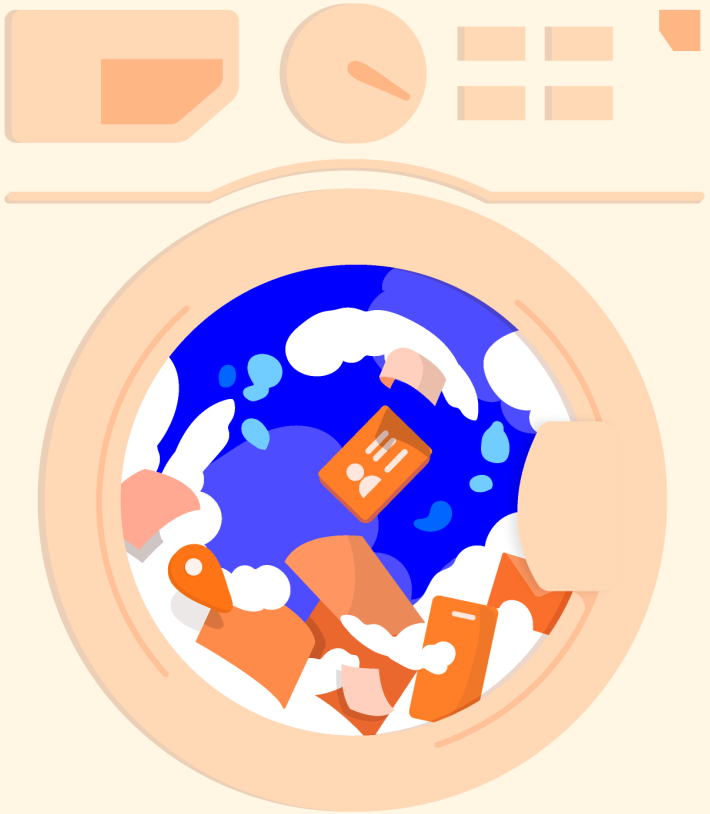
What data should we worry about?

- Private data:
  - Personally identifiable information e.g. name, address, email…
  - Quasi-identifiers
  - User input data
- Sensitive data
  - Health-related personal data
  - Proprietary data

The simplest way to keep data private

…don't collect it in the first place

Wash away your personal data



https://blog.concurlabs.com/wash-away-your-personal-data-b8faa262c755

Wash away your personal data

Lyft Thanks for riding with Joel August 31, 2018 at 2:53 PM Ride Details Lyft fare (17.39mi, 32m 42s) $44.67 Tip $5.00 Visa $49.67 This and every ride is carbon neutral Learn more Pickup Seatac Airport Acrd, SeaTac, WA Drop-off 509 Dewey Pl E, Seattle, WA Help Center Receipt #1174449750415856646 Map data © OpenStreetMap contributors Facebook Twitter Instagram Work at Lyft Become a Driver © Lyft 2018 185 Berry Street, Suite 5000 San Francisco, CA 94107

Lyft Thanks for riding with <NAME> August 31, :Details Lyft fare (17.39mi, 32m 42s) $44.67 Tip $5.00 Visa $49.67 This and every ride is carbon neutral Learn more Pickup <ADDRESS>, SeaTac, WA Drop-off <ADDRESS>, Seattle, WA Help Center Receipt #© OpenStreetMap contributors Facebook Twitter Instagram Work at Lyft Become a Driver © Lyft, <ADDRESS> CA 94107

But without collecting the data

we can't build the model

we don't know if our model is fair

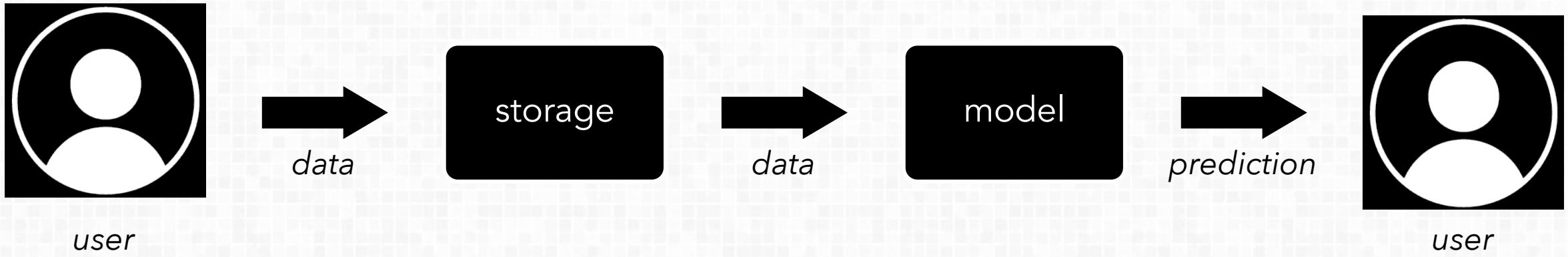
It seems like privacy and ML are opposed

But actually, privacy and ML can have the same goals

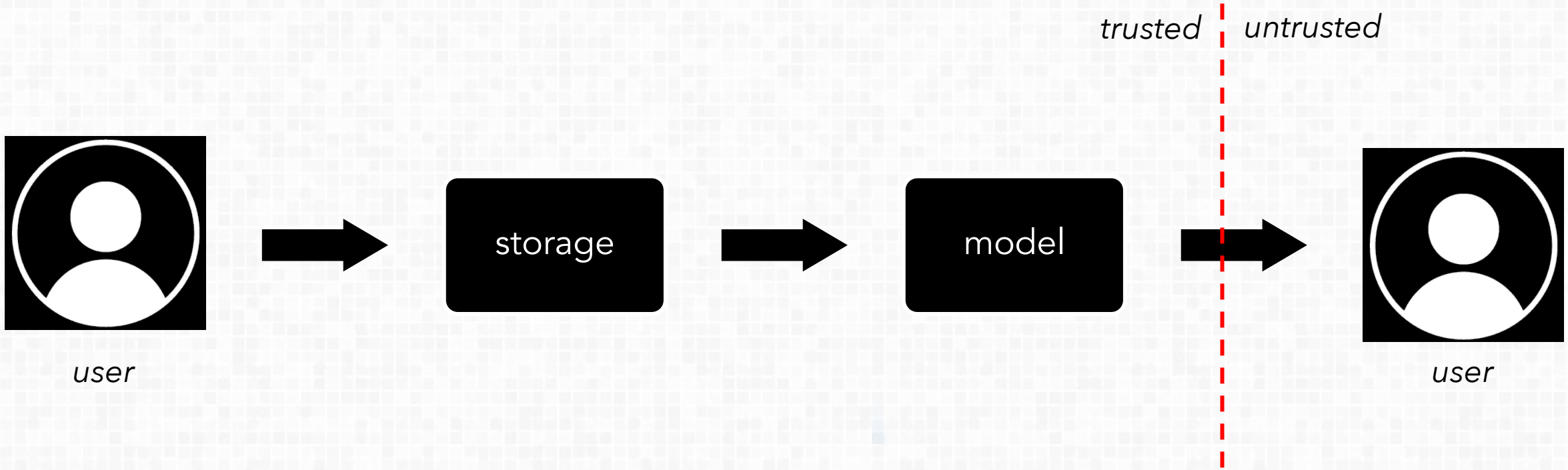
And there are privacy-preserving ML technologies to help

But how do I do this in Python?

Who do you trust?


storage
model
data
data
prediction
user
user

Differential privacy

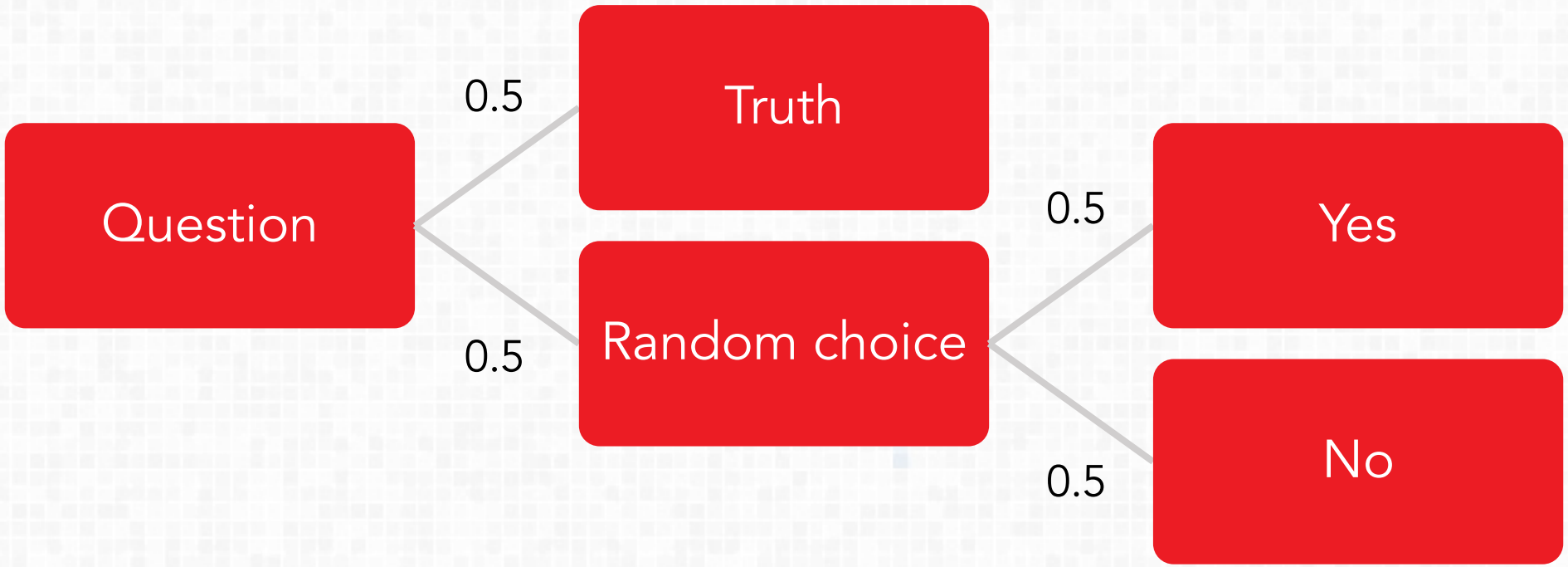


Differential privacy

A formalization of the idea that

a query should not reveal whether a person is in a dataset

Differential privacy



An example of randomized response

TensorFlow Privacy

430

Introducing TensorFlow Privacy: Learning with Differential Privacy for Training Data

TensorFlow Follow

Mar 6 · 7 min read

Posted by Carey Radebaugh (Product Manager) and Ulfar Erlingsson (Research Scientist)

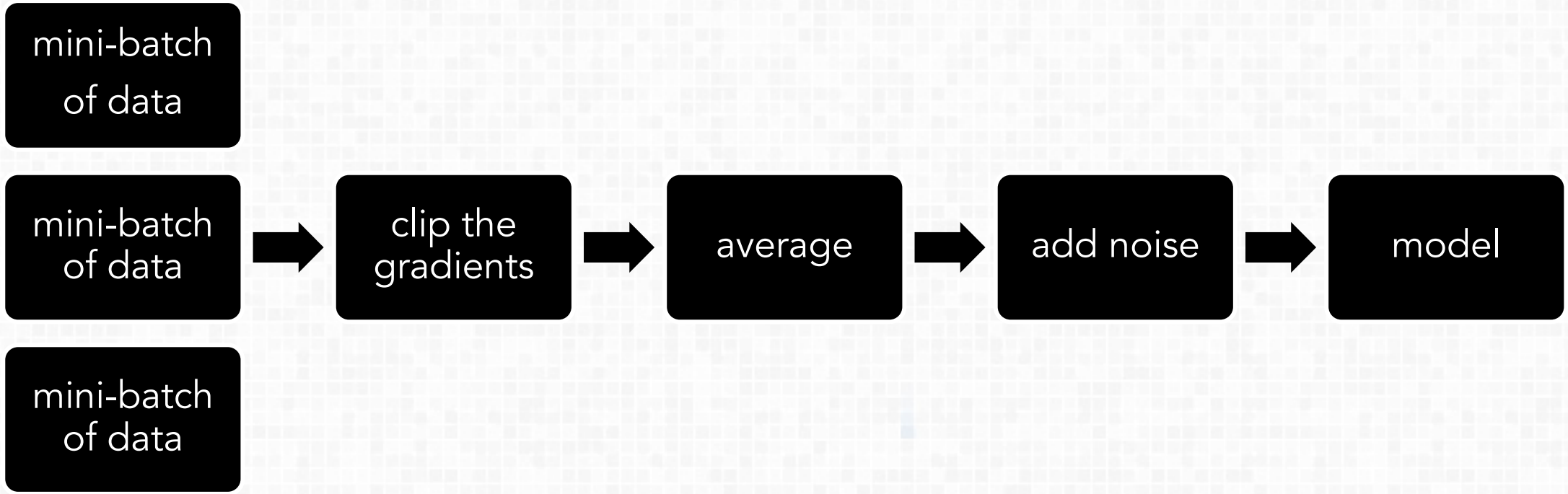
Today, we're excited to announce TensorFlow Privacy (GitHub), an open source library that makes it easier not only for developers to train machine-learning models with privacy, but also for researchers to advance the state of the art in machine learning with strong privacy guarantees.

https://medium.com/tensorflow/introducing-tensorflow-privacy-learning-with-differential-privacy-for-training-data-b143c5e801b6

```
import tensorflow as tf

model = tf.keras.models.Sequential([
   tf.keras.layers.Dense(128, activation='relu’),
   tf.keras.layers.Dense(128, activation='relu’),
   tf.keras.layers.Dense(1, activation='sigmoid’)
   ])
```

```
from tensorflow_privacy.privacy.optimizers.dp_optimizer
import DPGradientDescentGaussianOptimizer

optimizer = DPGradientDescentGaussianOptimizer(
   l2_norm_clip=1.1,
   noise_multiplier=1.3,
   num_microbatches=32,
   learning_rate=0.001)

loss = tf.keras.losses.BinaryCrossentropy(
   from_logits=True,
   reduction=tf.losses.Reduction.NONE)
```

```
from tensorflow_privacy.privacy.optimizers.dp_optimizer
import DPGradientDescentGaussianOptimizer

optimizer = DPGradientDescentGaussianOptimizer(
   l2_norm_clip=1.1,
   noise_multiplier=1.3,
   num_microbatches=32,
   learning_rate=0.001)

loss = tf.keras.losses.BinaryCrossentropy(
   from_logits=True,
   reduction=tf.losses.Reduction.NONE)
```


mini-batch of data
mini-batch of data
mini-batch of data
clip the gradients
average
add noise
model

```
model.compile(optimizer=optimizer,
              loss=loss,
              metrics=['accuracy'])

model.fit(X_train,
          y_train,
          epochs= 10,
          validation_data=(X_test, y_test),
          batch_size=32)
```

The epsilon concept

A mathematical guarantee of privacy

e^{ε} is the maximum difference between the outcomes of two transformations

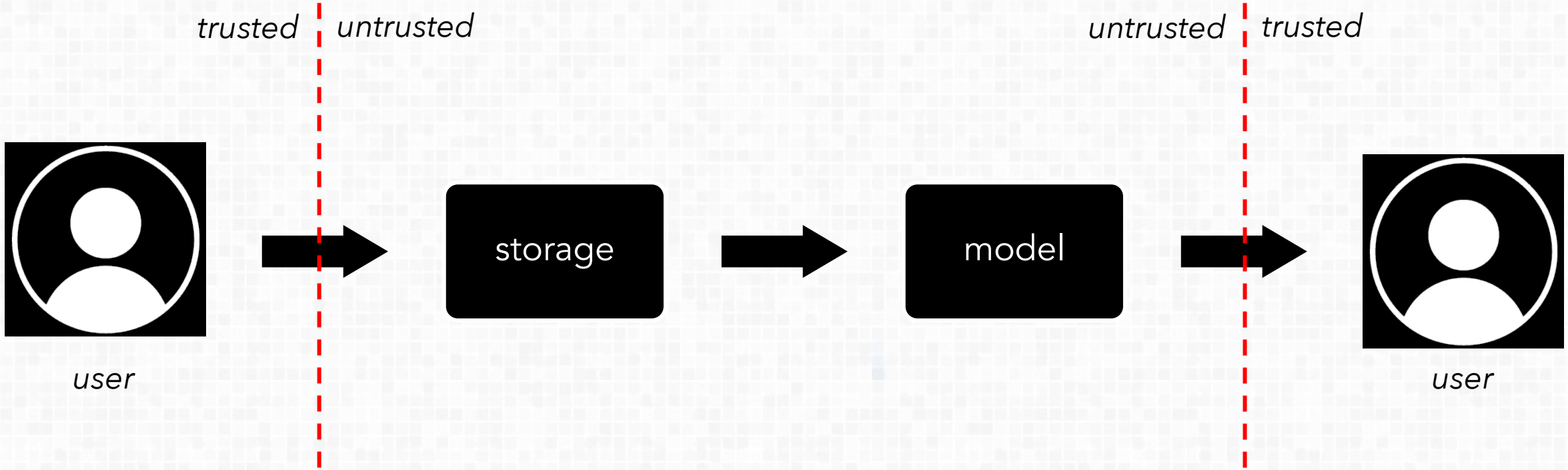
If ε is small, the transformation is more private

```
from tensorflow_privacy.privacy.analysis import
compute_dp_sgd_privacy


compute_dp_sgd_privacy(n=POPULATION_SIZE,
                       batch_size=32,
                       noise_multiplier=1.1,
                       epochs=10,
                       delta=1e-4)
```

Differential privacy

- Good when you don’t want to expose predictions
- But you still need to collect the data in the first place
- Provides mathematical definitions and guarantees of privacy

Encrypted ML



- Handles encryption for you
- Keeps the Keras interface
- Allows sharing of sensitive data

https://tf-encrypted.io/

Encrypted ML

Define a function that provides batches of training data

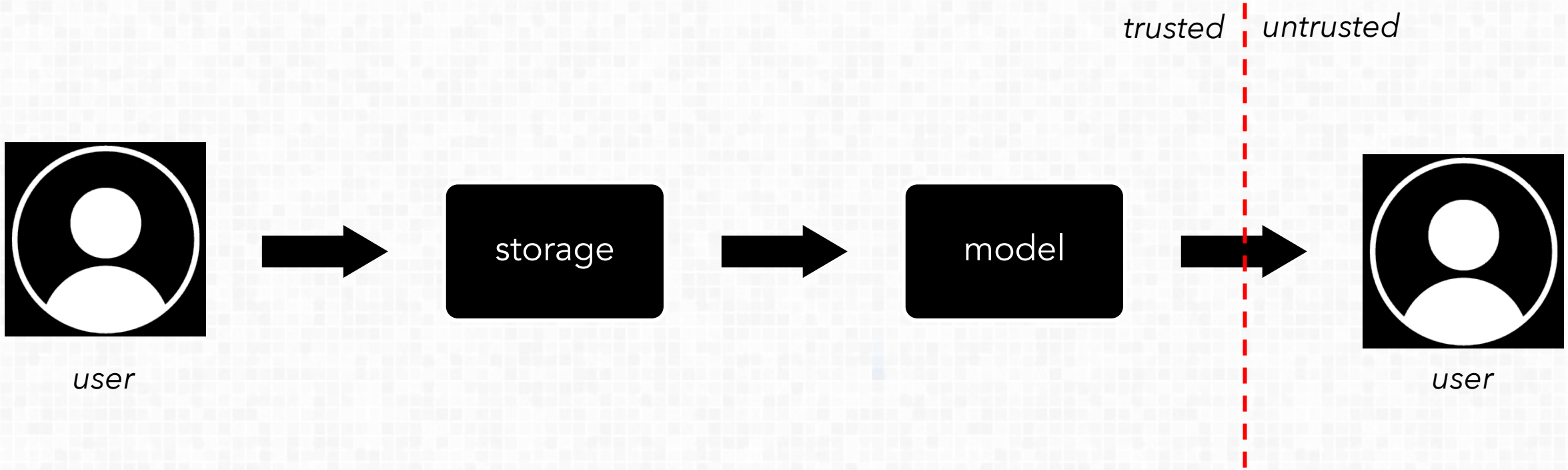
Then simply add the decorator:

```
@tfe.local_computation
```

https://bit.ly/encrypted-training

```
import tf_encrypted as tfe

model = tfe.keras.Sequential()
model.add(tfe.keras.layers.Dense(1,
batch_input_shape=[batch_size, num_features]))
model.add(tfe.keras.layers.Activation('sigmoid'))
```

Encrypt a trained model



Encrypt a trained model

Convert to an encrypted model for serving:

```
tfe_model = tfe.keras.models.clone_model(model)
```

Makes encrypted predictions

https://bit.ly/encrypted-training

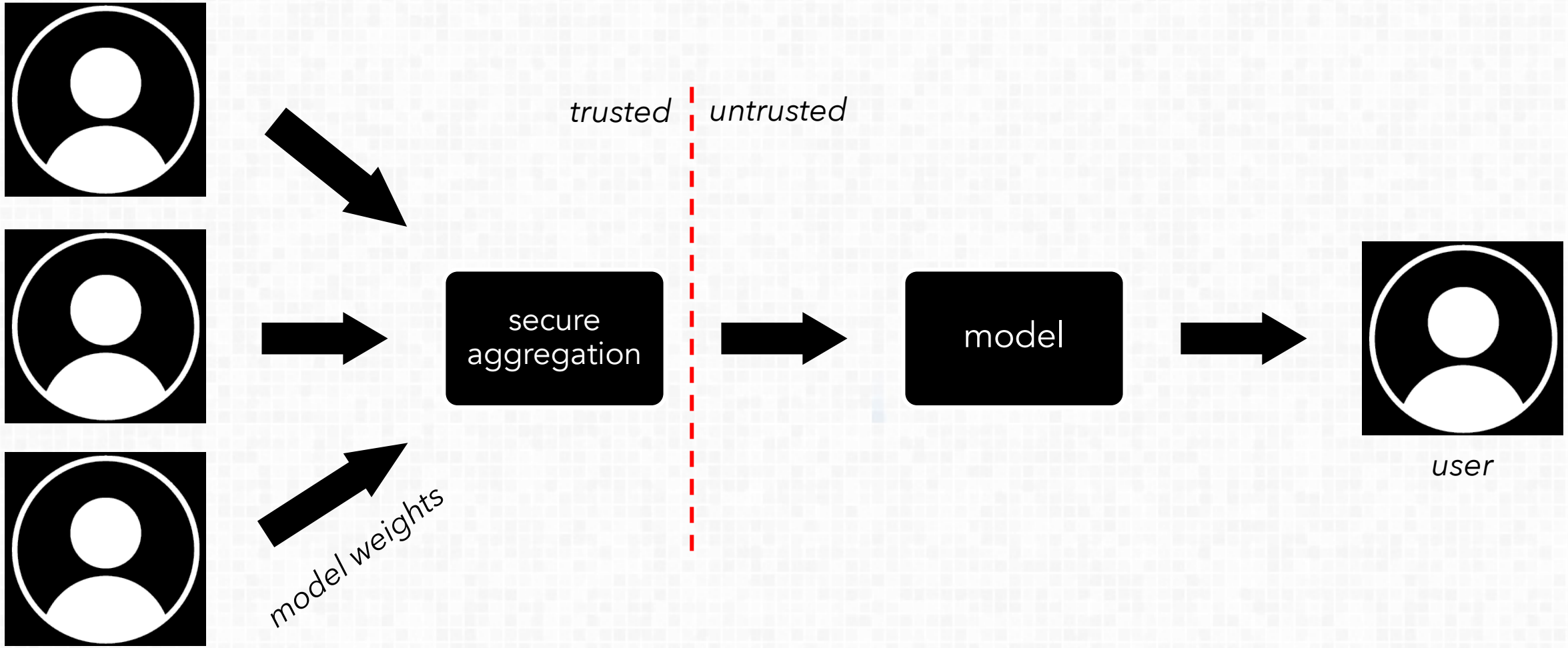
Encrypt a trained model

1. Load and preprocess the data locally on the client.
2. Encrypt the data on the client.
3. Send the encrypted data to the servers.
4. Make a prediction on the encrypted data.
5. Send the encrypted prediction to the client.
6. Decrypt the prediction on the client and show the result to the user.

When to use encrypted ML

- Two options – encrypt the training data, or just the model
- Encrypt the data if the model owner is not trusted
- Encrypt the model if the training data is public but inference is private

Federated learning



Federated learning

PySyft (for PyTorch and TensorFlow)

https://github.com/OpenMined/PySyft

https://github.com/OpenMined/PySyft-TensorFlow

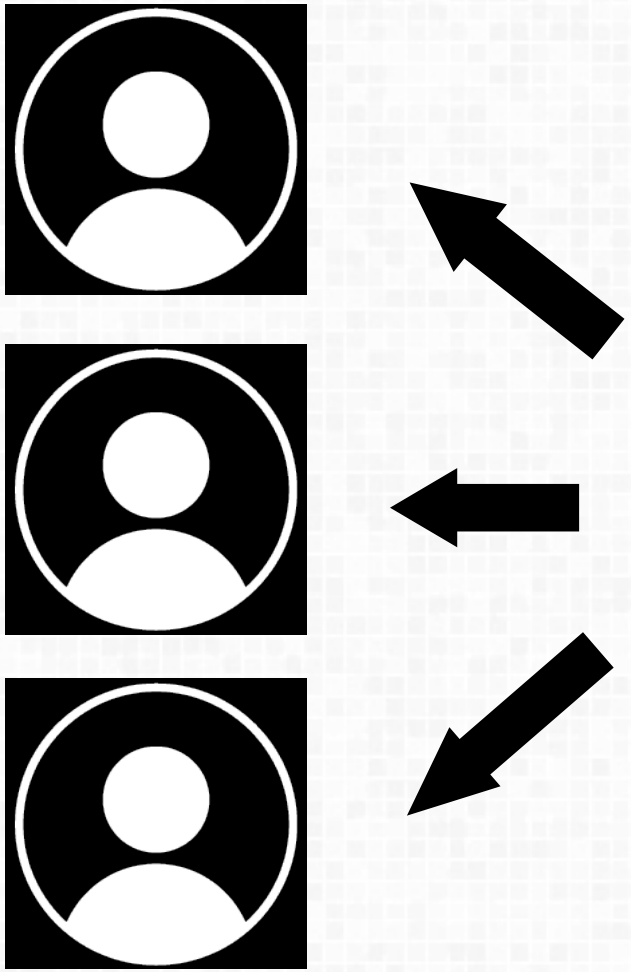
TensorFlow Federated

https://www.tensorflow.org/federated

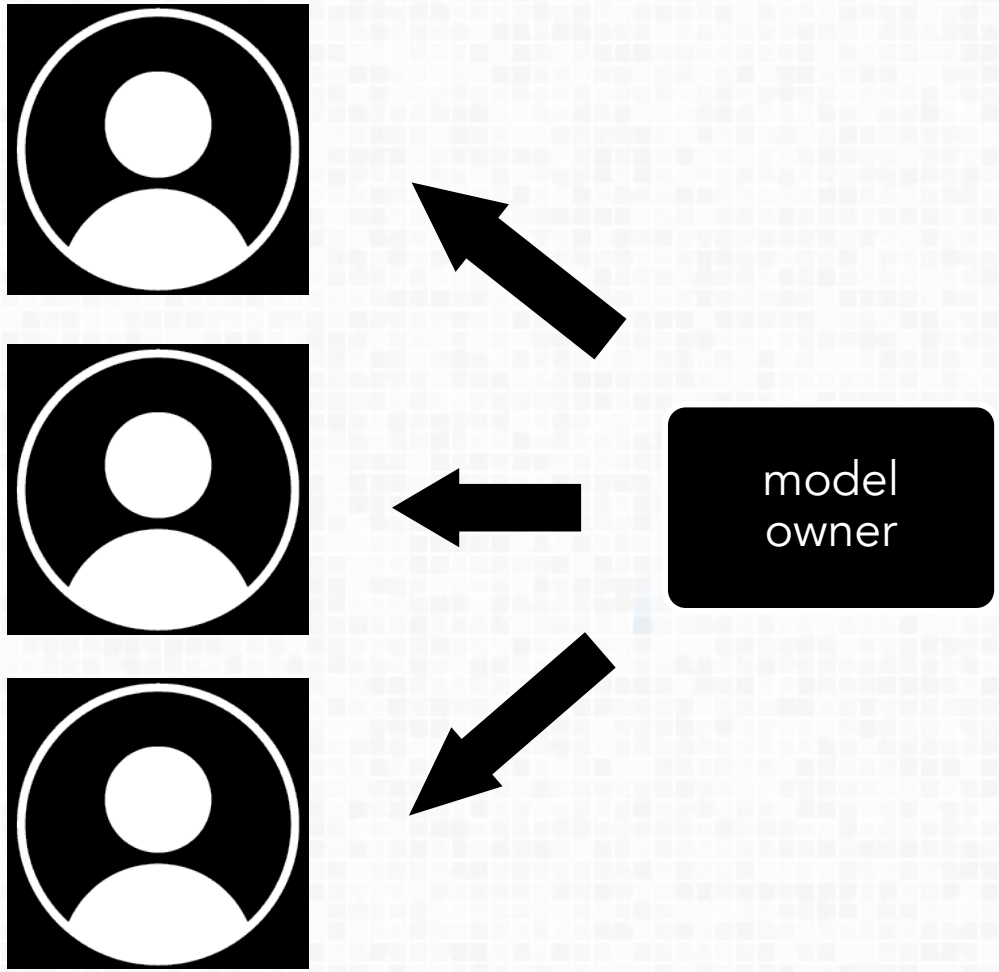
1. Create virtual workers

https://github.com/OpenMined/PySyft/tree/master/examples/tutorials

2. Get pointers to the training data on each worker

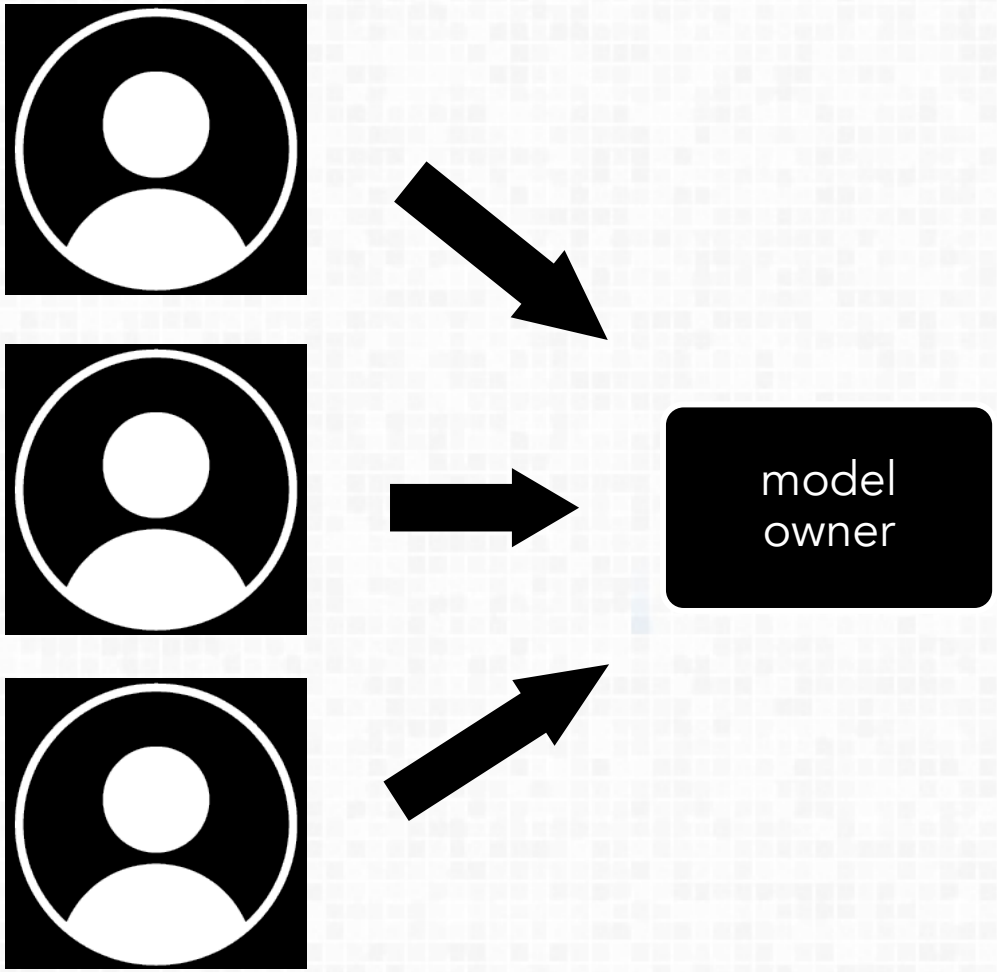


3. Send the model weights to each worker

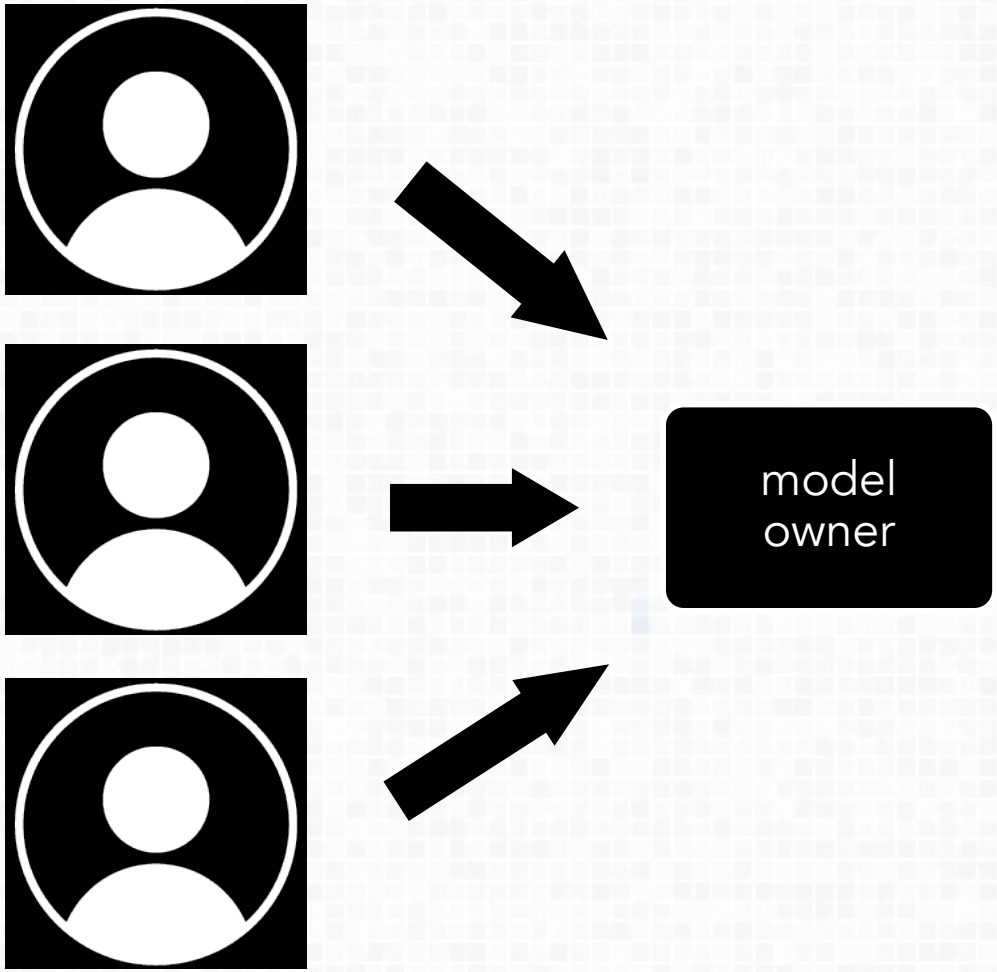


4. Train the model on each worker

5. Send the weights back to the model owner



6. Send the loss back to the model owner



What's missing?

- No privacy yet – we can often infer the original data from the model weights
- Need to add some secure way of averaging the model weights before they reach the model owner
- Scale!

https://github.com/OpenMined/PySyft/tree/master/examples/tutorials

When to use federated learning

- Personalized models
- The data is already decentralized
- The data is sensitive
- No new labels need to be added

Who do you trust?

What to choose?

If you trust the model owner, they can use the raw data but offer private predictions

If you don’t trust the model owner, encrypt the data or keep it on each user’s device

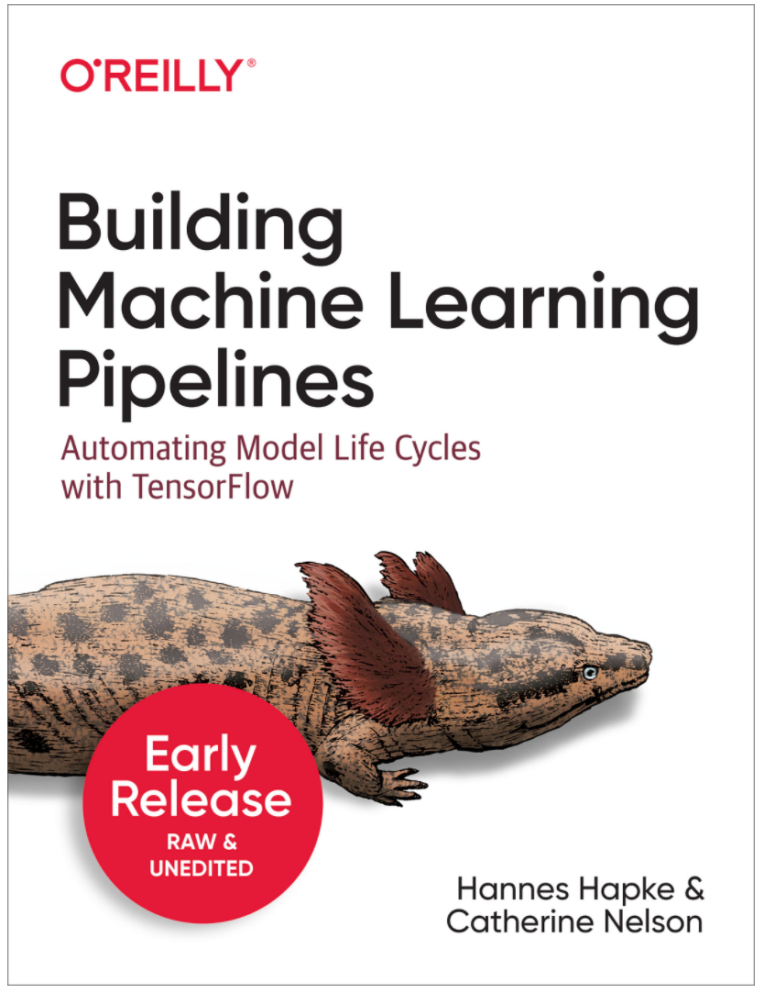
Caveats

There's a cost to privacy: lower accuracy, longer training times, or more complex infrastructure

Don't assume you're providing complete privacy for your users

This won't save you from all ethical issues

Next steps



www.buildingmlpipelines.com

Support the open source projects tf-encrypted, PySyft, TensorFlow Privacy

Questions? @DrCatNelson

thank you

concurlabs.com

#PyCon2020

@DrCatNelson