

MIXED-METHODS DIGITAL ORAL HISTORY

Enfolding semantic web technologies and historical-interpretative analysis



RESEARCH TEAM

Prof Dr Julianne Nyhan
PI, TU Darmstadt
Dr Marco Humbel
Postdoc, TU Darmstadt

Dr Andreas Vlachidis
PI, UCL
Dr Frieda Schmidt
Postdoc, TU Darmstadt

Dr Andrew Flynn
CO-I, UCL
Dr Jiajie Zhang
Postdoc, UCL

OBJECTIVES

Generate New Historical Knowledge

Articulate the individual and collective human agency that has shaped the emergence and institutionalisation of Digital Humanities since 1949.

Innovate Methodologically

Pioneer the use of Semantic Web technologies and digital methods to conduct high-resolution analysis of narratives within and across a multilingual corpus of oral history interviews.

Create Reusable Digital Artefacts

Develop a conceptual schema (ontology) for the oral history domain and an interoperable Knowledge Graph of interconnected entities.

Champion FAIR Data Principles

Devise and implement a dissemination framework that promotes the Findable, Accessible, Interoperable, and Reusable (FAIR) sharing of oral history data.

FOREGROUNDING INTERSUBJECTIVITY AND NARRATIVE

Oral history interviews are co-creations shaped by interviewer, narrator and context. We foreground this intersubjective space through mixed-methods: **qualitative coding of power dynamics and positionality**, computational detection of hedging, stance and dialogic turns, and network motifs of speaker alignment.

These lenses illuminate how authority is negotiated, how memories converge or clash, and how marginal voices surface. The analysis feeds back into our ontology, ensuring that nuance, rather than only 'facts' is encoded and queryable.

CORPUS ANALYSIS: TEXT MINING THE 'HUMANIST' MAILING LIST

The Humanist mailing list (1987–present) offers a unique longitudinal window on Digital Humanities discourse. Through keyword-in-context, temporal topic modelling, and co-citation analyses, we **integrate findings from the Humanist corpus** with oral history narratives from the Hidden Histories Project.

This comparative approach contextualises personal recollections within the field's broader scholarly dialogue, illuminating the geographical, technical, disciplinary, and institutional forces shaping Digital Humanities.

BACKGROUND

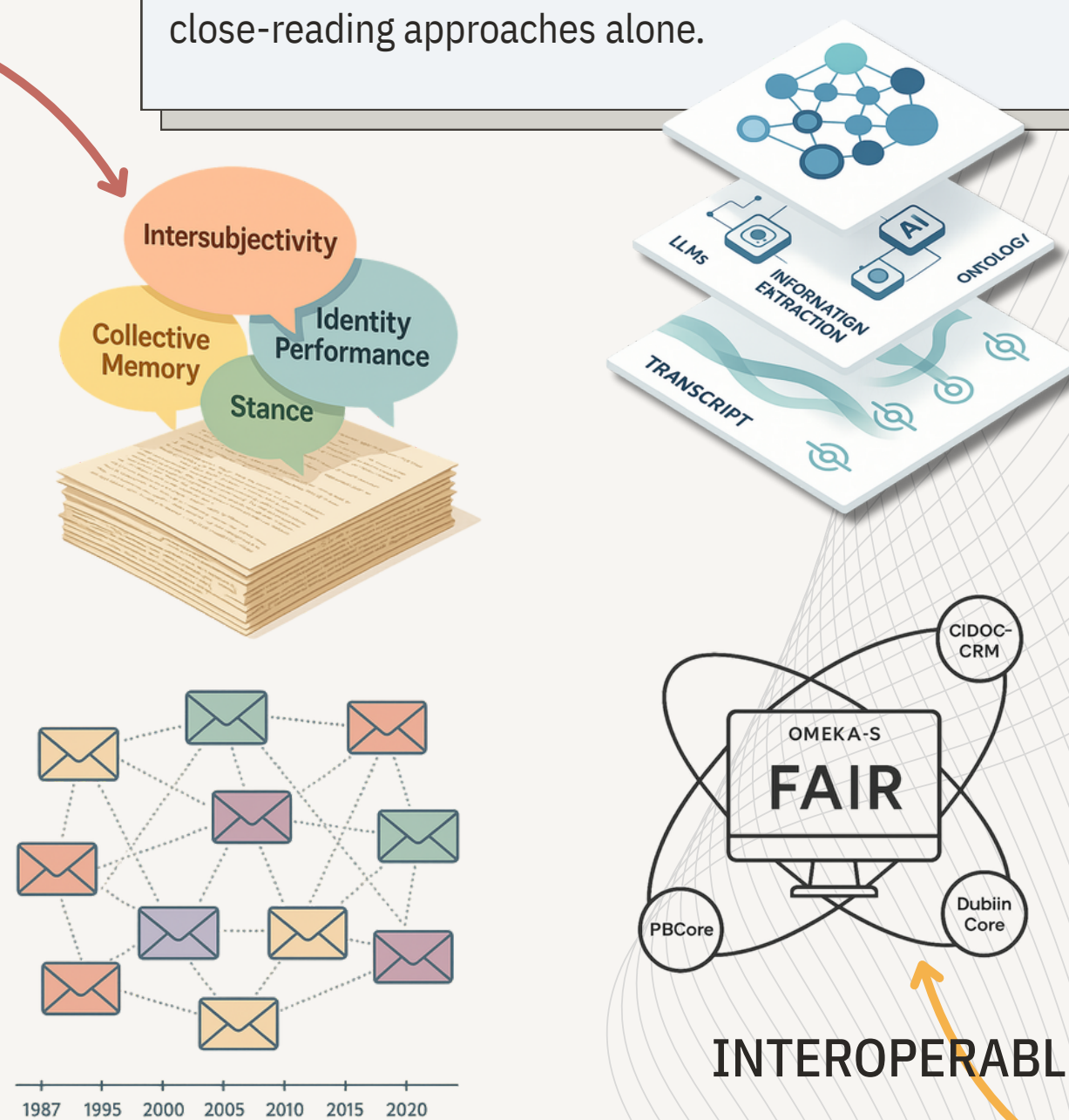
Oral history has long used technology to capture essential, often unheard voices. Yet, engagement with digital technologies has typically focused on dissemination and basic cataloguing rather than deep analytical exploration. Semantic web technologies, capable of uncovering richer interconnections within and across sources, remain underexploited.

The MeDoraH project bridges this gap by integrating **oral history**, **interpretative historical analysis**, and **knowledge graph** technologies. Using the history of Digital Humanities as a case study, MeDoraH pioneers a novel 'digital methodological-hermeneutical' approach, providing new insights into the narratives of formation, disruption, and transformation in the field.

ENABLING A DIGITAL HERMENEUTIC TURN WITH NLP

MeDoraH applies advanced NLP methods to oral history research. Leveraging multilingual NLP pipelines and large language models (LLMs), we systematically **extract nuanced semantic structures** from the Hidden Histories Project interviews, transforming rich narrative data into an interconnected, queryable Knowledge Graph.

This structured semantic framework allows researchers to perform sophisticated analyses, uncovering subtle narrative patterns and interrelationships that would remain obscure through conventional close-reading approaches alone.



INTEROPERABLE DESIGN & FAIR DATA

We have designed a comprehensive, tiered metadata schema for the entire oral history lifecycle, from planning to dissemination. For **maximum interoperability**, this schema is based on the Oral History Association's (OHA) guidance on Metadata Elements and aligned to established standards such as CIDOC-CRM, Dublin Core, and PBCore. Implemented in Omeka-S we provide a **centralised access point** for all project data, including downloadable assets in standard semantic web formats, ensuring consistency, reusability, and accessibility.