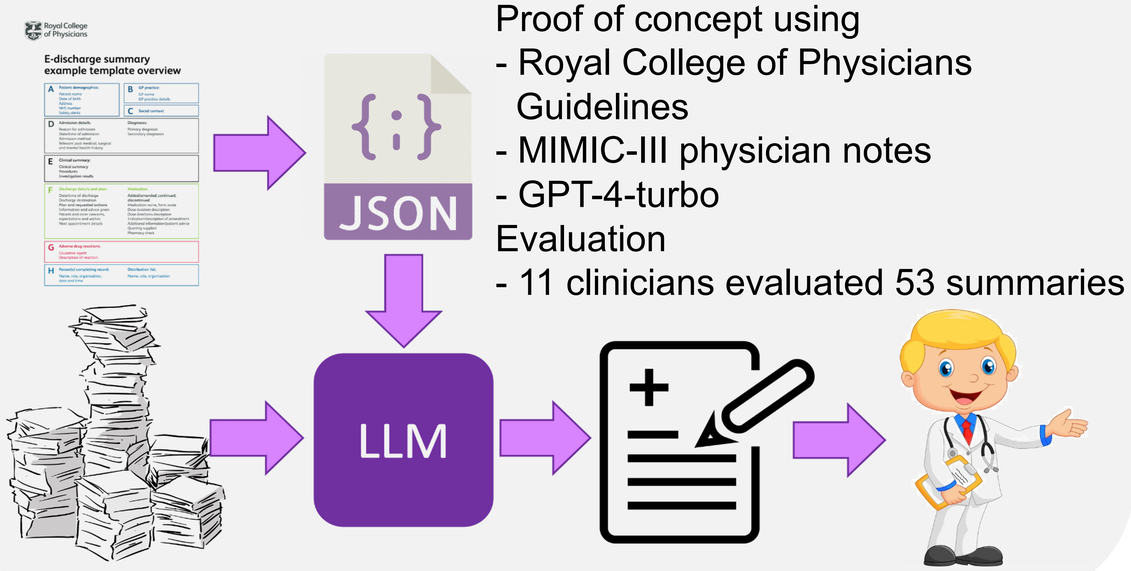
# Automated Generation of Hospital Discharge Summaries Using Clinical Guidelines and LLMs

Simon Ellershaw[1], Christopher Tomlinson[1,2,3], Oliver Burton[4], Thomas Frost[1], John Gerrard Hanrahan[4,5], Danyal Z Khan[4,5], Hugo Layard Horsfall[4,5], Mollie Little[4], Evaleen Malgapo[1], Joachim Starup-Hansen[4], Jack Ross[4], George Woodward[4], Martinique Vella-Baldacchino[6], Kawsar Noor[1,2,3], Anoop D Shah[1,2,3] Richard JB Dobson[1,2,3,7]

1 Institute of Health Informatics, University College London, UK
2 National Institute for Health and Care Research Biomedical Research Centre, University College London Hospitals National Health Service Foundation Trust, UK
3 Health Data Research UK, London, UK
4 University College London Hospitals NHS Foundation Trust, London, UK
5 Wellcome/EPSRC Centre for Interventional and Surgical Sciences, University College London, UK
6 MSK Lab, Imperial College London, London, UK
7 Department of Biostatistics and Health Informatics, King's College London, UK

## Method




Proof of concept using
- Royal College of Physicians Guidelines
- MIMIC-III physician notes
- GPT-4-turbo

Evaluation
- 11 clinicians evaluated 53 summaries

## Findings

- Clinical guidelines make good prompts
- Few shot learning overcomes training data scarcity
- LLMs are good at manipulating language but poor at making clinical judgements

## Results

| Section | Recall | Precision | Acc |
|---|---|---|---|
| Admission Details | 0.90 | 0.95 | 0.85 |
| Allergies And Adverse Reaction | 0.98 | 1.00 | 0.98 |
| Clinical Summary | 0.76 | 0.92 | 0.71 |
| Diagnoses | 0.84 | 0.94 | 0.80 |
| Discharge Details | 0.93 | 0.96 | 0.89 |
| Patient Demographics | 1.00 | 0.84 | 0.84 |
| Plan And Requested Actions | 0.90 | 0.88 | 0.80 |
| Social Context | 0.96 | 0.88 | 0.84 |
| Macro Average | 0.91 | 0.92 | 0.84 |
| Micro Average | 0.86 | 0.92 | 0.81 |

## Path to Clinical Adoption

Improved evaluation
- Data sources
- Clinical framework (low annotator agreement 59.72%)
- Automated evaluation for methodological hill climbing

Data governance-compliant LLM

Workflow integration
- Human in the loop -> automation bias
- Handling of error cases
- EHR Integration

Paper

Code

Contact

UK Research and Innovation