Development of a recommender system for the selection of software architecture methods

Florian Mittrücker, Master Thesis - Initial Presentation

Software Engineering für betriebliche Informationssysteme (sebis)
Fakultät für Informatik
Technische Universität München

www.matthes.in.tum.de
Outline

Background and motivation
Research question & roadmap
Current state
Next steps
Outline

Background and motivation

- Research question & roadmap
- Current state
- Next steps
Siemens is one of the largest companies in electrical- and electronic engineering worldwide!

### Key facts

#### Facts
- Established in 1847 in Berlin
- 343,000 employees worldwide

#### Office Locations
- In 190 countries & 125 in Germany

### Service & product profile

#### Activities in different sectors like
- Healthcare / Industry Automation
- Power generation

#### Product & Services
- X-ray machines / Power plants
- PLM Software

### Company development

#### Profit after tax

<table>
<thead>
<tr>
<th>Year</th>
<th>Profit after tax (million €)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2014</td>
<td>5.507</td>
</tr>
<tr>
<td>2013</td>
<td>4.409</td>
</tr>
<tr>
<td>2012</td>
<td>4.282</td>
</tr>
</tbody>
</table>

#### Employees

<table>
<thead>
<tr>
<th>Year</th>
<th>Employees (thousand)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2014</td>
<td>343</td>
</tr>
<tr>
<td>2013</td>
<td>348</td>
</tr>
<tr>
<td>2012</td>
<td>352</td>
</tr>
</tbody>
</table>

Leads to…

- usage of several thousands of applications and a huge IT-landscape
- high demand on software architecture knowledge & management
Background and motivation

Siemens internal software architecture definition and management department is in charge of

**In-house consulting**
- Code / Architecture reviews
- Guidelines for development
- IT-Project management

**Research**
- EU sponsored projects
- Cooperation with academic institutes
- Internal projects for improvements & innovations

**Consulting**
- Software Architects are involved in internal- and research projects

**Expertise**
- Gather new expertise and experiences through projects

**Knowledge need**
- Requires knowledge about software architecture methods

**Knowledge sources**
- Minds of architects
- Intranet
- Literature

**Problems**
- No central knowledge (method) source for architecture methods
- No assessment of known methods
- No IT supported recommendation mechanism for method selection
Siemens started an internal project to develop an architecture management toolbox which supports software architects in their daily work.

**AMELIE**

AMELIE = Architecture Management Enabler for Leading Industrial softwarE

**Objectives of AMELIE**

- Guide architects to perform architecture management.
- Ensure that software architecture, business strategy and innovation go hand in hand.
- Be “in control” of architecture development and “in sync” with the business.
- Foster experience sharing inside Siemens

**AMELIE ecosystem**

Value added services of AMELIE

- Knowledge Management
- Recommendation Mechanism
- Architecture Tracking
- Simulation
- Cost and Benefit Estimation
Background and motivation
AMELIE - The architecture management workbench

User Stories (US) describing the scope of the AMELIE workbench

- **US-01**: Step-wise guidance for executing SW-Development project
- **US-02**: Provide list of recommended methods
- **US-03**: Best practice recommendation
- **US-04**: “Real” instantiation of method / artefact
- **US-05**: Reference to an expert

AMELIE

Facets
- Business Case
- Requirements
- F’ Arch
- T’ Arch

Modeling Workbench

Methods
- BizMo
- **Business Model Canvas**
- …

Experts
- Bartholdt
- Hassel
- Wengatz

>> MagicDraw

Florian Mittrücker - Master Thesis
**Background and motivation**

Practical context / Examples for software architecture methods

---

### Terms & relations

**Facet**
- is like a development step (e.g. *Business Case*, *Requirements Elicitation*), which is passed through architecture development.

**Topic**
- is like a category (e.g. *Situation Analysis*) which contains several architecture methods to realise the topic specific objectives.

**Architecture method**
- is a method (e.g. *5C Business Analysis*) to achieve objectives of a topic

---

**Example: 5C Business Analysis**

Belongs to topic *Situation Analysis* which is part of the facet *Business Case*.
- Focuses on the business environment of the product
- It covers collaborators, customers, competitors, own company resources as well as the context of the current technology available in the market.
- Results are e.g. roadmaps, business cases and decisions on possible solution variants.

---

**My tasks**
- Structuring of methods and context information
- Development of a recommendation mechanism for software architecture methods
Outline

Background and motivation

**Research question & roadmap**

Current state

Next steps
How should a recommender system be designed in order to be appropriate for knowledge management of software architecture methods as well as for active recommendation during the development process?

1) Literature review
   - Recommender systems
     - in general and for software architecture methods
     - Deliverable: Literature classification scheme

2) Recommendation method/concept identification
   - Analysis of development process at Siemens
   - Method selection, based on existent process and literature
   - Deliverable: Appropriate recommendation method

3) Method instantiation
   - Concept development for recommendation system
   - Method instantiation (prototype) and qualitative evaluation
   - Deliverable: Evaluated concept and prototype

4) Writing + buffer time
   - Detailed writing & correction
   - Submission date 15.09.2015
   - Deliverable: Final thesis
Outline

Background and motivation
Research question & roadmap

Current state

Next steps
Objectives

- **Gather information** about concepts for recommender systems
- Establish **understanding** for **concepts** its functionalities and characteristics
- Create **classification scheme** to get a overview of most relevant articles

Research concept

Search terms

- Determination of search terms by means of initial research and existing literature reviews
  - Collaborative filtering / Contents filtering / Personalization system / Recommendation system
    Recommendation platform / Recommender system / Preference systems

Databases

- Several areas are involved:  Information retrieval / Forecast theories / Marketing …
- Five databases of EBSCO / Science Direct / Google Scholar

Analyse steps

1. Title & Abstract screening, Number of citations ➔ potentially relevant (Y/N)
2. Available (Y/N)
3. Analysis of article content, for- / backward search ➔ relevant (Y/N)
4. Classification of article

Source(s): (Armstrong, 2001 / Lilien, Kotler & Moorthy, 1992 / Park, Kim, Choi & Kim / 2012; Salton, 1989)
Classification scheme

- Structure of classification scheme was built up based on the abstracts of potentially relevant articles. Refinement was performed while reading.
- Most relevant content of the articles is described in my thesis

<table>
<thead>
<tr>
<th>No.</th>
<th>Journal/Source</th>
<th>Title</th>
<th>Author</th>
<th>Year</th>
<th>RS class description</th>
<th>Clustering approaches</th>
<th>Context of application</th>
<th>Focus of article</th>
<th>Filtering approaches</th>
<th>Properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>Knowledge and Data Engineering, IEEE Transactions on</td>
<td>Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions</td>
<td>Adamovicius, Gediminas; Tuzhilin, Alexander</td>
<td>2005</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>ACM Trans. Inf. Syst. Communications of the ACM</td>
<td>Item-based top-N recommendation algorithms</td>
<td>Deshpande, Mukund; Karpi, George</td>
<td>2004</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Communications of the ACM</td>
<td>Fab: content-based, collaborative recommendation</td>
<td>Balabanov, Marko; Shoham, Yeh</td>
<td>1997</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: Classification scheme (extract)

Added value
- Structured overview of most relevant articles
- Possibility to find articles very easy regarding a certain topic
Some knowledge for further actions - recommender / filtering approaches

We need profiles of elements and/or users. Profiles include a kind of preference or item evaluation which is used to generate useful recommendation.

Content recommendations
• User will be recommended items similar to the ones the user preferred in the past.

Collaborative recommendations
• The user will be recommended items that people with similar tastes and preferences liked in the past.

⇒ The utility of an unseen item is calculated based on passive or active feedback which is stored in profiles.

Hybrid approaches
• Combines elements of both collaborative and content-based methods.

Outline

Background and motivation
Research question & roadmap
Current state

Next steps
**Next steps**

**Methodology for further action**

<table>
<thead>
<tr>
<th>Design science (DS)</th>
<th>Action design research (ADR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>• DS is captured in the “build and then evaluate” cycle</td>
<td>• Reflects the premise that IT artefacts are ensembles shaped by the organizational context during development and use</td>
</tr>
<tr>
<td>• Focus on building artefacts and relegate evaluation to a subsequent &amp; separate phase</td>
<td>• Building and evaluating artefacts goes hand in hand</td>
</tr>
<tr>
<td>• Scant attention to the shaping of IT artefacts by the organizational context</td>
<td>• Various forms of the organizational context can be inscribed into the artefact during its development and use.</td>
</tr>
<tr>
<td>• Fails that the artefacts emerges from interaction with the organization</td>
<td>• Provides guidance for combining building, intervention and evaluation</td>
</tr>
</tbody>
</table>

➔ Need for a research method that explicitly recognizes artefacts as emerging from design, use and ongoing refinement in context.

**Added value**

• Design, use and ongoing refinement in context
• Methodology focuses practical use of artefact as well as a scientific approach
Next steps

- Completion of literature review
- Further analysis of internal development process
- Definition of data model for structuring architecture methods
- Development of recommendation concept
- Instantiation and final evaluation of concept (proof of concept/prototype)
Thank you for your attention!
Any questions?
Research concept

Search terms

- Determination of search terms by means of initial research and existing literature reviews (Park, Kim, Choi & Kim, 2012, P. 10060).
  - Collaborative filtering / Contents filtering / Personalization system / Recommendation system Recommendation platform / Recommendation engine / Recommender system / Preference systems

Databases

- EBSCO Business Source Premier / EBSCO EconLit / EBSCO Education Source / EBSCO ERIC / EBSCO Library, Information Science & Technology Abstracts / Science Direct /Google Scholar

Criteria (Initial Search):

- Must haves
  - Publication in academic journal
  - Publication date between 01.01.2000 and 31.03.2015 (Park, Kim, Choi & Kim, 2012, P. 10060);
  - Article must contain at least one of the search terms

- Exclusion criteria
  - Dissertations, unpublished working papers, textbooks, newspaper articles
  - Article is not completely in English or German
  - Article is not for free with TUM access rights

Source: (Park, Kim, Choi & Kim, 2012, P. 10060)
Design science vs. action design research

**DS in IS research - framework**

<table>
<thead>
<tr>
<th>Environment</th>
<th>Relevance</th>
<th>IS Research</th>
<th>Knowledge Base</th>
</tr>
</thead>
<tbody>
<tr>
<td>People</td>
<td></td>
<td>Develop/Build</td>
<td>Foundations</td>
</tr>
<tr>
<td>Organizations</td>
<td></td>
<td>Justify/Evaluate</td>
<td>Methodologies</td>
</tr>
<tr>
<td>Technology</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: (Hevner, March, Park, & Ram, 2004)

**ADR method: stages and principles**

1. **Problem Formulation**
   - Principle 1: Practice-Inspired Research
   - Principle 2: Theory-Inggrained Artifact

2. **Building, Intervention, and Evaluation**
   - Principle 3: Reciprocal Shaping
   - Principle 4: Mutually Influential Roles
   - Principle 5: Authentic and Concurrent Evaluation

3. **Reflection and Learning**
   - Principle 6: Guided Emergence

4. **Formalization of Learning**
   - Principle 7: Generalized Outcomes

Source: (Sein, Henfridsson, Purao, Rossi, & Lindgren, 2011)
Prototype

Objectives of prototype
• Proof of concept
• Demonstration of the concept, data model
• Demonstration of rudimental assessment and recommendation functionalities
• Evaluation of functionality and determination of future improvements

Not objectives of prototype
• Administration of content or users

Technology
• JAVA

User Interactions
• By means of console input/output

Source: (Park, Kim, Choi & Kim, 2012, P. 10060)