Eiropas Sociālā fonda projekts
“Datorzinātnes pielietojumi un tās saiknes ar kvantu fiziku”
Nr.2009/0216/1DP/1.1.1.2.0/09/APIA/VIAA/044

Research Directions of OLAP Personalization

Natalija Kozmina & Laila Niedrite

ISD 2010

Outline

- Why OLAP Personalization is important?
  - Definition
  - Motivating Example
- Classification of Existing Approaches
- A Closer Look at Main Directions of OLAP Personalization
  - Preference Constructors
  - Visual OLAP
  - Dynamic Personalization
  - Recommendations with User Session Analysis
  - Recommendations with User Profile Analysis
- Comparative Analysis of OLAP Personalization Approaches
- Conclusions & Future Work

Personalization

- *Personalization* is a process of providing users with selected information on their specific needs*.

* - BNET Business Dictionary.

http://dictionary.bnet.com/definition/personalization.html

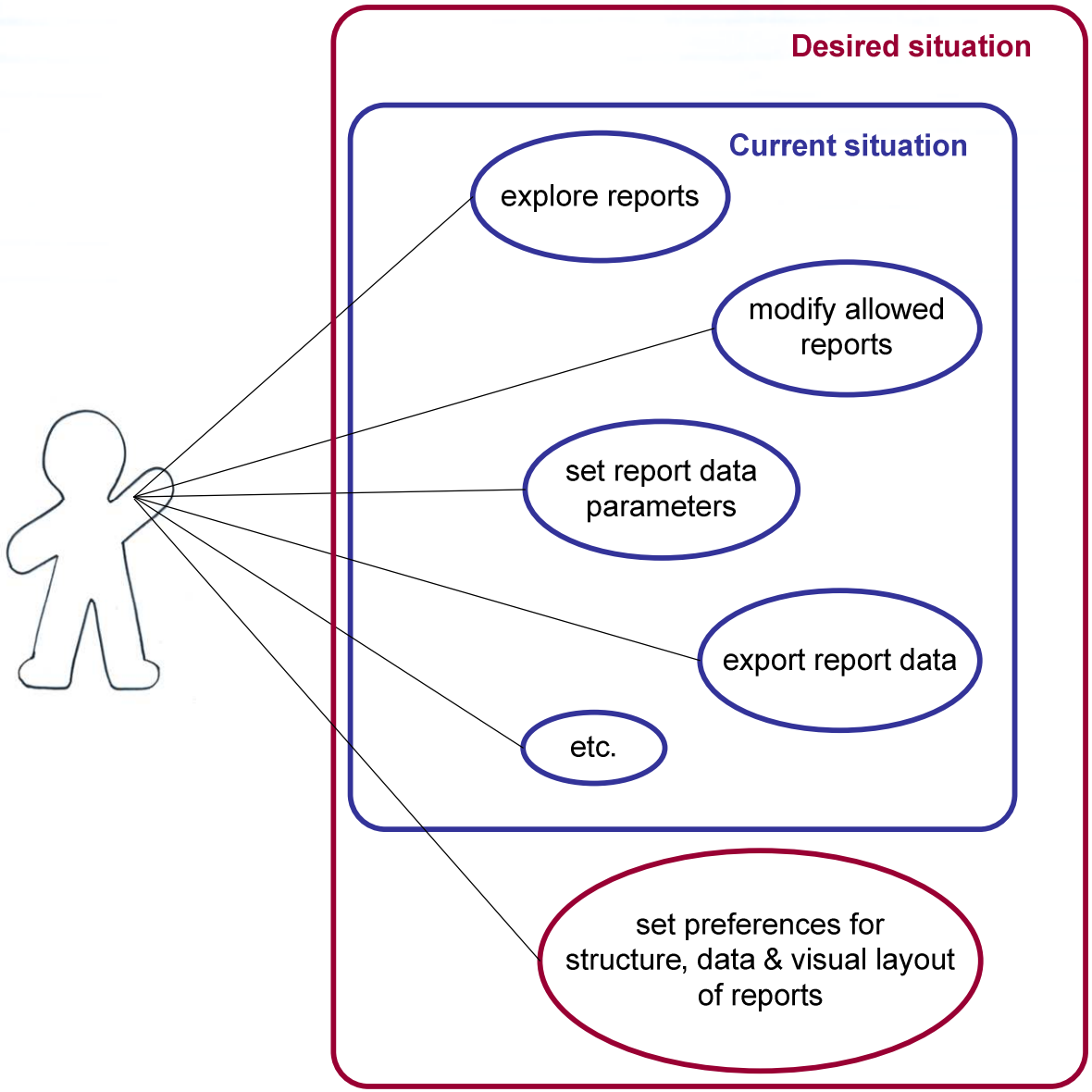
Why is OLAP personalization important?

- Typical problems in DW field:
  - Large volumes of data,
  - Burdening data exploration,
  - While exploring previously unknown data, the OLAP query result may highly differ from expectations.
- Possible solution – introducing personalization in the field of data warehousing.

Motivating example (1/2)

- DW report management tool
- Different groups of users (e.g., students, professors, workers of the University, etc.).
- Each group or particular user has...
  - different needs for reports,
  - interest for different contents of the report,
  - different reports' layout preferences.

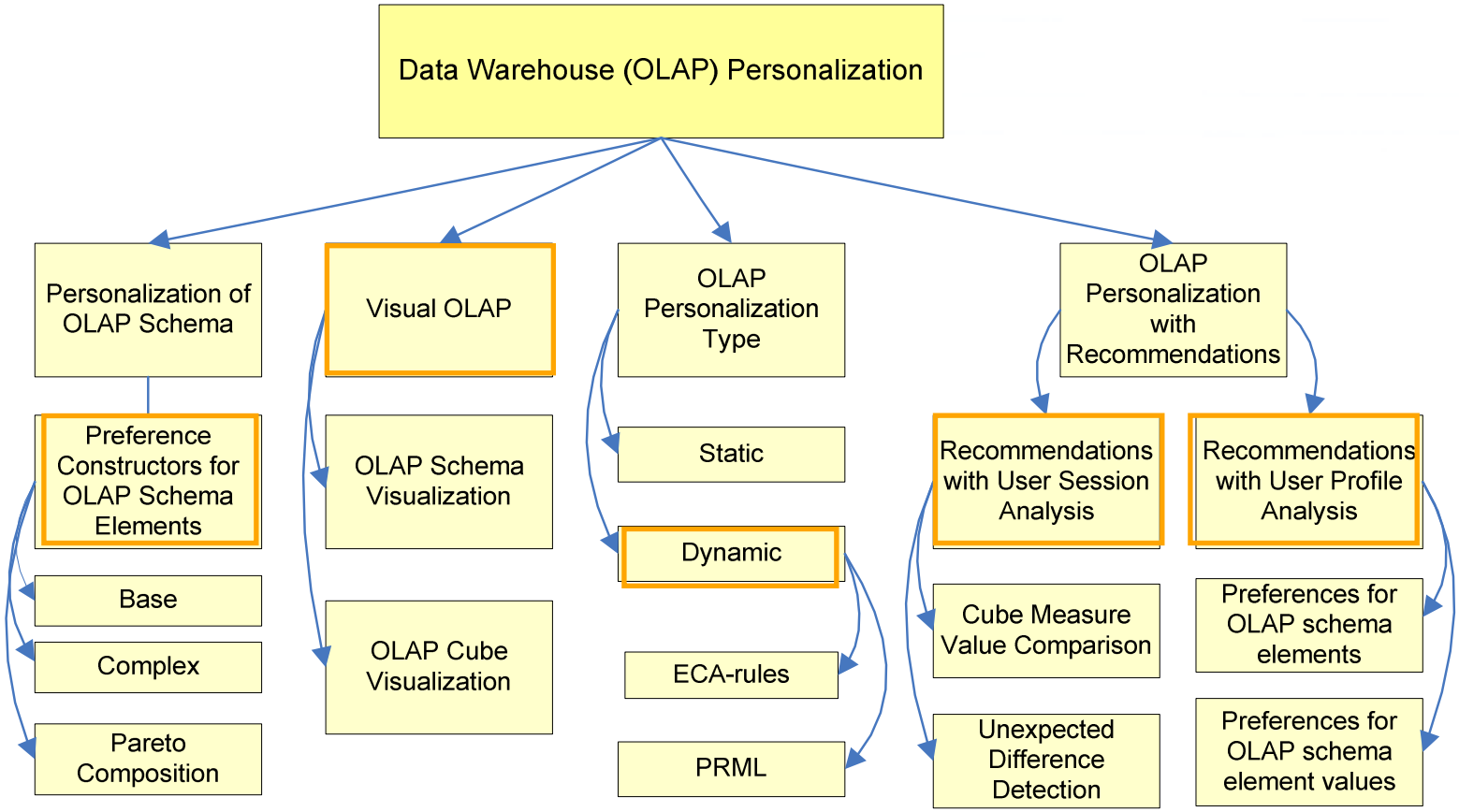
Motivating example (2/2)



How to provide user with the data he/she was looking for?

Introduce personalization into report management tool

Classification of Existing Directions of OLAP Personalization



Goal of the Research

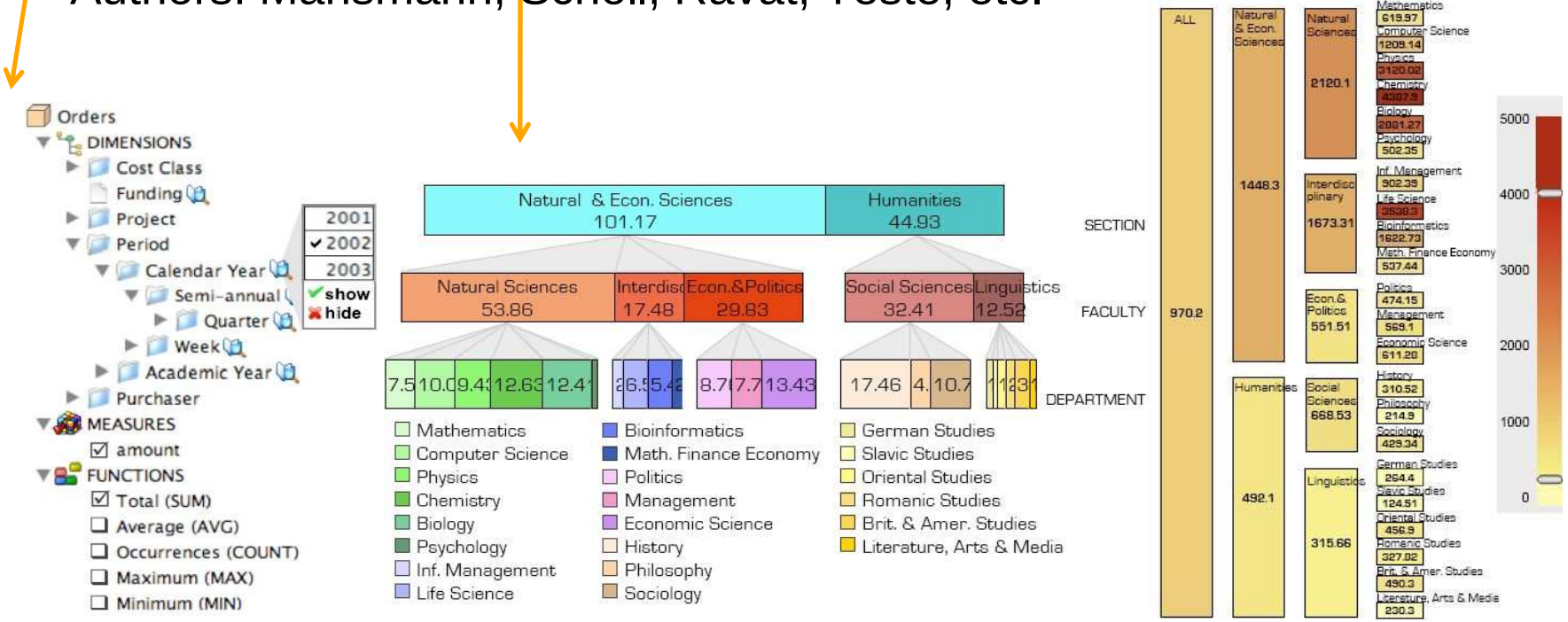
- Classify the ideas that have been already proposed in the field of OLAP personalization in order to find the most suitable personalization approach to be adjusted, developed and implemented into report management tool.

Preference Constructors (PC)

- *Base* preference constructors are applied to attribute, measure, hierarchy level.
- *Complex* preferences consist of combination of base preferences.
  - < expr >::= < baseConstr > | < expr > ⊗ < baseConstr > <baseConstr> ::=POS|NEG|BETWEEN|LOWEST| HIGHEST|CONTAIN|NEAR|COARSEST|FINEST
- Examples:
  - POS(Month, 'Jan-10') ,
  - BETWEEN(AvgIncome, 700, 1500),
  - etc.
- Authors: Golfarelli, Rizzi, etc.

Visual OLAP (VO)

- Users are able to...
  - navigate in dimensional hierarchies using a schema-based data browser,
  - select measures and aggregation functions,
  - set dimension attribute values as filters.
- Decomposition tree
- Authors: Mansmann, Scholl, Ravat, Teste, etc.

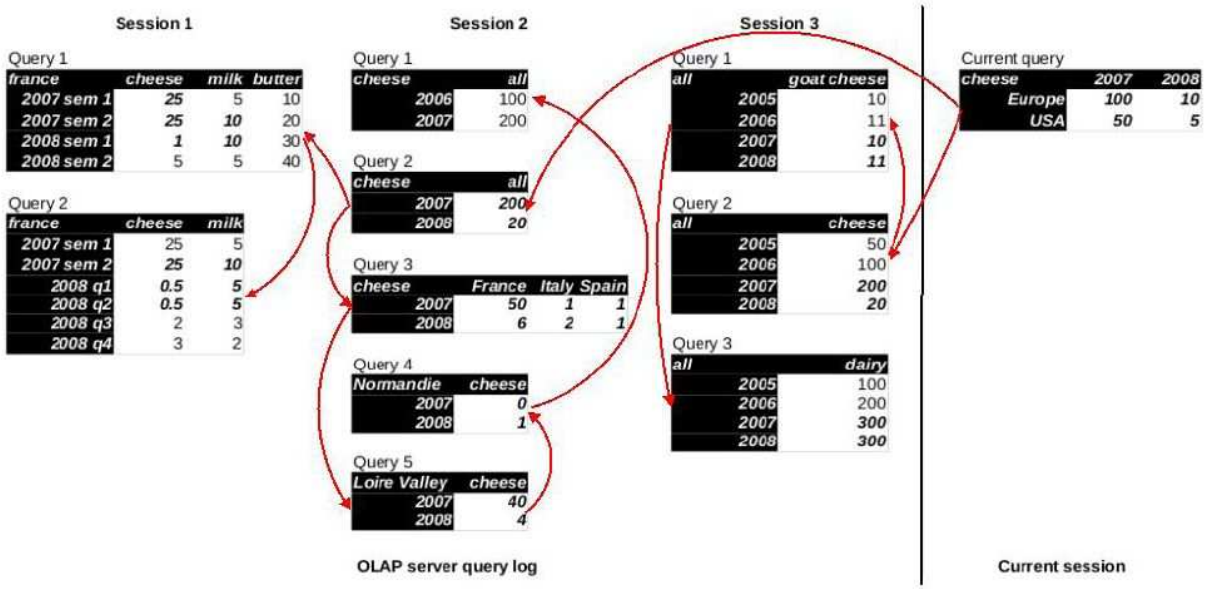


Dynamic Personalization (DP)

- Dynamic OLAP personalization – an adapted OLAP cube is created during the execution time according to the needs and performed actions of the user.
- ECA-rules structure:
  - When *event* do
    If *condition* then *action* endIf
    endWhen
- Two kinds of actions to be used in personalization rules:
  - hide,
  - set.
- Authors: Garrigós, Pardillo, Mazón, Trujillo, Thalhammer, Schrefl, Mohania, Gómez, etc.

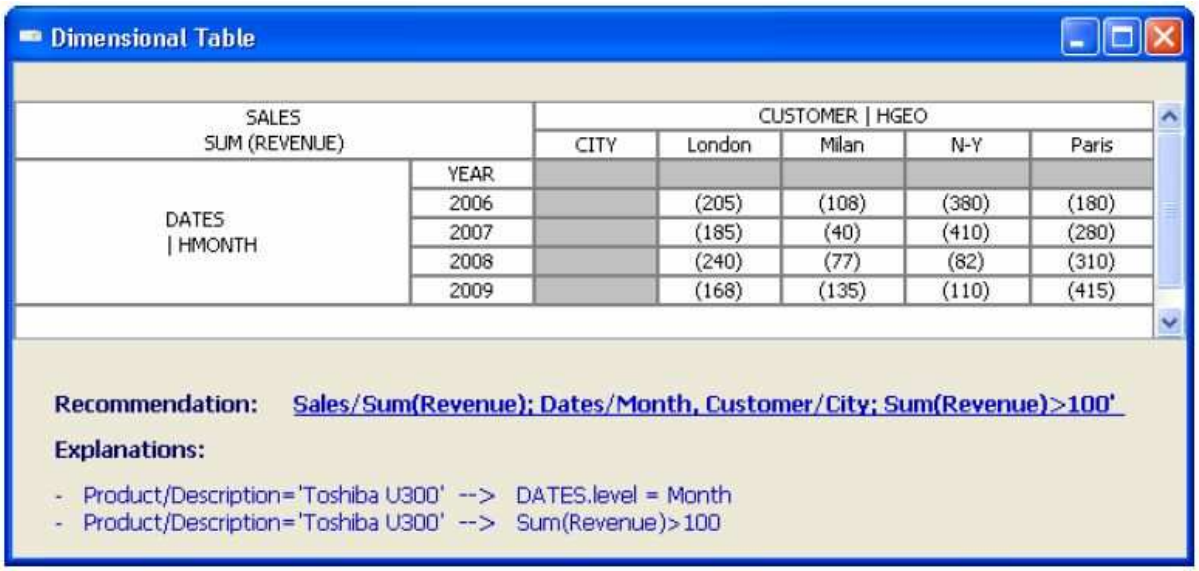
Recommendations with User Session Analysis (RUSA)

- Users' previous sessions' data patterns
  - OLAP server query log
- Cube measure values are being compared
- Unexpected difference in data detected
- Queries from all sessions, where user found the same unexpected data as in current session
- Authors: Giacometti, Marcel, Negre, Soulet, etc.



Recommendations with User Profile Analysis (RUPA)

- Analysis context consists of two disjoint sets of elements: OLAP schema elements – cubes, measures, dimensions, attributes, etc. and its values.
- Preferences are stated in the user profile and ranked with relevance score (a real number, [0; 1])
- Generation of recommendations
- Authors: Jerbi, Ravat, Teste, Zurfluh, Bellatreche, Giacometti, etc.



Comparative Analysis of OLAP Personalization Approaches (1/6)

- Applicability of different personalization types to OLAP objects
  - "A" – applicable
  - "D" – derivable
  - "-" – there is no information

Pers. Type / Pers. Object	Dimension	Dimension attribute	Hierarchy	Hierarchy level	Cube	Cube measure	Aggregation function	Select	Drilldown	Rollup	Rotate
PC	-	*A*	-	*A*	-	*A*	-	*D*	*D*	*A*	-
DP	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A, D*	*A*	*A*	*A, -*
VO	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*
RUSA	*A*	*A*	*A*	*A*	*A*	*A*	*D*	*A*	*A*	*A*	-
RUPA	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*A*	*D*	*D*	-

Comparative Analysis of OLAP Personalization Approaches (2/6)

- Personalization of OLAP schema elements is mostly present in all proposed OLAP personalization types, except for preference constructors (PC)
  - The way of expressing user preferences for dimensions, hierarchies, cubes as well as aggregate functions, is not described
- Preferences on OLAP operations such as Select, Drilldown and Rollup are not always expressed explicitly.
- There is a lack of information about personalization options, considering Rotate OLAP operation.

Comparative Analysis of OLAP Personalization Approaches (3/6)

- OLAP Personalization types and applied constraints

Personalization Type / Method	Hard Constraints	Soft Constraints	Other
PC	-	+	-
DP	+	-	-
VO	+	-	-
RUSA	-	-	+
RUPA	-	+	-

Comparative Analysis of OLAP Personalization Approaches (4/6)

- *Preference Constructors* use soft constraints as there is a possibility to express user's likes and dislikes.
- In *Visual OLAP* users' navigation events such as clicking and dragging are translated to valid SQL-queries with WHERE-clause, which is a hard constraint.
- There are hard constraints in *DP* with Event-Condition-Action-rules.
- The idea *Recommendations with User Session Analysis* is to find unexpected difference in the data and generate further recommendations with the same unexpected data. For that purpose a special operator is used.
- In *Recommendations with User Profile Analysis* soft constraints appear in user profiles, because user may express the extent of liking or disliking as a relevance score that is associated with analysis element.

Comparative Analysis of OLAP Personalization Approaches (5/6)

- Preference obtaining and user information collection methods, used in different types of OLAP personalization

	Preference Obtaining Method						*User Information Collection Method*	
Personalization Type	**Q&A**	**MI**	**CB**	**UKB**	**C**	**D**	**Explicit**	**Implicit**
PC	+	-	-	+	-	-	+	-
DP	-	-	+	+	-	-	-	+
VO	-	-	+	-	-	-	+	-
RUSA	-	-	-	+	-	-	-	+
RUPA	-	-	+	-	-	-	+	-

Comparative Analysis of OLAP Personalization Approaches – Conclusions (6/6)

- User preferences are stated explicitly in three approaches (*PC, VO, RUPA*) and implicitly in two approaches (*DP, RUSA*).
- Three out of six preference obtaining methods (i.e. *questions & answers, content-based and utility & knowledge-based*) are applied in considered types of personalization and the remaining three methods (*mixed initiative, collaborative and demographic*) are not applied.
- What could be improved:
  - Involve collaborative method for generating recommendations of queries, based on similarity of *other* users' likes and dislikes.

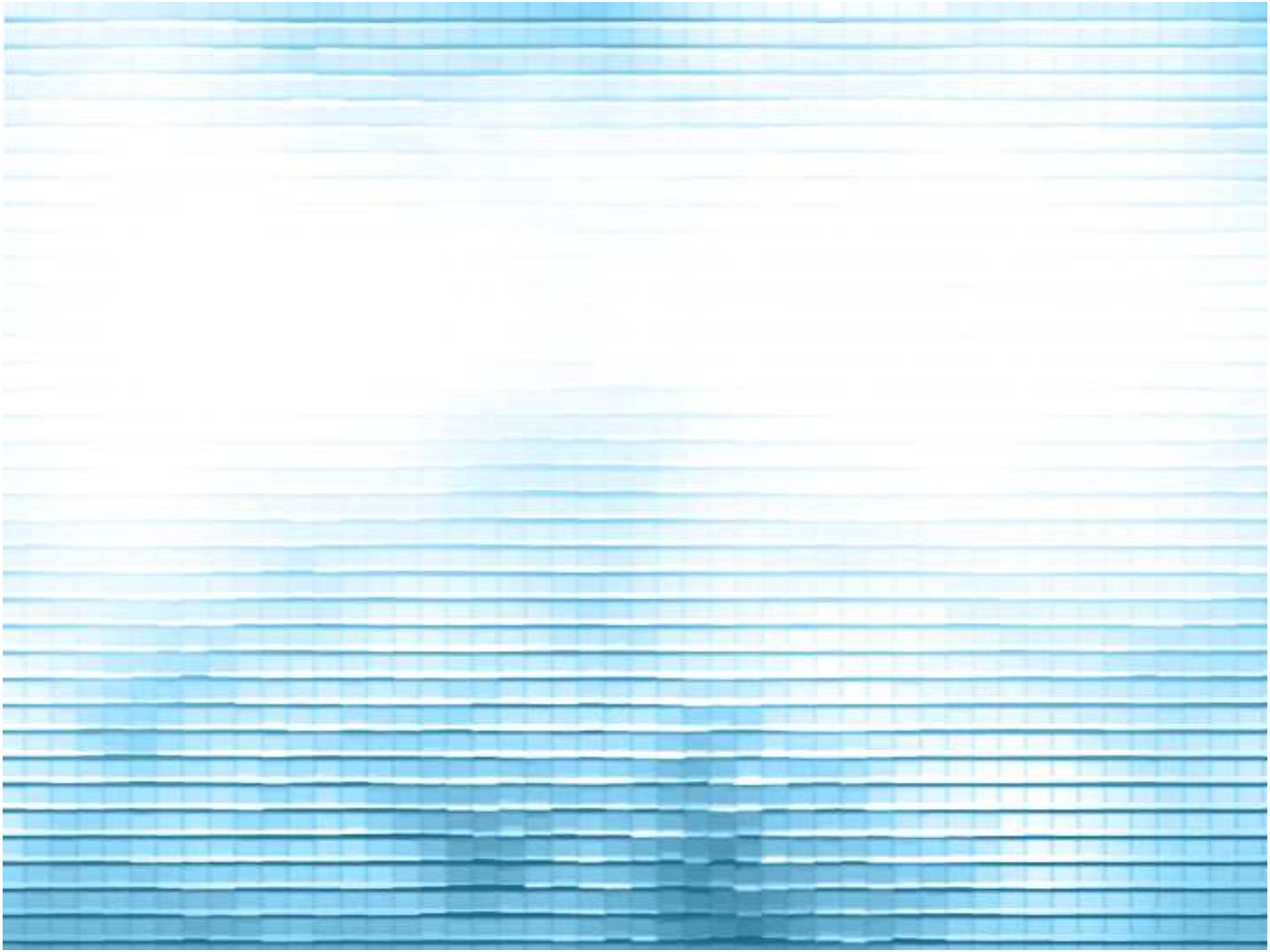
Conclusions & Future Work (1/2)

- Five approaches for introducing personalization in OLAP have been highlighted & comparative analysis proposed
- Taking into account the variety of different data warehouse report management tool users and their needs, as well as characteristics of considered personalization OLAP approaches, the Ideas of *RUPA* OLAP personalization approach are taken as a basis for introducing personlization into report management tool of the University of Latvia as
  - it allows to apply personalization to OLAP schema elements,
  - user is able to state his/her preferences in a profile explicitly,
  - soft constraints help avoiding empty result sets.

Conclusions & Future Work (2/2)

- We propose to expand *RUPA* approach by adding collaborative method for generation of recommendations to gain knowledge about other users' preferences.
- As a future work, a new method is being developed, which provides exhaustive description of interaction between user and data warehouse.
  - Mentioned method is a subject of a separate paper by N. Kozmina & L. Niedrite: "OLAP Personalization with User-describing Profiles".

Thank you!



OLAP
On-line analytical processing. OLAP is defined as providing fast access to shared multi-dimensional data. OLAP is a term used to generically refer to software and applications that provide users with the ability to store and access data in multi-dimensional cubes.

Cube
Also known as an OLAP cube. Data stored in a format that allows users to perform fast multi-dimensional analysis across different points of view.

Dimension
Structural attribute of a cube describing the data, for example, ‘Time’, ‘Product’ or ‘Geography’. A dimension acts as an index for identifying values within a multi-dimensional array.

Hierarchy
Dimension's members organized into parent-child relationships, for example ‘Quarter’ is a parent ‘ Month’ in the ‘Time’ dimension.

Star Schema
A set of relational tables comprising of central fact tables each surrounded by de-normalized dimensions.

User-describing profile diversity

Question	Description	Profile Type
What is the user expecting to get as a result?	User preferences data	*Preferential*
Who is the user?	Basic user data (personal data, session, activity, rights, etc.)	*User*
Where is the user located?	User physical location data & geolocation, according to user IP-address	*Spatial*
When does the user interact with the system?	Time characteristics of user activities	*Temporal*
How does the user & system interaction happen?	Characteristics of user device (i.e. PC, laptop, mobile phone, etc.), which is used for signing in as well as user software (e.g. web browser) characteristics	*Interaction*
Why the user is interested in this particular system?	User preferences are being gathered and analyzed. Recommendations are generated, according to user characteristics and preferences..	*Recommendational*

User-describing profile data sources