Extending the OLAP Technology to Handle Non-Conventional and Complex Data

Svetlana Mansmann

Universität Konstanz

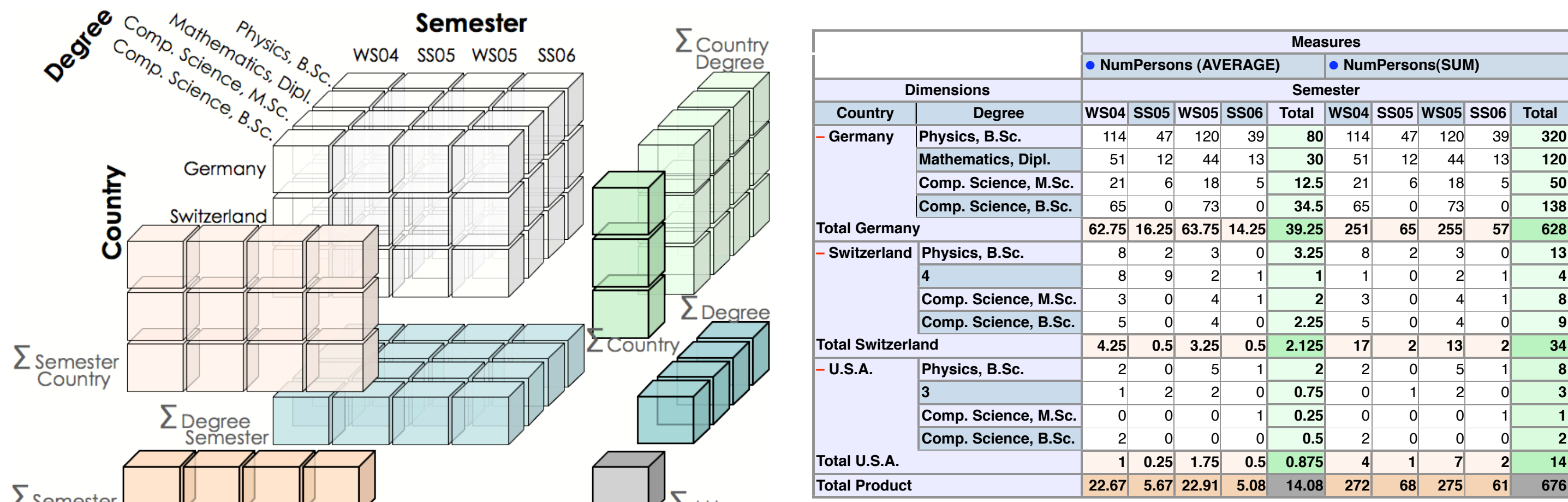Explorative Analysis and Visualization of Large Information Spaces

1 Background

In (*OLAP* (*On-line Analytical Processing*)), huge data volumes are analyzed as numeric facts, distributed along descriptive dimensions.

A sample 3-dimensional cube (fragment) with student enrollment numbers (left) and a pivot table view of the data (right)

2 Research Questions

The multidimensional data model and its summarizability constraint are too restrictive for handling complex and non-conventional usage scenarios. Besides, state-of-the-art frontend analysis tools have limited support for comprehensive analysis tasks.

Thesis contributions:
◈ An extended conceptual data model
◈ A conceptual-to-relational mapping of the extended model
◈ Comprehensive graphical modeling notation x-DFM
◈ Multidimensional analysis for Business Process Intelligence
◈ Visual OLAP framework
◈ Hierarchical visualization techniques for OLAP

3 The Roadmap of Dimensional Modeling
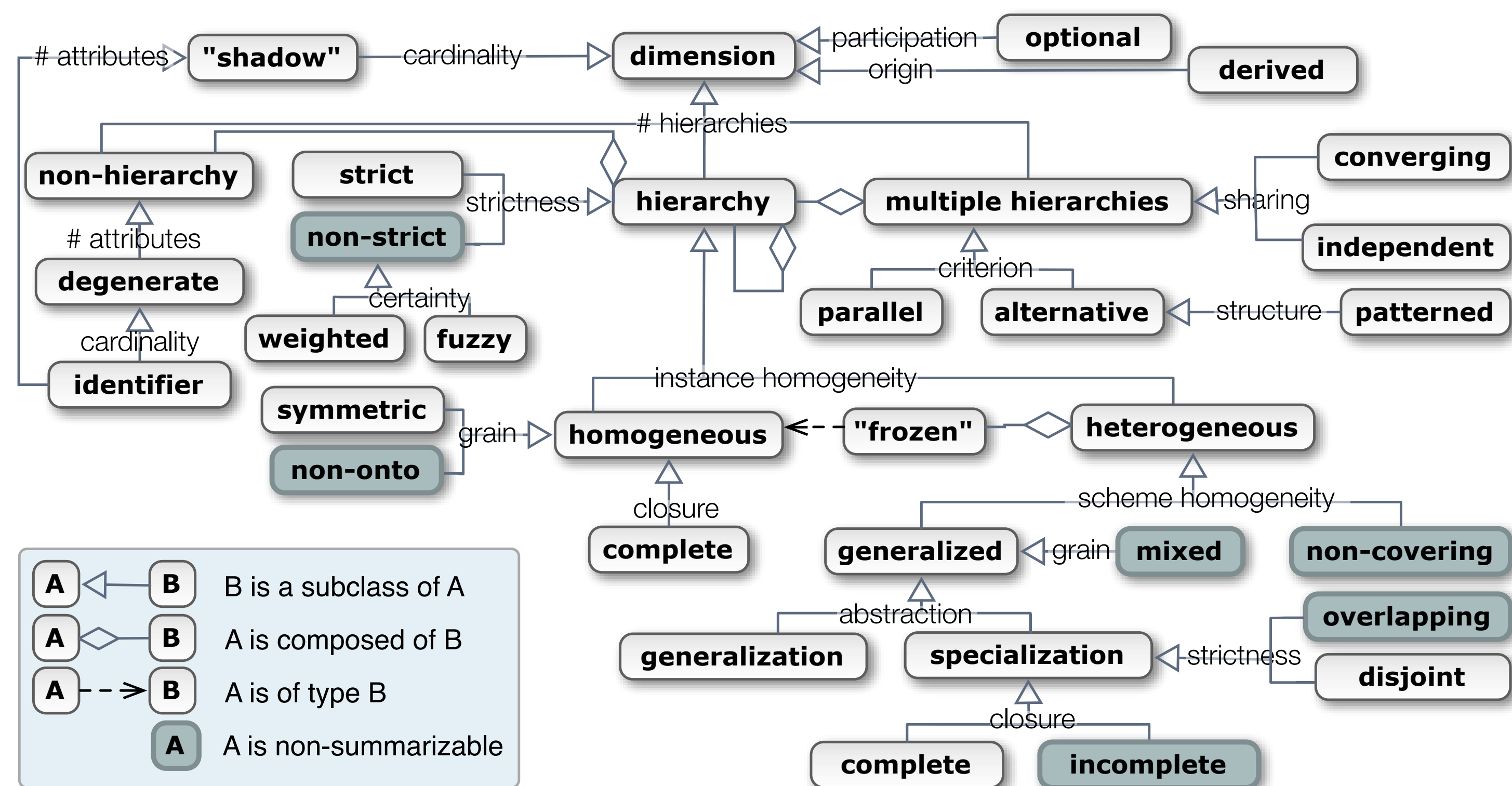
Sample Fact-dimensional scheme modeled in x-DFM:
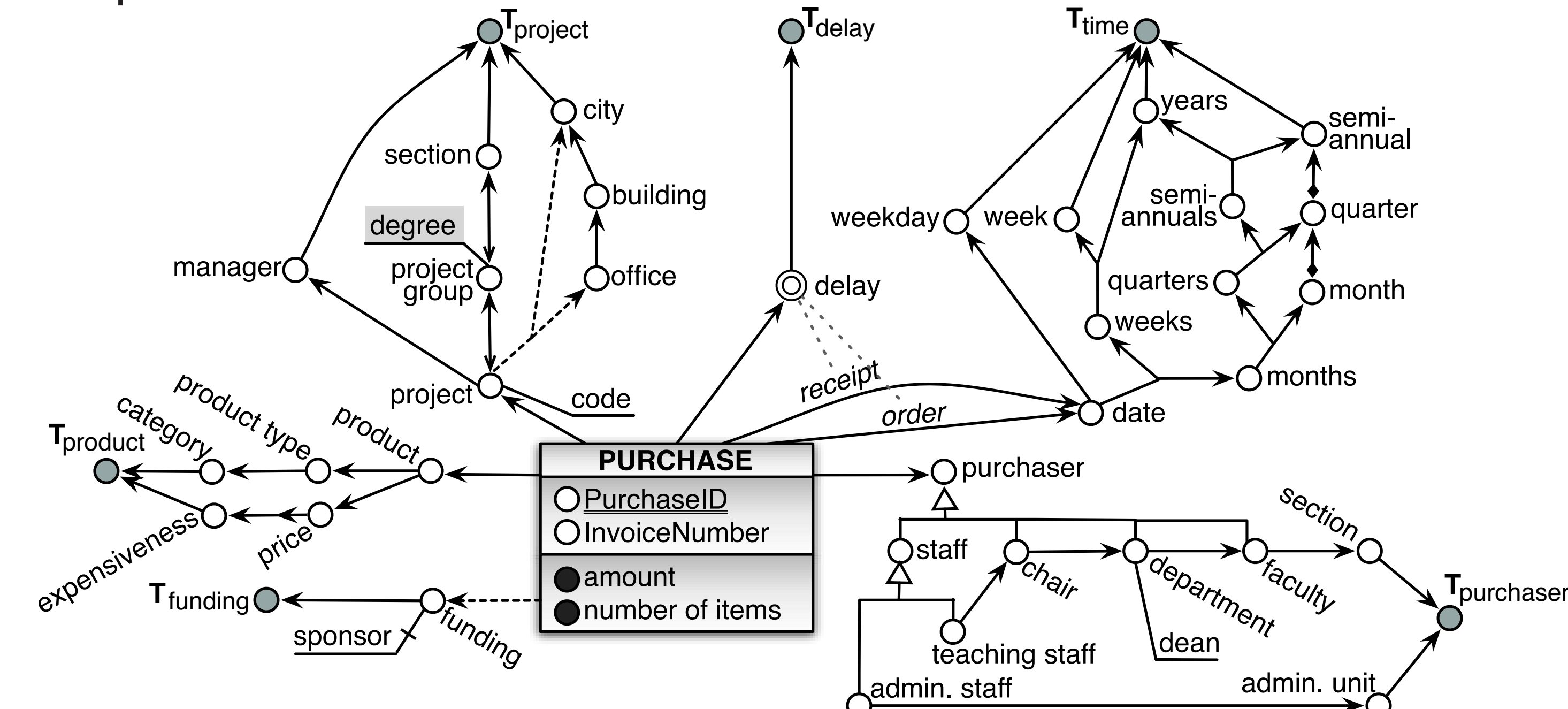

4 OLAP for Surgical Workflow Analysis

Sample Query: *For each surgery of type discectomy and each bone ablating instrument, return the average duration of a work step, in which that instrument was used by a surgeon.*

Generated SQL statement and its results as a pivot table:

```
SELECT S.SurgeryID, IT.Name, I.Name,
       COUNT(*) AS Occurrence, AVG(Duration) AS Average_duration
FROM WORKSTEP WS, INSTRUMENT I, INSTRUMENT_TYPE IT, ACTIVITY A,
     WORKFLOW WF, SURGERY S, PARTICIPANT P, POSITION PS
WHERE S.SurgeryID IN
      (SELECT SurgeryID FROM SURGERY_DISCIPLINE WHERE DisciplineID IN
          (SELECT DisciplineID FROM Discipline WHERE name = ''discectomy''))
      AND WS.InstrumentID = I.InstrumentID AND I.TypeID = IT.TypeID
      AND IT.Name = 'bone ablating' AND WS.ActuatorID = P.PersonID
      AND P.PositionID = PS.PositionID AND WS.ActivityID = A.ActivityID
      AND PS.Name = 'surgeon' AND WF.Run = 0
      AND A.RecordID = WF.RecordID AND WF.SurgeryID = S.SurgeryID
GROUP BY ROLLUP (SurgeryID, IT.TypeID, IT.Name, I.InstrumentID, I.Name)
```

		Measures							
		◉ Occurrence				◉ Average duration			
Dimensions		SurgeryID							
Instrument Group	Instrument	A	B	C	D	A	B	C	D
- bone ablating	mallet/chisel	0	3	1	1	00:00	00:23	00:34	00:50
	punch	9	22	10	9	02:38	00:35	00:46	01:27
	trephine	3	0	7	0	02:18	00:00	00:43	00:00
bone ablating Total		12	25	18	10	02:33	00:33	00:45	01:24

5 Enhanced Decomposition Trees
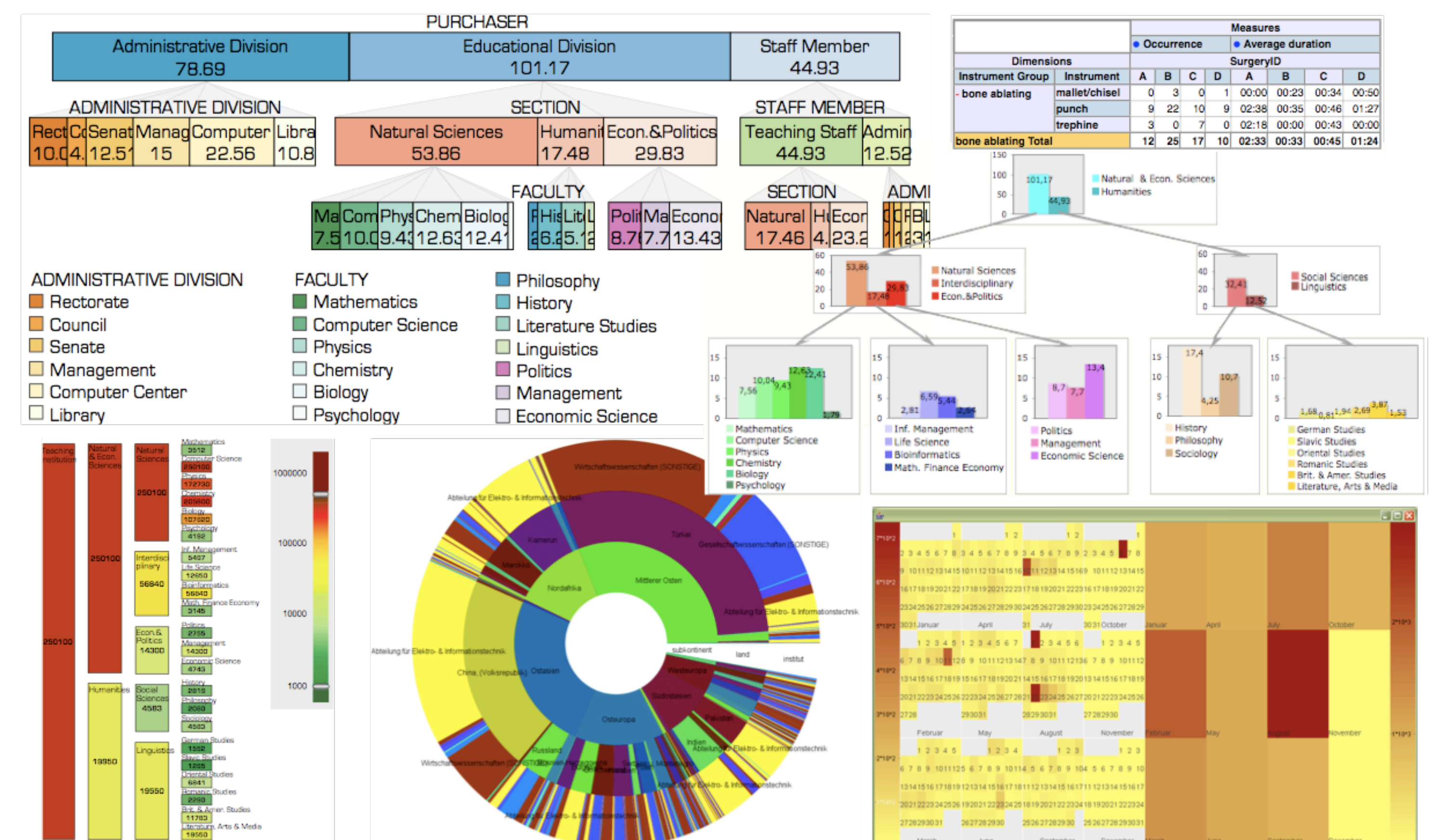
References

[1] S. Mansmann, T. Neumuth, O. Burgert, M. Röger, and M. H. Scholl. Conceptual Data Warehouse Design Methodology for Business Process Intelligence. In *Complex Data Warehousing and Knowledge Discovery for Advanced Retrieval Development: Innovative Methods and Applications*. 2008. (Book chapter).

[2] S. Mansmann, T. Neumuth, and M. H. Scholl. Multidimensional Data Modeling for Business Process Analysis. In *ER 2007: Proceedings of the 26th International Conference on Conceptual Modeling*, 2007.

[3] S. Mansmann and M. H. Scholl. Empowering the OLAP Technology to Support Complex Dimension Hierarchies. *International Journal of Data Warehousing and Mining*, 3(4), 2007. (Invited as DaWaK'06 Best Paper).

[4] S. Mansmann and M. H. Scholl. Extending the Multidimensional Data Model to Handle Complex Data. *Journal of Computing Science and Engineering*, 1(2), 2008. (Invited tutorial paper).

[5] T. Neumuth, S. Mansmann, M. H. Scholl, and O. Burgert. Data Warehousing Technology for Surgical Workflow Analysis. In *CBMS 2008: Proceedings of the 21st IEEE International Symposium on Computer-Based Medical Systems*, 2008.

[6] M. H. Scholl and S. Mansmann. Visual OLAP (Online Analytical Processing). In *Encyclopedia of Database Systems*. 2008.

DFG Colloquium Konstanz

Svetlana Mansmann — PhD Track — Associated Member since 01.09.2004 Work Group — Databases, Data Mining and Visualization

26 June, 2008

Research Training Group 1042 (GK) — Explorative Analysis and Visualization of Large Information Spaces