Master 2 Internship offer – Engineer

DocInnovate:
Multimodal Interactive Exploration of Heritage Documents
and Historical Archives

Keywords : Digital Humanities, Multimodal Document Analysis, Indexing of Heterogeneous Contents

Additional Information :

- **Supervisor:** BAKKALI Souhail
- **Team:** Images and Content (IC)
- **Internship Start Date:** February/March 2024
- **Internship Duration :** 6 months
- **Remuneration:** 600 € net per month (legal amount)
- **Workplace :** L3i, La Rochelle

Application Elements :

- — CV
  — Cover Letter
  — Latest Academic Transcripts
- Please send these documents to the following:

  Object: [L3i – poste] M2 Internship
  Email: souhail.bakkali@univ-lr.fr

1. About L3i Laboratory :

The L3i Laboratory, established in 1993 at the University of La Rochelle, brings together researchers in computer science and signal processing from various faculties. L3i combines the expertise of its researchers to address the challenges of enhancing digital content from a systemic perspective. This involves a cross-exploitation of interactive applications, content indexing, and knowledge representation.

The laboratory is structured around three scientific themes (Knowledge Engineering, Content Analysis and Management, Interactivity and Dynamic Systems), all focused on the common goal of interactive and intelligent management of digital content.

2. Brief Description of the Project:

The Master 2 internship is part of a project integrated into the Sustainable Intelligent Urban Coastal Institute at La Rochelle University. Its main objective is to address the major challenge of information retrieval across various documents using innovative approaches from digital humanities. This project aligns with the dynamics of La Rochelle University by extracting, structuring, and aggregating information from institutional and heritage archives.

The initiative aims to simplify information management, particularly within La Rochelle University, by proposing innovative approaches to extract semantically enriched data. It will contribute to the preservation of digital heritage by extracting information from historical archives, ancient manuscripts, research contracts, thus fostering knowledge sharing and strengthening research activities. Specifically, the project falls within the realm of digital document management by contributing to the generation of summaries and descriptions of documents. These interventions aim to efficiently organize digital resources, preserve document integrity, and enhance the search and understanding of visual information.

3. Scientific Context :

The project is part of a comprehensive initiative aimed at addressing the significant challenge of generalizing analysis and understanding models of documents in the face of the constant expansion of online collections. Despite the progress in machine learning methods, the challenge persists for multimodal deep learning models integrating complex data such as images and text. It focuses on two crucial and less-explored aspects: the generalization capability of models and the improvement of interactions between different sensory modalities.

The goal is to design multimodal deep learning models capable of effectively leveraging information across modalities, thereby enhancing the quality of learned representations. This innovative approach represents a significant contribution to research in digital humanities,

facilitating advanced research within vast documentary archives and improving the relevance of results, especially in the context of historical documents.

To comprehensively address the complex challenge of managing diverse documents in an administrative context, the project presented in this initiative unfolds in a two-fold approach:

• Adaptation of the integration model for comprehensive online evaluation: In this first phase, the Master 2 intern will be tasked with adapting the pre-training integration model to specifically meet the requirements of crucial tasks such as generating document summaries and descriptions. The recognition of the importance of this adaptation stems from the need for a comprehensive evaluation, integrating aspects specific to administrative reality. By jointly integrating language analysis and visual cues, the goal is to develop a versatile system capable of effectively processing diverse documents. The expected outcome is the creation of a generic model, applicable to various tasks, thereby strengthening the project's ability to adapt to complex document challenges.

• Set of digital tools for the scientific exploitation of documents: The second phase of the project focuses on the development of digital tools dedicated to the scientific exploitation of documents. This will include the creation of a search engine capable of retrieving documents visually and semantically based on the query content. These tools will be designed with the specificities of digital humanities in mind, ensuring adaptation to real administrative issues. The Master 2 intern will be involved in the design, development, and validation of these tools, contributing to strengthening the project's efficiency in the field of scientific exploitation of documents.

4. Prerequisites and Specific Constraints:

• Master 2 / Engineer level

• The student should have a strong foundation in image processing (Computer Vision), natural language processing (Natural Language Processing), and deep learning.

• Proficiency in Python programming and familiarity with Tensorflow/Keras and/or Pytorch would be highly appreciated.

• Experience in GPU programming could be an advantage.

• Good written and oral proficiency in English is highly desirable.

5. Bibliographical References :

[1]. Bakkali, S., Biswas, S., Ming, Z., Coustaty, M., Rusiñol, M., Terrades, O. R., & Lladós, J. (2023). TransferDoc: A Self-Supervised Transferable Document Representation Learning Model Unifying Vision and Language. arXiv preprint arXiv:2309.05756.

[2]. Widyassari, A. P., Rustad, S., Shidik, G. F., Noersasongko, E., Syukur, A., & Affandy, A. (2022). Review of automatic text summarization techniques & methods. Journal of King Saud University-Computer and Information Sciences, 34(4), 1029-1046.

[3]. Mahalakshmi, P., & Fatima, N. S. (2022). Summarization of text and image captioning in information retrieval using deep learning techniques. *IEEE Access*, *10*, 18289-18297.

[4]. Kuo, C. W., & Kira, Z. (2022). Beyond a pre-trained object detector: Cross-modal textual and visual context for image captioning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 17969-17979).

[5]. Dang, H., Benharrak, K., Lehmann, F., & Buschek, D. (2022, October). Beyond text generation: Supporting writers with continuous automatic text summaries. In Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology (pp. 1-13).

[6]. Altmami, N. I., & Menai, M. E. B. (2022). Automatic summarization of scientific articles: A survey. Journal of King Saud University-Computer and Information Sciences, 34(4), 1011-1028.