

NICTA eHealth initiatives on clinical speech and text processing

Hanna SUOMINEN ^{a,1}, David MARTINEZ ^b, Mike WU ^c,
Michelle ANANDA-RAJAH ^d and Lawrence CAVEDON ^b

^a *PhD; Canberra Research Laboratory, NICTA, National ICT Australia; College of Engineering & Computer Science, Australian National University; Canberra, ACT*

^b *PhD; Victoria Research Laboratory, NICTA; Computing and Information Systems, The University of Melbourne; Melbourne, VIC*

^c *MEng; Neville Roach Laboratory, NICTA; Sydney, NSW*

^d *Dr; Infectious Diseases Unit, Alfred Health; Melbourne, VIC*

Introduction

We discuss three initiatives from the *NICTA eHealth Business Team* towards developing ICT techniques for improving healthcare via:

1. capturing more clinical data;
2. converting them to processable information;
3. recording this as comprehensive electronic documents;
4. generating visualisations to support end-user' situational awareness and decision making.

The initiatives have been conducted with the proper research permissions and favourable ethics reviews.

Technology Brief and Implementation Processes

The *Clinical Speech to Text* initiative aims to automatically transcribe verbal communication from nursing handovers into written drafts for clinical review, editing, and addition to electronic health records. We are developing ICT for generating:

1. speech-to-text conversion,
2. standardisation with respect to controlled thesauri, and
3. structuring in accordance with documentation standards.

We will evaluate the usability, accuracy and acceptability in challenging settings characterised by high ambient noise levels and multiple speakers. We have:

1. built a strategic consortium of

¹ Corresponding and Presenting Author: Researcher, PhD; NICTA Canberra Research Laboratory, Locked Bag 8001, Canberra ACT 2601, Australia; E-mail: hanna.suominen@nicta.com.au.

- ICT providers (Cerner, Lingsoft, Telehealth Research and Innovation Laboratory, Universities of Western Sydney and Wollongong),
 - clinical researchers (Centre for Applied Nursing Research, Clinical Excellence Commission),
 - linguists (The University of Melbourne), and
 - end users (Local Health Districts of Sydney and South Western Sydney);
2. demonstrated the suitability of our content structure for clinical handover by successfully implementing this documentation model across three major teaching hospitals and into six wards in Sydney; and
 3. designed all technologies, evaluations, surveys and clinical scenarios for the in-lab phase.

The *Text Mining for In-Hospital Surveillance* initiative is developing ICT which can automatically mine clinical narratives (e.g., radiology reports) to detect conditions of interest. Its implementation provides:

1. automatic detection of specified conditions and
2. visualisation capabilities to support hospital-wide analysis and inspection of individual narratives.

We have:

1. collected over 2,000 authentic records from our partners at Alfred Health, Melbourne Health, and Peter MacCallum Cancer Centre;
2. expert-annotated a subset of approximately 400 and 100 records from diagnosed and control patients respectively; and
3. built a machine learning system on the annotated subset and used it to detect high-risk patients from the records of the over 400 remaining patients with sensitivity/specificity of 0.97/0.55.

Using the annotated subset and ten-fold cross-validation at a patient level to build and evaluate the system, it

1. detects high-risk reports with sensitivity/specificity of 0.95/0.72 and
2. identifies high-risk patients earlier than our expert-annotators in 10% of cases.

The *Privacy-Preserving Benchmarking and Text Mining* initiative addresses two challenges of clinical language processing:

1. dissemination of data, resources and ICT and
2. systematic benchmarking of methods and resources on a rich variety of clinical datasets.

Its implementation includes:

1. a meta-framework for language processing and
2. its demonstrations.

Outcomes include the NICTA *Epicure* meta-framework and web-server. The server provides a control layer for interfacing with third-party ICT systems. Its storage layer is used for saving and accessing data to be analysed, resources to be used in the analysis, and gold-standards to be used in the evaluation. An execution layer invokes processing methods or visualisation tools. Currently, we use it to demonstrate our second initiative.

Conclusion

Our initiatives have the potential to:

1. improve clinical documentation without interrupting the workflow of healthcare workers;
2. support early diagnosis of important diseases and enable automated surveillance across a hospital system;
3. contribute to research, evaluation and development of ICT for Health; and
4. strengthen collaboration along the eHealth value chain comprising practitioners, policy-makers, researchers and industry.

We are seeking partnerships to trial our technologies, integrate them into hospital information systems, and scale them to other tasks.

Keywords. Clinical Decision Support Systems, Electronic Health Records, Information Storage and Retrieval, Natural Language Processing, Speech Recognition Software

Acknowledgements. NICTA is funded by the Australian Government as represented by the Department of Broadband, Communications and the Digital Economy and the Australian Research Council through the ICT Centre of Excellence program. Hanna Suominen was partially funded by the Academy of Finland (decision number 136653). We thank Sumithra Velupillai (Stockholm University) and Karl Kreiner (AIT, Austrian Institute of Technology) for their contribution to the initiatives during their research visits at NICTA. We express our gratitude to Leif Hanlen (NICTA) for directing the initiatives, to Jim Basilakis and Diana Jefferies (University of Western Sydney), Linda Dawson (University of Wollongong), Maree Johnson and Bronwyn Overs (Centre for Applied Nursing Research), Barbara Kelly (The University of Melbourne) and Anthony Yeo for bringing multi-disciplinary expertise to the first initiative and to Monica Slavin and Karin Thursky (Alfred Health, Peter MacCallum Centre, Royal Melbourne Hospital) for providing domain expertise to the second initiative.