Wireless Broadcasting for Efficiency and Accuracy in Federated Learning

by Jonas Wessner

Mar 26th 2024, Taipei, TW

香港大學
THE UNIVERSITY OF HONG KONG

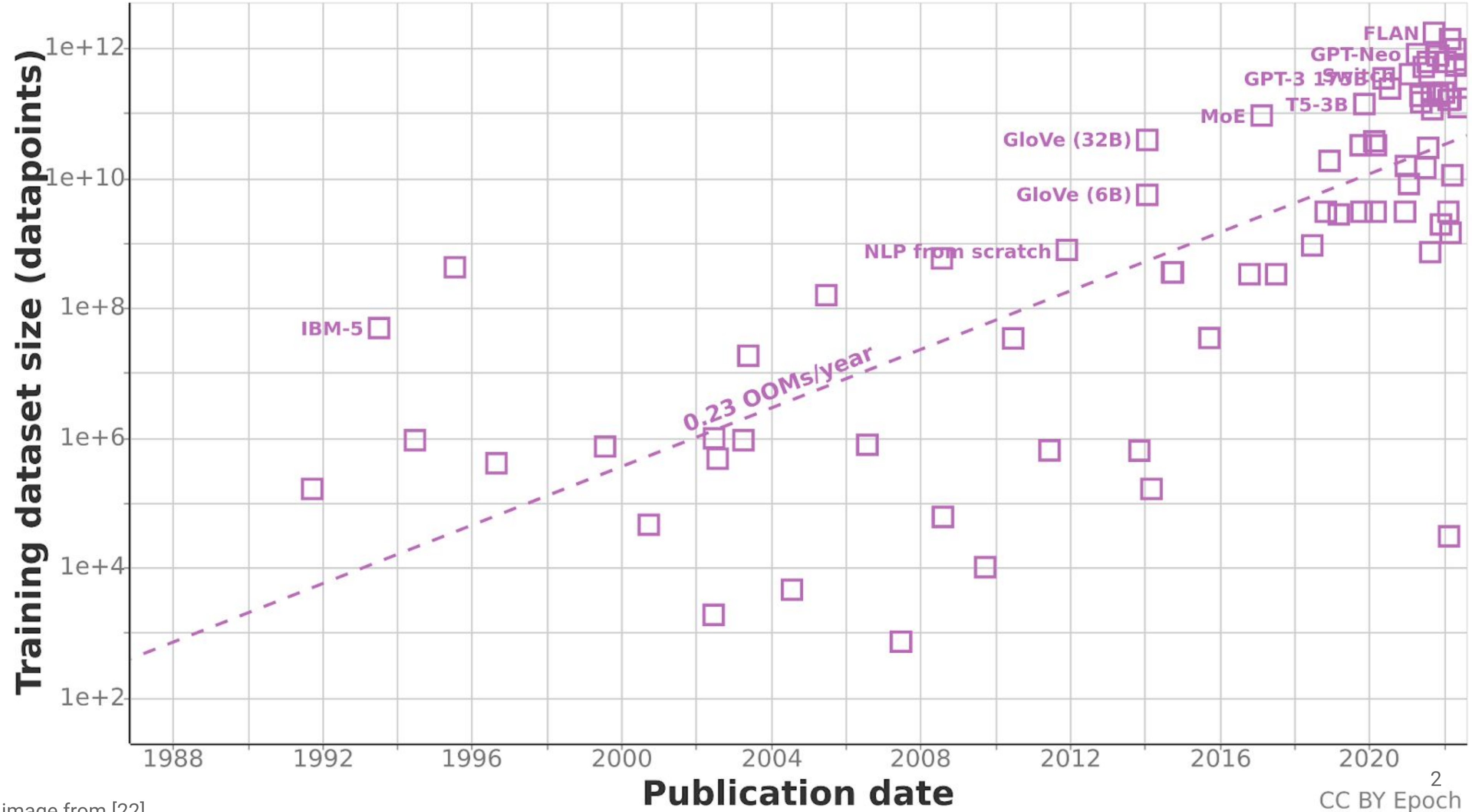


image from [22]

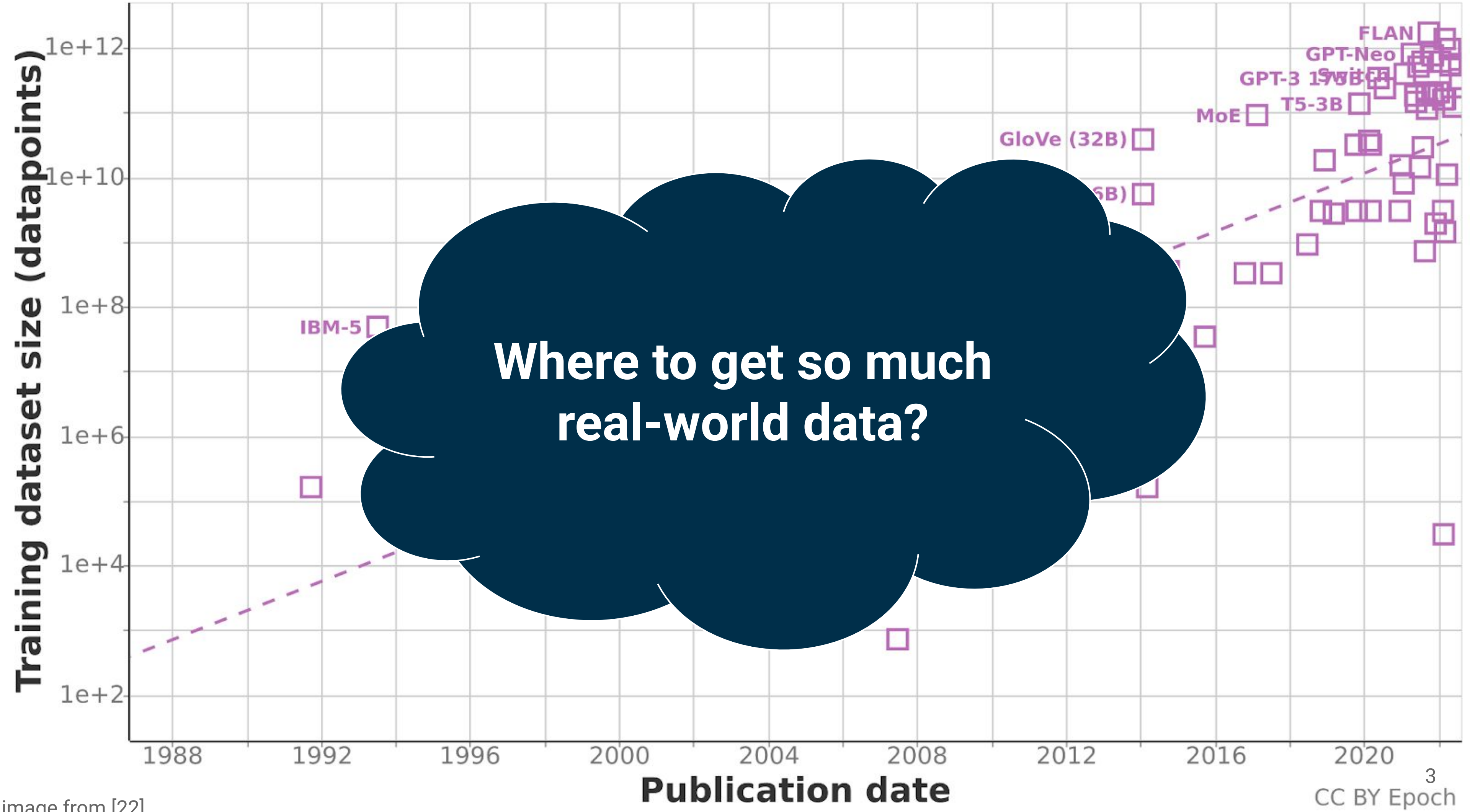
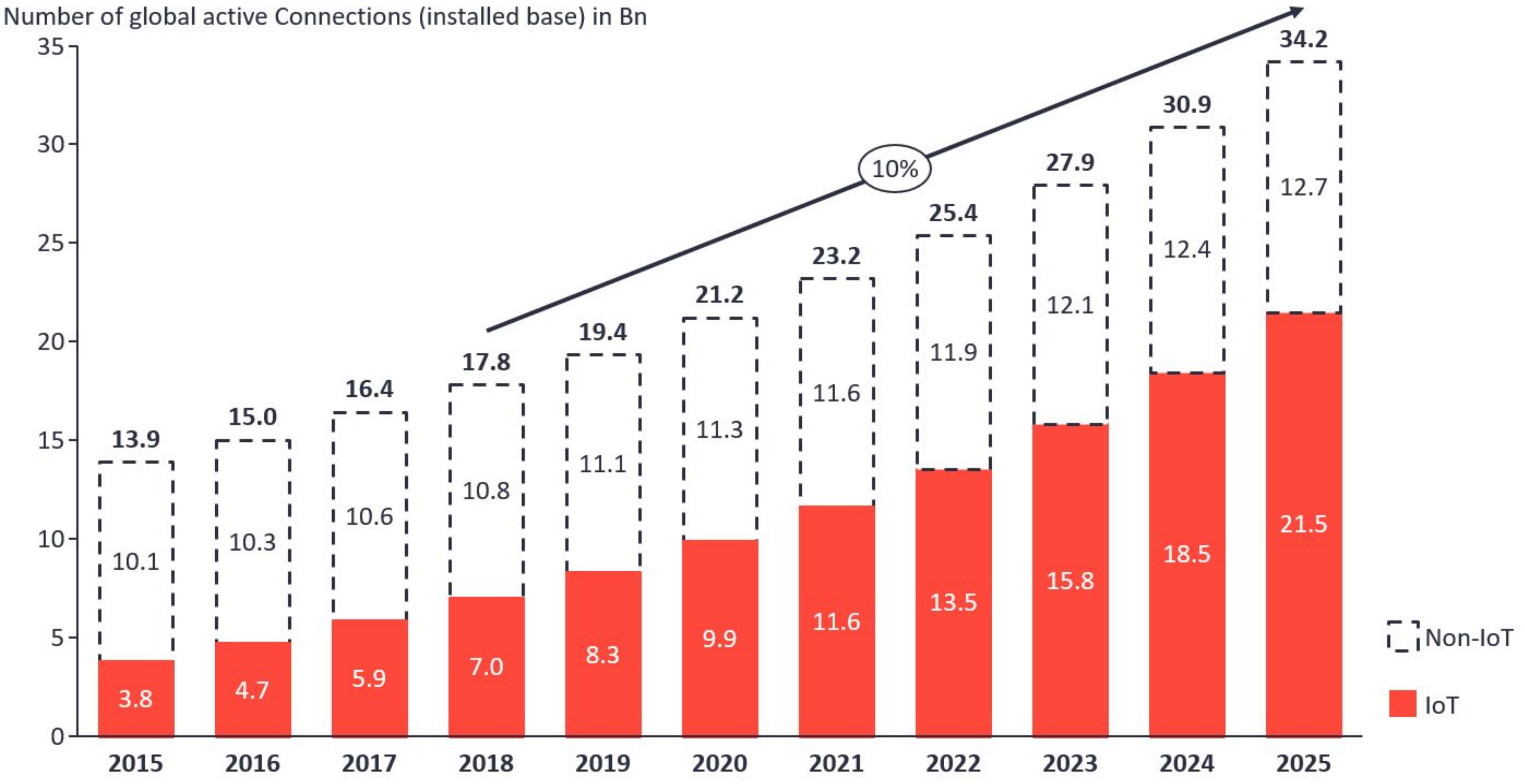


image from [22]

Total number of active device connections worldwide



Note: Non-IoT includes all mobile phones, tablets, PCs, laptops, and fixed line phones. IoT includes all consumer and B2B devices connected – see IoT break-down for further details
Source: IoT Analytics Research 2018

image from [23]

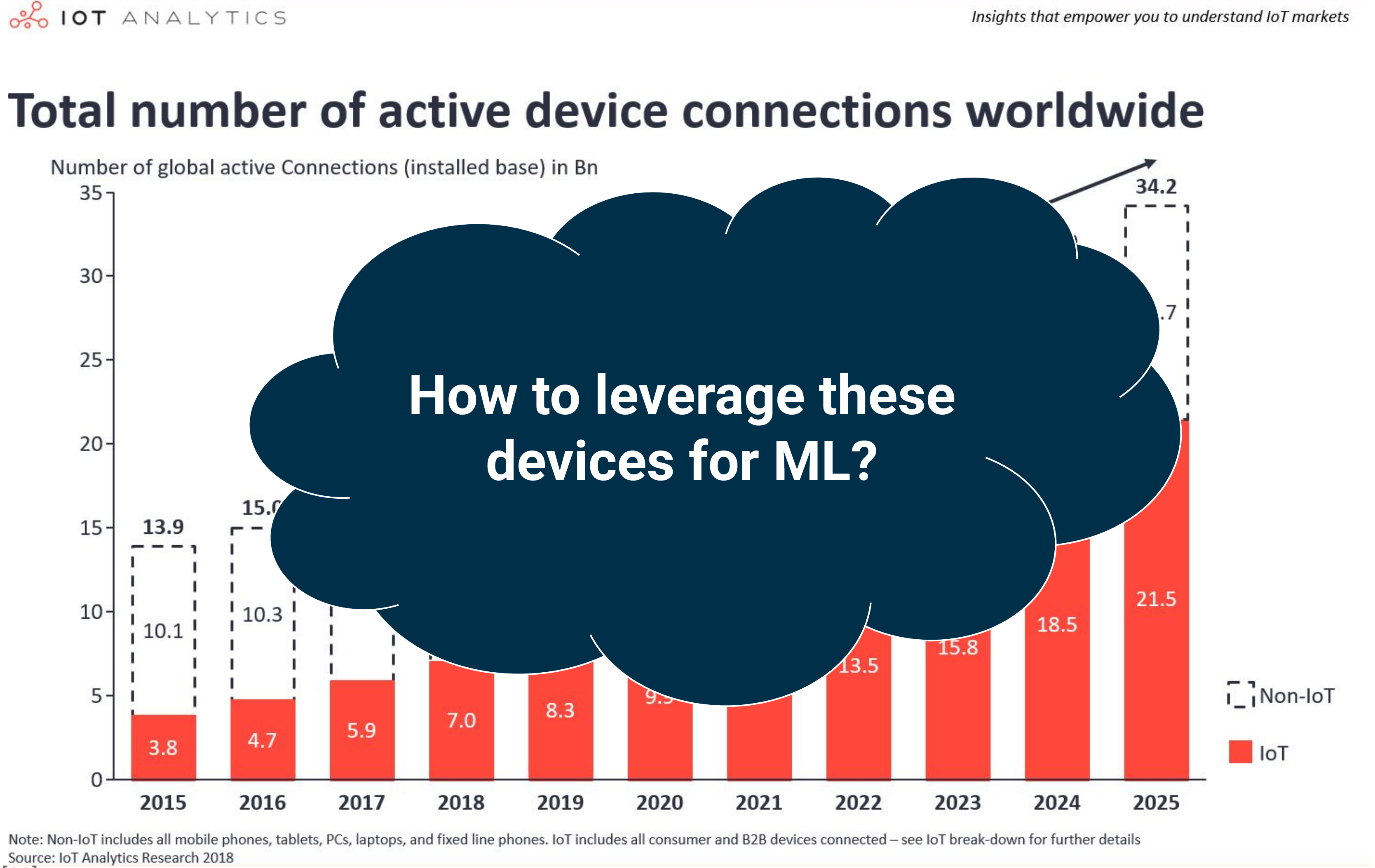
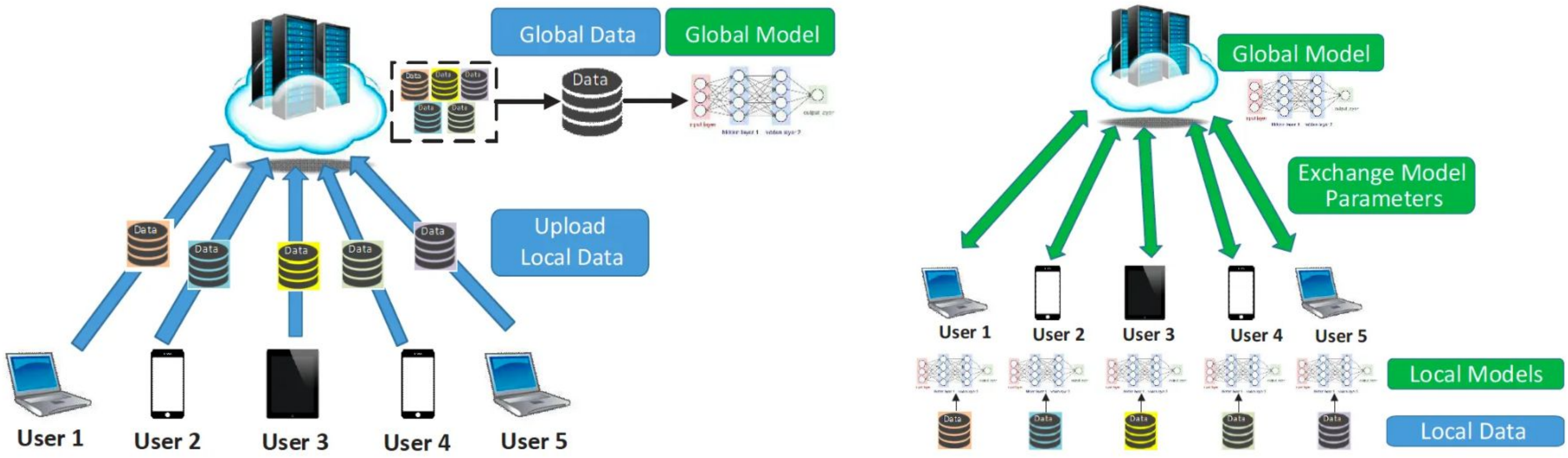


image from [23]

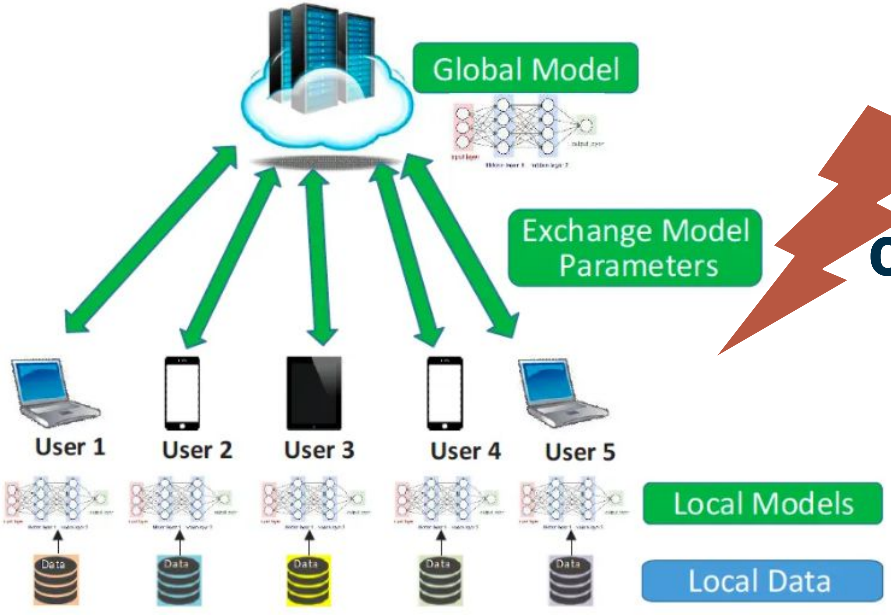
Federated Learning (by Google in 2016)



centralized learning **federated learning**

image from [24]

Federated Learning: Challenges



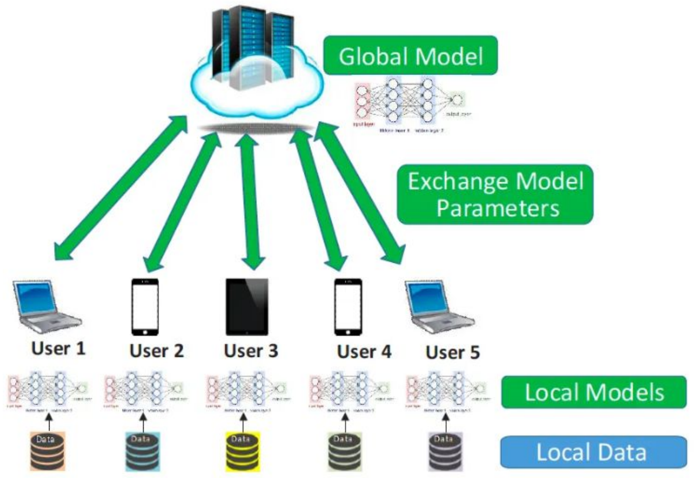
communication cost

non-IID

Current Research Topics

Communication Cost:

➔ Hierarchical Server Architecture [12]
➔ Optimizing number of local rounds [13]
➔ Merging of independently trained models [14]



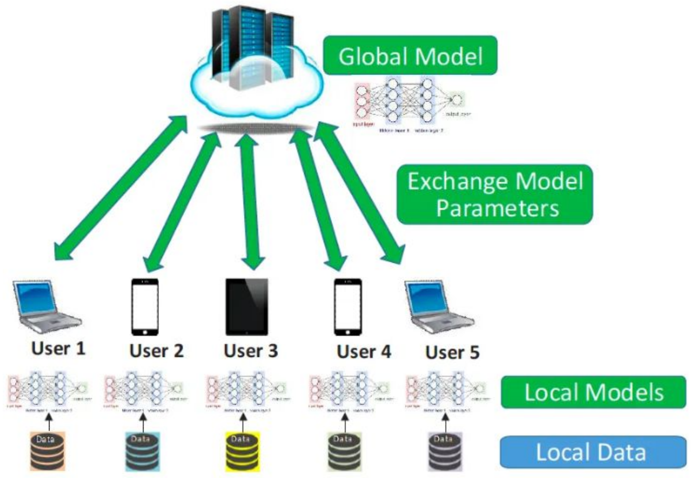
Current Research Topics

Communication Cost:

- ➔ Hierarchical Server Architecture [12]
- ➔ Optimizing number of local rounds [13]
- ➔ Merging of independently trained models [14]

Handling non-IID Data:

- ➔ K-means clustering of data to train K models [22]
- ➔ Sharing part of the data [21]
- ➔ Identifying a global trend to get rid of outliers [10]

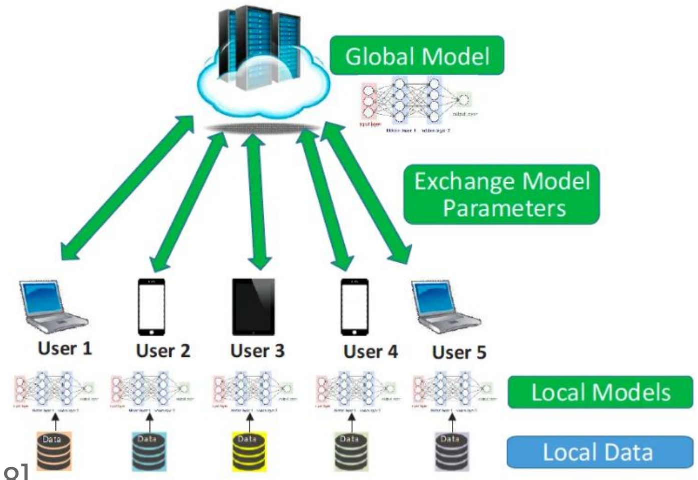


Current Research Topics

Communication Cost:
- ➔ Hierarchical Server Architecture [12]
- ➔ Optimizing number of local rounds [13]
- ➔ Merging of independently trained models [14]

Handling non-IID Data:
- ➔ K-means clustering of data to train K models [22]
- ➔ Sharing part of the data [21]
- ➔ Identifying a global trend to get rid of outliers [10]

Dependency on Central Server:
- ➔ Aggregate functions in P2P networks [17]
- ➔ Asynchronous communication in P2P networks with enh. privacy [18]



Current Research Topics

Communication Cost:

➔ Hierarchical Server Architecture [12]
➔ Optimizing number of local rounds [13]
➔ Merging of independently trained models [14]

Handling non-IID Data:

➔ K-means clustering of data to train K models [22]
➔ Sharing part of the data [21]
➔ Identifying a global trend to get rid of outliers [10]

Dependency on Central Server:

➔ Aggregate functions in P2P networks [17]
➔ Asynchronous communication in P2P networks with enh. privacy [18]

What about leveraging 5G+ networks?

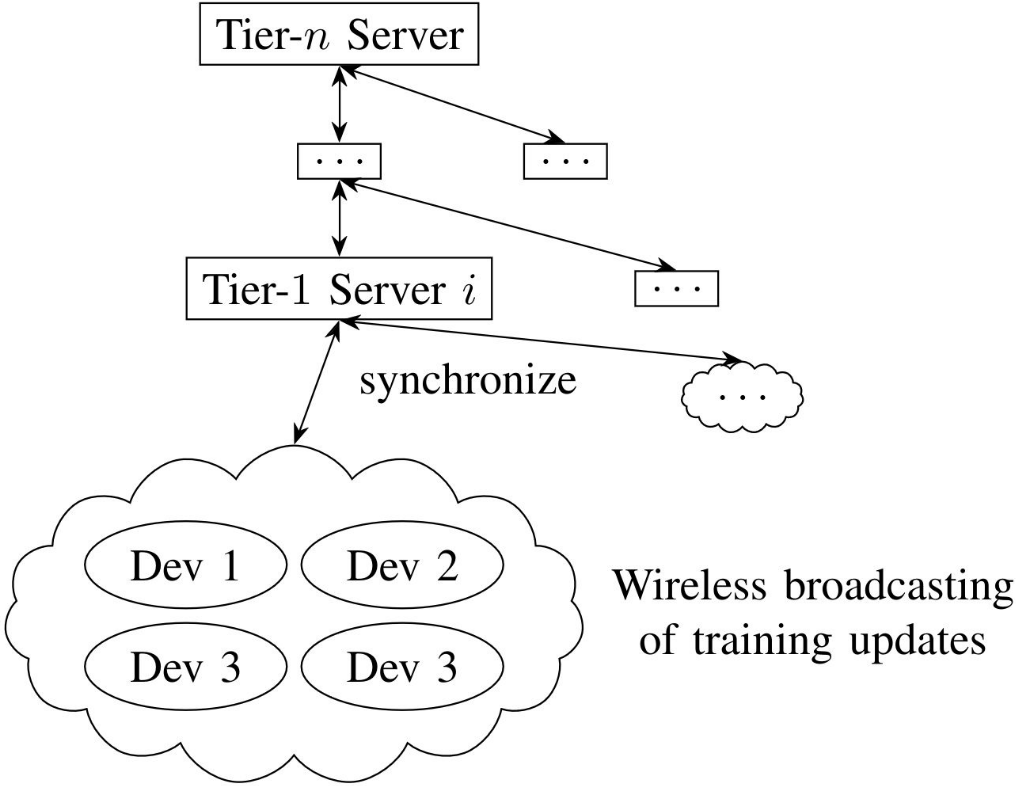
images from [25-27]

Let's use that for FL!

images from [25-27]

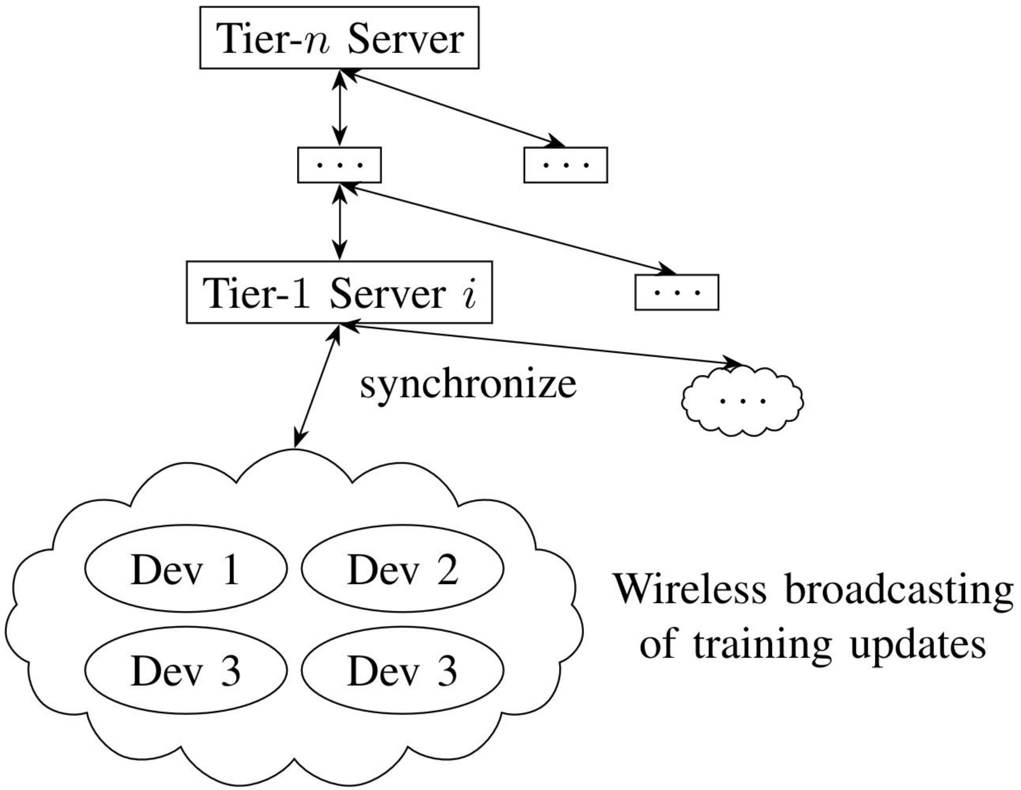
Hybrid Protocol: Wireless D2D & Hierarch. Server

- Decrease server load & communication cost
- Prevent parameter diverging



Hybrid Protocol:

- Devices broadcast their weight deltas and merge received deltas
- Clients have a random timer that triggers sync with server for one client per group

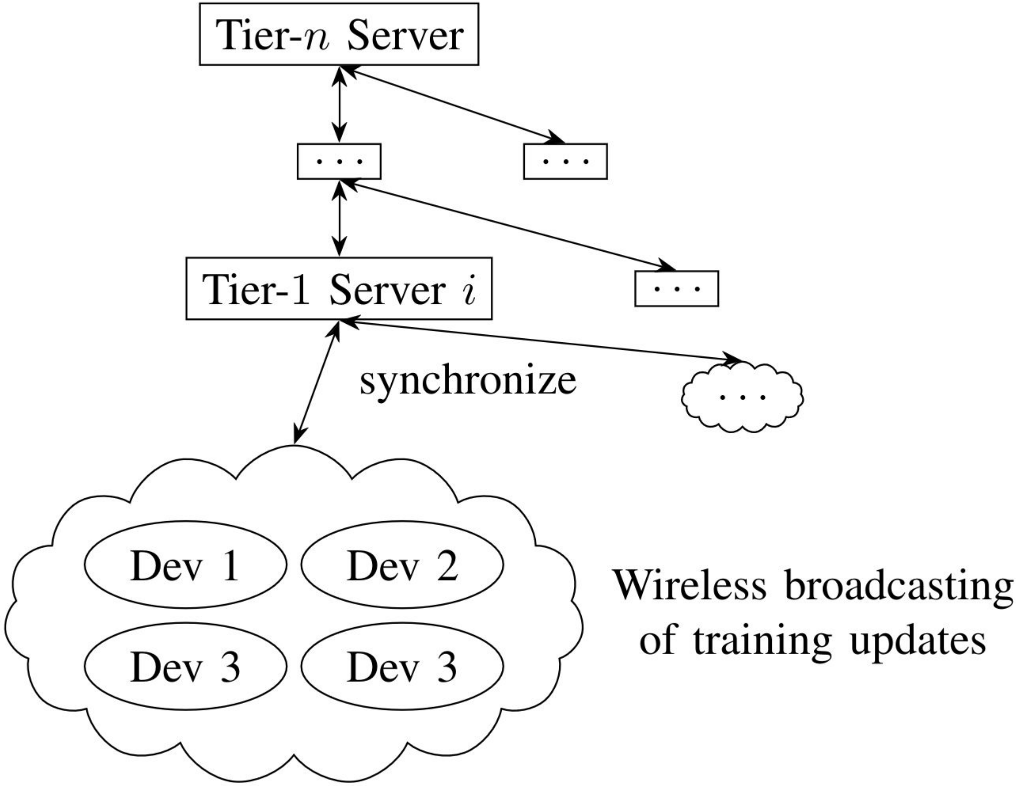


Hybrid Protocol:

- Devices broadcast their weight deltas and merge received deltas
- Clients have a random timer that triggers sync with s

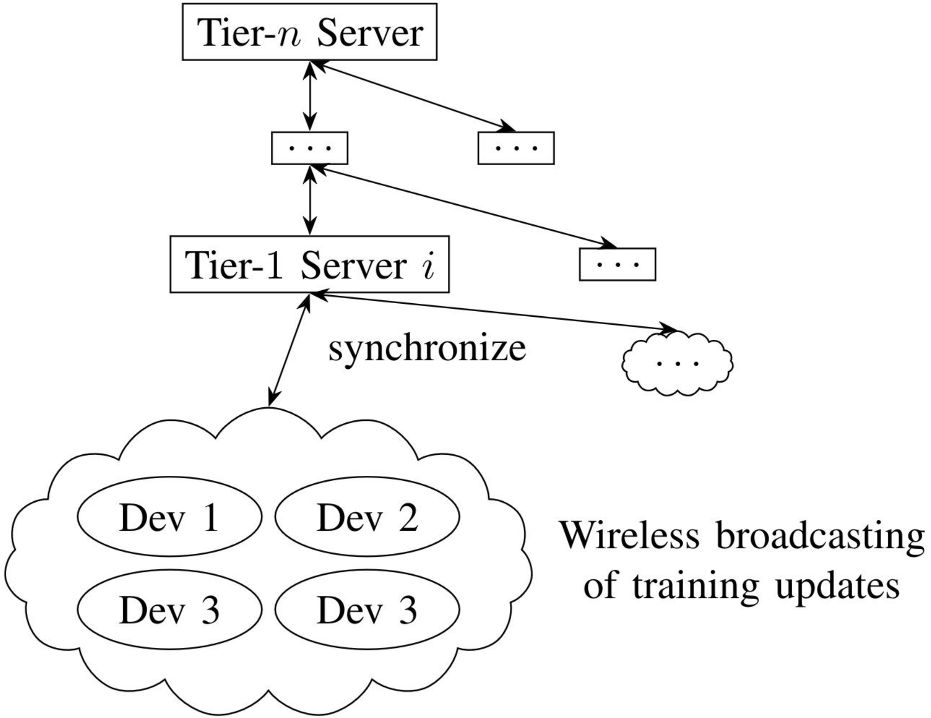
→ Less communication with server

→ Early migration causes more uniform data distribution



Hybrid Protocol: Future Work

- Training convergence analysis
- Sync frequency optimization
- D2D transmit power control



Conclusion

- 5G+ characteristics are an opportunity for dynamic loosely coupled low-cost D2D architectures
- Useful to reduce parameter divergence and server involvement in FL
- Future Work:
  App-independent implementation

References

[1] M. I. Jordan and T. M. Mitchell, “Machine learning: Trends, perspectives, and prospects,” Science, vol. 349, no. 6245, pp. 255–260, 2015.

[2] R. C. Moore and W. Lewis, “Intelligent selection of language model training data,” in Proceedings of the ACL 2010 conference short papers, 2010, pp. 220–224.

[3] H. N. Saha, A. Mandal, and A. Sinha, “Recent trends in the internet of things,” in 2017 IEEE 7th annual computing and communication workshop and conference (CCWC). IEEE, 2017, pp. 1–4.

[4] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, “Communication-efficient learning of deep networks from decentralized data,” in Artificial intelligence and statistics. PMLR, 2017, pp. 1273–1282.

[5] Q. Yang, Y. Liu, Y. Cheng, Y. Kang, T. Chen, and H. Yu, Horizontal Federated Learning. Cham: Springer International Publishing, 2020, pp. 49–67.

[6] Q. Yang, Y. Liu, T. Chen, and Y. Tong, “Federated machine learning: Concept and applications,” ACM Transactions on Intelligent Systems and Technology (TIST), vol. 10, no. 2, pp. 1–19, 2019.

[7] L. Li, Y. Fan, M. Tse, and K.-Y. Lin, “A review of applications in federated learning,” Computers & Industrial Engineering, vol. 149, p. 106854, 2020.

[8] A. Hard, K. Rao, R. Mathews, S. Ramaswamy, F. Beaufays, S. Augenstein, H. Eichner, C. Kiddon, and D. Ramage, “Federated learning for mobile keyboard prediction,” arXiv preprint arXiv:1811.03604, 2018.

[9] S. Silva, B. A. Gutman, E. Romero, P. M. Thompson, A. Altmann, and M. Lorenzi, “Federated learning in distributed medical databases: Meta-analysis of large-scale subcortical brain data,” in 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), 2019, pp.270–274.

[10] W. Luping, W. Wei, and L. Bo, “Cmfl: Mitigating communication overhead for federated learning,” in 2019 IEEE 39th international conference on distributed computing systems (ICDCS). IEEE, 2019, pp. 954–964.

[11] T. Brants, A. C. Popat, P. Xu, F. J. Och, and J. Dean, “Large language models in machine translation,” 2007. [12] L. Liu, J. Zhang, S. Song, and K. B. Letaief, “Client-edge-cloud hierarchical federated learning,” in ICC 2020-2020 IEEE International Conference on Communications (ICC). IEEE, 2020, pp. 1–6.

[13] S. Wang, T. Tuor, T. Salonidis, K. K. Leung, C. Makaya, T. He, andK. Chan, “Adaptive federated learning in resource constrained edge computing systems,” IEEE Journal on Selected Areas in Communications, vol. 37, no. 6, pp. 1205–1221, 2019.

[14] M. Yurochkin, M. Agarwal, S. Ghosh, K. Greenewald, N. Hoang, and Y. Khazaeni, “Bayesian nonparametric federated learning of neural networks,” in Proceedings of the 36th International Conference on Machine Learning, ser. Proceedings of Machine Learning Research, K. Chaudhuri and R. Salakhutdinov, Eds., vol. 97. PMLR, 09–15 Jun 2019, pp. 7252–7261.

References (cont.)

[15] H. Li and T. Han, “An end-to-end encrypted neural network for gradient updates transmission in federated learning,” arXiv preprint arXiv:1908.08340, 2019.

[16] D. A. Lelewer and D. S. Hirschberg, “Data compression,” ACM Computing Surveys (CSUR), vol. 19, no. 3, pp. 261–296, 1987.

[17] J. L. Salmeron, I. Ar´evalo, and A. Ruiz-Celma, “Benchmarking federated strategies in peer-to-peer federated learning for biomedical data,” Heliyon, 2023.

[18] A. G. Roy, S. Siddiqui, S. P¨olsterl, N. Navab, and C. Wachinger, “Braintorrent: A peer-to-peer environment for decentralized federated learning,” arXiv preprint arXiv:1905.06731, 2019.

[19] T. Wink and Z. Nochta, “An approach for peer-to-peer federated learning,” in 2021 51st Annual IEEE/IFIP International Conference on Dependable Systems and Networks Workshops (DSN-W). IEEE, 2021, pp. 150–157.

[20] A. Bellet, R. Guerraoui, M. Taziki, and M. Tommasi, “Personalized and private peer-to-peer machine learning,” in International Conference on Artificial Intelligence and Statistics. PMLR, 2018, pp. 473–481.

[21] Y. Zhao, M. Li, L. Lai, N. Suda, D. Civin, and V. Chandra, “Federated learning with non-iid data,” arXiv preprint arXiv:1806.00582, 2018.

[22] image from https://epochai.org/assets/images/posts/2022/trends-in-training-dataset-sizes.png

[23] image from https://iot-analytics.com/wp/wp-content/uploads/2018/08/Number-of-global-device-connections-2015-2025-Number-of-IoT-Devices.png

[24] image from https://miro.medium.com/v2/resize:fit:4800/format:webp/1*OxozfyCYZQ7Fm5tw2AWFLA.png

[25] image from https://webhostinggeeks.com/blog/wp-content/uploads/2023/09/data-transfer-rate-980x927.png

[26] image from https://assets-global.website-files.com/64ab7ea969c1429ab6efa588/65372b26e037d67147cb56ba_Edge-computing.jpg

[27] image from https://contents.mediadecathlon.com/s942935/k$44afc4a6679eb8610cab14481b5cbd6f/3500pt2334/6223xcr4667/default.jpg?format=auto