MDENet *Seedcorn project*

NLP and ML for automated requirements formalization

Principal Investigator: Chenghua Lin, University of Sheffield
Co-Investigator: Kevin Lano, King's College London
Research Assistant: Shekoufeh Kolahdouz-Rahimi, University of Roehampton

Project Aim

- Construct a framework for requirement formalisation (RF) whereby different NLP and ML tools can be combined to provide alternative RF pathways
- Evaluate different approaches on set of standard tasks, e.g., mining of unstructured documentation; formalisation of semi-structured documents; formalisation of structured documents.
- Define guidelines for selection of RF techniques appropriate for different RF tasks.

Project outcomes

We carried out systematic survey of existing RF approaches, including NLP and deep learning (DL) approaches across wide range of applications.

- **250** publications were examined
- **49** specific publications selected for deeper analysis

Survey Findings

In our systematic survey, we identified

- Heuristic NLP approaches are most common RF technique in research, primarily operating on structured and semi-structured data
- Deep learning techniques *not* widely-used, instead classical ML techniques such as decision trees and Support Vector Machine (SVM) are used in the surveyed studies.
- There is a lack of standard benchmark cases for RF and therefore it is difficult to compare the performance of different approaches.

Survey Findings

From this survey we concluded that the **RF field remains at experimental stage**:

- **Evaluation is not performed in systematic manner** and difficult to compare different approaches or reproduce results.

- **Performance and completeness of heuristic approaches typically not evaluated on broad range of input cases**, and not possible to determine which are best to use in different situations. Heuristic approaches inevitably have incompleteness, may only work well for inputs restricted to particular formats or linguistic styles.

- **Under-use of DL techniques** which could be relevant and applicable to RF tasks, and help to avoid limitations of heuristic approaches, especially for unstructured source data. Systematic comparison of DL and heuristic solutions would help to identify appropriate techniques for particular kinds of RF tasks.

- **Standard benchmarks and evaluation criteria need to be established**, in order to systematically compare different RF cases. E.g., there are well-established benchmarks in

  - Natural Language Generation (GEM https://gem-benchmark.com/)
  - Speech Processing (SUPERB https://superbbenchmark.org/)

Project Outputs

NLP DSLs: (1) To strengthen RF research, we developed a DSL for NLP pipelines, based on SQLite grammar; (2) enables high-level definition of NLP pipelines for RF, independent of any particular NLP platform such as NLTK, Stanford NLP or Apache OpenNLP.

- A transformation from the DSL to Python was defined to support implementation in NLTK.
- An OCL-based version of DSL was defined, together with supporting tools.
- DSLs were evaluated by applying them to specify & implement typical NLP pipeline tasks.

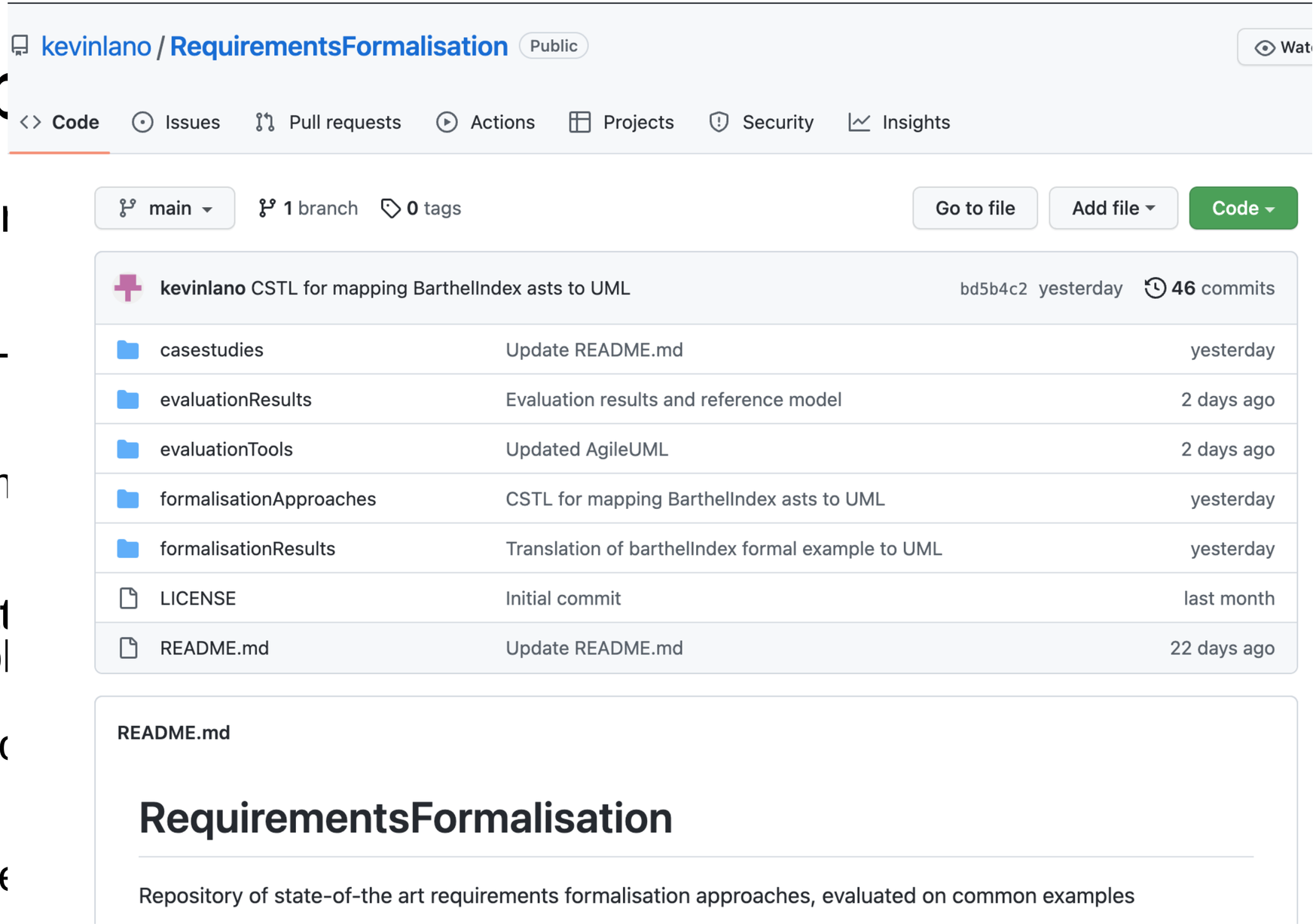
Project Outp

A RF repository: to pr
established repository

- links to state-of-the-
evaluations
- resource for RF con
world problems.

Standard set of evaluat
approaches, need establ

- Selected 25 cases
of user stories
- Also real-world case

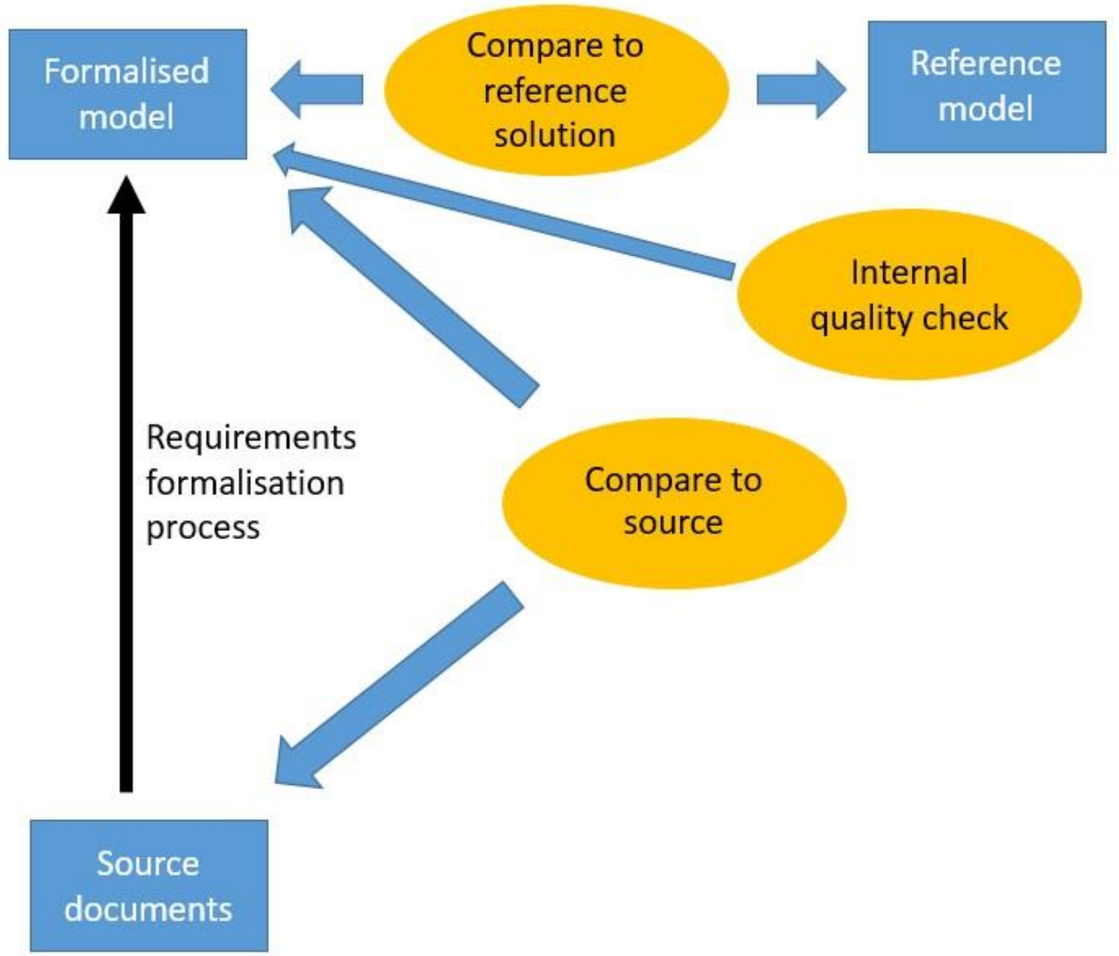


- Evaluated results for several different approaches applied to these cases.

Project Outputs

Developed tools to compare and evaluate RF approaches.

- compare formalized models produced by an approach to manually constructed reference models
- compare formalized models to source document, to check completeness
- evaluate internal quality of formalized model.



Future work

- We will develop RF repository with more case studies, evaluation tools and evaluations, and publicise in MDE forums and invite contributions from RF researchers.
- Work of comparative analysis will be continued by a PhD student at King’s College.
- An EPSRC research grant will be prepared on the topic, to be submitted in Spring 2023.