

# PRINCIPLES OF DATA ORGANISATION

Státnicové okruhy / State final exam



# STÁTNICOVÉ OTÁZKY / STATE FINAL EXAM

## 📌 Specializace: **Databáze a web**

### ✦ **2. Datový management**

Datové formáty. Datové modely pro strukturovaná data, příklady užití. Formáty pro grafová, hierarchická a tabulková data, datová schémata a jazyky pro transformaci dat. Formáty pro geodata. Základy grafických, multimediálních a tiskových formátů. Sémantický popis dat, slovníky. Procesy zpracování dat. Katalogizace dat, metadata. Základy šifrování a komprese dat. **Základy indexování. Typy organizace souborů, přímé/nepřímé indexování, primární/sekundární index. Hashování na vnější paměti. Hierarchické indexování. Indexování v prostorových databázích, prostorové spojení, prostorové dotazování.**

## 📌 Specialization: **Databases and Web**

### ✦ **2. Data Management**

Data formats. Data models for structured data, use-cases. Graph, hierarchical, tabular, and geodata data formats. Data schemas and data transformation languages. Basics of graphics, multimedia and print formats. Data vocabulary, data semantics. Data transformation, catalogization and metadata. Basics of data encryption and compression. **Basics of indexing. File organization techniques, direct/indirect indexing, primary/secondary index. Hashing in external memory. Hierarchical indexing. Indexing for spatial databases, spatial join, spatial query.**



# DETAIL

- ✂ Typy organizace souborů
  - ✂ Hromada, (ne)uspořádaný sekvenční soubor, index-sekvenční soubor, ...
- ✂ Základy indexování, přímé/nepřímé indexování, primární/sekundární index
  - ✂ Indexovaný soubor, bitmapové indexy, ...
- ✂ Hashování na vnější paměti
  - ✂ Cormack
  - ✂ Larson & Kalja
  - ✂ Fagin
  - ✂ Lineární hashování
- ✂ Hierarchické indexování
  - ✂ B strom
  - ✂ B+ strom
  - ✂ B\* strom
- ✂ Indexování v prostorových databázích, prostorové dotazování
  - ✂ Z křivka, Hilbertova křivka
  - ✂ Quad strom, k-d strom
  - ✂ R strom, R+ strom, R\* strom
- ✂ Prostorové spojení
  - ✂ Aproximace, filtrování
  - ✂ Vnitřní/vnější paměti
- ✂ File organization techniques
  - ✂ Heap file, (un)sorted sequential file, index-sequential file, ...
- ✂ Basics of indexing, direct/indirect indexing, primary/secondary index
  - ✂ Indexed file, bitmap index, ...
- ✂ Hashing in external memory
  - ✂ Cormack
  - ✂ Larson & Kalja
  - ✂ Fagin
  - ✂ Linear hashing
- ✂ Hierarchical indexing
  - ✂ B-Tree
  - ✂ B+-Tree
  - ✂ B\*-Tree
- ✂ Indexing for spatial databases, spatial query
  - ✂ Z-Curve, Hilbert Curve
  - ✂ Quad-Tree, k-d-Tree
  - ✂ R-Tree, R+-Tree, R\*-Tree
- ✂ Spatial join
  - ✂ Approximation, filtering
  - ✂ Internal/external memory

