

Analysing Cultural Heritage Documents

HTR/OCR, Information Extraction and Textual Variation

9 October 2026

TOURS | CCSR, University of Tours

INTERNATIONAL WORKSHOP | Salle Rapin

9h-9h30 **Opening presentation**

9h30-10h15 **Session 1: Language processing for historical documents**

Numériser le breton historique : l'océrisation par VLM entre gains massifs et limites comportementales

Morgane BONA-PELISSIER (Université Paris Nanterre)

RCnum: A Semantic and Multilingual Online Edition of the Geneva Council Registers from 1545 to 1550

Pierrette BOUILLON, Christophe CHALAZON, Sandra CORAM-MEKKEY, Gilles FALQUET, Mathilde FONTANET, Johanna GERLACH, Stéphane MARCHAND-MAILLET, Laurent MOCCOZET, Jonathan MUTAL, Hélène de RIBAUPIERRE, Raphael RUBINO, Marco SORBI (University of Geneva)

COFFEE BREAK

11h-13h **Session 2: ATR, methodology and challenges**

Unlocking a New Data Source: Digitizing Manuscript Medical Records

Simon GABAY, Justine FALCIOLA, Laurence TOUTOUS TRELUS, Myriam LAMRAYAH (University of Geneva)

From Tool to Infrastructure: The Digital Servitization of ATR with eScriptorium

Chiara AIOLA (Net7 Srl), Nicola BAGLINI (Net7 Srl), Federico BOSCHETTI (CNR-ILC), Angelo Mario DEL GROSSO (CNR-ILC), Luca DE SANTIS (Net7 Srl), Monica MONACHINI (CNR-ILC)

Analysing Cultural Heritage Documents: HTR/OCR, Information Extraction, and Textual Variation

Liri SOKOL (University of Jerusalem)

Opening HTR to Old Swedish Manuscripts: Methodology and Challenges

Paola PERATELLO, Adonis Dulaurent, Caterina Carli (École nationale des chartes-PSL)

Valorizing and indexing: some problems in transcribing finding aids of the Alinari documentation

Pamela KRZEMIEN (IMT - Scuola Alti Studi Lucca)

13h-14h

LUNCH BREAK

14h-15h30

Session 3: Computational philology

Mapping the Textual Fluidity of Rashi's Torah Commentary: HTR Transcription and Two-Tier Comparison Across Penkower-Selected Witnesses

Avichai LEVY (EPHE-PSL and Tel Aviv University)

De l'HTR à la collation : chaîne semi-automatique pour le projet d'édition critique numérique CreNum

Théo ROULET (Université de Grenoble)

Projet d'édition numérique des registres de comités de la Comédie-Française

Marine TIGER (Sorbonne Université)

Fabliaux Alignment Protocol. From the Integration of Semi-Automatic Methods into Philological Workflow to its Valorization

Jules NUGUET (ENS de Lyon), Laura BONANNO (Université Jean Moulin Lyon3- CIHAM UMR 5648)

COFFEE BREAK

16h-17h30

Session 4: Information extraction and data mining

Extraction et structuration automatique de rapports de fouilles archéologiques : pipeline local pour l'import dans Omeka-S

Valentin COTTA (Université Paris 1 Panthéon-Sorbonne)

Self-Supervised Pre-Training for Join Retrieval of Cairo Genizah Manuscripts

Gal GRUDKA-JUSTER, Sharva GOGAWALE, Omer VENTURA, Berat KURAR-BARAKAT, Daria VASYUTINSKY-SHAPIRA, Nachum DERSHOWITZ (Tel Aviv University)

ICDAR 2026 FalsID Competition on Falsification and Imitation Detection

Dominique STUTZMANN
(L'Institut de recherche et d'histoire des textes-CNRS)



Grant n° 101142242



Centre d'études supérieures de la Renaissance
59 rue Néricault-Destouches, 37020 Tours Cedex 1
Tel : 02 47 36 77 61 - cesr.cnrs.fr

Organising committee

Elena Pierazzo (CESR – University of Tours)

Serena Carlamaria Crespi (CESR – University of Tours)

Carlos-Emiliano González-Gallardo (LIFAT – University of Tours)

Thierry Brouard (LIFAT – University of Tours)



Registration required

prima.hypotheses.org

Contact : erc.prima@univ-tours.fr