Information Logistics in Networked Organizations

Issues, Concepts and Applications

ICEIS, June 2007

Kurt Sandkuhl
Jönköping University

Fraunhofer Institut Software- und Systemtechnik ISST

JÖNKÖPING UNIVERSITY

Overview

Motivation
• information disasters
• information overflow

Information Logistics
• Definition
• Related Areas

Information Demand Modelling

Evaluating Content

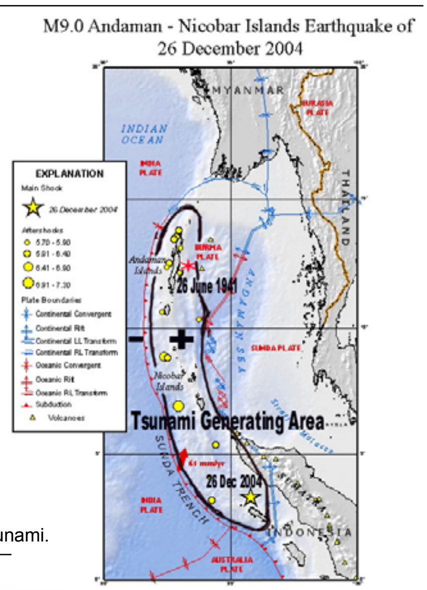
Selected Application Examples

Fraunhofer Institut Software- und Systemtechnik ISST

JÖNKÖPING UNIVERSITY

Information Disasters

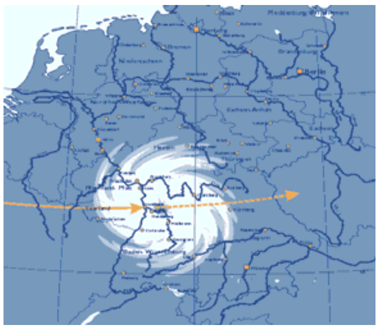
Tsunami of December 26, 2004



- South-East Asia and Indian Ocean
- Surprise for the people living in the affected areas
- But not for scientist aware of tectonic interactions in the region
- Seismic networks recorded incident
- No organisational infrastructure to pass warnings
- 220 000 + death toll

Source:
http://www.drgeorgepc.com/Tsunami2004Indonesia.html
Dr. George Pararas-Carayannis web page detailing the Sumatra tsunami.

Storm Disaster of December 1999



Bildquelle: Bayerisches Landesamt für Wasserwirtschaft

The storm disaster „Lothar“ of December 1999

- First warnings had been available on the Internet 13 hours prior to the storm
- Warnings over radio and television came too late for any precautions
- Damage could have been avoided if individuals had been warned early enough
- “Lothar” was also an information disaster

Fraud Disasters

San Francisco Chronicle

Tyco's Kozlowski guilty of fraud

Aide convicted also; jury in second trial finds they plotted to steal millions

Andrew Ross Sorkin, New York Times
Saturday, June 18, 2005

(06-18) 04:00 PDT New York -- Dennis Kozlowski, the former chairman and chief executive of Tyco International, and his top lieutenant were convicted Friday on fraud, conspiracy and grand larceny charges, bringing an end to a three-year-long case that came to symbolize an era of corporate greed and scandal.

Kozlowski and Mark Swartz, Tyco's former chief financial officer, were convicted by a New York State Supreme Court jury on all but one of 31 counts of grand larceny, conspiracy, falsifying business records and securities fraud.

"Today is a day of disappointment, but there's still hope," Kozlowski's attorney, Stephen Kaufman, said outside the Manhattan courthouse. "We look forward to the appeal, and we have confidence in that appeal."

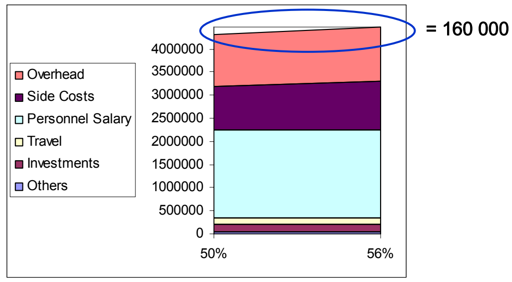
Fraunhofer Institut Software- und Systemtechnik
JÖNKÖPING UNIVERSITY

© Kurt Sandkuhl

Small Information Disasters (1)

Calculation of project costs

- Wage costs changed (from 50% to 56% surcharge)
- Project manager and CEO did not get this information
- Effect: organisation ”lost” 4% of the budget already at project start



Small Information Disasters (2)

Change in product requirements

- minor change in product requirements (other background color)
- change is not communicated to sub-supplier of one of the suppliers
- Effect:
  - acceptance test fails
  - product needs adjustment

Some Commonalities between the Cases

The information needed was electronically available
(Internet, information system, electronic document archive, etc.)

The person in need did get the information

Time is a critical factor
- Quick response to events or
- Change in information over time or
- Different and changing information sources

Demand of the person in need is decisive

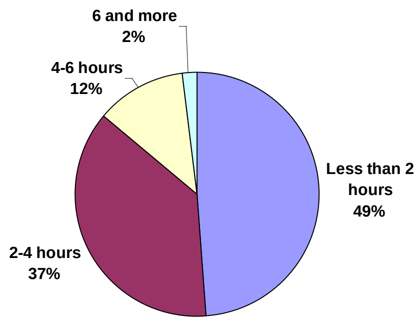
Challenge: improve information supply!

The other side of the Coin: Information Overflow

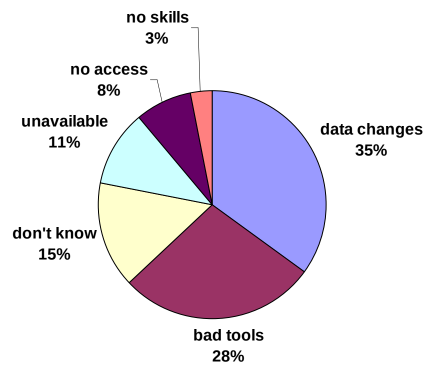
There is too much information
 5000 new books every day,
 total amount of information doubles every 5 years,
 550 billion web pages (life time: 50 -100 days), etc.

We are dependant on information
 information is an important production factor

Some Studies in this Context (1)



estimated time business professionals spend searching for information



Perceived impediments for locating information

Source: Perspectives on Information Retrieval. Delphi Group, 2002

Some Studies in this Context (2)

Gartner Group (Study on e-mail use 2003):
• On average 49 minutes per day for sorting e-mails

DNV (Technology Outlook 2004, ICT):
The bottleneck is no longer the availability of data (is growing exponentially) but rather information processing and consumption, which is growing only linearly.

Information Logistics

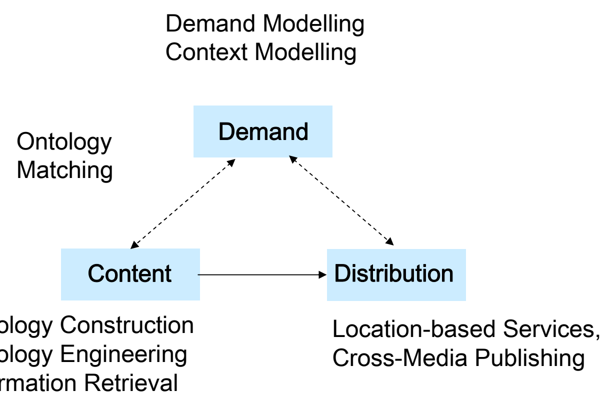
Information Logistics

Definition:

The main objective of Information Logistics is optimzed information provision and information flow. This is based on needs and demands with respect to the content, the time of delivery, the location, the presentation and the quality of information. The scope can be a single person, a target group, a machine/facility or any size of networked organisation.

The research field Information Logistics explores, develops and implements concepts, methods, technologies and solutions for the above mentioned purpose.

Initial Motivation: Information Logistics



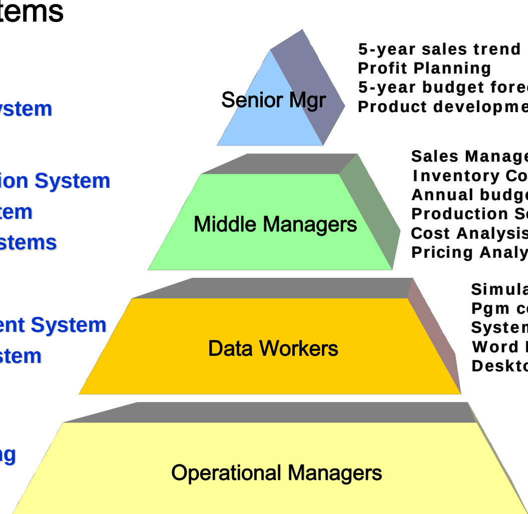
Information Systems

Executive Support System

Management Information System
Decision Support System
Intelligent Support Systems

Knowledge Management System
Office Automation System

Transaction Processing System



5-year sales trend
Profit Planning
5-year budget forecasting
Product development

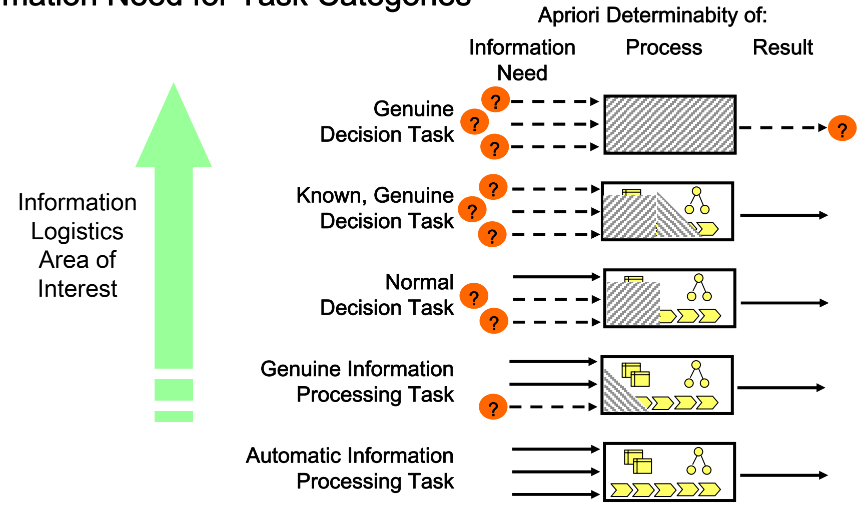
Sales Management
Inventory Control
Annual budget
Production Scheduling
Cost Analysis
Pricing Analysis

Simulation
Pgm coding
System support
Word Processing
Desktop Publishing

Order Processing
Fulfillment
Material Movement
A / R, A / P, GL
Payroll
POS

Source: Turban, Leidner, McLean, Wetherbe: Information Technology for Management, John Wiley & Sons, 5th Edition, 2005.

Information Need for Task Categories



Source: Task complexity affects information seeking and use.
Katriina Byström and Kalervo Järvelin
University of Tampere, Finland

Individual Information Need

Based on personal profile:
take different personal characteristics into account, like qualification and experience

Situation-based:
depending on current events and environment

Differentiate between role and individual

Active information supply:
not the user has to search for the information but the system provide the information actively

Adequate communication means:
right way to transmit information depending on situation

Fraunhofer Institut Software- und Systemtechnik
JÖNKÖPING UNIVERSITY

Demand Modeling

Information Demand

Information Need (Saracevic, 1975)
Information need is a psychological state associated with uncertainty, and with the desire to know the unknown.

Information Demand (Lundqvist et al. 2004)
Information Demand is the constantly changing need for current, accurate, and integrated information to support (business) activities, when ever and where ever it is needed.

Information Demand Modeling aims at capturing and formalizing all information relevant for demand oriented information supply

Selected approaches for Demand Modeling
- User profile based
- Situation based
- Context based

User Profiles

Some Examples

- Roaming profiles in MS-Windows
- Cookies
- Profiles in mobile phones
- W3C's Composite Capability / Preference Profiles

Definition:

A structured set of data representing preferences of and information about a user in a machine understandable way

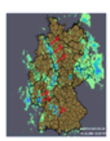
Typical characteristics

- User provides input for the profiles
- Capture the different perspectives of information demand
- Limits in covering dependencies between perspectives

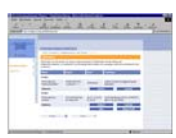
Example: Weather Information on demand

Example: Precise weather warning based on user profiles and radar prognosis

Storm



Forecast



Logistic

User

Information flow

Situation

Situation
- time interval and action occurring in this time interval
  Example: taxi ride, lecture, video conference, lunch
- Action is described by topics / set of concept paths

Approach primarily applied for Message Rating
- message: free-text content (natural language) in a known data format
  Example: e-mail, SMS

Different Manifestations of Relevance

System or algorithmic relevance:
Relation between a query and information objects in a system retrieved by a given algorithm

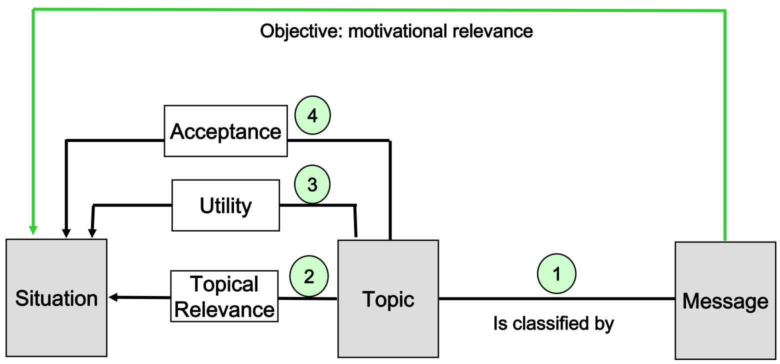
Topical or subject relevance:
relation between subject or topic expressed in a query, and topics or subjects covered by the retrieved information object

Cognitive relevance or pertinence:
relation between the cognitive information need of a user and the information objects retrieved

Situational relevance or utility:
relation between situation, task or problem at hand and information objects retrieved

Motivational or affective relevance:
relation between intents, goals, and motivations of a user information objects retrieved

Different Relevance Relationships

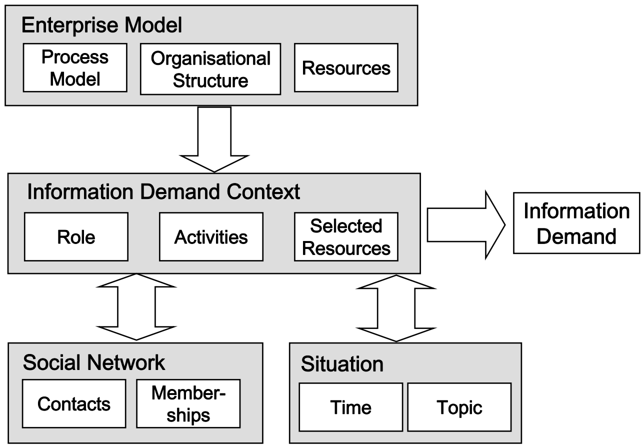


Context

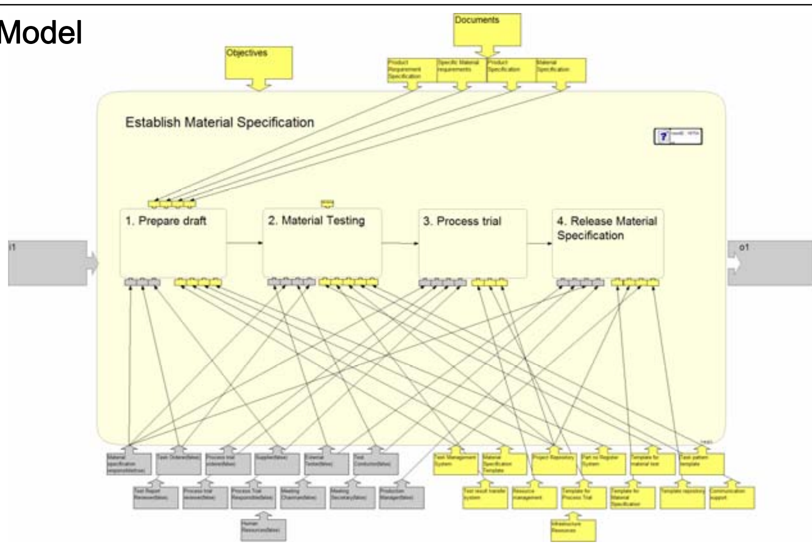
One of many definitions (Dey, 2001)
Context is any information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves.

A system is context-aware if it uses context to provide relevant information and/or services to the user, where relevancy depends on the user's task.

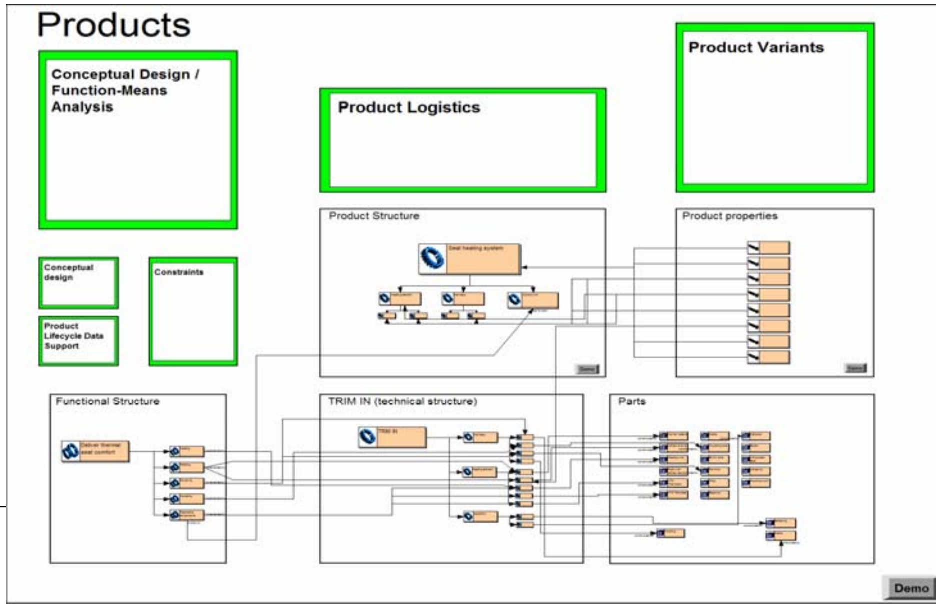
Context-based Information Demand



Example Model



Example Model (2)



But did we really capture everything?

Fraunhofer Institut Software- und Systemtechnik

JÖNKÖPING UNIVERSITY

Source: Thomas Norrby
Creative consultant www.vioni.se

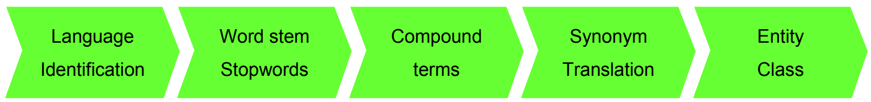
Evaluating Content

How to decide whether information is relevant?

Different information sources and communication media

Different aspects
- Metadata
- Structure
- Content
- Source / location
- Recommendations
- etc.

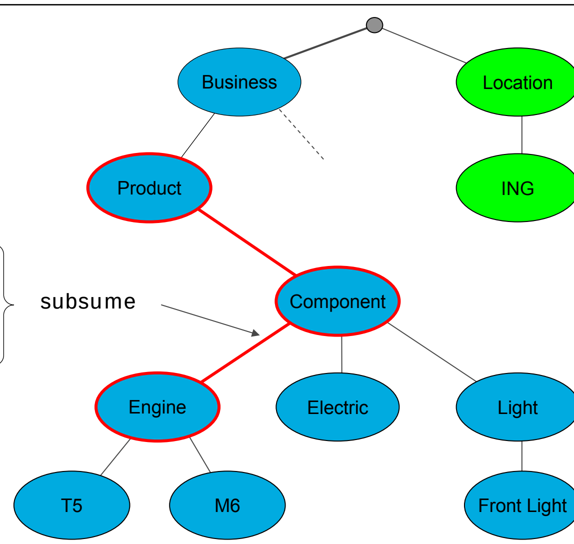
Typical Approach from IR: Topic Extraction



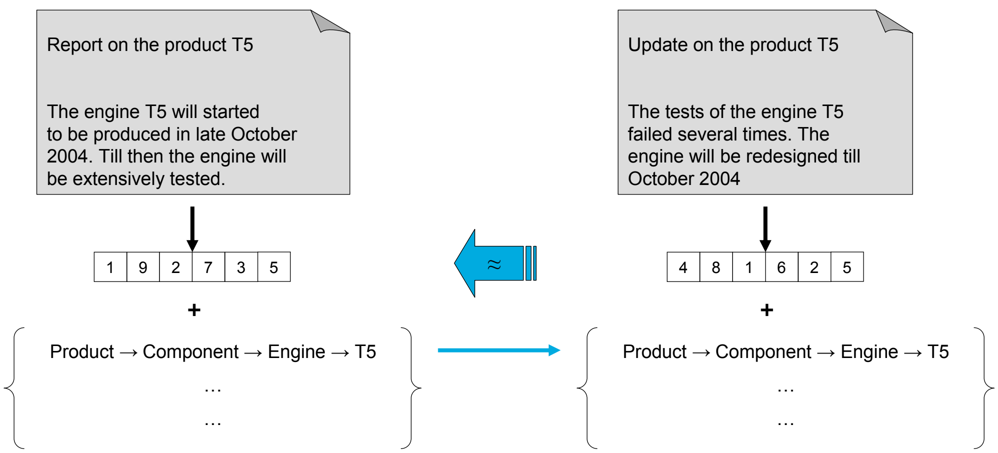
- Topics describe the content and used as features of a text
- Words occur in various forms, some words are irrelevant
- „Information Logistics“ is one term
- „car“, „auto“, „motorcar“, „bil“ means the same
- Location (City, Island), Personal name, Company name, Date, active ingredient

Semantic Path Sets

Product → Component → Engine
Events → Meeting → Board Meeting
Location → ISST → Berlin → R6.30



Semantic Similarity



Distance between Concepts in Ontologies

String Similarity:
- edit-distance like functions (e.g. Levenstein distance, Monger-Elkan distance, Jar-Winkler distance)
- token-based distance functions (e.g. Jaccard similarity, TFIDF or cosine similarity, Jense-Shannon distance).

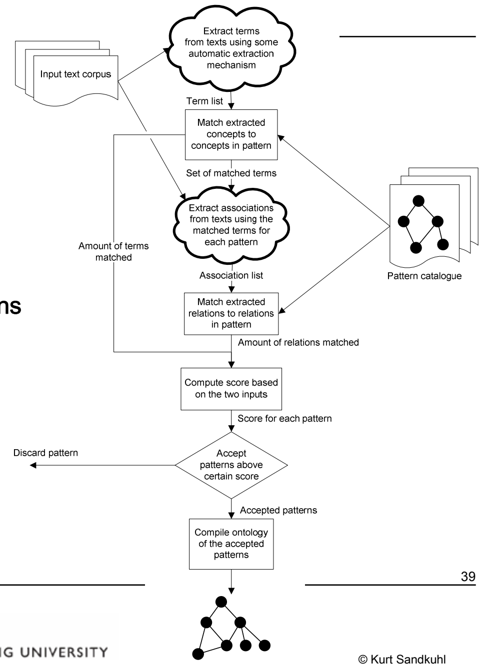
Synonyms
- detect the problem of using different terms in the ontologies for the same concept

Structure Similarity:
- usually based on is-a or part-of hierarchy of the ontology.
- Example: similarity flooding matching graphs to find corresponding nodes based on a fix-point computation.

Based on instances:
- GLUE uses multiple machine learners and is exploiting information in concept instances and taxonomic structure of ontologies.
- FCA-Merge: Formal Concept Analysis to produce a lattice of concepts which relates concepts from the source ontologies.

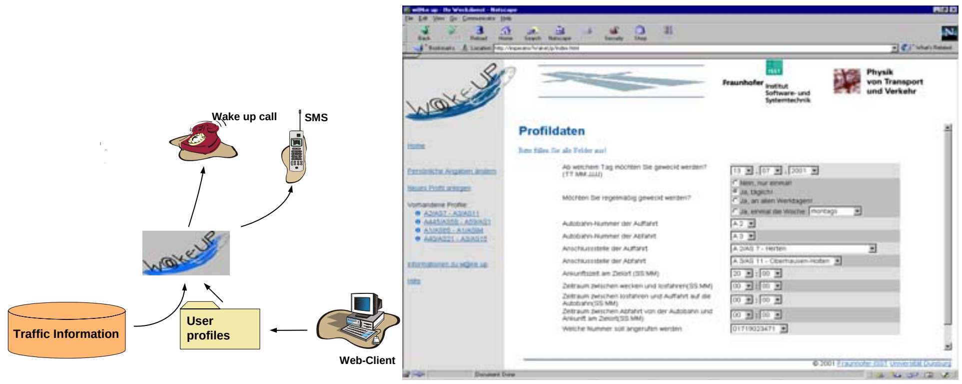
Automatic ontology construction

- Starting-point
  - Use existing tools
  - Fully automatic method
- Input
  - Text corpus
    => Extracted terms and relations
  - Pattern catalogue
- Output
  - Finished ontology
- Process
  - Match terms and relations
  - Accept or reject patterns
  - Build ontology



Application Examples

Application Example: Traffic Information / Wake-up Service



Weather Information on Demand (WIND)

September 2000	First prototype for demonstrations
May 2002	Pilot project with an insurance company (5.000 users)
July 2003	Start as commercial service for 16 insurance companies throughout Germany (50.000 users)
July 2004	More than 100.000 users
July 2005	Nearly 500.000 users

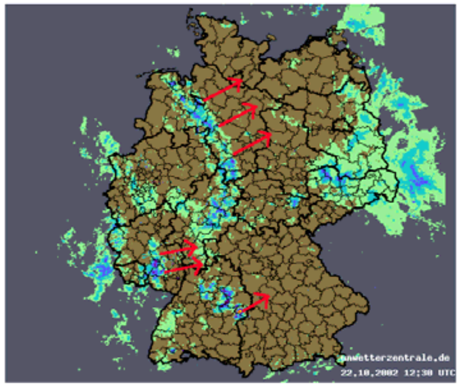
Partners

Versicherungskammer Bayern

meteoMedia AG

Fraunhofer ISST

Forecast: Radar picture and prognosis every 15 minutes



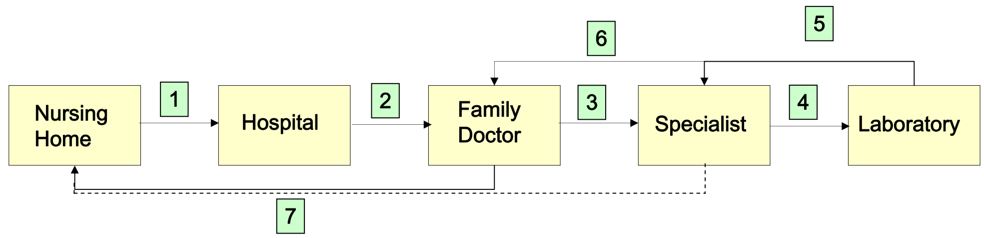
Information Demand Patterns

Industrial Partners:
- Kongsberg Automotive, Motala
- Proton Group, Värnamo

Work area: engineering change management
Objective: develop infomation demand patterns

Application Example: Cross-Institutional Communication in Heathcare

Example: Patient from Nursing home with leg fracture and diabetis



(1) Leg Fracture
(2) Release from Hospital
(3) Referral to Specialist
(4) Consultation with Laboratory
(5) Result of Laboratory
(6) Therapy Recommendation
(7) Coordination of Care Process

Summary

Information Logistics

Challenge: Not any information to any person any time, but only the right information for a demand

Contemporary technologies (information systems, search engines, knowledge management, semantic technologies) are important contributions but don't fully solve the problem

Change the perspective on information supply:
• Start from demand, actively support users
• Important for decision and problem solving processes

Ongoing Research

Content

Efficient development of ontologies
• Methodology for ontology development
• Automatic Ontology Development based on Patterns

Integration of heterogeneous information sources
• On-demand ontology integration based on P2P approaches

Ontology Evolution

Demand

Information Demand Modelling
• Methodology for modeling information demand

Information Demand Patterns

Ontology Matching based on Polygons

Fraunhofer Institut Software- und Systemtechnik JÖNKÖPING UNIVERSITY

© Kurt Sandkuhl

Thank you for your attention!

Questions?

More information about Information Logistics:

- www.hj.se/cenit
- infoeng.hj.se
- www.informationslogistik.se
- www.informationslogistik.org
- www.isst.fraunhofer.de

Thanks to all colleagues from Jönköping University and Fraunhofer ISST (Berlin and Dortmund) for contributing.