InfoCodex
Semantic Technologies

Phone +41 (0)81 750 53 00
E-mail support@infocodex.com
Web www.infocodex.com

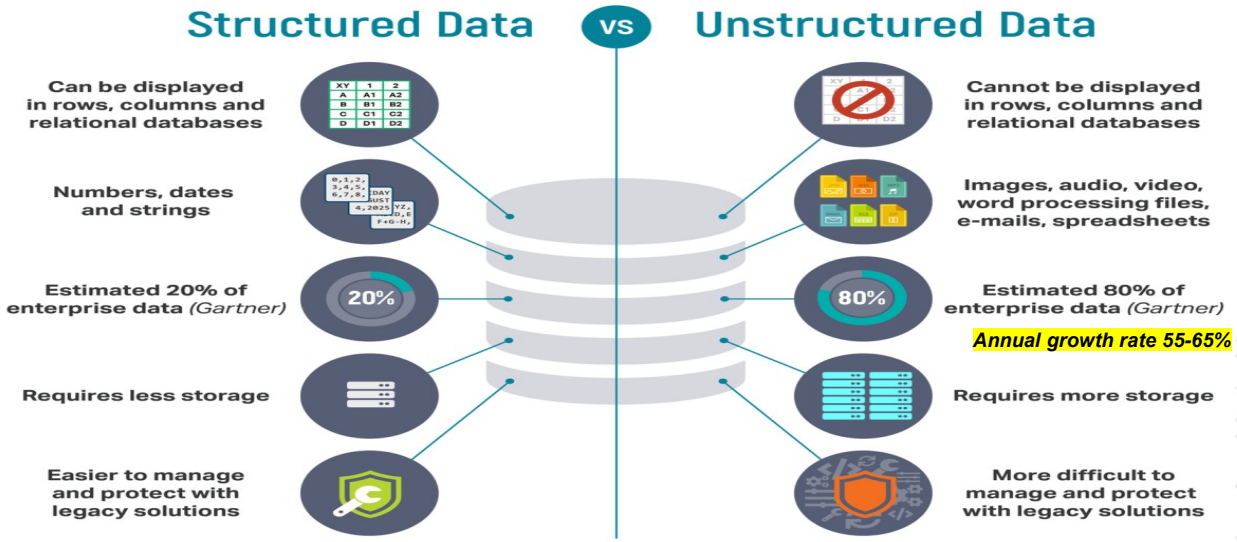
InfoCodex - a Semantic Artificial Intelligence Tool for *Collecting, Structuring, and Content Analysis of Unstructured Free Text* information

Turning Information into Knowledge

InfoCodex
Semantic Technologies
Phone +41 (0)81 750 53 00
E-mail support@infocodex.com
Web www.infocodex.com

Importance of Unstructured Data (Gartner, IBM)

Results of an Analysis by Gartner Group (2020)



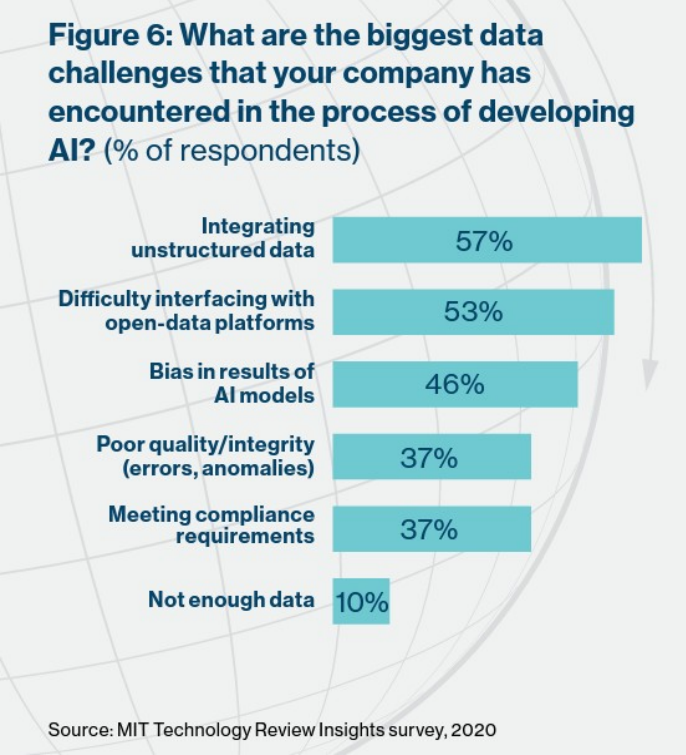
Predominance of free texts

Difficult to evalute

InfoCodex
Semantic Technologies
Phone +41 (0)81 750 53 00
E-mail support@infocodex.com
Web www.infocodex.com

Biggest Challenges in Developing Artificial Intelligence (AI)

Survey of the MIT (2020) (2020)



Integration of free texts

Interfaces to free text sources

InfoCodex
Semantic Technologies

Phone +41 (0)81 750 53 00
E-mail support@infocodex.com
Web www.infocodex.com

Conclusion: Today's Information Management Problems

Volume/growth:

- **80%** of the relevant data are **unstructured free texts** (Gartner, IBM)
- But their **evaluation is difficult** (is different from the traditional evaluation methods)

Processing:

- Biggest challenges in the AI are: **Integration/usage of free texts** (MIT 2020) (not the rule-based artifical intelligence such as robotics)
- Particular difficulties: Interfaces to free text sources

But: Are the unstructured free texts really used in practice?

InfoCodex
Semantic Technologies
Phone +41 (0)81 750 53 00
E-mail support@infocodex.com
Web www.infocodex.com

Some Remarks on the Effective Usage of Free Texts

1. Prime example: The crash case of Kodak (2005: 60‘000 employees, 2011: insolvency)
2. 

German proverb
3. Decision making is simpler in the case of structured data

with structured data it is much easier

on a free text basis it‘s more demanding

4. The competent software company SAP bought Qualtrics for $ 8 billion

InfoCodex
Semantic Technologies

Phone +41 (0)81 750 53 00
E-mail support@infocodex.com
Web www.infocodex.com

InfoCodex: 20 Years Experience in Semantics + Artificial Intelligence

The company

Swiss Software Company
Founded: 2002
Competences: Sem. Tech./ AI
Headquarter: Buchs SG
Employees: 7

Originated from:

MSI Dr. Wälti AG founded 1981 with core competences in high performance databases and data analysis (Big Data)

The product

Innovative software tool for the handling and evaluation of unstructured free text, based on semantics and artificial intelligence

Potential applications and added values

These are illustrated by practical use cases on the following slides

InfoCodex
Semantic Technologies

Phone +41 (0)81 750 53 00
E-mail support@infocodex.com
Web www.infocodex.com

Use Case 1: Knowledge Transfer / Don‘t Reinvent the Wheel

- there is a great, but hidden potential in this use case
- in R&D, in contract management, in sales, in marketing etc.
- difficulty: learning from others (in the unstructured knowledge repertoire)

A statement:
A new R&D project started today has already been processed in the last 20 years in the same company with a probability of 80%

(Head of R&D of Cerberus)

InfoCodex
Semantic Technologies

Phone +41 (0)81 750 53 00
E-mail support@infocodex.com
Web www.infocodex.com

Use Case 2: Knowledge Management / Overview

- How can I extract the essentials from 10,000 documents?
- How can I avoid the risk of disregard facts without having an overview?
- Or how do I recognize hidden opportunities in the flood of information?

The German ex-investment banker Rainer Voss said:

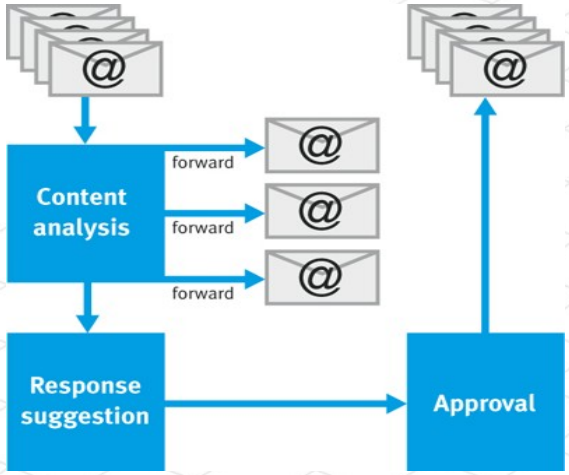
Imagine a filing cabinet with 5000 contract folders, each having 1000 pages.

How is one supposed to achieve transparency and gain an overview with justifiable effort to avoid the risk associated with non-compliance with the contracts?

InfoCodex
Semantic Technologies
Phone +41 (0)81 750 53 00
E-mail support@infocodex.com
Web www.infocodex.com

Use Case 3: Response Management

- Cross-lingual rapid assessment of customer complaints, inquiries, and suggestions
- Filter out spam and forwarding to the correct addressees
- Basis for complaint management + preventive measures in production



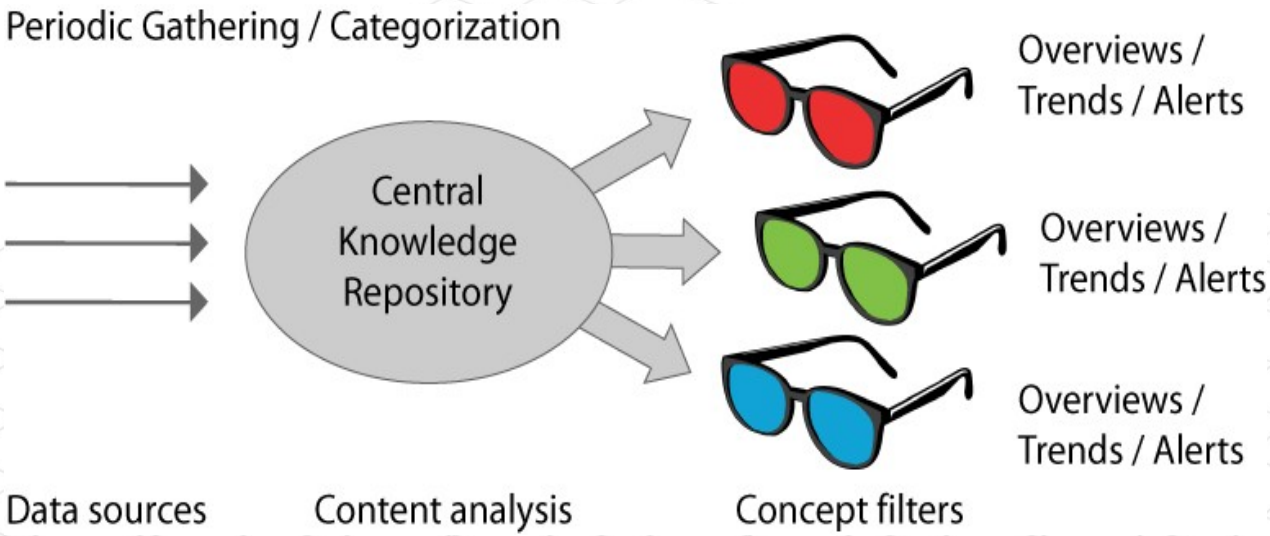
Added Values:

The speed and quality of the answers have a significant impact on the company's image

InfoCodex
Semantic Technologies
Phone +41 (0)81 750 53 00
E-mail support@infocodex.com
Web www.infocodex.com

Use Case 4: Market Intelligence / Recognition of Trends

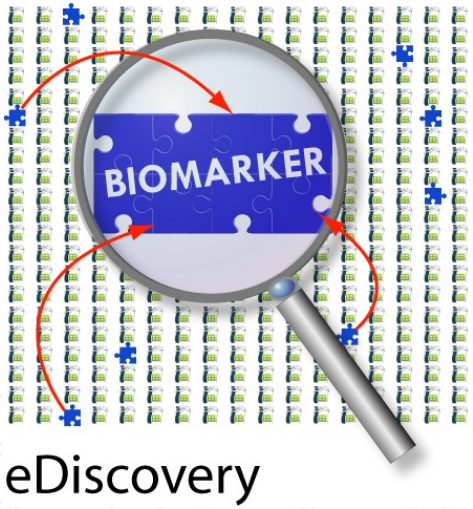
- Surveillance of new technological developments
- Observation of competitors and/or customers
- Early recognition of trends, risks and opportunities



InfoCodex
Semantic Technologies
Phone +41 (0)81 750 53 00
E-mail support@infocodex.com
Web www.infocodex.com

Use Case 5: Knowledge Discovery

- Discovery of new, unknown relationships and facts hidden in large text text collections (literature, news media)
- Early recognition of „emerging risks“ and chances



eDiscovery

The objective of Merck's benchmark:

- Test pure machine intelligence for "semantic" drug research

The tasks:

- Discover novel biomarkers for diabetes by analyzing 120'000 medical publications (PubMed)

This functionality, the "Holy Grail of Text Mining", can definitely **not be provided by NLP** or Pattern Recognition.

Further applications:

- Forensics/ fraud discovery, detection of cyber risks
- Compliance/ detection of hidden violations
- Research in general/ detection of new technologies

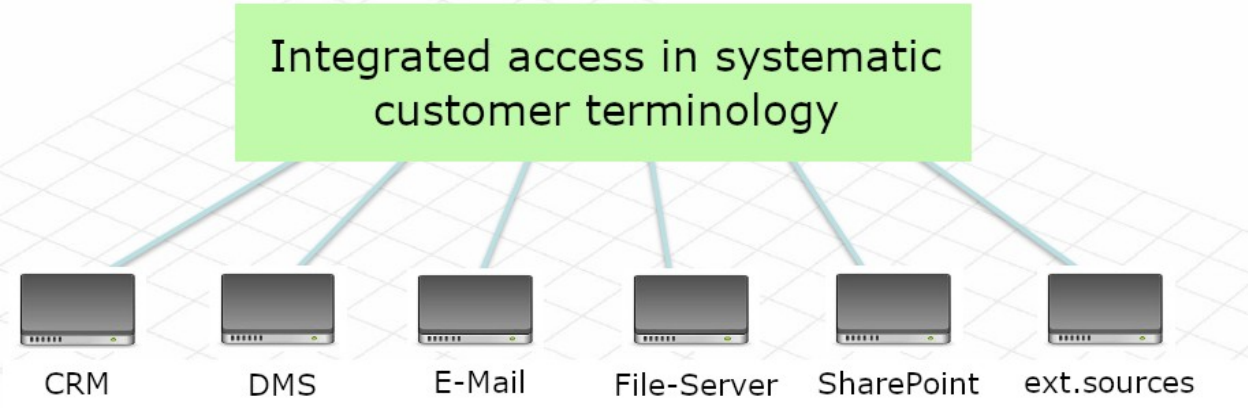
InfoCodex
Semantic Technologies

Phone +41 (0)81 750 53 00
E-mail support@infocodex.com
Web www.infocodex.com

Use Case 6: Efficient Information Retrieval / Controlled Tagging

Statement of the problem:

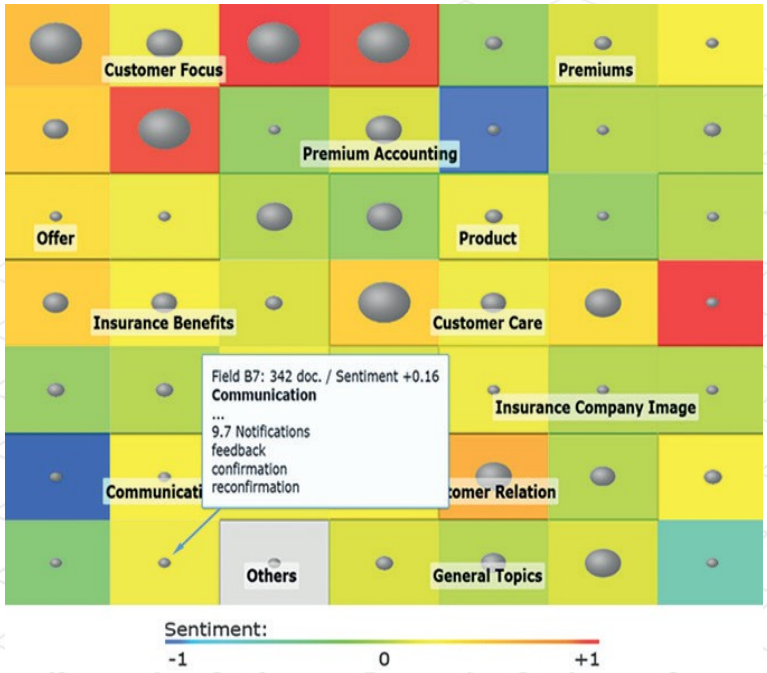
- The relevant documents consist often of unstructured documents scattered over various sources
- For an efficient retrieval, the documents should be automatically categorized and provided with keywords in a systematic customer-specific terminology



InfoCodex
Semantic Technologies
Phone +41 (0)81 750 53 00
E-mail support@infocodex.com
Web www.infocodex.com

Use-Case 7: Report and Sentiment Analysis

- Unstructured customer feedback or reports of collaborators often lie fallow and are not used
- Generation of added values through automatic categorization and thematic content evaluation of free texts (sentiment analysis, e.g. positive/negative or pro/contra)



Examples:

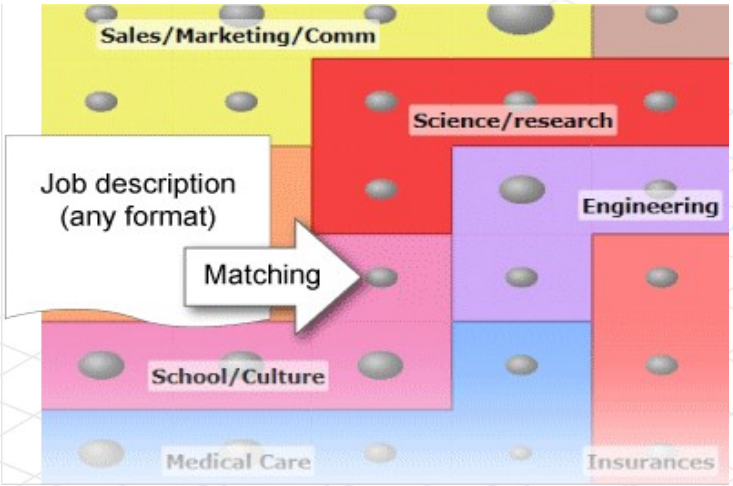
- Assessment of customer feedback
- Assessment of hotline messages from branch offices and news media
- Evaluation of daily/weekly reports

InfoCodex
Semantic Technologies

Phone +41 (0)81 750 53 00
E-mail support@infocodex.com
Web www.infocodex.com

Use Case 8: Profile Matching

- Find similar documents by comparing their contents
- Cross-language matching of text content,
  free of subjective and laboriously created keywords



Examples:

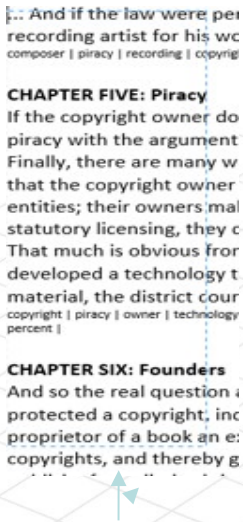
- Matching job decriptions with CVs (curriculum vitaes)
- Placing advertising texts in Webpages

Note:
In 2010, three American economists received the Nobel Prize in economy, because they had proved that the concurrently high levels of open positions and unemployment are partly due the missing matching quality.

InfoCodex
Semantic Technologies
Phone +41 (0)81 750 53 00
E-mail support@infocodex.com
Web www.infocodex.com

Use Case 9: Generation of Document Summaries

- Generate headlines of News/Reports
- Brake down into main sections in the case of large documents



Subdivison of long articles

Examples:

- Summary of E-Books
- Summary of medical/juridical reports (50-500 pages)
- Summary of short or long Internet publications

InfoCodex
Semantic Technologies

Phone +41 (0)81 750 53 00
E-mail support@infocodex.com
Web www.infocodex.com

Why can InfoCodex Solve these Applications?

This is because InfoCodex has the followings USPs:

1. **No training of the knowledge structure required**

   It can be applied immediately even in new and unknown unknown situations (in contrast to NLP-based technologies).

2. **Knowledge discovery capability**

   It can discover hidden relationships and new facts from analyzing large amounts of free text documents (in contrast to NLP-technologies with their sentence-by-sentence analysis).

The handling of these applications needs these funtionalities.

InfoCodex
Semantic Technologies
Phone +41 (0)81 750 53 00
E-mail support@infocodex.com
Web www.infocodex.com

Reasons for these Unique Features

The USPs are the result of the combination

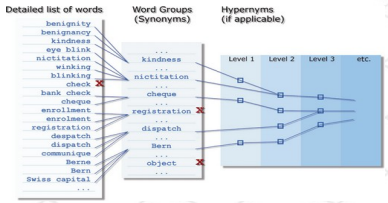
Universal linguistic knowledge repository

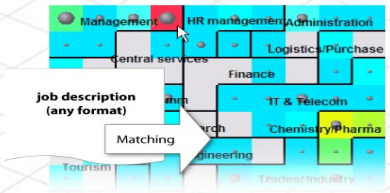
+

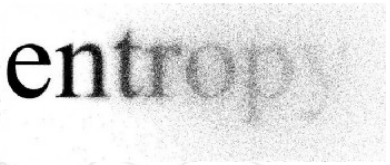
Self-organizing neural network (Kohonen map)

+

Math.-statistical analyses







Cerebrum (experience, memory)

Cerebellum (intelligence, AI)

Data analysis

These new technologies are needed for handling innovative and creative situations.

InfoCodex
Semantic Technologies

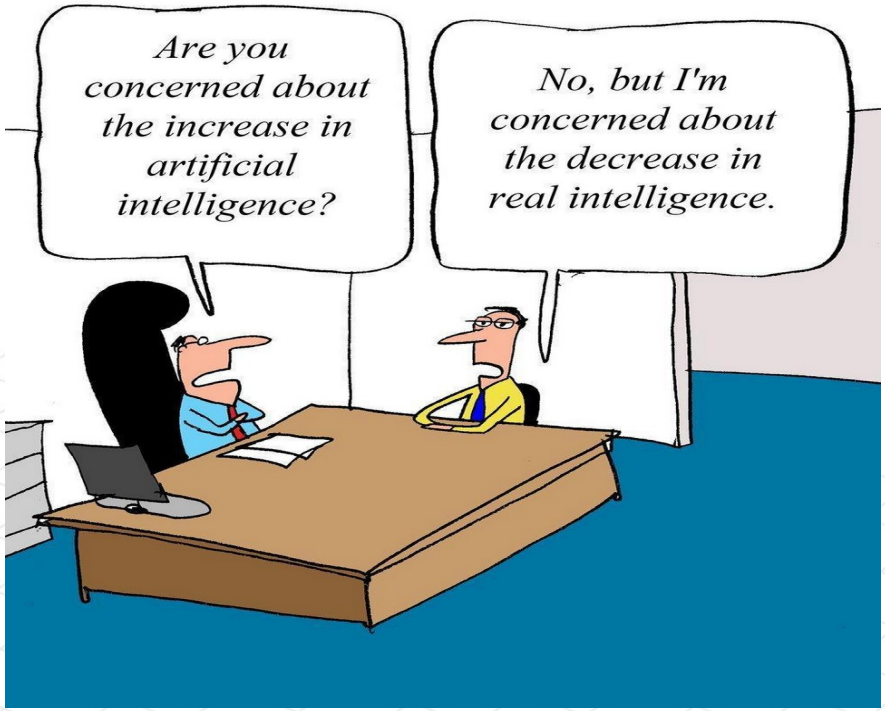
Phone +41 (0)81 750 53 00
E-mail support@infocodex.com
Web www.infocodex.com

Backup Slides for the Answering Possible Questions

InfoCodex
Semantic Technologies

Phone +41 (0)81 750 53 00
E-mail support@infocodex.com
Web www.infocodex.com

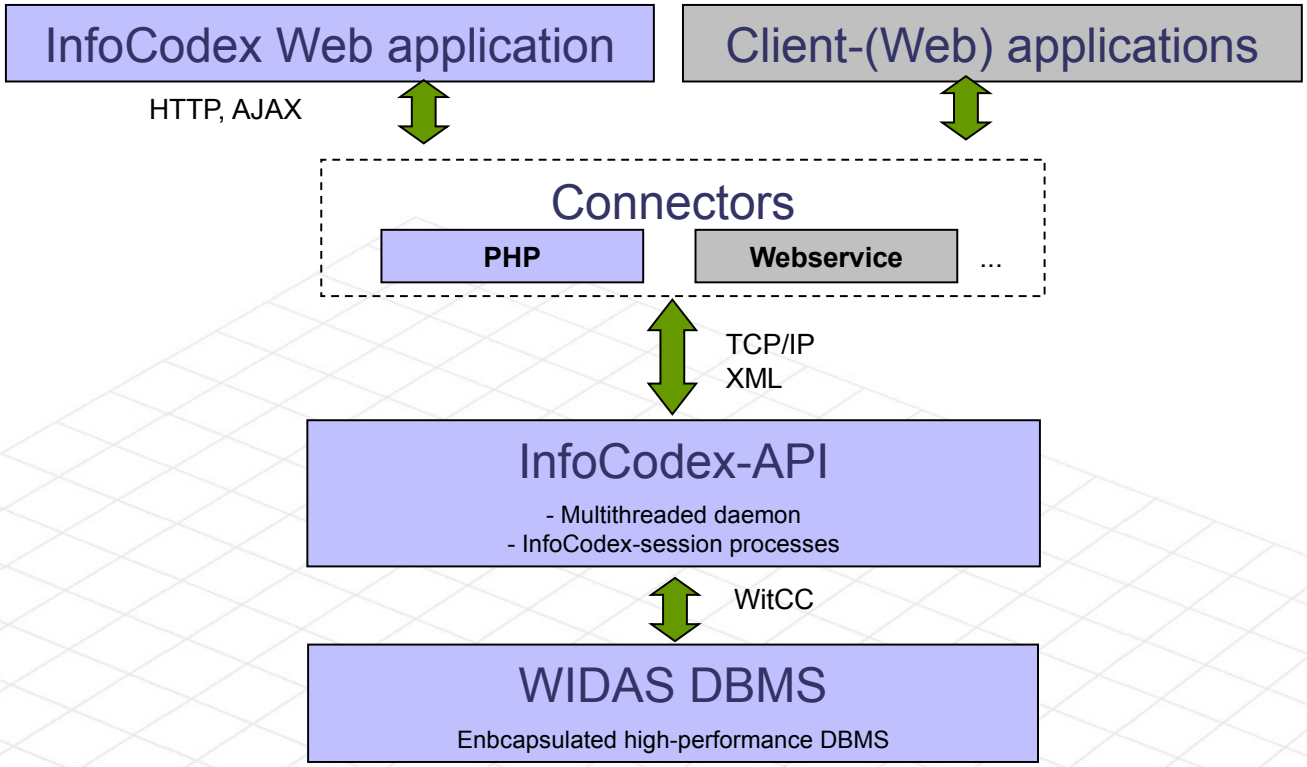
Artificial Intelligence on the Rise



InfoCodex
Semantic Technologies
Phone +41 (0)81 750 53 00
E-mail support@infocodex.com
Web www.infocodex.com

II. Technical Properties

System Architecture of InfoCodex



InfoCodex
Semantic Technologies

Phone +41 (0)81 750 53 00
E-mail support@infocodex.com
Web www.infocodex.com

II. Technical Properties

Spider Agents for Document Import

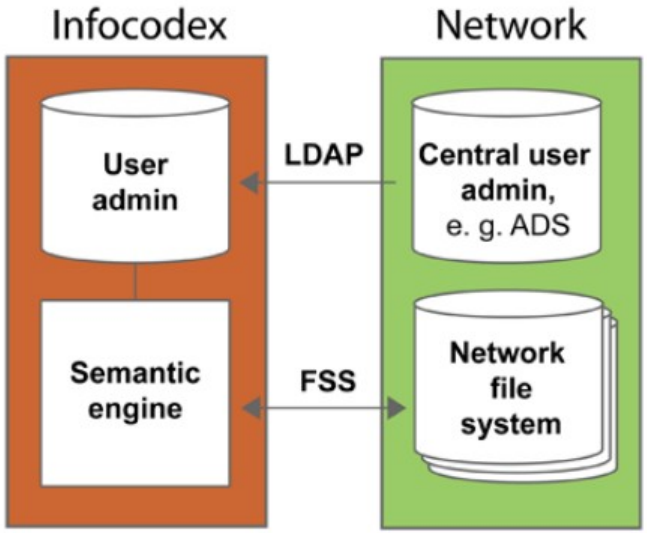
File Server and Web	Common document formats such as MS Word, PDF, Excel, PPT presentations, PostScript, RTF, TXT, HTML, XML, RSS, MSG, EML (in original or zipped form); other file formats for which an i-Filter is available
Mailboxes	Outlook, Outlook Express, Thunderbird, Exchange Server, Lotus Notes (including attachments of the e-mails in original or zipped format)
MS SharePoint	Standard interface available
DBMS/DMS	Needs individual connectors (e.g. using ODBC)

InfoCodex
Semantic Technologies
Phone +41 (0)81 750 53 00
E-mail support@infocodex.com
Web www.infocodex.com

II. Technical Properties

Privacy and Security

InfoCodex can be linked to the central user administration (e.g. ADS) by the LDAP Protocol. It fully respects the access rights given in ADS: File System Security



InfoCodex
Semantic Technologies

Phone +41 (0)81 750 53 00
E-mail support@infocodex.com
Web www.infocodex.com

II. Technical Properties

Scalability

The software implements multithreadi techniques and supports distributed processing features.

It allows the load to be spread across a range of dedicated servers.

The incorporated encapsulated DBMS is a purpose-built a high-performance system.

The size of the document collections is limited only by the available hardware resources (64-bit software).

InfoCodex
Semantic Technologies

Phone +41 (0)81 750 53 00
E-mail support@infocodex.com
Web www.infocodex.com

II. Technical Properties

International ICT Standards

Standards Recognized and Followed by InfoCodex:

Web side:	PHP, HTML, AJAX, HTTP, HTTPS; W3C-Standards OWL, RDF
Web servers:	Apache2 or IIS
Programs:	ANSI-C (GNU compiler), C++, C#
Interfaces:	XML, SOAP, LDAP (for coordination with a central user administration)
Linguistics:	WordNet of the the Princeton University, EuroVoc, DIN and others