

# HUM INFRA



# Huminfra

## Annual Report 2023



SPRÅKBANKENTAL

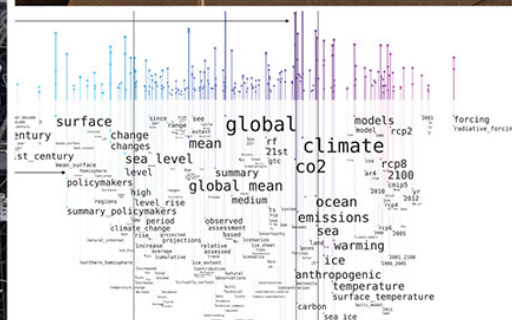


#GRIDH

SPRÅKBANKENTEXT



Humanities Lab  
JOINT FACULTIES OF HUMANITIES  
AND THEOLOGY



## Preface

**W**elcome to the second annual report of Huminfra, a Swedish national research infrastructure supporting digital and experimental research in the Humanities and beyond. Huminfra is a distributed infrastructure which unites 12 nodes across 11 universities and organisations.

Huminfra aims to provide users with a single entry point for finding existing Swedish materials, research tools and expertise via the web based information platform huminfra.se, to develop national training opportunities in the techniques available, and to prepare Swedish membership in a European network for cultural heritage data, DARIAH-EU. By bringing together expertise and instruments in a national consortium, Huminfra allows us to pool resources, strengthen collaborations nationally and internationally, and hopefully inspire and support new lines of research. This report provides an overview of achievements and progress made in Huminfra's second year.

We extend our thanks to the collaborating institutions and the Swedish Research Council for their support.

**Marianne Gullberg, Director of Huminfra**

Professor at the Centre for Languages and Literature,  
Lund University. Director of Lund University Humanities Lab.



## Summary

**H**uminfra is a distributed national infrastructure led by Lund University, which brings together 12 nodes in the domains of digital and experimental Humanities. In its second year, 2023, Huminfra launched [www.huminfra.se](http://www.huminfra.se), a website which collates and links to Swedish digital and experimental research resources (instruments, tools, expertise, materials, and training events) drawing on an inventory made at the nodes. It also links to other national and international infrastructures, and to ethics and data management expertise. This is the main feature of Huminfra. The consortium also held a number of training events, national and local. The work preparing for Swedish membership in *DARIAH-EU* also continued. Finally, new developments in external communication included the launch of a monthly Huminfra Newsletter to complement the social media accounts. Huminfra also organised a public webinar on the role of AI for research in the Humanities.



## Description of operations

### Time plan

The goals and deliverables outlined for Huminfra Y2/2023 have been well met. Despite initial delays in Y1 caused by slow recruitment and late appointments, the work in the infrastructure is on schedule with key deliverables provided.

### Construction, development and operation of the infrastructure

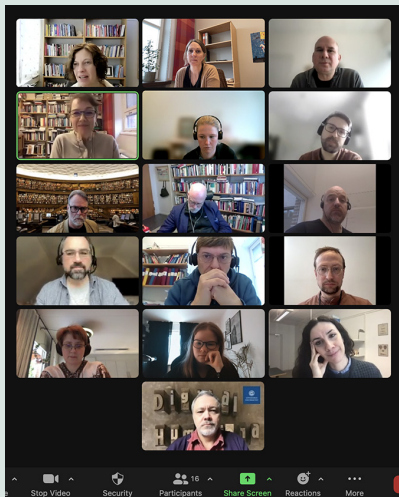
#### Module I Organisation and leadership

In Y2/2023, Huminfra held a range of *meetings* at different organisational levels (Operative management team: 11 virtual, 1 on site; Steering committee: 3 virtual, 1 on site; Node coordinators: 3 virtual; Consortium: 1 on site in Lund).

The members of the *Scientific Advisory Board* were appointed and met online in December 2023.

In February 2023, an *extension application* was submitted to the Swedish Research Council (SRC) for Huminfra phase 2, 2025–2028. A positive outcome was announced in September 2023. Since Huminfra was awarded a lower sum than applied for, Q4 involved a revision of activity plans and budget.

Preparatory work for *Swedish membership in DARIAH-EU* included discussions with the SRC (the National representing entity in the membership application) and DARIAH-EU management. As part of the preparations, two nodes (Linnaeus, Uppsala) became independent Cooperative partners and made their resources available on DARIAH-EU sites. The extension funding from the SRC included earmarked funds for national membership fees. An application for Swedish membership will therefore be submitted in the spring of 2024. If accepted, it will allow Huminfra to host DARIAH-SE for a five-year-period (the stipulated membership duration of DARIAH-EH), 2024–2028.



Huminfra's node coordinators met online three times during 2023 ...



... and the whole consortium was gathered for an onsite meeting in Lund in May.



Huminfra's Steering committee met in Lund in August.



The announcement that Huminfra was awarded extended funding for phase 2, 2025–2028, by the Swedish Research Council was

celebrated in September. The funding included earmarked funds for national membership in DARIAH-EU which meant that preparatory work for the membership application could begin in earnest.

3 300 unique users  
219 open & available resources  
170 new users per week

### 3 resource examples:

## The Swedish Lion I

... is a cutting-edge AI model for Swedish historical texts, capable of accurately interpreting handwritten texts in Swedish from the 17th, 18th, and 19th centuries, achieving an impressive accuracy rate of 95%. This innovative model has been trained using an extensive dataset of transcribed

texts, encompassing 3.3 million text lines and a total of 15.6 million words. This AI model is the result of a collaborative effort, involving the National Archives in both Sweden and Finland, in partnership with the Stockholm City Archives, Jämtlands läns fornskriftsällskap, citizen science volunteers and researchers from Stockholm and Uppsala Universities.

## Motion Capture

... is a technology that enables human bodily movements in 3D to be recorded with high spatial and temporal resolution. This is often used by researchers

studying gestures but also movements of objects and humans in a broader sense. The system exports 3D datasets in many formats. One option is to export a full body tracking as a file for direct animation of avatars in VR. During 2023 Lund University Humanities Lab collaborated with Riksteatern CREA in a project creating a real-time sign language avatar for the production "The Canterville Ghost" which will be on tour from January, 2025.

## Word Rain

... is a development of the classic word cloud. This development of a software library for text visualisation uses the semantic information encoded in a distributional semantics-based language model – reduced

into one dimension – to position the words along the x-axis. Thereby, the horizontal positioning of the words reflects semantic similarity. Font size is still used to signal word prominence, but this signal is supplemented with a bar chart, as well as with the position of the words on the y-axis.

Word Rain is a collaboration between Centre for Digital Humanities and Social Sciences, Uppsala (CDHU), the National Language Bank of Sweden/CLARIN Knowledge Centre for the Languages of Sweden at the Language Council of Sweden, and the iVis group at Linköping University.



## Module 2 Web-based information platform (WIP)

The beta version of Huminfra's *web-based information platform (WIP)*, *huminfra.se*, was launched in January 2023. During Q1-Q2, additional resources from the nodes and other relevant information was added. In June 2023 a complete, fully functional version of the WIP was released. The system developers now work on maintenance, continued development and improvements, including ensuring accessibility, and developing the search function. An expert group involving several of Huminfra's nodes started work on developing *controlled vocabulary* for the platform, work that will continue in Y3/2024.

The WIP is also used to collect information about activities at the nodes internally (so-called *node reports*) via a standardised web-based form, developed in 2023.

By December 2023, the WIP contained **219 open and available resources** at Huminfra's 12 nodes. More than **3,300 unique users** visited the platform in 2023, and the number steadily increases with approximately 170 new users per week.

## Module 3 Training (and other joint activities)

The training activities are reported below under the heading *Training, outreach and user support*.

Beyond the joint training activities, other important evidence of *collaboration* across the nodes includes nodes visiting other nodes for study purposes, planning meetings, and joint applications for research funding. The Swedish National Archives have also hosted two trainees linked to Huminfra, from the University of Gothenburg and Uppsala University.

The consortium also prepared for *Huminfra's first conference, HiC 2024*, to be held in Gothenburg in January 2024, hosted by Språkbanken Text and Gothenburg Research Infrastructure for Digital Humanities (GRIDH), both at the University of Gothenburg. By December 2023, **42 peer reviewed contributions** had been accepted, and proceedings were in preparation. The Linnaeus node also planned for a DARIAH Day point in the programme including a keynote by the DARIAH-EU President, and a panel discussion.

Joint work started on an (electronic) *Huminfra handbook*.

## Module 4 Outreach and communication

The outreach work is reported below under the heading *Training, outreach and user support*.

## Collaboration with other infrastructures

Huminfra works closely with a number of other infrastructures. These connections are important to Huminfra's goal to link and render visible relevant resources in one place in order to avoid information loss or duplication.

At the *international level*, Huminfra is indirectly linked to and participates in CLARIN ERIC through the membership in Swe-Clarín by several of the nodes. Individual nodes also have links to relevant international infrastructures (e.g., the Oxford Research Centre in the Humanities (TORCH), EUROPEANA, the iSchools Organisation, etc.).

At the *national level*, 4 national infrastructures are nodes in Huminfra (the *Swedish National Archives*, the *National Library of Sweden* via KBLab, *Språkbanken Text* and *Språkbanken Tal*), and Huminfra interacts with the *Swedish National Data Service* (SND) and *Nationella Språkbanken* since several nodes are members. Several nodes are also members of other national infrastructures such as *InfraVis* and *SweDigArch*. Individuals at local nodes also hold positions of trust in other national entities of relevance (e.g. *WASP-HS*, *AI Sweden* and the Swedish National Heritage Board (*Riksantikvarieämbetet*)). Links at the local level of the nodes are too numerous to list.

## Data management and supporting e-infrastructure

Since Huminfra does not in itself generate or handle research data, it does not require the constructions of common computational resources or resources for data storage but draws on local resources and computational networks. Huminfra only points researchers to local resources and facilities through the WIP which describes or indicates local instruments, tools, data sets, databases, registers, and archives with search engines (e.g., Swedish National Archives). Since Huminfra does not generate data and data are stored locally, local guidelines at individual nodes for data policy, and data management (plans) instead apply. Huminfra encourages and supports open data policies according to the FAIR principles, and open science policies more generally. Importantly, however, data types in the Humanities vary and may include personal or sensitive data requiring layered access. It is therefore important that access decisions are governed by the Swedish Ethical Review Authority. Guidelines concerning ethics and data management are available at SND.

## Changes in organisation

No changes were made in 2023 to the organisation relative to the funding decision. However, in May 2023 Mats Fridlund (Gothenburg Research Infrastructure in Digital Humanities, *GRIDH*, University of Gothenburg) replaced Cecilia Lindhé as a member of the Operative Management Team.

A Scientific Advisory Board (SAB) was appointed in September 2023. Members: Prof. **Steven Connor**, King's College, London; Dr. **Peter Leonard**, Stanford University; Assoc. Prof. **Marianne Ping Huang**, Aarhus University; Prof. **Asli Özyürek**, Max Planck Institute for Psycholinguistics. A first, virtual meeting was held in December 2023. A visit on site is planned for 2024.

## Steering group activities

Huminfra's Steering committee held 4 meetings (3 virtual, 1 on site) in Y2/2023. The January meeting focused on establishing routines, and the extension application submitted to the SRC in February. The March meeting focused on policy documents, the annual budget, and activities in Y2. A physical meeting in Lund in August focused on strategic issues connected to the extension application, and the November meeting targeted the revised budget and activity plan for phase 2, 2025–2028, the DARIAH application process, and budget issues.

## Operative Management Team



**Marianne Gullberg,**  
Director

Professor at the Centre for Languages and Literature, Lund University.  
Director of Lund University Humanities Lab.



**Coppélie Cocq,**  
Dep. Director

Professor at Humlab, Umeå University.  
Dep. Director of Humlab, Umeå University.



**Mats Fridlund**

Researcher and Associate Professor at the Department of Literature, History of Ideas and Religion, University of Gothenburg.  
Dep. Director of GRIDH, University of Gothenburg.



**Koralja Golub**

Professor at the Department of Cultural Sciences, Linnaeus University.  
Director of ilnstitute, Linnaeus University.

## Steering Committee



**Göran Blomqvist, Chair**

Former CEO of the Bank of Sweden's Tercentenary Foundation (Riksbankens Jubileumsfond).  
Corresponding member of The Royal Swedish Academy of Letters, History and Antiquities.



**Kerstin Lidén**

Professor at the Department of Archaeology and Classical Studies, Stockholm University.



**Mikko Tolonen**

Professor at the Department of Digital Humanities, University of Helsinki.



**Johanna Berg**

Project manager, The National Museums of World Culture.



**Stefan I. Lindberg**

Professor at the Department of Political Science, University of Gothenburg.



**Margareta Sollenberg**

Researcher at the Department of Peace and Conflict Research, Uppsala University.

## Scientific Advisory Board



**Steven Connor**

Professor of English, Director of Research of the Digital Futures Institute, King's College, London.



**Peter Leonard**

PhD, Assistant University Librarian for Research Data Services at Stanford Libraries, Stanford University.



**Marianne Ping Huang**

Associate Professor at the School of Communication and Culture, Aarhus University.



**Asli Özyürek**

Director of Multimodal Language Department of Max Planck Institute for Psycholinguistics.  
Professor at the Donders Institute for Brain, Cognition and Behavior.



## Financial performance

Huminfra accrued a financial surplus in Y1 due to delays in recruitment and restricted travel in 2022/ Y1 relative to the budget. Importantly, the activities for 2023 (joint training workshops, courses, the outreach events, travel to physical meetings (consortium and steering committee) better aligned the Y2/2023 budget with planned expenditure, but some nodes still had restricted activities in 2023 (KTH, SU). The activity plan for the entire consortium will help align expenditure to budget.

Note that Huminfra does not have any user fees. Access to Huminfra's WIP is free. Any costs associated with access to the local resources (instrumental or human) and user support indicated at the WIP are set by the local nodes themselves (cf. SFS 2022:1378). Insofar as local nodes apply user fees, they are handled locally and not within Huminfra itself.

## Key performance indicators

Huminfra's key performance indicators are linked to its users, defined as unique visits to the WIP, and as unique participants at training events organised by Huminfra.

Between January 1 and December 31, 2023, the WIP had **3,300 unique visitors** (defined as unique IP addresses). Of these, 512 had found the site via a search engine; 540 came from one of the nodes; 422 from social media sites; and the remainder (the majority) had typed the URL directly into their browser.

The number of participants in Huminfra events in 2023 amounts to **1,385**.

## Equality

Huminfra's equality plan is implemented and pursued in daily operations. Gender balance is observed in all units (the Operative management team 3 women/1 man; the Steering committee 3/3; the Scientific Advisory Board 2/2). The consortium has more men than women (4/12 node coordinators female). Although this is typical of the technology-driven part of the humanities, the consortium monitors the situation carefully. Further, gender balance and avoidance of gender stereotyping is observed in data compilations, evaluations, workshops, symposia, conferences, meetings, text, and images. No key performance indicators (KPIs) can be reported for users' gender, since users are defined only as unique IP numbers (as per the consortium agreement).

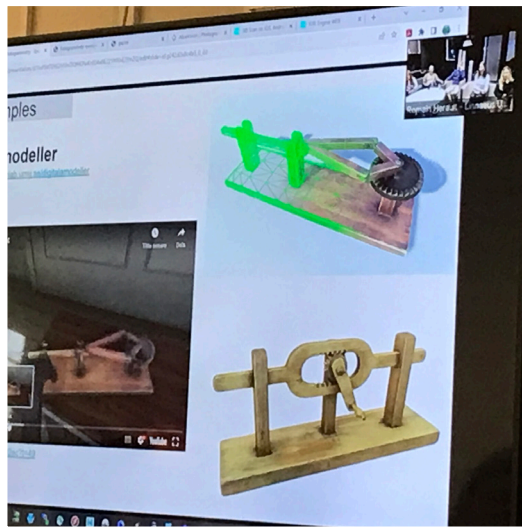
## Risk analysis

The extension application to the SRC 2023 yielded an updated risk analysis. Three main areas were identified: 1) Recruiting and keeping top talent. Employment conditions are carefully monitored and HR support engaged at the nodes. 2) Failed internal/external communication with nodes and (new) users, respectively. The communication plan helps identify strategic and adaptive solutions. 3) Maintenance of cutting-edge hard- and software, and expertise. Operational solutions already available at the nodes mitigate hard- and software risks, and the monitoring of recent method development is part of module activities to prevent knowledge stagnation.

## Training, outreach and user support

### Training (Module 3)

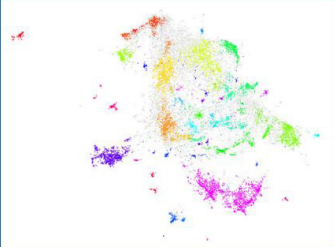
Training activities at Huminfra's nodes gained momentum in Y2/2023, with **59** educational activities and other events (courses, workshops, tutorials, seminars, demos, and talks) arranged for participants nationally under the label Huminfra. Several training events were organised jointly between two or more nodes in creative formats. For example, a workshop on photogrammetry was arranged simultaneously on-site at the Umeå and Linnaeus Universities, with joint discussions and knowledge exchange achieved through a hybrid solution. As of 2023, national courses will generally be multi-node events. A number of such events are already planned for Y3/2024, e.g. an online workshop on AI and the library (KBLab, Linnaeus, and Borås Universities) in January 2024.




Photogrammetry workshop arranged simultaneously at Linnaeus University and Humlab, Umeå University in October, 2023. Discussions and knowledge exchange were carried out through a hybrid solution.


←

## Topic modeling with BERTopic





**Introductory workshop with KBLab**



↑ In December, 2023, KBLab held a digital workshop within Huminfra on topic modeling using BERTopic. Forty-seven participants from a range of countries, and spanning academia, memory institutions and public authorities, received an introduction to how AI can help offer new thematic entrance points to large collections of text data.

Lund University Humanities Lab arranged an introductory workshop on experimental methods in the humanities, in October, 2023. Participants were given demos and hands-on sessions in a range of fields, such as motion capture, eye-tracking, language technology tools, and audio/video equipment.

→



In December, 2023, Humlab, Umeå University organised the conference 'Advancing Research Communication – the role of Humanities in the Digital Era'. The aim of the conference was to bring together researchers in the humanities and related disciplines to discuss the role of research communication in a contemporary world characterised by disinformation, online hate, the climate crisis and the rapid development of AI.

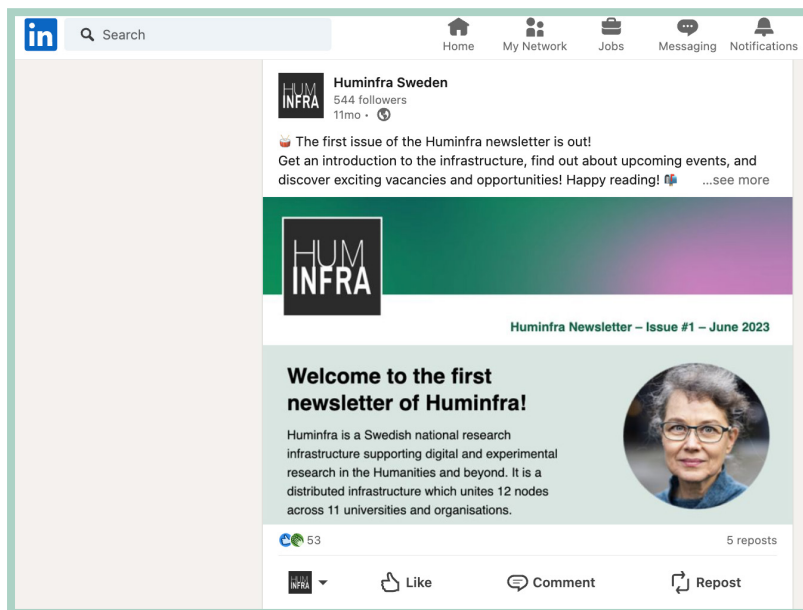
→



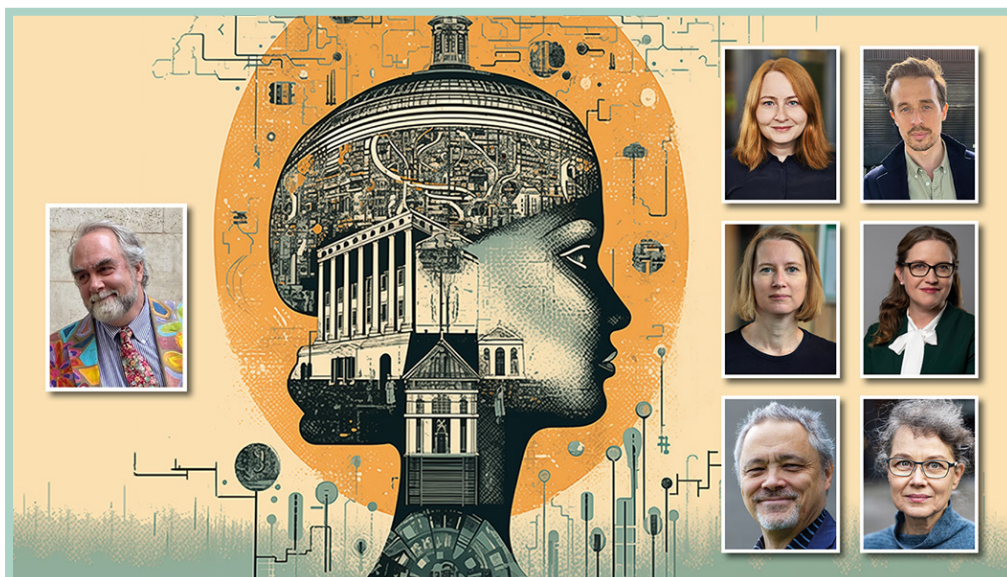


## Outreach (Module 4)

Outreach activities in 2023 largely involved the establishment and development of the WIP and the social media accounts. The launch of the full version of the WIP, huminfra.se, combined with the launch of Huminfra's monthly newsletter (both in June 2023) resulted in a visible increase in followers and engagement in Huminfra's social media, in particular on LinkedIn, where the number of followers grows steadily. By December 2023, the newsletter had more than **120 subscribers**. The newsletters also affect user statistics for huminfra.se with peaks in visitors after the publishing of each new edition. Huminfra was also presented in invited talks and demos locally at the nodes to university, faculty, departmental or organisational leadership, to researchers, students, to SND, and other potential user groups. In November 2023, Huminfra organised a public webinar on AI and research in the Humanities. The webinar consisted of a keynote by Prof. Geoffrey Rockwell, University of Alberta, and a subsequent panel discussion in Swedish between four researchers working in Swedish academia, cultural heritage institutions, and politics. Over **100** participants followed the webinar. Huminfra will continue to host such public webinars.



*Huminfra's monthly newsletter was launched in June, 2023. Through its compilation of news, events, and opportunities from Huminfra's 12 nodes, and collaborative partners, it has become an appreciated channel for keeping up-to-date with what's on in the domains of digital and experimental Humanities in Sweden.*



*The public webinar on 'Research in the Humanities in the wake of ChatGPT' which Huminfra organised in November, 2023, consisted of a keynote by Prof. Geoffrey Rockwell, University of Alberta, and a subsequent panel discussion between the researchers Lovisa Brännstedt, Eric Cullhed, Karin Danielsson, and Maja Fjaestad. Mats Fridlund moderated the panel, and Marianne Gullberg hosted the webinar.*



# HUM INFRA

Supporting **digital and  
experimental research**  
in the Humanities



[huminfra.se](https://huminfra.se)



[info@huminfra.se](mailto:info@huminfra.se)



Lund University Humanities Lab/Huminfra



+46 (0)46-222 30 91