

## Post-Doctoral Researcher in NLP for Historical Texts and Knowledge Graphs VOICES project

**Position:** Post-Doctoral Researcher in NLP for Historical Texts and Knowledge Graphs

**Duration:** Specific purpose, One year full-time contract (VOICES project; funded research position)

**Location:** ADAPT Research Centre, Trinity College Dublin, Ireland

**Application Deadline:** 12pm on 13.03.2026

**Start Date:** As soon as possible

**Salary Scale:** PD1: €46,805 – €53,391 per annum IUA scales, commensurate with experience and in line with TCD/Irish funding agency scales

### Project and Role Overview

We are seeking a post-doctoral researcher to join the ERC funded VOICES project (<https://voicesproject.ie/>), which focuses on developing tools to support historians in analysing historical texts for references to women and their roles in society. The successful candidate will work at the intersection of natural language processing, large language models, and digital humanities, building end-to-end pipelines that extract structured information from historical English texts and expose it through usable interfaces and knowledge graphs.

The project will involve:

- Designing and adapting named entity recognition (NER) and related information extraction models for historical and domain-specific English.
- Integrating large language model (LLM)-based approaches with traditional NLP pipelines.
- Developing interfaces and infrastructure that allow historians to interact with the models and inspect the extracted knowledge.

### Key Research Themes

The research will centre on several technical themes:

#### Information extraction and NER

- Detecting and classifying entities such as people, places, dates, events, and other relevant concepts in heterogeneous historical sources.
- Handling older and non-standard forms of English, OCR noise, and variable spelling.

#### Domain adaptation and model development

- Adapting existing NER and sequence labelling models to historical and domain-specific corpora (fine-tuning, prompt-engineering, data augmentation, etc.).
- Combining LLM-based methods (e.g. instruction-tuned models, in-context learning) with more classical NLP architectures (e.g. CRFs, BiLSTM-CRF, transformer encoders).

#### NLP pipelines and tooling

- Building robust NLP pipelines for pre-processing, annotation, and evaluation across diverse document types (letters, newspapers, administrative records, etc.).
- Using established toolkits and APIs (e.g. spaCy, Hugging Face, Stanza, or similar) and extending them where necessary.

#### HOST INSTITUTIONS



#### PARTNER INSTITUTIONS



## Interfaces and infrastructure

- Wrapping models with APIs and user-facing tools (e.g. web dashboards, annotation/inspection tools) suitable for non-technical historians.
- Ensuring reproducible deployment, basic monitoring, and maintainability of the developed tools.

## Knowledge graphs and semantic enrichment

- Transforming extracted entities and relations into structured representations and ingesting them into a knowledge graph.
- Linking extracted entities to the VOICES knowledge graph and external resources where appropriate (e.g. authority files, gazetteers), and supporting historians in exploring the data.

This position offers the opportunity to collaborate closely with historians and digital humanities researchers, contribute to high-impact publications, and help shape emerging methods for computational history and gender studies.

## Responsibilities

The successful candidate will:

- Conduct research on information extraction and NER for historical English texts, focusing on entities such as women, places, and dates in historical sources.
- Design, implement, and evaluate NER and related models using both LLM-based approaches and traditional NLP methods, including domain adaptation to historical corpora.
- Build end-to-end NLP pipelines using standard toolkits and APIs, including data pre-processing, model training, evaluation, and error analysis.
- Develop interfaces (e.g. web applications or interactive tools) and supporting infrastructure that allow historians to query, visualise, and critique model outputs.
- Map model outputs into a structured knowledge graph and contribute to the design and population of the knowledge graph repository.
- Collaborate with an interdisciplinary team of historians, digital humanities scholars, and computer scientists, including requirements gathering and iterative tool design with end users.
- Publish research findings in reputable conferences and journals in NLP, AI, and digital humanities, and contribute to open-source software and datasets where feasible.
- Contribute to project documentation, reporting, and dissemination activities (e.g. workshops, demos, talks).

## Qualifications

### Essential

- PhD in Computer Science, Artificial Intelligence, Natural Language Processing, Computational Linguistics, or a closely related discipline (or thesis submitted by the start date).
- Strong background in NLP, including experience with sequence labelling tasks such as named entity recognition, relation extraction, or similar.
- Demonstrated experience with deep learning frameworks (e.g. PyTorch, TensorFlow) and modern NLP toolkits (e.g. Hugging Face Transformers, spaCy, or equivalent).

#### HOST INSTITUTIONS



#### PARTNER INSTITUTIONS



- Experience in training, fine-tuning, or adapting language models (e.g. transformer-based architectures such as BERT-like models, encoder-decoder models, or LLMs accessed via APIs).
- Strong programming skills in Python and experience with software engineering best practices (version control, testing, documentation).
- Excellent problem-solving skills, with the ability to work independently and as part of an interdisciplinary team.
- Strong communication skills in English, including the ability to explain technical concepts to non-technical collaborators.

### Desirable

- Prior experience with historical or noisy texts, domain adaptation, or low-resource NLP.
- Familiarity with using and adapting large language models in practical applications (e.g. prompt design, retrieval-augmented generation, instruction tuning).
- Experience building APIs and/or simple web applications (e.g. using FastAPI, Flask, Django, or similar) to wrap models and provide user-facing tools.
- Experience with knowledge graphs and semantic technologies (e.g. RDF, OWL, SPARQL, graph databases such as GraphDB or Neo4j).
- Previous work with digital humanities projects or collaboration with humanities or social science researchers.
- Track record of publications in relevant venues in NLP/AI or digital humanities.

### Application Process

Interested candidates should submit the following as a single PDF file [here](#)

- Curriculum vitae (max. 4 pages), including a list of publications.
- Cover letter (max. 2 pages) describing how your experience fits the role and your interest in working with historical texts and digital humanities.
- Contact details for two referees (no reference letters required at application stage).

Informal enquiries about the role can be sent to Prof John Kelleher [john.kelleher@adaptcentre.ie](mailto:john.kelleher@adaptcentre.ie), with the subject line "**VOICES Post-Doctoral Researcher enquiry**". The application closing date is 12pm on 13.03.2026; late applications will not be accepted.

#### HOST INSTITUTIONS



#### PARTNER INSTITUTIONS

