Faculty of Science

Course:	CSCI 3030U: Database Systems and Concepts Presentation Seminar

1. Each group is supposed to create a presentation (in Google Slides) about the chosen topic related to databases and present it during the class.
2. Division into **teams of 7** will be done over link provided below. Provide your team by 8th of November. After that the distribution of teams will be random. Choose 3 "team presenters" that will present in the class. (Mark will be done mostly based on the quality of slides created by the entire team.)

https://docs.google.com/document/d/1I8HZRTUya3PViJkFIXZ5_VfihhBa9ai9CaVpGolZ3fc/edit

3. Each group will be given **10 minutes** to present (plus around 1-2 minutes for questions).
4. Submit the presentation as PDF (each student has to submit independently that participated in the task) on Canvas by **December 2nd, midnight (and presented in the last week of classes: Nov 22nd, Nov 24th and Nov 29th)**. Include student names in the presentation on the first page and title.
5. Chosen topic should not be covered in the class but can be an extension of the topic covered in the class. Potential sources include online articles, books on databases (including course book) and research articles from data science conferences (VLDB, SIGMOD, EDBT, ICDE and CIKM).
6. Cite used references at the last slide of the presentation.
7. Grade for the presentation will be part of the participation mark (5%).
8. Quality of the slide and the presentation is important (i.e., no reading text from slides, phone and cards). It is recommended to do practice talk(s).
9. Include examples and figures in your presentation. Try to limit number of text.
10. Criteria for evaluating presentations:
    - Organization (title, outline, examples, effective use of images, etc.) – 3 marks
    - Clarity and coherence of the content – 5 marks
    - Presentation within the assigned time limits – 1 mark
    - Answering questions if any – 1 mark

Sample topics (but not limited to) include:

- Peer-to-peer database systems
- Data Streams
- Query rewriting using views
- Database tuning
- Text mining
- Database security
- Database Privacy
- Privacy-preserving data mining
- Linked Open Data
- Database crowdsourcing
- Data curation
- Data cleaning
- Data Profiling
- Deduplication in databases
- ETL

- Web Data Extraction
- Data Storage and Indexing in Sensor Networks
- Data Clustering
- Scientific Databases
- Geographical Information Systems
- Metadata Management
- Health Informatics Databases
- Visual data mining
- Data Replication
- Distributed Databases
- Storage Structure and Systems
- Views and View Management
- Schema Mapping
- Database reverse engineering
- Text indexing techniques
- Web search and web crawling
- Multimedia databases
- Data Warehouses
- Workflow Management
- Query processing and optimization
- Query languages
- Temporal databases
- Stream data mining
- Database Indexing
- Graph database management
- SPARQL/RDF
- Data provenance
- Ranking queries
- Web Search
- Responsible Data Science
- Fairness in Raking
- Team Formation