# Curs 13
# Federated Learning

What is Federated Learning?

Federated Optimization

Privacy for Federated Learning and Analytics

Probleme deschise și altele subiecte

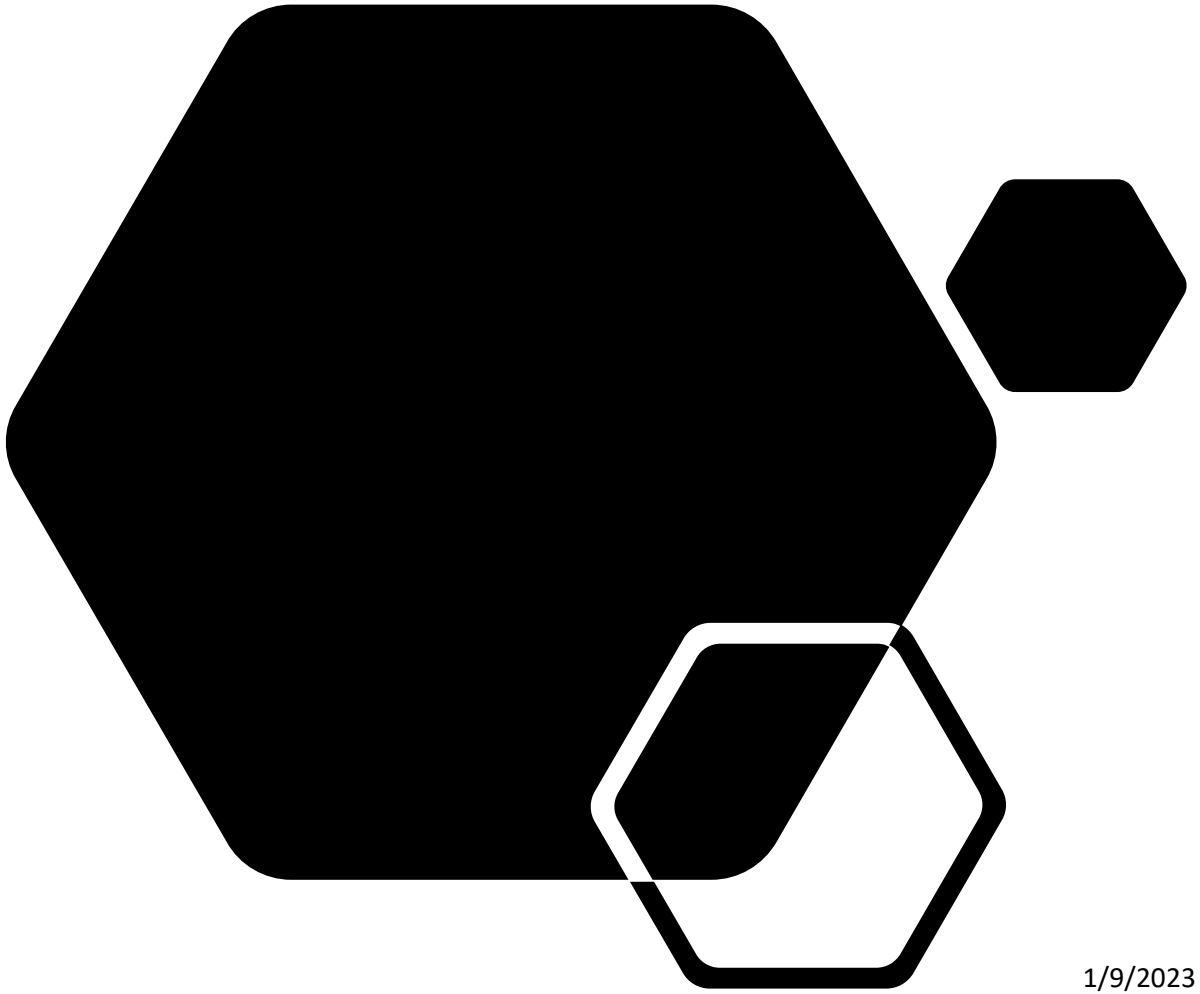
1/9/2023

# Ce este Federated Learning?

Edge data

Cross-Device Federated Learning

Federated Analytics



# Datele se nasc în dispozitivele edge

- Miliarde de telefoane și dispozitive IoT generează în mod constant date

- Datele permit produse mai bune și modele mai inteligente

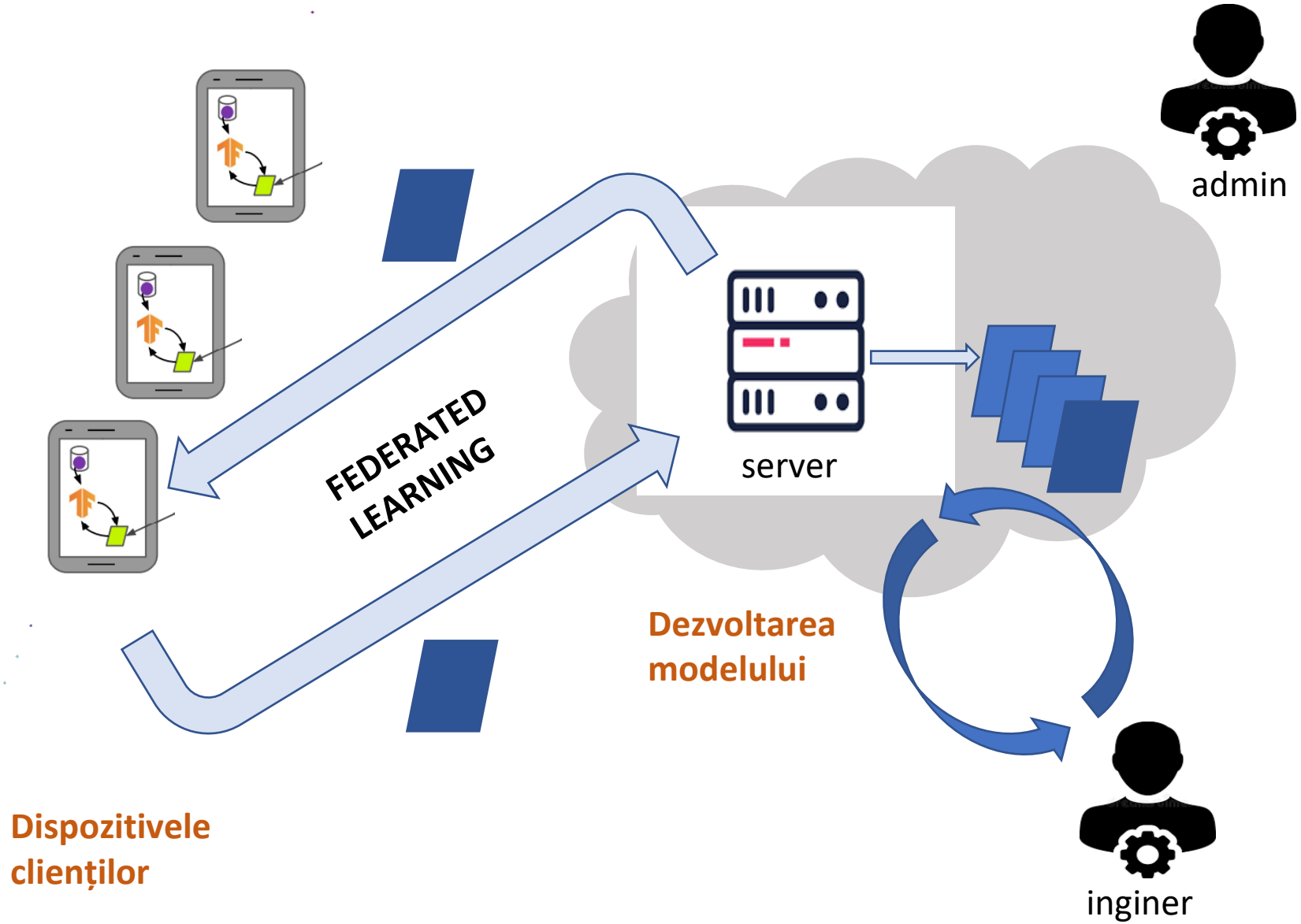
# Pot datele să trăiască in dispozitivele edge?

- Procesarea datelor este mutată pe dispozitiv:
  - Latență îmbunătățită
  - Funcționează offline
  - Durată mai bună a bateriei
  - Avantaje privind confidențialitatea

- De exemplu, inferență pe tastaturi și camera de pe telefoanele mobile.

!!!
Invățare?
Analiză?

# Federating Learning pe mai multe dispozitive



# Aplicații ale FL

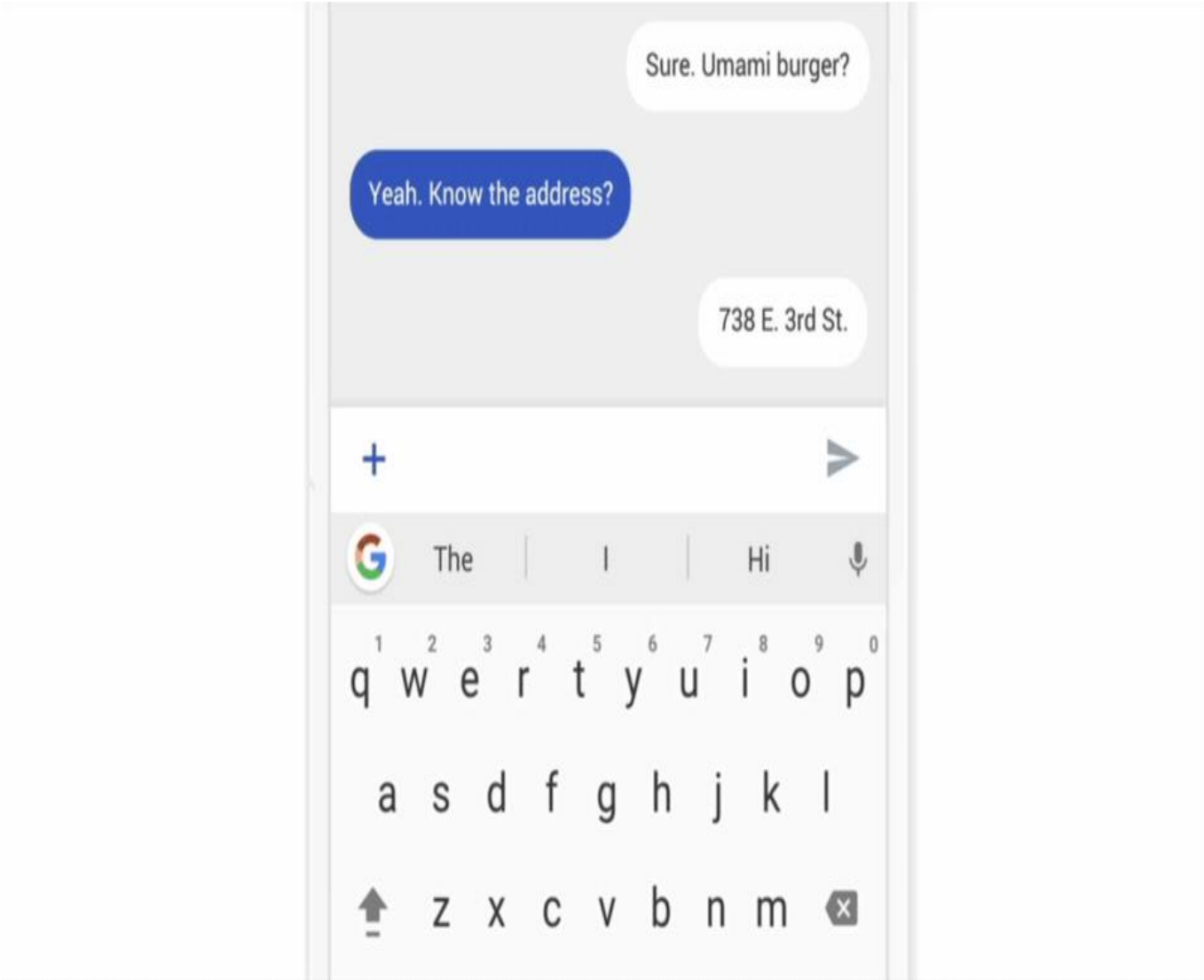
- Ce face o aplicație bună?
  - Datele de pe dispozitiv sunt mai relevante decât datele proxy de pe partea serverului
  - Datele de pe dispozitiv au nivelul de confidențialitate mai mare
  - Etichetele pot fi deduse în mod natural din interacțiunea utilizatorului
- Exemple de aplicații
  - Modelare lingvistică pentru utilizarea tastaturi telefoanelor mobile și recunoaștere vocală
  - Clasificarea imaginilor pentru prezicerea ce fotografii oamenii vor distribuii
  - ....

# Gboard: predicția cuvântului următor

•**Federtated RNN** (comparativ cu **modelul N-gram**):

•● Precizie mai bună a predicției cuvântului următor: **+24%**

•● Predicții mai utile: **+10% mai multe clickuri**



# Alte modele de FL în Gboard

**Predicția emoji**

- Predicții emoji cu 7% mai precise
- Clickuri pe banda de predicții cu +4% mai mult
- Cu 11% mai mulți utilizatori distribuie emoji-uri

**Predicția acțiunii**

Când este util să sugerezi un gif?

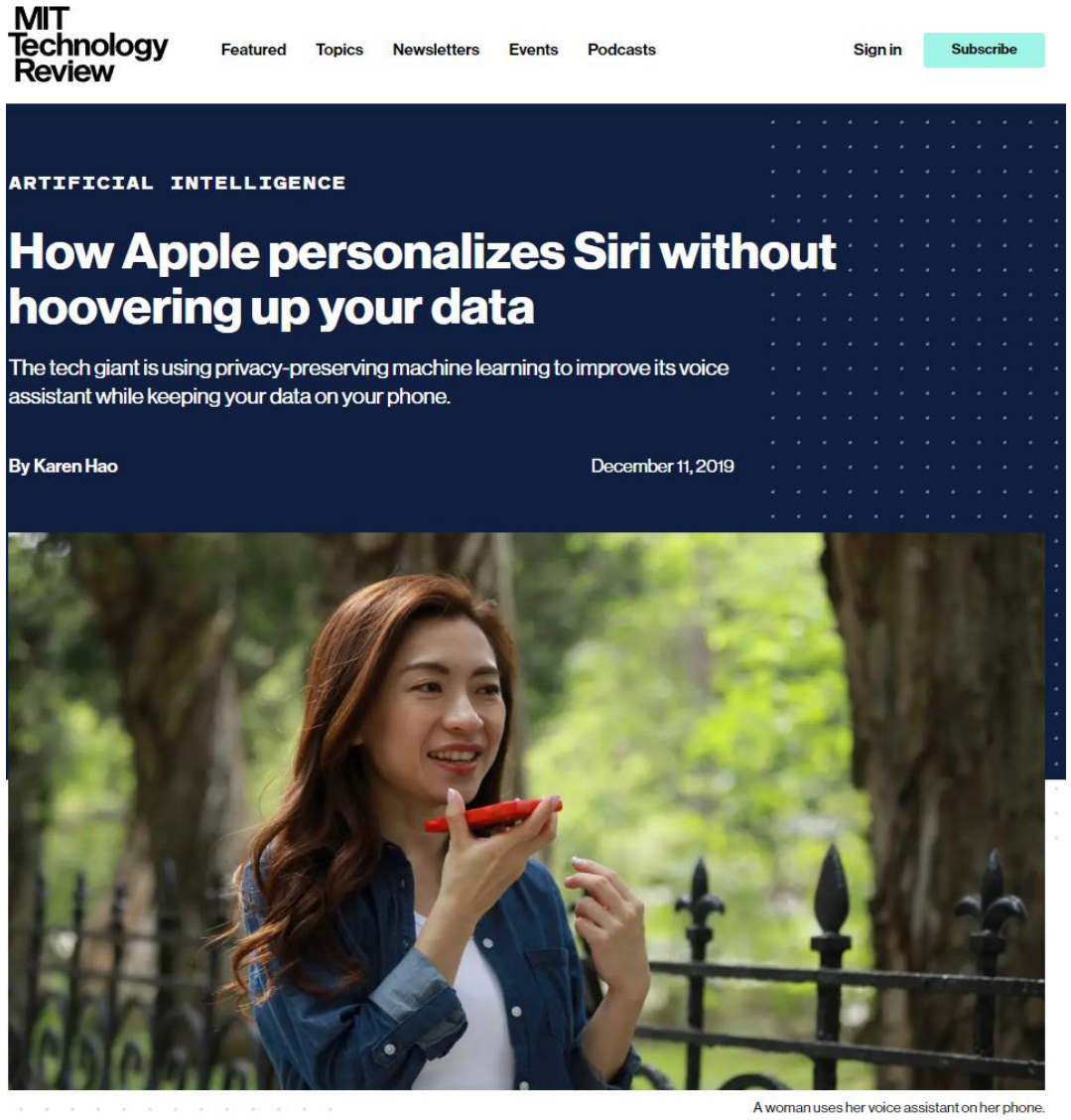
- Reducere cu 47% a sugestiilor nefolositoare
- Creșterea generală a emoji, gif și autocolant acțiuni

**Descoperirea de cuvinte noi**

- Descoperirea federată a ceea ce vorbesc oamenii tastează că Gboard nu știe.

# FL pe mai multe dispozitive la Apple

•*“Instead, it relies primarily on a technique called federated learning, Apple’s head of privacy, Julien Freudiger, told an audience at the Neural Processing Information Systems conference on December 8. Federated learning is a privacy-preserving machine-learning method that was first introduced by Google in 2017. It allows Apple to train different copies of a speaker recognition model across all its users’ devices, using only the audio data available locally. It then sends just the updated models back to a central server to be combined into a master model. In this way, raw audio of users’ Siri requests never leaves their iPhones and iPads, but the assistant continuously gets better at identifying the right speaker.”*




https://www.technologyreview.com/2019/12/11/131629/apple-ai-personalizes-siri-federated-learning/

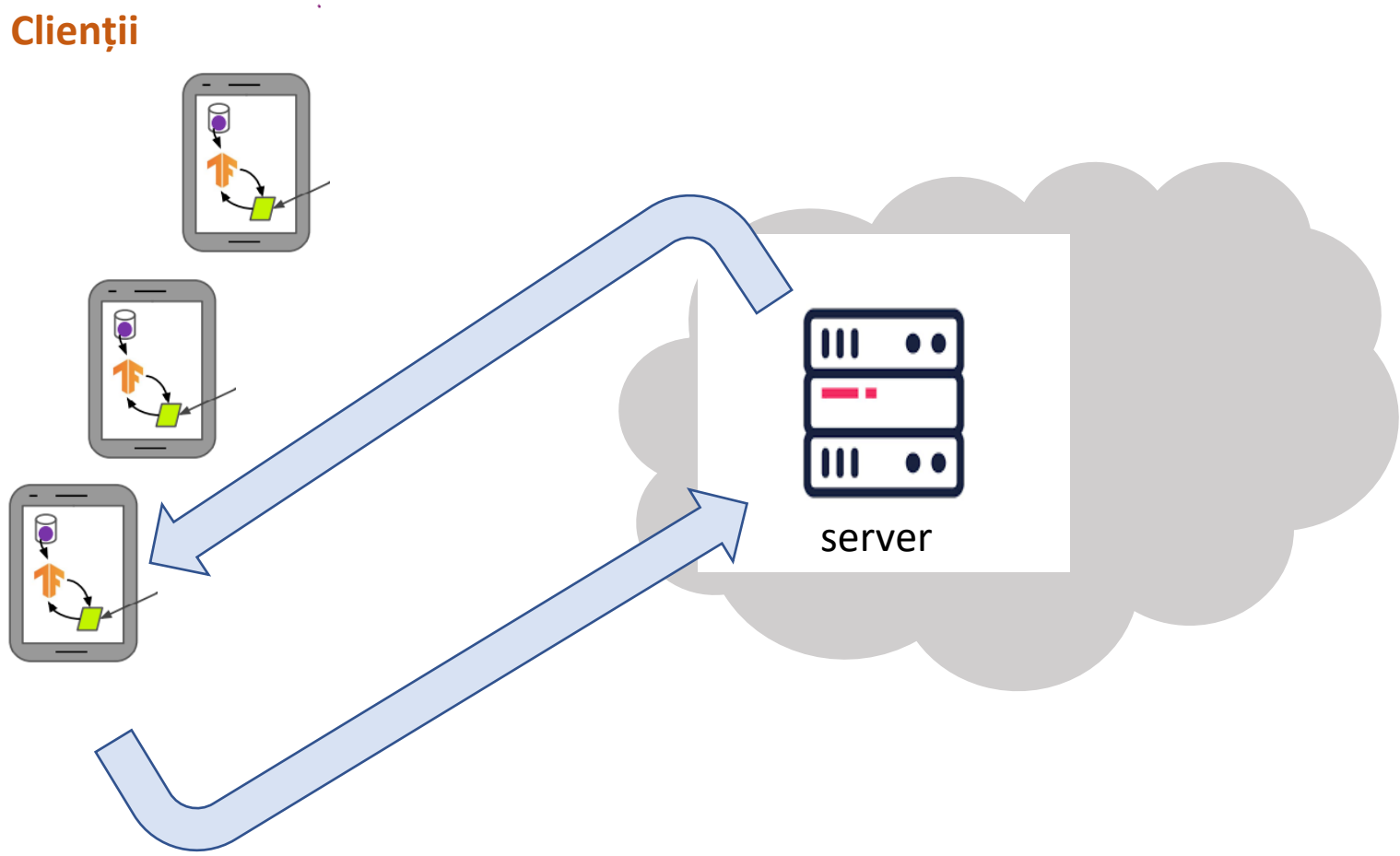
# Federated Learning

- **Federated learning** este un cadru de învățare automată în care mai multe entități (clienți) colaborează pentru a rezolva o problemă de învățare automată, sub coordonarea unui server central sau a unui furnizor de servicii. Datele brute ale fiecărui client sunt stocate local și nu sunt schimbate sau transferate; în schimb, actualizări destinate agregării imediate sunt folosite pentru a atinge obiectivul de învățare.

definiție propusă în
***Advances and Open Problems in Federated Learning***

# Federated learning - caracteristici definitorii

- Datele sunt generate local și rămân descentralizate.
- Fiecare client își stochează propriile date și nu poate citi datele altor clienți.
- Datele nu sunt distribuite independent și identic.
- Un server/serviciu central de orchestrare coordonează instruirea, dar nu vede niciodată datele brute.



# Terminologia FL

- **Clienți** - Noduri de calcul care dețin și date locale, de obicei aparținând unei entități:
  - Dispozitive IoT
  - Dispozitive mobile
  - Centre de date în diferite regiuni geografice
- **Server** - Noduri de calcul suplimentare care coordonează procesul FL, dar nu accesează datele brute. De obicei, nu o singură mașină fizică.

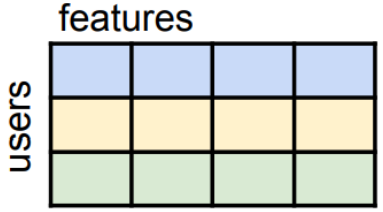
# Cross-device FL

*VS*

# Cross-silo FL

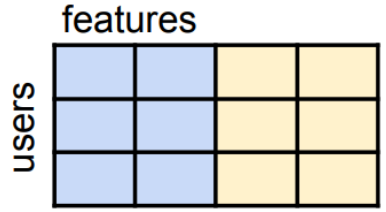
## Cross-device FL

- milioane de dispozitive client disponibile intermittent
- clienții nu pot fi indexați direct (nu se folosesc identificatori de clienți), actualizările sunt anonime
- serverul poate accesa doar un eșantion aleatoriu de clienți în fiecare rundă
- populație mare = majoritatea clienților participă o singură dată
- comunicarea este adesea principalul blocaj
- date partiționate orizontal




## Cross-silo FL

- numar mic de clienti (institutii, organizatii), disponibilitate mare
- fiecare client are o identitate sau un nume care permite sistemului să-l acceseze în mod specific
- majoritatea clienților participă la fiecare rundă
- clienții pot rula algoritmi care mențin starea locală pe parcursul rundelor
- comunicarea și calculul ar putea fi principalul blocaj
- date partiționate orizontal sau vertical



ARTIFICIAL INTELLIGENCE, DIAGNOSTICS

# UPenn, Intel partner to use federated learning AI for early brain tumor detection

The project will bring in 29 institutions from North America, Europe and India and will use privacy-preserved data to train AI models. Federated learning has been described as being born at the intersection of AI, blockchain, edge computing and the Internet of Things.

By ALARIC DEARMENT

1 Comment / May 11, 2020 at 10:03 AM

The University of Pennsylvania and chipmaker Intel are forming a partnership to enable 29 heatlhcare and medical research institutions around the world to train artificial intelligence models to detect brain tumors early.

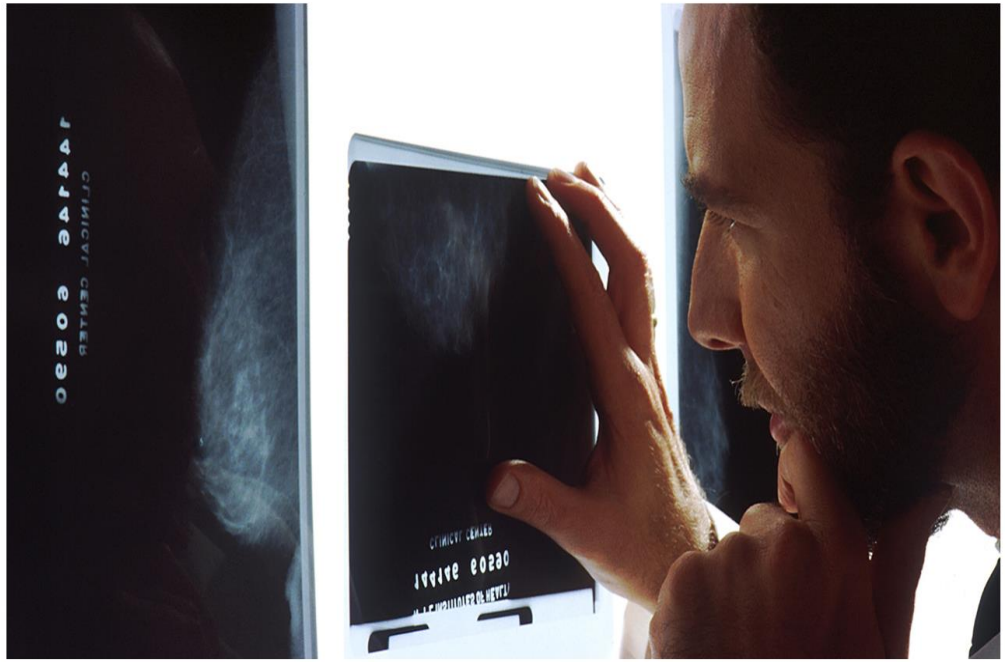
NVIDIA

HOME AI DATA CENTER DRIVING GAMING PRO GRAPHICS AUTONOMOUS MACHINES HEALTHCARE STARTUPS AI PODCAST

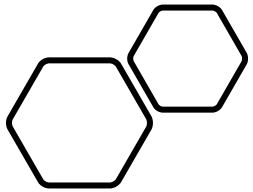
# Medical Institutions Collaborate to Improve Mammogram Assessment AI with NVIDIA Clara Federated Learning

In a federated learning collaboration, the American College of Radiology, Diagnosticos da America, Partners HealthCare, Ohio State University and Stanford Medicine developed better predictive models to assess breast tissue density.

April 15, 2020 by MONA FLORES

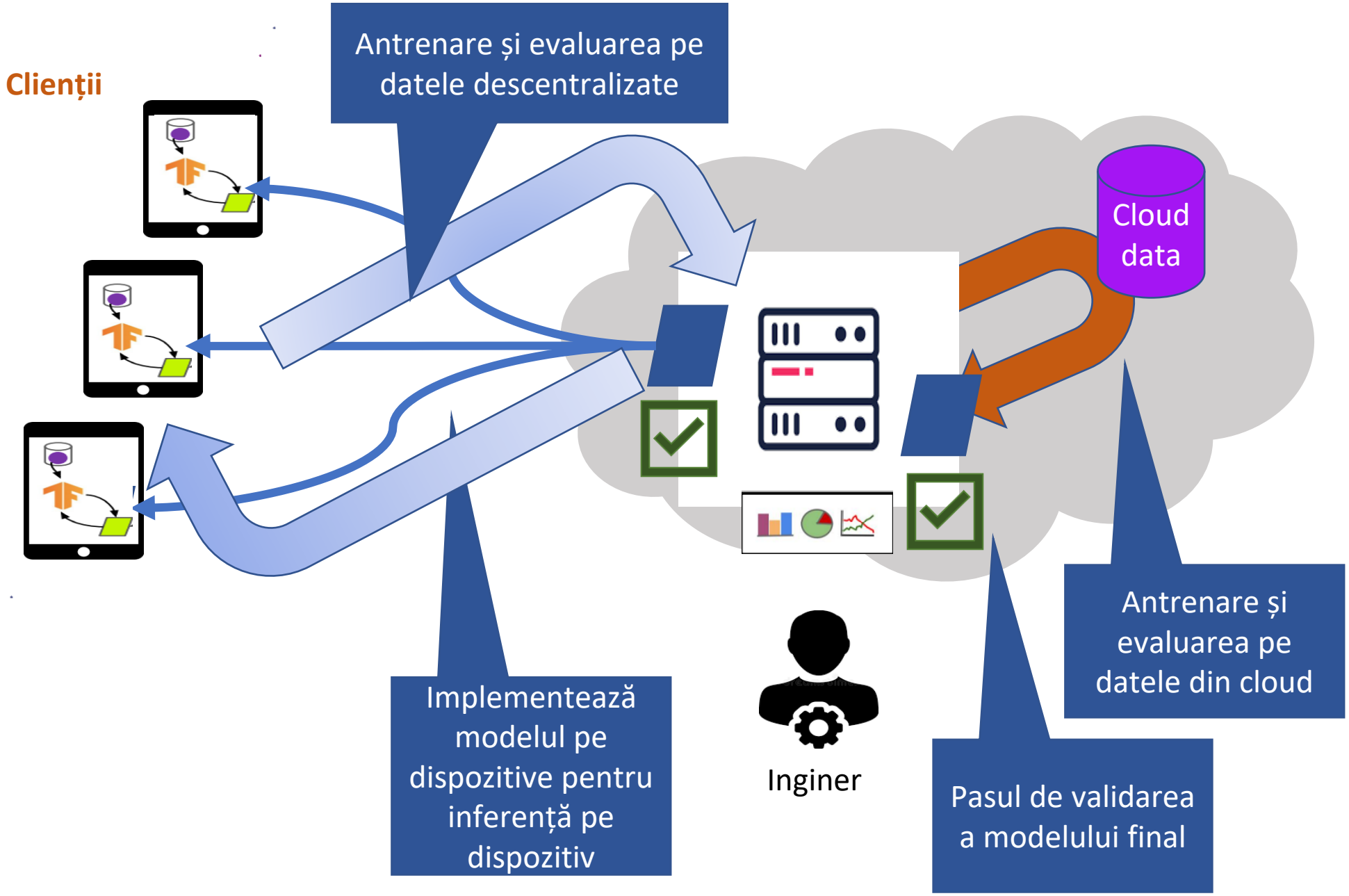


**Cross-silo federated learning pentru Intel/NVIDIA**

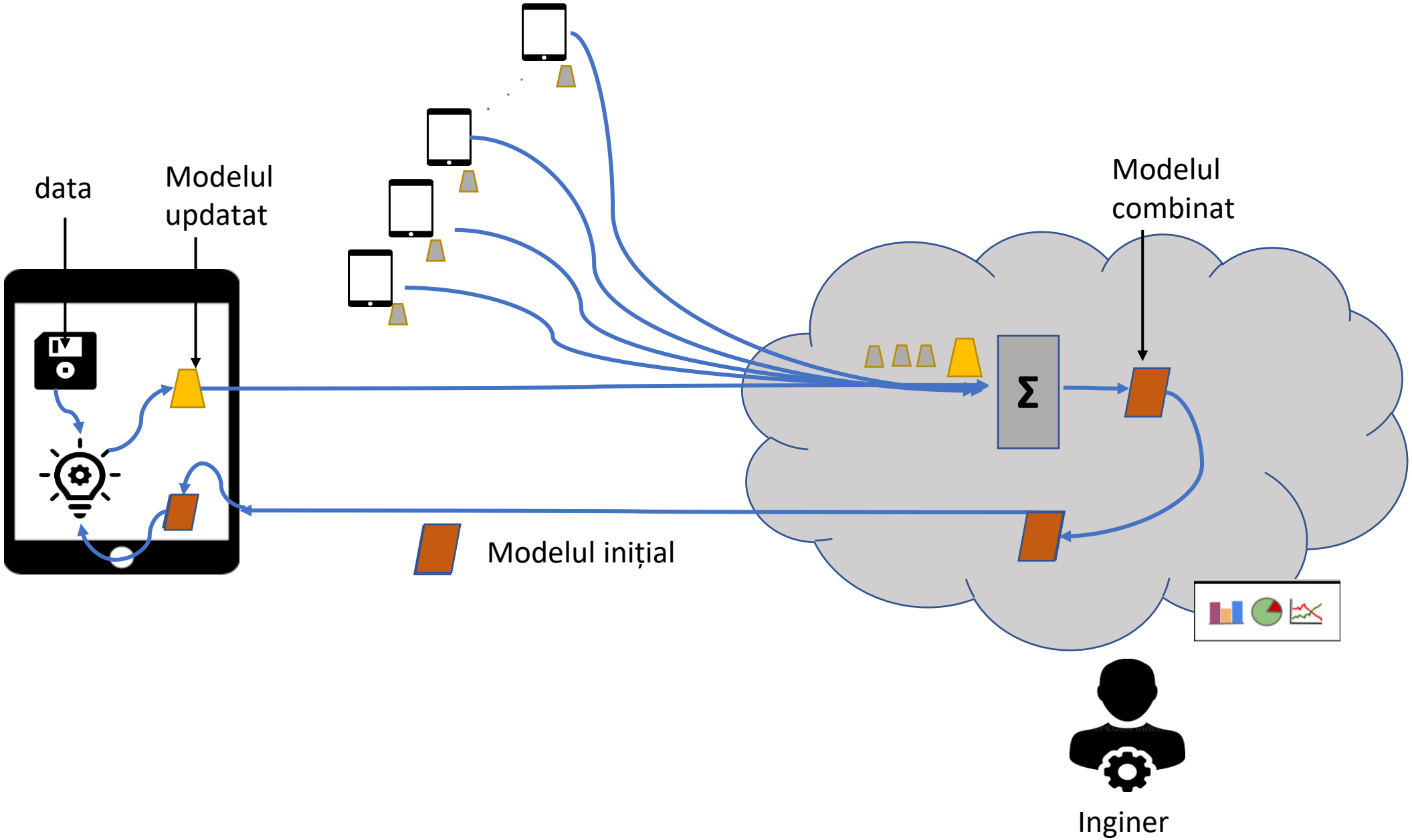


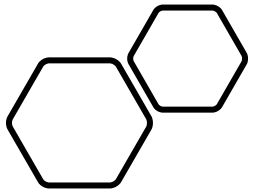
# Cross-Device Federated Learning

# Workflow-ul de dezvoltarea a modelului



# Federated training

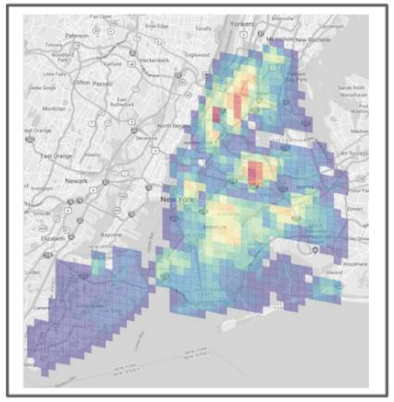




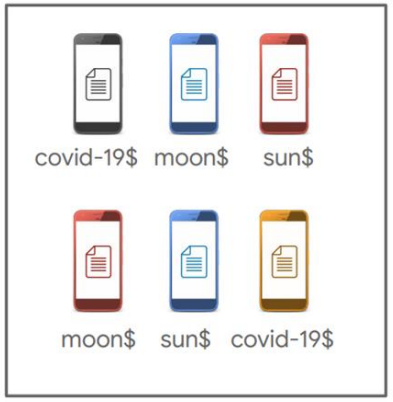
# Federated Analytics

# Federated Analytics

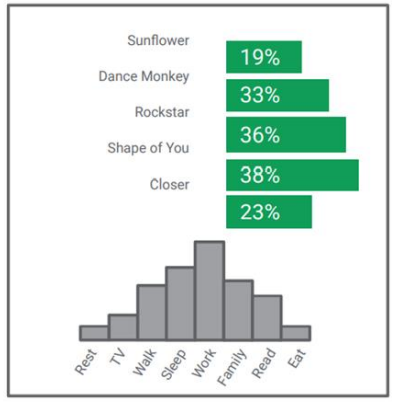
*"**Federated analytics** este practica de aplicare a metodelor științei datelor la analiza datelor brute care sunt stocate local pe dispozitivele utilizatorilor. La fel ca **FL** , funcționează prin rularea calculelor locale pe datele fiecărui dispozitiv și punând la dispoziția inginerilor doar rezultatele agregate - și niciodată datele de pe un anumit dispozitiv. Spre deosebire de **FL**, totuși, **federated analytics** își propune să sprijine nevoile de bază ale științei datelor. "*



Geo-location heatmaps




Frequently typed
out-of-dictionary words



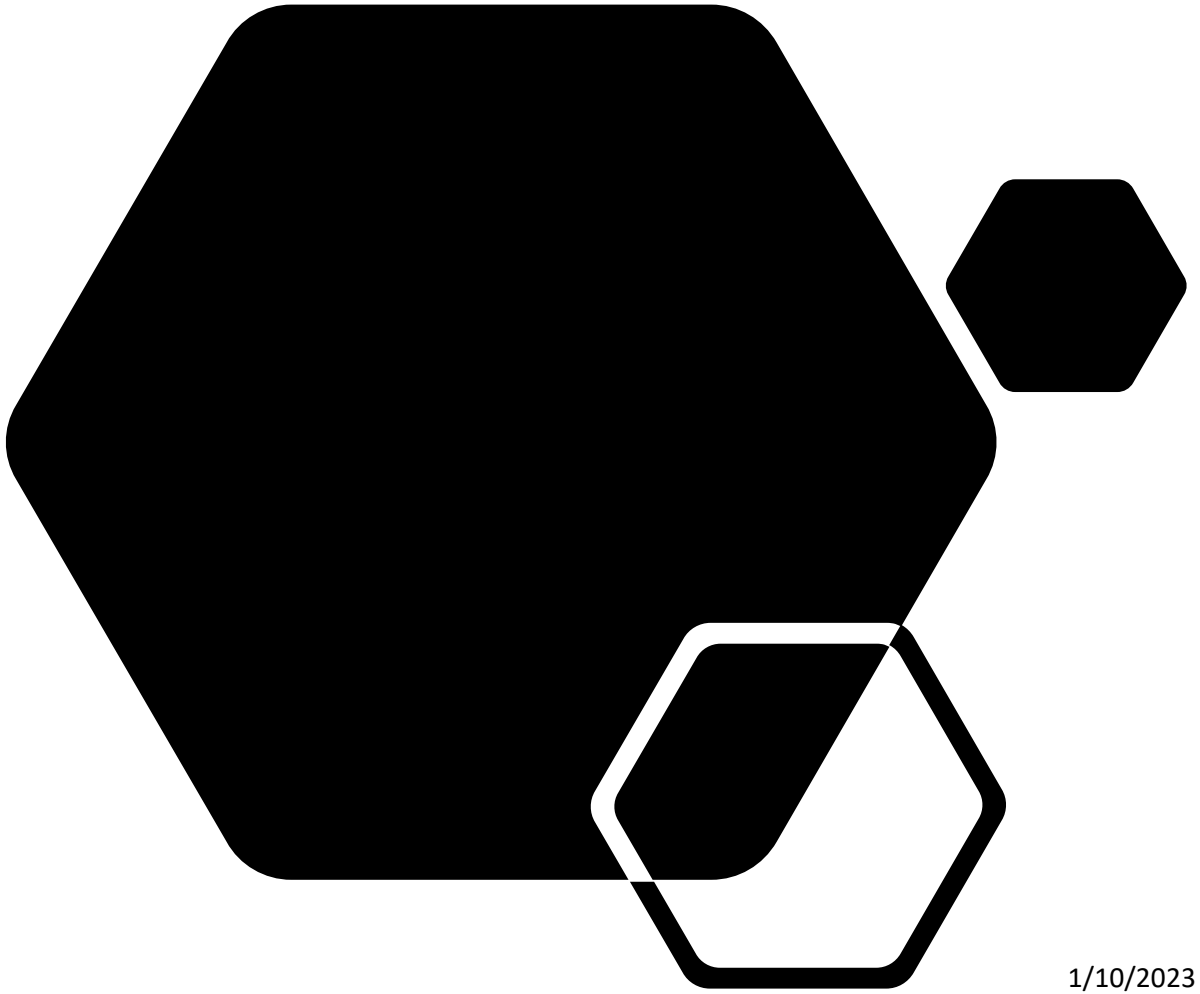

Popular songs, trends,
and activities

# Federated Optimization

Workflow și provocări

Federated Averaging (FedAvg)

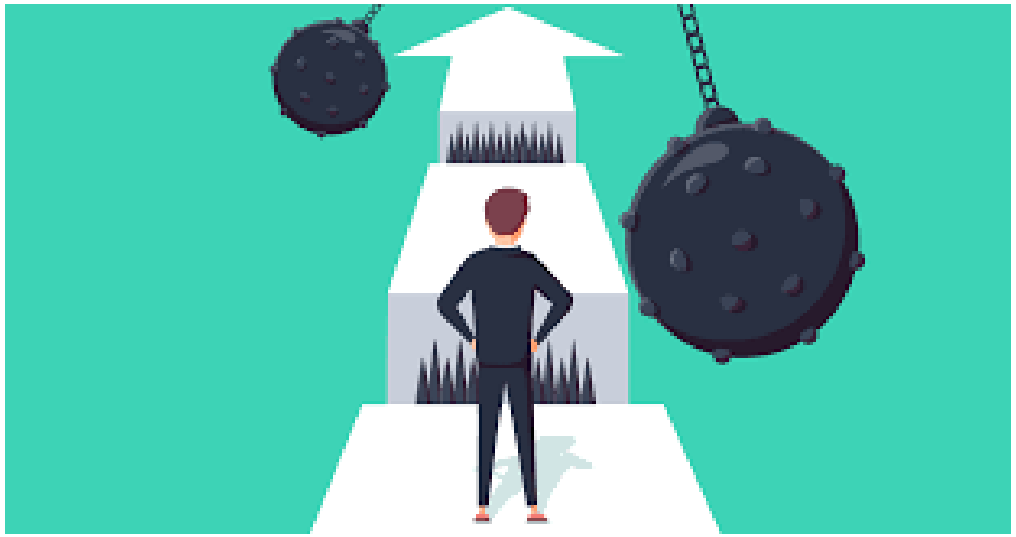
Hands-On Federated Optimization



1/10/2023

# Workflow și provocări

- **Goal**: antrenează modele de învățare automată la limită
- **De ce?**
  - ✓ reduce presiunea asupra rețelei
  - ✓ confidențialitate
  - ✓ încorporează rapid date noi



# Provocări

Datele și **sistemele eterogene** (distribuite neidentic) pot influența procedurile de optimizare

***comunicare costisitoare***

***preocupări legate de confidențialitate***

***eterogenitatea statistică***

***eterogenitatea sistemelor***

# Federated Averaging (FedAvg)

- Reduce comunicarea prin:
  - efectuarea de actualizări locale
  - comunicarea cu un subset de dispositive
- De ce este util să fie efectuată **actualizare locală**?
  - Poate efectua mai multe calcule locale (mai mult decât un singur mini-lot)
  - Încorporați actualizări mai rapid (aplicați imediat informații despre gradient)

✓ Poate duce la convergerea în mai puține runde de comunicare

✗ Poate afecta convergența dacă nu este reglat corespunzător

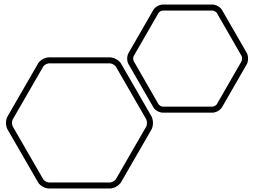
- **Cum diferă FedAvg de SGD distribuit?**
  - Actualizarea locală (**FedAvg**) poate reduce rundele de comunicare cu **~100x** față de **SGD**

# Concluzii

- Metodele de optimizare federate care efectuează actualizări locale pot în mod semnificativ reduce rundele de comunicare necesare pentru convergență

- Cu toate acestea, eterogenitatea poate duce la:
  - convergenţă mai lentă, stabilitate redusă, divergenţă

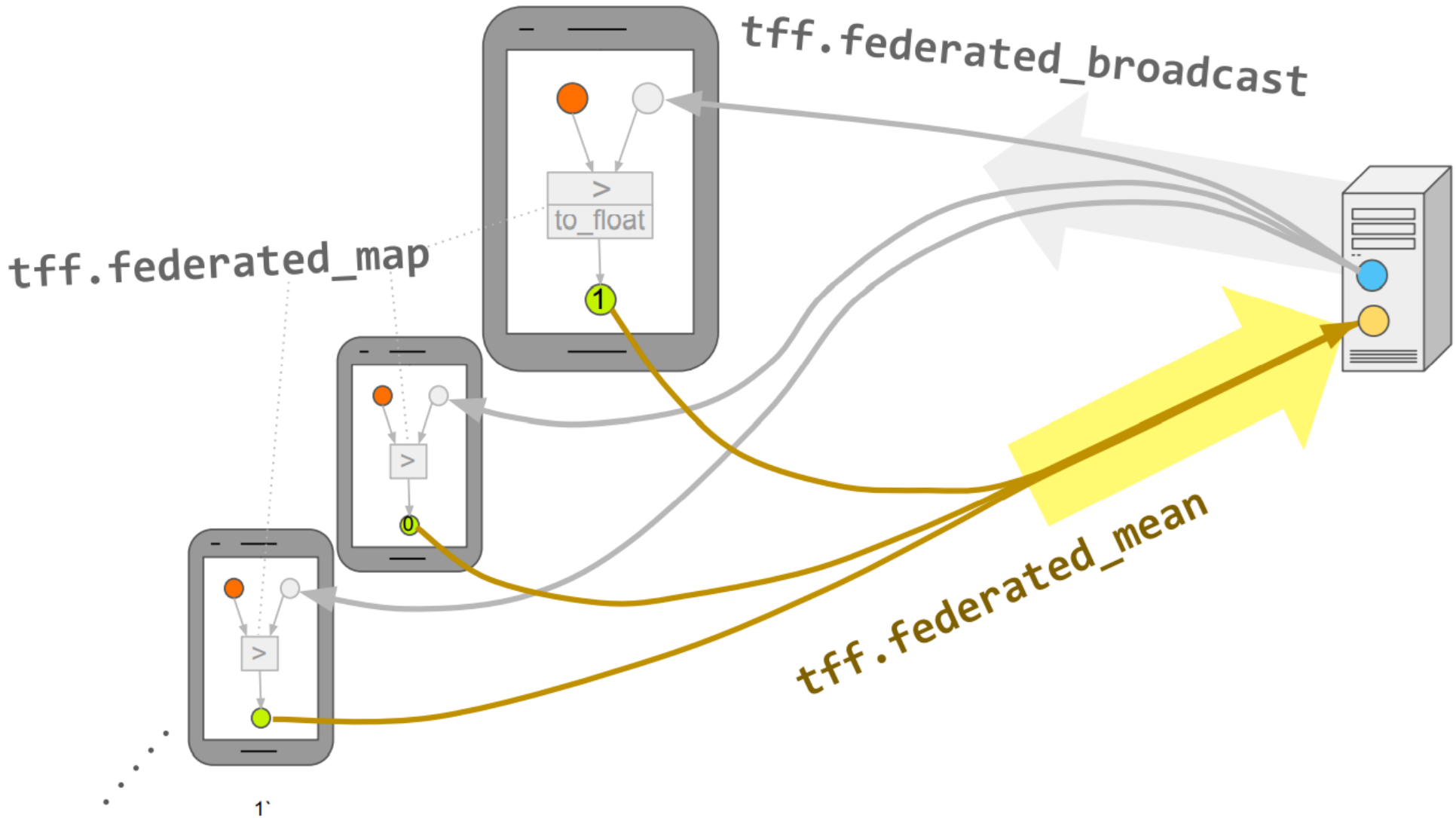
- Esențial pentru a analiza și a evalua metodele federate cu:
  - date non-IID, participare parțială/variabilă



# Hands-On Federated Optimization

# TensorFlow Federated (TFF)

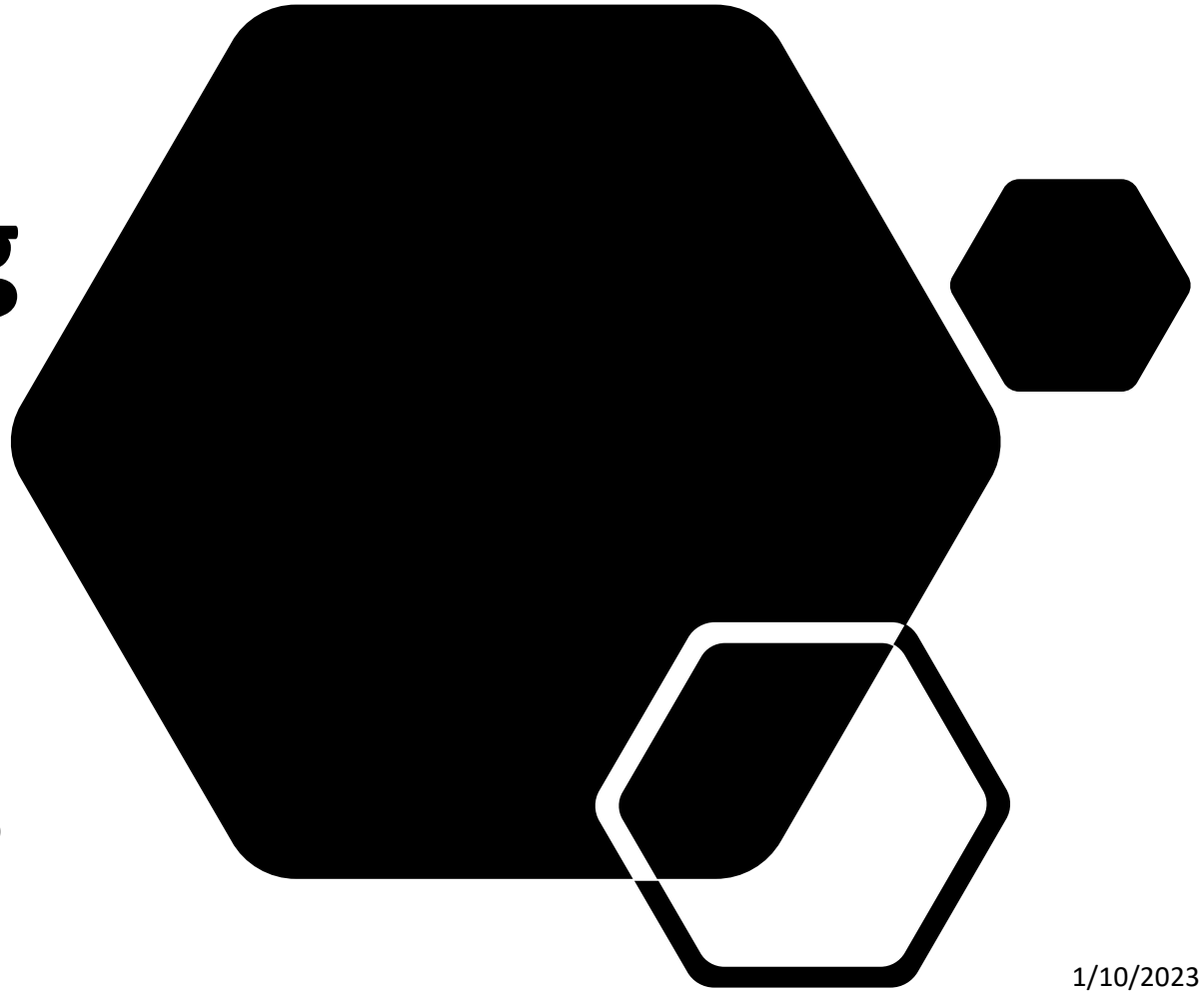

- API declarativ și limbaj pentru definirea calculelor federate
- Timp de rulare al centrului de date distribuit de înaltă performanță, oferind execuție paralelă scalabilă a calculelor complexe per utilizator
- Biblioteci de calcul pentru învățare și analiză federate, implementând algoritmi de optimizare sofisticați și DP agregări prin TF Privacy.


tff.federated_broadcast
tff.federated_map
tff.federated_mean
to_float
1
0
1`

# Privacy for Federated Learning and Analytics

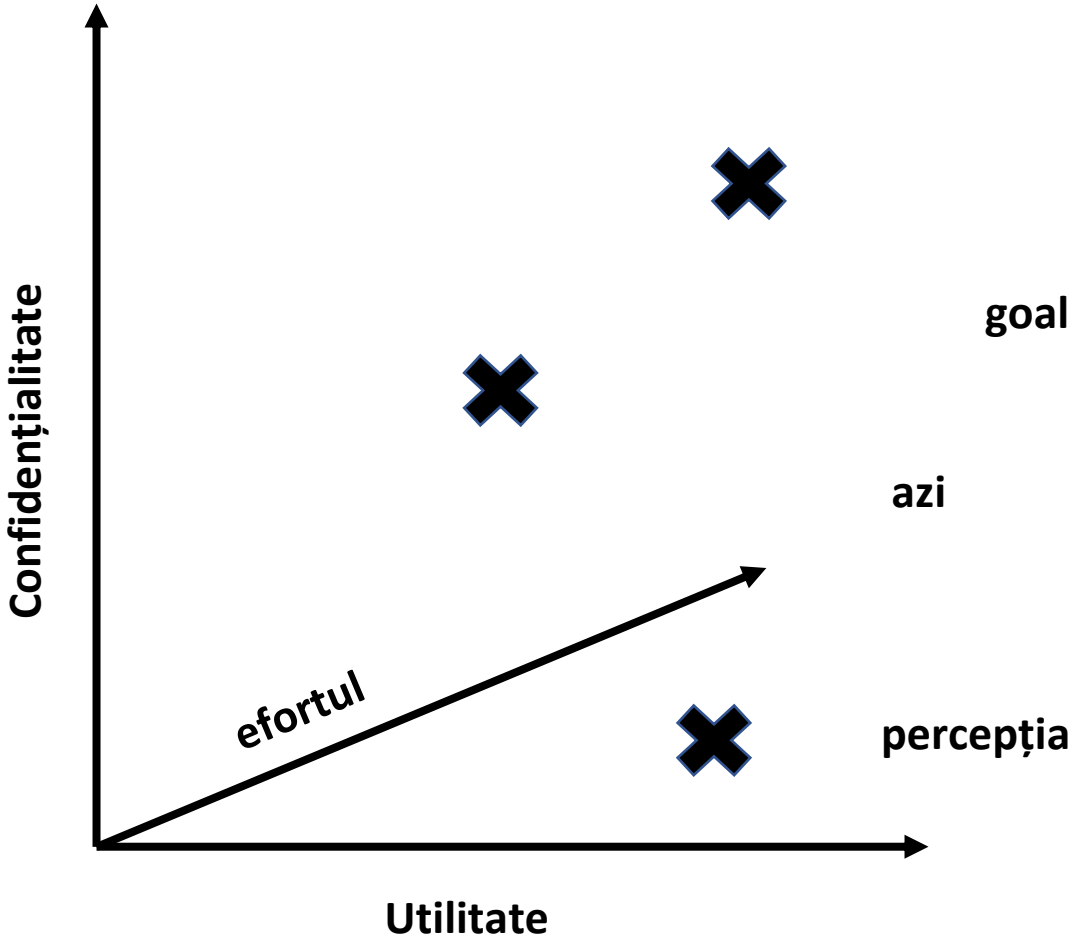
ML pe date sensibile:
confidențialitate vs utilitate

Ce informații private ar putea afla un actor?



1/10/2023

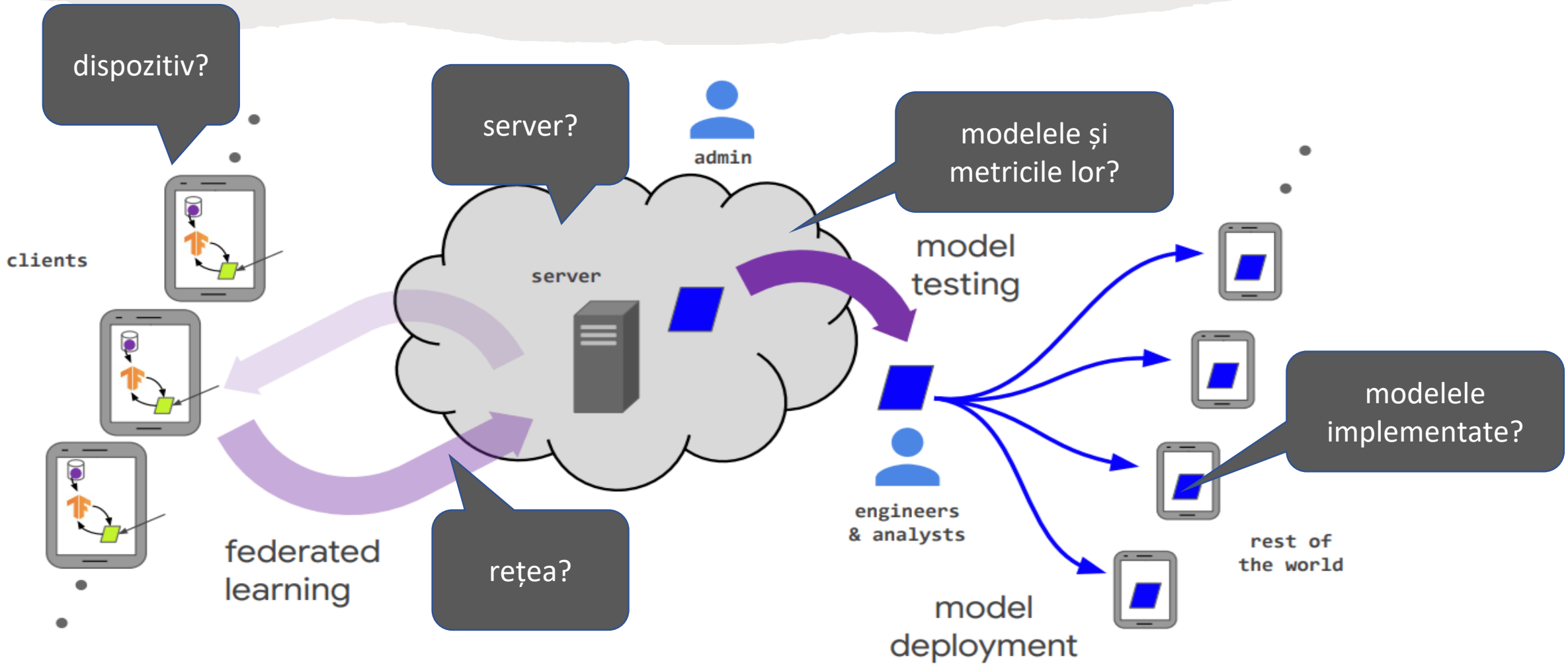
# ML pe date sensibile - confidențialitate vs utilitate



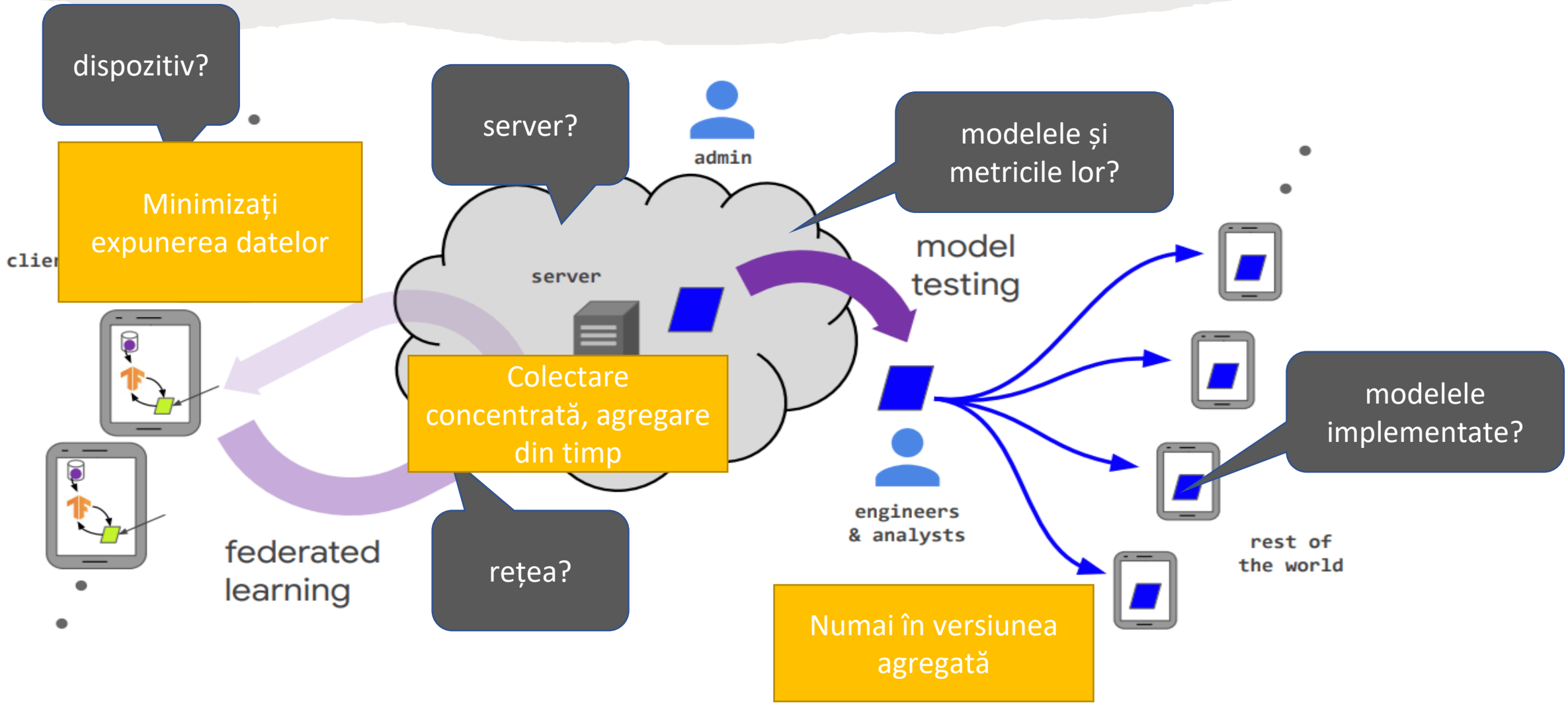

- Împingem frontiera cu o tehnologie mai bună.
- Faceți posibilă obținerea unei confidențialități și utilități ridicate cu mai puțină muncă.

# Ce informații private ar putea afla un actor?

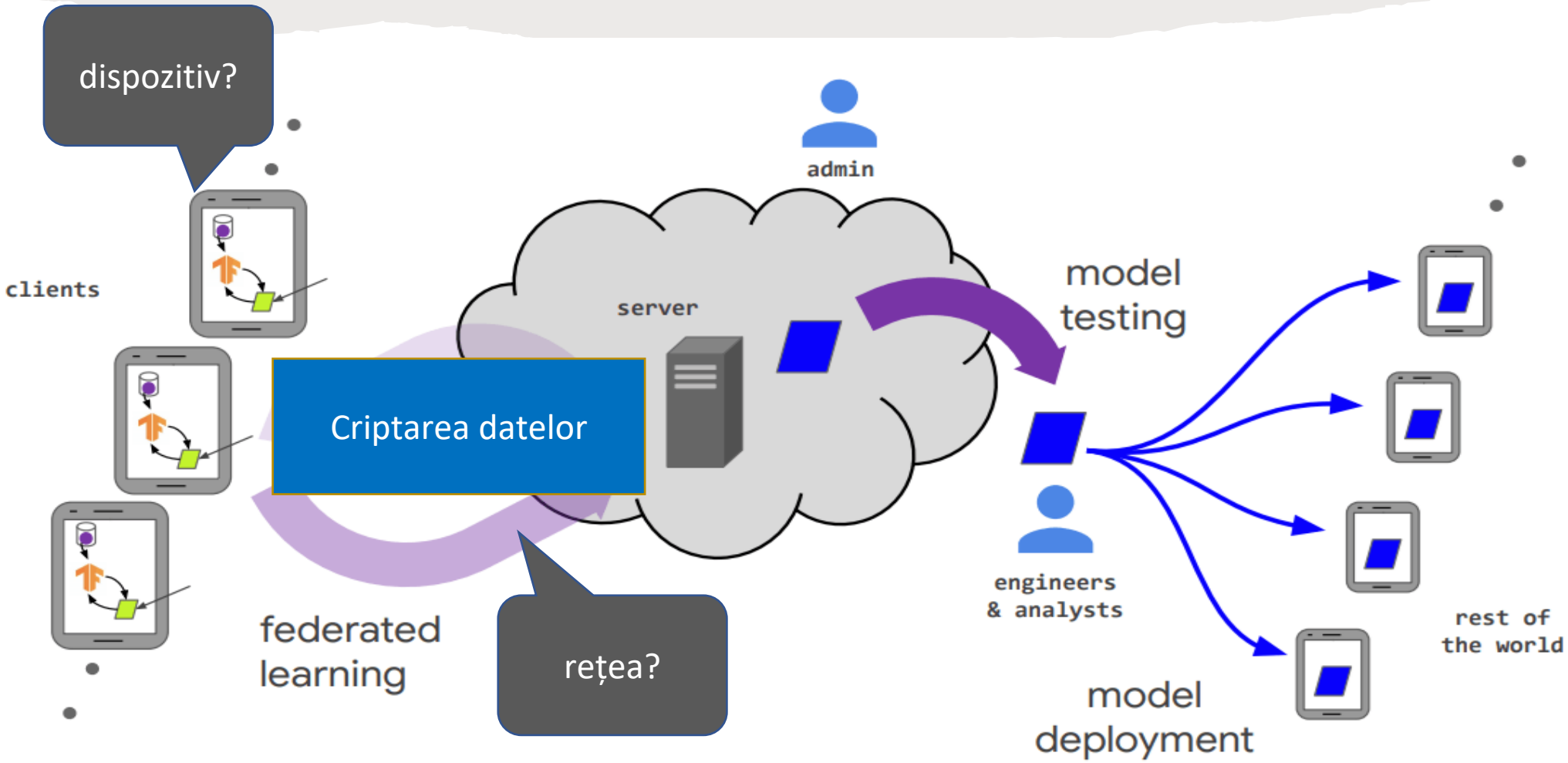
# Ce informații private ar putea afla un actor?
# Câtă încredere trebuie să avem?



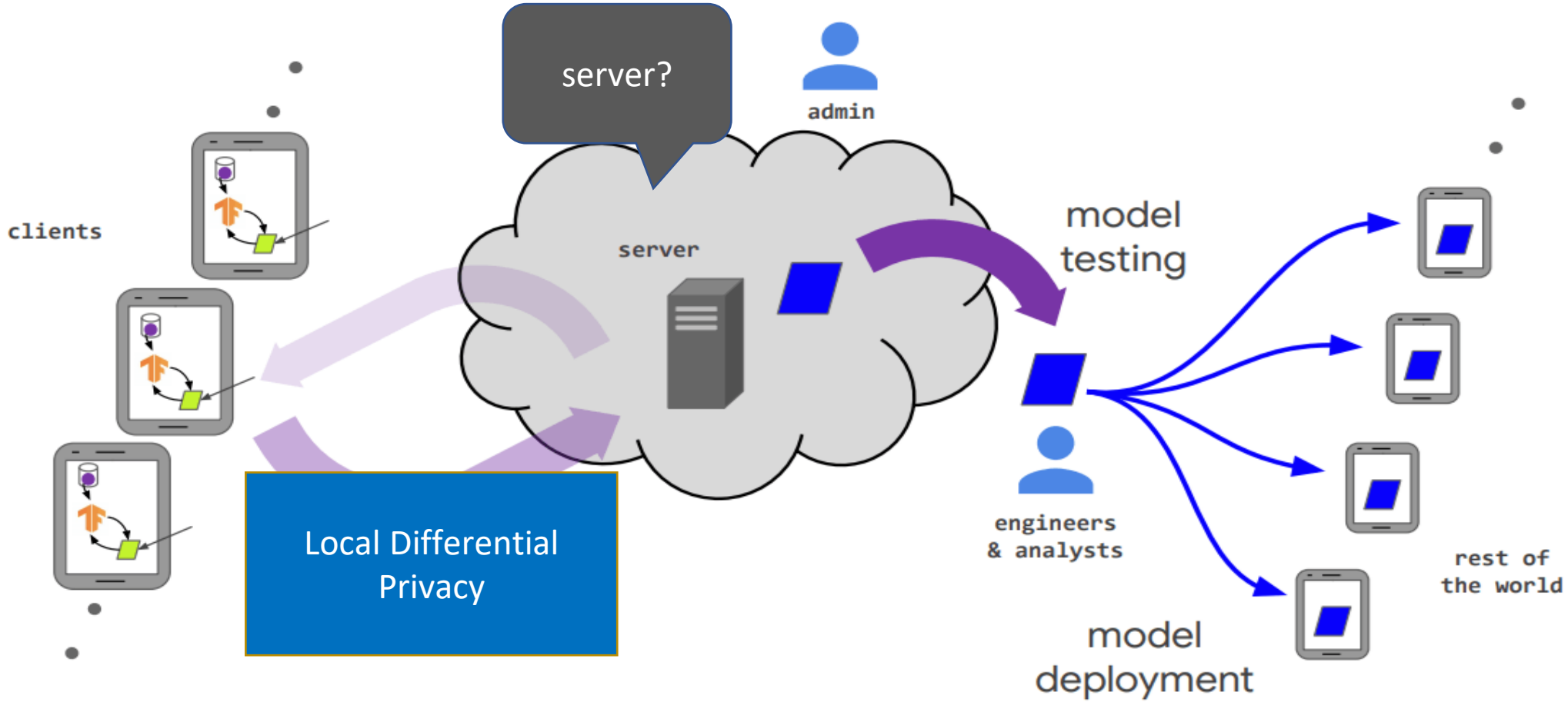
# Principiile de confidențialitate urmărite în FL



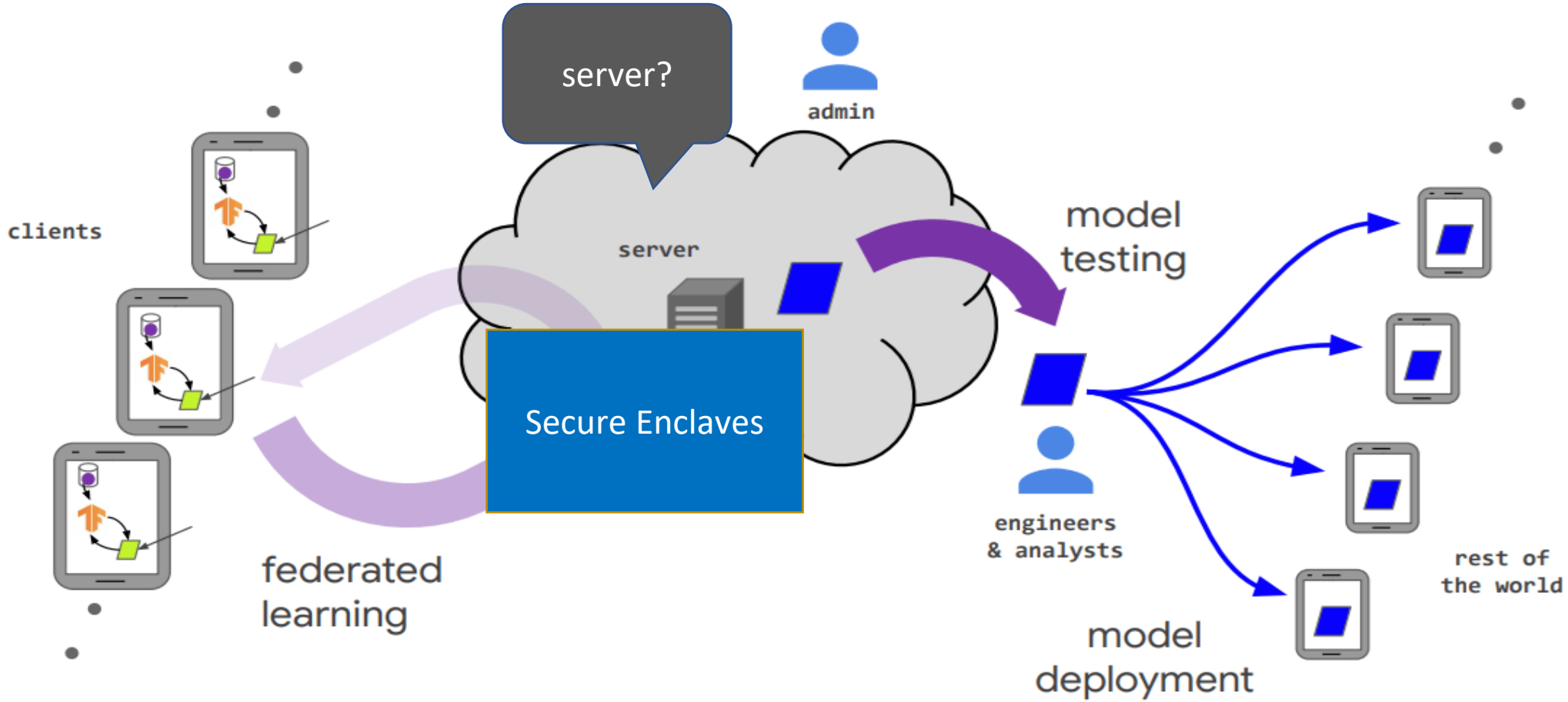
# Principiile de confidențialitate urmărite în FL



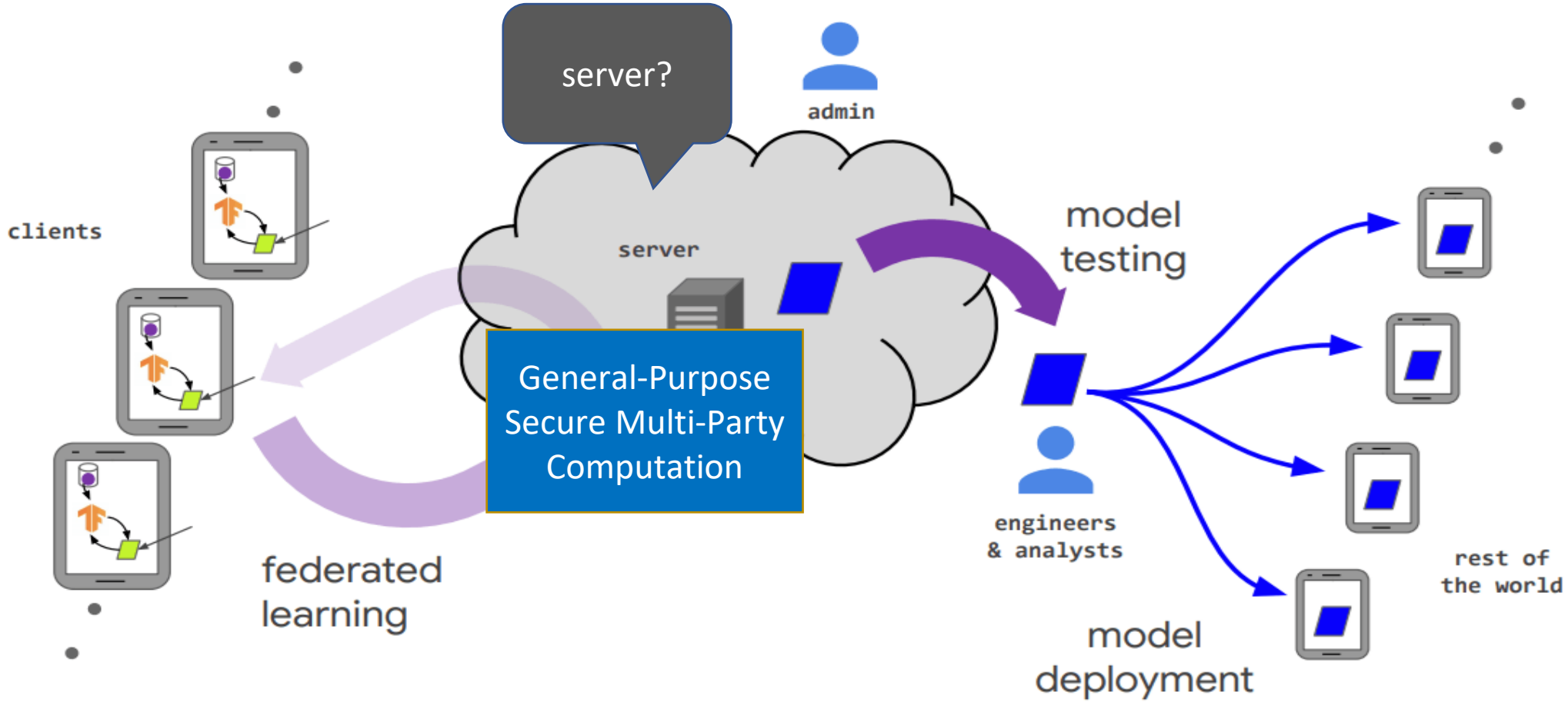
# Principiile de confidențialitate urmărite în FL



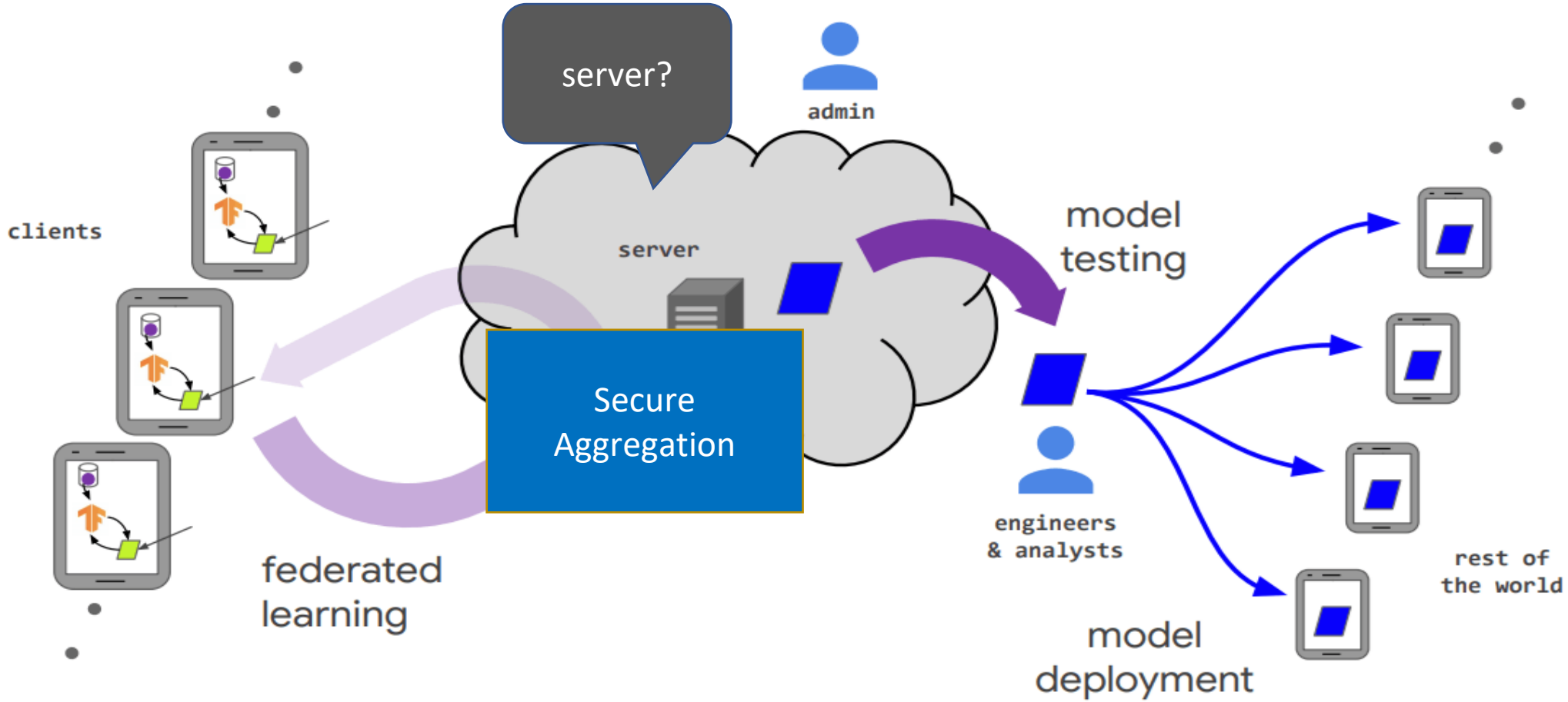
# Principiile de confidențialitate urmărite în FL



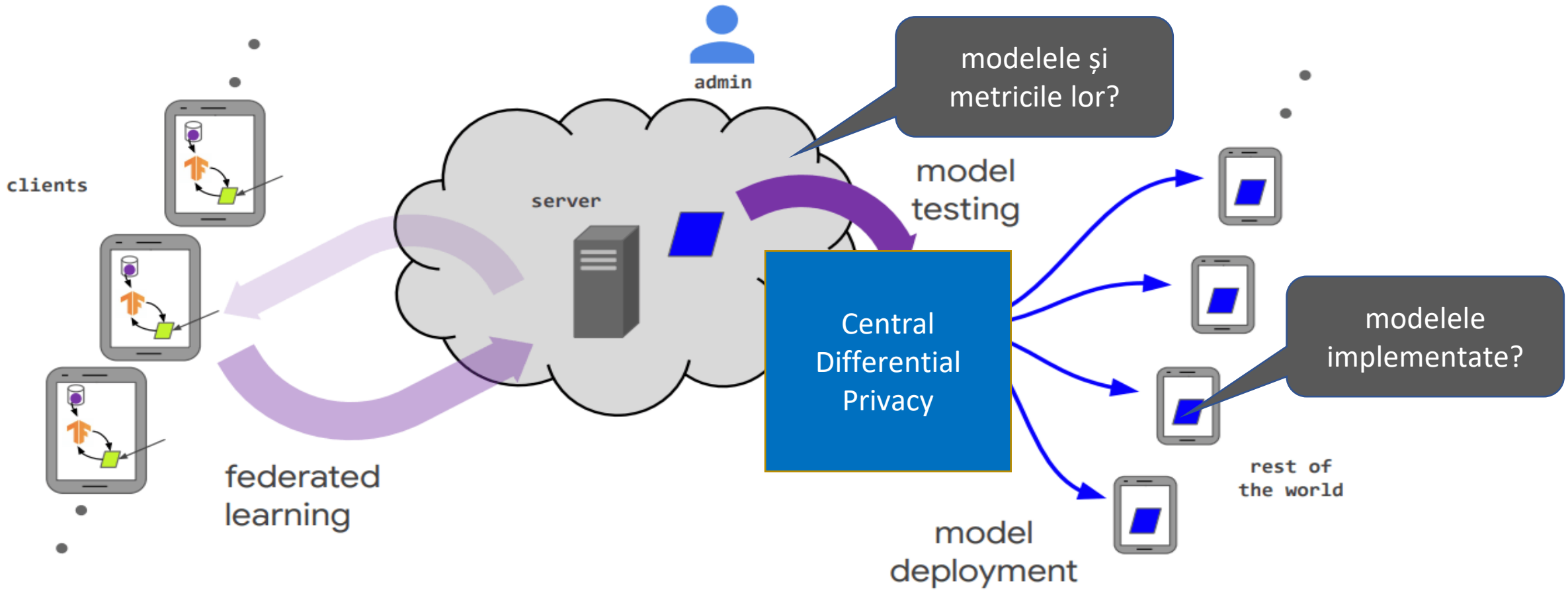
# Principiile de confidențialitate urmărite în FL



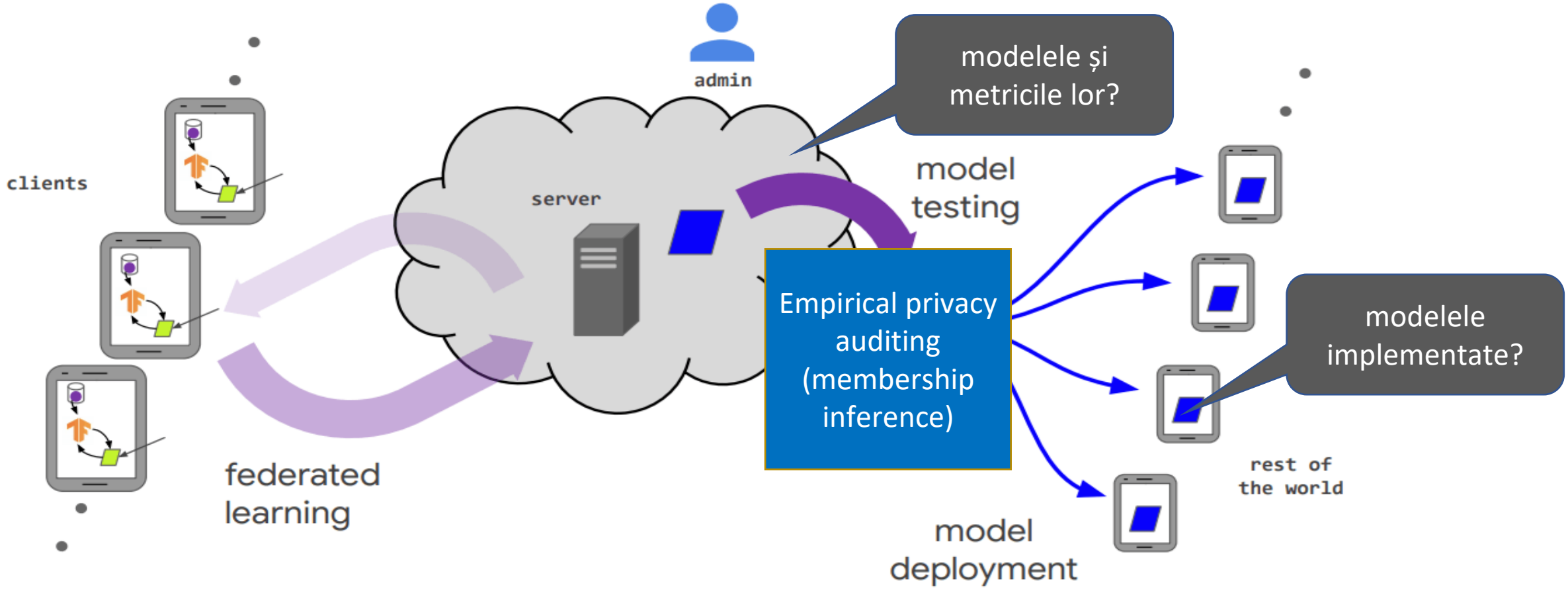
# Principiile de confidențialitate urmărite în FL



# Principiile de confidențialitate urmărite în FL



# Principiile de confidențialitate urmărite în FL



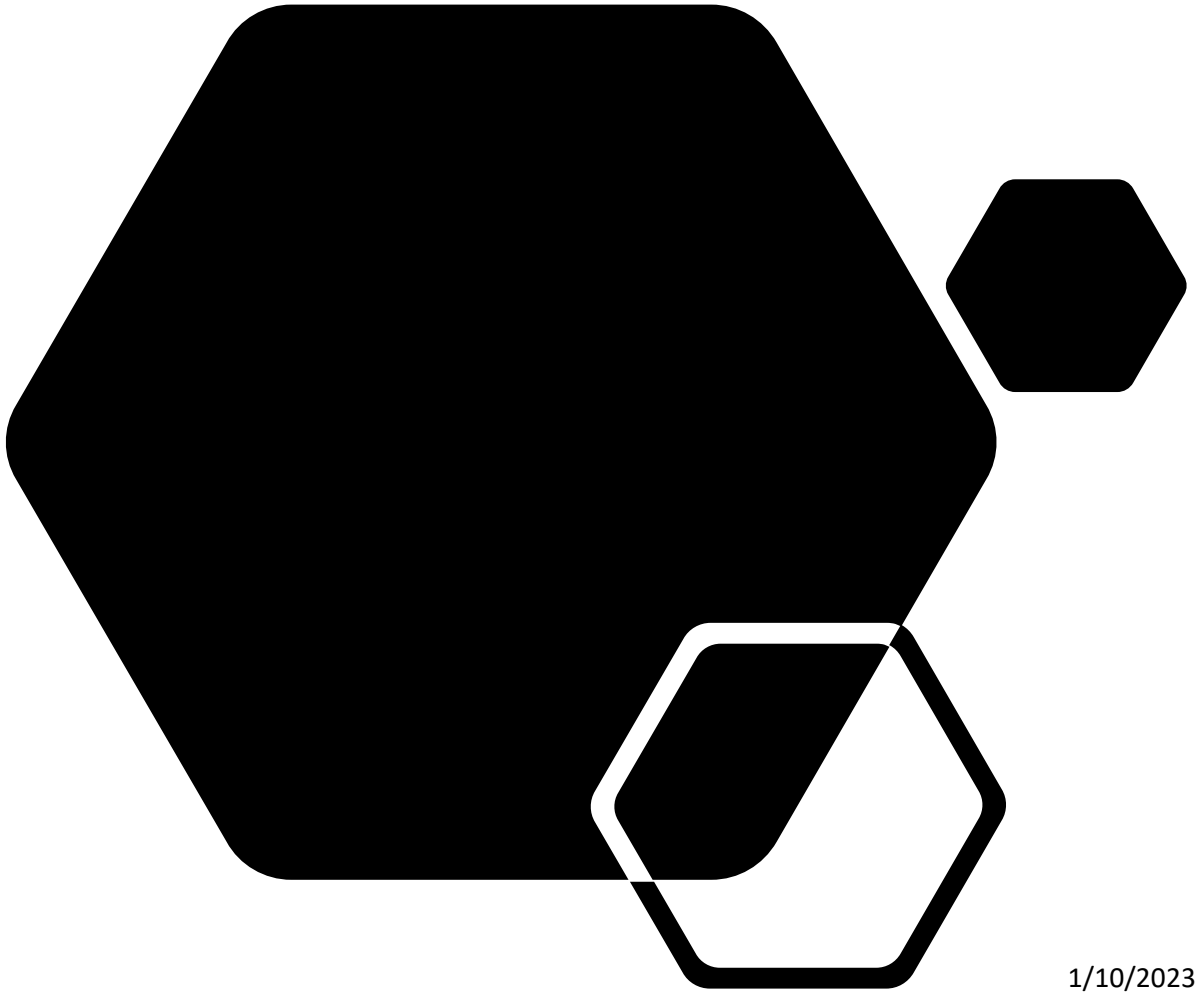
# IV. Probleme deschise și altele subiecte

Îmbunătățirea eficienței și eficacității

Asigurarea echității și abordarea surselor de părtinire

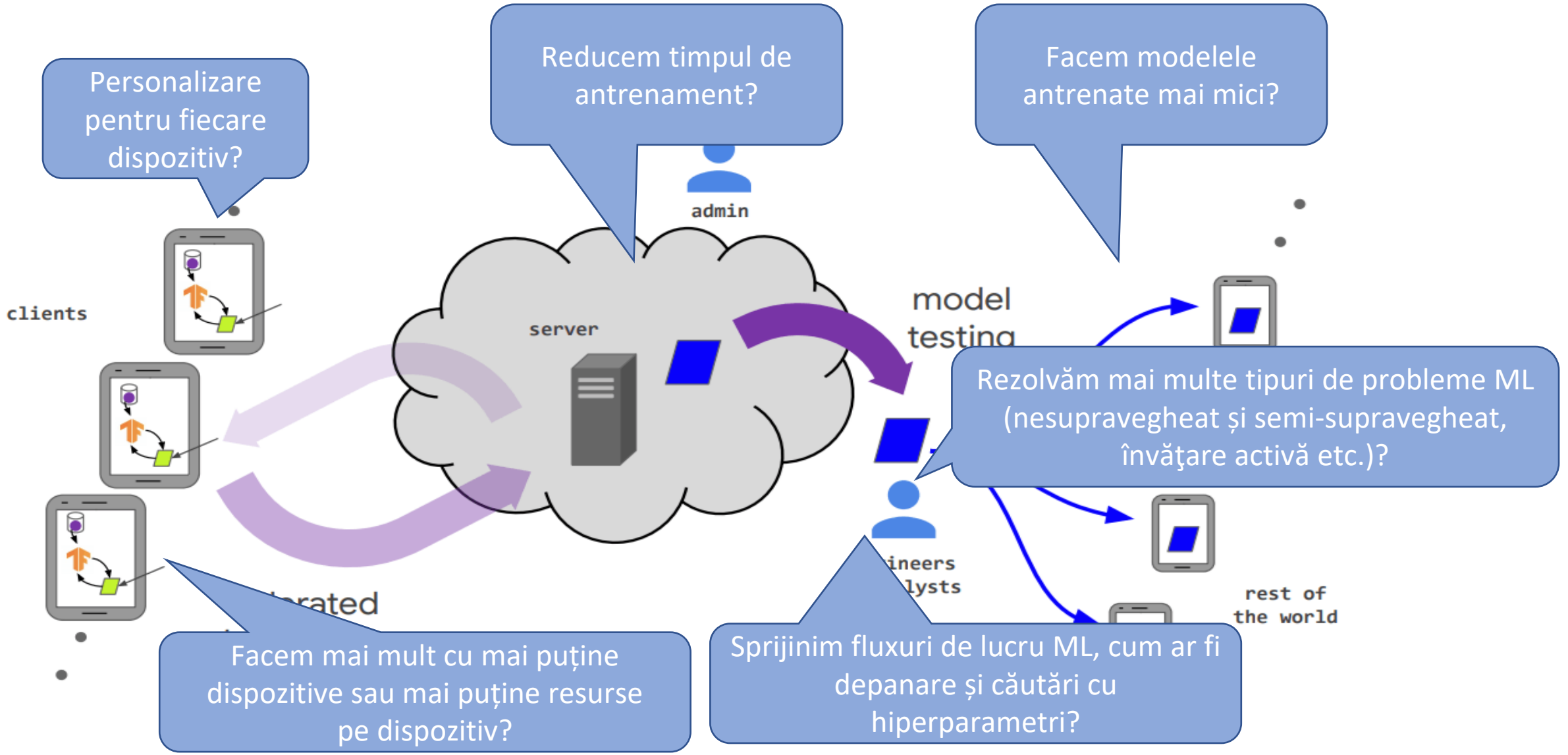
Rezistență la atacuri și eșecuri

Provocări de sistem

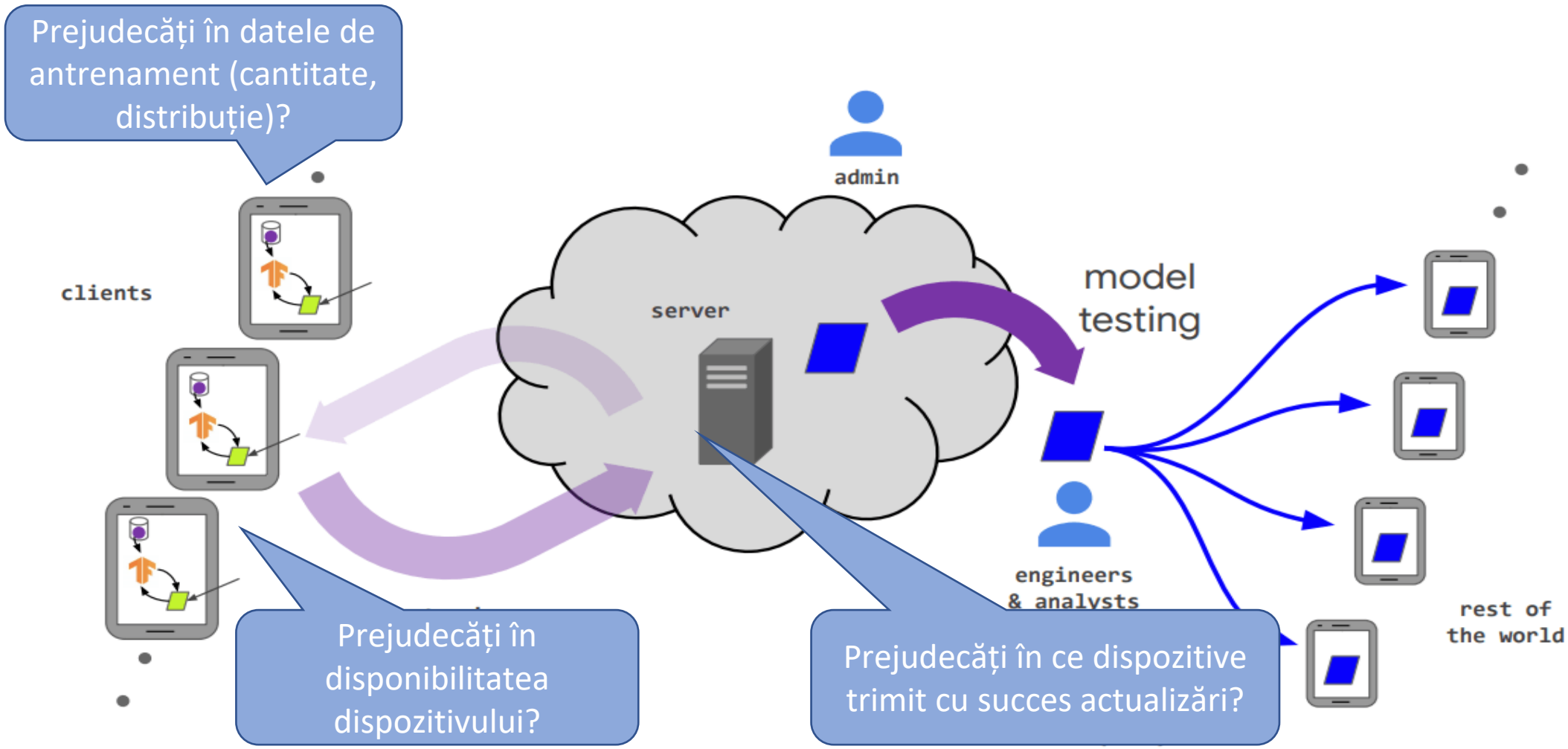


# Îmbunătățirea eficienței și eficacității

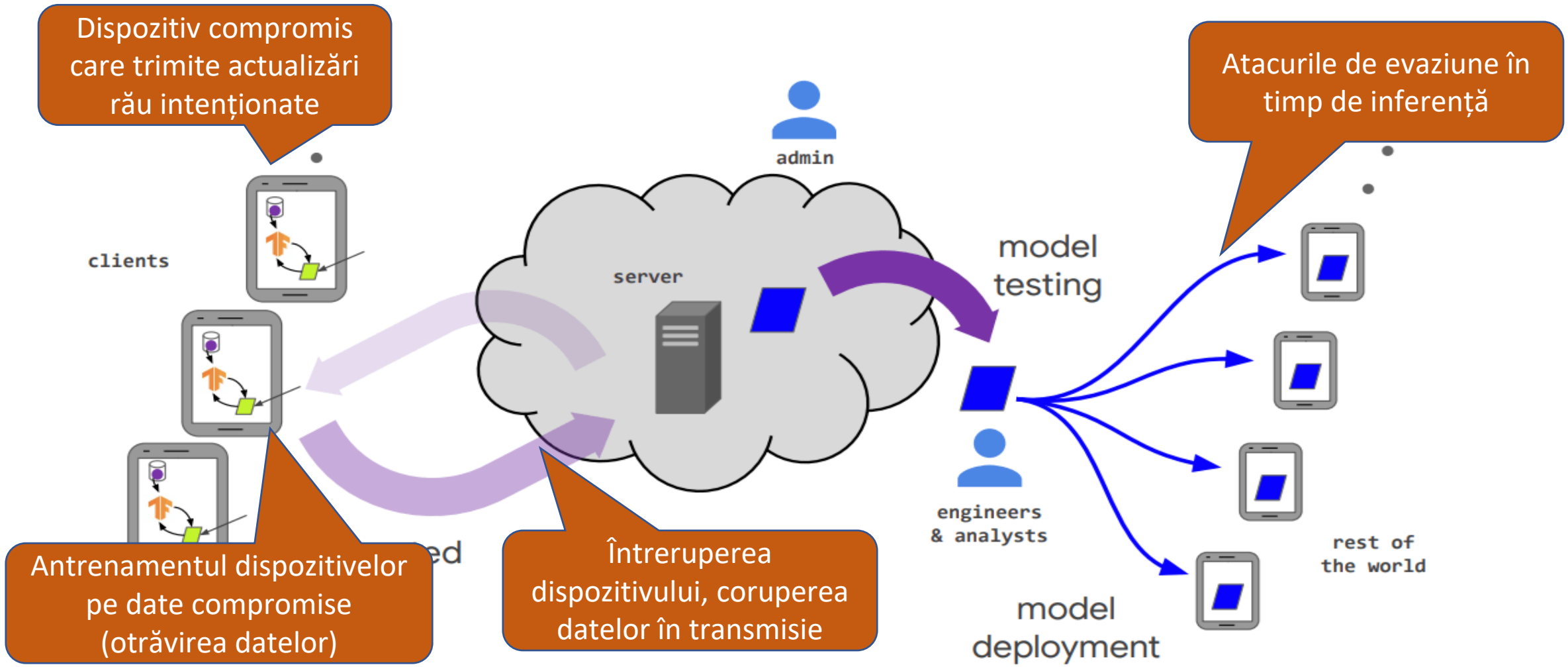
# Îmbunătățirea eficienței și eficacității



# Asigurarea echității și abordarea surselor de părtinire



# Rezistență la atacuri și eșecuri



# Can you backdoor attack federated learning?

- Obiective de atac:
  - **Atacurile bizantine** (neținte):
    - Degradarea performanței la task-ul principală
  - **Atacurile backdoor (țintite)**
    - Introducerea unui task backdoor menținând în același timp o precizie generală bună
- Strategii de atac:
  - **Data poisoning**:
    - Introducerea unor puncte de date otrăvite
  - **Model poisoning**:
    - Trimiterea unor gradienți otrăviți în mod arbitrar

# Provocări de sistem în FL pe mai multe dispozitive

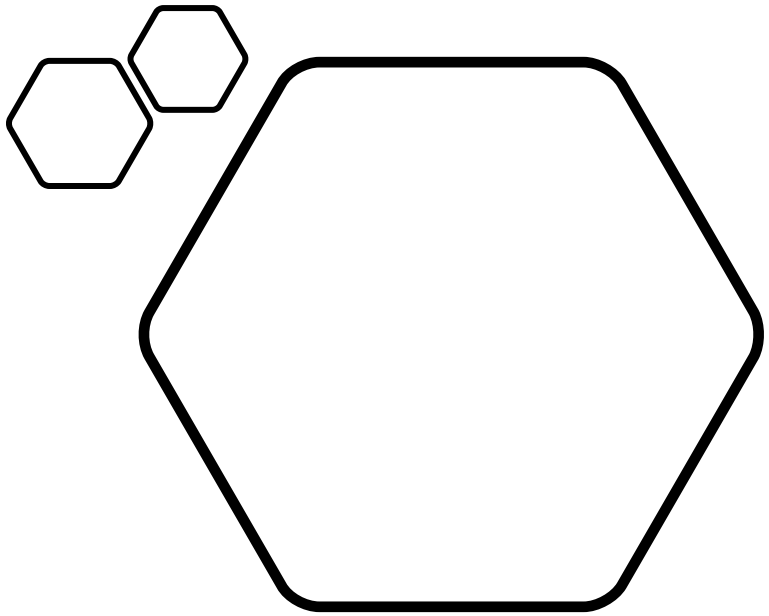
- Masiv paralel, până la $\mathbf{10^{10}}$clienți.
- Doar o fracțiune de clienți sunt disponibili la un moment dat.
- Foarte nesiguri **5%** sau mai mulți dintre clienții care participă la o rundă de calcul sunt de așteptat să eșueze sau să renunțe (de exemplu, deoarece dispozitivul devine neeligibil atunci când cerințele de baterie, rețea sau inactivitate pentru antrenament/calculare sunt încălcate)

**Provocări operaționale**

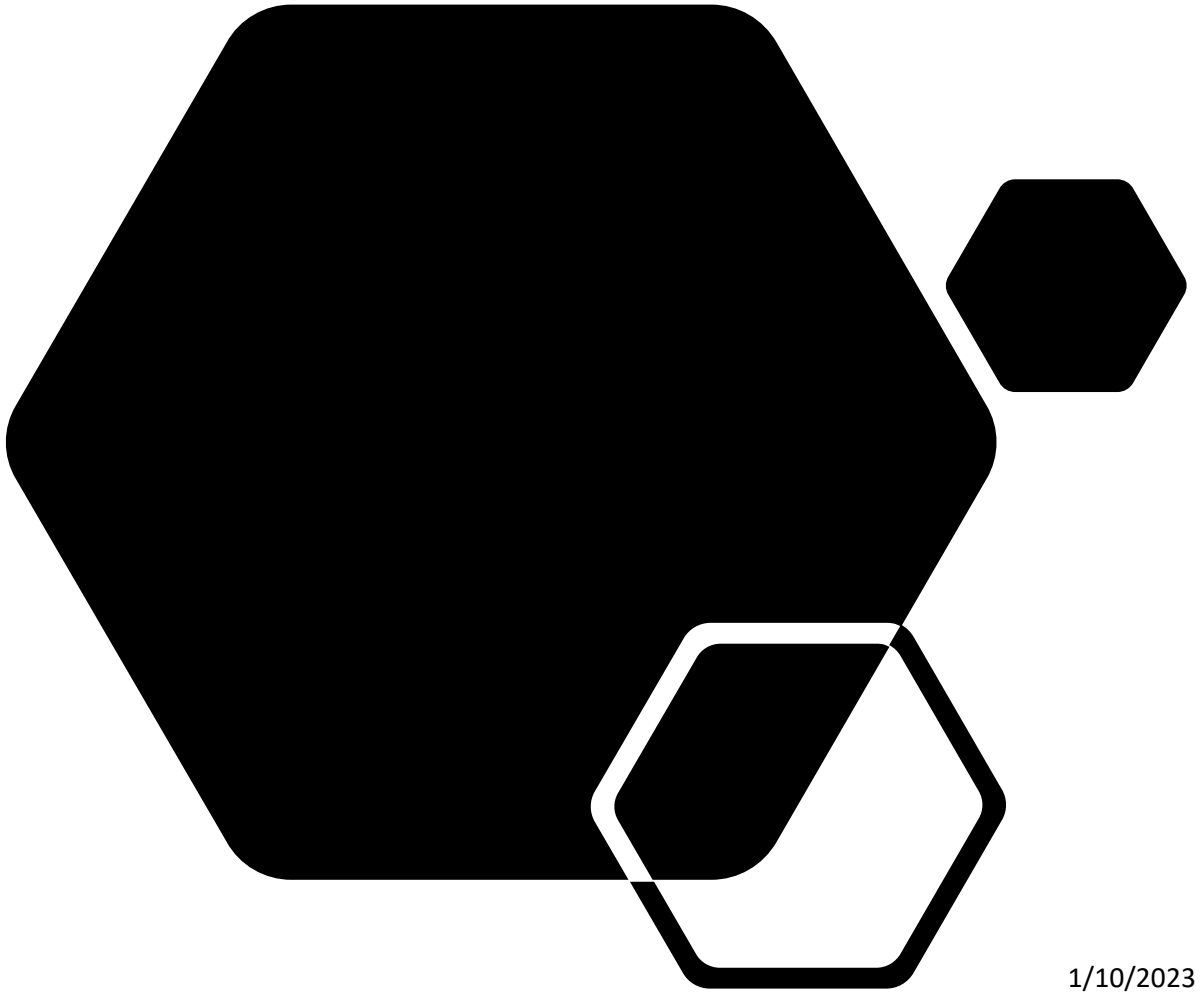
- **Asigurarea compatibilităţii codului** implementat (de exemplu, grafice TensorFlow) cu multiple versiuni ale runtime-ului implementat.
- **Asigurarea stabilității dispozitivelor** utilizatorului în toate circumstanțele: fără impact asupra performanței, bateriei sau utilizarea rețelei.
- **Furnizarea unui serviciu multi-tenant** utilizat de mai multe echipe, fiecare antrenând multe modele.
- **Sprijinirea fluxurilor de lucru**: analiză, depanare model, căutarea hiperparametrului etc

# Take away

- Ce este Federated Learning (FL)?
  - Edge data
  - Cross-Device Federated Learning
  - Federated Analytics
- Federated Optimization
  - Workflow și provocări
  - Federated Averaging (FedAvg)
  - Hands-On Federated Optimization
- Confidențialitate în Federated Learning
  - ML pe date sensibile - confidențialitate vs utilitate
  - Ce informații private ar putea afla un actor?
- Probleme deschise și altele subiecte
  - Îmbunătățirea eficienței și eficacității
  - Asigurarea echității și abordarea surselor de părtinire
  - Rezistență la atacuri și eșecuri
  - Provocări de sistem



# Extra time



# Referințe


- Ramaswamy, et al. Federated Learning for Emoji Prediction in a Mobile Keyboard
- T. Yang, et al. Applied Federated Learning: Improving Google Keyboard Query Suggestions
- M. Chen, et al. Federated Learning Of Out-Of-Vocabulary Words
- UPenn, Intel partner to use federated learning AI for early brain tumor detection
- Medical Institutions Collaborate to Improve Mammogram Assessment AI with NVIDIA Clara Federated Learning
- McMahan, et al. Communication-Efficient Learning of Deep Networks from Decentralized Data. AISTATS 2017
- Federated Learning: Challenges, Methods, and Future Directions

- tensorflow.org/federated
- github.com/tensorflow/federated
- Advances and Open Problems in Federated Learning
- Federated Learning: Challenges, Methods, and Future Directions