

## Research design and proposal writing

### **Research topic: Knowledge Graphs for Machine Learning**

**Description of topic:** State-of-the-Art analysis and report on current knowledge graph implementations and frameworks which could be considered as a preprocessing step to benefit data preparation for machine learning tasks. The project would include the implementation of a use case and validation with an open source data set. The goal is to show that usage of a knowledge graph improves accuracy for predictions and provides better benchmarks compared to the baseline.

### FIRST ASSIGNMENT

- Prepare a set of slides - **Exactly 8 slides (studentnumber\_first\_assignment.pdf)** (marks will be given according to how focused, precise, rigorous and relevant the presentation is, as well as the replicability of what is proposed)

**Note:** use all the space in each slide in a way that the content is readable when printed (use white background).

- **slide 1**

**Title:** 'Research question, hypothesis and preliminary design'

**content:** \_name, surname, programme number, stream, student number

- **slide 2**

**Title:** "Domain, scope, assumptions, limitations and delimitations of research - ACM 2012"

**Content:** Describe the DOMAIN of the research using the categories and concepts from the [ACM 2012 classification system](#) (example of one line: **A: Human-centered computing → Human computer interaction (HCI) → HCI design and evaluation methods → Usability testing** ). You must describe your research with 5 (and only 5) lines as a bullet point list (A, B, C, D, E). Cite the relevant articles that you have read (7+) across these 5 branches, according to the [APA7 style](#). Additionally, add below this list the following:

- a focused sentence (only one) for best describing the SCOPE of your research
- a focused sentence (only one) for best describing the ASSUMPTIONS of your research (if you have multiple assumptions, delimitate them with a semicolon ';')
- a focused sentence (only one) for best describing the LIMITATIONS of your research (if you have multiple limitations, delimitate them with a semicolon ';')
- a focused sentence (only one) for best describing the DELIMITATIONS of your research (if you have multiple delimitations, delimitate them with a semicolon ';')

**Marks (total 3%):**

0-> domain AND scope AND limitations, AND assumptions, AND delimitation are vague;

0.5->four of the above are vague and ONLY one is well-defined and clear;

1->three of the above are vague AND ONLY two are well-defined and clear;

2->two of the above are vague AND ONLY three are well-defined and clear;

2.5->only 1 is vague and four are well-defined and clear;

3-> domain, scope, assumptions, limitations and delimitations are all well-defined and clear;

○ **slide 3**

**Title:** "Gaps in the literature and research question"

**content:** Write about:

1. precise and concise gaps extracted from literature you have identified by reading 7+ articles. You can write multiple sentences/paragraphs, organise them as you wish (by theme, gap, discipline, etc.) and justify your claims by using all the relevant articles you have found.
2. precise research question (one 'question' ending with a question mark)

Marks (total 6%):

0-> gaps are unknown or very vague; it is impossible to anticipate a research question from them;

1-> gaps are somehow vague AND not all the identified relevant articles have been cited or contextualised; it is difficult to anticipate the research question from the gaps;

2-> gaps are not vague but can be more precise OR not all the relevant articles have been cited or contextualised; a research question can be imagined, but with some degree of uncertainty;

3-> gaps are not vague and precise, backed up by ALL the relevant articles identified from the literature review, and this can lead the reader to imagine a possible coherent research question from them;

+

0-> research question is unknown or very vague; it is impossible to imagine how to design a research experiment to answer it;

1-> research question is valid BUT it is vague; it is very difficult to imagine how to design a research experiment from it;

2-> research question is not vague but can be more precise; it is in somehow sufficient to imagine how to design some component of a research experiment to answer it, BUT not all;

3-> research question is very precise AND complete AND valid; it is straightforward to imagine how to design a research experiment from it;

○ **slide 4**

**Title:** "hypothesis "

**content:** Define the null and alternate hypotheses for tackling the research question (max 1 sentence each). Note that the null hypothesis does not need to be in the form of 'IF...THEN...', but as a statement reflecting the commonly accepted fact.

Marks (total 5%):

0-> it is not a valid hypothesis; it is impossible to imagine which research tasks need to be implemented to test it;

1-> it is valid but VERY vague; it is VERY difficult to imagine which research tasks need to be implemented to test it;

2-> it is valid but vague; it is difficult to imagine which research tasks need to be implemented to test it;

3-> it is valid AND not vague; it is ONLY sufficient to anticipate SOME of the research tasks that need to be implemented to test it;

4-> it is valid AND not vague; it is sufficient to anticipate ALMOST all the research tasks that need to be implemented to test it;

5-> it is valid AND not vague AND precise. ALL the research tasks that need to be implemented to test it can be clearly anticipated;

○ **slide 5**

**Title:** "Feasibility of the study"

**content:** Briefly describe the activities/tasks you have planned towards testing the

research hypothesis and a rough time required for implementing each of them (BE REALISTIC)

Marks (total 3%):

0-> undefined plan to test hypothesis;

1-> vague plan to test hypothesis with minimal description of the research activities, leaving doubts on how to implement them;

2-> fair plan to test hypothesis with research activities sufficiently described, but NOT fully complete, leaving the reader with some doubts on how to implement some of them;;

3-> good plan to test hypothesis with complete description for EACH research activity, allowing the reader to FULLY implement them, without doubts;

- **slide 6 + 7**

**Title:** "Bibliography"

content: list the articles you have read so far (this should be always an ongoing activity).

Overall you must have 7+ bibliographic peer-reviewed entries. items) **STRICTLY** using the [APA7 style](#)

Mark (total 2%):

0-> the APA7 style has not been used;

1-> the APA7 style has been partially used; some errors OR incompleteness exist in in-text citations OR in the bibliographic section;

2-> correct AND complete use of the APA7 style across in-text citations AND in the bibliographic section;

- **Slide 8**

**Title:** "Performance metrics of your experiment"

Content: Provide details about the performance metrics you are going to use in your experiment (names, ranges, type, etc.). Overall, provide as many details as possible to inform the reader on the performance metrics you will be using in your experiment and its evaluation. Examples - for Data Science students (DS), these metrics include (among many more) precision, recall, f1-score, ROC curves, AUC curves, etc. For Advanced Software Development (ASD) students, these metrics include (among others) time, resource consumption, memory, Signal-to-noise ratio, usability metrics, user experience metrics, etc.

Mark (total 1%):

0-> no performance metrics have been identified;

0.5-> some performance metrics have been identified, BUT within minimal description, leaving readers with doubts;

1-> performance metrics have been clearly identified AND clearly described AND all the details of each relevant metric have been provided; the description FULLY help readers to understand how the experiments, aimed at testing the research hypothesis, can be evaluated/validated with performance metrics;

- 

**Note** - While building your slides, take into account the following questions

- Is the scope of the research clear? Is the motivation convincing? Are the gaps well defined? Is the research question focused or still ambiguous, too narrow, or too broad?
- Is the hypothesis clear and convincing? How can it be better expressed?
- Do you believe the study is feasible given the time frame provided? (master or phd?)