



# UiT The Arctic University of Norway

<b>Short home exam in:</b>	Inf- 3701
<b>Date:</b>	Monday 16.05.2022
<b>Time (duration):</b>	4 hours
<b>Course coordinator:</b>	Randi Karlsen
<b>Number of pages:</b>	2
<b>Support:</b>	You may call 776 20 880 for support.
<b>Weighting of questions, or other information:</b>	
<b>Important information about this exam</b>	<ol style="list-style-type: none"><li>1. This is an individual examination and must be completed without cooperation of any kind.</li><li>2. You are allowed to use any kind of sources (lecture notes, lectures in pdf. format, books, internet-sources) in answering the topics of the exam.</li><li>3. All exam answers delivered in Wiseflow are automatically checked for plagiarism. Be aware that copying sources, web pages, other students or literature without reference is not allowed, and will be considered as cheating.</li></ol>

**Important:**

*Answer all questions in your own words (i.e. do not copy and paste)*

*Use references – to the books / publications your answer is based on.*

*In case of inconsistencies between info sources: use publications in the course's Reading List for correct answers.*

*If there are any uncertainty concerning a question. State your assumptions concerning the uncertainty before answering.*

**Task 1: (25%)**

- a) Which are the challenges of modern data management applications compared to traditional applications of relational database? Describe the Big Data challenges and the main properties of Big Data.
- b) Describe what is OLAP, what is a data cube, and which are the 4 main operations within OLAP systems.

**Task 2: (25%)**

- a) Describe the classification hierarchy as a directed acyclic graph (DAG) within information retrieval systems. Describe what a directory is in the context of information retrieval systems.
- b) Describe the challenges of false negatives and false positives occurring in document ranking strategies. Describe the two metrics to measure retrieval effectiveness, precision and recall.

**Task 3: (25%)**

- a) Discuss the difference(s) between providing replication consistency and database consistency.
- b) Describe how the quorum consensus protocol can be used to express both strict and weak consistency. Explain the consequences of allowing weak consistency.

**Task 4: (25%)**

- a) Describe objective(s) and challenge(s) of data partitioning in homogeneous distributed databases.
- b) Describe how semi-structured data can be used to overcome interoperability problems.