

Natural Language Processing Research Assistant

Full time or part time

Complexity Science Hub Wien

€ 2.082,80* (40h)

* salary negotiable based on experience and qualifications



COMPLEXITY
SCIENCE
HUB
VIENNA

The Complexity Science Hub Vienna is a young research institution dedicated to a deeper quantitative and predictive understanding of complex systems for the betterment of society. We provide an exciting, creative, bureaucracy-free environment for open-minded visionaries who want to make a change and are brave enough to step out of mainstream science. We search for a

Natural Language Processing Research Assistant

THE PROJECT

The appointed candidate will join an interdisciplinary research team modeling and analyzing the evolution of complex human societies and identifying factors causing state breakdown and collapse. The empirical arm of this project is the Seshat Databank that aggregates the knowledge of historians and archeologists about long-term evolution and dynamics of past societies. The candidate can also become involved in the work of other research groups. The main objective of this project is to help develop an algorithm that can help expand the Seshat Databank. In particular, the algorithm should parse through history and anthropology books and articles and detect paragraphs that contain relevant historical information for a given civilization.

YOUR RESPONSIBILITIES

The expected tasks are (i) create a pipeline to automate text extraction through optical character recognition (OCR) from scientific books and articles into python (ii) help organizing and refine labels of the Seshat Repository so it can be used as labeled dataset for information extraction from text (iii) develop and train a large language models to automatically detect data-candidates i.e. paragraphs that contain relevant historical information for a given civilization.

YOUR PROFILE

Required:

BSc or MSc in Computer Science, Engineering, related fields or work/research experience in data science and text analysis

- Fluency in English
- Python coding expertise
- You are proactive, diligent, self-motivated, inquisitive and team-oriented.

Ideal or willingness to learn:

- Expertise in NLP and OCR, including using Hugging Face models and libraries such as Python-tesseract
- Excellent communication skills, ability to explain complex topics and code in simple words to a non-expert audience.
- Familiarity with version control systems, especially GIT
- Experience managing and working with large datasets
- You are curious about history, complex systems, and science in general

WE OFFER

A position in a dynamic, creative, internationally oriented, and open working environment. Mentoring and a collaborative environment. The variety of projects at the Hub will allow you to engage in new and hot topics and to collaborate with world renowned scientists.

APPLICATION

Please send your application to applications@csh.ac.at with the subject line "Natural Language Processing Research Assistant". Review of applications will start on the 15th of January; later applications will be reviewed on a rolling basis until the position is filled. Application material includes a CV, and a one-page cover letter stating your motivation and fit with the job qualifications, a copy of your degree/s, and any additional material showing your computational skills and creativity. CSH is committed to the principle of equal employment opportunity for all applicants. All employment decisions are therefore based on job requirements, qualifications, merit, and organizational needs. We strongly encourage individuals from underrepresented groups to apply. The CSH Vienna processes your personal data in accordance with the law (www.csh.ac.at/data-protection/)