



Image: Intellexus

Computational Methods for Idea and Concept Extraction and Matching in Ancient Manuscripts

September 18, 2025
Safed Academic College, Israel

Ancient manuscripts preserve a wealth of ideas and concepts, reflecting the intellectual, cultural, and historical dynamics of the societies that produced them. These texts often embody layers of borrowing, supplementation, and adaptation, evolving philosophical propositions, and intricate networks of scholars and communities. Advances in natural language processing (NLP) enable us now to computationally trace these ideas and concepts across large corpora, uncovering connections and patterns that are difficult to detect through traditional methods alone. By combining philological expertise with computational tools, we can explore the geology of texts, the genealogy of ideas and concepts, and the intellectual ecosystems that shaped them.

This workshop will focus on innovative computational approaches to extracting and matching ideas and concepts in ancient manuscripts, with a particular interest in low-resource languages and crosslingual settings. Case studies may include, for example, Buddhist Sanskrit and Classical Tibetan corpora, but contributions are welcome from any cultural or textual tradition. Topics include crosslingual semantic matching, concept alignment,



Image: Intellexus

network visualisation, and interdisciplinary methods integrating historical, linguistic, and computational perspectives. The goal is to bring together philologists, computer scientists, and digital humanists to advance large-scale, computationally supported research on ancient intellectual traditions.

The workshop will take place on September 18, 10:00–15:30, in Safed (Israel) as part of the Intellexus Annual Nexus 2025, a 5-day hackathon where teams of UHH and RUNI researchers collaborate to tackle computational challenges in Sanskrit and Tibetan.

Organizers: Orna Almogi (University of Hamburg), Harunaga Isaacson (University of Hamburg), Dorji Wangchuk (University of Hamburg), Nachum Dershowitz (Tel Aviv University), Shai Fine (Reichman University), Kfir Bar (Reichman University), Orr Zwebner (Reichman University)

Program

10:30 - Coffee + Gathering

11:00 - Workshop part A:

Dr. Orna Almogi - Opening

Mr. Yehonatan Nachshoni - *The concept of “Ein Sof”*

Mr. Kai Golan - *MWE and Idiom Identification Using LLMs*

Dr. Hallel Baitner - *A Computational Approach to Qumran Hebrew: A Comparative Analysis of Its Linguistic Position in the History of Hebrew*

Prof. Nachum Dershowitz - *Finding nonparallel texts in classical Hebrew anthologies*

13:00 - Lunch (Safed Academy cafeteria)

14:00 - Workshop part B:

Prof. Dorji Wangchuk - *Genealogy of Buddhist Concepts: Selected Cases*

Dr. Kfir Bar - *Breaking Down Documents: Embedding Text Chunks for Deeper Understanding*