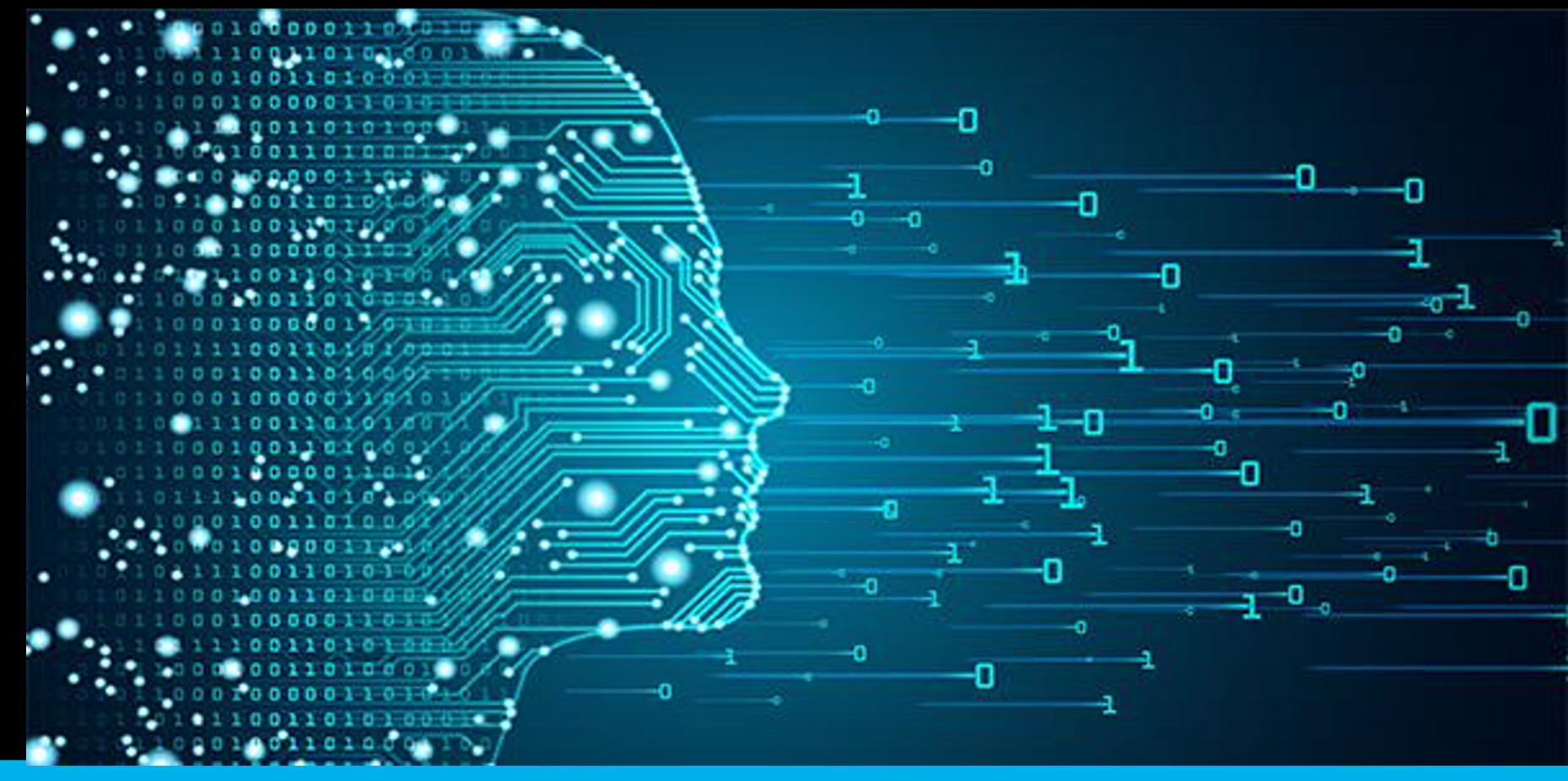
Patient-Reported Outcomes: A Privacy-Centric and Federated Approach to Machine Learning



Sara Jordan[1], Rachele Hendricks-Sturrup[1], [1]Future of Privacy Forum, Washington, DC, USA

Abstract

Machine learning developers, healthcare practitioners, ethicists, and privacy professionals can all agree: health data requires special protections. Should certain types of health data warrant specific protections? In this scientific position paper, we answer this question affirmatively and stake the position that patient reported outcomes measures (PROs/PROMs) data requires specific protections, particularly when it is used in or informed by machine learning regimes. We posit that a federated learning architecture is appropriate, both when model components are sent or gradients received, for ensuring privacy of users' data. However, use of a federated learning approach for privacy without attention to security dimensions of private data management may compromise users' data unexpectedly.

Scientific Position

- Specific forms of health data require specific protections by stating the claim that PROM data should be protected in specific ways.
- When collected through user's mobile devices, PROM data should be protected by use of federated data architecture, while ensuring high standards of quality in reporting and managing patient data and consent to privacy policy terms and conditions.
- When such data is part of machine learning enabled PROM applications, a federated approach to machine learning can be essential to protect users' privacy, to protect the investments in valid PROM instruments, and to protect investment in model development.

PROs & PROMs

- PROs help clinicians, researchers, medical device and drug manufacturers, and governmental stakeholders overseeing medical device and drug development, distribution, and safety monitor, understand, and document, in a readable format, patients' symptoms, preferences, complaints, and/or experiences following a clinical intervention.
- PROMs gather quantitative and qualitative data reflecting the health status of a patient, allowing for the data to be used to correlate or predict serious adverse events like hospitalizations for acute and life-threatening symptoms.
- Many entities are exploring ways to capture longitudinal PROMs using mobile health technology and ways to integrate and combine PROMs with data from multiple sources like wearable/mobile device data.

Privacy Concerns

- A systematic literature review of benefits and disadvantages to the electronic capture of PROs identified and described privacy protection, or lack thereof, as a key disadvantage to the systematic and routine collection of PROMs.
- Since PROM data shared via mobile apps can reveal fine grained information about a patient or caregiver, systems using this data should adopt a nuanced approach to creating and implementing privacy policies and machine readable patient consent to uphold patient trust.

Privacy Policies and Machine-Readable Consent

- Machine-readable privacy policies ensure that meta-apps are useful and operate well.
- Machine-readable consent terms allows for consent tools to become data themselves and searchable queries and can capture any spectrum of patient preferences when express consent is offered in a nuanced way.

Federated Approaches to Privacy-Preserving Machine Learning

- On-device PROM tools that use machine learning may employ the best privacy policies or consent mechanisms, but may ultimately leave key components, such as the security dimensions, of privacy up to the user.
- On both client-side and server-side learning, there must be a reasonable expectation that regular transmission patterns will not open the models for attacks, whether upon the model or upon users' data.

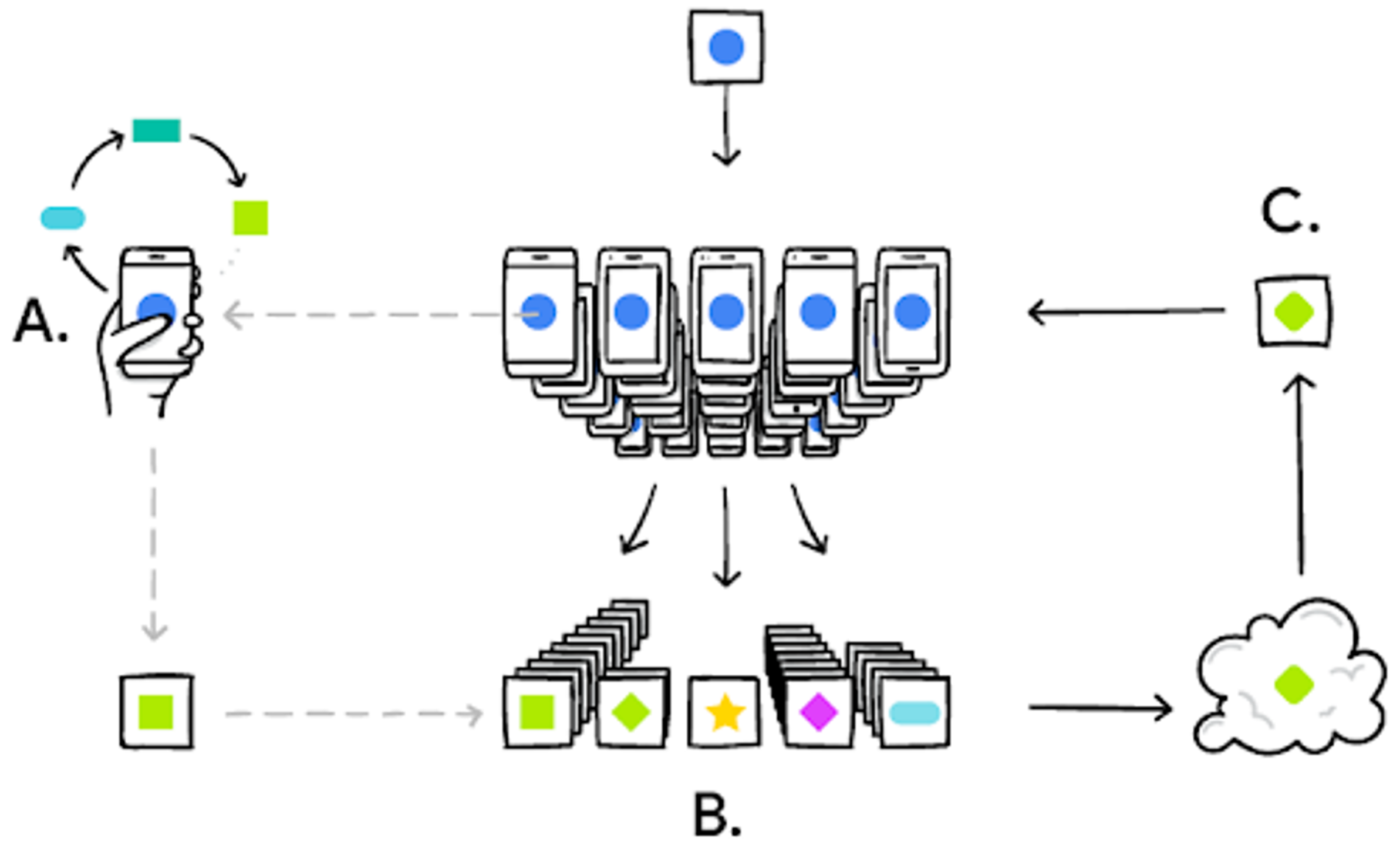


Figure 1. Google AI illustrates federated learning as "Your phone personalizes the model locally, based on your usage (A). Many users' updates are aggregated (B) to form a consensus change (C) to the shared model, after which the procedure is repeated"

Conclusion

- Developers should ensure that:
  - Choices about models do not open users to attack or undue influence and thus do not open practitioners to liability for interpretation of false responses;
  - Models are not compromised and valuable machine learning spending lost to competitors; and
  - Models are tested and validated to ensure quality of unstructured PROM data versus influencing or skewing PROs concerning safety, symptoms, and other important outcomes.

References

Scan QR code:

FUTURE OF PRIVACY FORUM