

PhD Annual Progress Report

School of Computing & Information Science · University of Edinburgh

Student	Amara J. Osei
Student ID	s2184739
Programme	PhD Computer Science
Primary Supervisor	Prof. Elena V. Marchetti
Second Supervisor	Dr. Ryu Tanaka
Year of Study	Year 2 of 4
Reporting Period	January 2025 – June 2025
Submission Date	4 July 2025

Research Objectives

My thesis investigates **privacy-preserving federated learning for clinical NLP**: specifically, how language models can be fine-tuned across hospital siloes without centralising patient data. The three core objectives are:

1. Design and evaluate a differentially-private federated fine-tuning protocol for clinical BERT-style models across heterogeneous data distributions.
2. Quantify the utility–privacy trade-off on two real-world cohorts (MIMIC-III discharge summaries; NHS Scotland free-text notes).
3. Develop a lightweight audit mechanism enabling regulators to verify differential-privacy guarantees without accessing raw model weights.

Progress This Period

Literature & baseline. Completed a systematic review of 78 papers on federated NLP and clinical text mining. Established a centralised (non-private) fine-tuning baseline: Bio-BERT achieves $F_1 = 0.847$ on NER of clinical entities in MIMIC-III held-out test set, confirming the upper bound for subsequent federated experiments.

DP-Fed implementation. Implemented DP-FEDAVG with Gaussian noise mechanism in PyTorch + FLOWER (fl.dev). Ran ablations over privacy budgets $\epsilon \in \{1, 2, 4, 8\}$ with $\delta = 10^{-5}$. At $\epsilon = 4$ the federated model recovers 96.1% of baseline F_1 .

Audit prototype. Drafted design for a zero-knowledge proof-based audit log; cryptographic primitives selected; implementation begins Q3 2025.

Milestones & Timeline

Milestone	Target Date	Status
Literature review & systematic survey	March 2025	Completed
Centralised baseline model	March 2025	Completed
DP-FedAvg implementation & ablations	June 2025	Completed
NHS Scotland data access agreement	August 2025	In Progress
ZK-proof audit prototype (v1)	October 2025	Planned
First conference paper submitted	November 2025	Planned
Transfer viva (18-month review)	January 2026	Planned
Full thesis draft	December 2026	Planned

Publications & Presentations

Workshop paper (under review) A. Osei, E. V. Marchetti, R. Tanaka. “Differentially-Private Federated Fine-Tuning of Clinical Language Models.” *PrivateNLP @ ACL 2025*, Vienna, Austria. (Notification: Aug 2025.)

Poster A. Osei. “Privacy Budgets in Federated Clinical NLP.” *Edinburgh Doctoral Research Conference*, March 2025. **Best Poster Award — School of Computing.**

Internal seminar Presented progress to the NLP & Healthcare reading group (May 2025, 25 attendees).

Challenges

- **Data access delays.** Negotiations with NHS Scotland’s Information Governance team have taken longer than anticipated. A revised data-sharing agreement is expected by July 2025; contingency plan uses synthetic EHR data if approval extends past August 2025.
- **Compute constraints.** Large-scale ϵ -sweep experiments require significant GPU time. EPSRC Cloud allocation (Tier 2) has been applied for; interim experiments run on the School’s 4×A100 cluster.
- **ZK-proof complexity.** Adapting Groth16 proofs to floating-point DP noise accounting is non-trivial; additional reading on arithmetic circuits planned for Q3.

Plan for Next Period (July – December 2025)

1. Finalise NHS Scotland data access and conduct cross-silo experiments on real clinical text (*target: September 2025*).
2. Complete ZK-proof audit prototype and benchmark overhead (*target: October 2025*).
3. Submit conference paper to PrivateNLP@ACL or EMNLP 2025 workshop (*target: November 2025*).
4. Begin writing Chapter 3 (federated protocol) of thesis (*target: December 2025*).
5. Prepare 18-month transfer viva documents (*target: January 2026*).

Supervisor’s Comments

Prof. Elena V. Marchetti

Primary Supervisor

Amara has made excellent progress during this reporting period, particularly in establishing a rigorous centralised baseline and completing a thorough systematic review ahead of schedule. The DP-FedAvg results are promising and clearly demonstrate methodological competence. I am satisfied with the quality and pace of work.

The principal concern for the next period is the NHS Scotland data access timeline; I encourage Amara to continue pressing Information Governance and to begin parallel experiments on the synthetic cohort so that momentum is not lost. The zero-knowledge audit extension is ambitious but well-motivated — I suggest scoping the initial prototype narrowly (single-layer noise accounting) to de-risk the November submission deadline.

Dr. Ryu Tanaka

Second Supervisor

Progress is on track. I concur with the above and add that the poster award recognition is a positive sign of communication quality. I look forward to reviewing the ZK-proof design document in September.

Supervisor signature: _____ *Date:* _____

Student signature: _____ *Date:* _____