



# **Year and bachelor projects**

## Department of Software Engineering



# Topics on KSI in general

- ❖ **software development, software engineering applications, data science**
  - ❖ multimedia (search engines), similarity modelling
  - ❖ open data
  - ❖ graph and multi-modal databases (social networks)
  - ❖ recommender systems
  - ❖ desktop, web and mobile platforms
  - ❖ parallel and high-performance systems
- ❖ **Bioinformatics**
- ❖ **HW-related projects (GPU)**
  - ❖ custom GPU+CPU cluster
- ❖ **machine learning techniques** (deep networks, etc.)

# Search in video

- ❖ **Development of algorithms for more accurate and faster searching in MMDB**
  - ❖ interactive search of large video collections
  - ❖ image modeling, ranking, visualization
  - ❖ participation in international competitions
  - ❖ VideoHunter@MFF workshop

Contact: **Jakub Lokoč**

**Text querying  
(also temporal)**

**Liked images**

**Rescoring**

**Display types**

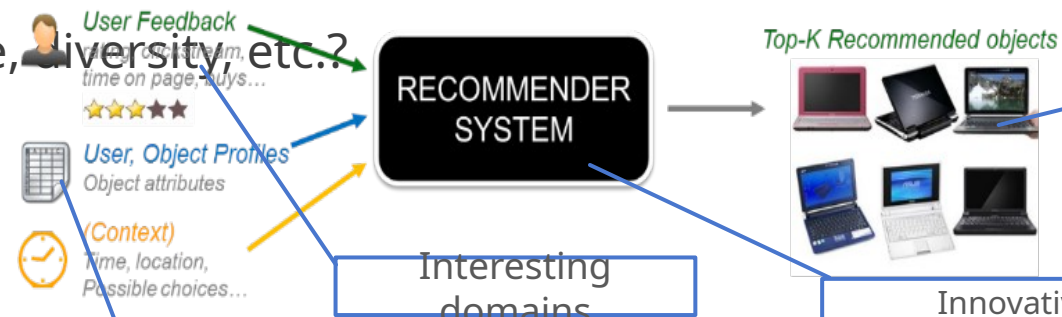
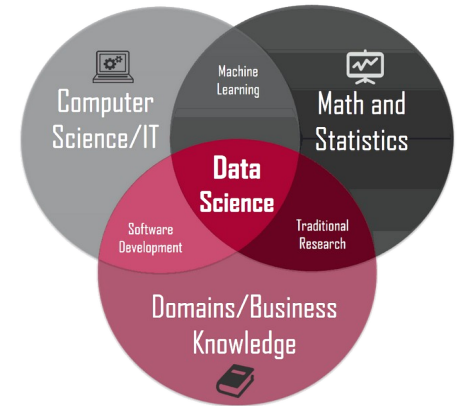
**History**

Query result in the selected display

The screenshot displays the SOM Hunter interface. On the left, there is a sidebar with several control panels: 'Text querying (also temporal)' with a search bar containing 'mountain' and a 'New search' button; 'Liked images' with a 'Likes:' section; 'Rescoring' with buttons for 'Rescore to Top N' and 'Rescore to SOM'; 'Display types' with 'SOM Display' selected and 'Top Scored' and 'Top Scored Context' options; and 'History' with an 'ACTIVE' button and a time range of '5:00:05 PM' to '5:00:11 PM'. Below these is a 'Bookmarks:' section. At the bottom left is the 'SOM HUNTER' logo featuring a Viking ship. The main area is a grid of 60 small video thumbnails, each showing a different mountain scene. Each thumbnail has a small green number in the top-left corner, representing a score or identifier. The grid is organized into 6 rows and 10 columns.

# Recommender systems

- ✓ Interested in data science?
- ✓ Do you want to create something practical?
- ✓ Do you like to combine knowledge from different disciplines?
  - ❖ Come explore/create recommender systems!
  - ❖ NSWI166 (Introduction, ZS 2/1) NDBI021 (Continuation, LS, 2/1)
- ❖ **Examples of year / bachelor / diploma projects**
  - ❖ Fairness and/or calibration in recommender systems
  - ❖ Using audio/video information to recommend films
  - ❖ Design of a domain-specific recommender (eco)system
  - ❖ How do people perceive relevance, diversity, etc.?
  - ❖ Winner of IT SPY 2023
- ❖ Contact **Ladislav Peška**
  - ❖ [www.ksi.mff.cuni.cz/~peska/projects.php](http://www.ksi.mff.cuni.cz/~peska/projects.php)



Comparative analyses  
Off-line vs. On-line  
Sequence-awareness  
Beyond Accuracy,  
Bias...

Innovative algorithms, proportional representation, long-term added value, group recommendation, package composition and other challenges

Integration of different data sources

# Bioinformatics applications and tools

## ❖ Web plugins

### ❖ Visualization of RNA structure

- ❖ To be used in RNAcentral
  - ❖ the world's largest repository of information on RNA molecules
- ❖ In collaboration with the European Bioinformatics Institute and Georgia State U

### ❖ Facelifting tools for protein active site prediction

- ❖ Used by hundreds of users per month

## ❖ Machine learning in bioinformatics

- ❖ Detection of active sites from protein sequences
- ❖ Classification of faulty layouts of RNA structures

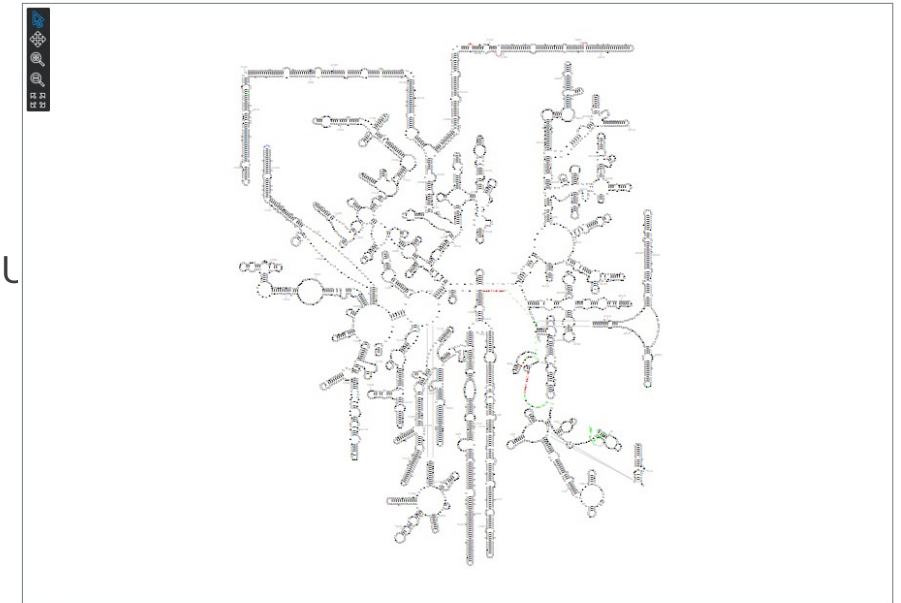
## ❖ Contact: David Hoksza

no prior knowledge of biology is assumed

### Secondary structure

Generated by RZDT using the *HS\_LSU\_3D* template provided by RiboVision. Learn more →

Toggle colours Toggle numbers Save PNG Save SVG Copy dot-bracket notation



#### Colour legend

- Same as the template
  - Modified compared to the template. Tip: Hover over green nucleotides for more details
  - Inserted nucleotides
  - Repositioned compared to the template
- Tip: Hover over the nucleotides to see nucleotide numbers



# Open data

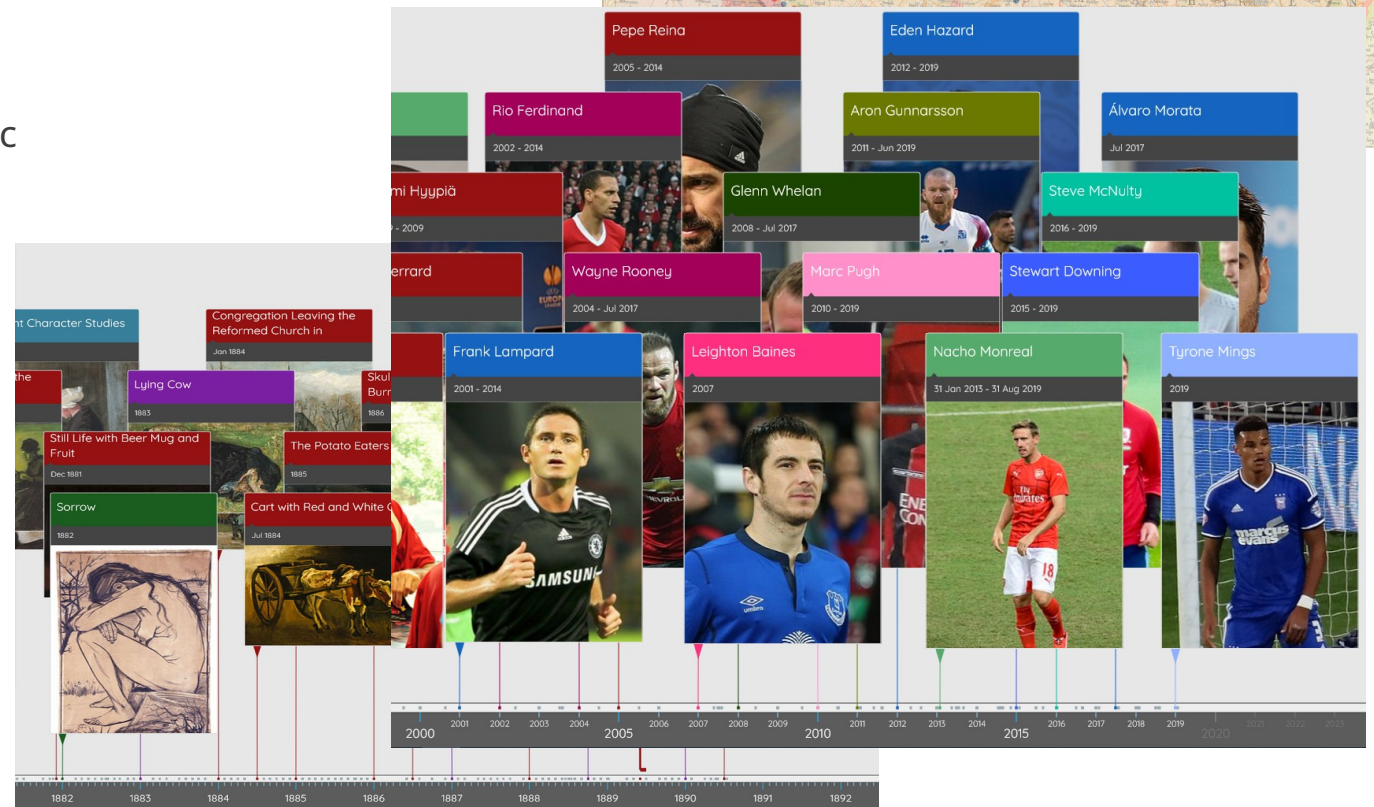
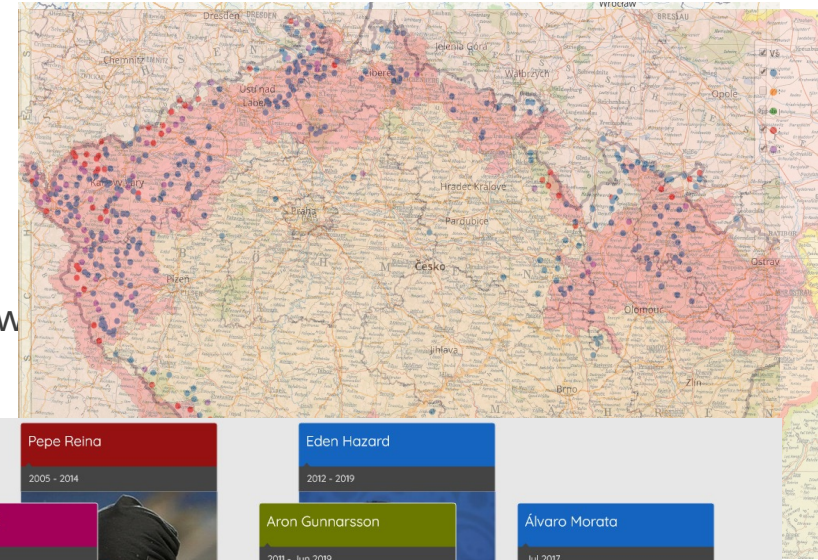
## ❖ open-source tools for working with open data

- ❖ visualisation of the history of settlement of the Czech Republic
- ❖ applications for teaching history, etc.
  - ❖ events of the world wars, the work of Vincent van Gogh, transfers of footballers between

## ❖ successful works will be used for

- ❖ <https://data.gov.cz>
  - ❖ official domain for open data of the Czech Republic
- ❖ <https://wikidata.org>
  - ❖ community database used in Wikipedia or Google

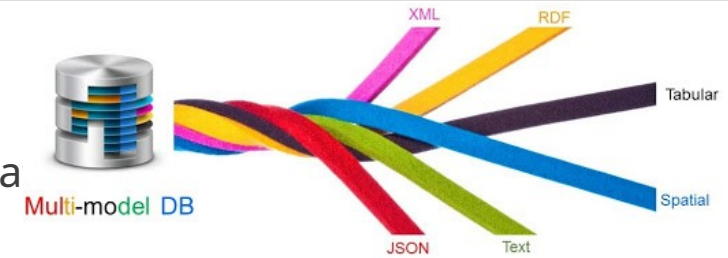
## ❖ Contact: **Martin Nečaský**



# Modern data models and database systems

## ❖ Popular data models

- ❖ Graph data (social networks), hierarchical data (JSON documents), multi-model data
- ❖ Multi-model data = represented in multiple linked data models
  - ❖ e.g. social network of customers (= graph), their orders (= relational data) and descriptions of ordered products (= JSON documents)



## ❖ Modern formats and databases

- ❖ Conceptual modelling and design of a multi-model database schema
- ❖ Visualization and querying over multi-model data
- ❖ Visualization and querying over graph data
- ❖ Comparative analysis of selected modern database systems

## ❖ The specific topic can be arranged according to student preferences

## ❖ Contact: Irena Holubová

