



NEOGLIA

The new infrastructure of healthcare AI

Executive summary

NeoGlia empowers researchers and enables payers and providers to monetize their data in a new way. Our technology allows data science teams to train machine learning models on our nationwide federated data network of payers and providers. Due to the size of this cumulative dataset, these models are bound to outperform the ones built within a single institute. The key difference between NeoGlia and existing data monetization programs is the guarantee that no participating institute can access or read data of any other party. Furthermore, participating institutions are remunerated everytime their data is used in the training of a new machine learning model. If the reader is familiar with PCORnet¹, one could describe NeoGlia as a modern version of it, made for machine learning with a revenue sharing business model.

Background

Machine learning research on EMR and Claims data has shown significant advances in recent years, producing astonishing results and proposing novel ways of learning from and utilising this complex and vastly heterogeneous data type². In one of the seminal papers of the field, it has been demonstrated that FHIR can act as a data standardisation layer between different institutes and this approach can lead to state of the art results on various prediction tasks³, e.g. readmissions.

Furthermore, researchers from MIT have pushed forward the original federated learning framework for training deep learning networks in a distributed manner⁴, and devised a new learning mechanism which they named splitNN⁵. This method allows for the training of a central network with several data sources, without the raw data ever leaving the containers or servers of the participating institutes. Due to these advancements, coupled with the government sponsored interoperability and the healthcare industry's increasing appetite of cloud based solutions and data storage, we believe the time is right for building NeoGlia.

Infrastructure

NeoGlia integrates both with cloud based and on-premise EMR systems and data warehouses. Our productionized split learning infrastructure bridges data from different hospitals and payers. The first layers of the neural network are built on the client side with only activations being sent over the network to the NeoGlia servers. The final deep layers of the neural network are trained in unison with activations from all participating institutions. For further details, please refer to Figure 2 or request our technical whitepaper.

Revenue Model

Figure 1 provides an overview of a multi-institute deep learning model, leveraging data from two hospitals. The vendor is billed for training models on the data of the participating two institutions. Revenue is shared between NeoGlia and the data providers, who are compensated based on their

contribution to the training cohort in terms of patient counts. For further details, please refer to the example scenarios below. Please note, that non-profit research projects will have special status not bound to this revenue model.

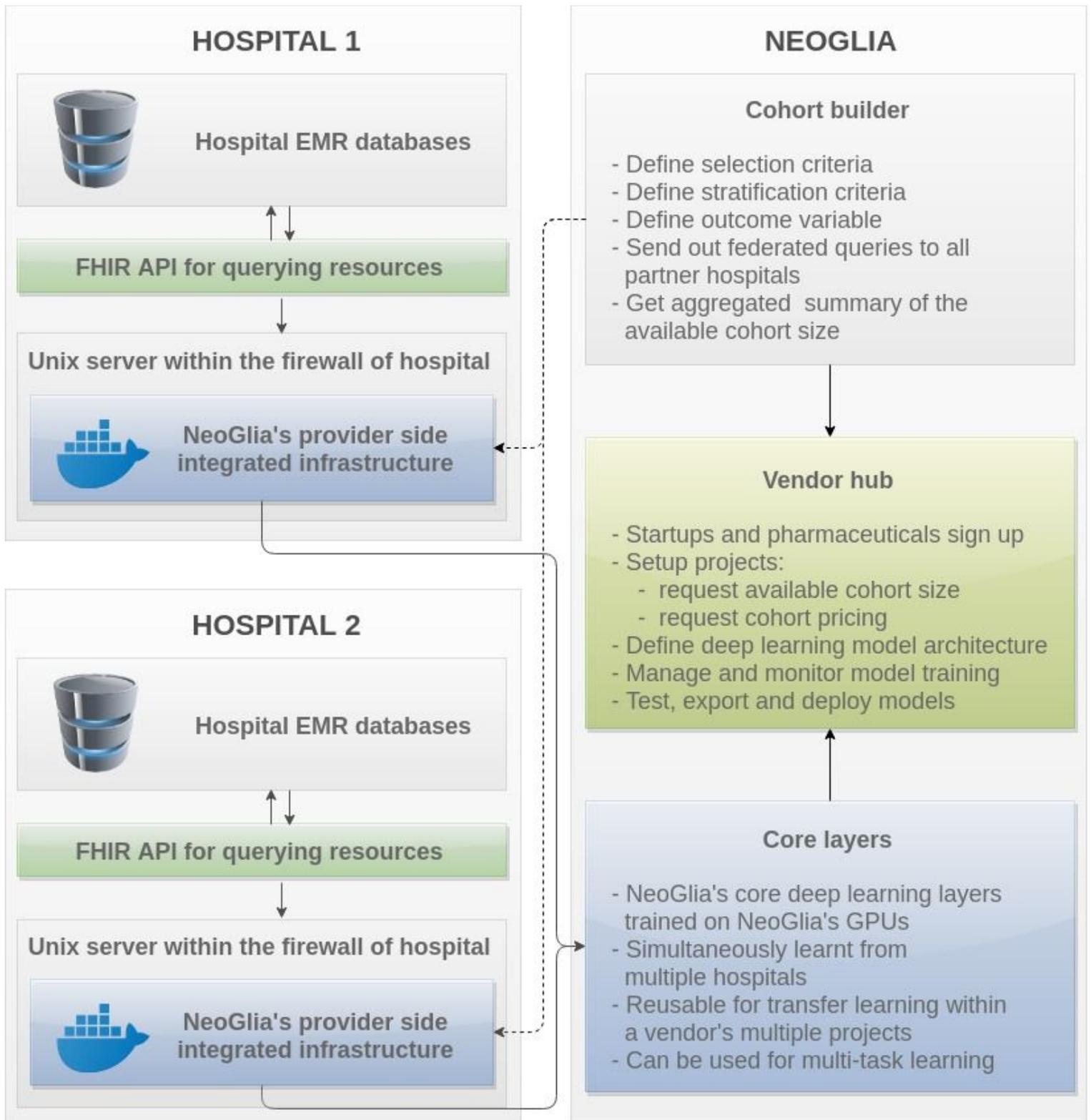


Figure 1 | Multi-institute federated deep learning model, leveraging EMR data from two sites.

Example scenario - static model

Pharmaceutical company X wants to build a line of therapy transition model to understand why some of their customers switch to the competitor's product. For this they require 1,000,000 patients which have been receiving their drug and one month of computation time for the training of a highly accurate model. Data network access is charged \$1,000,000. NeoGlia will take a 30% share for market making with 70% going to the data providers. If one data provider supplied 100,000 of the selected patients they would gain a \$70,000 pay out.

Example scenario - dynamic model

Researchers at a large integrated healthcare provider are aiming to build the most accurate readmission model. For this, they leverage data from several dozen payers and succeed in reaching state of the art prediction accuracy. The finished model is deployed in numerous hospitals around the US with metered billing per prediction. The revenue of the model is split between NeoGlia and all the institutions whose data was used in training of the model, i.e. a continued revenue stream opens up for the participating payers.

Team

Daniel Homola

Phd from Imperial College London in Bioinformatics. Experienced in building deep learning disease detection models on claims and EMR data.



daniel@neoglia.com



Ruben Wolff

MSc from ETH Zürich in Machine Learning. International consulting experience bringing machine learning applications into production.

ruben@neoglia.com

References

1. <https://pcornet.org/>
2. <https://arxiv.org/abs/1706.03446>
3. <https://www.nature.com/articles/s41746-018-0029-1>
4. <https://arxiv.org/abs/1610.05492>
5. <https://arxiv.org/abs/1812.00564>

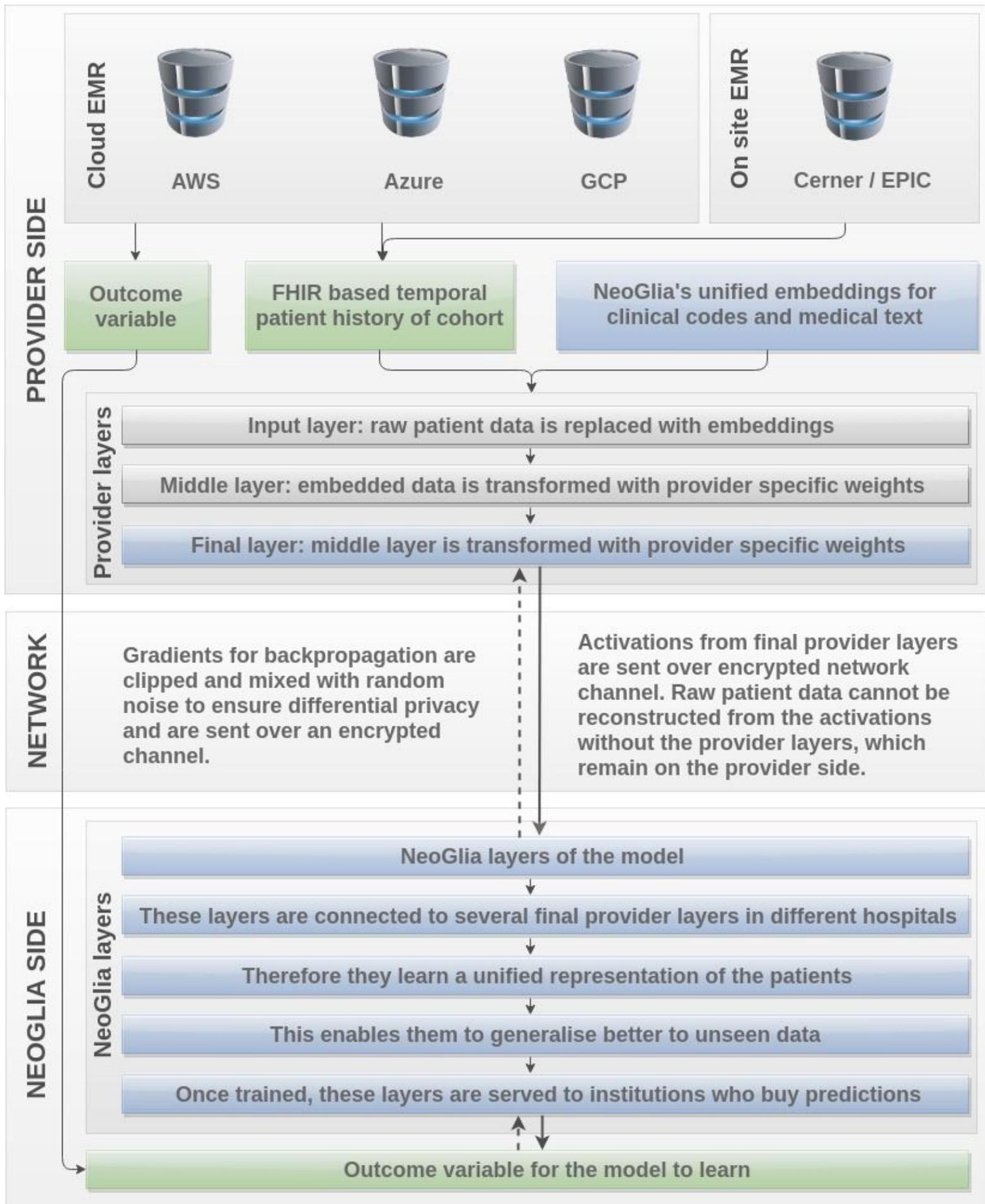


Figure 2 | Federated, privacy preserving deep learning.