University of Aveiro, Information Retrieval, 2024/2025

Assignment 3 – Monograph (individual work)

The monographs will be based on the research, analysis (critical summary), report, and presentation of one subject related to the course.

You must select a topic from this list (the topics are also listed below). To avoid duplicates, you must first check the already selected topics here. It is possible to change the selection by filling out the form again.

You must submit the monograph (PDF, up to 6 pages, excluding cover pages and references) and a small presentation (also PDF), expect to last 8-10 minutes.

Both the document and the presentation must be organized in the following aspects:

1- Introduction to the topic, to frame the audience with the problem/subject

2- presentation of the approaches studied in the articles, focusing on the central methods/models/strategies and the differences between the approaches

3- conclusions

Important notes

1- It is expected that you attend all the sessions

2- The works must be presented to the entire audience, in a way that allows them to understand the topic

3- Proactivity and interaction (questions/comments) on the part of those watching will be valued

Monographs topics

 # Indexer and Searcher:

1. Inverted Index Optimization Techniques
2. Query Expansion
3. Pseudo Relevance Feedback
4. Efficient Multi-lingual Search Indexing
5. Index Compression Techniques (Gamma Encoding, Elias-Fano, ...)
6. PageRank and Link-based Ranking Algorithms
7. Finite State Transducer (https://blog.burntsushi.net/transducers/)
8. WAND (weak and) and Block-Max-WAND
9. Machine Learning in IR (SVM, CNNs and RNNs)

 # NIR:

10. Word Embeddings and Representational Learning (Word2Vec, GloVe, fastText, ...)
11. Learning-to-Rank Techniques (RankNet, LamdaNet, LambdaMART, ...)
12. Dense Retrieval
13. Learned Sparse Retrievals
14. Differential Search Index
15. Retrieval Augmented Generation

16. Semantic Search and Embedding-based approaches to search
17. Transformer models encoder only for IR (BERT and others)
18. Transformer models encoder-decoder or decoder only for IR (T5 and others)
19. Attention Mechanisms and Their Role in NIR
20. Transfer Learning in NIR
21. Chatbots and Conversational Agents in Information Retrieval
22. Knowledge Graphs in Enhancing Search Capabilities
23. Multi-modal IR: Text, Image, and Video
24. Low-resource Language Retrieval Using Neural Methods

 # Others:

25. Web Crawlers: Design, Challenges, and Optimization
26. Data Augmentation Techniques in IR Training Datasets
27. Cross-lingual Information Retrieval with Neural Models
28. Query Performance Prediction
29. Recommended Systems
30. Interactive Neural IR Systems and User Engagement
31. Personalized Search: Balancing Relevance and Privacy
32. Ethical Considerations (bias, privacy, transparency, ...)
33. Temporal Dynamics in IR: Time-aware Retrieval Systems
34. Evaluating IR Systems Beyond Traditional Metrics (NDCG, MRR, etc.)
35. Research Trends and Future Directions

 # Tools/Case studies:

36. Tools: SPARe (https://github.com/ieeta-pt/SPARe)
37. Tools: PyTerrier (Terrier)
38. Tools: Pisa (https://github.com/pisa-engine/pisa)
39. Tools: PySerini (Anserini)
40. Tools: PyTorch vs. TensorFlow for NIR Development
41. Case Study: Google's Search Algorithms and Ranking Mechanisms
42. Case Studies: Web search engines, e-commerce, and digital libraries