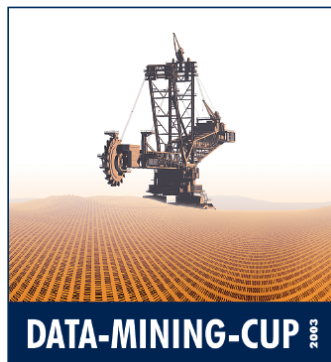


---

**Vortrag auf den 4. DATA-MINING-CUP Anwendertagen  
(Chemnitz, 18.-20. Juni 2003)**



<http://www.data-mining-cup.de>

Copyright-Hinweis:

Das Urheberrecht des folgenden Vortrags liegt beim Author. Verbreitung, Vervielfältigung und Kopie, auch auszugsweise, ist nur mit schriftlicher Genehmigung des Authors erlaubt.



## Explorative Data Mining for End Users – Pharmaco-Economy in 200 Dimensions



### About Panoratio

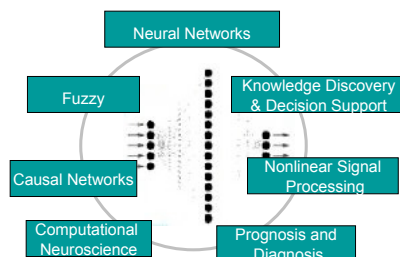
We provide a new technology to condense the analytical essence of Gigabytes of data to a “Database Image™” of just a few Megabyte.

=> Interactively access, experience and deploy the knowledge residing in ever-increasing “data silos” in an intuitive way.

=> Unparalleled performance allows to implement a complete new class of applications and unleash the potential for “Real-Time Information Enterprise”.

#### Panoratio Technology Roots:

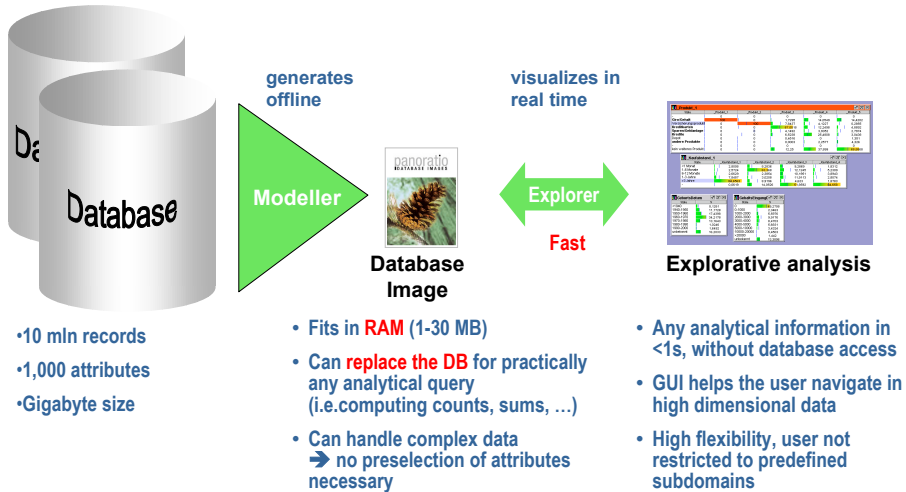
Siemens Corporate Technology, one of the worlds largest R&D labs for Neural Networks and Machine Learning





## Database Images

99% of the information relevant for analytics in 1% of the size



## Database Image Technology

### Technical USPs

- **Speed and Dimensionality** - up to 1,000 times as fast as common OLAP applications – practically no limitations with respect to number of dimensions
- **Compactness and Portability** - Database Images in MB size (as compared to GB OLAP-cubes) can be easily distributed (even via e-mail) and run on standard PC hardware.
- **Separation of Storage and Analysis** – optimal storage (requiring a relational DB scheme) and analysis (optimally supported by Database Images) is separated => less load on the DB, less expensive hardware
- **Data Security** – no access to DB and individual records, only statistics

### Generic User Benefits

- **Interactive 360° View** – decision makers may exploit themselves all relevant data in an intuitive way.
- **Communications** – Database Images can be easily exchanged to deliver e.g. “interactive reports” or to market analytical results without comprising data privacy.
- **Seamless Processes in a Real-Time Enterprise** – End-users are enabled to directly turn data into knowledge with little need for support from business analysts and Datamining experts => less friction and better informed decisions
- **One consistent approach** - over all levels of analysis allowing user to smoothly migrate from one level to the next.
- **No Mutual Limitation of Technology and Business** – no goal specific data modelling required, no need to predefine analyses



## Analysis of Pharmaco-Economical Data



### Siemens Health Services

- leading hospital IT outsourcing partner in the US
- administrates billing records and medical files of 150 mln. transactions per day
- offers studies based on their data to pharmaceutical companies

### Project Goals

- shorten data analysis process
- allow pharmaceutical companies to steer information selection more directly

### Presented here: Drug comparison study

- compare use of two competing drugs
- discover new sales potential
- discover unexpected medical application



## Our Solution

### source data

#### 200 dimensions

- medical information (diagnoses, symptoms, ...)
- social and demographic info (sex, age, ...)
- financial information (insurance type, costs, ...)
- hospital information (type, size, region, ...)

#### 2.5 million records

Offline  
compression



Database  
Image  
10MB

Interactive  
response times  
for any query

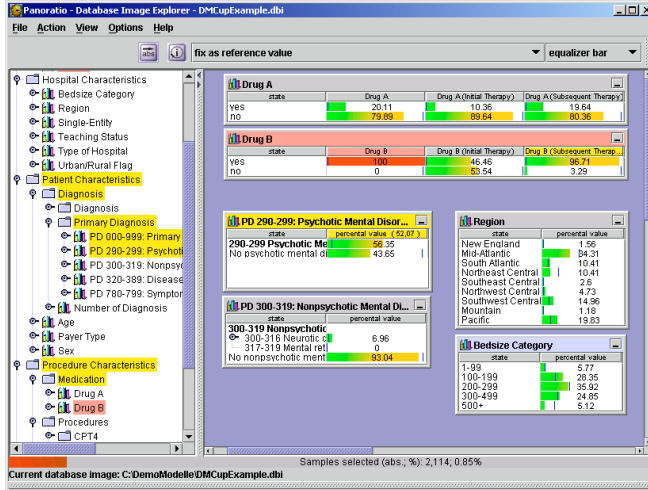
### flexible explorative analysis



- user typically watches 10 variables at one time
- user can open any of the remaining 190 variables any time
- tool alerts the user which variables might be relevant in the current context



## Example Discoveries

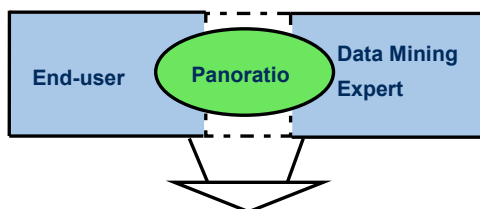
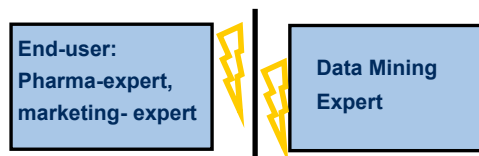


- many patients receive *both* drugs
- both drugs used under similar medical conditions – with interesting exceptions
  - ➔ different medical properties or marketing issue?
- both drugs less popular in large than in small hospitals. (even when conditioned on medical condition, region, or cost)
  - ➔ marketing/sales-issue?



## Database Image is delivered to the end user

➔ efficient cooperation instead of lengthy iterations between domain- and data mining experts



Valuable information / decision support

### Traditional approach

- End-user predefines specific questions
- Experts perform analysis limited to the pre-defined scope
- End-user has follow-up questions
- ➔ Time consuming iterations to gather meaningful information

### Panoratio approach

- End-users empowered to explore data interactively
- Instantaneous response times encourage end-user to use data widely and wisely
- Follow up questions can be answered by end-users themselves



## Advantages

- **Fast**  
even in the context of large amounts of data and many dimensions.
- **Flexible**  
All information is available instantaneously; no need to restrict oneself to a low number of quantities ahead of time to avoid technical performance issues.
- **Verifiable**  
All information displayed has a well defined meaning that can be verified without analytics expertise.
- **Intuitive**  
analysis tool can be used directly by domain experts – no time consuming iterations with analysis experts.



## Examples from other Application Areas

### Siemens Westinghouse: Analysis of Gas Turbine Humming

“Database Images” ...

- ... allowed to interactively explore the performance of turbines.
- ... allowed to find optimal operation modes, optimize performance
- ... and, hence, save costs.

### KarstadtQuelle:

“Database Images” ...

- ... make knowledge from huge amounts of sales data directly available to marketing departments.
- ... enable “online list broking” for sales teams.
- ... speed up analysis cycles and the communication between datamining experts and marketing departments.

### Direktanlagebank

- ... analyze customer behavior and typical product sequences.
- ... identify “sleeping” customers with a high customer value.

### Webmining for siemens.com and Siemens Employee Portal:

“Database Images” allow to

- ... interactively explore the traffic on web sites.
- ... identify “weak spots” with respect to the usage of websites.

**Thank You!**



**Dance with Your Data.**

For more information contact:

Panoratio Database Images GmbH  
München

[info@panoratio.com](mailto:info@panoratio.com)  
<http://www.panoratio.com>