

# Revue d'Etudes Tibétaines

**From Print to Pixels:  
Building Digital Tools for Modern Tibetan Textual  
Analysis**

Edited by  
Franz Xaver Erhard, Robert Barnett,  
and Nathan W. Hill



numéro soixante quatorze — February 2025

# Revue d'Etudes Tibétaines

numéro soixante quatorze — February 2025

ISSN 1768-2959

**Directeur** : Jean-Luc Achard.

**Comité de rédaction** : Alice Travers, Charles Ramble, Marianne Ginalski, Jean-Luc Achard.

**Comité de lecture** : Ester Bianchi (Università degli Studi di Perugia), Fabienne Jagou (EFEO), Per Kværne (University of Oslo), Rob Mayer (Oriental Institute, University of Oxford), Fernand Meyer (CNRS-EPHE), Françoise Pommaret (CNRS), Ramon Prats (Universitat Pompeu Fabra, Barcelona), Charles Ramble (EPHE, CNRS), Françoise Robin (INALCO), Alice Travers (CNRS), Jean-Luc Achard (CNRS).

**Périodicité**: La périodicité de la *Revue d'Etudes Tibétaines* est généralement bi-annuelle, les mois de parution étant, sauf indication contraire, Octobre et Avril. Les contributions doivent parvenir au moins six (6) mois à l'avance. Les dates de proposition d'articles au comité de lecture sont Novembre pour une parution en Avril, et Mai pour une parution en Octobre.

**Participation**: La participation est ouverte aux membres statutaires des équipes CNRS, à leurs membres associés, aux doctorants et aux chercheurs non-affiliés. Les articles et autres contributions sont proposés aux membres du comité de lecture et sont soumis à l'approbation des membres du comité de rédaction. Les articles et autres contributions doivent être inédits ou leur réédition doit être justifiée et soumise à l'approbation des membres du comité de lecture. Les documents doivent parvenir sous la forme de fichiers Word, envoyés à l'adresse du directeur (jeanluc.achard@sfr.fr).

**Comptes-rendus**: Contacter le directeur de publication, à l'adresse électronique suivante : jeanluc.achard@sfr.fr

**Langues** Les langues acceptées dans la revue sont le français, l'anglais, l'allemand, l'italien, l'espagnol, le tibétain et le chinois.

La *Revue d'Etudes Tibétaines* est publiée par l'UMR 8155 du CNRS (CRCAO), Paris, dirigée par Matthias Hayek.

**Hébergement**: <http://www.digitalhimalaya.com/collections/journals/ret/>





# Revue d'Études Tibétaines

*numéro soixante quatorze — February 2025*

## **From Print to Pixels: Building Digital Tools for Modern Tibetan Textual Analysis**

Edited by

Franz Xaver Erhard, Robert Barnett,  
and Nathan W. Hill

### **Robert Barnett and James Engels**

Assembling a Digital Toolkit: An Introduction to Text-  
mining for Modern Tibetan

pp. 5–43

### **Franz Xaver Erhard**

The Divergent Discourses Corpus: A Digital Collection  
of Early Tibetan Newspapers of the 1950s and 1960s

pp. 44–80

### **Christina Sabbagh**

Enhanced ATR Accuracy for Tibetan Historical Texts:  
Optimising Image Pre-processing for Improved  
Transcription Quality

pp. 81–127

### **Franz Xaver Erhard**

Text and Layout Recognition for Tibetan Newspapers  
with Transkribus

pp. 128–171

- Franz Xaver Erhard und Xiaoying 笑影**  
Foreign Names and Places in Tibetan Newspapers  
of the 1950s and 1960s pp. 172–186
- Yuki Kyogoku, Franz Xaver Erhard, James Engels, and  
Robert Barnett**  
Leveraging Large Language Models in Low-resourced  
Language NLP: A spaCy Implementation for Modern  
Tibetan pp. 187–220
- Ronald Schwartz and Robert Barnett**  
Religious Policy in the TAR, 2014–24: Topic Modelling  
a Tibetan-Language Corpus with BERTopic pp. 221–260
- James Engels and Robert Barnett**  
Developing a Semantic Search Engine for Modern  
Tibetan pp. 261–283
- Natalia Mikhailova**  
*The Tibet Mirror*, “Friends” of Tibet, and the Inter-  
nationalisation of the Tibet Question pp. 284–328

