

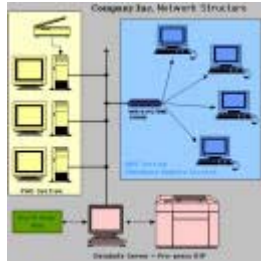
Workflow Orchestration and Mining for Integrated Asset Management in Smart Oilfields

Presenter : Fan Sun
Tao Zhu
Yinglong Xia
Muhammad Murtaza

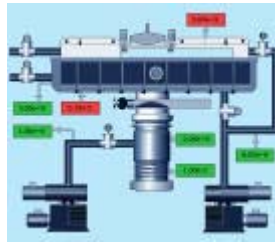
- Introduction to CiSoft
- Overview of Integrated Asset Management (IAM)
- Motivation for Workflow Mining
- Workflow Mining
- Event Logging
- Ontology Mapping

- Center for **I**nteractive **S**mart **O**ilfield **T**echnologies
- Research areas include:
 - **Integrated Asset Management**
 - Well Productivity Improvement
 - Robotics and Artificial Intelligence
 - Embedded and Networked Systems
 - Reservoir Management
 - Data Management Tools
 - Immersive Visualization

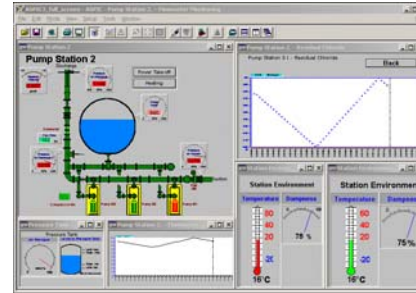
Integrated Asset Management (IAM)



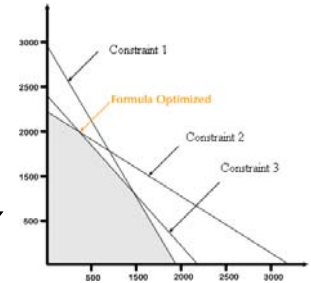
Historic Data, Databases connected over networks



Automation and Control

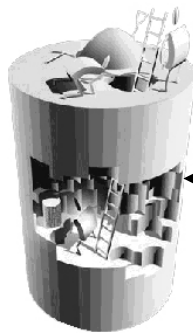


Visualization



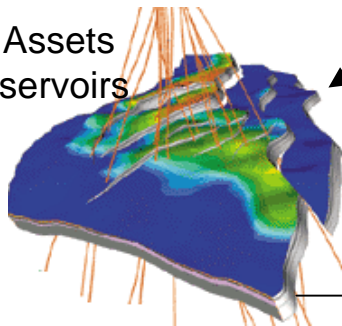
Optimizers
Eg: Fuzzy Logic

Data Mining Techniques

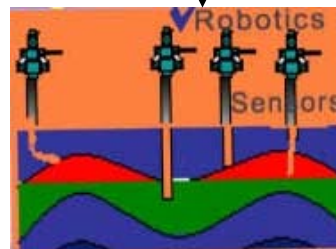


Integrated Asset Management

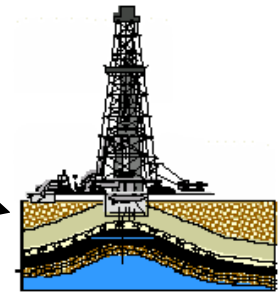
Physical Assets
Wells, Reservoirs



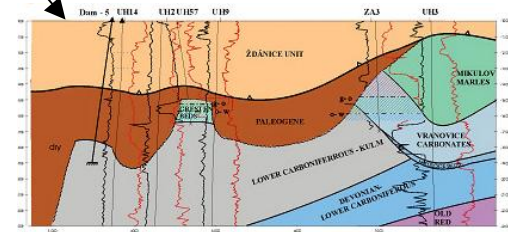
Real - time Data



Wireless Sensors

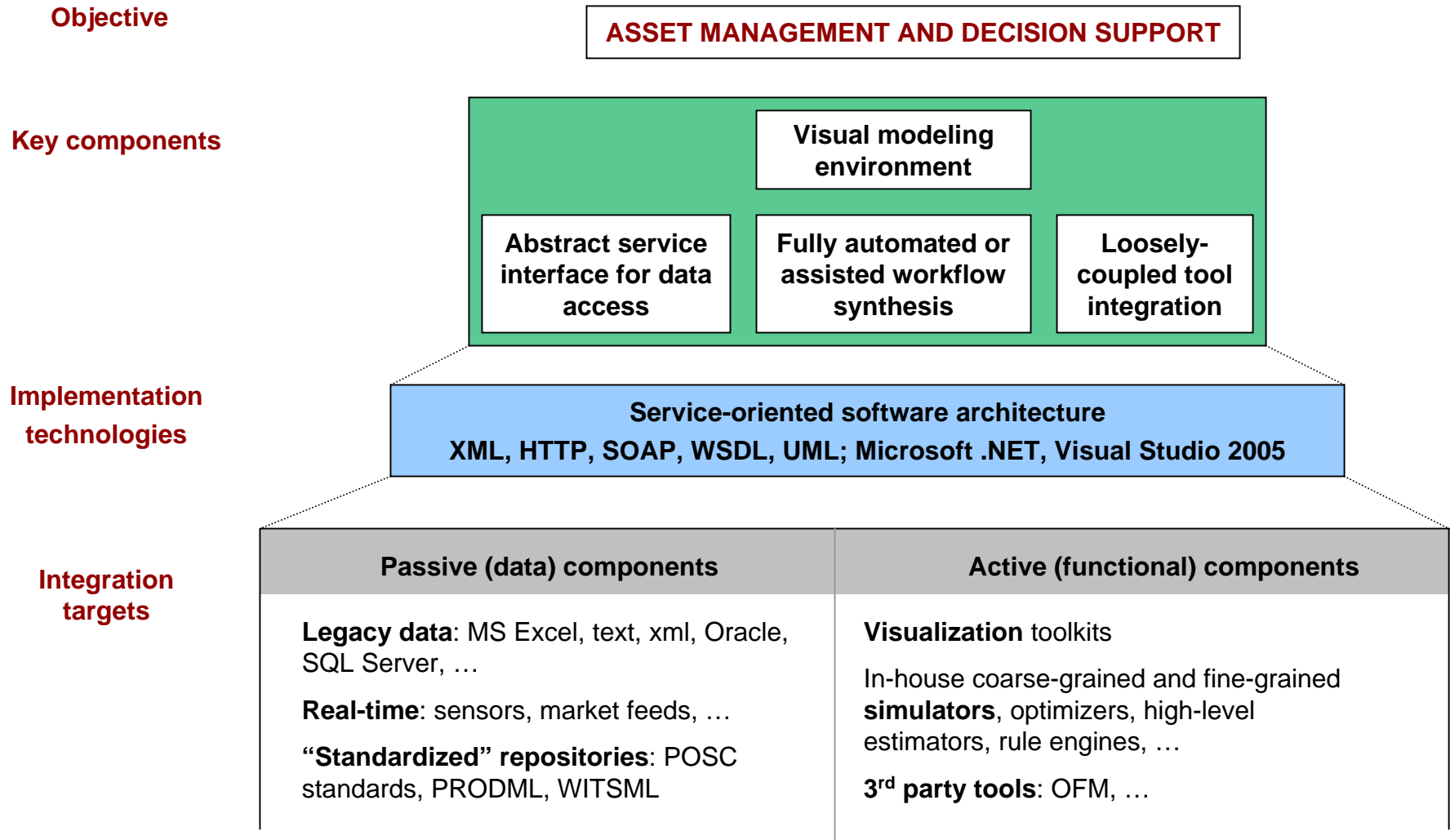


Well Simulator



Surface Facility Simulator

- Managing oilfield assets involves
 - Continuous decision-making
 - Multiple interactions
- Asset management decisions require
 - Interactions among multiple domain experts
 - Coupling between multiple scientific and business applications
- IAM objective: Enable better and faster decision making
 - ❖ on-demand access to **information** from a wide variety of sources
 - ❖ automate repetitive tasks and improve **productivity**
 - ❖ enable **what-if scenario** analysis
 - ❖ facilitate **collaboration** between groups and applications



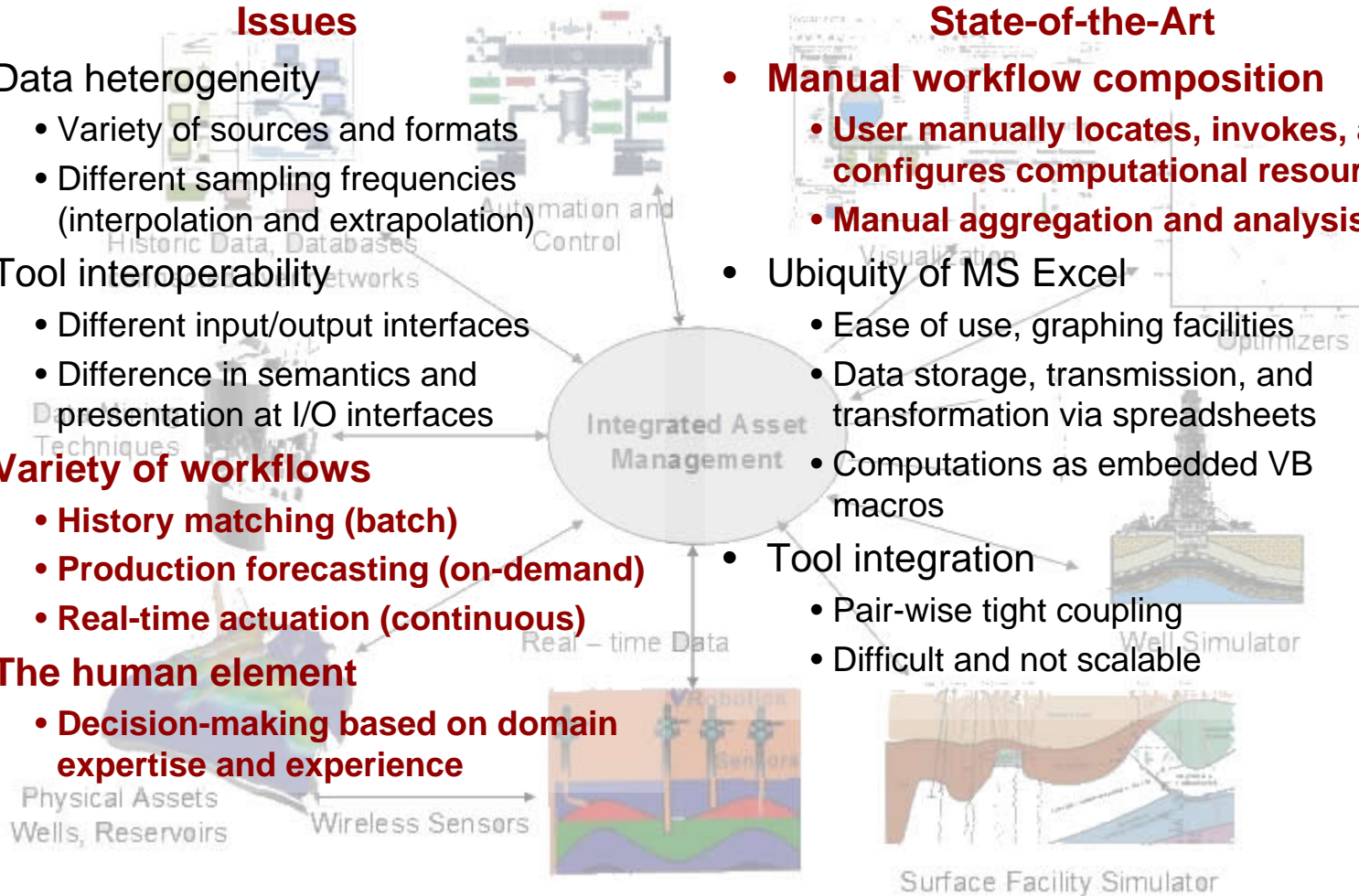
Challenges in IAM

Issues

- Data heterogeneity
 - Variety of sources and formats
 - Different sampling frequencies (interpolation and extrapolation)
- Tool interoperability
 - Different input/output interfaces
 - Difference in semantics and presentation at I/O interfaces
- **Variety of workflows**
 - **History matching (batch)**
 - **Production forecasting (on-demand)**
 - **Real-time actuation (continuous)**
- **The human element**
 - **Decision-making based on domain expertise and experience**

State-of-the-Art

- **Manual workflow composition**
 - **User manually locates, invokes, and configures computational resources**
 - **Manual aggregation and analysis**
- Ubiquity of MS Excel
 - Ease of use, graphing facilities
 - Data storage, transmission, and transformation via spreadsheets
 - Computations as embedded VB macros
- Tool integration
 - Pair-wise tight coupling
 - Difficult and not scalable



Why workflow mining in IAM?

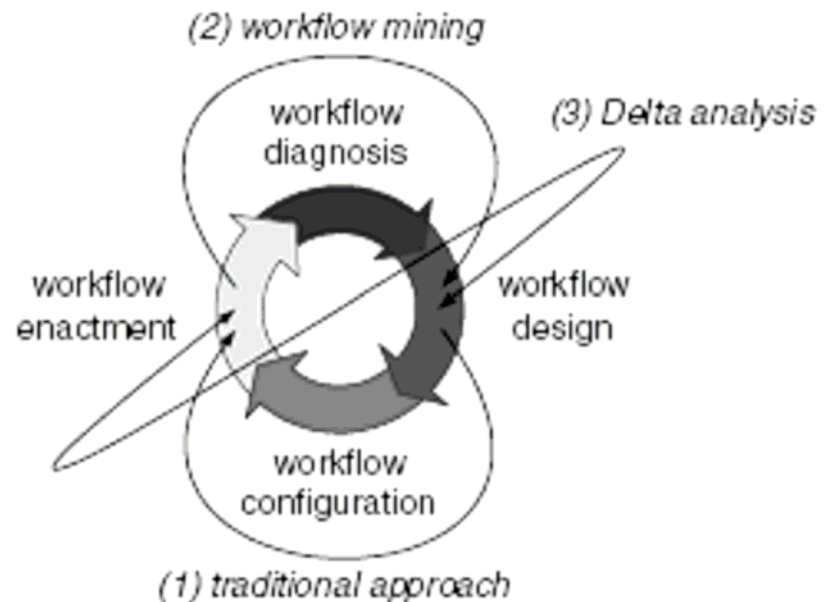
- Technology aspect
 - Continuous optimization at the asset level
 - Shared situational awareness for decision making
 - Similar efforts
 - Defense: Net-centric warfare, Joint Battlespace Infosphere
 - eBusiness: The zero-latency enterprise
- Human aspect
 - **2500**: Enrollment in U.S. petroleum engineering programs in 2004 – down from **12,000** in 1982
 - **60%**: Percentage of experienced managers expected to retire from the oil and gas industry by 2010
 - **49 years**: Average age of a petroleum engineer
- Goal: **To capture the domain knowledge**

- Many of today's information systems are driven by explicit process models.
- Workflow management systems are configured on the basis of a workflow model specifying the order in which tasks need to be executed.
- Workflow mining supports workflow design.
- Starting point for workflow mining is a so-called "workflow log" containing information about the workflow process as it is actually being executed.

Workflow Life Cycle

- The workflow life cycle consists of four phases:
 - (A) workflow design,
 - (B) workflow configuration,
 - (C) workflow enactment, and
 - (D) workflow diagnosis.

- The goal of workflow mining is to reverse the process.



- **The objective way of modeling is to use data related to the actual events that took place.**
- **Closely monitoring the events taking place at runtime also enables Delta analysis, i.e., detecting discrepancies between the design constructed in the design phase and the actual execution registered in the enactment phase.**
- **Workflow mining results in an “a posteriori” process model which can be compared with the “a priori” model.**
- **The goal of workflow mining is to extract information about processes from transaction logs.**
- **Assumption is that it is possible to record events such that (i) each event refers to a task (i.e., a well-defined step in the workflow), (ii) each event refers to a case (i.e., a workflow instance), and (iii) events are totally ordered.**
- **These workflow logs are used to construct a process specification which adequately models the behavior registered.**

Converting Staffware log to Workflow Log



Staffware log

Table 1

A Staffware log

Directive description	Event	User	yyyy/mm/dd hh:mm
<i>Case 10</i>	Start	bvdongen@staffw_e	2002/06/19 12:58
Register	Processed to	bvdongen@staffw_e	2002/06/19 12:58
Register	Released by	bvdongen@staffw_e	2002/06/19 12:58
Send questionnaire	Processed to	bvdongen@staffw_e	2002/06/19 12:58
Evaluate	Processed to	bvdongen@staffw_e	2002/06/19 12:58
Send questionnaire	Released by	bvdongen@staffw_e	2002/06/19 13:00
Receive questionnaire	Processed to	bvdongen@staffw_e	2002/06/19 13:00
Receive questionnaire	Released by	bvdongen@staffw_e	2002/06/19 13:00
Evaluate	Released by	bvdongen@staffw_e	2002/06/19 13:00
Archive	Processed to	bvdongen@staffw_e	2002/06/19 13:00
Archive	Released by	bvdongen@staffw_e	2002/06/19 13:00
	Terminated		2002/06/19 13:00

Workflow Log

Table 2

A workflow log

Case identifier	Task identifier
Case 1	Task A
Case 2	Task A
Case 3	Task A
Case 3	Task B
Case 1	Task B
Case 1	Task C
Case 2	Task C
Case 4	Task A
Case 2	Task B
Case 2	Task D
Case 5	Task A
Case 4	Task C
Case 1	Task D

- Workflow logs will typically contain noise, i.e., parts of the log can be incorrect, incomplete, or refer to exceptions.
- if the model exhibits alternative and parallel routing, then the workflow log will typically not contain all possible combinations.
- workflow logs can be used to systematically measure the performance of employees. The legislation with respect to issues such as privacy and protection of personal data differs from country to country.

Workflow logs: Format

- A common XML format

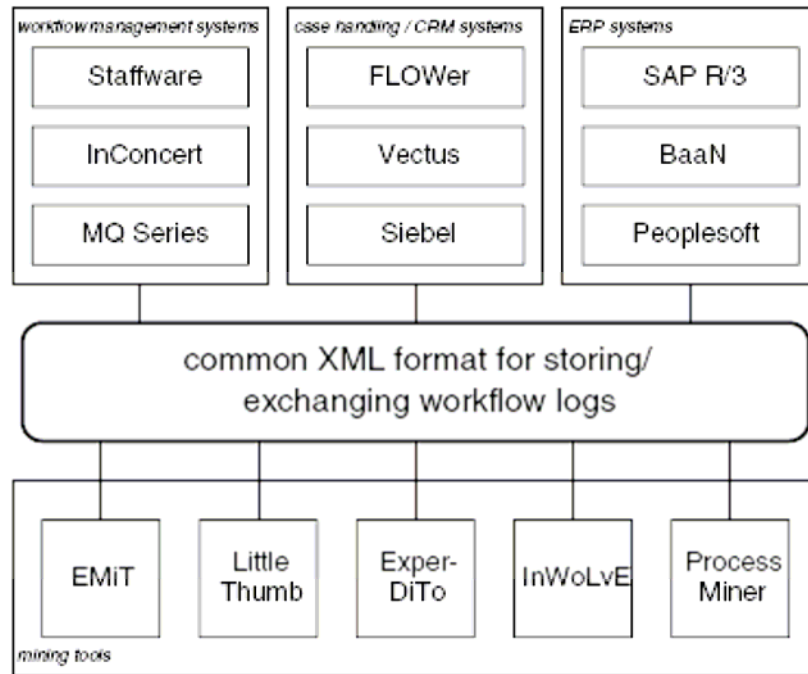
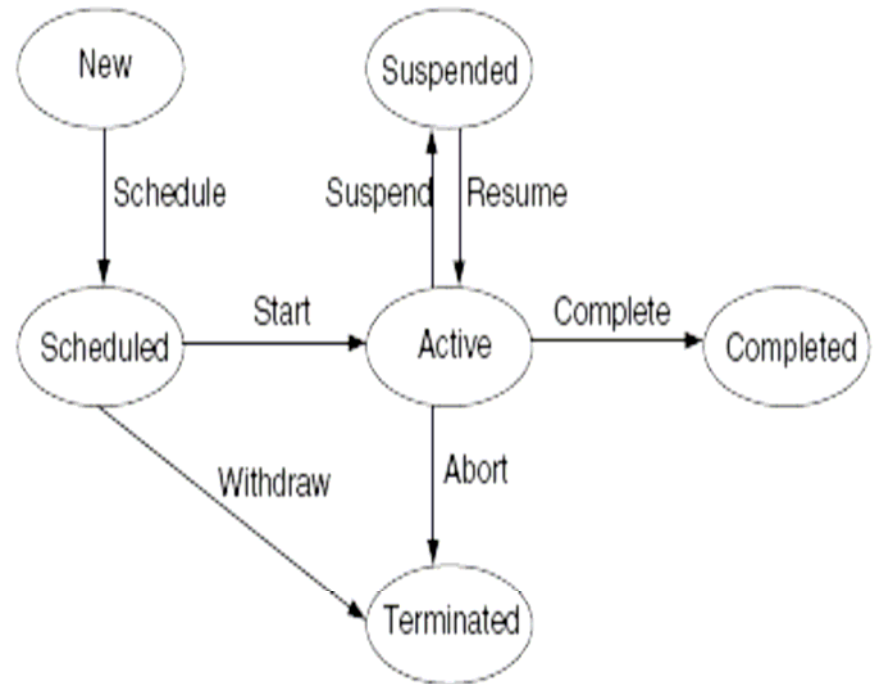


Fig. 4. The XML format as the solver/system independent medium.

Events and Types

- **There are eight types of events:-**
 - normal,
 - schedule,
 - start,
 - withdraw,
 - suspend,
 - resume,
 - abort, and
 - complete.

- **Arrows show the possible transitions as atomic events.**



- Heuristic approaches
- Three mining steps:
 - Step (i) the construction of a dependency/frequency table (D/F-table),
 - Extracting from logs:-
 - overall frequency of task ‘a’,
 - the frequency of task a directly preceded by task b,
 - the frequency of a directly followed by task b,
 - the frequency of a directly or indirectly preceded by task b but before the previous appearance of b,
 - the frequency of a directly or indirectly followed by task b but before the next appearance of a,
 - a metric that indicates the strength of the causal relation between task a and another task b.

- Heuristic approaches.....

- Step (ii) the mining of the basic relations out of the D/F-table (the mining of the R-table),

- we can determine the basic relations ($a \rightarrow_w b$, $a \#_w b$, and $a ||_w b$) out of the D/F-table.

IF $((\#A \rightarrow B \geq N) \text{ AND } (\#A > B \geq \theta) \text{ AND } (\#B < A \leq \theta))$ THEN $a \rightarrow_w b$ |

- where N is noise in log,
- Threshold of N for induction process.

$$\theta = 1 + (\text{Round}(N \times \#L) / \#T)$$

- #L is the number of trace lines in the workflow log, and #T is the number of elements (different tasks).

- Step (iii) the reconstruction of the WF-net out of the R-table,

- Use of Alpha algorithm as in formal approach.

Comparison and open problems

- Tools such as EMiT, Little Thumb, InWoLvE, and Process Miner are driven by different problems.
- EMiT shows which class of workflow processes can be rediscovered.
- Little Thumb to show how heuristics can be used to tackle noise problems.
- Concept tools in InWoLvE deal with duplicate tasks.
- Process-Miner exploiting the properties of block-structured workflows through rewriting rules.

Comparing EMiT, Little Thumb, InWoLvE, and Process Miner

	EMiT	Little Thumb	InWoLvE	Process Miner
Structure	Graph	Graph	Graph	Block
Time	Yes	No	No	No
Basic parallelism	Yes	Yes	Yes	Yes
Non-free choice	No	No	No	No
Basic loops	Yes	Yes	Yes	Yes
Arbitrary loops	Yes	Yes	No	No
Hidden tasks	No	No	No	No
Duplicate tasks	No	No	Yes	No
Noise	No	Yes	Yes	No

- Human Involved workflow
- Data driven
- Support fast decision making

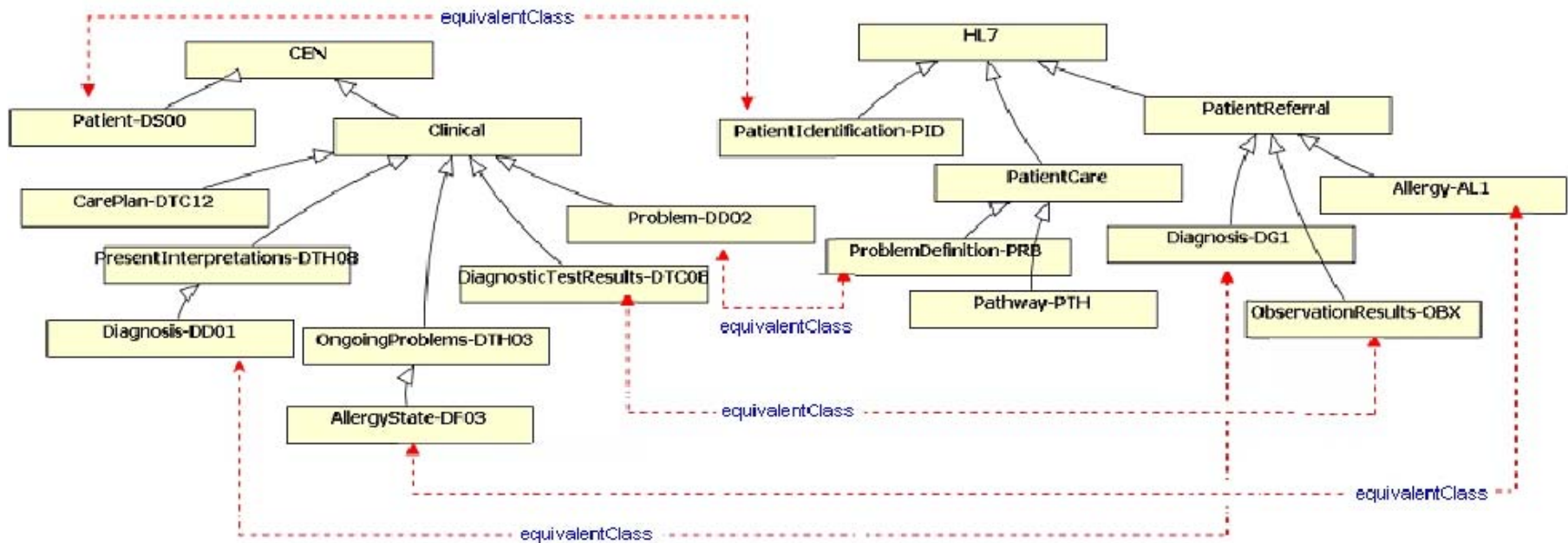
- Event modeling
- Ontology mapping
- Workflow mining using semantic web technologies

- In order to two parties to understand each other, they should use the same **formal representation** for the **shared conceptualization** (so the same ontology)
- Unfortunately it is not easy to make everybody to agree on the same ontology for a domain
- And when you have different ontologies for the same domain the problem shows up.
 - Parties with different ontologies do not understand each other.

Here comes the ontology mapping into the play

Ontology Mapping

- Ontology Mapping is the process whereby two ontologies are semantically related at conceptual level, and the source ontology instances are transformed into the target ontology entities according to those **semantic relations**.



- Three dimensions of ontology mapping:
 - **Discovery**: manually, automatically or semi-automatically defining the relations between ontologies
 - **Representation**: A language to represent the relations between the ontologies
 - **Execution**: Changing instance of a source ontology to an instance of target ontology

- “A **MA**pping **FR**amework for Distributed Ontologies”, developed at Univ. Karlsruhe
- One of the main contributions is the definition of “**Semantic Bridges**” (SB) between ontologies which establishes correspondences between entities from source and target ontology.
- Defines “**Semantic Bridge Ontology**” which is an ontology of mapping constructs.
- Includes functionality for all of the three dimensions of ontology mapping (discovery, representation, execution)

MAFRA Conceptual Architecture

