



Buddhist Benchmarking & Evaluation Webinar

JUL 31 Friday, July 31
5:00 PM - 7:00 PM GMT+3

Zoom

You have manage access for this event.

Manage

Hosted By

- Chris Scammell
- Jake Moore
- Kai Golan

93 Going



Sonam Dorje Dolma, mitsu and 91 others

Contact the Host

AI

LIVE

You're In

Join Event Invite a Friend

No longer able to attend? Notify the host by [canceling your registration](#).

Get Ready for the Event
Profile Complete · Reminder: SMS & Email

About Event

How good are LLMs at Buddhist translation? Can an LLM help with philology research? How do you measure "wisdom" in LLMs?

Measuring how good an AI is at any task requires benchmarking and evaluation. This webinar provides an overview of how these processes work, and uses Tibetan-English translation as a working example to explore how to benchmark LLMs with a concrete use case.

Though this will be a more technical presentation, the session is open to everyone - academics, philosophers, computer scientists, and general

enthusiasts. *We especially welcome those who'd like to get involved in collaborative benchmarking, as a discussion of possible next steps will follow the talks.*

Programme (July 31, 7:00am PST)

- **Evaluation methodology** (45 min) — Kai Golan Hashiloni (Intellexus): how to evaluate LLMs in general.
- **Break** (10 min)
- **Translation evaluation** (45 min) — Jake Moore (Khyentse Vision Project): a specific case of assessing translation quality.
- **Q&A and Possible Breakout Rooms** (20 min)

Moderated by Chris Scammell, Buddhism & AI Initiative.