Tibetan AI/NLP Workshop

Summary

We have just concluded the April 16-19 Tibetan Translation and AI workshop, held in Padua, Italy. The conversation was facilitated by Mind & Life Europe (MLE), Italian Buddhist Union Research Center (UBI CS), and the Buddhism & AI Initiative, with generous financial support from Khyentse Foundation.

Bottom line up front: we think this event was a great success, both for community-building and for coming up with concrete next steps to move the (Tibetan/Buddhist/AI/NLP) ecosystem forward on benchmarking, data, tooling, training, and shared principles. Participants expressed that they were *"shocked to the degree of consensus,"* able to *"really feel the benefit of openness more,"* and recognize that *"all of this [the technical projects this community faces] is so solvable – we're not trying to reinvent the wheel here, just make sure it fits together."*

This document is an overview of the workshop, written to serve as a record of dialogue and hopefully to guide someone who wasn't at the workshop through the same discussion and considerations that were followed in person. It includes:

- **Summary:** An overview of details and conversations at the workshop
- **Immediate Next Steps:** Action items and projects that will begin immediately, and how to get involved
- **Discussion:** A theme-by-theme summary of discussion topics with relevant pictures and charts from the workshop
- **Possible Projects and Next Steps:** Full "next step" details including longer term projects worth considering
- **Links:** Links to raw notes, meeting recordings, relevant documents, and any other materials from the workshop

Participants

Approximately 30 people participated in the workshop, in person or remotely. Participants represented a cross-section of the field, including:

- Several major translation and digitization organizations working on Buddhist textual material
- Funders and foundations actively supporting the space
- Academic and research centers (Buddhist studies, computational linguistics, philology)
- Independent technologists and AI/NLP practitioners working on Tibetan-language tooling
- Translators, scholars, and contemplative practitioners
- Publishers focused on Buddhist material

To preserve participant privacy, individuals and their affiliations are not named in this version of the document. Throughout the writeup, contributors are referred to generically (e.g., "one participant," "a translation-platform team," "a funding organization").

Schedule

The schedule below provides an overview of topics and cadence. Sessions were a mix of remote-friendly workshops, in-person workshops, and unconference sessions.

Remote-friendly workshop | In-person workshops | Unconference sessions

<table>
<tr><th></th><th>Thursday April 16</th><th>Friday April 17</th><th>Saturday April 18</th><th>Sunday April 19</th></tr>
<tr><td>07:00 - 09:00</td><td rowspan="3">Arrival and checking into rooms</td><td>Breakfast</td><td>Breakfast</td><td>Breakfast</td></tr>
<tr><td>09:00 - 12:00</td><td>Workshop 2.</td><td>Workshop 3.</td><td>Workshop 4.</td></tr>
<tr><td>12:00 - 14:00</td><td>Lunch</td><td>Lunch</td><td>Closing / Lunch</td></tr>
<tr><td>14:00 - 15:30</td><td rowspan="3">Optional welcome aperitif and hot pools</td><td>Unconference 1.</td><td>Unconference 3.</td><td rowspan="5">Departure

Optional Venice trip</td></tr>
<tr><td>15:30 - 16:00</td><td>Break</td><td>Break</td></tr>
<tr><td>16:00 - 17:30</td><td>Unconference 2.</td><td>Unconference 4.</td></tr>
<tr><td>17:30 - 19:30</td><td>Workshop 1.</td><td rowspan="2">Trip to Padua city center and dinner in town</td><td>Break and hot pools</td></tr>
<tr><td>19:30 - 21:00</td><td>Dinner</td><td>Dinner</td></tr>
</table>

Day 1

Workshop 1: The opening discussion focused on existential questions for the ecosystem such as “Why do we still need translators, scholars, teachers?”, and crowdsourced answers to questions around ecosystem health such as “What is lacking this field (AI and translation), and what is going well?” This discussion set the stage for the rest of the workshop by highlighting the need for collaboration, sharing resources, and building technology that could support rather than erode Buddhist transmission.

Day 2

Workshop 2: The first full day jumped straight into technical conversation, featuring an overview of the updated Inventory Document to align on the technologies and research questions in scope for discussion. Participants then broke out into smaller groups to brainstorm possible new projects that could benefit the ecosystem–a “blue sky” idea generation session that began to surface projects like benchmarking, training programs, and data collection opportunities that became further refined over the course of the workshop.

Unconference 1: The first participant-led unconference was a “show and tell” session featuring a participant’s Translation Agent Pipeline and another organization’s research benchmarking Tibetan translation.

Unconference 2: The second unconference featured two sessions, the first focused on a participant's work on a digital archive of teachings and implications for the ecosystem; and a second conversation on the existential and philosophical concerns of the ecosystem.

Day 3

Workshop 3: The second full day began with a plenary discussion on questions relating to how open the ecosystem should be with data sharing with respect to (1) each other, by sharing data freely (2) the public, by completely open-sourcing all data, and (3) Big Tech, by actively pushing for partnerships to get Buddhist data into AI models. The discussion that followed highlighted the difficult tension of the need for the ecosystem to maintain sovereignty over its data and tools, while at the same time needing to position itself to benefit from frontier technology. The second half of the conversation split participants into small groups to discuss benchmarks, data, tools/product, and ecosystem sustainability (e.g. funding, GPU access, technical training).

Unconference 3: The afternoon unconference featured another "show and tell" featuring an OCR benchmarking project, a translation platform's CAT tool, a collaborative project on translating the Bodhicaryāvatāra, and a text alignment app.

Unconference 4: Participants chose to stay as a larger group to further refine the four discussion topics from the morning, spending twenty minutes on each subject to detail out possible projects and next steps. That evening, the notes from each conversation track were cleaned up into project proposals with suggested next steps and points of contact.

Day 4

Workshop 4: The final morning began with three breakout groups discussing projects around (1) Benchmarking and Data, (2) Modular Technology, and (3) Community Norms and Whitepaper. Groups refined next steps into clearer action items, and returned to share them as a large group before closing with individual reflections on takeaways from the event.

Immediate Next Steps

Immediate follow-up action items from the workshop are below. Anyone who wishes to be included in follow-up communications should contact the workshop coordinators (contact details available on request).

- **WhatsApp Group |** Contacts have been shared with participants of the event. Ongoing dialogue will continue in a WhatsApp group for anyone to join.
- **Meetings to progress projects/topics:**
    - **Benchmarking |** Two participants will coordinate a webinar on how to do benchmarking well. This call will serve as the basis for future coordination on benchmarking/evaluation projects.
    - **Trainings for Buddhist Studies scholars and students |** There is a summer school opportunity, among other training options, to teach AI skills to Buddhist studies experts and students. Several participants are interested in following up on a call for thinking about potential trainings with universities. The suggestion is that this group has a first meeting to figure out next steps and think about how to include non-conference attendees after that.

    - **Manifesto** | A small group will lead a follow-up meeting to discuss whether or not to release an ecosystem manifesto, and what content should/shouldn't be included.
- **Mini Projects:** (no big-group meeting needed before moving forward)
    - **Data List** | Two participants will kick off a small effort on data mapping, extending the Inventory Document. Will loop in other participants as needed.
    - **Documentation of Translation Tool Workflows** | A participant (with close support from a translation-platform team) will put together a map of the translation workflow and how various tools/APIs fit together, to get a picture of modularity (how well different organizations' technologies can fit together). Several other organizations to be consulted.
    - **Norm Shift** | For everyone in the ecosystem: consider the points on open-sourcing, APIs, show-and-tell, etc. and make adjustments to your approach as you see fit. A list of points that were emphasized in the workshop is included below.

Discussion

The below is an overview of discussion points at the workshop in rough chronological order.

Why do we still need translators, scholars, teachers? What is the role of a human in this field in the age of AI?

The power of AI systems and their use in many of the tasks relevant to this field (translation, philology, even writing Buddhist dharma talks) calls into question the role of a human vis-à-vis AI.

Participants were quick to note the ineffable qualities of transmission that go beyond the level of text–the difference felt from in-person interaction with a teacher, meditating in community, or the atmosphere of a monastery/retreat setting. At the same time, historical Buddhist teachings have acknowledged that there are multiple valid forms of teachers, including texts featuring words of the Buddha but also post-canonical texts, appearances such as experience and what arises from it, and ultimate reality itself. Practically, AI is being used widely in the Buddhist context, not just for translation but also by teachers for writing dharma talks, analyzing texts and commentaries, and more. Beyond AI, technology is increasingly supporting Buddhism, with senior teachers conducting empowerments over videocall.

Some participants could imagine a time in the future without the need for human translation: if AI was good enough at producing the exact same translation as a human, what would be the difference? Others felt that human translators were critical to ensure that continuity of a mindstream where transmission could be conveyed–taking a human out of the process could undermine the point and power of translation. Others still acknowledged that translation arose from a need–two people who do not share a language–and believed there may be a time in the future where this issue has been overcome and the need for translation at all does not exist. Translators are also critical in building a Buddhist culture that will ensure the longevity of transmission.

Through all of this, however, it was Buddhism's aim to alleviate suffering (via human transformation, or healthy human cultural forms and institutions) that was emphasized as the "end goal," and AI does not shift this.

What challenges does AI pose to Buddhism and this ecosystem?

The philosophical questions around what is and is not possible with technology are complemented by a set of challenging practical realities for the ecosystem: AI is getting increasingly capable of most human tasks, Big Tech is pouring hundreds of billions of dollars into training these models, and the Buddhist community is only just beginning to wrap its head around this. The people in the workshop and a small number of other folks in the ecosystem not present represent the majority of contributors to this field, at least as it relates to Tibetan Buddhism, though takeup isn't much further in other traditions.

Practical Risks: The rapid advance of this technology threatens to disrupt the ecosystem with very real consequences on Buddhism and on human lives: the translators whose jobs may be displaced; scholars whose work is being used largely without consent in training these systems; Tibetan diaspora community members (from Buddhist teachers to data annotators to technologists) who leave because there are increasingly few opportunities to build a life within the ecosystem. And for the wider general population of Buddhists and non-Buddhists alike, AI may become one of the primary ways people access information on Buddhism and meditation, yet there has been little assessment of how reliable or supportive it is.

Dharmic Risks: Some voiced that dharma is potentially at risk of being "flattened" as technology secularizes Buddhism and further divorces language about Buddhism from the lived experience of it. As one participant put it:

> *"We've seen over the last 150–200 years the process of European scholars of Buddhism interpreting Buddhism, and this going back to Asia and then shifting how Buddhists see their tradition. I have a concern around AI and the secular mindset it's coming out of from Big Tech as a continuity of this kind of thing. The goal for them is to make money, and perhaps to control humans, but this is not the Buddhist goal. Through these systems, all Buddhist literature starts to sound the same, and the texture of different traditions is lost. The flattening is produced by a worldview that is not fundamentally Buddhist."*

Participants have already experienced things like learning their translations have been used to train AI models without consent (which is hurtful, but moreover disconnects the mindstream of transmission without making a link to it). Another type of flattening could be caused by dogmatic Buddhist sects using AI to proliferate themselves without concern for other traditions. There is also of course the philosophical question of what we allow to be a container for dharma. As one participant from a translation-platform team asked:

> *"In terms of big philosophical questions along the lines of transmission from the perspective of Dharma transmission, what boundaries are we setting in terms of beings who can transmit dharma? Other beings besides humans theoretically could transmit dharma, no? Could AI beings meet necessary qualifications for transmitting authentically dharma?"*

Societal Risks: Conversations throughout the workshop also emphasized risks outside of the Buddhist space: some participants were concerned with the possible risk of extinction from advanced AI, others with the risk to human meaning-making, others with the risk of authoritarian uses of AI.

What opportunities exist? How can Buddhist teachings flourish in the age of AI?

While AI poses significant challenges to dharma and this ecosystem, it also poses great opportunities–which is why this meeting was convened between technologists, scholars, and practitioners.

The direct work of this ecosystem: One of the greatest successes of the workshop was the shared sense of camaraderie that comes from meeting face-to-face, particularly between individuals dedicated to such a closely shared cause. We discuss in significant detail the possible projects from this ecosystem and how they could support Buddhism in the age of AI more broadly below–and use the rest of this subsection to talk about other approaches slightly beyond the focus of the workshop.

Healthy technology: Significant inspiration was found in the question of developing truly "Healthy Technology." By approaching technology through an ecology of attention, we can evaluate positive uses of AI and other technology already: things like video calls which support meditators sitting together, AI that can help conduct innovative research (one participant shared that a student wrote an exemplary thesis on the intersection of Zen practice and embodied cognition with AI support), or AI that can help Buddhist teachers better understand texts (one participant shared an anecdote of senior monastics using AI to analyze texts, coming away saying "oh, now I really understand this!"). Participants noted that, just as people are willing to pay more money for handmade pottery, there will may continue to be interested in human translation works, or at the least, technologies that sacrifice quality to preserve the most essential human elements (one participant mentioned that "just because we're delivering things digitally doesn't mean we need to be sucking our users into this" and suggested that any AI tech around Buddhism could try to do things like linking audio sound bites from the source). Considering the "end user," conversation focused on the idea that ultimately, transformation is the goal–the technology we are building should ultimately circle back to support realization and liberation.

Use of narratives: Given the extreme and diverse views about how AI will play out in the future, some participants suggested that we ought to focus on the present conditions and not worry so much about the larger arc of the future. Others, however, noted that the future is largely conditioned on these narratives and visions and that creating healthy ones is critical–and importantly something that Buddhists can contribute to! In all scenarios of risk, Buddhism has a role in addressing the alleviation of suffering. But importantly, Buddhists should be engaged in thinking about what a bright future looks like and how to get there.

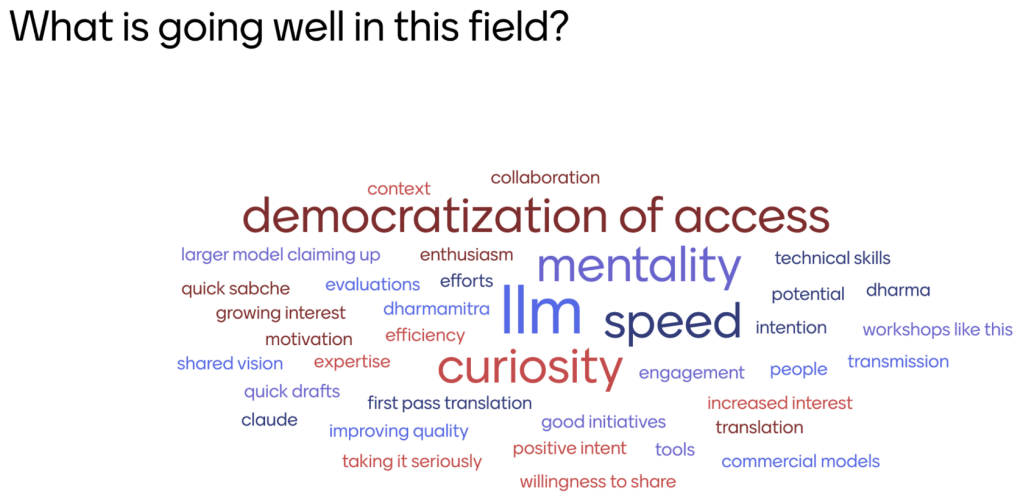
Lab and policy influence: As AI conversations grow in public awareness, there are increasing opportunities for Buddhists to engage in multifaith dialogue, and also interdisciplinary dialogue with AI labs and policymakers. In order to make the most of these opportunities, (at least some) Buddhists need to be engaged with technology and aware of how Buddhist ideas can help shape frontier technology and regulation.

What is going well in this ecosystem? What is going poorly?

We conducted some quick crowdsourcing to generate a word cloud of first impressions to these questions. Here is a visual of what folks think this ecosystem is lacking:



And here is a visual of what folks think this ecosystem is doing well:



Unsurprisingly, the final question of "What should be discussed in the workshop" mirrored these themes and set a strong agenda for the conversations ahead rooted in collaboration.



At the end of the first workshop, three short presentations underscored these points of collaboration: one on a collaborative data digitization process between three organizations, one on a participant's process for building a translation platform meant for the whole ecosystem, and one offering an overview of the Technical Inventory that had been put together for this workshop.

A Message from a Senior Contributor

A major contributor and pioneer to the field of AI and Buddhism wasn't able to attend the workshop, but they sent through helpful notes to guide conversations which are repeated here:

Identify the target audience for our tools and data: *I've had prolonged debates about tools only to learn late in the conversation that one person is talking about serving "all sentient beings" and another is addressing "academic scholars." It seems unlikely one system will serve all users, so it seems important to identify your target audience.*

Be specific and quantitative: *Generalizations about AI abound, yet every frontier model is different and has a variety of workflows. For example, with respect to translation capacity: comments like "AI is unable to translate and edit more than a few pages of text at a time" need to be specified relative to model and workflow. I have been able to do a draft translation of 25+ pecha pages at a time using a recent agentic coding tool. It's a recent approach for me, but I haven't hit a limit yet. I'm not saying this solves the problem for everyone. I'm simply saying that it's another data point. With respect to OCR quality: it is good to get specific word error rates of different systems on different scripts.*

What are our implicit assumptions about the translation process? *What's the implicit translation process that we're presuming, where are we spending most of our time, and which step in that process are our tools aimed to best address? I don't think the particulars of any given model are as important as having a common vocabulary.*

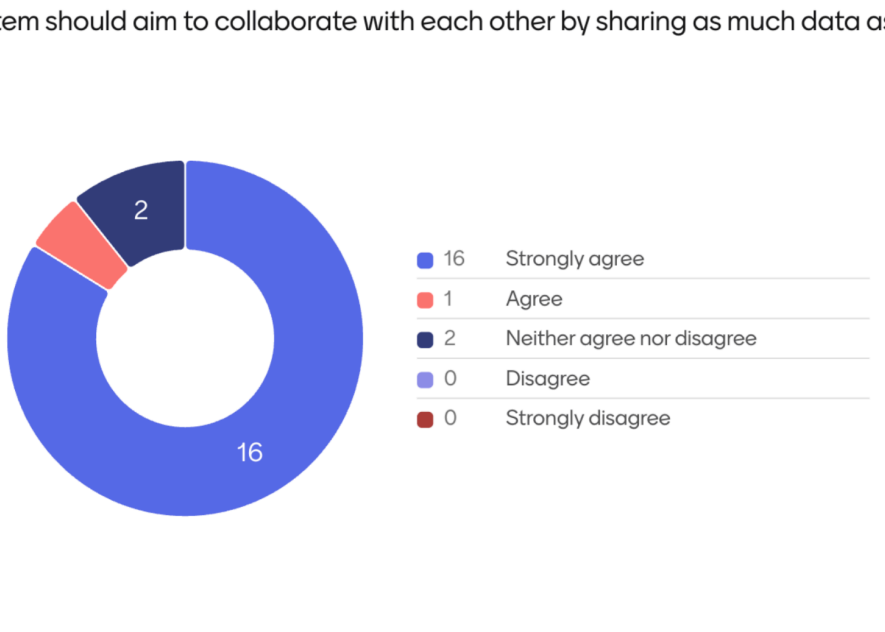
What's really working now? *Tibetan studies has a long history of ambitious projects that fail to materialize. Visions of future systems are wonderful, but let's be sure to identify what's really working, who is creating it, and how are people using it.*

How “open” should this ecosystem be with its data and tools?

By the midway point in the workshop when we began to discuss this question directly, there was already a strong shared understanding that openness and collaboration in the ecosystem are essential. But we wanted to be sure, so we conducted a poll. Each question asked participants to select how much they agreed with a statement on a scale of 1-5.

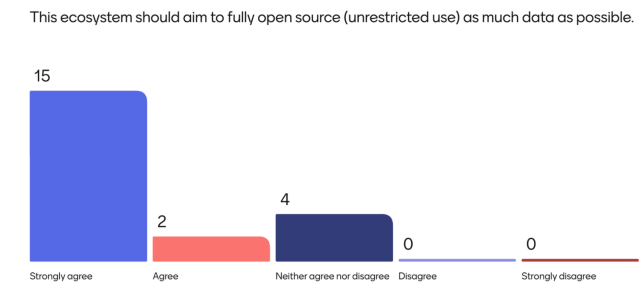
Below, we show the results for each question from *before* and *after* group discussion. I.e., we polled first before any dialogue, then spoke for 90 minutes as a large group, and then voted a second time.

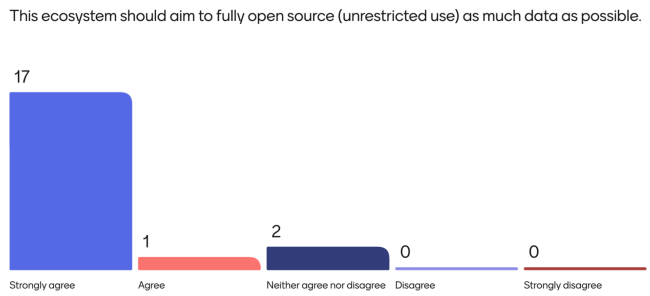
This ecosystem should aim to collaborate with each other by sharing as much data as possible.



This ecosystem should aim to fully open source (unrestricted use) as much data as possible.

Before Dialogue | *After Dialogue*





Votes overall shifted slightly towards "strongly agree" after dialogue.

The results showed near-unanimous support for collaborating with each other and the idea to "open source (unrestricted use) as much data as possible." Note that this threshold is higher than just pairwise or community-level sharing; the sentiment in the room was that an open ecosystem isn't just helpful for the existing participants in the field, but can create easier pathways for newcomers to begin contributing to projects.

To quote one participant's memorable framing:

> *"This is the sociological problem that a meeting like this should solve. Just put ALL the data on Github! Maybe it's a mess. Maybe someone will steal it and write something before me. So what?! We should have a bodhisattva vow of benefiting all beings with our digital stuff. And anyways, nobody, just nobody is prowling the internet looking for data to steal. The level of fear may be justified in nuclear physics where these things can happen, but in Buddhology they just won't!"*

Current blockers to open-sourcing and sharing that came up throughout the workshop are as follows:

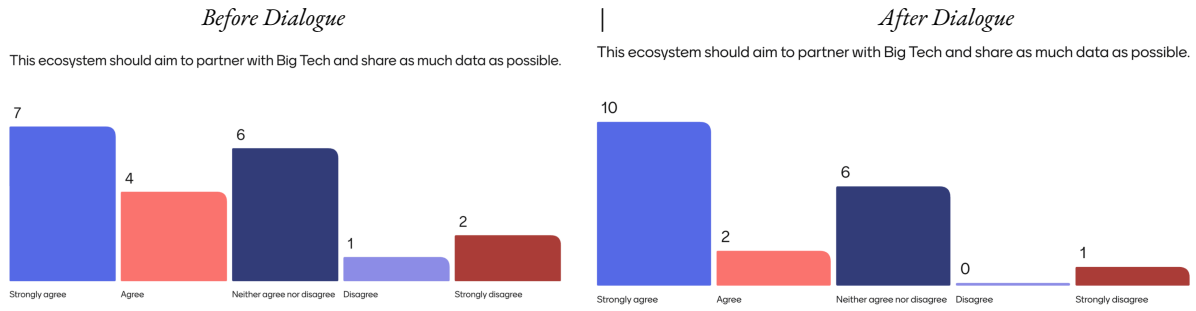
- Time cost of open-sourcing and sharing.
- The risk of putting up bad data and tools, since low quality can be a reputational cost.
- Laziness – multiple people said they need to be nagged to open-source or share their work, since uploading and cleaning and documentation is annoying.
- Previous failures to share, including times where the community has said "We've made a site where we are going to put all our stuff on there!" but nobody ever uploaded anything.
- Publication incentives and barriers in academia which make researchers sit on datasets waiting to publish.
- No clear platform or process that's currently being used, and no critical mass of engagement to get started.
- Some data needs to be protected for junior researchers in more vulnerable career positions.
- Copyright of the dataset itself.

Elsewhere in this overview we discuss in more detail approaches to overcoming these blockers, and what sharing could be like for this ecosystem–specifically through projects around benchmarking, data, tools, and principles.

How should this ecosystem coordinate–or not–with Big Tech?

We also polled the workshop participants on the a question of collaboration with Big Tech, and got more mixed results:

This ecosystem should aim to partner with Big Tech and share as much data as possible.



Votes overall shifted slightly towards "strongly agree" after dialogue and away from "disagree".

These results make it appear like this topic was highly contentious, but in reality the subject had very high alignment around the risks and opportunities. We believe that the diversity of responses reflects more a different weighting of the tradeoffs discussed below, rather than a fundamental disagreement about *whether or not these are the tradeoffs* (i.e. essentially nobody denied the risks; essentially nobody denied the opportunities).

Buddhism does not have the resources to compete. The top-performing AI models are run by companies funded on the order of $100Bs. Our ecosystem is working with something on the order of low $10Ms. The wider Buddhist ecosystem may have more funding, but never on the order of direct technical competition. The same disparity is true for compute resources and technical talent.

Top models are a distribution channel. Dharma will be disseminated primarily from top models, not custom models or apps made by this ecosystem or other Buddhists. Our community should acknowledge this and work to evaluate and improve the quality of the most popular models for serving dharma.

Top models perform best on the tasks we care about. AI benchmarks universally support the idea that larger models are capable of achieving higher performance on a variety of tasks. AI agents take this further, with more advanced models capable of solving problems that previously would have required stitching together results from many different smaller models. This matters to our ecosystem: we should build technology that is modular enough to be able to "slot in" the best model or agent for any use case.

Open-source and smaller models are an option. There are some exceptions to the previous point; one participant showed an example of a smaller model performing better on a Tibetan translation task than a larger frontier model. Another showed smaller custom models outperforming as well for OCR. Being able to select the right model for accuracy/price point is important; one doesn't want to always use the most expensive model. In the realm of open-source and smaller models, costs of training systems become much more reasonable and there

is the possibility that the ecosystem should try to build more sophisticated small models or open-source finetunes.

The price of "intelligence" may decrease–or not. Smaller models today can outperform larger and more expensive models from a few years ago. This trend suggests that there may be a point when small, cheap, open-source models can do "everything we want" from the current models, driving down costs. Compute may also be more widely accessible, leading to price reduction. But the opposite may also be true: increased demand for intelligence may drive up the price of accessible compute, AI costs could go up as companies stop subsidizing model prices, and the tasks we need to solve as an ecosystem may become more complicated–and therefore rely on frontier technology that cannot be replaced by smaller models.

They're going to take your data anyway. Big Tech is scooping up so much data from the internet that some participants noted that there was little need to proactively engage with them: so long as the data was in a clean format and on accessible repositories, chances are high it would already make its way into AI pretraining. (On the other hand, there are other reasons to engage AI companies besides to get data to them.)

Support could get dropped at any minute. In addition to costs ballooning, there are other reasons commercial AI models–or other types of partnerships with AI companies–could be dropped. One major company previously ran a "No Language Left Behind" program, and the team was scrapped after two years. Another well-known AI company used to be run by a nonprofit; some shenanigans ensued. AI companies are currently private; geopolitical tensions could change their allegiance and aims. For a long list of reasons from the mundane to the scary, AI support could stop at any minute.

Big tech's motivations are questionable. AI companies' aims–according to the beholder–range from commercial/secular, to transhumanist, to industrial/militaristic (perhaps a few people believe these companies' aims are truly benevolent). In most likely scenarios the aims of an AI company are orthogonal to Buddhism's aims, and likely even antithetical.

Sovereignty is important. For all of the preceding reasons, no matter how much the ecosystem relies on or collaborates with Big Tech, sovereignty around technology, data, and practice is essential. What this means today is the need to build competency (technical skill in AI training and software development, and projects at the intersection of Buddhism & AI), grow resources (funding, GPUs, human capital), and eventually consider building more powerful technology that lessens the risk of reliance on Big Tech.

All of this nets out to a nuanced landscape of tradeoffs, where participants disagreed on the right "general approach." If we were to try to generalize a "net" opinion from the dialogue, it was likely that one should engage with frontier technology for the opportunities it presents to spread the dharma and support our ecosystem, while maintaining extreme caution–via backup plans with open-source technology, building internal competency, and never forgetting the human element at the center of this work (e.g. always linking projects back to the lineage of transmission and understanding how engagement eventually tracks to a healthier "end user" of technology).

What is needed to make this ecosystem sustainable?

With the emphasis placed on engaging with technology, building capacity, and maintaining sovereignty—the sustainability of this ecosystem is an essential question. Four aspects of sustainability got the most attention:

Funding. The question of funding in this space is challenging, as there are primarily only two major funders supporting the organization of the workshop and their immediate ecosystem. Heavy reliance on these funders is

a risk, but there is also an opportunity to build greater collaboration with funders on shared strategy, expectations, and incentive schemes.

(Currently it appears one funder is leaning more towards collaborative projects driven by multiple organizations in the ecosystem, while another is leaning towards privately seeking out a set of partners to help advance a specific concrete agenda for Buddhist digitization and technology, both of which are helpful).

Historically, the funding model in this space and in Buddhist studies more broadly is very hesitant to pay people and instead tends to pay for projects, which can lead to administrative overhead and "extreme volunteerism." If people move into academia, incentive structures shift away from what benefits the ecosystem (as one participant put it: "Useless research must be allowed to cease!"). The shift that funders have communicated in being open to collaborative projects is a good sign, though a question comes up around continuity of funding: to build technical capacity, a long enough runway to guarantee stability is necessary.

Some conversations touched on governance models, but ultimately the theme that kept coming up was the idea that funders should set requirements for openness and collaboration. Participants agreed that funders should require grantees to open-source data and technical artifacts and collaborate—and that some portion of funds should be withheld until deliverables are uploaded. Comparisons to existing endangered language documentation programs illustrated that there are successful models to follow to incentivize this.

Bringing in more talent. There are very few people in the ecosystem with the trifecta of skills between (1) strong software engineering / AI, (2) language skills, and (3) Buddhist studies or dharma training. Increasing this would be good but is not trivial–not only because recruiting is a challenge, but because training would require the small number of people in the ecosystem who already have these skills to commit time away from their existing projects. A possible "lighter lift" project of offering training as part of a summer institute was suggested and will be explored further.

It was emphasized that attracting and retaining people requires providing motivation, which can come from joy of work, vision, or impact on users. Starting from the impact on end users is a good way, though for-profit vs. nonprofit vs. academic vs. spiritual motivations are very different, from impact, to cultural preservation, to solving interesting research problems, to Buddhist goals (such as aiding all sentient beings in the temporary alleviation of suffering and the ultimate attainment of liberation, and building cultures that can enable authentic transmission and thereby ensure the longevity and wide reach of our efforts to benefit sentient beings).

Supporting current talent, particularly (1) in the Tibetan diaspora community, and (2) among technologists. The current strong talent in this space needs to be retained. According to workshop conversations, many individuals have faced funding challenges that have required shifts in plans or career paths to less-aligned but more stable opportunities, and organizations have faced employees leaving because of job security or the need to also take on less-aligned projects in order to sustain themselves.

The problem above is particularly challenging for the Tibetan diaspora community in India, and one participating organization shared a number of examples of retention being an issue. Many Tibetans consider immigration their highest priority, and this is true not just of programmers but also annotators, and teachers: senior monastics are now moving to other countries and taking careers such as driving rideshare vehicles; once this has happened it is nearly impossible to bring them back, even if paid a really good salary. (This trend parallels a decline in monastic attendance across the Tibetan diaspora community). One possible solution is higher salaries, and also being able to offer Tibetans jobs in a place that feels like a "destination" (e.g. Thailand as a stepping stone, hopefully Western countries eventually). A significant amount of work in the NLP/tech space

around Tibetan Buddhism is done by Westerners, not Tibetans, and so the projects driven by Tibetans ought to be particularly protected and supported by this ecosystem.

But retention is also a problem for Westerners, including the strongest technical talent in this space. General consensus at the workshop is that there was a universal experience of instability of funding among technical talent, where project-by-project funding meant that it was challenging to do any forward-thinking work necessary for scaling. Individuals have not been incentivized to create new organizations, recruit ambitiously, or think about long-term and big plans because even their own funding is unstable. A few anecdotes were shared of past funding decisions leading to some relationships falling out, which is understandable but seems particularly likely to happen in an ecosystem where funding is scarce.

Open-source. In addition to matters of funding and human capital, a healthy "tech commons" of open-sourcing, documentation, APIs, etc. is also an important part of sustainability. For example, really good documentation creates continuity. Stability and reliability of data also matter. Not every university or organization can guarantee long-term URI stability, so routing data through a public foundation (e.g. Wikimedia for person-readable text) is a good model. HuggingFace, CommonVoice, and Zenodo provide other platforms or datasets.

Participants agreed that publishing is critical – both open-sourcing, and also publishing academic papers. The academic process was defended: "if you publish a paper, it at least forces some communication." Even technical organizations now have an incentive reward for paper publishing. Again, funding standards can force documentation of datasets and increased visibility.

There was some discussion around format standardization (e.g. TEI for text data). This idea overall seems to have been voted against, though standardization within a project is good. Some communities have agreed on shared formats, which makes collaboration easier. But some people find this overly restrictive, and others noted that AI tooling may make rigid format standardization less critical, since raw data can often be processed and converted.

Possible Projects and Outcomes

All of the discussion above happened in parallel with the workshop participants advancing ideas around collaborative projects to take action against the concerns and opportunities this ecosystem faces. Below is a summary of all projects, with notes about leadership and how to get involved.

Benchmarking and Evaluation

We need to understand the degree to which AI can support–or replace–work in this field. Benchmarks help compare: (1) AI capability vs human, (2) which AI model is best, (3) solutions across organizations (AI, tools, pipelines) by creating a standard way to measure task performance. Without benchmarks, selection of tools & processes relies solely on anecdotal evidence, it becomes difficult for the community to know if they should lean on each other's tools and AI solutions, there is less incentive for a "positive competitive environment" in the ecosystem, and it is unclear if the ecosystem is collectively progressing on its goals.

What we are trying to measure here is broad. There are many domains that can be measured: different traditions, skills, knowledge, genre, etc. But for an illustrative example, translation can be broken down into a "set of tasks a

human translator would do in their day-to-day job," just as the task a Buddhist philologist can be broken down into a set of tasks–as one participating team has already begun with their published benchmark.

To construct a benchmark, there is a somewhat standard pipeline from idea to evaluation:

- Interested folks who want to discuss tasks should come together, representing both computer science and humanities.
- First identify how a human solves the tasks, and try to assess how to solve it computationally.
- Each task then requires a multidisciplinary group to collect data evidencing that task, ~100s of samples per task.
- Then design a pipeline for evaluation (solely computer science) to assess "Each model scores X on Y task."
- All tasks summed up to get a final score of "how good is the model in the target domain."

This process isn't trivial, and creating even a small benchmark can take time and money. There are computational approaches that can help (e.g. pre-annotation of data for a benchmark using AI can cut down time) but the most difficult aspects of benchmark creation are at the point where humanities questions meet technical questions: what is the precise skill that matters, and what is a good measurable proxy for it? Participants are planning to follow these next steps to progress as a community on this:

- **Webinar on Benchmarking** | There will be a webinar on Machine Translation evaluations and other evaluations that will be open to the whole ecosystem, covering how to construct a benchmark, and metrics used for scoring translation and other tasks. This is aimed at a wider audience than just the workshop, raising awareness around benchmarking for NLP tasks that can apply to any tradition/field/subdomain.
    - Leading: Two participants will co-lead.
    - Supporting: Several other participants and organizations have expressed interest. Could include talks from external speakers from other languages.
    - Next step if you want to be involved: Contact the leads. Look out for an event date in the communication channels.
- **Working Group / Scoping Call** | One organization is very interested in working with several others on the question "Is it possible to continuously assess the quality of AI translations?", and other scholars are interested in benchmarking philology and "digital Buddhology" tasks. Building these benchmarks is a medium-effort project. Before the project kicks off, there will be a scoping call to explore what directions to first explore and how much overlap there is in translation/philology benchmarks, helping determine whether these should be 1 vs 2 next projects, and determining if grant funding is needed for the next stage of work or not.
    - Leading: Three participants from different organizations.
    - Supporting: Interest from many others at the workshop. Should link up with funder representatives interested in this direction of work.
    - Next step if you want to be involved: Join the webinar first; that conversation will flow into this one. If you miss the webinar, reach out to one of the project leads to ensure an invite to this call/planning process.

Later more detailed work could progress towards:

- **Basic Translation/Philology Benchmarking** | Ideally a medium-sized, grant-funded, collaborative effort to (1) break down the role of a Buddhist translator into tasks, (2) figure out how to evaluate each or a sample of those tasks, (3) construct small benchmarks with initial data for each task (e.g. for translation, sentence pairs from 1–2 texts in each genre for 5 genres). The result should be capacity-building for how to benchmark within the field, preliminary leading results, and a clear understanding of what a "full" project looks like.
- **(In Parallel) Benchmark Hub:** Several participants have existing benchmarks. There is a question of whether they should be located on the same website or not, and if there are standards for how these benchmarks should be approached. It is likely that this will unfold naturally from the progression of conversations and projects here, but it is possible to already start uploading shared benchmarks to a shared Tibetan AI Benchmarks repository. One participant is planning to contribute their Q&A list once it crosses 250 entries. Could also explore ELO and community ratings in an AI arena to incentivize healthy competition.
    - Pros and cons: There is a spectrum from piecemeal (benchmarks hosted solo, or privately) to a central hub.
    - Reasons for piecemeal: Benchmarks aggregate very different tasks; a single platform risks ideological fixing over time; individual benchmarks can be customized more (e.g. a participant's OCR benchmark looked great and was easy to follow).
    - Reasons for hub: There are many shared tasks in the Buddhist studies or translation domains; centralization helps people know where to find information; it is labor-intensive to create the tasks and datasets, and a central hub makes collaboration easier.
- **Harder Benchmarks and Future Work** | There are many possible extensions of the above project that will be determined as the group progresses, including: full translation/philology benchmarking; benchmarking across methods and meta-evaluation; additional Buddhist domains and "wisdom" benchmarks (Can AI models accurately reflect/transmit Buddhist Wisdom? Can they answer ethical questions from a Buddhist perspective? Can they support meditation?); and annotation pipelines for more reliable capacity for data annotation among experts and monastics across domains/traditions.

Data Mapping

The foundation of this ecosystem is a shared set of Buddhist data–texts across traditions, translated content, audio files of recorded dharma talks, images of sacred relics–and datasets of this data that have made cuts and additions to this for various use cases. A shared data commons is critical for this ecosystem, both in digitizing and sharing the current resources, as well as in strengthening the collection with new data.

Currently, practices are all over the place for data sharing, though several participating organizations are demonstrating what a "good job" in this space looks like. Participants discussed the degree to which data sharing should happen–the consensus was essentially "the more the merrier," though in terms of next steps, it seems wisest to start with the smallest version of a project rather than becoming overly ambitious about immediate aggregation:

- **Data List.** Participants agree that a shared data location would be amazing, but that an even better minimum starting point would just be a shared data list–hosted on Github or another platform–covering datasets, where to access them, and information on sharing conditions (what

organizations are willing to share, and what they want in return). A designated "documentarian" could put this together once and be funded to keep this updated by actively coordinating with groups in the ecosystem. List could also be extended to also track benchmarks, APIs, models and tools.

    - Leading: Two participants will discuss this and a process to get feedback from the community. If anyone else wants in on the early stages, please reach out.

- **A Data Hub** | Beyond "the list," this project could extend to a shared site to host datasets (e.g. a HuggingFace repository), which holds (1) raw data, (2) filtered/preprocessed/deduplicated data, (3) metadata-tagged or task-specific data, or could be extended to also track benchmarks, APIs, models and tools.
    - Next steps: None. The data list is first, and nobody raised their hand to coordinate this piece in particular. In the interim, extremely strong endorsement from all in attendance to open source data, document it, and put it on sites like Zenodo, HuggingFace, CommonVoice (for audio), and others.
- **More Data; All the Data** | At multiple points participants discussed collecting more data from the places it's known to be but not easily collated and accessible (across tradition, source type, location, private/public collection, etc.). In particular for Tibetan data, the vast majority of high-quality data is coming from Chinese sources, and there are very weak connections East-West here.
    - Next steps: None. There was broad support for this but no clear objective outlined – this may need to be a top-down project from a funder if there is interest in model-training since this is out of scope of any one organization; many people could be rallied for further contribution here.

Modular Technology

Various groups and organizations are building technology where there are some shared goals–but not entirely shared use cases. How can the community get the most of collaboration without sacrificing the need for specificity? Participants suggested that a path forward was through "modular technology" where certain features of tools could be toggled on/off, and tools could support adding other organizations' work via APIs.

For example: one organization has a translation platform; another has an excellent tool for dealing with source texts; others have a Dossier/Railroad package to help provide context for AI prompts; another organization has interest in a translation tool but has a different pipeline... how can all of this related work best support each other?

A general principle of supporting technology sharing via API was emphasized. In addition, there was a consideration to use benchmarks as a way to standardize quality assessment across tools and agent approaches. Overall, concrete next steps here included:

- **Documentation of translation tool workflows** | Organizations were interested in building a translation framework that could have modular plug-in APIs for different phases. A central place for documentation will be created that tries to capture different needs around "what needs to be made into an API."
    - Leading: One participant (with close support from a translation-platform team) will put together a map of the translation workflow and how various tools/APIs fit together.
    - Supporting: Several other organizations to be consulted.

- **Pairwise API integration meetings** | Following documentation of the possibilities of integration, organizations have agreed to follow-up meetings to work on integrating APIs. Examples discussed: turning one product into a more interoperable workflow enabling aspects to be tailored to different organizations that want to try the interface; integrating one team's tools into other teams' workflows; integrating an adaptation tool with a verse-handling tool; making dictionaries API-friendly/dockerized.
    - Next steps: Pairwise meetings organized between organizations.
- **Version Reconciliation** | A few participants expressed the need for tools and projects to do better version reconciliation of texts: competing sources are often collapsed into a single textus receptus that isn't able to show the variation. Multiple organizations have tools which could be supportive for this.
    - Next steps: None. This project was discussed as interesting and important, but it was not raised as higher priority than the other projects pursued.

Training and Technical Skill Continuity

In terms of practical steps on ecosystem sustainability, participants agreed that ambitious approaches here seem perhaps more important, more expensive, and much more difficult than some of what is laid out technically above. Nonetheless, there is a starting point of possibility in collaborating around a summer program that gives the community a clear goal for next steps:

- **Trainings for Buddhist Studies scholars and students** | There is a summer school opportunity, among other training options, to teach AI skills to Buddhist studies experts and students. Several participants are interested in following up on a call for thinking about potential trainings with universities.
    - Next steps: Suggest this group has a first meeting to figure out next steps and think about how to include non-conference attendees after that.
- **Pairwise Talent Support** | Two organizations to have a follow-up conversation on the possibility of collaboration on supporting Buddhist/junior-developer roles (one organization has needs, the other has a good recruiting pipeline). Other organizations with technical talent needs are welcome to join this conversation.
- **Future Opportunities** | There is a lot more that can be done in this space – the ideas below are ones that were discussed at the workshop but no specific plans were made to take forward:
    - Funding expansion: Funders–alongside members of this community–can look to build a vision/plan to attract further funding into the space via collaboration with adjacent funding organizations or by bringing funding from new domains into the space.
    - GPU Resources: A working group could convene and research the possibility of securing GPUs for the ecosystem via academic partnerships, pooling, or get a dedicated cluster for the space (a "GPU Stupa!"). Cost estimates and possible options would help.
    - Recruiting and Training: Explicit projects to bring in new talent and train them up into existing organizations. Obviously of benefit, but has the current cost of requiring buy-in/significant dedicated time from the small number of current technical staff in the ecosystem; summer school voted a more accessible near-term option.
    - Ensuring funding for technical skill continuity: Either by grants for individuals or to organizations, funding could support longer runways for people with technical skill.

Particularly essential where deep domain knowledge (e.g. AI training) has been acquired since it cannot be replaced immediately by hiring; rebuilding can take years.

Norm Shift and Whitepaper

Finally, participants discussed the best way to share the principles that emerged from the workshop with the broader ecosystem. In addition to acting on the spirit of collaboration more directly, an idea was raised to create a whitepaper or "manifesto" that enshrines some of these principles in a document that can be signed and shared more broadly.

- **Norm Shift:** A number of points were raised at the workshop in the spirit of "we as an ecosystem should act more like this." The below is a list that individuals and organizations can consider adopting immediately. Many of these points are also salient to the "manifesto" project below.
    - Open data ("open-source without restriction") should be the default; bilateral/private data sharing between organizations also strongly encouraged.
    - APIs are essential for sharing tools easily – when possible, support these for community use.
    - Documentation is critical for ease of use and continuity; please create this for data, tools, etc. and consider multimodal documentation, e.g. video tutorials and podcasts.
    - Show-and-tell for community projects is great – let's do more of this, as well as providing user feedback to each other (e.g. one participant has asked for "some ruthless feedback!" for their translation agent; regular ecosystem participation in this and ongoing testing is extremely supportive).
    - Learning from other fields may be particularly helpful across many projects, and this community should spend effort to learn from what has and hasn't worked. E.g. from other low-resource languages, from other religions like Christianity that has incredible translation resources, from other academic fields like classics (Greek/Latin) where philology work has deep history.
    - Sustainability and technical skill continuity: Beyond all of these points is the importance of the people in this space, particularly around preserving the role of technical staff and annotators who are in vulnerable positions. This is an especially heightened risk in Tibetan/Indian communities where retention is challenging. The skillset for data work, model training, and more cannot quickly or easily be rebuilt and so it's worth preserving.
    - Funding requirement: Funders should consider helping enforce these norms by requiring open-sourcing of data or technical artifacts as a condition of grants, with a certain amount of funding withheld until delivery.
- **Ecosystem Statement ("Manifesto"):** Workshop participants are considering writing "An Open Buddhist Tech Manifesto" that contains many of the points above and could be signed first by participants of the workshop, and then by beneficiaries and users of this technology: academics, translators, or other Buddhists, to showcase cooperation. The statement/manifesto would lean on models of "fair data" from other communities; aim to summarize the meeting; explain values committed to (sharing with each other, and sharing with the wider ecosystem); make shared resources clear (e.g. a landscape website with data, benchmarks, etc); and help score "cooperation" metrics (perhaps via a 5-star rating system).

- Next steps: A small group of participants will meet and have a first meeting to figure out next steps and think about how to include non-conference attendees after that. A first pass of some of the ideas here has already been drafted.

.