A User-Centered Methodology for the Evaluation of (Semantic) Web Service Discovery and Selection

Friederike Klan, Birgitta König-Ries

friederike.klan|birgitta.koenig-ries@uni-jena.de
Service Oriented Architectures

- Service consumer
- Semantic matchmaking
- Invocation
- Service provider
Web Service Discovery and Selection – B2B Scenario

inventory management of a retailer

invocation

semantic matchmaking

manufacturer

invocation

manufacturer

manufacturer
Web Service Discovery and Selection – B2B Scenario

- inventory management of a retailer
- invocation
- semantic matchmaking

- service requirements known in advance
- automatic selection by a machine
Web Service Discovery and Selection – B2C Scenario

flight booking web portal

invocation

semantic matchmaking

provider

provider

invocation
Web Service Discovery and Selection – B2C Scenario

- Flight booking web portal
- Invocation
- Semantic matchmaking
- Service requirements NOT known in advance
- Manual selection by a human

Manual selection by a human
### Changed System Requirements

- **service requirements NOT known in advance**
  - R1 interactive and incremental requirements elicitation
  - R5 maintain accurate requirements model

- **manual selection by a human**
  - R2 provide incentives for requirements construction and encourage to make tradeoffs
  - R3 educate about relevant service alternatives
  - R4 avoid inconsistent selection decisions
Evaluation Methodology
Evaluation Methodology

- **SYSTEM**
  - Initial system model: $SM_{init}$
  - Final system model: $SM_{fin}$
  - System usage

- **USER**
  - Initial mental model: $MM_{init}$
  - Mental model after system usage: $MM_{use}$
  - System usage
  - Browse
  - Mental model after exhaustive search: $MM_{exh}$
  - Selection after system usage: $Sel_{use}$
  - Selection after exhaustive search: $Sel_{exh}$

- Expressed requirements
Verifying Requirements – Requirements Construction (R2)

SYSTEM

SM\text{init} initial system model

SM\text{final} final system model

USER

MM\text{init} initial mental model

MM\text{use} mental model after system usage

MM\text{exh} mental model after exhaustive search

expressed requirements

system usage

browsing

Sel\text{use} selection after system usage

Sel\text{exh} selection after exhaustive search

R2
Verifying Requirements – Service Alternatives (R3)

SYSTEM:
- **SM_{init}**: initial system model
- **SM_{fin}**: final system model

USER:
- **MM_{init}**: initial mental model
- **MM_{fin}**: mental model after system usage
- **MM_{exh}**: mental model after exhaustive search

With expressed requirements, the process unfolds through:
- **Selection**
  - After system usage: $Sel_{syst}$
  - After exhaustive search: $Sel_{exh}$

System usage and selection activities are connected as follows:
- **Browsing**
- **Initial System Model**: $SM_{init}$ to **Initial Mental Model**: $MM_{init}$
- **System Usage**
- **Final System Model**: $SM_{fin}$ to **Final Mental Model**: $MM_{fin}$
- **Exhaustive Search**
- **Selection**
  - After system usage: $Sel_{syst}$
  - After exhaustive search: $Sel_{exh}$
Verifying Requirements – Consistent Selection (R4)

SYSTEM

- \( \text{SM}_{\text{init}} \) initial system model
- \( \text{SM}_{\text{fin}} \) final system model

USER

- \( \text{MM}_{\text{init}} \) initial mental model
- \( \text{MM}_{\text{use}} \) mental model after system usage
- \( \text{MM}_{\text{exh}} \) mental model after exhaustive search

EVALUATION

- \( \text{Sel}_{\text{init}} \) selection after system usage
- \( \text{Sel}_{\text{exh}} \) selection after exhaustive search
- \( \text{Ranking}_{\text{MMuse}} \) ranking based on MM

expressed requirements

- Browsing
- R4
Verifying Requirements – Accurate Model (R5)

SYSTEM
- SM_{ini} (initial system model)
- SM_{fin} (final system model)
- Ranking_{SM_{fin}} (ranking based on SM_{fin})

USER
- MM_{ini} (initial mental model)
- MM_{use} (mental model after system usage)
- MM_{exh} (mental model after exhaustive search)
- Sel_{use} (selection after system usage)
- Sel_{exh} (selection after exhaustive search)

EVALUATION
- Ranking_{MM_{use}} (ranking based on MM_{use})
- Ranking_{MM_{exh}} (ranking based on MM_{exh})
## Conversational Service Selection

<table>
<thead>
<tr>
<th>service</th>
<th>EFFECT</th>
<th>entity</th>
<th>from</th>
<th>to</th>
<th>carrier</th>
<th>departure</th>
<th>price</th>
</tr>
</thead>
<tbody>
<tr>
<td>service1</td>
<td>Owned</td>
<td>Transportation Ticket</td>
<td>Jena</td>
<td>Paris</td>
<td>Train</td>
<td>12.09.2012 10:38</td>
<td>156EUR</td>
</tr>
<tr>
<td>service2</td>
<td>Owned</td>
<td>Transportation Ticket</td>
<td>Jena</td>
<td>Paris</td>
<td>RailAndFly</td>
<td>12.09.2012 10:50</td>
<td>571EUR</td>
</tr>
<tr>
<td>service4</td>
<td>Owned</td>
<td>Transportation Ticket</td>
<td>Jena</td>
<td>Paris</td>
<td>RailAndFly</td>
<td>12.09.2012 11:50</td>
<td>571EUR</td>
</tr>
<tr>
<td>service5</td>
<td>Owned</td>
<td>Transportation Ticket</td>
<td>Jena</td>
<td>Paris</td>
<td>RailAndFly</td>
<td>12.09.2012 12:20</td>
<td>less</td>
</tr>
<tr>
<td>service6</td>
<td>Owned</td>
<td>Transportation Ticket</td>
<td>Jena</td>
<td>Paris</td>
<td>Train</td>
<td>12.09.2012 12:38</td>
<td>156EUR</td>
</tr>
</tbody>
</table>

91% of the offers fulfill your requirements.

**Available attributes ...**

- airline (49%)
- trip duration (100%)
- intermediate hops (82%)

**Available subtypes ...**

- carrier → Train (46%)
- → RailAndFly (49%)
- → Bus (5%)

**Tradeoff opportunities ...**

- trip duration ↑, # intermediate hops >1 (82%)
- trip duration ↑, departure ↓ (70%)
- airline!=Lufthansa (100%)
Case Study – Evaluation Setting

- study participants
  - 10 test users (6 males, 4 females, age 25 – 58 years)
  - 9 experience with online purchasing
  - 2 of them familiar with Web Services

- service offers
  - generated from structured information about computer items
  - from 8 categories (desktop PC, notebook, e-book reader, …)

- participants had to choose from 200 services of one category
Case Study – Assessing Service Requirements

- asked participants
  - to indicate service aspects that are important to them
  - specify their requirements on these aspects and
  - to weight the indicated service aspects against each other

- put no restrictions
  - on type of these requirements or
  - the way of specifying them
Selected Results – Requirements Construction (R2)

Respondents added/revised and abandoned requirements

SYSTEM

USER

expressed requirements

SYSTEM

USER

Respondents added/revised and abandoned requirements

initial mental model

after system usage

final mental model

after exhaustive search

initial system model

system usage

final system model

system usage

expressed requirements

browsing

system usage

Respondents added/revised and abandoned requirements

SM

MM

SM

MM

SM

MM

Sel

Sel

inil

init

exh

use

use

R2

R2

after system usage

after exhaustive search

after system usage

after exhaustive search

respondents added/revised and abandoned requirements
Selected Results – Requirements Construction (R2)

expressed requirements

SYSTEM

USER

no hidden requirements

SM

initial system model

SM

final system model

MM

initial mental model

MM

mental model after system usage

MM

mental model after exhaustive search

Sel

selection after system usage

Sel

selection after exhaustive search

R2

expressed requirements

system usage

browsing

R2

expressed requirements

no hidden requirements
Selected Results – Service Alternatives (R3)

[Schematic diagram showing the process of selecting service alternatives, with labels for system and user models, expressed requirements, and selection outcomes after system usage and exhaustive search.]
Selected Results – Consistent Selection (R4)

mean rank was 8.33 ±5.33 (out of 200)
Selected Results – Accurate Model (R5)

- Mean rank difference of the top ten offers was 30.98 ± 8.50
- 23.04 ± 7.88 for the 5 fairly consistent weightings
Summary

- identified requirements to service selection in B2C scenarios
- presented a user-centered evaluation methodology w.r.t. these requirements
- case study demonstrated feasibility and appropriateness of the suggested approach