# Blind Learning: Facilitating Data Access for Informatics Innovators while Protecting Providers Information

**Gharib Gharibi, PhD[1], Babak P. Gilkalaye, MS[1], Riddhiman Das, MS[1], Greg Storm, PhD[1], Suraj Kapa, MD[1], I. Zachi Attia, PhD[2], Paul A. Friedman, MD[2], Craig Gentry, PhD[1]**

**[1] TripleBlind [2] Department of Cardiovascular Medicine, Mayo Clinic**

**Motivation and Challenges.** Machine learning (ML) methods and tools have made significant advances in the informatics domain. This success heavily relies on the accessibility of large amounts of *diverse* data—which is critical for training high-performance, fair, and generalizable models. However, existing methods for accessing multi-institutional and multi-national healthcare data still face two major challenges. First, they typically rely on conventional privacy methods, such as anonymization, which has proven insufficient[1]. Second, emerging privacy techniques lack adequate tool support, which puts them out of reach for most informatics innovators.

**Contributions.** We present an automated, privacy-preserving toolset, named *Blind Learning (BL)*, that aims to *facilitate* and *accelerate* secure access to decentralized data without moving it outside the owner's infrastructure. Specifically, *BL* provides a set of high-level, automated APIs (application programming interface) that enables the training of sophisticated AI models from decentralized data while preserving its privacy. *BL* employs a combination of Federated Learning, Split Learning, and Secure Multi-Party Computation (MPC) techniques for model training. It also meets the HIPAA's de-identification standard. We further discuss the underlying technical details during the demo session[2].

**Methods.** We evaluate the accuracy and privacy of *BL* on decentralized data compared to a locally trained baseline on centralized data in cleartext (without privacy). We used a real-life medical imaging case study for training a binary Pneumonia classification model with a realistic architecture `ResNet-26` on $5,932$ chest X-rays. In the *BL* case, we divided the dataset across three data owners located on separate cloud instances, and the training took place over the public Internet. We then trained the same architecture locally in cleartext over the whole dataset. We used standard training configurations, including `binary cross entropy` as a loss function, `Adam` optimizer, and a learning rate of $1 \times 10^{-3}$. The goal is to show that models trained securely on distributed data using *BL* are on par with the locally trained models–while drastically reducing data leakage measured in the success accuracy of the membership inference attack[3]. We further demonstrate the ease-of-use of our system during the demo session*.

**Results and Discussions.** Our thorough evaluation demonstrates that *BL* can securely produce models from decentralized datasets equivalent to the models trained in cleartext on centrally pooled data. For example, the model trained on three datasets using *BL* achieved an overall accuracy of $84.7\%$ while the locally trained model achieved $85.1\%$ accuracy (both averaged over three runs). *BL* also mitigates against membership inference attack and reduces its success accuracy to $55\%$ from $75.1\%$ on the locally trained model. It is also important to note that distributed training using *BL* is more computationally and timely efficient than other distributed learning methods, including Federated Learning and Split Learning, as illustrated previously[2].

**System Demonstration Plan.** (1) Motivate the need for automated privacy-preserving tools convenient for healthcare professionals. (2) Explain the underlying methodology and algorithms implemented in our toolset. (3) Demonstrate the usage of our toolset using several examples (classification, multi-modal regression). (4) Invite the audience to interact with our system during the conference using an active, online development environment, Jupyter Notebooks. *

**Tool Maturity.** Our tool is currently being used by medium to large-sized healthcare providers from around the globe. For example, Mayo Clinic uses our toolset to (1) train cutting-edge deep learning models from decentralized data owned by Mayo Clinic and other partners and (2) validate others' proprietary AI models on Mayo Clinic's data.

## References

[1] Schwarz CG, Kremers WK, Therneau TM, Sharp RR, Gunter JL, Vemuri P, et al. Identification of anonymous MRI research participants with face-recognition software. New England Journal of Medicine. 2019;381(17):1684-6.

[2] Gharibi G, Patel R, Khan A, Gilkalaye BP, Vepakomma P, Raskar R, et al. An Automated Framework for Distributed Deep Learning–A Tool Demo. In: IEEE 42nd International Conference on Distributed Computing Systems. IEEE; 2022. p. 1302-5.

[3] Shokri R, Stronati M, Song C, Shmatikov V. Membership inference attacks against machine learning models. In: 2017 IEEE Symposium on Security and Privacy (SP). IEEE; 2017. p. 3-18.

*A technical system demo is also available at `https://tripleblind.ai/icdcs22/`

*Accepted and presented at AMIA Clinical Informatics Conference, May 23-25, 2023.