

# Federated Learning

A collaboration on anti-money laundering

Edvin Callisen  
Research Engineer, AI Sweden  
[edvin.callisen@ai.se](mailto:edvin.callisen@ai.se)

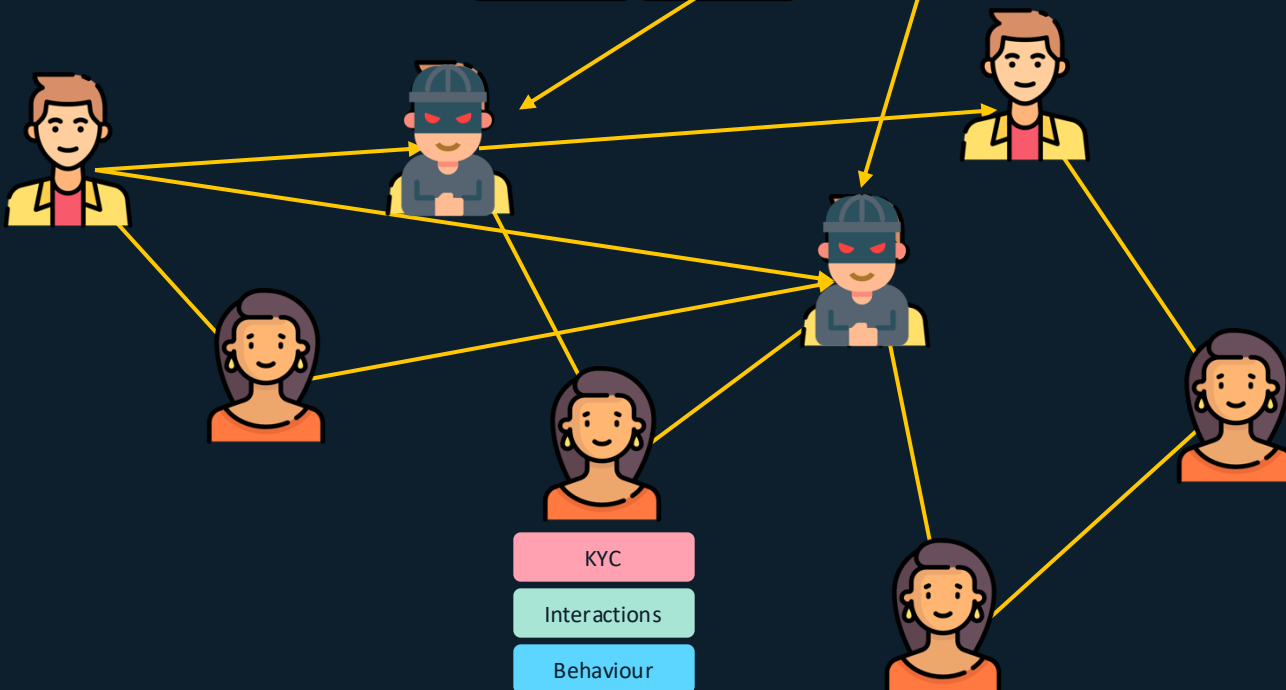
# Agenda

- The anti-money laundering (AML) problem
- Our holistic pipeline to attacking the AML problem
- Synthetic data
- Federated Learning
  - Requirements
  - Challenges
  - Opportunities
- Preliminary results

# The Money Laundering Problem



Want to identify these!



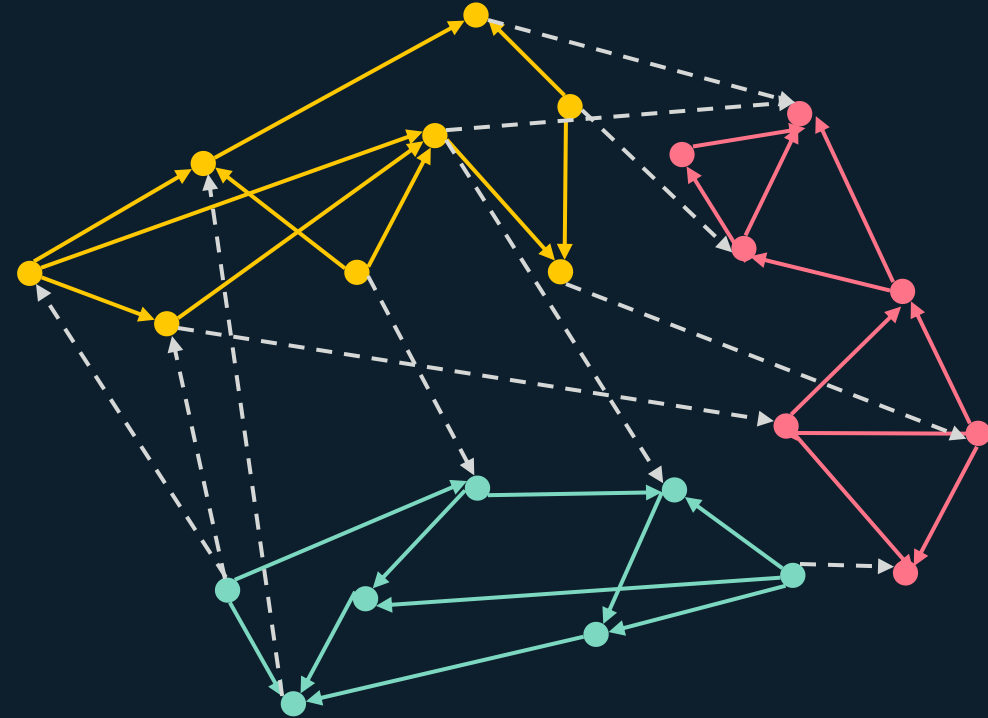
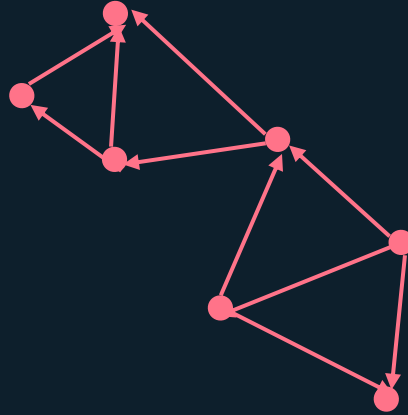
- ! UN estimates €1.87 trillion laundered annually
- ! Significant resources are spent on AML
- ! High false-positive rate, estimated at ~98%
- ! In 2019, 58 cases of AML penalties total fines: \$8.14 billion.

# A typical AML pipeline<sup>17</sup>



<sup>17</sup> F. Johannesen and M. Jullum., "Finding Money Launderers Using Heterogeneous Graph Neural Networks", arXiv:2307.13499, 2023

# The Global Transactional Network



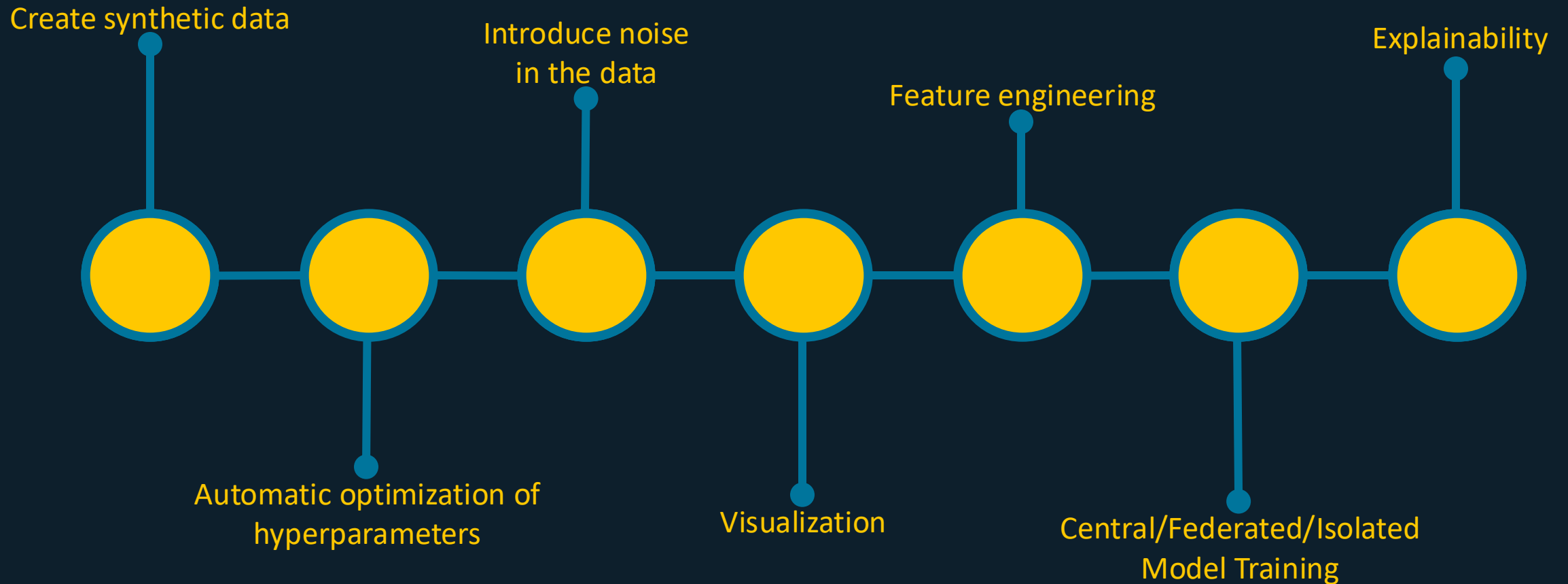
**Banks are blind to what happens outside the local transaction network!**

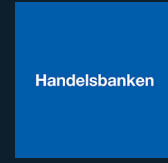
# We ask ourselves:



- ? Can federated learning be used to collaborate between banks to detect money laundering?
- ? How can we collaborate without sharing sensitive data?
- ? How to federate the learning in practice?
- ? How to create realistic synthetic transaction data?
- ? What impact does unreliable data have?
- ? Does performance improve by exploiting the inherent graph-structure in the data?
- ? Can we use explainable AI to leverage deep learning?

# A Holistic Pipeline

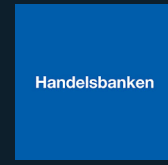




# Synthetic Data

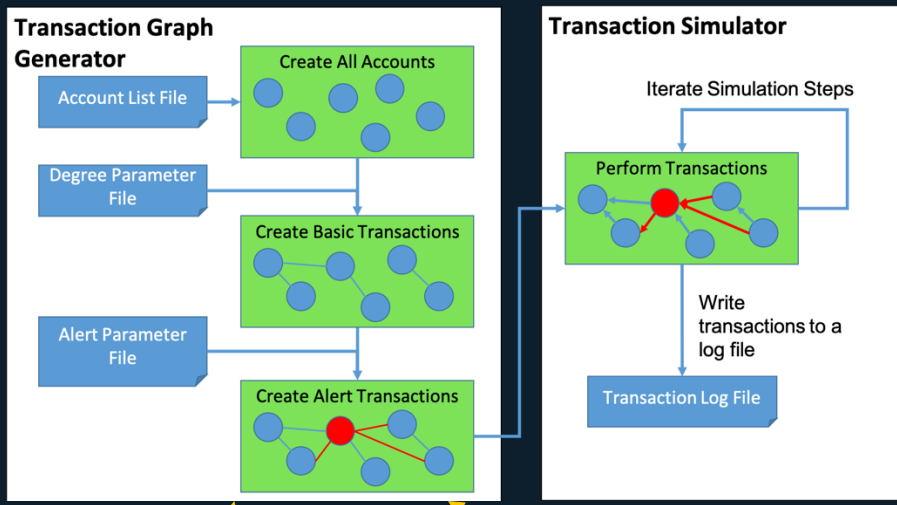


# Synthetic Data: Engine<sup>18</sup>

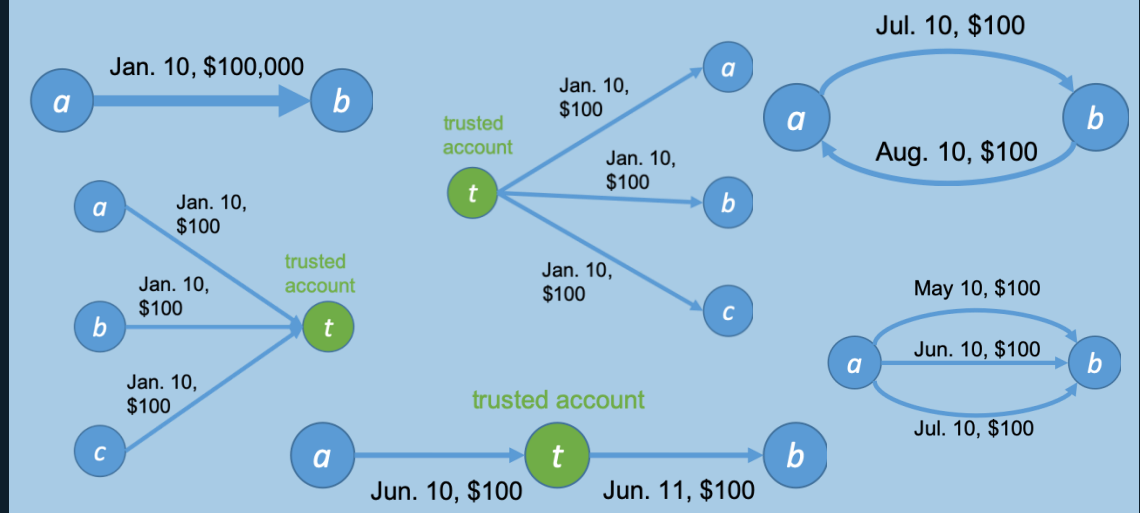


Python

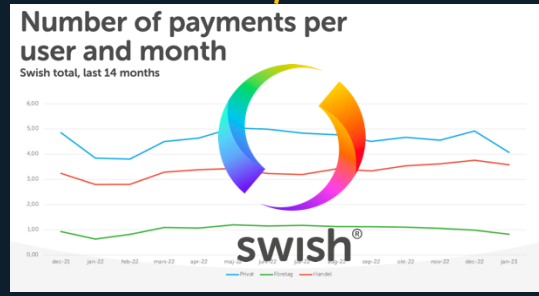
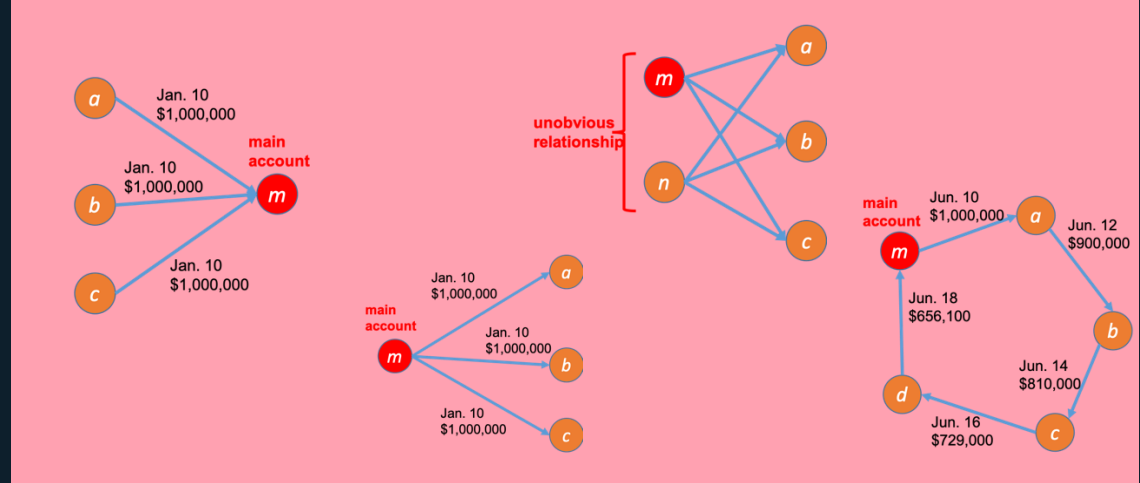
Java



## Normal Transactions



## Alert Transactions



Genomsnittlig månadslön 1973–2022

| År   | Privat sektor |         |        | Offentlig sektor |         |        | Regionaltillda |         |        |
|------|---------------|---------|--------|------------------|---------|--------|----------------|---------|--------|
|      | Samtliga      | Kvinnor | Män    | Samtliga         | Kvinnor | Män    | Samtliga       | Kvinnor | Män    |
| 2022 | 50 700        | 47 200  | 52 300 | 38 900           | 34 100  | 41 300 | 39 900         | 47 700  | 46 600 |
| 2021 | 48 800        | 45 400  | 50 300 | 38 900           | 33 400  | 40 200 | 38 300         | 46 600  | 45 200 |
| 2020 | 47 400        | 44 100  | 48 900 | 38 900           | 32 600  | 38 900 | 37 100         | 45 200  | 44 800 |
| 2019 | 47 100        | 43 400  | 48 600 | 38 900           | 32 000  | 38 300 | 36 400         | 44 800  | 44 400 |
| 2018 | 46 300        | 42 200  | 48 000 | 38 900           | 31 200  | 37 400 | 35 500         | 44 400  | 43 600 |
| 2017 | 44 700        | 40 800  | 46 300 | 36 900           | 29 800  | 30 500 | 34 600         | 43 600  | 42 700 |
| 2016 | 44 000        | 40 100  | 45 600 | 36 000           | 28 700  | 29 700 | 33 700         | 42 700  | 42 200 |
| 2015 | 43 800        | 39 400  | 45 400 | 35 200           | 27 700  | 29 000 | 34 800         | 42 200  | 41 600 |
| 2014 | 42 700        | 38 100  | 44 400 | 34 500           | 26 900  | 27 200 | 33 900         | 41 600  | 41 600 |

<sup>18</sup> M. Weber, et al., "Scalable Graph Learning for Anti-Money Laundering: A First Look", arXiv1812:00076, 2018

# Synthetic data

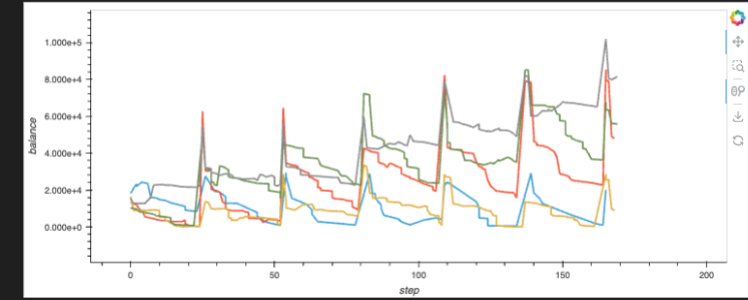
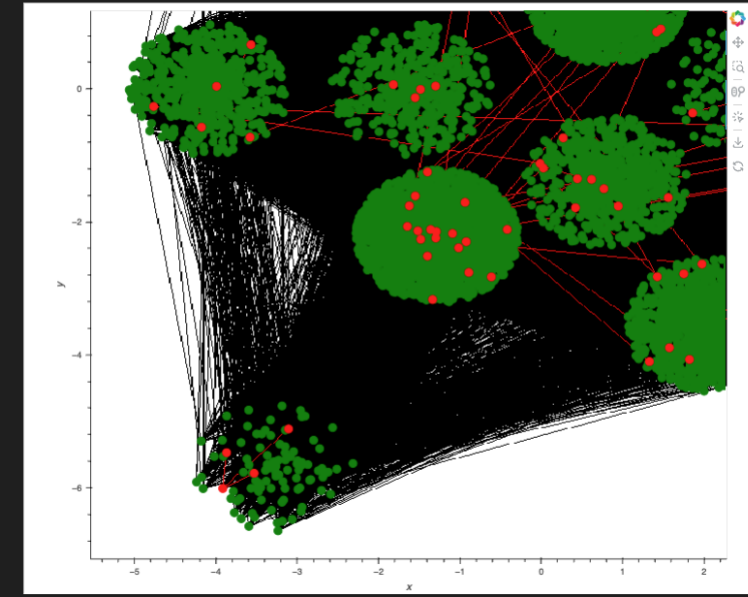
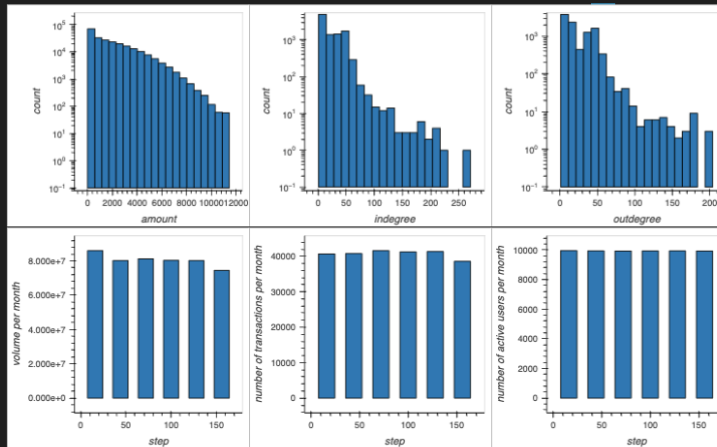
523843 14, TRANSFER, 506.1, 11637, c1, 14, 0, 21168.38, 20662.27, -1, s  
 523844 14, TRANSFER, 466.51, 26034, c2, 14, 0, 78507.9, 78041.39, -1, s  
 523845 14, TRANSFER, 447.33, 25228, c2, 14, 0, 93403.85, 92956.53, -1, s  
 523846 14, TRANSFER, 492.5, 14717, c1, 14, 0, 165702.42, 165209.92, -1, s  
 523847 14, TRANSFER, 420.99, 12423, c1, 14, 0, 169737.06, 169316.06, -1, s  
 523848 14, TRANSFER, 648.72, 20907, c2, 14, 0, 146715.98, 146067.25, -1, s  
 523849 14, TRANSFER, 552.32, 5237, c0, 14, 0, 173834.85, 173282.53, -1, s  
 523850 14, TRANSFER, 457.76, 26947, c2, 14, 0, 24121.91, 23664.15, -1, s  
 523851 14, TRANSFER, 506.6, 22503, c2, 14, 0, 75712.35, 75205.75, -1, s  
 523852 14, TRANSFER, 401.9, 6015, c0, 14, 0, 35482.05, 35080.14, -1, s  
 523853 14, TRANSFER, 472.86, 11827, c1, 14, 0, 137134.37, 136661.51, -1, s  
 523854 14, TRANSFER, 481.02, 24100, c2, 14, 0, 133435.03, 132954.0, -1, s  
 523855 14, TRANSFER, 358.65, 13327, c1, 14, 0, 20303.14, 19944.48, -1, s  
 523856 14, TRANSFER, 564.29, 24286, c2, 14, 0, 131110.93, 130546.64, -1, s  
 523857 14, TRANSFER, 629.17, 3181, c0, 14, 0, 117720.02, 117090.85, -1, s  
 523858 14, TRANSFER, 502.32, 13674, c1, 14, 0, 176062.34, 175560.01, -1, s  
 523859 15, TRANSFER, 742.65, 20001, c2, 15, 0, 155026.1, 154283.45, -1, s  
 523860 15, TRANSFER, 442.78, 28673, c2, 15, 0, 86478.2, 86035.42, -1, s  
 523861 15, TRANSFER, 916.25, 16491, c1, 16, 0, 33273.07, 32356.81, 2876  
 523862 15, TRANSFER, 892.88, 16491, c1, 16, 0, 32356.81, 31463.92, 1076  
 523863 15, TRANSFER, 429.28, 14088, c1, 15, 0, 36964.7, 36535.42, -1, s  
 523864 15, TRANSFER, 367.54, 17825, c1, 15, 0, 191650.31, 191282.76, -1, s  
 523865 15, TRANSFER, 412.17, 11147, c1, 15, 0, 97173.57, 96761.4, -1, s  
 523866 15, TRANSFER, 603.85, 22293, c2, 15, 0, 54695.99, 54092.13, -1, s  
 523867 15, TRANSFER, 537.3, 962, c0, 15, 0, 114621.55, 114084.24, -1, s  
 523868 15, TRANSFER, 597.94, 18919, c1, 15, 0, 179478.18, 178880.25, -1, s  
 523869 15, TRANSFER, 475.55, 10360, c1, 15, 0, 123899.24, 123423.68, -1, s  
 523870 15, TRANSFER, 522.88, 29476, c2, 15, 0, 142923.57, 142400.68, -1, s  
 523871 15, TRANSFER, 348.32, 27118, c2, 15, 0, 153985.15, 153636.82, -1, s  
 523872 15, TRANSFER, 396.5, 14500, c1, 15, 0, 43844.47, 43447.97, -1, s  
 523873 15, TRANSFER, 606.94, 6187, c0, 15, 0, 72482.49, 71875.53, -1, s  
 523874 15, TRANSFER, 428.79, 14494, c1, 15, 0, 183155.53, 182726.73, -1, s  
 523875 15, TRANSFER, 405.0, 1223, c0, 15, 0, 54502.5, 54097.5, -1, sink  
 523876 15, TRANSFER, 411.7, 13076, c1, 15, 0, 152069.92, 151658.21, -1, s  
 523877 15, TRANSFER, 699.92, 20975, c2, 15, 0, 30828.94, 30129.01, -1, s  
 523878 15, TRANSFER, 563.27, 2671, c0, 15, 0, 88612.42, 88049.15, -1, s  
 523879 15, TRANSFER, 592.45, 2671, c0, 16, 0, 88049.15, 87456.7, 29977  
 523880 15, TRANSFER, 522.83, 2407, c0, 15, 0, 153886.06, 153363.23, -1, s  
 523881 15, TRANSFER, 521.57, 21546, c2, 15, 0, 122583.67, 122062.09, -1, s  
 523882 15, TRANSFER, 597.84, 12364, c1, 15, 0, 188310.23, 187712.39, -1, s  
 523883 15, TRANSFER, 389.11, 9528, c0, 15, 0, 71826.74, 71437.62, -1, s  
 523884 15, TRANSFER, 490.31, 18588, c1, 15, 0, 42782.89, 42292.58, -1, s  
 523885 15, TRANSFER, 666.58, 20877, c2, 15, 0, 192097.37, 191430.79, -1, s  
 523886 15, TRANSFER, 297.38, 26633, c2, 15, 0, 80318.29, 80020.9, -1, s  
 523887 15, TRANSFER, 469.08, 11921, c1, 15, 0, 27181.41, 26712.32, -1, s



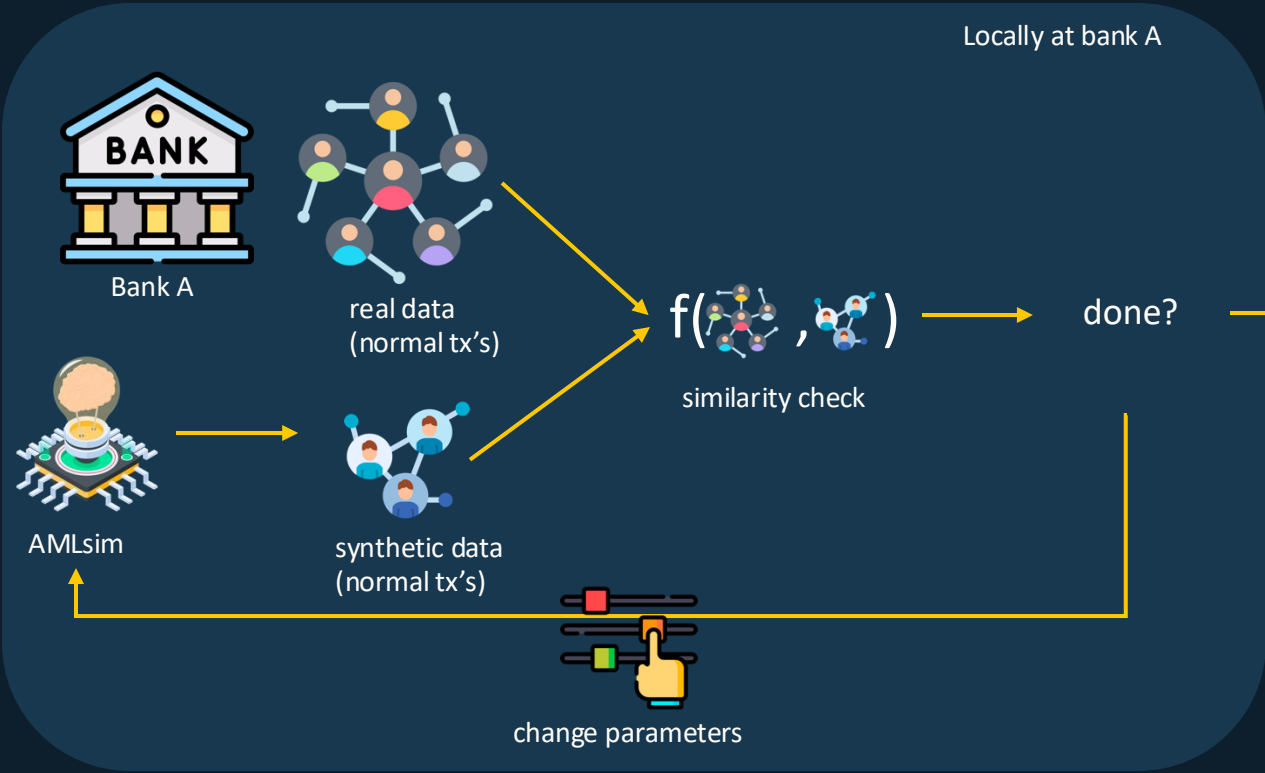
## Transaction Network Explorer

| Information                                     | Banks                      | Legitimate models   | Laundering models                |
|---|----------------------------|---------------------|----------------------------------|
| total number of accounts: 10000                 | nordea<br>länsförsäkringar | single<br>fan-out   | fan-out<br>fan-in                |
| total number of legitimate transactions: 241557 | ica<br>sparbanken          | fan-in<br>forward   | cycle<br>bipartite               |
| total number of laundering transactions: 75     | swedbank<br>skandia        | mutal<br>periodical | stacked<br>random                |
| steps: 1 - 170                                  | svea<br>handelsbanken      |                     | scatter-gather<br>gather-scatter |
| Edge homophily: 0.9893                          | marginalen                 |                     |                                  |
| Node homophily: 0.9914                          | danske<br>seb              |                     |                                  |
| Class homophily: 0.0297                         | älandsbanken               |                     |                                  |

### Steps



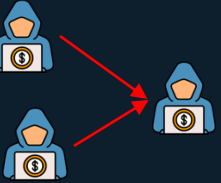
# Synthetic Data: Boosting Fidelity



Bank B

synthetic data

$g(\cdot, \cdot)$   
similarity check



AML patterns

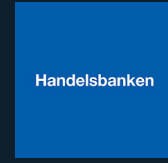


Bank A



Bank B

consistent  
AML patterns



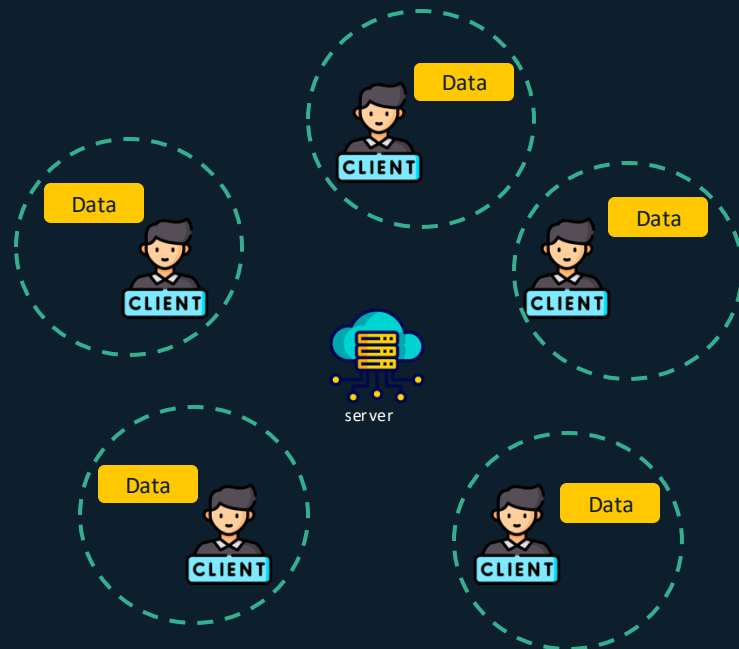
# Federated Learning

# Federated Learning



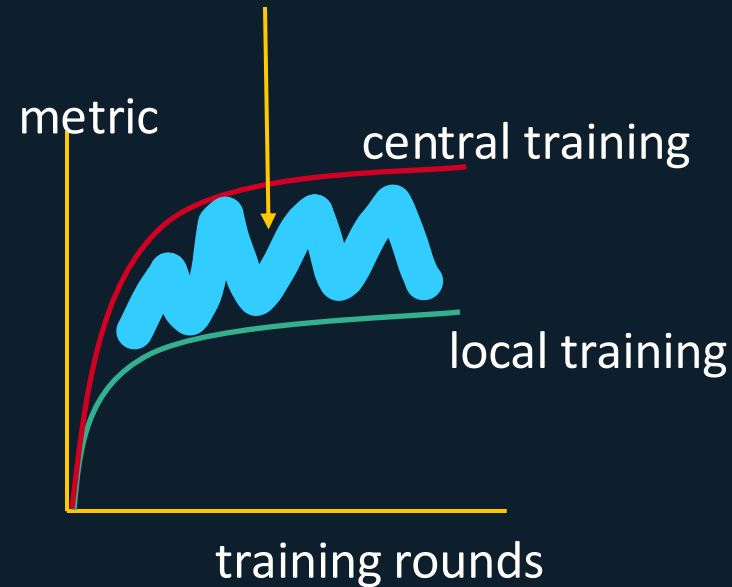
## Local training

- ✗ dataset may be small
- ✓ dataset remain confidential



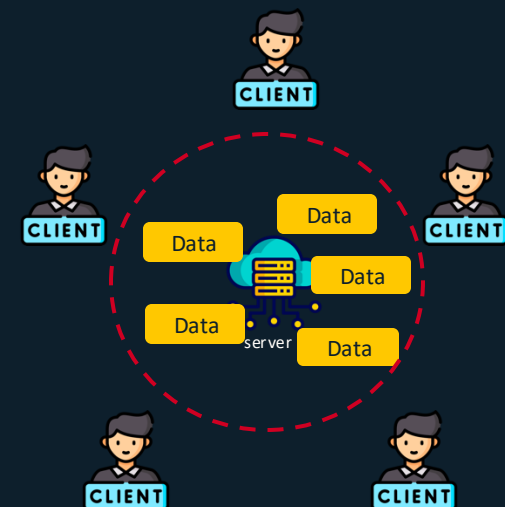
## Federated training

- ✓ Larger pool of data
- ✓ dataset remain confidential

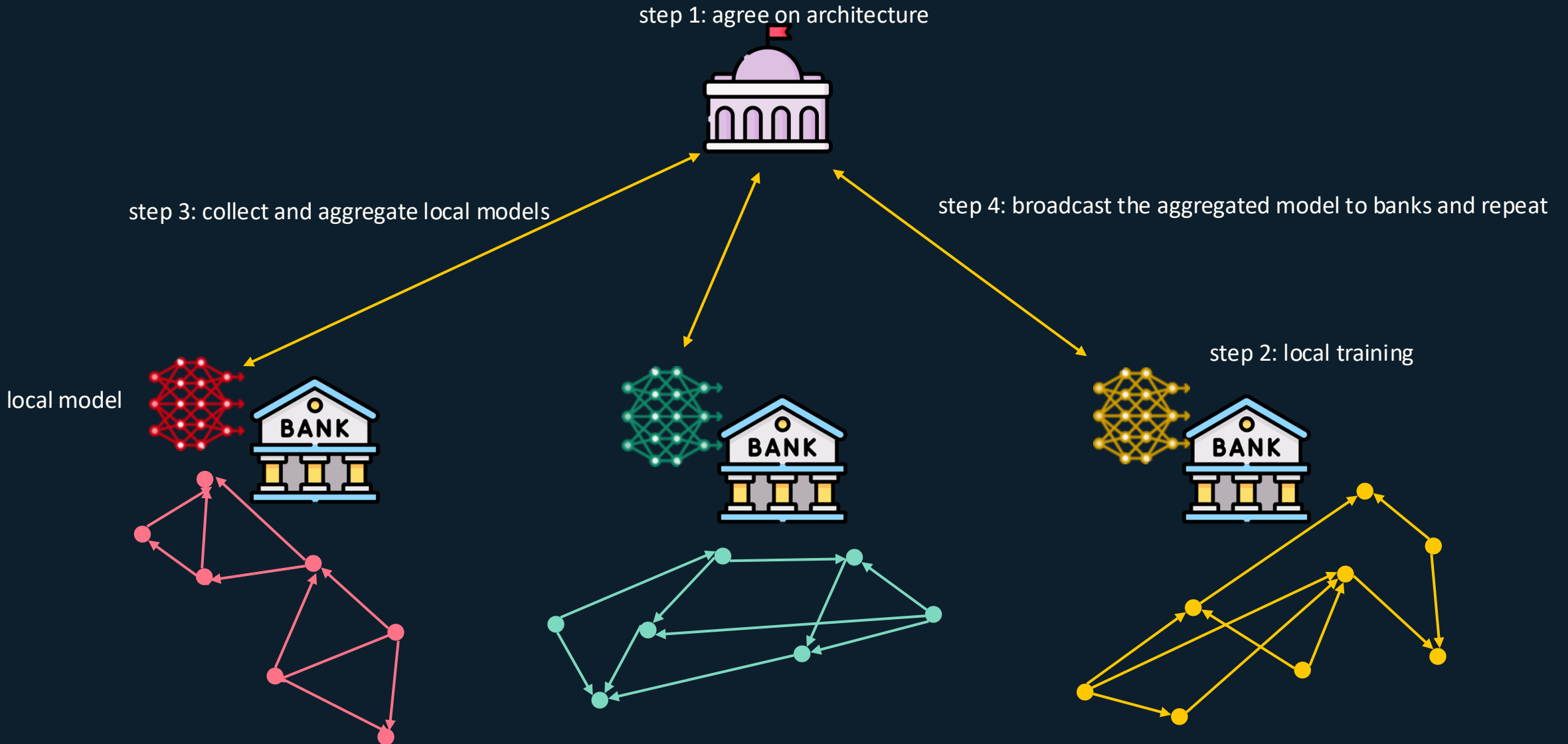


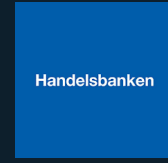
## Central training

- ✓ Larger pool of data
- ✗ Data privacy violated



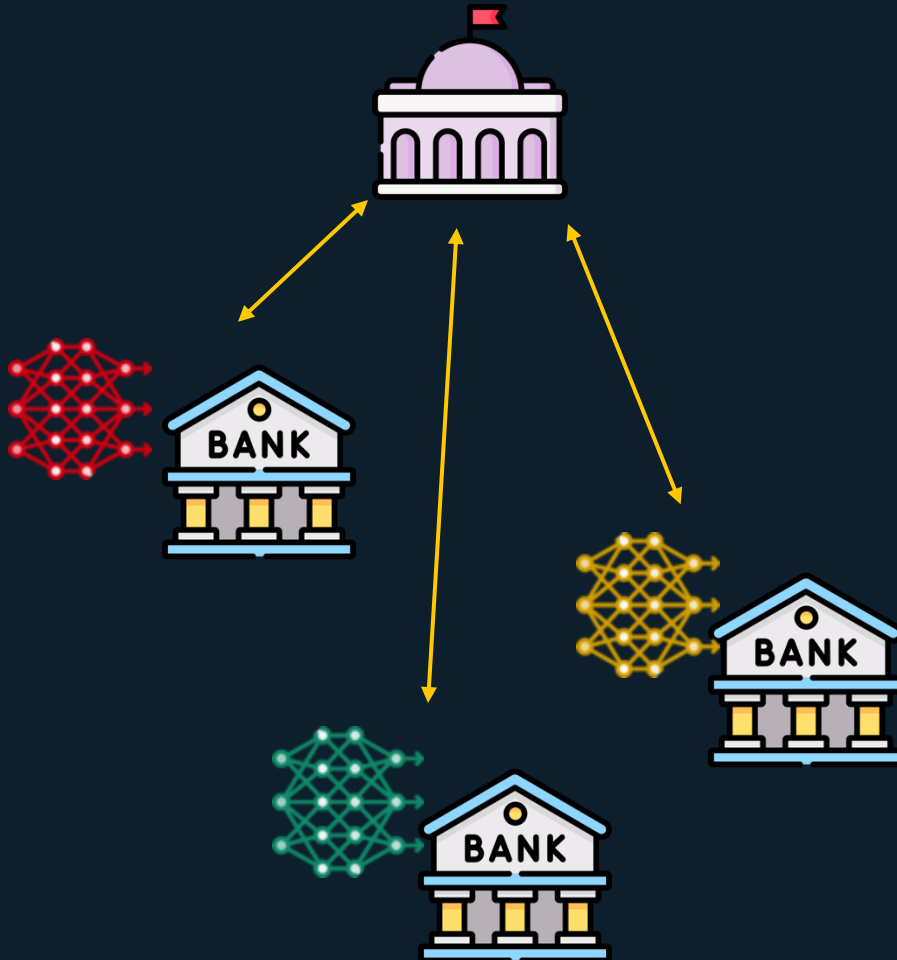
# Federated Learning



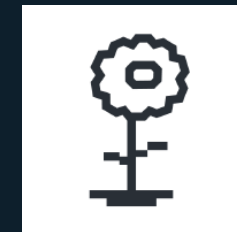
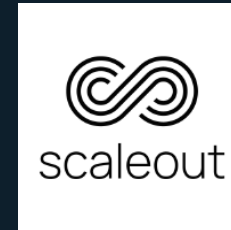


# Federated Learning in practice

# Requirements



- Infrastructure
  - Central server
  - Clients with computational power



Flower

- Alignment on data structure
- People
  - Domain knowledge
  - Machine learning

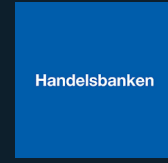


# Challenges

- Heterogenous data distributions
- Privacy concerns

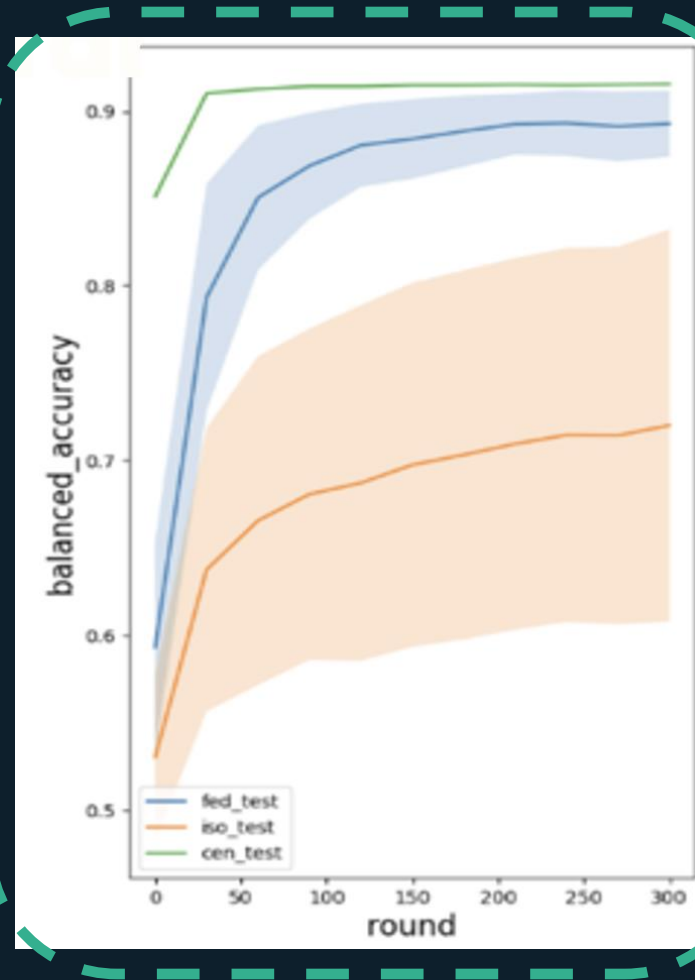
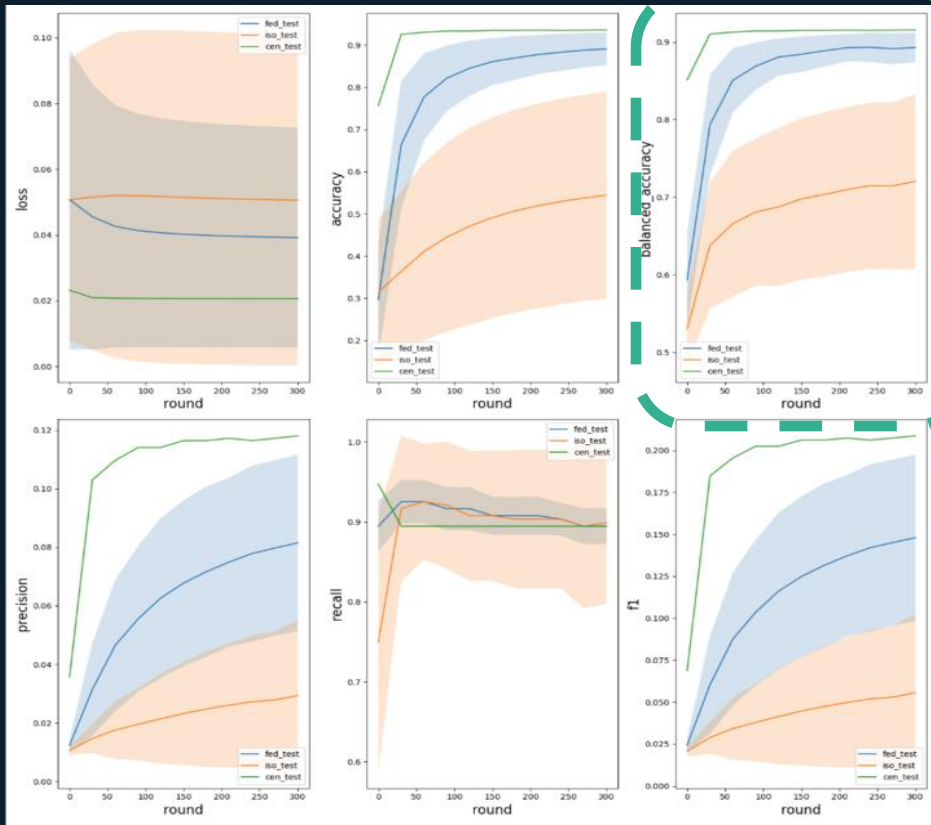
# Opportunities

- Exposing the model to more data
  - Better performance
  - Better generalization



# Results

# Preliminary experiments



Centralized  
Federated


Isolation


# Thank you!

Edvin Callisen


Research Engineer, AI Sweden

[edvin.callisen@ai.se](mailto:edvin.callisen@ai.se)

 [ai.se](https://ai.se)

 [my.ai.se](https://my.ai.se)

 [ai.se/newsletter](https://ai.se/newsletter)

 [youtube.com/c/aisweden](https://youtube.com/c/aisweden)

 [linkedin.com/company/aisweden](https://linkedin.com/company/aisweden)

**AI**  
SWEDEN

