



Hackathon in the Arava

February 14–18, 2016
Kibbutz Lotan, Israel

Organizers

- Orna Almogi (Hamburg University, Indian & Tibetan Studies)
- Nachum Dershowitz (Tel Aviv University; Paris Institute for Advanced Study)
- Dorji Wangchuk (Hamburg University, Asia-Africa Institute)
- Lior Wolf (Tel Aviv University, Computer Science)

The objective of the hackathon is to develop digital tools that will facilitate the identification of shared passages—passages found in two or more texts as a result of borrowing or citing—and thus enable the comprehensive study of the evolution of canonical Buddhist corpora, including scriptures, their commentaries, and related treatises. The specific corpus to be studied is the Tibetan Buddhist canon, consisting mostly of Indic Buddhist literature in Tibetan translation. This corpus was formed over a period of more than a thousand years and contains various layers of materials; its translation into Tibetan spanned several centuries.

We will probe methods of finding inexact quotations and borrowed texts within the corpus, simultaneously comparing numerous different textual units, with the aim to further our apprehension of the evolution and

emergence of individual texts. Moreover, the tools also aim at allowing a better understanding of the processes of translation and revision of translation, and of editorial policies as well. Particular challenges include the monosyllabic nature of Tibetan, the omission or addition of grammatical particles without changing the meaning, varying orthographies, and the plethora of homophones, which must all be taken into consideration.

The importance of the planned application lies first and foremost in that it will allow scholars to uncover details of the emergence and history of transmission of individual texts and textual corpora and to better understand the intellectual history of Buddhism.

The intertextual tools developed for Tibetan should have wide applicability to other languages and corpora.