Private Deep Learning for Healthcare

CERN openlab Technical Workshop 2021

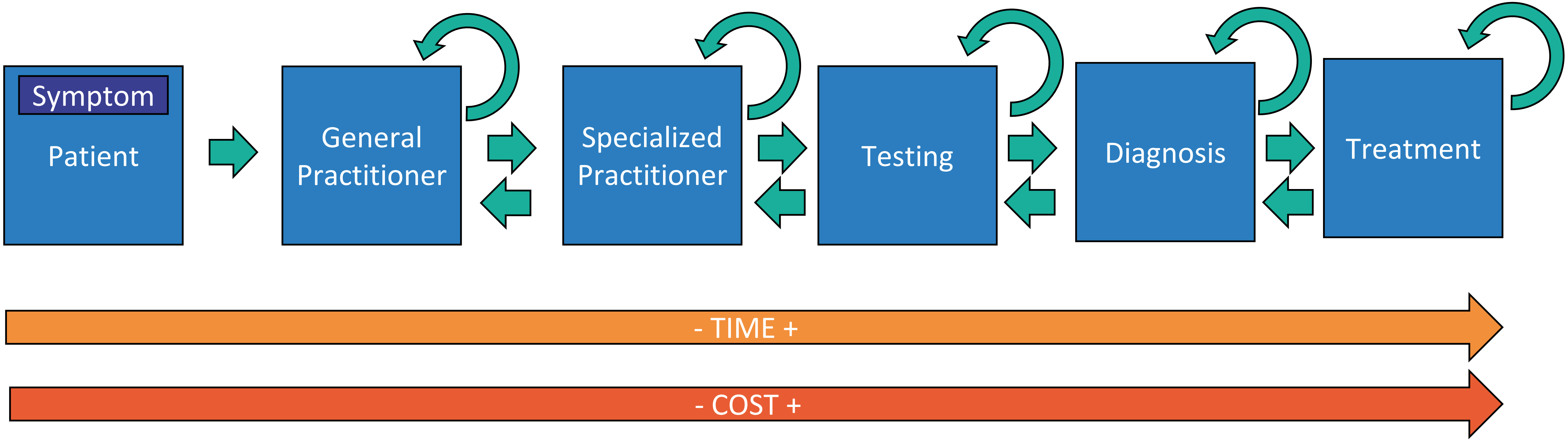
José Cabrero Holgueras (CERN, Universidad Carlos III de Madrid)

jose.cabrero.holgueras@cern.ch

10/03/2021

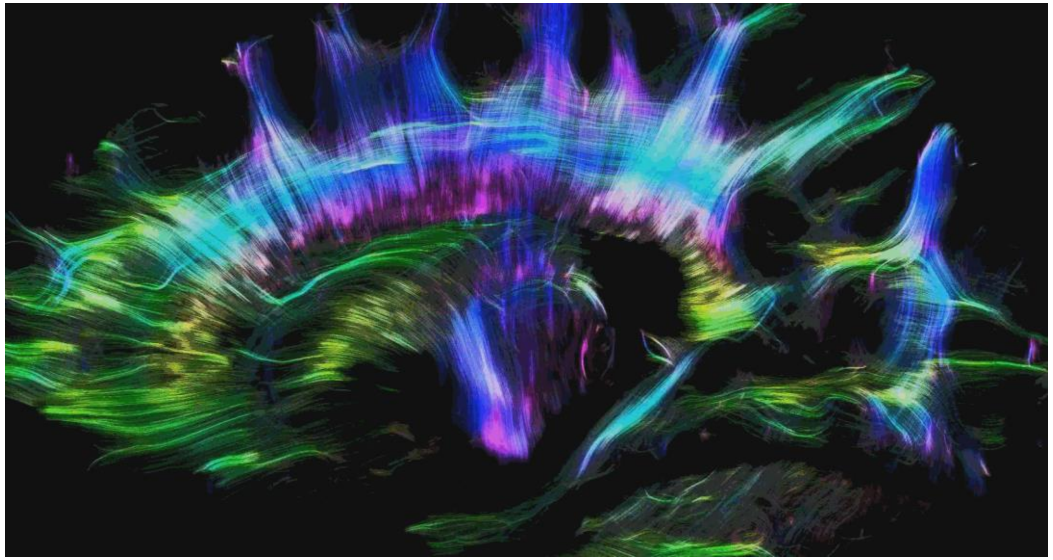
The Present of Healthcare

Diagnosis Pipeline

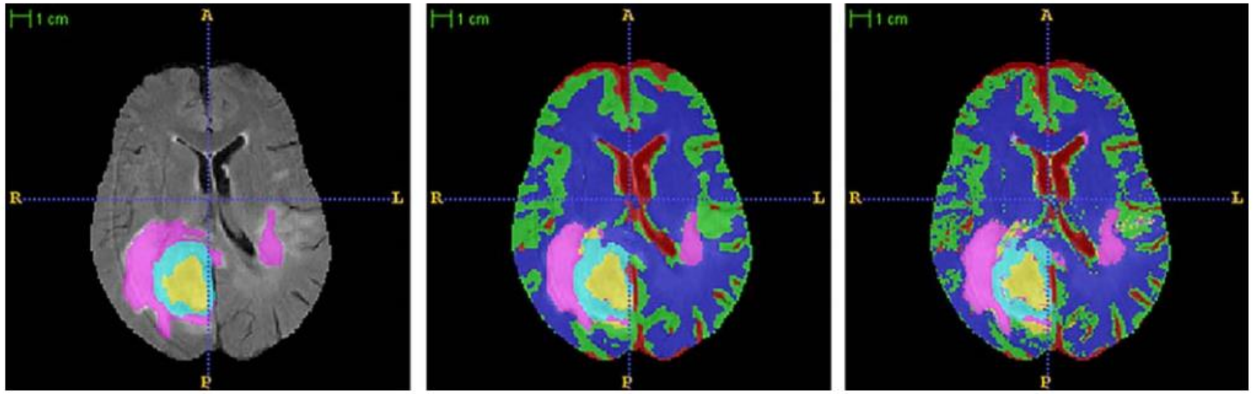


Enhancing Healthcare through Artificial Intelligence

- Artificial Intelligence (AI) and Deep Learning (DL) can be a helpful to medicine [1]:
  - Enhancing reliability, reducing errors of diagnosis.
  - Better understanding of treatments and drug effects.
  - Health outside hospitals.



Source:(Alfred Anwander - https://www.youtube.com/watch?v=jrC8iY6_aZQ, CC BY-SA 4.0, https://commons.wikimedia.org/w/index.php?curid=44329691



Source:(Bauer et al., 2011 MICCAI)

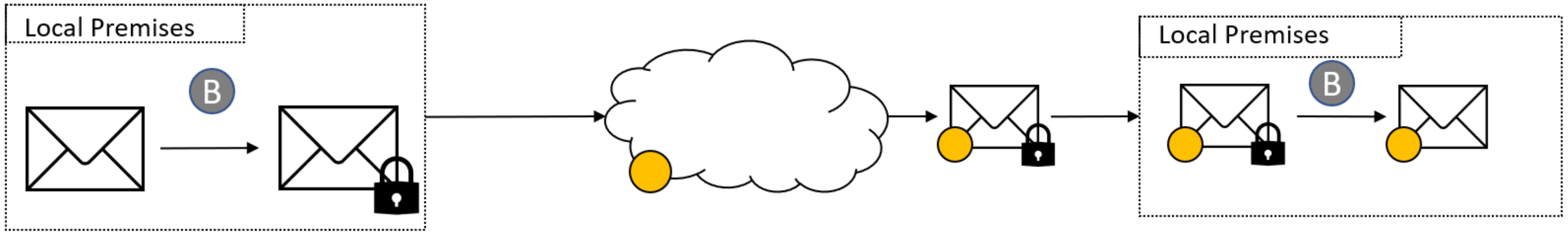
Privacy Issues of DL in Healthcare

- Deep Learning 101:
  1. Get expensive computing hardware.
  2. Get data, the more, the better.
  3. Choose a DL architecture for your data.
  4. Train the model with the data until it “works”.
  5. Use that model for inference (assisted diagnosis).

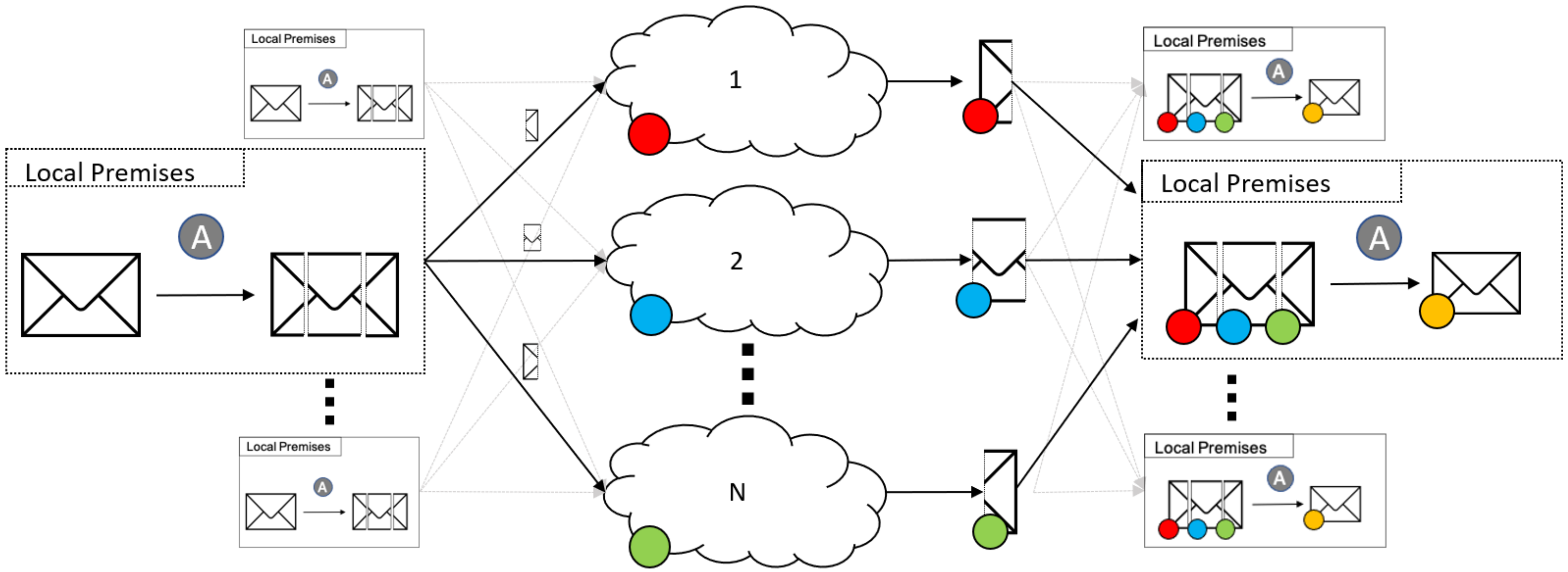
- Problems:
  - Hardware Acquisition may be cost-effective in certain healthcare environments.
    - Cloud Computing and the untrusted third party.
  - Availability of data: Insufficient amounts for DL.
    - Sharing data between institutions and the problem of , and the Data Privacy Regulations.

Privacy Preserving Computation Techniques

Homomorphic Encryption



Secure Multiparty Computation



Applying Privacy to Deep Learning

Not that straightforward…

- An analysis of over 40+ state-of-the-art contributions.

- Most of the techniques suffer from performance deficiencies.
  - CPU Runtime.
  - Memory allocation.
    - 1 GB Shared among 3 parties = 3 GB.

- Lack of open-source implementations.
  - Not integrated in common frameworks.

- Limited operation set.
  - Only addition and multiplication are available.

Next steps:

Building a system for inference:

- Being able to perform a secure diagnosis on medical data with DL.
- Scenario:
  - Client input: Medical Image
  - Server input: DL Model
  - Requirement: Neither learn about each other inputs.

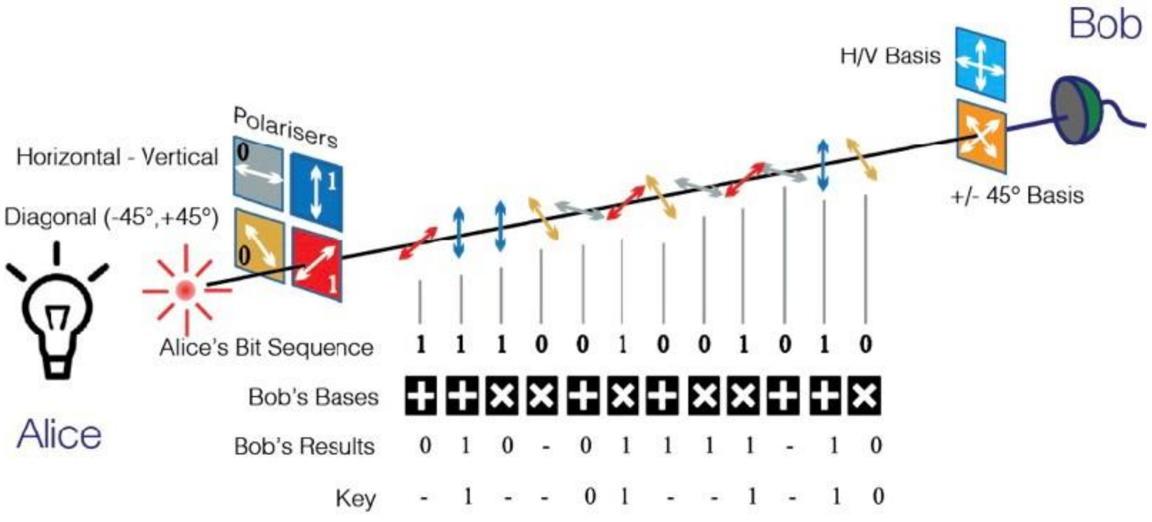
Extend the system for training:

- Data sharing between institutions for DL training
- Scenario:
  - Client input: Encrypted Medical Datasets split N participants.
  - Server trains the model privately.
  - Requirement: The participants learn nothing about the input of other participants.

- Open Source
- Adapted to common DL frameworks.

Quantumacy, OpenQKD and Quantum Key Distribution

- Quantum Key Distribution (QKD) enables the creation of a symmetric key profiting from quantum properties.
- Enhancing the privacy of our system through QKD to enhance Privacy Preserving Processes.
  - E.g., communication between healthcare institutions.
- Enabling quantum-secure key-exchange-like processes.



Source: (Quantum Flagship)

Conclusions

- New technologies which are changing paradigm of data processing.
- Enable privacy approaches for data analytics is crucial.
- Open questions:
  - Efficiency improvements.
  - Convergence of training with approximations.

References

- [1] Topol, Eric J. "High-performance medicine: the convergence of human and artificial intelligence." Nature medicine 25.1 (2019): 44-56.
- [2] Saritha, Saladi, and N. Amutha Prabha. "A comprehensive review: Segmentation of MRI images—brain tumor." International Journal of Imaging Systems and Technology 26.4 (2016): 295-304.
- [3] Bauer, Stefan, et al. "A survey of MRI-based medical image analysis for brain tumor studies." Physics in Medicine & Biology 58.13 (2013): R97.

QUESTIONS?

jose.cabrero.holgueras@cern.ch