TDDD43 Advanced Data Models and Databases

http://www.ida.liu.se/~TDDD43

6hp

Teachers

- Examiner: Patrick Lambrix
- Lectures: Patrick Lambrix, Olaf Hartig, Valentina Ivanova,
- Labs: Valentina Ivanova, Huanyu Li
- Director of studies: Patrick Lambrix

Course literature

- Articles (on web/handout)
- Lab descriptions (on web)

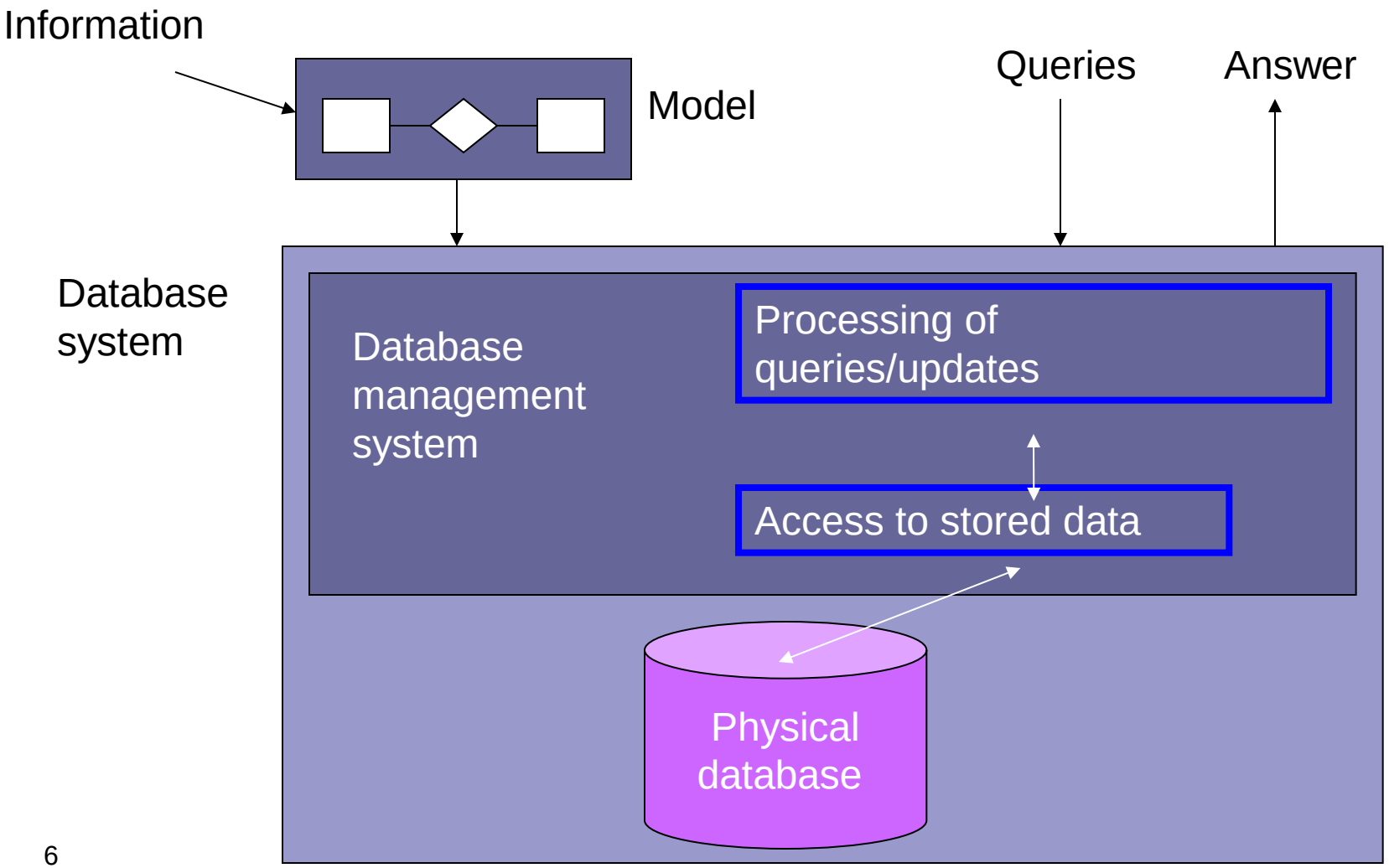
Databases / Data sources

- One (of several) ways to store data in electronic format
- Used in everyday life: bank, hotel reservations, library search, shopping

Databases

- Database management system (DBMS): a collection of programs to create and maintain a database
- Database system = database + DBMS

Databases / Data sources



What information is stored?

- Model the information
  - Entity-Relationship model (ER)
  - Unified Modeling Language (UML)

What information is stored? - ER

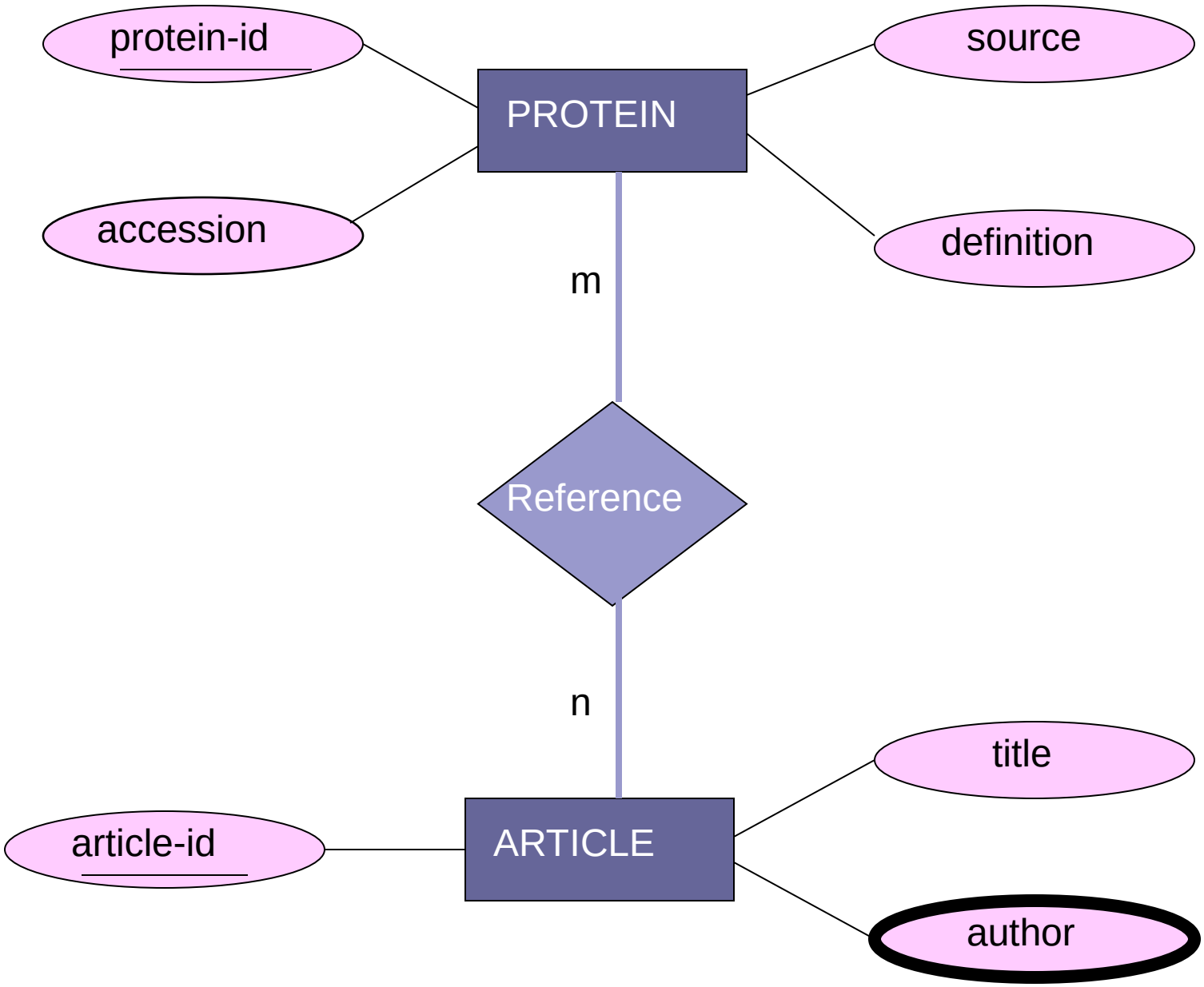
- entities and attributes
- entity types
- key attributes
- relationships
- cardinality constraints

- EER: sub-types

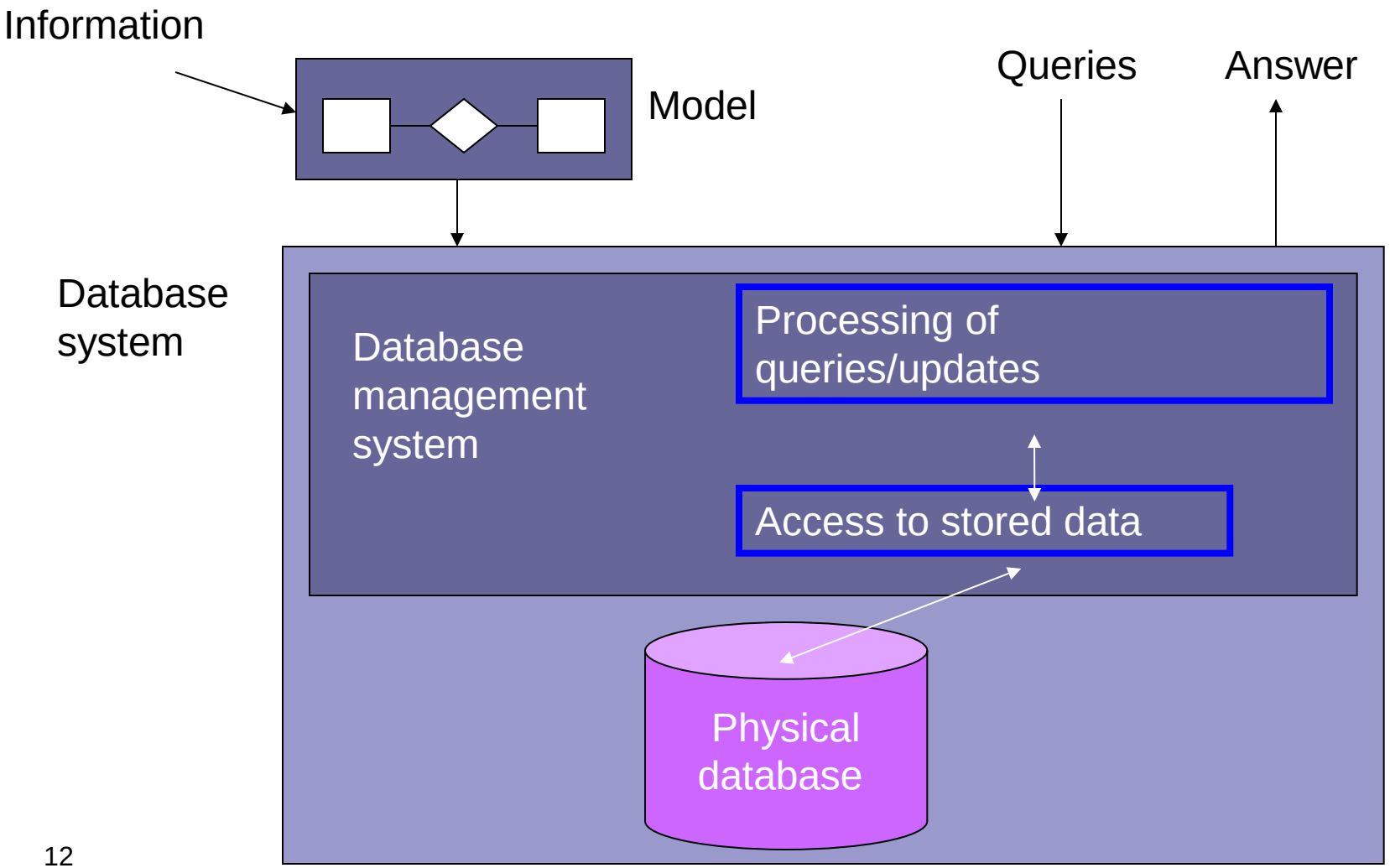
```
   1 tgctacccgc gcccgggctt ctggggtgtt ccccaaccac ggcccagccc tgccacaccc
  61 cccgcccccg gcctccgcag ctcggcatgg gcgcgggggt gctcgtcctg ggcgcctccg
 121 agcccggtaa cctgtcgtcg gccgcaccgc tccccgacgg cgcggccacc gcggcgcggc
 181 tgctggtgcc cgcgtcgccg cccgcctcgt tgctgcctcc cgccagcgaa agccccgagc
 241 cgctgtctca gcagtggaca gcgggcatgg gtctgctgat ggcgctcatc gtgctgctca
 301 tcgtggcggg caatgtgctg gtgatcgtgg ccatcgccaa gacgccgcgg ctgcagacgc
 361 tcaccaacct cttcatcatg tccctggcca gcgccgacct ggtcatgggg ctgctggtgg
 421 tgccgttcgg ggccaccatc gtggtgtggg gccgctggga gtacggctcc ttcttctgcg
 481 agctgtggac ctcagtggac gtgctgtgcg tgacggccag catcgagacc ctgtgtgtca
 541 ttgccctgga ccgctacctc gccatcacct cgcccttccg ctaccagagc ctgctgacgc
 601 gcgcgcgggc gcggggcctc gtgtgcaccg tgtgggccat ctcggccctg gtgtccttcc
 661 tgcccatcct catgcactgg tggcgggcgg agagcgacga ggcgcgccgc tgctacaacg
 721 accccaagtg ctgcgacttc gtcaccaacc gggcctacgc catcgcctcg tccgtagtct
 781 ccttctacgt gcccctgtgc atcatggcct tcgtgtacct gcgggtgttc cgcgaggccc
 841 agaagcaggt gaagaagatc gacagctgcg agcgccgttt cctcggcggc ccagcgcggc
 901 cgccctcgcc ctcgccctcg cccgtccccg cgcccgcgcc gccgcccgga cccccgcgcc
 961 ccgccgccgc cgccgccacc gccccgctgg ccaacgggcg tgcgggtaag cggcggccct
1021 cgcgcctcgt ggccctacgc gagcagaagg cgctcaagac gctgggcatc atcatgggcg
1081 tcttcacgct ctgctggctg cccttcttcc tggccaacgt ggtgaaggcc ttccaccgcg
1141 agctggtgcc cgaccgcctc ttcgtcttct tcaactggct gggctacgcc aactcggcct
1201 tcaaccccat catctactgc cgcagccccg acttccgcaa ggccttccag ggactgctct
1261 gctgcgcgcg cagggctgcc cgccggcgcc acgcgaccca cggagaccgg ccgcgcgcct
1321 cgggctgtct ggcccggccc ggacccccgc catcgcccgg ggccgcctcg gacgacgacg
1381 acgacgatgt cgtcggggcc acgccgcccg cgcgcctgct ggagccctgg gccggctgca
1441 acggcggggc ggcggcggac agcgactcga gcctggacga gccgtgccgc cccggcttcg
1501 cctcggaatc caaggtgtag ggcccggcgc ggggcgcgga ctccgggcac ggcttcccag
1561 gggaacgagg agatctgtgt ttacttaaga ccgatagcag gtgaactcga agcccacaat
1621 cctcgtctga atcatccgag gcaaagagaa aagccacgga ccgttgcaca aaaaggaaag
1681 tttgggaagg gatgggagag tggcttgctg atgttccttg ttg
```

DEFINITION	Homo sapiens adrenergic, beta-1-, receptor
ACCESSION	NM_000684
SOURCE ORGANISM human	
REFERENCE	1
AUTHORS	Frielle, Collins, Daniel, Caron, Lefkowitz, Kobilka
TITLE	Cloning of the cDNA for the human beta 1-adrenergic receptor
REFERENCE	2
AUTHORS	Frielle, Kobilka, Lefkowitz, Caron
TITLE	Human beta 1- and beta 2-adrenergic receptors: structurally and functionally related receptors derived from distinct genes

Entity-relationship

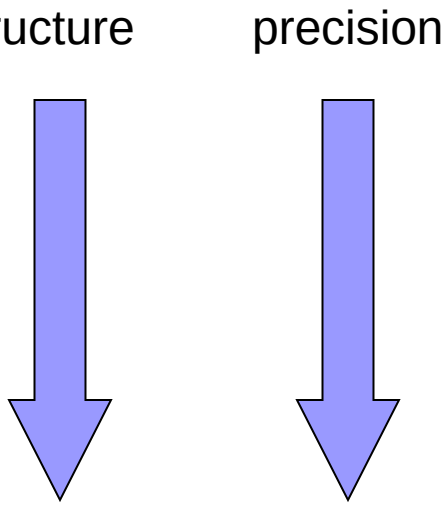


Databases / Data sources



How is the information stored? (high level)
How is the information accessed? (user level)

- Text (IR)
- Semi-structured data
- Data models (DB)
- Rules + Facts (KB)



Course overview

- Information Retrieval (HT1 – lectures)
- Semi-structured data, XML and RDF (HT1 - lectures + labs)
- NoSQL databases (HT2 - lectures + lab)
- Semantic Web, Ontologies, OWL (HT1+HT2 - lectures + lab)
- Data integration (HT1+HT2 – lectures + lab)

Info

- Results reported in connection to exams
- Info about handing in labs on web; strong recommendation to hand in as soon as possible
- Sign up for labs via web **in pairs**

Examination

- TDDD43
  - Written exam
  - Labs

Changes w.r.t. last year

- Update of lectures
- Minor clarifications in labs

My own interest and research

- Modeling of data
  - Ontologies
- Ontology engineering
  - Ontology alignment
    (Winner Anatomy track OAEI 2008 /
    Organizer OAEI Anatomy track since 2013 /
    Organizer OAEI Interactive track since 2015)
  - Ontology debugging
    (Founder and organizer WoDOOM/CoDeS 2012-2016)
- Former work: knowledge representation, data integration, knowledge-based information retrieval, object-centered databases
- http://www.ida.liu.se/~patla00/research.shtml