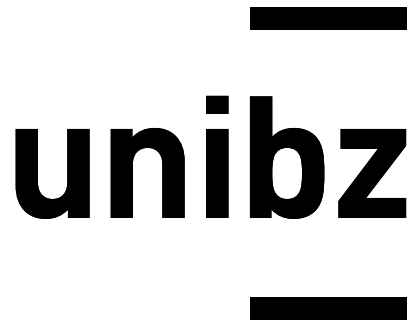
Data and Process Modeling

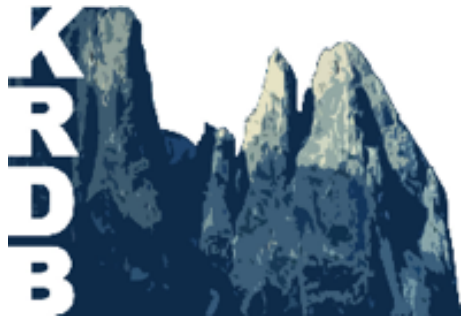
0.Overview

Marco Montali

KRDB Research Centre for Knowledge and Data
Faculty of Computer Science
Free University of Bozen-Bolzano

A.Y. 2014/2015





Basic Info

- Personal Info:
  - e-mail: `montali@inf.unibz.it`
  - home-page: `http://www.inf.unibz.it/~montali/`
  - office hours: by previous arrangement via e-mail.

- Web page of the course:

  `http://www.inf.unibz.it/~montali/dpm/1415`

  - Check it constantly!

Our Starting Point

Models and Reality

A model is an abstraction of reality according to a certain conceptualization. Once represented as a concrete artifact, a model can support communication, learning and analysis about relevant aspects of the underlying domain. [...] a represented model (a *dusty diagram*) created by an unknown predecessor is a medium to preserve and communicate a certain view of the world, and can serve as a vehicle for reasoning and problem solving, and for acquiring new knowledge (maybe having striking new *ideas*!) about this view of the world.
(Guizzardi, 2005)

Conceptual Modelling

The activity of formally describing some aspects of the physical and social world around us for the purposes of understanding and communication.
(Mylopoulos, 1992)

What this Course is About

Goal

Introduce, study and put into practice languages, methodologies and techniques focused on conceptual modeling for information systems.

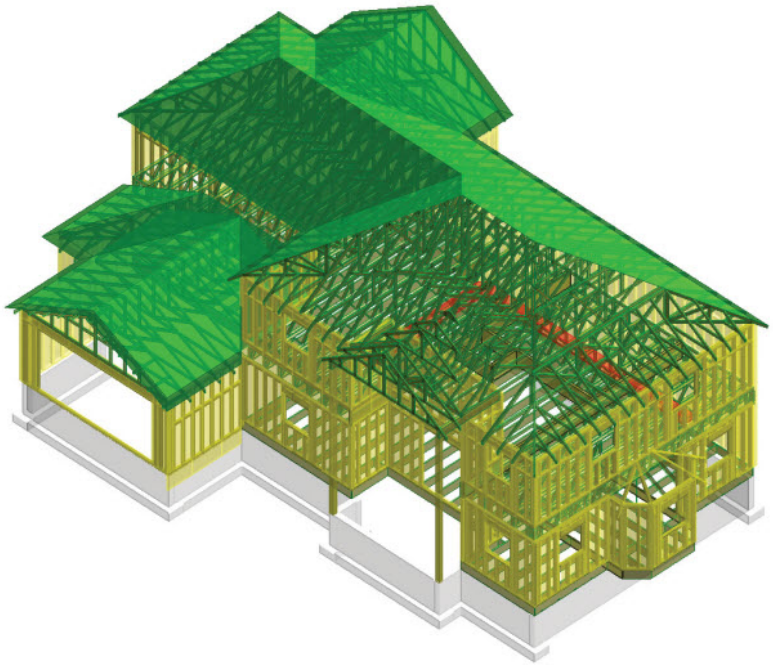
Information System

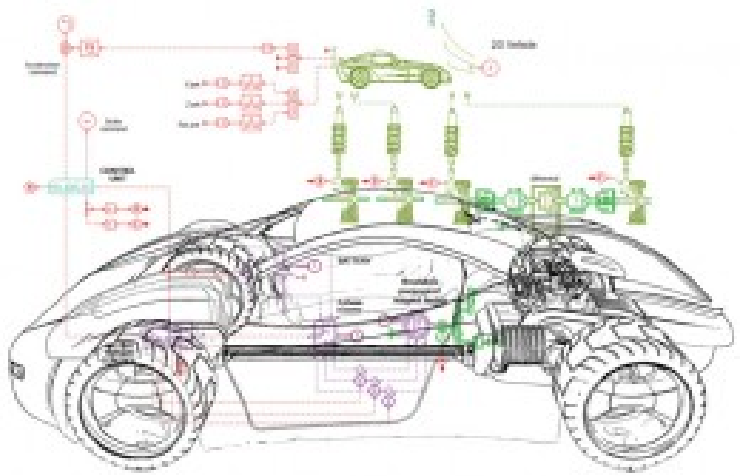
A system that collects, stores, processes, and distributes information about the state of a domain to facilitate planning, control, coordination, and decision making in an organization.

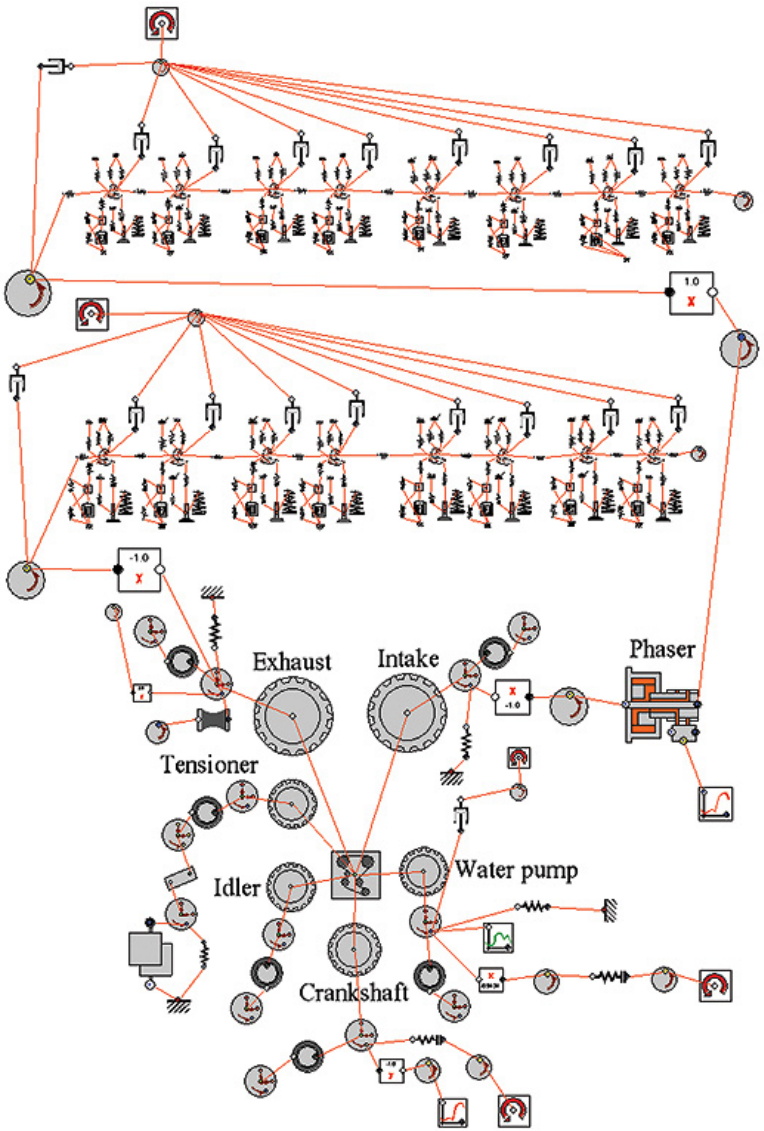
We are interested in conceptual models for:

- understanding, documentation and communication among humans;
- design and engineering of information systems to support humans in conducting their business and everyday work.

Models in Industrial Engineering







Architectural Plans

What do we Model? How?

Example

Brian wants to travel to Sydney. He decides to call a cab to get from his apartment, located in Threepwood Street 14, to the airport in Frankfurt. During the call, he is asked to provide his own name, and the address at which he wants to be picked up. The cab management system also tracks the timestamp of the call and Brian's telephone number. After 10 minutes, the cab arrives at his apartment. Then, the cab needs 30 mins for the 20 km to terminal C.
Once arrived, Brian checks in using the Self-Check-In terminal, by typing his passport ID and his ticket number (in alternative, he could also have provided the combination of flight number and booking code). Of course, he could have also checked in at the Check-In counter. He mistypes the ticket number once, so he receives back an error message. Then he retypes the correct ticket number, and gets back his boarding pass, which lists the departure gateway and his seat number.

What do we Model? How?

Example

Since he has no baggage to check in, he goes directly to the security screening that is located around 100 m left of the departure hall. Luckily, the queue is not very long. Already after 5 minutes he can go to the gate. Instead of relaxing in the frequent flyer lounge, he checks out the duty-free shops and buys a newspaper. To do so, he is requested to provide his credit card (or, alternatively, cash), together with the boarding pass. 15 minutes later, he goes back to the gate, waiting until boarding time. Another 10 minutes later, he finally boards the airplane, by showing his boarding pass and passport to the boarding officer of the airline company.

1. Read the example carefully.
2. Try to contrast the particular history of Brian, with the general case.
3. Try to understand the different relevant aspects to be captured.
4. Depict the (general) scenario graphically.

Obviously, in a real situation we should first ask ourselves: **why**?

What Do We Model?

Data

Facts accounting for the (important aspects of) a particular state of the domain under study.

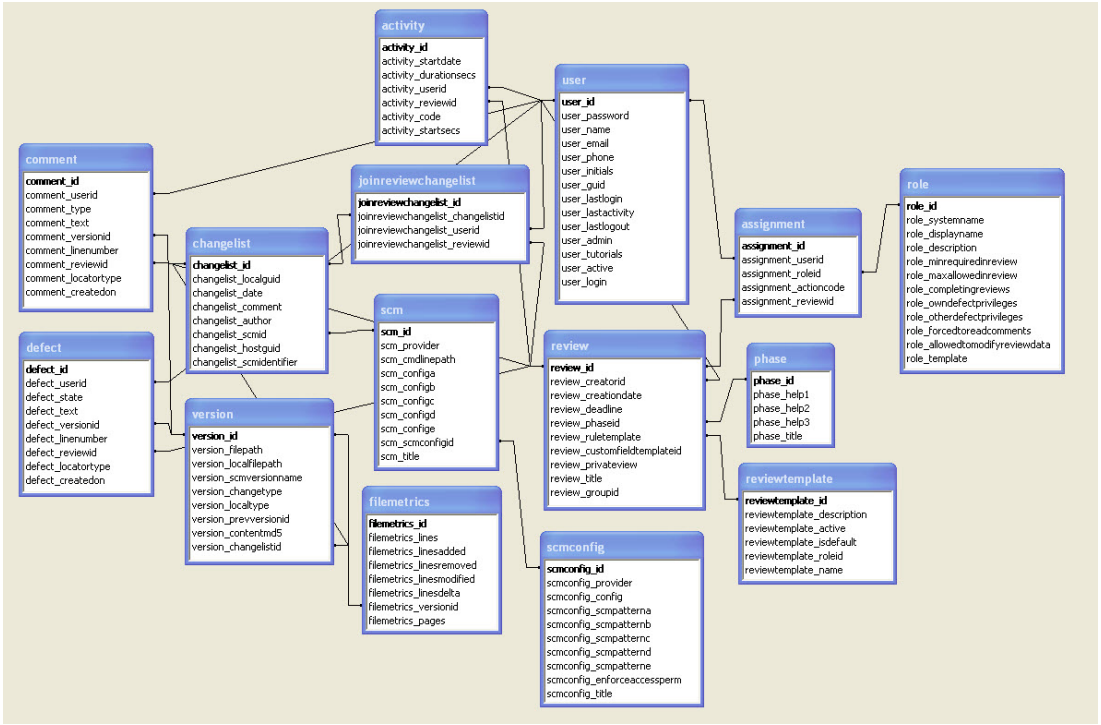
Processes

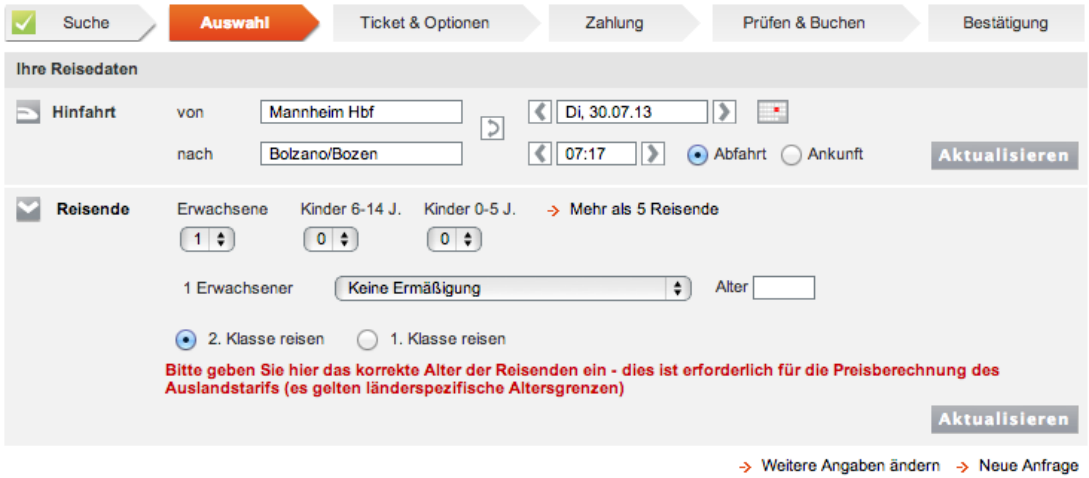
Collections of tasks describing how to structure the work in an organization, so as to achieve its strategic goals and produce value to its customer(s).

Data provide the basis for process execution.

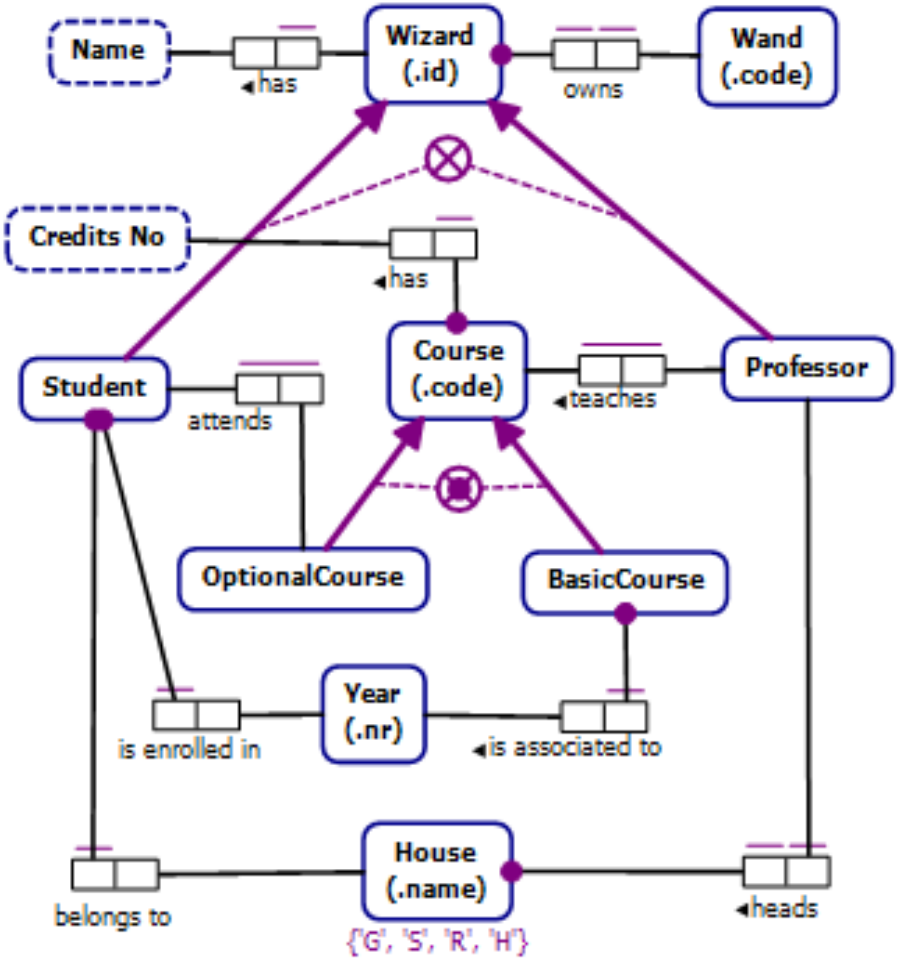
Processes use and manipulate data over time.

. . . for Information Systems

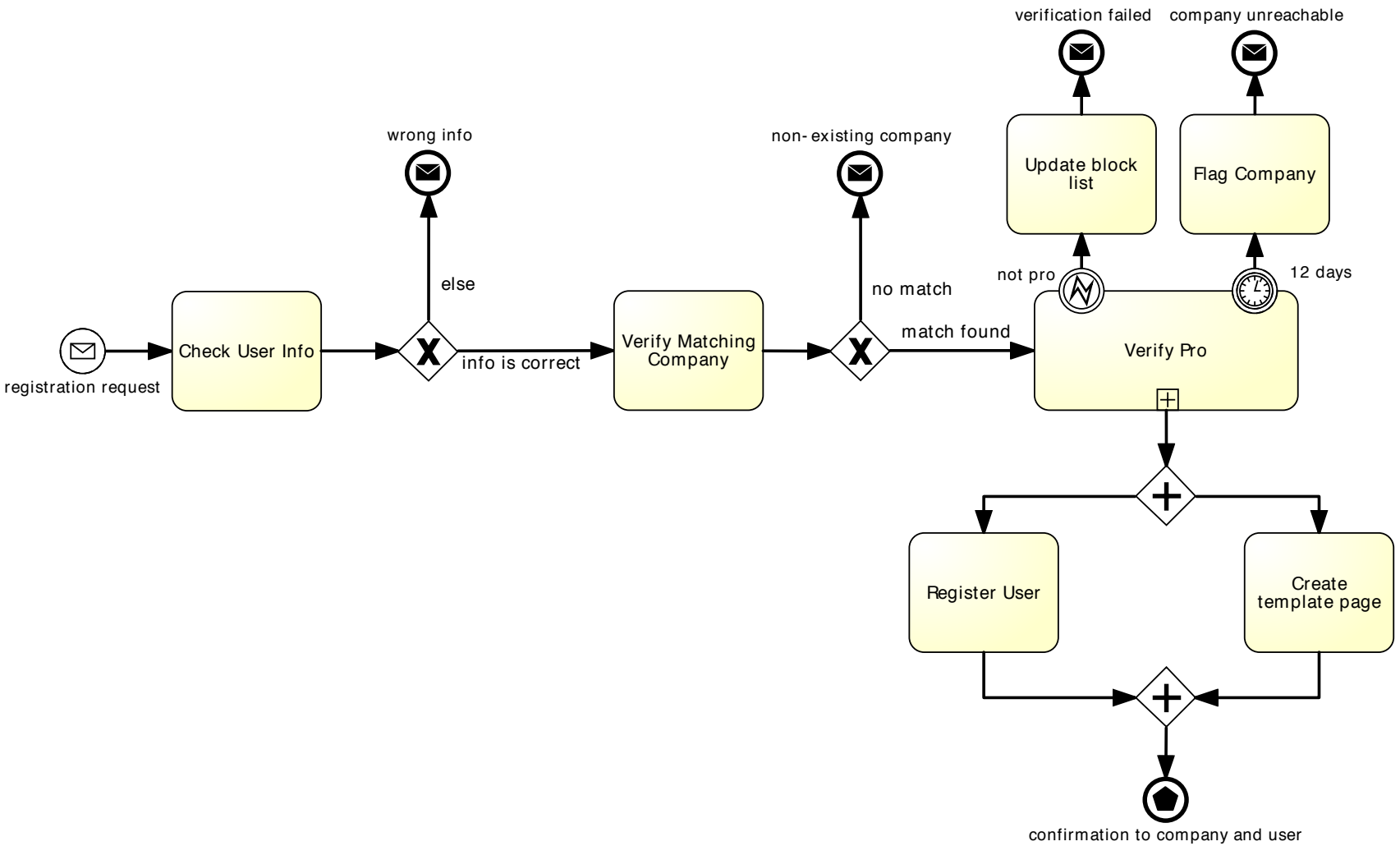




Data Modelling



Process Modelling



- Part A: Data Modelling [$\sim$5 weeks].
- Part B: Process modelling [$\sim$5 weeks].
- Part C: Combining data and process modelling [1 week]

- Part A: Data Modelling [$\sim$5 weeks].
  1. A fact-oriented approach: Object-Role-Modeling (ORM) notation.
  2. CSDP methodology: from unstructured information sources to an ORM schema.
  3. Relational mapping: from an ORM conceptual schema to a logical data schema.
  4. Quality of conceptual schemas, and insights from formal ontology.
- Part B: Process modelling [$\sim$5 weeks].
- Part C: Combining data and process modelling [1 week]

- Part A: Data Modelling [$\sim$5 weeks].
- Part B: Process modelling [$\sim$5 weeks].
  - Business Process Management.
  - Control-flow patterns.
  - Modelling processes with the OMG standard Business Process Model and Notation (BPMN).
  - Formal foundations of process modelling and analysis: transition systems and Petri nets.
  - Process mining: extracting process models and other useful insights from event logs.
- Part C: Combining data and process modelling [1 week]

- Part A: Data Modelling [$\sim$5 weeks].
- Part B: Process modelling [$\sim$5 weeks].
- Part C: Combining data and process modelling [1 week]
  - Case management, business artifacts.
  - Foundational challenges.

Labs

Two main goals:

1. Experiment the studied techniques.
2. Bridge the gap between modelling and software development.
   - Seeing (refined) conceptual models as live software artifacts!

What are we going to experiment with:

- NORMA for Visual Studio to model data and do relational mapping.
- Object-relational mapping with Hibernate.
- Oryx Signavio to model and analyse BPMN processes.
- Process engines for running executable business processes.
- Process mining tools.

Material

- Slides!!!
- Web-site (see also the linked papers).
- Main books:
  Data Modeling
  - Halpin, T., and Morgan, T. Information Modeling and Relational Databases, 2nd Edition. Morgan Kaufmann Publishers, Elsevier, 2008.

  Process Modeling
  - Weske, M. Business Process Management: Concepts, Languages, Architectures. Springer, 2007.
  - Dumas, M., La Rosa, M., Mendling, J., and Reijers, H. A. Fundamentals of Business Process Management. Springer 2013.
- Other books:
  - Olivé, A. Conceptual Modeling of Information Systems. Springer, 2007.
  - van der Aalst, W. M. P. Process Mining *Discovery, Conformance and Enhancement of Business Processes*. Springer 2011.

Exam

- Written exam (mark: 0-25 points).
  - Questions on the entire course material.
  - Practical exercises on modelling and analysis of conceptual schemas.
  - We will have special classes dedicated to exercises.

- Project (mark: 0-7 points).
  - Application of languages and techniques seen in the course to a non-trivial, concrete case study.
  - To be discussed with the lecturer.

Final mark: sum of the two.

- If > 30: *cum laude*.