

Privacy in Federated Learning

presented by Marcel

The dilemma of data aggregation



Use public data

Protect private data

Use public data



Protect private data

What data **should be public?**

What data **should be private?**

**Aggregations
of private data
can be useful.**



The ethics of data collection are tricky.

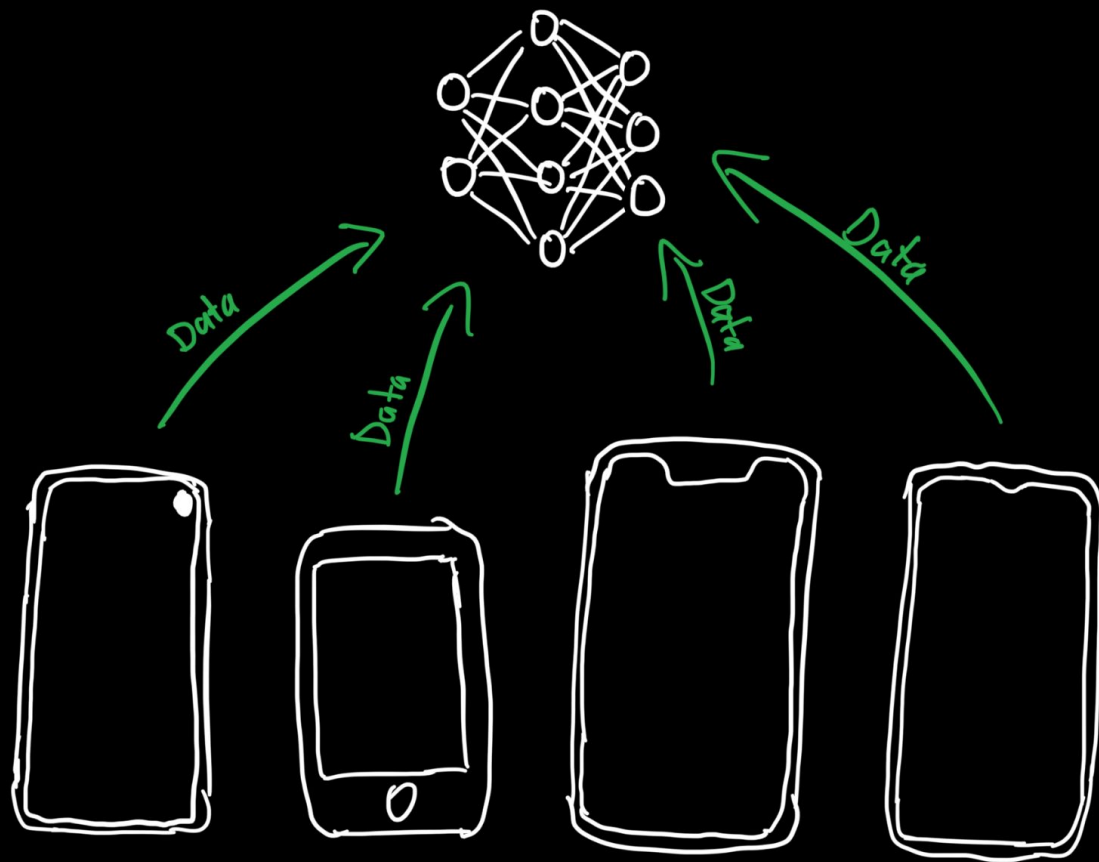
Privacy
for the individual



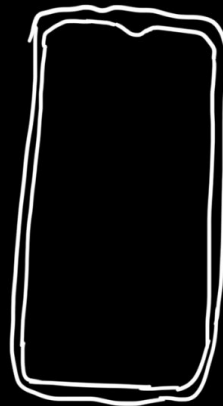
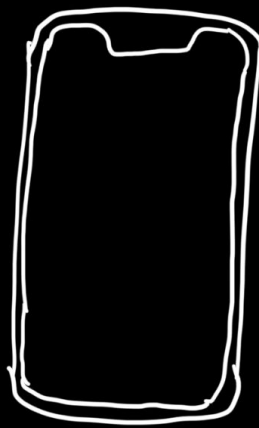
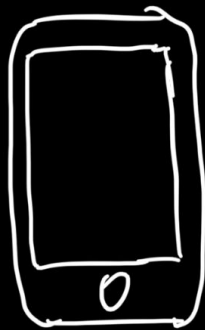
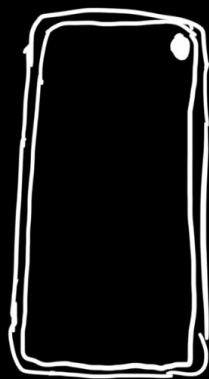
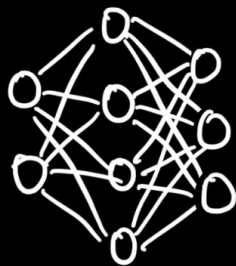
Utility
for everyone

The architecture of typical Machine Learning

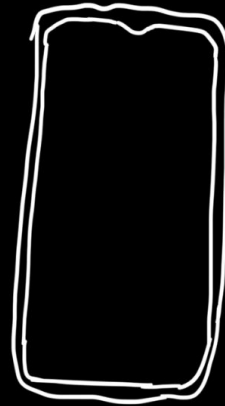
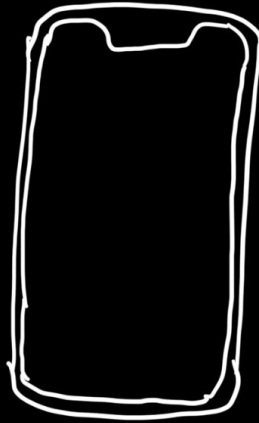
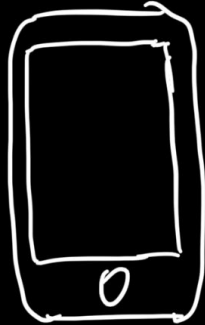
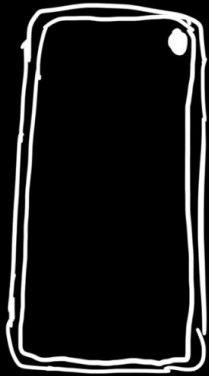
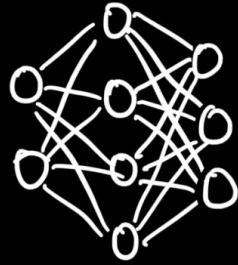


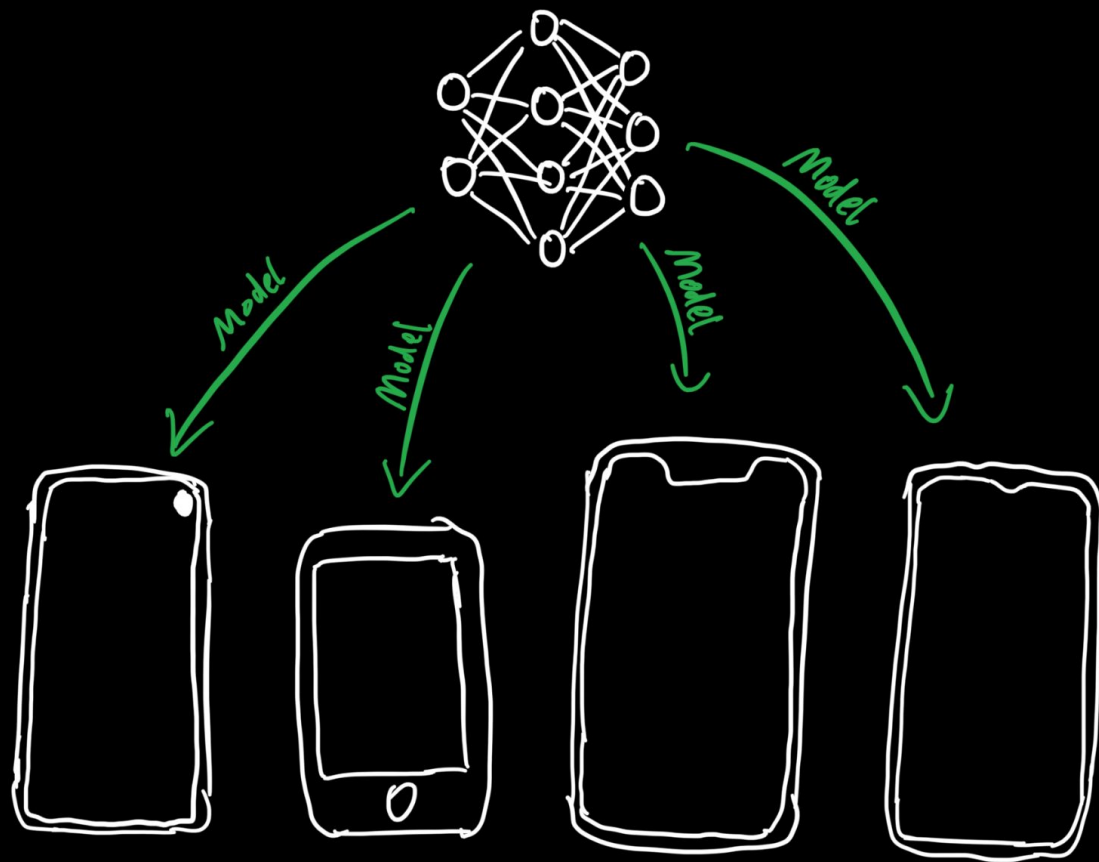


Train



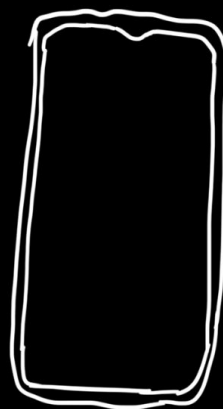
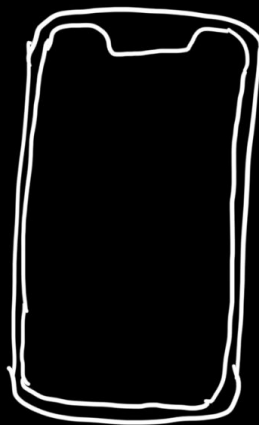
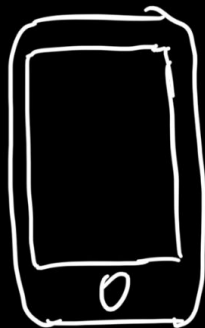
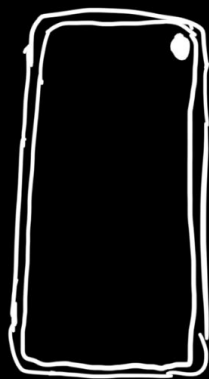
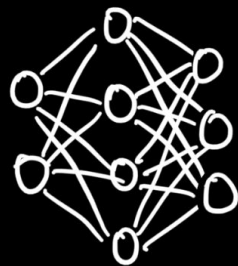
Test

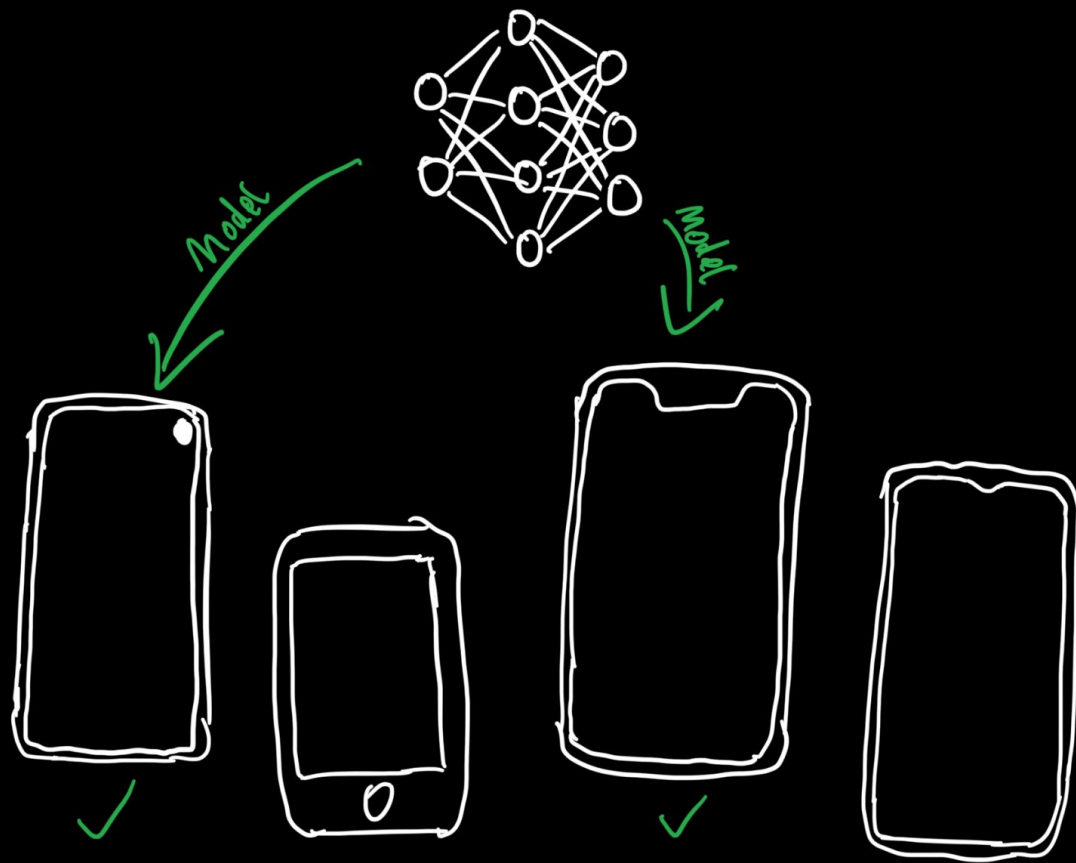


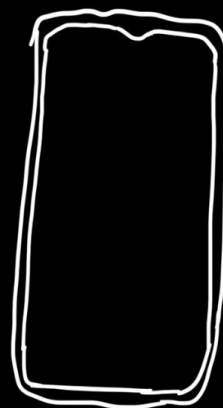
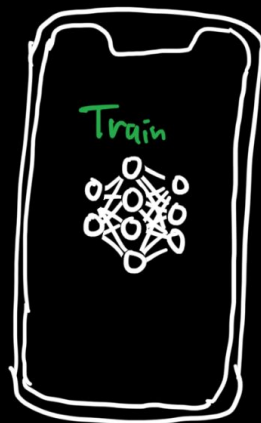
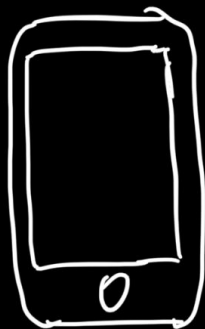
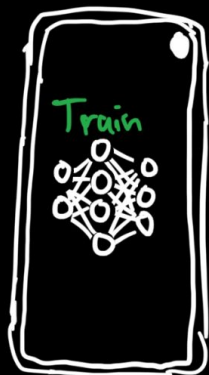
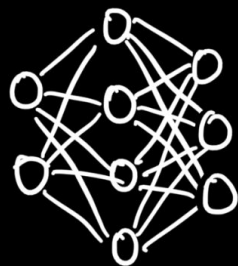


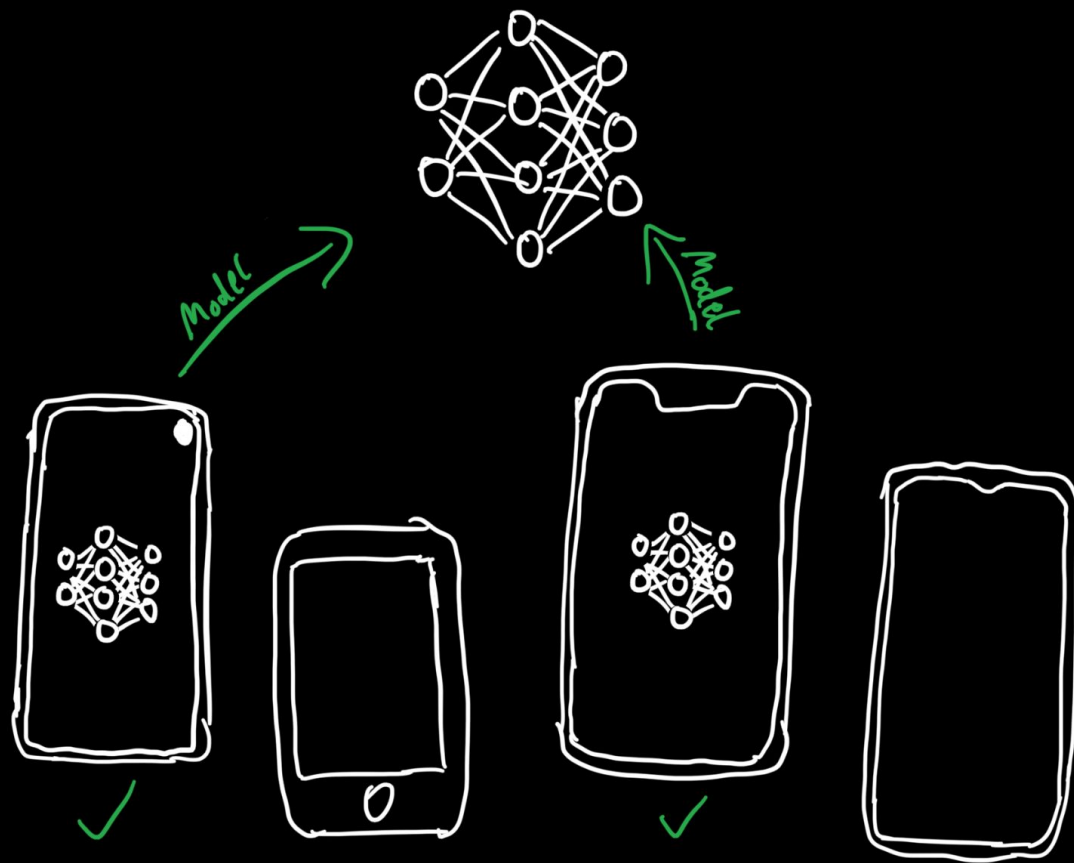
The architecture of Federated Learning

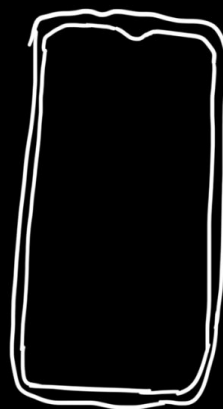
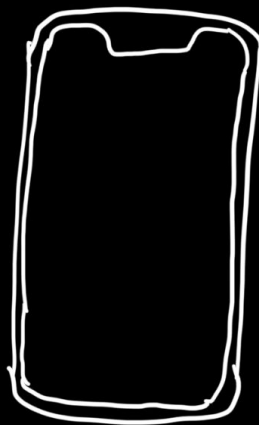
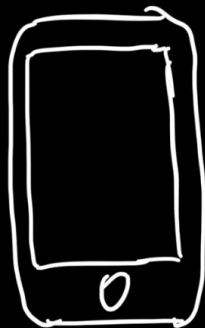
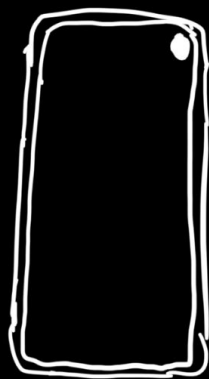
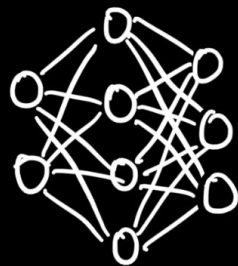


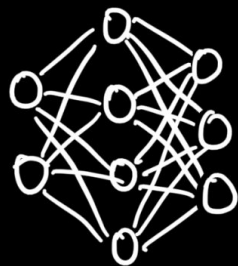






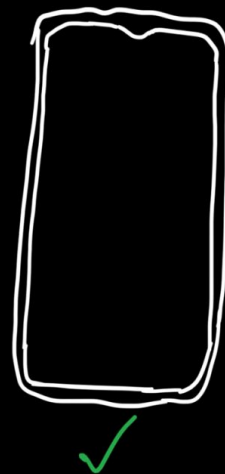
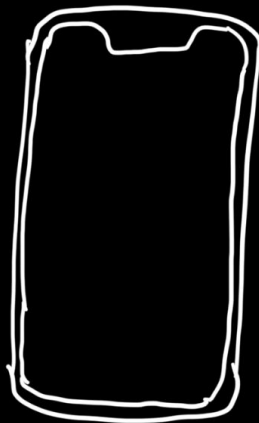
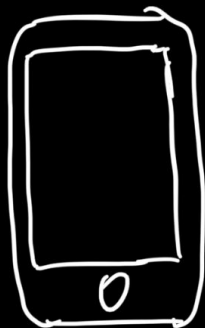
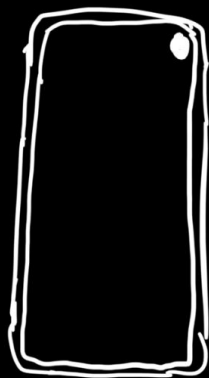


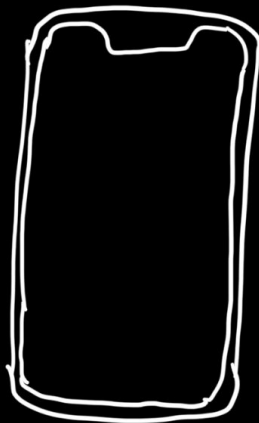
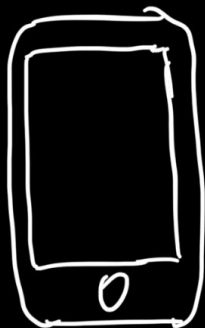
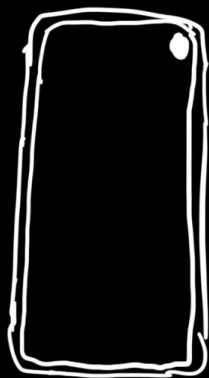
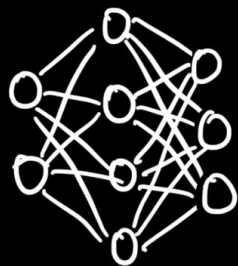


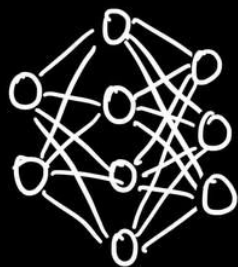


Model

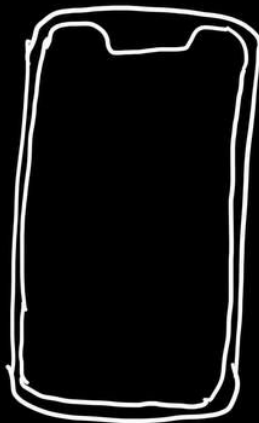
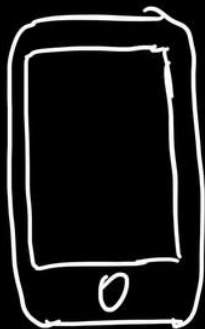
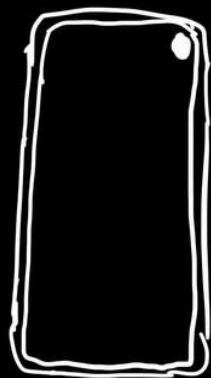
A green curved arrow pointing from the neural network diagram down to the fourth smartphone sketch, indicating the model's output or prediction.

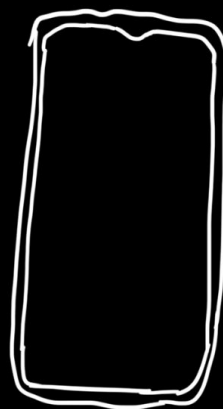
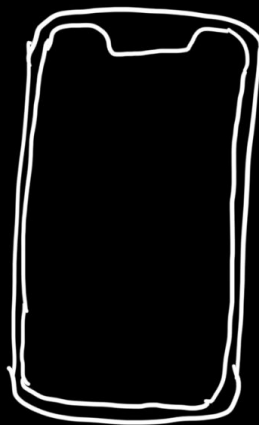
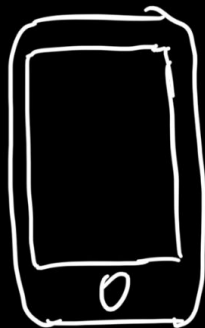
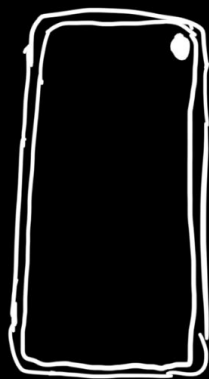
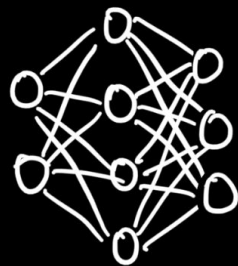


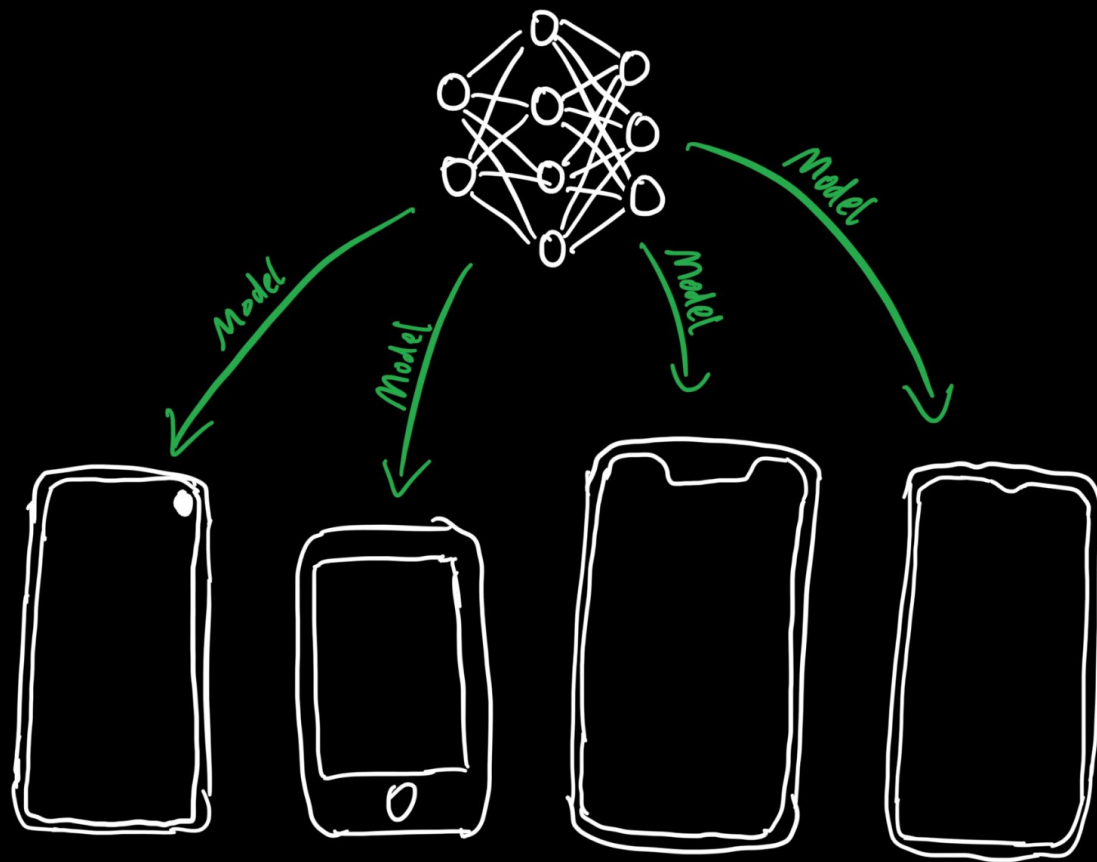




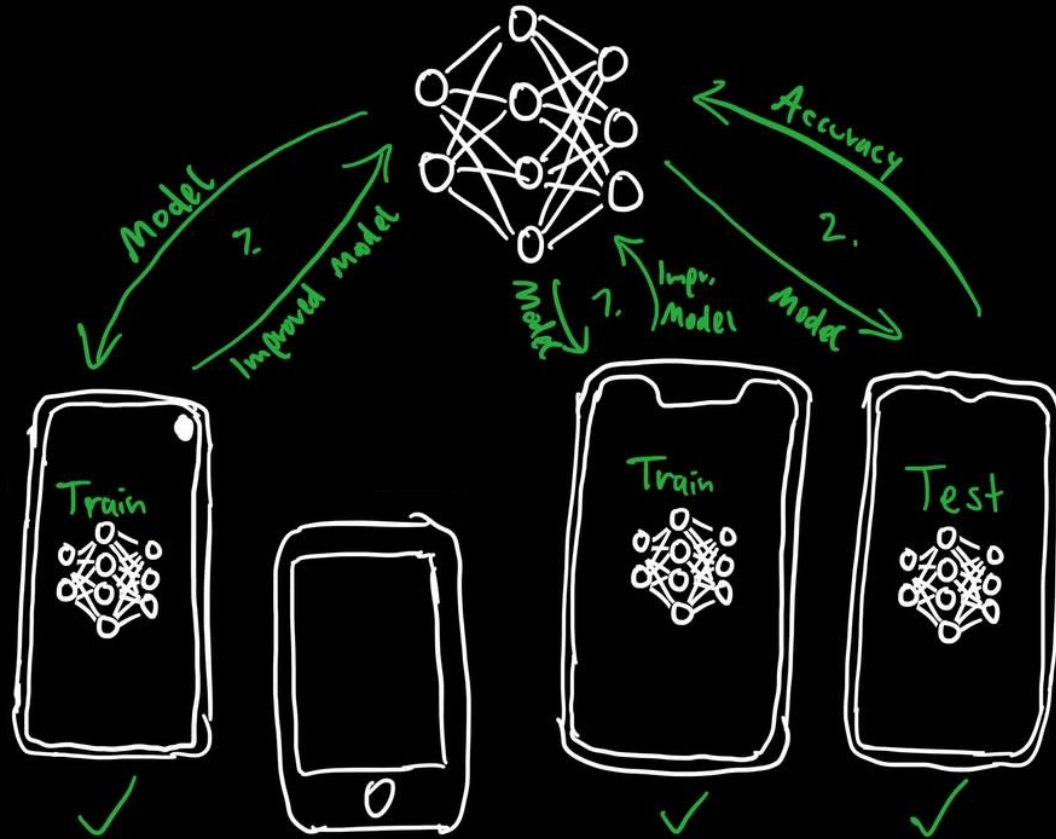
Accuracy





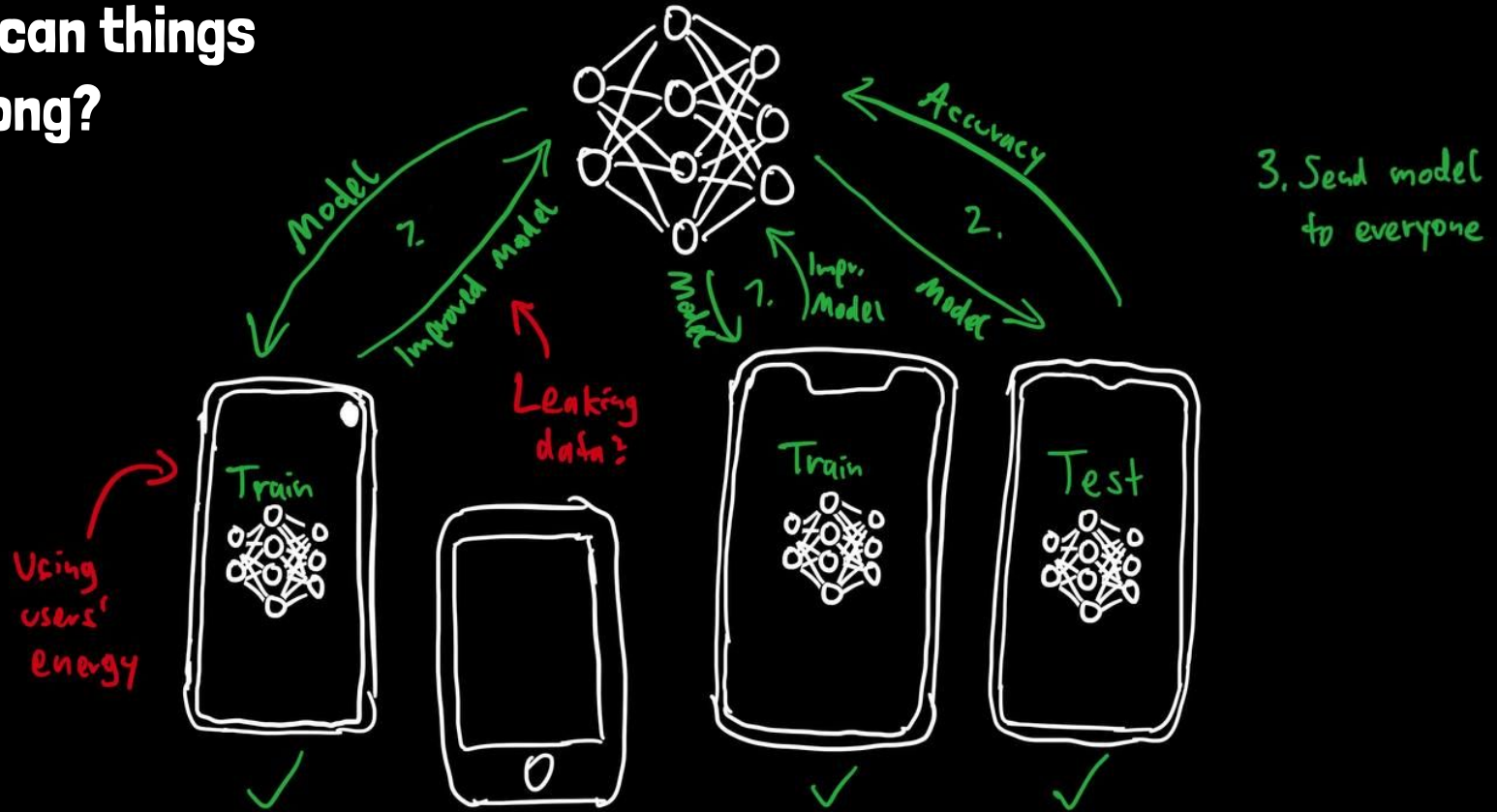


On one slide



3. Send model to everyone

Where can things go wrong?



Using users' compute resources



**Only some devices
are eligible**



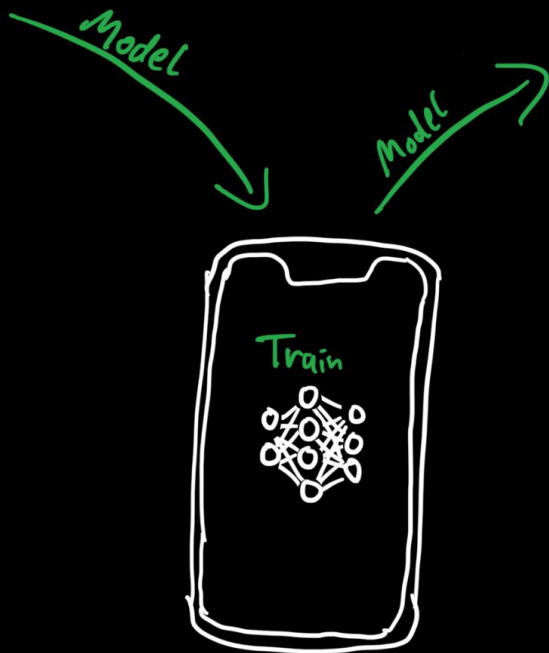
Is this really private?



No.

The trained model is not just random,
but encodes trends in the data.

Reconstruction Attacks



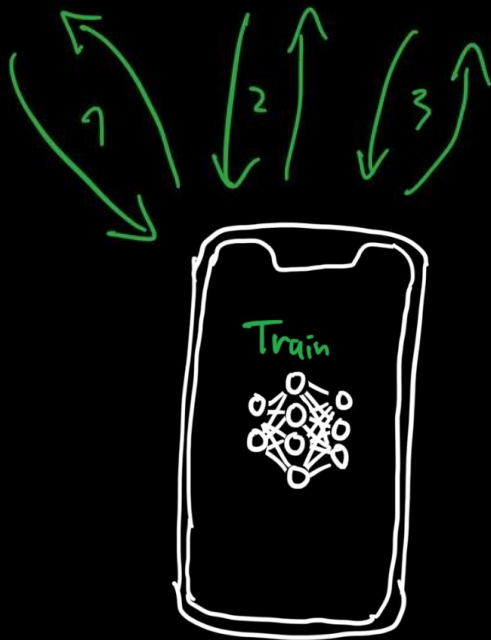
Difference between models ("gradient updates") necessarily leaks information about trained concepts.

Example: Autocomplete for "Let's meet at"

Before: 6, 7, 8

After: 8, 6, 7

Model Inversion Attacks



Multiple trainings can leak information about training data: Look for inputs that would cause the model to change in these ways.

Example: Learn to distinguish faces.

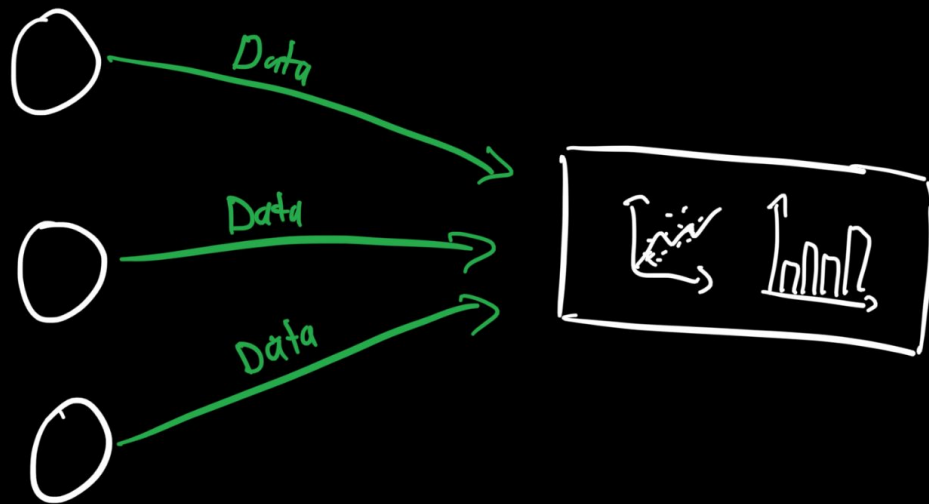


Extracted from victim by repeatedly claiming it is another face.

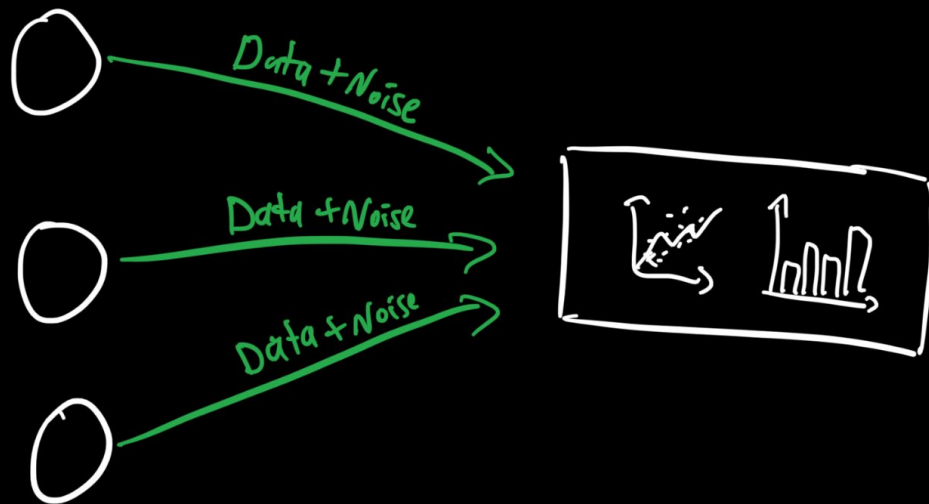
Differential Privacy

How to do statistics
without compromising privacy

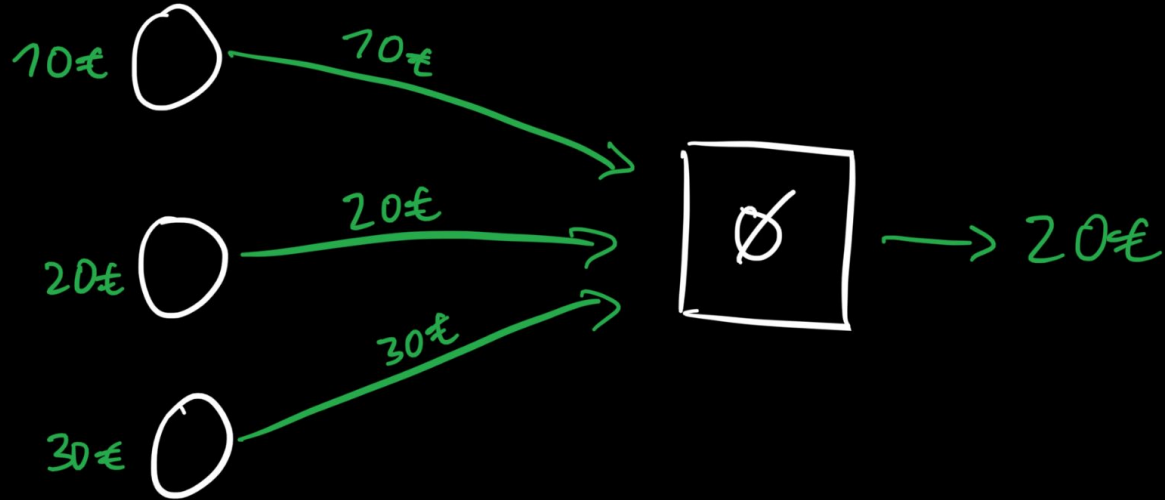
Differential Privacy



Differential Privacy

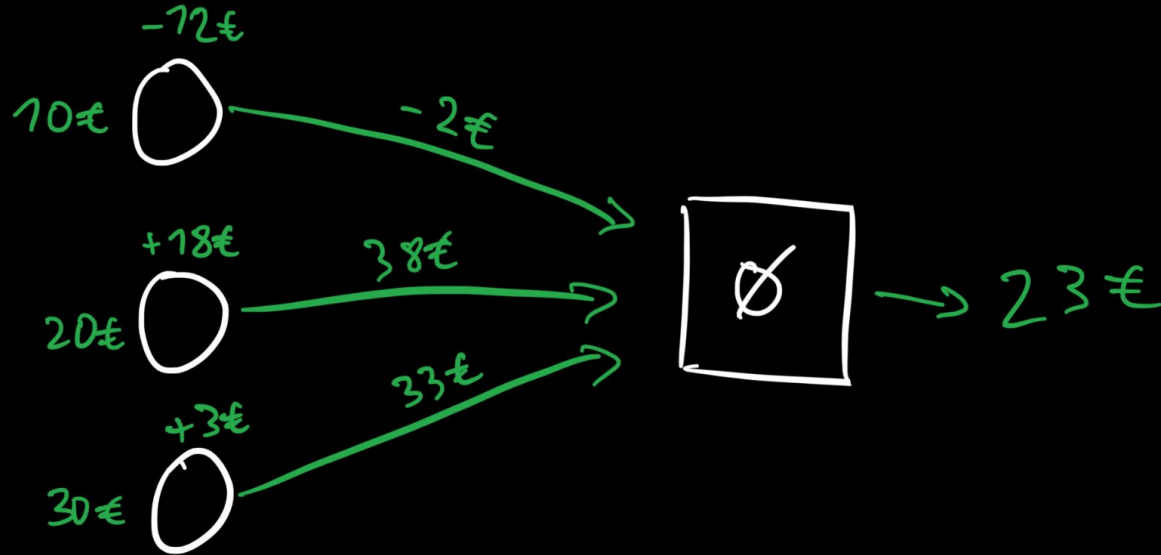


Example: Salary



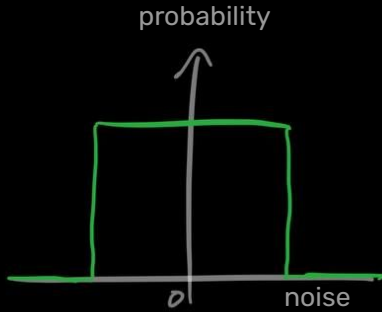
"Tell me your salary!"

Example: Salary

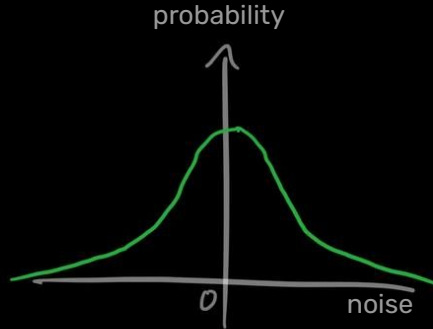


"Tell me your salary, but before telling me, add a random number between -20 € and +20 €!"

Better: Gaussian Noise



Noise in a range



Gaussian Noise

Every situation can reach every result!
I know **nothing for certain.**

The Caveat™

Privacy
for the individual



Utility
for everyone

More noise
Less accurate results
Need more data

Less noise
More accurate results
Need less data

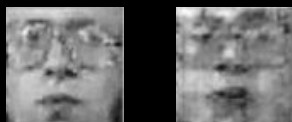
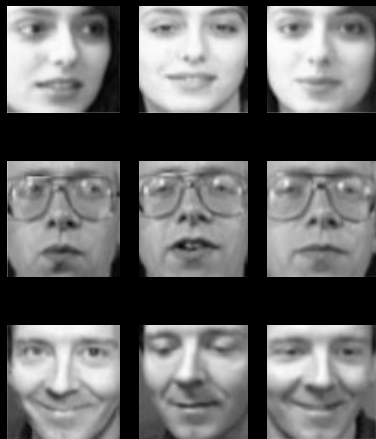


Amazingly, this works on Neural Networks!

(but not as well)

The Caveat™

Example: Distinguishing faces



without & with
differential privacy

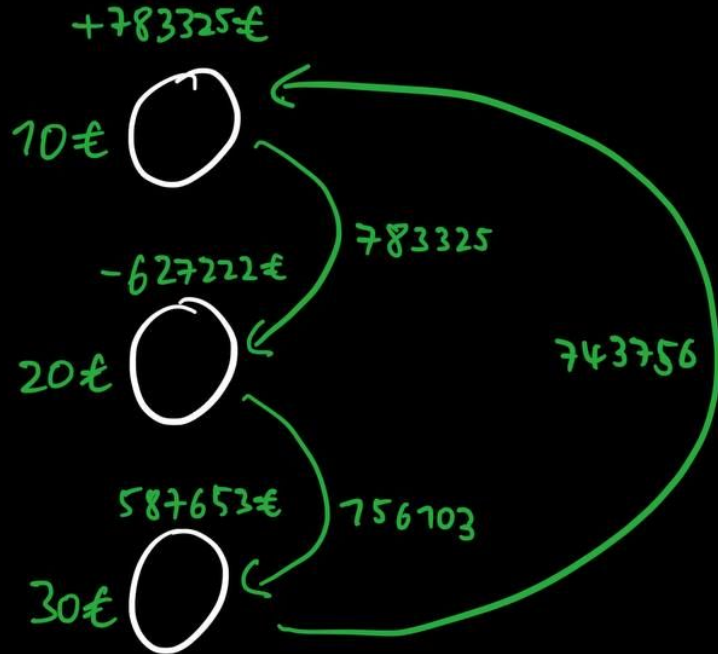
Example: Synthesizing speech

"The victim's device contains speech recordings. The GAN will generate **babbling, with lots of fictitious word-like sounds** [...], thus there is no privacy violation. However, it may be possible to infer the **language** used (e.g., English or Chinese) or whether the speaker is **male or female**, and this leaked information may constitute a privacy violation."

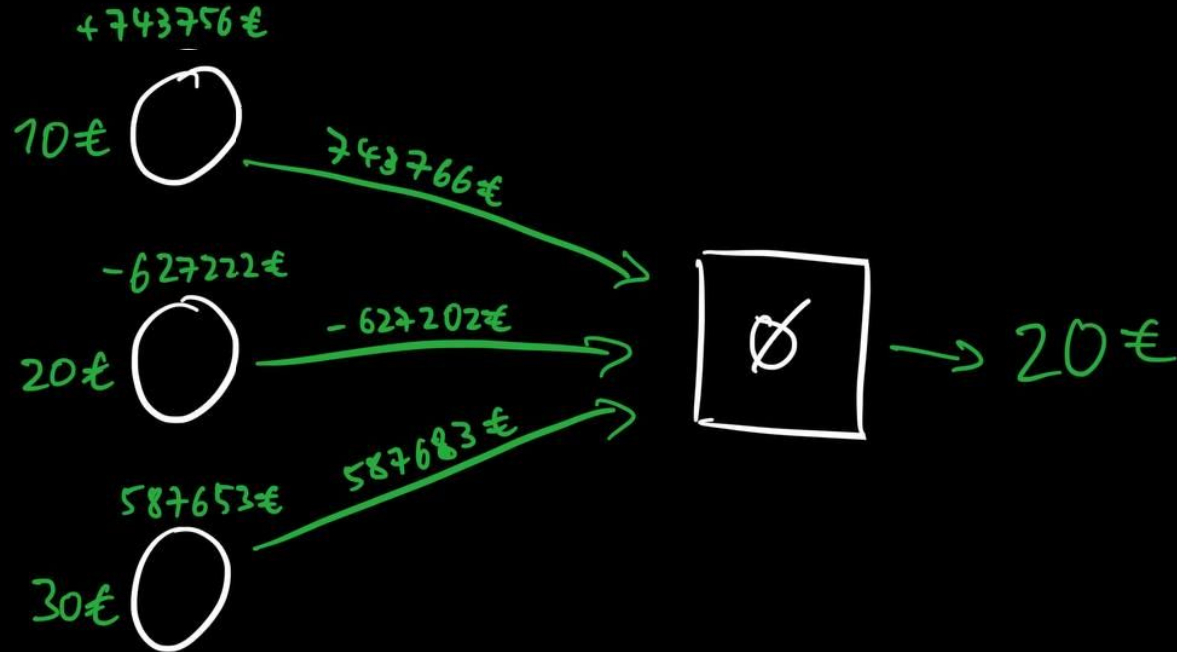
Secure Aggregation

How to prevent the server from analyzing individual responses

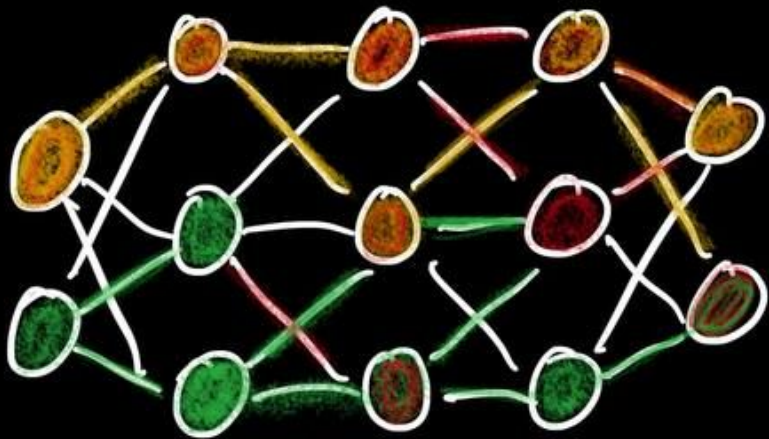
Secure Aggregation



Secure Aggregation



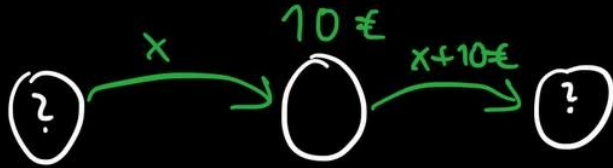
The Caveat™



Real models are huge
(millions of parameters)

- Room for non-overlapping concepts
 - Private data may leak
 - Especially if users have heterogeneous data
- Outlier data can't be ignored
 - Model Poisoning easier

The Caveat™



- Ground level of trust still required
- Requires knowing peers separately from central entity
 - Requires public key infrastructure
 - Doesn't work on users' devices

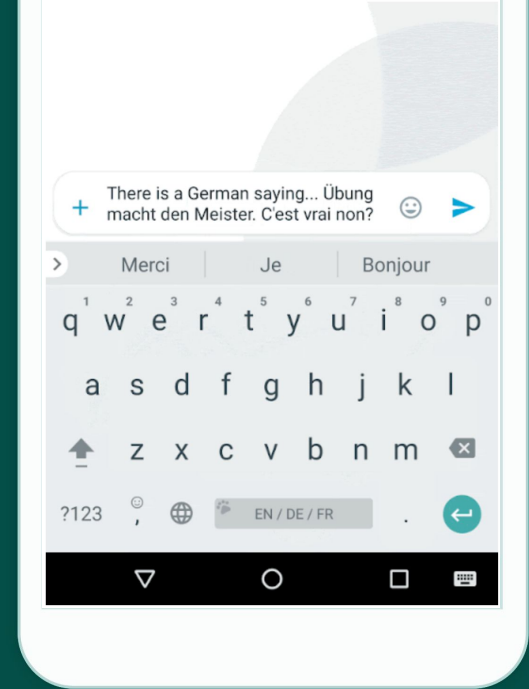
Use in the real world



Gboard

Federated Learning improves autocomplete and dictation.

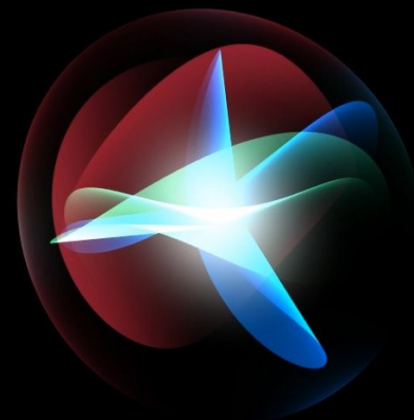
“[...] A technology called federated learning helps Gboard learn new words and phrases without sending the text you speak or type to Google. What Gboard learns might be sent to Google services, without including what you typed or spoke, where it will be combined with learnings from other users to create better speech and typing models. [...]”



Siri

Federated Learning improves on-device speaker verification.

Only third-party literature of AI conferences.
No official information about it, not even in the privacy policy.



Healthcare

- Patient data is very private
→ few people donate data
- The risk of sharing is high,
even with pseudonymization
- Federated Learning could
enable more diverse datasets
that better represent all
humans

“[A] commonly used algorithm to determine enrollment in specific health programs was **biased against African Americans**, assigning the same level of risk to healthier Caucasian patients. [...] [O]ne way to alleviate the risk for such biased algorithms is the ability to learn from **EHR data that is more representative of the global population and which goes beyond a single hospital or site**. [Other researches] conducted both patient representation learning and obesity comorbidity phenotyping in a federated manner and got good results.” **Federated Learning for Healthcare Informatics**

**Potential
societal impact**



Dilutes privacy



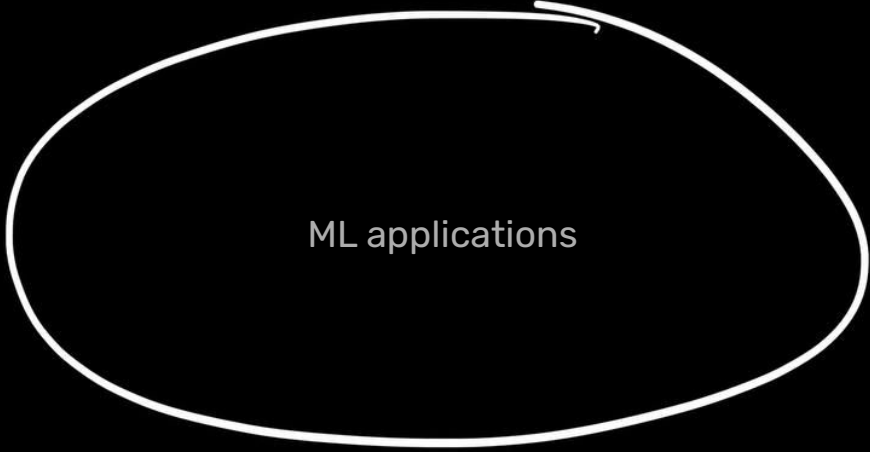
Federated Learning **complicates and dilutes the concept of privacy.**

Most **users won't understand** what's happening to their data.



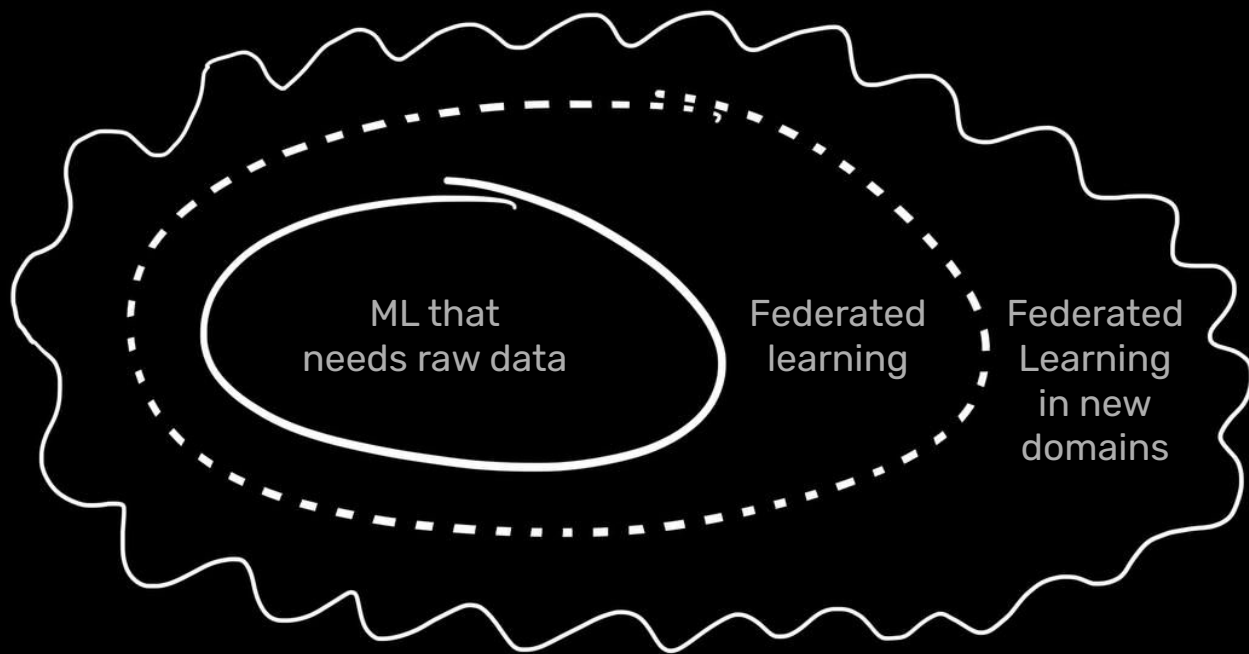
No! It's used to train a local ML model and then the resulting parameters, which only encode general concepts of our private data, are collected to a central location in a way that no single dataset can be inspected individually.

Decreases privacy? Increases utility?



ML applications

Decreases privacy? Increases utility?



Makes it socially more ethical to **use private data** that was previously tabu.

But that data can be used **for good**.

Summary



What we covered

The dilemma of data aggregation: Privacy for the individual vs. Utility for everyone

Traditional ML vs. Federated Learning

Problems, solution attempts, and caveats

- Using users' resources
 - Only in certain situations
 - Slippery slope
- Leaking private data
 - Differential Privacy
 - Re-introduces trade offs
 - Doesn't work on heterogeneous data
 - Secure Aggregation
 - Not applicable for most user-facing software
 - Makes model poisoning easier

Lots of research, sparingly used in the real world

Potential societal impact: dilutes privacy, may reduce overall privacy, but increase utility

Sources

mgar.us/federated-learning-slides

Hackerethik principles by the *Chaos Computer Club*; ccc.de/de/hackerethik

Federated Learning comic by *Google AI*; federated.withgoogle.com

Deep Models Under the GAN: Information Leakage from Collaborative Deep Learning paper by *Briland Hitaj, Giuseppe Ateniese, and Fernando Perez-Cruz*; users.wpi.edu/~kmus/ECE579M_files/ReadingMaterials/p603-hitajA.pdf

Inverting Gradients – How easy is it to break privacy in federated learning? paper by *Jonas Geiping, Hartmut Baumeister, and Hannah Dröge*; arxiv.org/pdf/2003.14053.pdf

Differential Privacy at the US Census podcast episode by *Simson Garfinkel* and *Kyle Polich* from *Data Skeptic*; dataskeptic.libsyn.com/differential-privacy-at-the-us-census

Practical Secure Aggregation for Privacy-Preserving ML paper by *Keith Bonawitz, Vladimir Ivanov, Ben Kreuter, Antonio Marcedone, H. Brendan McMahan, Sarvar Patel, Daniel Ramage, Aaron Segal, and Karn Seth*; eprint.iacr.org/2017/281.pdf

Gboard device frame image from *Google Play*; play.google.com/store/apps/details?id=com.google.android.inputmethod.latin

Trilingual input typing in Gboard from *The Machine Intelligence Behind Gboard*; ai.googleblog.com/2017/05/the-machine-intelligence-behind-gboard.html

Learn how Gboard gets better support ticket from *Google*; support.google.com/gboard/answer/9334583

Siri image from *Apple*; apple.com/siri

How Apple personalizes Siri without hoovering up your data article by *Karen Hao* from *Technology Review*; technologyreview.com/2019/12/11/131629/apple-ai-personalizes-siri-federated-learning

Federated Learning for Healthcare Informatics paper by *Ye Xu, Benjamin S. Glicksberg, Chang Su, Peter Walker, Jiang Bian, and Fei Wang*; link.springer.com/content/pdf/10.1007/s41666-020-00082-4.pdf

Discussion

- How would you feel about federated learning happening on your private data? What about keyboard data? Medical data?
- Is the energy usage ethical? Is there an ethical difference between mining Bitcoin and training a machine model?
- Last week we talked about GDPR's informed consent. Can that work if the tech isn't widely understood? What's your opinion on Apple's and Google's approach?
- Do you think Federated Learning is beneficial to society overall? If you had the option to "undiscover" it, would you do that?
- No raw data means that for companies/researchers, it's difficult to adjust for biases and outliers (see model poisoning). Are there use cases where data quality trumps privacy and you think raw data should be collected?