SIGHUM 2027 (a.k.a. LaTeCH-CLfL)
The 11th Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature

to be held at EACL in March 2027 in Athens, Greece as a one-day workshop

Organizers: Diego Alves, Yuri Bizzoni, Stefania Degaetano-Ortlieb,
Anna Kazantseva, Janis Pagel, Stan Szpakowicz

First Call for Papers (with apologies for cross-posting)

SIGHUM 2027 is the eleventh in a series of meetings for NLP researchers who work with data from the broadly understood arts, humanities and social sciences, and for specialists in those disciplines who apply NLP techniques in their work. The workshop continues a long tradition of annual events which also host the SIGHUM business meetings.

Workshop site

https://sighum.wordpress.com/sighum-2027/

Important dates

Submission deadline: December 15th, 2026
Notification of acceptance: January 20th, 2027
Camera-ready paper due: January 29th, 2027

Description

The community of the broadly understood Digital Humanities (DH) has witnessed remarkable growth and transformation, fuelled by the rapid advancements in NLP. There is a steady interest in, and a high demand for, NLP methods of semantic and structural annotation, intelligent linking, discovery, querying, cleaning and visualization of primary and secondary data. Even so, the heterogeneous landscape of the DH – with their diverse, often multilingual or multimodal sources – can be a challenge for NLP. Consider, for example, the growing interest in historical language data and in under-resourced languages.

There are unique obstacles in developing comprehensive language models in aid of the linguistic diversity in DH. The handling of noisy and non-standard data, and the need

for domain adaptation and intensive annotation, stay at the forefront of research effort in the community. The literary studies, which have witnessed substantial progress in the application of NLP methods, bring their own similar problems. Navigating forms of creative expression requires more than the typical information-seeking tools. A case in point might be the study of literature of a certain period, author or sub-genre, the recognition of certain literary devices, or the quantitative analysis of poetry.

The Large Language Models (LLMs) expand the DH toolkit. There is support for automatic text cleaning and annotation, creation of semantic resources, analysis of narrative, genre and literary style, and linking information across sources. LLMs can support historical or low-resource languages, particularly when complemented with domain-specific fine-tuning and careful evaluation. One must note, however, that even with careful adaptation, curation and attention to interpretability, LLM outputs remain prone to errors, bias and lack of transparency; rigorous assessment is required to ensure their suitability for scholarly research.

There is growing emphasis on the importance of explanation in NLP models. That applied equally to DH, whose various domains enjoy the effect of NLP. Transparency and clarity of the results are critical if one is to accept the processed data and gain valuable insights. That is why one must carefully consider a balance between raw performance scores and interpretability, in keeping with the specific research objectives.

For many years now, this broad research context has drawn together NLP experts, data specialists and researchers in Digital Humanities who work in and across their domains. Our long-standing series of workshops has shown that cross-disciplinary exchange supports work in the Humanities, Social Sciences and Cultural Heritage communities. It encourages the Computational Linguistics community to build rich and effective tools, and – above all – explainable models.

Topics

Our workshops attract original work on a wide variety of topics, including – but as usual not restricted to – these:

- adaptation of NLP tools to Cultural Heritage, Social Sciences, Humanities and literature;
- automatic error detection and cleaning of textual data;
- complex annotation plans, tools and interfaces;
- creation (fully- or semi-automatic) of semantic resources;
- creation and analysis of social networks of literary characters;
- discourse and narrative analysis/modelling, notably in literature;
- emotion analysis for the humanities and for literature;

- generation of literary narrative, dialogue or poetry;
- identification and analysis of literary genres;
- information/knowledge modelling in the Humanities, Social Sciences and Cultural Heritage;
- interpretability of large language models output for DH-related tasks (explainable AI);
- linkage and retrieval of information from different sources, media, and domains;
- low-resource and historical language processing;
- modelling dialogue literary style for generation;
- profiling and authorship attribution;
- search for scientific or scholarly literature;
- work with linguistic variation and non-standard or historical use of language.

Information for authors

We invite papers on original, unpublished work in the topic areas of the workshop. We will consider long papers, short papers and system descriptions (demos). We also welcome position papers.

A long paper, which presents completed work, may have up to eight (8) pages of content plus additional pages of references (just two, please, if possible). The final camera-ready versions of accepted long papers will get one additional page of content (up to 9 pages), so that reviewers' comments can be taken into account.

A short paper or a demo, which presents work in progress or the description of a system, may have up to four (4) pages of content plus additional pages of references (one if you can, please). Upon acceptance, short papers will get five (5) content pages in the proceedings.

A position paper – clearly marked as such – should not exceed eight (8) pages including references.

All submissions are to follow the *ACL paper styles (for LaTeX / Overleaf and MS Word) available at https://github.com/acl-org/acl-style-files. Papers will be submitted electronically, only in PDF, via the SIGHUM 2027 submission website on the SoftConf pages – *we will post the link as soon as we have it.*

Reviewing will be double-blind. Please do not include the authors' names and affiliations, or any references to Web sites, project names, acknowledgments, and so on – anything that immediately reveals the authors' identity. Please keep references to your own work at a reasonable minimum, and do not include anonymous citations.

In general accordance with the Multiple Submission Policy at EACL 2027, we will not consider any paper simultaneously under review in a journal or at another conference.

During the review period, papers submitted to our workshop cannot also be submitted elsewhere.

The reviewers will look out for any signals of the use of AI in the writing of your paper. In particular, any hallucinated bibliographic reference means immediate rejection.

Non-archival poster presentation of ACL Findings papers

For non-archival poster presentation at the workshop, we will consider papers accepted as Findings of ACL. To qualify, a paper must align very closely with the scope of SIGHUM. We aim to bring together research relevant to language technologies for the Humanities and Cultural Heritage, otherwise distributed across ACL areas. That should foster exchange in our research community. We will base selection on the already reviewed papers' thematic relevance to the workshop and on the availability of (few) presentation slots. Our decisions will be final.