PANEL - CONTENT / PATTERNS 2013

Advances on Information Mining

Samuel Kosolapov, Hans-Werner Sehring, Alfred Zimmermann
Introduction

1. What are the Advances in Information Mining we have to discuss today?

1. Goal of our discussion:

   We want to discuss about our experiences and our positions, coming from both academia and practice
Panelists

Self-introduction of panelists with short information about each person, their work and expectations to our discussion

- Samuel Kosolapov – Israel
- Hans-Wener Sehring – Germany
- Alfred Zimmermann – Germany
Definition of Information Mining

1. **Origin**: Data Mining or Knowledge Discovery from Databases
2. **Orientation**: towards high structured data
3. **Analysis of heterogeneous information sources**
4. **Often** combined with **AI technologies**, like fuzzy logic
5. **Classical methods** from **statistics, decision trees, neural networks**
6. **Challenges**: mining of texts, models, patterns, image & sound data, medical data, spacial data, other temporal data
7. **Information Mining**: Extension of Data Mining to identify understandable patterns in heterogeneous information sources
Tasks of Knowledge Discovery

1. **Deviation Filtering** (Anomaly Detection) of unusual instances
2. **Associations**: Dependency modeling based on monitored associations and association rules learning
3. **Clustering**: discovering type and classes of similar instances
4. **Classification**: mapping of instances to clusters of predefined classes
5. **Regression**: find a function which models the data with least error
6. **Summarization**: compact representation for visualizations or reports
7. **Pattern Mining**: Find sets of instances that occur frequently together
Panel: Advances on Information Mining

Positions from Panelists

1. Samuel Kosolapov:
   ### YourContributionTitle

2. Hans-Wener Sehring:
   ### YourContributionTitle

3. Alfred Zimmermann:
   Extracting and Integrating Metamodel-Information from Capability Maps of IT Enterprise Architectures
Synthesis of Opinions

1. Questions from the audience and answers from panelists
2. What are our main conclusions and hypotheses for future work in Information Mining?
3. What are the big challenges we have to face?
ESAMI – Enterprise Services Architecture Metamodel Integration

Prof. Dr. Alfred Zimmermann
ESARC defines an original holistic classification scheme for cyclic diagnostics and optimization of eight types (viewpoints) of service-oriented enterprise software architectures.

ESARC substantiates the TOGAF standard with other models to provide a useful mapping foundation of a reference architecture, defining main architecture artifacts and their relationships.
ESAMI – Enterprise Services Architecture Metamodel Integration
Integration Process

1. Analyze each Base Model with Concept Maps
2. Extract the Base Capability Model from each model:
   Capability, Model, Element, [Example]
3. Initialize the Reference Architecture Model from Base Capability Models:
   Capability, Model, Element
4. Analyze Correlations (Concept Mappings) between Base Models and Reference Architecture Model, and optionally conclude transitive correlations (Min)
5. Determine Integration Options for the resulting Capability Integration Model
6. Develop the Synthesis Metamodel from Base Metamodels
7. Consolidate the Reference Architecture Model according to the Joined Metamodel, and finally readjust Correlations and Integration Options
8. Develop Capability Map and Ontology of the Reference Architecture Model
9. Develop Correspondence Rules between Model Elements
10. Develop Patterns for Architecture Diagnostics and Optimization
Integration Scenario of Business Architectures from ArchiMate® 2.0 and TOGAF 9.1

<table>
<thead>
<tr>
<th>Capability</th>
<th>Model</th>
<th>Element</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>BusinessActor</td>
<td>BusinessActor</td>
<td>BusinessActor =</td>
<td>BusinessActor = Language</td>
</tr>
<tr>
<td>(Viewpoint)</td>
<td>BusinessRole</td>
<td>BusinessRole =</td>
<td>Insurance Department</td>
</tr>
<tr>
<td></td>
<td>BusinessRole</td>
<td>BusinessRole =</td>
<td>Travel</td>
</tr>
<tr>
<td></td>
<td>Process</td>
<td>Process =</td>
<td>Seller</td>
</tr>
<tr>
<td></td>
<td>Service</td>
<td>Service =</td>
<td>travel insurance</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Service =</td>
<td>Offering travel insurance</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Service =</td>
<td>offering travel insurance</td>
</tr>
</tbody>
</table>

Open Group Standard
ArchiMate® 2.0 Specification.
The Open Group 2009-2012,
ISBN 1-937218-00-3

Open Group Standard
TOGAF 9.1.
The Open Group 2011,
ISBN 978-9087536794
## ArchiMate 2.0
### Base Capability Map for BusinessActivator: BusinessActor

<table>
<thead>
<tr>
<th>Capability</th>
<th>Model</th>
<th>Element</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>BusinessActivator</td>
<td>BusinessActor</td>
<td>BusinessActor</td>
<td>BusinessActor = Luggage Insurance Department</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BusinessRole</td>
<td>BusinessRole = Travel Insurance Seller</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Process</td>
<td>Process = take out travel insurance</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Service</td>
<td>Service = Offering travel insurance</td>
</tr>
</tbody>
</table>
34.4 Content Metamodel Extensions

As discussed earlier, the TOGAF content metamodel supports a number of extension modules that allow for more in-depth consideration for particular architecture concerns. Figure 34-9 shows the core content metamodel and predefined extension modules.
# TOGAF 9.1
## Base Capability Map for Organization: Actor

<table>
<thead>
<tr>
<th>Capability</th>
<th>Model</th>
<th>Element</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Organization</td>
<td>Actor</td>
<td>Actor</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Role</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>OrganizationUnit</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Function</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Process</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>BusinessService</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Event</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Location</td>
<td></td>
</tr>
</tbody>
</table>
## ESAMI – Enterprise Services Architecture Metamodel Integration
### Consolidated EAM Reference Architecture: Analysis and Integration

<table>
<thead>
<tr>
<th>Reference</th>
<th>EAM Reference Architecture</th>
<th>Correlation Index</th>
<th>Integration Options</th>
<th>Documents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Origin</td>
<td>Capability</td>
<td>Model</td>
<td>Element</td>
<td>ArchiMate</td>
</tr>
<tr>
<td></td>
<td>Actor</td>
<td>2</td>
<td>2</td>
<td>p</td>
</tr>
<tr>
<td></td>
<td>Role</td>
<td>3</td>
<td>2</td>
<td>m</td>
</tr>
<tr>
<td></td>
<td>Collaboration</td>
<td>3</td>
<td>0</td>
<td>m</td>
</tr>
<tr>
<td></td>
<td>Organizational Unit</td>
<td>1</td>
<td>3</td>
<td>p</td>
</tr>
<tr>
<td></td>
<td>Business Function</td>
<td>3</td>
<td>3</td>
<td>m</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Organization Location</th>
<th>Correlation</th>
<th>Integration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Location</td>
<td>0</td>
<td>reject</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>partially</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>mandatory</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>(leading model)</td>
</tr>
</tbody>
</table>

**Archimate Specific. and TOGAF Standard**

**Reference Documents**
- File: 25-32
- Pages: 87-88
- Authors: 20130505-ESAMI
ESAMI – Enterprise Services Architecture Metamodel Integration
SOA Ontology Typed Metamodel of Business Reference Architecture
What are the Advances in Information Mining we have to discuss today?

Goal of our discussion: We want to discuss about or experiences and our positions coming from both academia and practice.

Let me first introduce our panelists and ask them about some insight information about your person, your work and your expectations to our discussion.

Samuel Kosolopov from Israel

Hans-Werner Sehring from Germany

Alfred: Extracting and Integrating Metamodel-Information from Capability maps of IT Enterprise Architecture

What are the big challenges we have to expect?

Questions from the audience and answers from panelists

What are our main Conclusions and Hypotheses for future work in Information Mining?

Information Mining

Origin: Data Mining or Knowledge Discovery from Databases

Orientation: towards high structured data

Analysis of heterogeneous information sources

often combined with AI technologies, like fuzzy logic

classical methods from statistics, decision trees, mural networks

challenges: mining of texts, models, patterns, image & sound data, medical data, spacial data, other temporal data

Short Definition of IM: Extension of data mining to identify understandable patterns in heterogeneous information sources

Tasks of Knowledge Discovery

Deviation Filtering (anomaly detection) of unusual instances

Associations: Dependency modeling based on monitored associations and association rules learning

Clustering: discovering type and classes of similar instances

Classification: mapping of instances to clusters of predefined classes

Regression: find a function which models the data with least error

Summarization: compact representation for visualizations or reports

Pattern Mining: Find sets of instances that occur frequently together

Let me first introduce our panelists and ask them about some insight information about your person, your work and your expectations to our discussion.

I'm the moderator of this panel

my name az, Professor for Software Architecture & EAM at RTU, Germany.

I'm also responsible for my research group at the ARL – Architecture Reference Lab – of the SOA Innovation Lab, which is the big research & innovation consortium in EAM of major industrial companies and public organizations in Germany and Switzerland.

I suggest to start first with a short self presentation of each panelist to understand each position coming from the research and practical background.

Now I would ask, what are you doing and what is your main position with examples from your work?

Samuel

Hans-Werner

Alfred: Extracting and Integrating Metamodel-Information from Capability maps of IT Enterprise Architecture
PANEL - CONTENT / PATTERNS 2013

Advances on Information Mining

Samuel Kosolapov, Hans-Werner Sehring, Bijan Raahemi, Alfred Zimmermann
Introduction

1. What are the Advances in Information Mining we have to discuss today?

2. Goal of our discussion:

   We want to discuss about our experiences and our positions, coming from both academia and practice
Panelists

Self-introduction of panelists with short information about each person, their work and expectations to our discussion

- Samuel Kosolapov
- Hans-Wener Sehring
- Bijan Raahemi
- Alfred Zimmermann
Definition of Information Mining

1. **Origin**: Data Mining or Knowledge Discovery from Databases

2. **Orientation**: towards **high structured data**

3. **Analysis of heterogeneous information sources**

4. **Classical methods** from Statistics, Decision Trees, Neural Networks

5. **Often** combined with **AI Technologies**, like Fuzzy Logic, Concept Learning

6. **Challenges**: mining of texts, models, patterns, image & sound data, medical data, spacial data, other temporal data

7. **Information Mining**: Extension of Data Mining to identify **understandable patterns** in heterogeneous information sources
Tasks of Knowledge Discovery

1. **Deviation Filtering** (Anomaly Detection) of unusual instances
2. **Associations**: Dependency modeling based on monitored associations and association rules learning
3. **Clustering**: discovering type and classes of similar instances
4. **Classification**: mapping of instances to clusters of predefined classes
5. **Regression**: find a function which models data with least error
6. **Summarization**: compact representation for visualizations or reports
7. **Pattern Mining**: Find sets of instances that occur frequently together
Positions from Panelists

1. Samuel Kosolapov
   **Addition of Rich Data from Smartphone Sensors to Image Metadata**

2. Hans-Wener Sehring
   **Adapting Content Based on Users’ Behavior**

3. Bijan Raahemi
   **Machine Learning with Applications in Intrusion Detection**

4. Alfred Zimmermann
   **Extracting and Integrating Metamodel-Information from Capability Maps of IT Enterprise Architectures**
Synthesis of Opinions and Discussion

1. **Questions** from the **audience** and **answers** from **panelists**

2. **What** are our **main conclusions** and **hypotheses** for **future work** in Information Mining?

3. **What** are the **big challenges** we have to face?
Addition of rich environment, location, motion, and orientation data from smartphone sensors to Image Metadata will provide better scene analysis.

Are current Camera and Image Analyzing applications ready for the change?

Samuel Kosolapov
CONTENT/PATTERNS Panel.
Google Street View in Compass Mode: Compass that displays your location and direction

I want Sensors Info to be included into the Image Metadata but in a standardized way. Yes, everyone can add a TAG, but in order to provide metadata search some standardization is required.
Galaxy Nexus. List of Sensors

- **Camera** (Frons & Back) + Optical proximity sensor: GP2A
- **Microphone**
- **TouchScreen sensor**: *Melfas MMSxxx touchscreen*
- **Triaxial acceleration sensor** *BOSCH BMA250*
  - Measurement of accelerations in 3 perpendicular axes
  - Used to sense tilt, motion, shock and vibration. (range +2g... +16g; 10 bit !)
- **Triple Axis MEMS Gyroscope**: *InvenSense MPU3050*
  - 3-axis gyroscope with an embedded Digital Motion Processor™
  - Digital-output X-, Y-, and Z-Axis angular rate sensors (gyros)
    - range of ±250 to ±2000°/sec
  - Used to sense: gesture recognition, panning, zooming, scrolling, zero-motion detection, tap detection, and shake detection
  - Digital-output **temperature sensor**
- **Geomagnetic sensor**: tri-axial *Yamaha YAS530*
  - Resolution: 0.15uT (X,Y), 0.3uT (Z)
  - Used to sense: Absolute Direction (compass)
- **Barometric pressure sensor**: *BOSCH BMP180*
  - (300...1100 hPa, first of its kind in a smartphone)
  - Used to sense: height over the sea level
Why would I want more sensors data inside the image?

- We have a huge amount of photos and videos in the cloud. But search images by content is still very slow and is not always feasible.
- Metadata can provide additional info which can be used to narrow the search.
- Example: Leaning Tower of Pisa
  “Engineers” announced that the Tower had been stabilized such that it had stopped moving for the first time in its history. They stated it would be stable for at least 200 years.

I want to analyze set of images taken from the same camera position, direction and orientation in order to check this claim.

- Actually, sensor information may help to reconstruct 3D scene from a plurality of 2D Images.
Adapting Content Based on Users’ Behavior. 
Panel on Advances on Information Mining. 
Computation World 2013, Valencia, Spain.

Dr. Hans-Werner Sehring, T-Systems Multimedia Solutions GmbH, Germany.
Scope of Information Mining Used in this Proposition.

- Information/content: purposeful data.
- Mining: Detect implicit relationships in data (similar: find structure in data).
- My field: Content Management.
  - Manually managed content (data).
  - Structure and relationships are well-known/prescribed.
- Observation: Content is perceived in …
  - a context ("Context is King") …
  - by a user/consumer/use/situation/…
Learning About Users.

- Information = purposeful data, in a context; here: content targeted at a person.
- Information mining here: learning about users and their use of content.
  - Need to describe persons.
  - No explicit data available.
    - Not willing to give information: consumer, …
    - Not possible to give info: diverse interest, subjective (non-explicable) views, …
  - Known, e.g., from personalization based on tracking/profiling.
- Known, e.g., from personalization based on tracking/profiling.
  - Tracking actions.
  - Allowing registered users to give information about themselves.
  - Observing registered users, eventually including anonymous tracking history.
Information Mining on the Web (I).

- In field of content management: Web Content Management now named
  - Web Experience Management (WCM)
  - Customer Experience Management (CEM, CXM)
- Currently of particular interest: marketing automation.
Information Mining on the Web (II).

- Topics, e.g.:
  - Web tracking (what did the customer do?).
  - Learn about customer journeys (how did the customer come to us?).
  - Segmentation (cluster/classify users).

- With the goals to …
  - Move focus from Point of Sale (PoS) to Point of Contact.
  - Iteratively measure success and improve a site’s appearance.
    - Personalization (for groups of users).
    - A/B testing.
  - Compute recommendations (based on segments).
  - Predictive marketing (based on user’s history).

- Overall goal of course: increase conversion rate by attracting the customers in the “right” context.
Information Mining on the Web (III).

- Challenges:
  - Getting to know about users.
    - How to identify anonymous users?
    - With different/changing devices?
    - Across different services provided by different service providers?
    - Within legal constraints?
  - How to distinguish different usage scenarios? E.g.,
    - Buy a new washing machine,
    - Having bought a new washing machine,
    - Looking for a washing machine to recommend to the mother in law.
  - How to incentivize users to give information about themselves?
  - …?

Drawing conclusions from knowledge about users.