# MS I Dr. Wälti AG

# Institut für angewandte Mathematik, Statistik, Informatik

Location: Buchs SG

Foundation: 1981

Subsidiary Firm: InfoCodex SA











i-expo award 2006 in Paris for:

The best and most innovative Economic Intelligence tool



## Applications / Benefit

### 1. Information Research

- Attain an overview
- Enhancement to internet search engines
- Thematic radar







## Applications / Benefit

### 2. Exploit Company Internal Knowledge

What is where
(automatic shelving of documents into well organized categories)

Coordination of knowledge resources

Company internal search engine



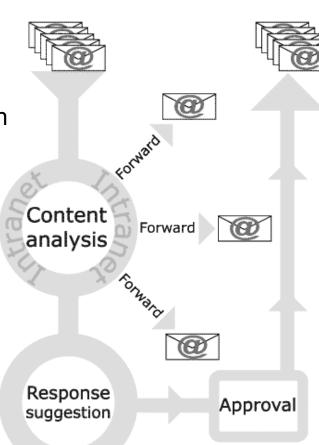




## Applications / Benefit

### 3. Process Automation

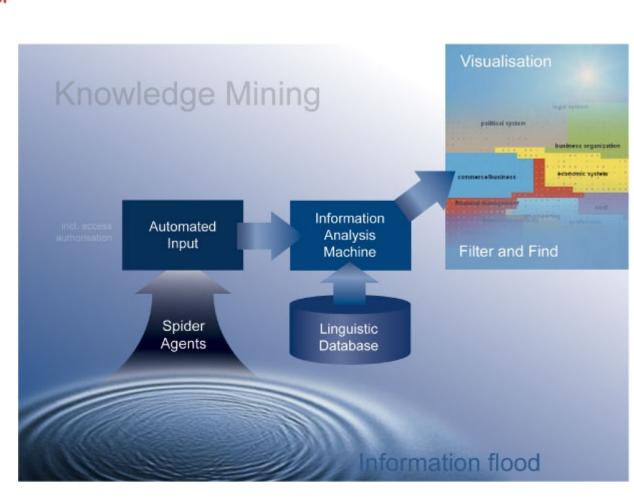
- Content analysis → categorization
- E-mail forwarding
- Response management







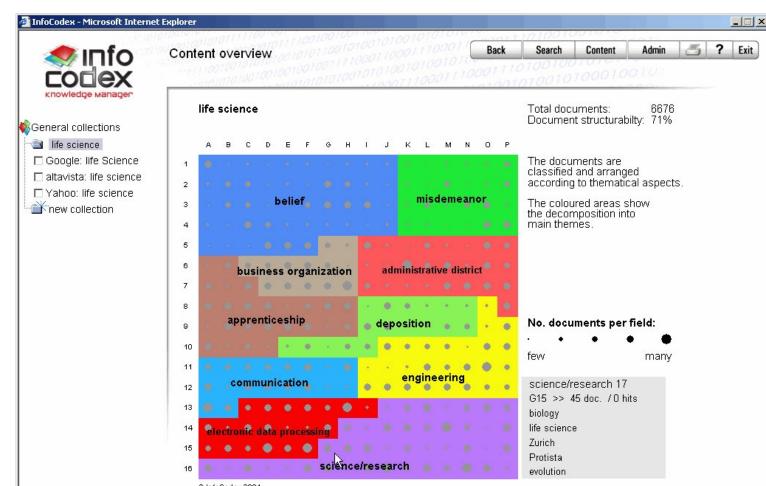
## Technology





## Simplicity Quality Performance

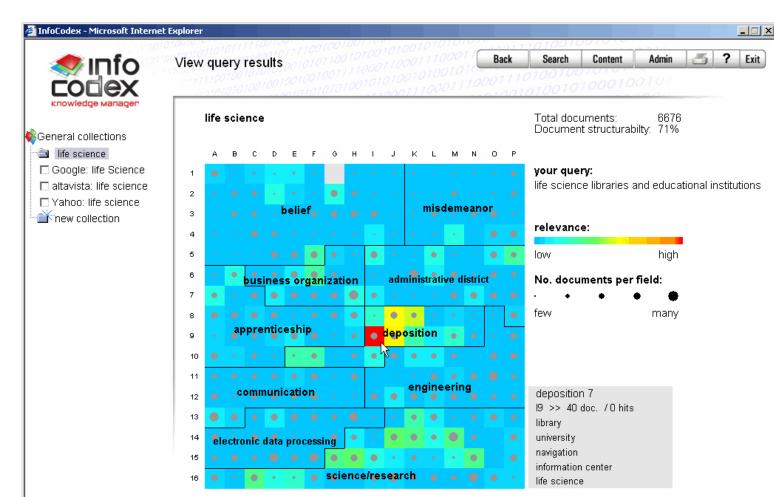
- Categorization of documents without any training
- Graphical overview over the whole collection





## Simplicity Quality Performance

Heat-map representation of search results (essential aid in the case of large result lists)





## Simplicity Quality Performance

### **Recognition of content:**

Linguistics + taxonomy + statistics + neural networks

### **Combination of languages:**

English + French + German + Italian

### **Search possibilities:**

Traditional full-text search Synonym search Similarity search Index search

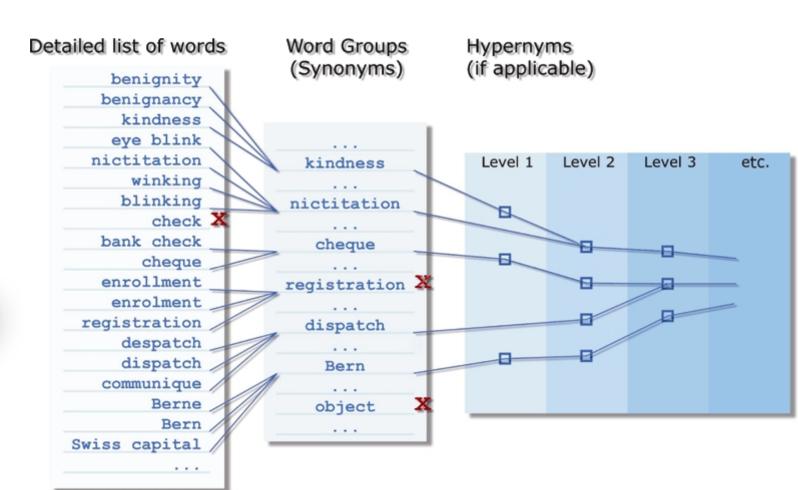




## Simplicity

## Quality

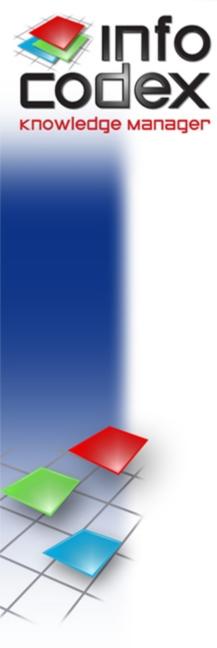
## Performance





- Based on the benchmark-winning WIDAS sytem
- Accurate categorization of millions of documents





Task

**Raw Data** 

105 GB

Simplicity Quality Performance

Benchmark of the SFSC

Statistical Data Warehouse

Metadata (code descriptions, structural information)

**Relational Database** 

for Detailed Microdata

(2)

Multidimensional DB for Aggregated Macrodata

Aggregation

**Ad-hoc Queries** 

from Multidimensional DB,

**Ad-hoc Queries** from Relational DB, Result Table Generation Results

200 \_

0

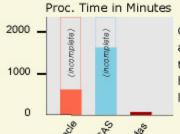
(1) Import

> Proc. Time in Seconds Sum for 600 queries 1-7 400

Joins / Table Generation Proc. Time in Seconds Sum for 100 queries 1-2

appe 50

Proc. Time in Minutes 1500 1000 . 500



Only Widas was able to complete the critical step (3) beyond the first level of aggregation







## Unique features of InfoCodex

InfoCodex is the only system having a multilingual knowledge repository (linguistic database linked to a taxonomy tree).

The projection of documents onto a customed content space for the deployment of self-organizing neural networks is patented in Europe and the USA.

### **Resulting Benefits**

immediate use

categorization of documents without any human training

synonym recognition

true cross-language retrieval

similarity search

find documents similar to an incoming paper (e.g. Email)

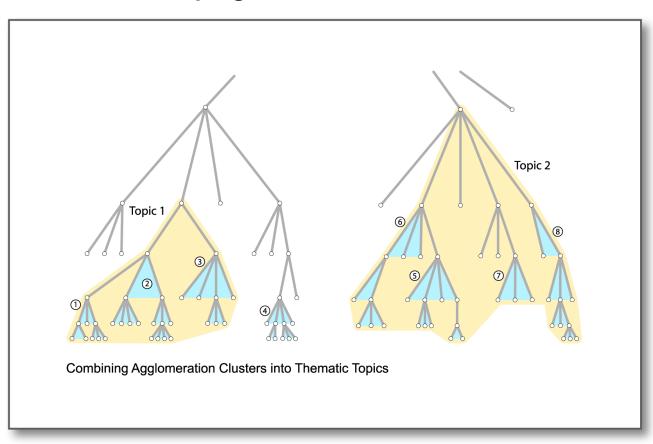
topic clustering

essential aid in the case of large search result lists



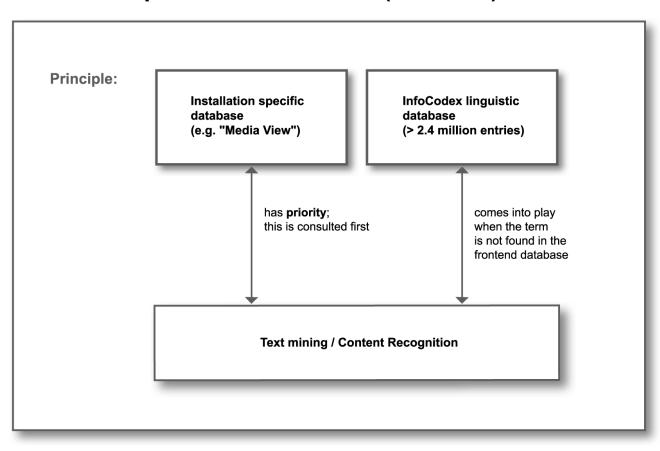


## **Thematical Grouping of Documents**



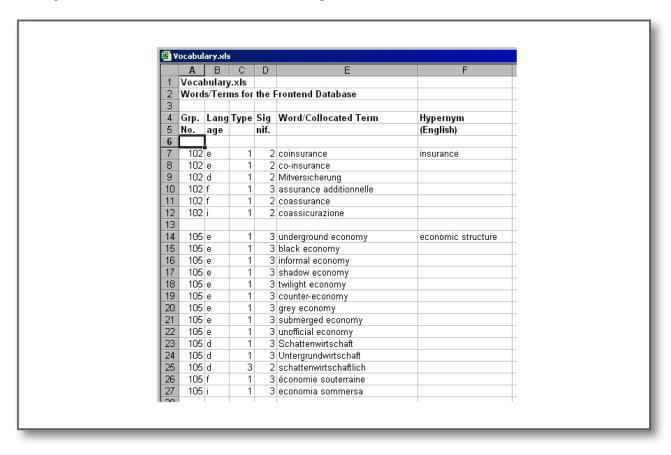


### **Customer Specific Vocabularies (Thesauri)**





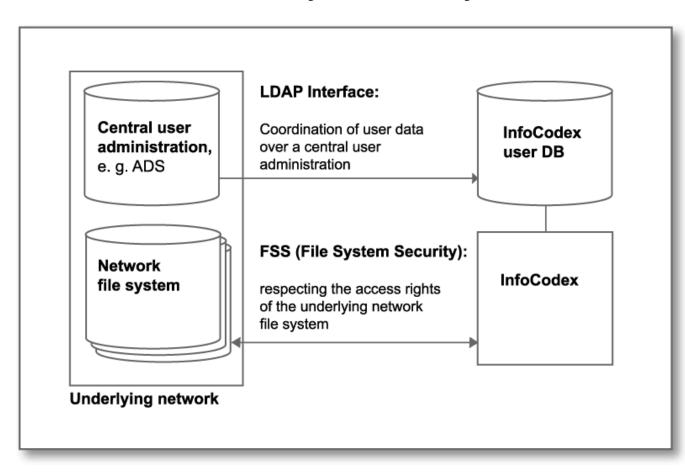
### **Preparation of Customer Specific Vocabularies**





## Security

### **LDAP Interface and File System Security**





## Security

## **Protected Regions / IC Domains**

### **Main IC Domain**

- Administrator: "sysadmin"
- Users: all InfoCodex users

### IC Domain "Biotech"

- Administrator: "Headbio"
- Users: all of group "bio"

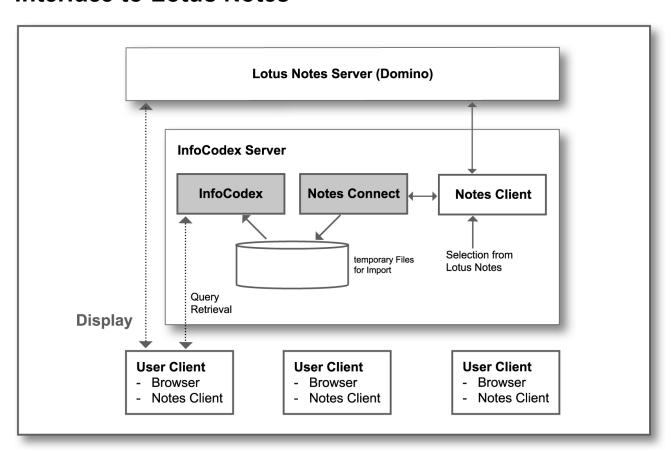
### IC Domain "BioMgm"

- Administrator: "Headbio"
- Users: only "Headbio"





### **Interface to Lotus Notes**





## Integration

# Integration into MS-Sharepoint

