REQUIREMENTS ENGINEERING FOR DATA WAREHOUSES

APRIL 2015

Presented by: Azadeh Nasiri
Problem:
GORE-based methods in the literature:

- Targeted various RE problems
  - Some of them focus on the understanding the organisational context of a DW
  - Some others focus on the information requirements of a DW ...
- Developed based on different principles
  - The i* framework
  - Toropos methodology
  - KAOS
  - URN including GRL and UCM

It is difficult to give a comprehensive GORE approach in the DW domain where a complete and consistent set of the DW requirements are taken into account.
Method engineering approach: a discipline to construct a method from existing ones

Current GORE-based Methods for DWs

Decision-making process

Phase 1

Phase 2

Phase 2

The proposed GORE-based Method
PROPOSED METHOD

Decision-making process

Phase 1: Searching for conditions that call for decision-making

Phase 2: Analysing possible courses of actions

Phase 3: Selecting a proper course of action from available options

GORE approach: using goals for requirements elicitation, requirements model and analysis, requirements negotiation and modification.

Goal: an objective that the system under consideration should achieve

Goal model: a graphical representation of the reduction of goals
  ✓ Elaborate how a goal is achieved
  ✓ Supports heuristic, qualitative or formal reasoning schemes during RE
PHASE 1: CONDITIONS FOR DECISION MAKING

- Translating a strategic goal to decisions that need to be made to achieve that goal

Objective

Modeling

Running example

- Providing a formal modeling foundation and proper representation of variables important for decision-making (developing a business conceptual model)
PHASE 2: ANALYSING POSSIBLE COURSES OF ACTIONS

- Evaluating the effect of decision alternatives on the strategic goal achievement
- Providing a proper analysis foundation to evaluate qualitatively or quantitatively the strategic goal satisfaction

Objective

Modeling

Running example

Legend:
- Situation
- Goal
- Refinement node: AND, OR, XOR
- Influence
- Decomposition
- Contribution relation
PHASE 2: analysing possible courses of actions

**Composite KPIs:** value of this KPI is obtained from other KPIs (component KPIs)

**Atomic KPIs:** value of this KPI is obtained from data sources

**Techniques to evaluate composite KPIs**
- Conversion factor technique
- Normalisation technique
- Qualitative technique

---

### Composite KPI
- Decrease AEs (G2)
  - Ws1 Cs1
- Increase R&D cost (S1)
  - cv = 10%
- Increment in cost (Xs1)

### Atomic KPI (Component KPIs)
- Increase patients' knowledge (G3)
  - Number of training hours (Xg4) cv = 20 hours
- Increase Patient follow-up (G4)
  - Number of communication channels (Xg5) Cv = 4 channels
- Increase drug quality (G5)
  - Number of AEs due to ineffective drug (Xg3) cv = 20 hours

---

**KPI**
PHASE 2: ANALYSING POSSIBLE COURSES OF ACTIONS

Techniques to evaluate composite KPIs

Conversion factor technique

$$X_{g2} = X_{g2}^e + W_{s1} \cdot C_{s1} \cdot X_{s1} + \sum_{i=3}^{5} W_{gi} \cdot C_{gi} \cdot X_{gi}$$

- $X_{gi}^c$ composite KPI
- $X_{gi}$ component KPI
- $X_{si}$ Situation-related component KPI
- $W_{gi}$ influence strength
- $W_{si}$ Situation-related influence
- $X_{gi}^e$ expected value of a composite KPI
- $C_{gi}$ conversion factor
- $C_{si}$ situation-related conversion factor

Composite KPI

- Decrease AEs (G2)
- Increase patients' knowledge (G3)
- Increase Patient follow-up (G4)
- Increase drug quality (G5)

Strategies:
- Strategy A: Increase in number of training hours (Xg4) $cv=20$ hours
- Strategy B: Increase in number of communication channels (Xg5) $Cv=4$ channels
- Strategy C: Increase in number of AEs due to ineffective drug (Xg3) $cv=20$ hours

Increase R&D cost (S1)

Increase in cost (Xs1) $cv=10\%$
PHASE 2: ANALYSING POSSIBLE COURSES OF ACTIONS

- Each KPI has a current value evaluated against:
  - Target, Threshold, Worst
- PL = (threshold - current)/(threshold - target) * 100
  - target < current value < threshold
- PL = (threshold - current)/(worst - threshold) * -100
  - threshold < current value < worst

\[ X_{g2} = C \cdot X_{s1} \cdot PL(X_{s1} + \max[PL(X_{g3}), PL(X_{g4}), PL(X_{g5})]) \]

- PL\(_{X_{gi}}\): performance level of composite KPI
- PL\(_{X_{si}}\): performance level of situation-related component KPI
- W\(_{gi}\): influence strength
- W\(_{si}\): situation-related influence strength

Techniques to evaluate composite KPIs

Normalisation technique
PHASE 2: analysing possible courses of actions

Mapping rules

<table>
<thead>
<tr>
<th>$cv \geq t$</th>
<th>$M \leq cv &lt; t$</th>
<th>$th \leq cv &lt; M$</th>
<th>$w \leq cv &lt; th$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$per^+=$“Full”</td>
<td>$per^+=$“Partial”</td>
<td>$per^+=$“None”</td>
<td>$per^+=$“None”</td>
</tr>
<tr>
<td>$per^-=$“None”</td>
<td>$per^-=$“Partial”</td>
<td>$per^-=$“Full”</td>
<td>$per^-=$“Full”</td>
</tr>
</tbody>
</table>

Techniques to evaluate composite KPIs

Qualitative technique

Propagation rules

$$
\begin{align*}
(X_i^a, X_j^a) & \rightarrow X_c^a \\
\min \left\{ \frac{per^+(X_i^a)}{P}, \frac{per^+(X_j^a)}{P} \right\} & \rightarrow \text{Positive performance} \\
\min \left\{ \frac{per^-(X_i^a)}{P}, \frac{per^-(X_j^a)}{P} \right\} & \rightarrow \text{Negative performance} \\
\text{Ranged from (Full, Partial, None)} \\
& \quad \text{F} > \text{P} > \text{N}
\end{align*}
$$

Influence strengths:
(+D), (-D), (++D), (- -D)
(+S), (-S), (++S), (- -S)

Composite KPI
$X_c^a$

Component KPI
$X_i^a$

Influence strength-satisfied

Influence strength-denied

Legend:
- Situation
- Goal
- Refinement node: AND, OR, XOR
- Target
- KPI
- Influence
- Refinement

Target value
Current value
Threshold value
Worst value

Number of training hours
Number of communication channels
Number of AEs due to ineffective drug

Increase R&D cost (S1)
Increase patients' knowledge (G3)
Increase Patient follow-up (G4)
Increase drug quality (G5)
PHASE 3: SELECTING A PROPER COURSE OF ACTION

Objective

◆ What data and in which form is of particular interest for decision makers to store in DWs

Modeling

◆ Adopting goal models to eventually represent the information in the Multidimensional (MD) schema with elements of facts (the center of analysis) and dimensions (the context of analysis)
CONCLUSION

Advantages:

- Taking advantage of the contribution of existing works in the RE for BI systems
- Giving a big picture of what a GORE approach needs to support in the RE for DWs
- Involving the decision-making process in the early phase of the system development
- Covering all phases of the decision-making process
- Extending the method with the dynamic part of the DW, where the requirements of operations on the DW are captured

- Extending goal models with UML to capture the interaction of users with a DW
THANK YOU FOR YOUR ATTENTION