

Project BoLI – Ethical Principle and Pledge Statement

Bolstering Linguistic Resources in the Indigenous Languages of India

Attribution

This document is authored and published by:

UnReaL-TecE LLP
and
Council for Diversity and Innovation (CoDIN)

In connection with the BoLI initiative.

License

This document is made available under the
Creative Commons Attribution 4.0 International License (CC BY 4.0).

You are free to:

- Share — copy and redistribute the material in any medium or format
- Adapt — remix, transform, and build upon the material for any purpose, even commercially

Under the following terms:

- Attribution — You must give appropriate credit to *UnReaL-TecE LLP* and *CoDIN*, provide a link to the license, and indicate if changes were made.
License link: <https://creativecommons.org/licenses/by/4.0/>

Machine-readable attribution:

CC BY 4.0 — © 2025 UnReaL-TecE LLP & Council for Diversity and Innovation (CoDIN)

Document Version

Version: 1.0 (2025)
Part of the BoLI Guidelines

BoLI Ethics Principle Statement

The BoLI (Bolstering Linguistic Resources in the Indigenous Languages of India) initiative is committed to developing language technology through ethical, inclusive, and community-centred research.

We affirm that linguistic data are not mere resources, but expressions of human identity, cultural heritage, and collective knowledge, to be treated with respect, care, and reciprocity.

Our ethics rest on the Four R's — Respect, Relevance, Reciprocity, and Responsibility — founded on the principle of *Do No Harm*. We recognise that every act of data collection and model building intervenes in real social worlds; therefore, we commit to minimising harm, preventing exploitation, and protecting the agency and dignity of all contributors.

- **Do No Harm** – Anticipate and prevent individual and community harm at all stages of research.
- **Obtain Informed and Ongoing Consent** – Treat consent as a dynamic, participatory process grounded in comprehension and revocability.
- **Ensure Respect and Cultural Integrity** – Uphold linguistic and cultural dignity, avoiding distortion or decontextualisation.
- **Uphold Fairness and Reciprocity** – Recognise contributors as co-creators; ensure mutual benefit and equitable compensation.
- **Promote Inclusion and Representation** – Prioritise underrepresented languages, dialects, and speaker groups to address systemic absence.
- **Protect Data and Rights** – Safeguard personal, intellectual, and moral rights through secure and transparent data management.
- **Maintain Transparency and Accountability** – Ensure openness, documentation, and independent oversight across all activities.
- **Foster Responsibility and Long-Term Stewardship** – Sustain linguistic resources as shared cultural assets through community partnership and care.

Together, these principles express BoLI's commitment to building technologies that are ethical by design — technologies that not only avoid harm but actively promote fairness, reciprocity, and respect for the diversity of human expression.

BoLI Ethics Pledge

We, the members and collaborators of BoLI, pledge to uphold the highest standards of ethical responsibility in all our work.

We pledge to:

- Respect every language, speaker, and community as custodians of living knowledge.
- Ensure the Relevance of our work to the needs, aspirations, and realities of the communities we serve.
- Uphold Reciprocity, ensuring that the benefits of research and innovation flow back to contributors through fair recognition, compensation, and shared ownership.
- Act with Responsibility, maintaining transparency, accountability, and care across every stage of data collection, processing, and reuse.

We affirm that “Do No Harm” is not a limit, but the foundation on which our ethical practice stands—a commitment to protect against exploitation, exclusion, and misrepresentation.

We commit, in letter and spirit, to upholding the BoLI Ethics Principles, ensuring these values guide every stage of our work—from field collection to governance and access. Through this pledge, we commit ourselves to developing ethical, inclusive, and human-centred language technologies that are grounded in dignity and diversity.