Research on NLP for RE at Utrecht University

A Report

Fabiano Dalpiaz and Sjaak Brinkkemper
Requirements Engineering Lab
Utrecht University, the Netherlands
March 18, 2019
1. Team overview
Requirements Engineering

Requirements Engineering (RE) is the discipline that is concerned with understanding, modeling and specifying, analyzing and evolving the requirements of software systems. The Requirements Engineering Lab (RE-Lab) at Utrecht University is involved in several research directions with the common objective to help people express better requirements in order to ultimately deliver better software. Our recipe involves the use of state-of-the-art, innovative techniques from various disciplines (computer science, logics, artificial intelligence, computational linguistics, social sciences, psychology, etc.) and to apply them to solve real-world problems in the software industry.
Research themes, illustrated
The RE-Lab team

Principal investigators

Researchers

External members

Master’s students
GitHub repository

https://github.com/RELabUU
2. NLP research at the RE-Lab
User story requirements:
As a NLP4RE attendee,
I want to see the presentations schedule,
so that I can skip Fabiano’s talk.
As a user,
I want to be able to select different types of recyclable waste, so I have and get a list of facilities that accept each type and their opening hours, so that I can find an optimal route and schedule.

As a Publisher,
I want to know if this site has a pricing plan and what the prices are, so that I can work out what …

I want to print a report, so that my customers consider me a professional consultant.
AQUSA: evaluation

- Five criteria implemented
  - Precision 72%
  - Recall 93%
  - **Original goal:** 100% recall

- Longitudinal study in three companies for two months
  - Better user stories
  - No improvements of project mgmt. metrics

---

**Improving Agile Requirements: The Quality User Story Framework and Tool**

Garm Lucassen · Fabiano Dalpiaz · Jan Martijn E.M. van der Werf · Sjaak Brinkkemper

---

**Improving user story practice with the Grimm Method: A multiple case study in the software industry**

Garm Lucassen, Fabiano Dalpiaz, Jan Martijn E.M. van der Werf and Sjaak Brinkkemper
Extraction of conceptual models

As a $\langle visitor\rangle_{ent}$,

$\langle I\rangle_{=visitor}$ want to $\langle choose\rangle_{rel}$ an $\langle event\rangle_{ent}$

so that $\langle I\rangle_{=visitor}$ can $\langle book\rangle_{rel}$ a $\langle ticket\rangle_{ent}$ for that $\langle event\rangle_{ent}$
Extraction of conceptual models

As a \( \langle \text{visitor} \rangle_{\text{ent}} \),
\( \langle I \rangle_{\text{visitor}} \) want to \( \langle \text{choose} \rangle_{\text{rel}} \) an \( \langle \text{event} \rangle_{\text{ent}} \)
so that \( \langle I \rangle_{\text{visitor}} \) can \( \langle \text{book} \rangle_{\text{rel}} \) a \( \langle \text{ticket} \rangle_{\text{ent}} \) for that \( \langle \text{event} \rangle_{\text{ent}} \).

As a visitor,
I want to filter on free events
so that I can save money.
Support tool: Visual Narrator

User story requirements → Visual Narrator → Conceptual model

- Conflicts
- Duplicates
- Incompleteness
- ...

Extracting Conceptual Models from User Stories with Visual Narrator

Garm Lucassen · Marcel Robeer · Fabiano Dalpiaz · Jan Martijn E.M. van der Werf · Sjaak Brinkkemper

RE Journal 17
Interactive Narrator (a rendering engine for the Visual Narrator)
Visual Narrator: results

- **Positive results**
  - High precision and recall in the extracted concepts (~90%)
  - Perceived useful for training learners by practitioners

- **Negative results**
  - Low cognitive scalability: we moved from large collection of user stories to large models
  - NLP issues
    - Hard to cope with compound nouns
    - Difficult to associate the right object to the verb
Terminological ambiguity

- Quasi-synonyms in user stories
  - Problem: are those two words referring to the same entity?
    - image gallery – gallery
    - image – picture
    - to view – to see
  - Idea: to combine semantic similarity with info. visualization
Terminological ambiguity

- The REVV-Light tool
  - Input = Visual Narrator’s output
  - Calculates semantic similarity between the terms
    - Semantic fingerprinting
  - Synonyms are possible ambiguities
Experiments with REVV-Light

- Quasi-experiment against manual inspection
  - 28 real-world data sets, 2,000+ requirements

- Results about our approach
  - Manual inspection was better in the time constrained setting
  - High usability expectations by the participants
  - The similarity algorithm needs context information!

- General finding: reaching consensus on ambiguity is hard!

Detecting terminological ambiguity in user stories: Tool and experimentation

Fabiano Dalpiaz\textsuperscript{a,*}, Ivor van der Schalk\textsuperscript{a}, Sjaak Brinkkemper\textsuperscript{a}, Fatma Başak Aydemir\textsuperscript{b}, Garm Lucassen\textsuperscript{c}

IST Journal 19
Requirements data sets (user stories)

Published: 28 Jul 2018 | Version 1 | DOI: 10.17632/7zbk8zsd8y.1
Contributor(s): Fabiano Dalpiaz

Description of this data

A collection of 22 data sets of 50+ requirements each, expressed as user stories. These were all found online, or retrieved from software companies with a permission to disclose.

The data sets have been originally used to conduct experiments about ambiguity detection with the REVV-Light tool: https://github.com/RELabUU/revv-light

Experiment data files

- g02-federalspending.txt (11 KB)
- g03-loudoun.txt (9 KB)
- g04-recycling.txt (7 KB)
- g05-openspending.txt (9 KB)

Latest version

Version 1
Published: 2018-07-28
DOI: 10.17632/7zbk8zsd8y.1

Cite this dataset

Dalpiaz, Fabiano (2018), “Requirements data sets (user stories)”, Mendeley Data, v1
http://dx.doi.org/10.17632/7zbk8zsd8y.1

Statistics

Views: 130  Downloads: 15
Requirements from competitors

The RE-SWOT method

1. Identify features and transform ratings

<table>
<thead>
<tr>
<th>Review</th>
<th>Feature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Review 1</td>
<td>Feature A</td>
</tr>
<tr>
<td>Review 2</td>
<td>Feature A</td>
</tr>
<tr>
<td>Review 3</td>
<td>Feature B</td>
</tr>
<tr>
<td>Review 4</td>
<td>Feature C</td>
</tr>
</tbody>
</table>

2. Calculate FPS

<table>
<thead>
<tr>
<th>Feature</th>
<th>App1</th>
<th>App2</th>
<th>App3</th>
<th>Avg</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feature A</td>
<td>0.2</td>
<td>0.5</td>
<td>0.3</td>
<td>0.3</td>
</tr>
<tr>
<td>Feature B</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>Feature C</td>
<td>-0.5</td>
<td>-0.5</td>
<td>-</td>
<td>-0.5</td>
</tr>
</tbody>
</table>

Reference app

App 2

3. Generate RE-SWOT matrix

4. Generate requirements

<table>
<thead>
<tr>
<th>Req 1</th>
<th>Req 2</th>
<th>Req 3</th>
<th>Req 4</th>
</tr>
</thead>
</table>

NLP: feature extraction
The RE-SWOT Matrix

<table>
<thead>
<tr>
<th>Feature Performance Score</th>
<th>App</th>
<th>Reference app</th>
<th>Competitor app</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive and above market average</td>
<td></td>
<td>Strength</td>
<td>Threat</td>
</tr>
<tr>
<td>Negative and below market average</td>
<td></td>
<td>Weakness</td>
<td>Opportunity</td>
</tr>
</tbody>
</table>
The RE-SWOT Matrix, visualized

Features in common
- lightroom: edit photo, raw file
- snapseed: editing tool, raw file, user friendly
- VSCO: raw photo, import photo, edit photo, iphone x, photo editing

TO START
Type your app name:
lightroom
Filter by SWOT
Multiple values
Filter by feature volume
Multiple values
Filter by quarter
2017 Q4

Legend - SWOT
- Opportun..
- Weakness
- Strength
- Threat

Features unique to competition
- lightroom: new version, mobile version, new interface, new UI, white background, favorite editing
- snapseed: favorite editing, new UI
- VSCO: upload video, social media, video editing, free filter

Features unique to lightroom
- lightroom: desktop version, mobile version, lightroom

Full talk on RE-SWOT at REFSQ’18
Session 2: Tuesday 19/3 at 13:30
3. Future directions
A. Linking reqs to architectures

- Establish traceability links via linguistic analysis
- Especially useful in software product companies
  - The linkage can be assisted by glossaries
B. Automated elicitation via chatbots

- RE-Lab’s research paradigm in the past few years
  - NLP
  - Information Visualization
  - Human analyst

- Future paradigm: Chatbot conversation
C. Synthesis of creative requirements

- Kano’s model

- Can we automatically synthesize creative/exciting requirements?

- Work-in-progress with
  - Semantic similarity
  - Semantic role labeling

- The challenge? Requirements that make sense!
Thanks from the Requirements Engineering Lab at Utrecht University!

Fabiano Dalpiaz
Sjaak Brinkkemper
Marcela Ruiz
Sietse Overbeek
Davide Dell’Anna
Eduard C. Groen
Gerard Wagenaar

f.dalpiaz@uu.nl