Practical 3: Research in Groups

IS5102 – Database Management Systems

Due date: Friday 18th November 2022 (week 10), 21:00
34.0% of continuous assessment for the module
(MMS is the definitive source for deadline and credit details)

Summary

This is a **group assignment**.

You will be assigned to groups, and your group will be assigned a topic. Your task is to research the topic as a group, and present your findings in a report.

You are expected to have read and understood all the information in this specification at least ten days before the deadline. You must contact the lecturer regarding any queries well in advance of the deadline.

Arrange to meet with your group members early and often. A key part of working in information systems is to harmonise different perspectives through team work, and this practical will help to build your skills in this area.

Purpose

The aims of this practical are to get experience in:

- database related issues, related to the module but not covered extensively in lectures;
- researching a topic to some detail;
- working as a small group;
- writing a report to present research findings at appropriate levels of detail.

Group membership

This assignment is to be done in groups. If you have any serious reasons as to why you should NOT be in the same group as someone else please contact the Student Welfare Officers by email at `student-welfare-cs@st-andrews.ac.uk` as soon as possible, and in any event before Wednesday 9th November, by which time the groups are expected to start working together. If you are unable or unwilling to raise a concern before that date, you can still contact the Student Welfare Officer(s) within one week of the groups having been announced. Note that in this case it may be much harder to make group changes.

Project allocation

Topic 1: Before Databases

Databases were first described by E.F. Codd in 1970, but people needed to organise data for centuries before that, for example in libraries and record offices, largely using paper records manipulated by hand. How did some of these pre-digital data management systems work, and what were the similarities and differences compared to modern database management systems?

Topic 2: Databases and the cloud

The cloud has brought new opportunities but also new challenges to database system design. What new opportunities arise in this context? What are the issues faced by an organisation trying to move databases to the cloud, and how can they be overcome?

Topic 3: Database Security

Database management systems are often accessed through a network, and requests may come from anyone anywhere in the world. What security features do modern databases have to ensure data can only be read and written by the right sources, and how do they work in practice?

Topic 4: Data Ethics

In recent decades, databases have made it much easier for corporations, governments and individuals to collect and store large amounts of information. What are the ethical concerns associated with this data gathering, and what attempts have been made to enforce responsible data policies? How do modern SQL databases support more ethical handling of personal data?

Topic 5: Tools for Data Analytics

Databases accumulate valuable information, which can be retrieved and processed by various tools for different aspects of data analytics: data mining, visualisation, reproducibility, high-throughput computing, etc. Identify and compare some popular and useful tools and systems for key aspects of data analytics, including their strengths and benefits.

Topic 6: Jobs in the Data Industry

The advent of Big Data has led to many professions starting with the word "Data", such as e.g."data analyst", "data scientist", "data steward", "data engineer", "data architect", etc. What do these roles involve, how do they differ, and how they can help organisations large and small?

Topic 7: Processes in the Data Industry

The advent of Big Data has led to many new terms starting with "Data" and describing dealing with information at various stages; to name just a few, there are "data cleansing", "data fusion", "data governance", "data mining", "data modelling", "data scraping", "data wrangling", etc. What do these processes involve, which tools exist to support them, and how can they be effectively used by organisations large and small?

Submission

Submit a single **PDF** file containing two reports. The first report is the **group report** and it should be the same for all members of a particular group. It should be followed by an **individual report** (that is, different for every group member).

Group Report

Please present the findings of your group research as a research report. It should be properly formatted and structured, and should go into deeper details on the material you discovered, with the context, background and findings. As a general guide, you might want to consider using a structure addressing the following points:

- Context; the problem being solved
- Background on alternative approaches
- Current open questions and/or key issues to consider
- Recommendations and/or conclusions

Feel free to illustrate your arguments with examples, tables, diagrams, or plots. You should assume that the report is intended for the audience consisting of the students currently taking IS5102 module.

The report should have a title page stating the module code, the group number, the title of the reported topic, and matriculation numbers of all participants of the group.

Each section or subsection of the report should have a named individual *lead author* (indicated by the matriculation number in the title of the section/subsection), who takes responsibility for coordinating contributions from other group members to its content, editing the final version of the text, and ensuring that it complies with Good Academic Practice. It is expected that each member of the group is a *lead author* of some part of the report. It is also expected that each member of the group is capable of answering high or medium level questions on *any* part of the report.

The report should be written in a formal third person style, and be carefully proofread to ensure that it uses correct grammar and spelling. It should contain properly formatted bibliography (for example, using IEEE[1] or APA[2] styles), listing the sources of information on which your report draws (including software citations, where applicable).

The advisory word limit for the group report is 6000 words, excluding the title page and the bibliography.

Again, you are expected to work on this task as a group, and all members of the group should submit the same group report. The only pages that differ in your group's submissions should be those with your individual reports.

Individual Report

Each student must also write an individual report, describing their contributions to the group work and reflecting on their experience. The individual report should cover the following questions:

- which (sub-)topic did you research?
- to which parts of the group report did you contribute, and how?
- how did you organise your own work?
- how the collaborative work was organised?
- what did you feel went well, in your own work, as well as in the group work?
- what did you find challenging?
- what problems did you encounter, if any, and how did you try to resolve those?
- what would you do differently, if anything?

and should contain any other relevant information. The advisory word limit for the individual report is 1000 words.

[1] `http://www.ieee.org/documents/ieeecitationref.pdf`
[2] `https://libguides.st-andrews.ac.uk/Referencing_styles/APA/7th`

Marking

The marking of this assignment will be based on the quality of research done; on the quality of presenting your research in the report; and on its suitability for the intended audience, assumed to comprise of students taking IS5102 module.

It is expected that all students in the same group will receive the same grade for their *group report*. Individual grades will be determined based on the *individual report*. If there are unresolved problems with teamwork within a group, these should be identified in the individual reports. After careful consideration, individual grades may be further amended to take such problems into account. Of course, it is in everybody's interests to flag up issues or problems early with the lecturer.

The final grade for this assignment will be determined as follows:

- 75% for the group report;
- 25% for the individual report.

Policies and Guidelines

Marking

See also the standard mark descriptors in the School Student Handbook, which apply as usual:
`https://info.cs.st-andrews.ac.uk/student-handbook/learning-teaching/feedback.html#Mark_Descriptors`

Lateness

The standard penalty for late submission applies (Scheme B: 1 mark per 8 hour period, or part thereof):
`https://info.cs.st-andrews.ac.uk/student-handbook/learning-teaching/assessment.html#Lateness_Penalties`

Good Academic Practice

The University policy on Good Academic Practice applies:
`https://www.st-andrews.ac.uk/students/rules/academicpractice/`