Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

The linguistic adaptation gap: connecting multilingual AI policy to climate early warning in South Asia

Mohammad Rahmatullah
Daffodil International University
Sonali Das
sonalidas@soa.ac.in

Siksha 'O' Anusandhan, Deemed to be University
Mousumi Dash
Siksha 'O' Anusandhan, Deemed to be University
Durgaprasad Mishra
Siksha 'O' Anusandhan, Deemed to be University

Article

Keywords: multilingual AI, climate adaptation, early-warning systems, language justice, policy integration, South Asia, corpus-assisted policy analysis, artificial intelligence governance

Posted Date: November 12th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10768578/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Abstract

National artificial-intelligence strategies and climate-adaptation plans are expanding at the same time, yet they are rarely studied as connected policy infrastructures. This article examines whether multilingual AI capacity is being linked to climate-risk communication and early-warning systems in Bangladesh, Nepal and Sri Lanka. It analyses six national policy instruments available by 21 July 2026: three AI policies or strategies and three climate-adaptation or climate-change instruments, comprising approximately 210,000 words. A corpus-assisted critical policy analysis first maps sentences concerning linguistic inclusion, language technology and early warning; close qualitative coding then evaluates five stages of integration: recognition, language-resource infrastructure, service linkage, participatory validation and accountability. The analysis identifies a recurring institutional split. AI instruments treat local languages as datasets, models and digital-public-service infrastructure, while climate instruments treat them mainly as channels for last-mile dissemination. Bangladesh explicitly plans Bangla language models and AI-enabled disaster alerts, but the two commitments are not joined. Nepal combines measurable local-language AI actions with unusually operational, people-centred early-warning programmes, yet assigns them to separate policy tracks. Sri Lanka gives the strongest attention to Sinhala- and Tamil-language AI resources, while its climate policy acknowledges language barriers without specifying a computational or institutional response. Across all three cases, no instrument establishes an end-to-end arrangement connecting language data, warning generation, community validation and accountable deployment. The article terms this disconnection the linguistic adaptation gap and proposes a five-stage Linguistic Adaptation Integration Framework. The findings recast multilingual warning not as a final translation task but as a cross-domain public infrastructure problem, with implications for language policy, AI governance, climate services and disaster-risk reduction.

1 Introduction

A warning is only a warning when it can be understood, trusted and acted upon. Technical accuracy is not enough – flood depth, the path of the cyclone or the likelihood of landslides could be accurate, but the message could be in a different language, be delivered in a register that is not understood, or be passed on through a channel where the intended recipients are not literate and connected. Disaster communication is not just about the transfer of meteorological data. A sociotechnical chain that is composed of three elements: scientific prediction, institutional authority, linguistic form, and media infrastructure, all of which intersect with situated interpretation. The availability of information has also long been differentiated from its usability in research on climate services (Dilling and Lemos, 2011; Lemos et al., 2012; Vaughan and Dessai, 2014). Multilingual disaster communication also indicates that language affects access to, trust in, and perception and response to disaster (Uekusa and Matthewman 2023). These literatures are united in their point: Information is not usable just because it is.

AI is now starting to impact a few components of this chain. These days, governments are seeing the possibility of machine learning systems that can predict hazards, incorporate environmental data, personalize alerts, automate public services and produce text or speech in local languages. Meanwhile, progress in natural-language processing (NLP) and large language models has sparked a new round of policy discussion on linguistic inclusion. The state of language technologies is still very unevenly distributed across languages of the world: many languages are well supported and have good data, benchmarks and model performance, whereas many others are poorly supported or are not even available (Blasi et al., 2022; Joshi et al., 2020). The challenge is not just a technical one. Who decides which languages are machine-readable, which communities provide data, and who validates outputs, and which varieties are considered “standard”, are matters of power and public authority (Bird, 2020; Blodgett et al., 2020; Mohamed et al., 2020). In high-stakes situations, these decisions can make the difference in being able to hear a public message at the time it is needed.

National policy has started to take account of both sides of this challenge, but typically in separate policy documents and administrative arenas. AI policies are increasingly signed up for local digital services, chatbots and inclusive digital services, domestic language models and local-language datasets. Plans for climate adaptation and disaster risk strategies include people-centred early-warning systems, communication channels and messages in suitable languages in the local community. It's close enough to be very impressive. The linkage is not. The training of a government-funded national language model can be used either with or without hazard terminology. A climate plan could call for notifications in the local language, but not specify the procedures or programs to evaluate and translate and provide the necessary datasets. The agencies that construct the national AI

infrastructure, and the agency responsible for warning dissemination, may be located in different places, and have no shared mandate, budget or performance indicator.

This article examines that disconnection looking at a comparative analysis of Bangladesh, Nepal and Sri Lanka. The three have analytically valuable variation. All three are at high risk to climate and disaster impacts, all three are linguistically diverse, and as of the date of the corpus closure (21 July 2026), each country had a national level AI policy or strategy and a current national climate-adaptation or climate-change policy instrument in English. Their documents are also different legally and in terms of policies. The National Artificial Intelligence Policy 2025 is in adoption in Nepal, whereas the National AI Policy 2026–2030 of Bangladesh and the National Strategy on AI of Sri Lanka are draft policies at the time of analysis. The comparison is not therefore, legal equating of the documents. A more defensible question is what types of policy imagination and implementation architecture emerge when each country's top-level AI and climate instruments are read together: what forms of policy imagination and implementation architecture become visible when each country's top-level AI and climate instruments are read together?

The study will be guided by the following three research questions:

1. 1. How are national AI policies and national policies on climate in Bangladesh, Nepal and Sri Lanka linguistic inclusive and multilingual in access?
2. 2. Where do these policies "manifest in data infrastructure, service delivery, participation and accountability" (as a gateway to language)?
3. 3. How do the AI and climate policy domains relate to each other in terms of local-language AI capacity and climate-risk communication and early-warning systems?

The article has one main argument. The language is recognized in the three cases, albeit with different institutional meaning in each case. AI policy sees language mainly as a tool for technology and development: a collection of corpora to be curated, a model to be developed, a public service to be deployed, and sometimes as a part of digital sovereignty. Climate policy takes a narrow view of language, and calls it a delivery problem, a means by which the communities can be reached with warnings, advisories and awareness materials. Since the resource side and the delivery side are not connected, there are no end-to-end arrangements for each case between the local language data and the local language models to the production of warnings, community validation and the responsible deployment. I refer to this patterned disconnection as the "linguistic adaptation gap.

It is not a mere replication of the conventional notion of a climate-information usability gap, but rather an adaptation of this. Lemos et al. (2012) report on the gap between the production and effective use of climate knowledge. The linguistic adaptation gap is a more specific cross domain failure: institutions might acknowledge linguistic diversity, have or plan language technologies, and by themselves undertake a commitment to inclusive warnings, but not have the policy integration to enable coordination of these capacities. The difference here is not based on whether the word ‘language' is used or not. It shows up when there is no chain of service design, participatory testing and accountability.

The article is providing four contributions. First, it puts language policy, AI governance and climate-adaptation research into a comparative perspective. Previous research was focused on multilingual risk communication, low resource NLP, national AI readiness and climate-policy integration; not a single study that engages in a systematic comparison of South Asian AI strategies and national climate-warning commitments was found. Second, it offers a clear corpus-supported approach for policy-rich, but limited document sets. Patterns are found by lexical mapping; interpretation is still based on close reading and document status and implementation schedule. Third, it builds the five-stage Linguistic Adaptation Integration Framework (LAIF), which outlines: recognition, development of language-resource infrastructure, linkage to language services, participatory validation and accountability. Lastly, the study reveals why multilingual warning should not be used as an additional and low-cost translation in the first steps of the early-warning system. It is to be governed as public infrastructure, starting from data creation to community action.

The subsequent section builds on the conceptual links between language justice, multilingual AI, and usable climate information and policy integration. Case selection, corpus construction and coding are given an explanation in section 3. The comparative results are shown in Section 4. The linguistic adaptation gap is discussed in section 5, as well as the LAIF framework and the policy implications. The ending defines what an integrated national approach would entail.

2 Conceptual background

2.1 Information delivery to linguistic actionability

The early-warning systems are frequently presented in terms of the ‘linked technical and institutional elements' namely: risk knowledge, monitoring and forecasting, communication and preparedness to respond. However, communication is not simply the bridge between prediction and action, but is an ongoing and social process, not a one-and-done information delivery (Moser, 2010). Messages must be able to convert probabilities, thresholds and model output to something a person can understand when under pressure. What has been found to be effective in climate-services scholarship is that information is only useful when it is shared and engaged by actors – producers, intermediaries and users – rather than one-way spread (Dilling and Lemos, 2011; Lemos et al., 2012; Vaughan and Dessai, 2014). Salience, credibility, legitimacy, timing and fit to decision contexts are key factors for usability. Scientifically better prediction may not be of public interest if institutional routines, communication practices or user needs are not considered.

This is a problem that is exacerbated by the use of the language because there are several different operations that can be implicit in the word translated as “translation.” Lexical translation from a named language to another may be required, adaptation for local varieties, conversion from technical to ordinary register, and spoken delivery for users with limited literacy and explanation in terms of culturally familiar signs, places or practices. Advance planning and trained intermediaries and institutional responsibility are crucial to multilingual crisis communication research, which emphasizes that translation should not take place ad hoc once a crisis begins (Uekusa and Matthewman, 2023). Language justice is a matter of concern that is not limited to the availability of an already produced message. It addresses questions of participation in knowledge production, as well as recognition of linguistic difference and the institutional distribution of communicative authority as opposed to simply translating outputs of the dominant language (Piller, 2016; Van Horne et al., 2023).

In this study, linguistic actionability is the ability of a message to enable prompt and informed action for intended recipients. It has a minimum of four dimensions. The first is semantic accuracy: hazard type, hazard severity, location and timing should be translated. The second is pragmatic clarity—the message should tell the receivers what they're supposed to do. The third, contextual intelligibility: there is a need to feature terms, place names, measurements and channels that are appropriate in the context of local communication practice. The fourth is institutional trust, which means that users need to identify the originator of the warning and know what the warning is telling them to do. The dimensions can't be deduced just from the language's presence. Promis of messages “in local languages” define a principle but not how accuracy, clarity, and contextual fit and trust will be created and measured.

This is important because Bangladesh's cyclone-warning experience shows. Research on warning systems has focused not only on the institutional development but has also noted a lack of consistency in the perception, trust and response of said warning providers and receivers (Haque, 1995; Roy et al., 2015). The results do not mean that language was the only or primary reason for non-response. Previous experience and perceived risk, livelihood, and access to shelter influence evacuation decisions. However, they do indicate that warning effectiveness is not just a function of the sending of a warning. Communication is an integral aspect of preparedness and governance.

2.2 Multilingual AI as public language infrastructure

The fast growth of NLP and generative AI is opening up new avenues to generate speech, text and conversational services at scale. It is also an economy of language that is highly imbalanced. Joshi et al. (2020) demonstrate that research interest and resources are focused on a small fraction of languages. The inequalities in language-technology performance are shown to be systematic, as shown by Blasi et al., (2022). Nominally supported languages can have different performances by script, domain, dialect, code switching and task. Translation for general purpose system that works well with news may not be effective with compressed emergency instructions, unusual place names or technical hazard terminology.

Multilingual AI cannot be confused with an interface that lets users choose a language. It is more likely to be called public language infrastructure: sets of interdependent data, lexicons, speech resources, annotation standards, models, benchmarks, institutional expertise and governance procedures to support public services in certain language contexts. Material and material history of this infrastructure. Collection of data from speakers, selection of categories, normalisation of varieties, distribution of

model errors. Sambasivan et al. (2021) explain what they call “data cascades” in high-stakes AI, as a result of neglected data work. Bird (2020) and Mohamed et al. (2020) caution that language technology can be used to perpetuate colonial relationships when communities are seen as data providers instead of actors with decision-making power around the purpose and application of the technology.

There is an alternative model, the participatory work in low-resource machine translation. Local knowledge and community engagement are not an afterthought in developing a system, but are a key part of the dataset's quality, language's legitimacy and sustainable capacity as demonstrated by Nekoto et al. (2020). For warning systems, participation has to be sustained after the datasets are created. Speakers and disaster-exposed communities need to make sure that messages generated from the model are accurate, comprehensible, culturally appropriate and actionable. The combining of BLEU, accuracy or generic language-model performance is not the benchmark to look at, however. It consists of realistic task-specific human task evaluation: Does the recipient recognize and identify the hazard correctly? Is the affected area easily found by the recipient? Have the needed action been explained? Is there an intelligible variety and pace in an audio version? Are mistakes preventable and fixable?

While some international ethics frameworks have included the principles of inclusion, cultural diversity and human oversight (United Nations Educational, Scientific and Cultural Organization, 2021), accountability in the age of AI scholarship also states that governance cannot be limited to general ethics principles. Auditing should be done throughout the system life cycle, assign responsibilities, establish avenues for contestation and correction (Raji et al., 2020). In multilingual public communication, this implies language-specific performance reporting, documentation of human control, version control of key language, incident reporting and redress. A policy statement of “inclusive datasets” is good but it lacks clarity on which language groups are represented, how their representation is calculated and what steps are taken if the warning model keeps failing for one group.

2.3 Policy integration and the linguistic adaptation gap

policy integration is not merely the two goals in the same document. Candel and Biesbroek (2016) present a concept of integration as a process of policy goals, instruments, subsystems and actor involvement. The method is applicable in the present study as multilingual AI and climate warning function are administratively spread out. ICT ministries/ AI centres can manage the computing infrastructure, datasets, and model governance. Forecasts are created by meteorological departments. Warnings and response are coordinated by disaster-management agencies. Delivery is mediated by local government, broadcasters, telecommunication companies, civil society and community-based volunteers. A commitment can be repeated on this landscape without integration.

The analysis thus identifies five stages:

1. 1. **Recognition**: Identification of linguistic diversity, language barriers, language-specific users
2. 2. **Language-resource infrastructure**: Refers to corpora, terminology, speech or text datasets, models, benchmarks or technical capacity.
3. 3. **Service linkage**: Linked with climate information, climate advisories, generation and dissemination of early-warning.
4. 4. **Participative validation**: Relevant language communities and users are engaged in design, testing, revision or evaluation of comprehension and actionability.
5. 5. **Accountability**: Policy provides for leadership, resources, deadlines, performance indicators, auditing or redress for multilingualism.

These stages are the Linguistic Adaptation Integration Framework in Section 5. They are analytic and cumulative in theory but not necessarily so in practice. Participative validation can reinforce recognition, accountability can ensure improved resource development and service experience can reveal needed terminology. However, the stages alert to the point where a policy chain is broken.

The linguistic adaptation gap is characterised as the institutional gap between declaring language access and implementing a system to ensure that language materials are linked to actionable climate-risk communication at the end-to-end system level. There are three clarifications of importance. First, it is not an accusation that any climate policy must call for a specific type of AI system. There are areas where non-AI communication and human translation still matter and automation can result in the spread of dangerous errors. Second, integration is not about turning over the power to the technical agencies from the disaster

professionals or communities. It involves the development of shared governance that enables language technology to support known and specific public-safety goals. Third, the framework doesn't assume that the more technology there is, the better for justice. A national language model can favour a standard language compared to minoritised varieties. Participation and accountability are essential components of integration, not just a direct translation from datasets to deployment.

This structure makes a more discriminating comparative question. Rather than asking questions about the presence of language, AI or early warning in the policy, the question it asks is what the policy is supposed to do, who owns it and how commitments link across domains.

3. Materials and methods

3.1 Case and document selection

The study uses a purposive, comparative policy design. Bangladesh, Nepal and Sri Lanka were selected because all three had, by 21 July 2026, a national AI policy or strategy and a current national climate-adaptation or climate-change instrument available in English. The countries also offer relevant variation in language regimes, hazard profiles and policy maturity. The aim is analytical comparison rather than regional representativeness. India and Pakistan, for example, have major AI and disaster-policy initiatives, but the document pairs available at the corpus closure date were not directly comparable to the six national instruments used here. The sample should therefore not be read as an exhaustive account of South Asia.

The corpus contains six documents totalling approximately 210,000 words (Table 1). Bangladesh is represented by the National AI Policy Bangladesh 2026–2030, Draft V2.0 and the National Adaptation Plan of Bangladesh 2023–2050. Nepal is represented by the adopted National Artificial Intelligence Policy 2025 and the National Adaptation Plan 2021–2050. Sri Lanka is represented by the July 2024 draft National Strategy on Artificial Intelligence and the National Policy on Climate Change 2023. The instruments are not identical in legal force, length or administrative form. This difference is recorded rather than flattened. The analysis concerns the architectures they propose, not their implementation outcomes.

Table 1
Comparative policy corpus

Country	AI instrument and status at corpus closure	Climate instrument and status	Pages	Approximate words
Bangladesh	*National AI Policy 2026–2030, Draft V2.0* (public-review draft)	*National Adaptation Plan 2023–2050* (national plan)	44 + 264	94,500
Nepal	*National Artificial Intelligence Policy 2025* (adopted)	*National Adaptation Plan 2021–2050* (national plan)	56 + 173	81,800
Sri Lanka	*National Strategy on AI, July 2024* (consultation draft)	*National Policy on Climate Change 2023* (national policy)	68 + 42	33,800

Note. The documents were obtained from government, UNFCCC or official policy-hosting sources. The corpus freeze is important because policy status can change. Bangladesh's and Sri Lanka's AI documents are analysed as drafts, and no claim is made that every provision had been adopted or funded by 21 July 2026.

3.2 Corpus preparation and machine-assisted mapping

To get the PDFs in plain text (UTF-8), I used pdftotext. Quoting and page verification were done with page rendered originals. Policy bodies and implementation appendices were added as they are not always part of the policy text. Where pages such as title pages, tables of content, acronym lists and reference sections scored as 'false positive', these were not included in interpretive frequency claims. Normed hyphenation and line-break artefacts. This results in a document list, a coding sheet, and a reproducible lexicon specification in the supplementary materials (not copyright-protected source PDFs are not redistributed).

The key consideration in selecting a method was to take a balanced approach between coverage and interpretation, hence a corpus-assisted approach was used. While lexical counts can provide some insights into the meaning of a policy, corpus methods can also be used to identify patterns of words and to retrieve dispersed passages in order to facilitate critical policy analysis (Baker et al. 2008). All documents were divided into sentences and then subjected to search, using three clear lexicons:

- **Linguistic inclusion**: In terms of language, linguistic, multilingual, local language, mother tongue, Bangla, Bengali, Nepali, Sinhala, Tamil, dialect, translation, interpretation, language barrier.
- **Language technology**: combinations of language terms with model, dataset, corpus, NLP, speech, text-to-speech, machine translation, LLM, chatbot, benchmark, localization.
- **Early warning**: early warning, warning system, forecast, alert, warning, evacuation instruction, and associated hazard-communication terms.

The mapping was purposefully conservative. It detects matched sentences, not independent thematic units; and a sentence can match more than one category. Counts were normalized for document length for comparison purposes, but inferential statistics were not applied given the corpus comprised of six policy instruments that were purposively selected instead of a probability sample. All sentences where the terms language, climate and the disaster or warning terms appeared together were checked by hand. Titles, acronyms or national slogans resulting in false matches were eliminated.

3.3 Qualitative coding

The qualitative phase involved structured content analysis and critical policy reading in which the policy texts were viewed as problem representations that structure roles and potential remediations (Bacchi and Goodwin 2016). All documents were coded on the five dimensions of the LAIF, which are outlined in Section 2.3. Table 2 contains questions on coding. When no provision was identified, an absent code was marked, when a principle or activity was mentioned but no chain of operations, a partial code was marked and when instruments, responsible actors, timelines, budgets, indicators were mentioned, an operational code was marked. These labels are not interval-level scores these are interpretive profiles.

Table 2
Coding dimensions for linguistic adaptation integration

Dimension	Guiding question	Evidence sought
Recognition	Whose linguistic difference is named and how is it framed?	Named languages, multilingual context, language barriers, vulnerable language groups
Resource infrastructure	What technical and knowledge resources are specified?	Corpora, datasets, glossaries, speech resources, models, benchmarks, computing or research capacity
Service linkage	Are language resources connected to climate services or warnings?	Localized forecasts, automated or assisted warning production, multilingual alert channels, workflow integration
Participatory validation	Who tests meaning and usability?	Co-design, community review, comprehension tests, drills, feedback, human translation or validation protocols
Accountability	Who is responsible and how is performance governed?	Lead agencies, cross-agency coordination, budgets, timelines, language-specific indicators, audits, redress
Note. Authors' own development		

Close reading focused on the position of language in grammar and institutions. Special focus was given to verbs and objects: policies which were promising to "support" languages, to "develop" models, to "disseminate" warnings, to "consult" communities or to "monitor" systems. Implementation tables were read in conjunction with policy prose, in order to discern between aspiration and assigned action. To confirm or disconfirm the emerging finding of domain separation, negative-case analysis was used to find passages in the data that might be regarded as negative cases.Negative-case analysis was conducted to confirm or disconfirm the emerging finding of domain separation by searching for passages that might be considered negative cases, such as explicit connections between local-language datasets and warning systems. The strongest evidence of provision was then compared across cases in cross case synthesis.

3.4 Reliability, scope and ethics

The machine assisted mapping was not an independent classifier, but rather a tool for retrieval and description. All central findings were checked against the PDF pages. To enable another researcher to repeat searches and question coding decisions, the codebook, lexicons and evidence table on the page level accompany the manuscript. Inter-coder reliability is not claimed as the study was done by one researcher. Coding categories were established in advance of the cross-case synthesis of responses,

ambiguous passages were coded as “partial,” and absent claims were limited to the six instruments as opposed to all government practice in order to minimize coding drift.

No human subjects, personal data or automated decision making are involved in the study, nor do public policy documents contain any personal or human subject data or automated decision making. An ethical approval was thus not necessary. The main ethical concern is representational: the nuances in the Bangla, Nepali, Sinhala or Tamil versions may not be fully captured in the English version, and the article itself may capture the assumption that there is a complete list of named languages and that this list reveals linguistic diversity. Those restrictions are not considered a technical note, but a part of the substantive issue.

AI-assisted manuscript development. ChatGPT from OpenAI was employed to help refine the research design, generate the drafts, revise their language, format their references, and create their documents. Not used as an independent class or as an author name. During the process of writing the manuscript, the source documents and the passages quoted from the policies were cross-checked with the underlying documents, as were the lexical outputs, coding profiles and quoted references. The design of the study, interpretations, claims and final text are the responsibility of the human author.

4 Results

4.1 Parallel policy vocabularies

There is a consistent division of labour in the lexical mapping (Table 3). AI tools include significantly more sentences related to the linguistic inclusion and language technology, as compared to the corresponding climate documents. Instruments for climate have only more words relating to early warnings or forecasts, and little language-technology vocabulary. This is nothing unique, it is not unusual that documents have various purposes. Its importance is related to co-occurrence and operation linkages. Following manual checking, the three AI tools did not include any commitment to local language in the same policy sentence or implementation action, nor any commitment to early-warning. Warning delivery was only explicitly linked to appropriate or local languages in the Bangladesh and Nepal adaptation plans. The passages that were identified for the climate were not NLP resources, language datasets, or AI models.

Table 3
Descriptive corpus mapping

Document	Linguistic-inclusion sentences	Language-technology sentences	Early-warning sentences	Verified language + early-warning co-occurrences
Bangladesh AI policy	18	7	5	0
Bangladesh NAP	7	0	33	2
Nepal AI policy	6	3	3	0
Nepal NAP	8	0	46	2
Sri Lanka AI strategy	10	5	0	0
Sri Lanka climate policy	1	0	4	0

Note. The counts do not show the order of inclusion. For example, the Bangladesh NAP has fewer language sentences than the AI policy, but provides language a direct role in early-warning dissemination. On the other hand, the AI plan of Sri Lanka has a significant number of proposals for Sinhala and Tamil language infrastructure, but no early-warning sentence. The qualitative results help to interpret the differences.

4.2 Bangladesh: two ambitious systems that do not meet

The draft AI policy makes the most explicit general statement of linguistic inclusion in the corpus in Bangladesh. It acknowledges the “multilingual context” of the country, the connection between meaningful language access and equity and public trust, and the need to focus on Bangla and other nationally relevant languages in AI models, datasets and digital public services (Information and Communication Technology Division, 2026, p.12). The provision does not just have to be symbolic recognition. The policy

suggests the use of Bangla chatbots in the public-administration section to provide service navigation, and mentions a national Bangla-language model to be developed to cater to the civil servants. It also emphasizes the use of a locally hosted language model and training data set for Bangladesh (p. 27). There is also an implementation provision, which designates the Bangladesh Computer Council as a part of the government-led national Bangla language-model program, to digitize and curate historical, literary and knowledge resources (p. 34).

A comprehensive AI agenda for disaster-risk reduction is also outlined in the document. It recommends early-warning systems based on AI, machine-learning modelling of hydrological and rainfall data, AI-backed emergency warning platforms, and personalized warnings via SMS, apps and public broadcasts (p. 32). These are unusually straight forward service pledges. However, Bangla, other languages, translation, speech, literacy, terminology or language-model evaluation are not mentioned in the disaster section. The language infrastructure and the warning infrastructure share no shared specification and lie side by side in the country's ambitions.

This isn't a lack of coordination. The draft policy proposes the development of public-sector AI platforms, all of which are centrally coordinated and have standardized interfaces under the ICT Division, which in principle could assist disaster agencies (p. 27). It also requires testing, monitoring and human oversight. What has happened is that these horizontal governance mechanisms haven't been translated into the needs of disaster in a language. No instruction is given for building a hazard corpus into the national language model, no named relationship is given between the Bangladesh Computer Council and warning authorities, and no indicator is provided for understanding of warnings by language group. Therefore, the AI policy is in operation in the following areas: Resource development and generic system governance, partially service linkage and participative validation.

The opposite setup is in the National Adaptation Plan. It has a socially and institutionally based warning system. The plan provides for using ICT-based early-warning dissemination and “readily understandable” advisory services in local languages (Ministry of Environment, Forest and Climate Change, 2022, p. 79). It has given national and local languages a place in translation, channel selection and targeting the message to a particular audience (p. 136). It also needs information, education and communication materials to be available in all ethnic, linguistic and disability groups, and in suitable language and modality for early-warning communication (p. 139).

These provisions go beyond the usual "the information should be accessible" statement. They relate language to warning and recognise differentiated users. But language still is a communication-stage requirement. The plan doesn't mention the technical infrastructure to sustain hazard language, generate quick multiple language versions, test automated assistance or report language-specific performance. Nor does it link the NAP's local language commitments to the other NAP elements, which were not mentioned in the national language-model element until later in the AI draft. Therefore, Bangladesh has the lowest degree of having both parts of an integrated system, that is, detailed language-AI infrastructure and explicit local-language warning commitments. The space is between them.

4.3 Nepal: measurable commitments in separate administrative tracks:

Nepal's AI policy is shorter and more explicit on its commitment to local languages than the draft policy from Bangladesh. It encourages the development of AI systems in local languages (emphasized) (Ministry of Communication and Information Technology, 2025, p. 16). In other places, the policy has NLP as a measure to mitigate learning inequality and references to AI applications in hydrometeorological forecasting and disaster management (p. 13). The timeline of its implementation is also significant, with the Ministry of Communication and Information Technology and the Language Commission sharing the responsibilities for local-language integration and the progress to be evaluated through the number of local languages integrated into AI systems (p. 54).

This is the most robust language-specific measure of accountability in all of the three AI tools. It refers to a linguistic competent institution and inclusion is measurable. But the metric doesn't include quality, nor does it include domain or user outcome. Language can be integrated within a system of poor high stakes communication. Also, the action is segregated from a Disaster and Hydrometeorological provision. The policy does not mandate the Language Commission to be involved in the design of early-warning systems or establish links between AI indicators of local language and meteorological and disaster agencies. Recognition and accountability are relatively good while linkage to services is still lacking.

The people centred warning programme in the corpus is most operational in Nepal's National Adaptation Plan. It calls for multi-hazard early-warning systems reaching local-language radio and TV, SMS and social media, with nationwide targets and lead agencies and significant cost estimates (Ministry of Forests and Environment, 2021, p. 121). Other programme entries call for gender equality and social-inclusion-responsive warning systems and communication languages and channels that are suitable for different groups (pp. 127–128). The plan does not just state a wish to communicate, but sets it as a programme.

The linguistic solution in the plan however, is again channel-centred. It describes where warnings must go, and the use of local languages, but it doesn't describe the language resources or production process being used for language production. National hazard terminology repository is not present, plan for speech resources is not present, protocol for evaluating machine-translation is not present, and there is no link to the local-language indicator of the AI policy. The Nepal case thus shows that there is no automatic linkage of administrative detail to cross-sector integration. There are no connected institutional objects between the country-level AI-language action and the multilingual EWS programme, which is unusually measurable. There are no measurable institutional objects between the country-level AI-language action and the multilingual EWS programme.

4.4 Sri Lanka: strong Sinhala and Tamil AI infrastructure, weak warning linkage

The national language capability has been highlighted in Sri Lanka's draft AI strategy. It demands sophisticated NLP in Sinhala and Tamil, Open source data, Research capacity, funding, or collaborations with other countries, language technologies (Committee on Formulating a Strategy for AI, 2024, pp. 37–38). It also suggests that in local languages, such as Sinhala and Tamil, there should be a mix of public awareness content and identifying local language Large Language Models and national chatbots as tools for access to public services and for cultural preservation (pp. 39, 62). Of the three AI instruments, this approach most explicitly recognizes the local-language resources as an innovation ecosystem of data, research and services.

The present question's weakness of the strategy is very clear: there is no early-warning programme. It broadly speaks to sustainability and sectoral AI opportunities, without linking Sinhala- and Tamil-language infrastructure to climate information and disaster alerts and emergency communication. This omission does not provide any assurance that disaster risk is being ignored in practice in Sri Lankan AI initiatives given that the document is a strategy draft and has many annexes. It does indicate that multilingual warning does not appear as a use case or governance test in the national strategy.

The National Policy on Climate Change 2023 has general statements regarding climate communication, climate information access, and sector-specific communication strategies to address climate change. Additionally, it promotes the adoption of local technologies and traditional knowledge such as traditional early-warning systems and the strengthening of forecasting and warning of extreme and slow-onset events (Ministry of Environment, 2023, p. 17, 24, 27). It most directly provides for climate action that is inclusive and addresses “language barriers” (p. 27).

Thus the document acknowledges the barrier but doesn't precisely define how it will be overcome. It does not mention Sinhala or Tamil in the language-specific warning provisions, differentiate written from spoken communication, identify responsible linguistic institutions and draw up language-specific indicators. The term technology, language equity and early warning are found as distinct policy lines. In contrast to Bangladesh and Nepal, Sri Lanka's climate instrument provides less operational detail on dissemination in multiple languages. It has no reference to the significant language-resource agenda already in place in another national policy area, as compared to its own AI strategy.

4.5 Cross-case synthesis: the same break at different stages

Table 4 provides a summary of the coding profiles. The three cases are different in terms of policy strength but the same structural break is repeated. The language-resource infrastructure is relatively strong in Bangladesh and Sri Lanka, while a language-specific AI indicator is strongest in Nepal. Bangladesh and Nepal are best on local-language delivery of warning information. None is linked these strengths in all 5 stages.

Table 4
Linguistic adaptation integration profiles

Document	Recognition	Resource infrastructure	Climate-warning service linkage	Participatory validation	Language-specific accountability
Bangladesh AI	Operational	Operational	Absent	Partial	Partial
Bangladesh NAP	Operational	Absent	Operational	Partial	Partial
Nepal AI	Operational	Partial	Absent	Partial	Operational
Nepal NAP	Operational	Absent	Operational	Operational/partial	Operational
Sri Lanka AI	Operational	Operational	Absent	Partial	Partial
Sri Lanka climate	Partial	Absent	Partial	Partial	Absent/partial

Note. Authors' own development

There are three comparative findings hereafter. Firstly, the meaning of language varies from one policy area to another. It is linked with models, datasets and digital sovereignty, innovation and public-service interfaces in AI instruments. It is correlated with dissemination, awareness, channels and vulnerable users in climate instruments. The first framing is a language-as-upstream resource; the second is a language-as-downstream-obligation framing.

Second, institutional proximity is equated with integration. Since the AI policy in Bangladesh is also integrated with language models and disaster AI, the two concepts are not presented in the same implementation chain. There are two programmes which are very measurable in Nepal, but their indicators and lead agencies do not overlap. Sri Lanka has a developed national language AI agenda, and a climate policy that acknowledges language barriers, but no cross-reference. No shared workflow due to repetition across policies.

Third, participatory validation is the poorest common denominator. Documents mention consultation, inclusion, awareness/community-based approaches, but none mention how speakers will assess warnings with the support of AI prior to deployment. There is no policy which demands language by language comprehension, actionability, quality of speech, rendering of place names, or false reassurance testing. The accountability is thus directed towards the system presence, its coverage or towards generic AI safety, rather than the communicative performance of warnings among specific users.

The space is best viewed as an institutional gap and not a language gap. There is no need for any more statements in which governments say that local languages are important. They should be provided with means to enable that language resources, disaster expertise and community knowledge would become one governed service.

5 Discussion

5.1 The linguistic adaptation gap

The main conclusion is that neither language nor warning have been neglected by South Asian governments. Rather, there are significant obligations in the documents. This disconnect is due to the fact that the various policy domains envision different end states. AI policy tends to stop when a local-language model, dataset or interface has been created. Climate policy typically starts when there's a message already formed that needs to be communicated to a community. But there's an institutional middle that is missing between hazard data and the creation of linguistically reliable, context-sensitive, and accountable messages: that process.

This is important because failures in upstream components impact failures in downstream components and vice versa. Warning agencies can ask for a rapid translation when there is no validated hazard glossary or speech model for the language they need to translate into. A language model can be learned from literary, administrative or web data, but may not have the technical terms or

pragmatic structures for instructions for evacuation. A message can be grammatically correct, but unsafe because it reduces the chances of survival, mispronounces a location or turns a warning category into a commonplace synonym. However, a technically correct translation can also be unsuccessful if the users are not convinced of the source or do not understand the recommended action or receive the translation in an accessible modality.

Linguistic adaptation gap is thus compounding the problem of usability of the climate information. Climate information use is sometimes scientific but not linguistic, and sometimes linguistic but not scientific. Sometimes climate information is scientifically relevant but institutionally available, and sometimes it is linguistically available but not scientifically relevant. It also gives a boost to the discussion on low-resource NLP. A language is not sufficiently included if it is possible to create some general text generation model for it. Inclusion needs to be evaluated from the perspective of the consequence tasks, public institutions and users. From this point of view, climate warning provides a rigorous challenge to the public value that national language-model projects make.

The discovery also calls into question the prevailing policy notion of the “last mile.” Last mile, in other words, suggests that scientific and technical work is finished before the communication starts, which means that the linguistic adaptation is the last (delivered) mile. This is deceptive, as the corpus proves. Upstream decisions are made for terminology, data, model architecture, geolocation, uncertainty and message length. Language skills can only be introduced at dissemination stage, where they have to fix decisions already made in the system. The entire warning chain should be designed to facilitate multilingual communication.

5.2 The Linguistic Adaptation Integration Framework

The results of the analysis are presented in Fig. 1 in the form of the Linguistic Adaptation Integration Framework (LAIF). The framework will not be used for maturity ranking. It asks if there is a national arrangement that links five functions and if there is an institutional bridge between AI/ICT bodies and bodies of climate, meteorological and disaster-management.

[Insert Fig. 1 here]

Stage 1: recognition. Policies should draw on language communities and communicative situations and not just on the category “local language.” The listing of national languages is a beginning, not a complete list. Minoritised languages, regional varieties, code switching, sign languages, literacy and disability should be taken into account in policies. But the recognition has to be place- and danger-sensitive since the language ecology may differ from district to district, from border to border, from coast to coast and from migration path to migration path.

Stage 2: language-resource infrastructure. National AI initiatives are required to keep open or well-governed hazard corpora, multilingual glossaries, place-name databases, parallel warning datasets, speech resources and task-specific benchmarks. These should contain expressions of uncertainty, severity levels, action verbs and culturally appropriate phrases. Data governance needs to record where the data came from, if it involved consent, data licensing, dialect coverage and revision. These resources are essential to a national language model and are not alternatives to it.

Stage 3: service linkage. There should be a link between resource development and actual warning processes. This could be machine translation or machine generation of templates, speech synthesis, multilingual chatbots, as well as localized advisories or decision support for trained translators. It doesn't need to be fully automated for warning issuance. With so much at stake, the use of constrained templates, approved terminology, human authorization and fallback procedures should be used in systems. Before disaster, interfaces between meteorological data and disaster procedures, as well as between disaster and language services, should be determined.

Stage 4: participatory validation. Language communities, local officials, broadcasters, disability organizations, volunteers and disaster exposed populations should test messages/channels. The evaluation should be based on comprehension, perceived urgency, recall, location recognition and intended action. Drills should contrast text, audio and visual differences and look for differences between language, gender, age, literacy and connectivity. Participation should involve the right to refuse misleading model output, not just provide data, and should be able to alter system goals.

Stage 5: accountability. Policies should specify lead and partner institutions, budgets, timelines and metrics in language. Some indicators might be the percentage of high-risk districts served by validated multilingual warnings, median time for generating

authorized language versions, comprehension rates in drills, failure rates for place names and severity categories, and number of language failures reported and corrected within a specified time. However, independent auditing and public redress are particularly crucial in cases where AI is helping with high-stakes communication.

The five stages in the figure are of equal importance, but the horizontal bridge is more important. There must be formal arrangements between AI centres, language commissions or academies, the meteorological departments, the disaster agencies, the telecommunications regulators, the public broadcasters and local government. These can be a co-programme, common budget line, interoperable data standard or a standing technical committee. If such a bridge does not exist, then each institution can perform its own function and the public service as a whole will be incomplete.

5.3 Country-specific implications

The near-term opportunity is the greatest for Bangladesh, as its draft AI policy has already mentioned an AI-based warning system and has also provided for a national Bangla language model, which are both needed for the NAP to include advisories that are understandable to the people in that country. A coordinated implementation plan could link the Bangladesh Computer Council, ICT division, disaster and meteorological agencies, public radio and television stations, and local-language experts. The national language-model project might involve also a controlled hazard corpus, but it does not need to be Bangla only as it does to meet the NAP's wider pledge to local languages. Pilot testing to be conducted in hazard-exposed districts, and consider regional usage, communities and community intermediaries; consider spoken formats.

Nepal has solid institutional foundation for measurement. The AI policy includes the Language Commission, which also incorporates a local-language indicator, could be broadened to include the assessment of task performance. Channels, agencies, coverage targets and major investment needs are also identified in the NAP. A new indicator might be based on the number of early-warning workflows that have had resources validated and tested in applicable languages, not the number of languages that are nominally spoken in unspecified AI systems. The federal setup of Nepal also demands provincial and local validation, as a message created centrally could not cater the linguistic and geographic contexts in all corners of the country.

One of the first steps forward for Sri Lanka is to identify multilingual climate communication as a priority use case for the proposed infrastructure in Sinhala and Tamil in the AI strategy. Climate policy acknowledges language barriers and these can be translated into specific actions, such as using shared terminology, building and sharing language datasets of speech and text, implementing human-in-the-loop warning tools and testing with Disaster Management Centre and language communities. As both Sinhala and Tamil have political and institutional value, governance must ensure that the quality of service provided to them is not inferior and must also incorporate other language needs and not consider bilingual provision as the end goal.

5.4 Implications for language policy and AI governance

The results place Language Policy in the new context of digital government. Language models are frequently romanticized as tools for cultural preservation in national AI documents, as well as for inclusion and service to people of color. Public value requires sectoral integration, those goals are important. In the case of climate warning, it is clear how a language resource gains its significance in administrative procedures, technical standards and accountability. Language policy scholarship can thus go beyond the status and corpus planning dichotomy and investigate computational resources in the high-stakes public service.

In the case of AI governance, the analysis reveals that the notion of “local language support” is insufficient as an ethical principle. It can be met with nominal coverage and cover large differences in quality. There should be domain-specific evaluation and reporting of performance in high stakes policy, with performance data disaggregated by language and modality. This is in line with lifecycle perspectives on algorithmic accountability (Raji et al., 2020) and critiques of data work that reveal the upstream omissions that manifest as downstream harms (Sambasivan et al., 2021). In defining an inclusive public model, the community should be alerted that the model's output can be used safely by relevant communities.

The article also warns about technological solutionism. While infrastructure is a key factor, human translators, local broadcasters, volunteer networks and community leaders are also crucial, particularly in the event of infrastructure failures. AI can help you be fast and scale up your content production, but it can also generate fluent misinformation or give you a false sense of security about your content coverage. So, the LAIF framework is not focused on automation but on integration. A country could meet the

framework by using carefully-managed machine support and robust human institutions. It will not meet the framework if it was deployed without participation, auditing and redress through an automated multilingual system.

5.5 Implications for climate services and disaster-risk reduction

People centred systems and effective communication are highlighted in the Early Warnings for All and the Sendai Framework (United Nations Office for Disaster Risk Reduction, 2015; World Meteorological Organization, 2022). The comparative evidence suggests that to achieve national implementation, the language infrastructure should be considered as an integral component of warning capacity. This changes budgeting. Translation and validation do not need to be dependent on ad hoc emergency expenditure, but need to be invested in datasets, staff, software maintenance, community drills and monitoring.

It also alters assessment. The extent of warning is commonly expressed in terms of geographic reach, channels or population estimates. When the recipients of these measures receive messages to which they cannot relate, they can see that they are being overprotected. While language-specific comprehension and actionability metrics would not substitute coverage measures, they would lend greater credibility to coverage measures. These measures could be incorporated into regular exercises, post-exercise analysis and maybe even into other research efforts instead of as stand-alone studies.

Last but not least, multilingual warning is a regional issue. Hazards migrate and the media cross national boundaries. A common language and nomenclature for hazards would facilitate regional collaboration and reserve room for national/local differences. But not a single standardized language form should be imposed by regional datasets. They should be well-governed in such a way that they retain their provenance and allow local adaptation.

5.6 Limitations and Future Scope

This study has five key drawbacks. First, it does not address policy implementation but rather policy design. There may be other existing modes of government agency collaboration that are not documented in the selected documents. Interviews, procurement records, technical standards and audit of warning-systems need to be done to determine whether the gap in language remains in reality.

Secondly, the document types are not the same. Bangladesh and Sri Lanka had draft AI instruments at the time of the corpus closure date and Nepal adopted it while climate instruments range from detailed adaptation plans to a shorter national policy. The comparison is not a league table of national performance, but a way of identifying patterns and opportunities.

Thirdly, the English translations might be a barrier to understanding meanings and institutional terms in Bangla, Nepali, Sinhala and Tamil documents. Future studies should involve multi-lingual teams to conduct comparison of official language versions and follow up on the influence translation has on policy commitments.

Fourth, lexical mapping is style and sentence segmentation dependent. The counts in the study were descriptive and addressed this issue by manual verification and close reading. More extensive laws, sectoral plans, procurement documents and operational manuals would enable more powerful computational analysis.

Fifth, user testing was not included in the study. The LAIF dimensions of comprehension and actionability are not derived from experiments with warnings but instead by literature review and policy analysis. The subsequent step in research should involve co-designing and testing multilingual prototypes of warning systems in hazard-prone communities. This type of work must not take data out of emergency situations, and must have consent, compensation, community control, and safe failure protocols in place.

6 Conclusion

There is no language, AI or early warning policy recognition gap at the policy level in Bangladesh, Nepal or Sri Lanka. They don't have any sort of institutional chain that makes those recognitions do their job. Their AI tools envision local language as data, models and public service interfaces. Their climate instruments envision language as a medium of dissemination. In between these two visions lies the linguistic adaptation gap, which is a lack of a reliable workflow that is shared between language resources and warning generation, participative validation and accountable delivery.

The gap must not be filled by passing emergency communication over to a language model. It signifies the management of multilingual communication as infrastructure! National systems require hazard corpora and terminology, limited and human-assisted tools, community testing, named institutional partnerships, budgets, language-specific metrics and correction mechanisms. A way to evaluate whether these elements are a chain or are isolated commitments is offered by the Linguistic Adaptation Integration Framework.

The overall lesson is quite simple but meaningful. A warning is not linguistically inclusive because it can be translated after it is written; a policy for AI is not linguistically inclusive because an AI model can generate text in a named language. Inclusion is not just in theory but when institutions can show that people receive, understand, trust and act on critical information in the languages and the form in which they live.

Declarations

Competing interests

The authors declare no competing interests.

Funding

Funding information has been redacted for double-anonymized review and supplied in the cover letter.

Author Contribution

M.R. has contributed to conceptualization, overall preparation and editing of the paper. S.D., M.D. and D.M. have helped in formal analysis, reviewing, supervising the study.

Data Availability

The policy documents analysed in this study are publicly available from the issuing governments, the UNFCCC and official policy portals. A supplementary document register, page-level evidence table, coding framework, lexicons and aggregated matched-sentence counts accompany the manuscript. Source PDFs are not redistributed where third-party terms apply. The data can be accessed at the following URLs:1.https://mode.gov.lk/notices/AI%20Sri%20Lanka%20Sri%20Lanka%E2%80%99s%20National%20Strategy%20on%20AI%20Open%20for%20Public%20Consultation2.https://aipolicy.gov.bd/docs/national-ai-policy-bangladesh-2026-2030-draft-v2.0.pdf3.https://giwmscdnone.gov.np/media/pdf_upload/National%20AI%20Policy-Final_uxc94vg.pdf4.https://env.gov.lk/web/index.php/en/23-downloadsOthers have been provided in the reference section.

Code availability

The text-extraction and lexicon-matching procedure has been documented in the supplementary materials. The lexicon-matching script, document register and aggregated coding outputs accompany the manuscript; they may also be deposited in a DOI-bearing repository at submission.

Ethics declaration

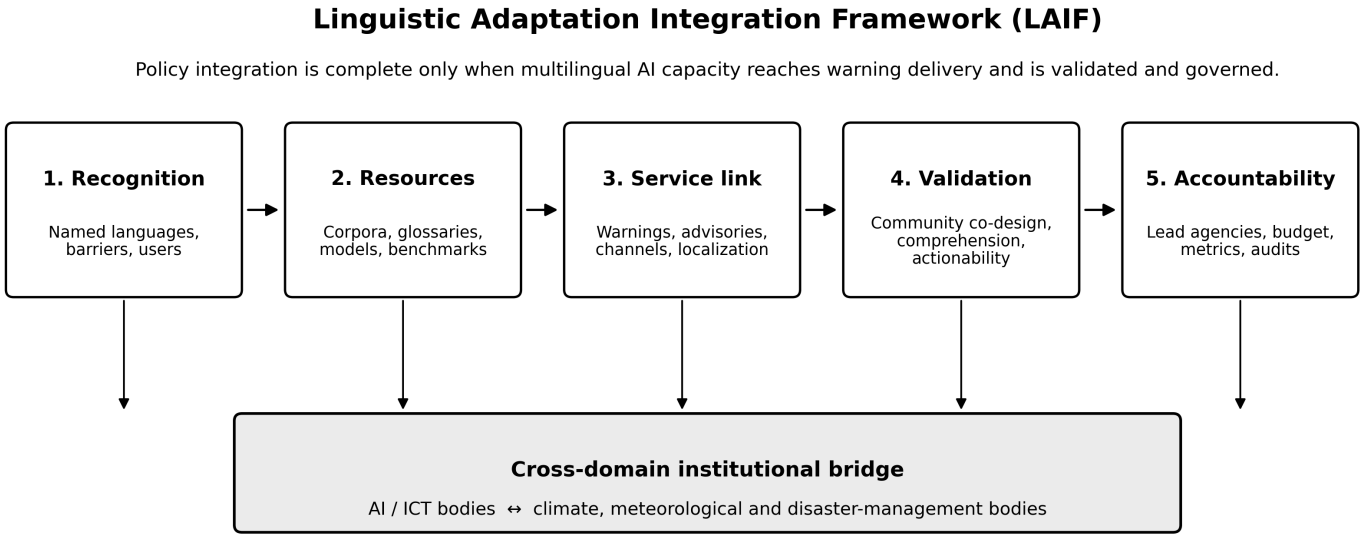
This study has analysed publicly available policy documents and involved no human participants, personal data or intervention. Institutional ethical approval and informed consent were therefore not required.

References

1. Bacchi C, Goodwin S (2016) Poststructural policy analysis: a guide to practice. Palgrave Macmillan, New York. https://doi.org/10.1057/978-1-137-52546-8
2. Baker P, Gabrielatos C, KhosraviNik M, Krzyzanowski M, McEnery T, Wodak R (2008) A useful methodological synergy? Combining critical discourse analysis and corpus linguistics to examine discourses of refugees and asylum seekers in the UK press. Discourse Soc 19(3):273–306. https://doi.org/10.1177/0957926508088962
3. Bird S (2020) Decolonising speech and language technology. In: Proceedings of the 28th International Conference on Computational Linguistics, pp 3504–3519. https://doi.org/10.18653/v1/2020.coling-main.313
4. Blasi DE, Anastasopoulos A, Neubig G (2022) Systematic inequalities in language technology performance across the world's languages. Sci Adv 8(26):eabm9128. https://doi.org/10.1126/sciadv.abm9128
5. Blodgett SL, Barocas S, Daume H III, Wallach H (2020) Language (technology) is power: a critical survey of bias in NLP. In: Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, pp 5454–5476. https://doi.org/10.18653/v1/2020.acl-main.485
6. Candel JJL, Biesbroek GR (2016) Toward a processual understanding of policy integration. Policy Sci 49(3):211–231. https://doi.org/10.1007/s11077-016-9248-y
7. Committee on Formulating a Strategy for AI, Government of Sri Lanka (2024) National strategy on artificial intelligence: AI Sri Lanka 2028, draft, July 2024. Government of Sri Lanka, Colombo. https://mode.gov.lk/notices/AI%20Sri%20Lanka%20 Sri%20Lanka%E2%80%99s%20National%20 Strategy%20on%20AI %20Open%20for%20Public%20Consultation. Accessed 21 July 2026
8. Dilling L, Lemos MC (2011) Creating usable science: opportunities and constraints for climate knowledge use and their implications for science policy. Glob Environ Change 21(2):680–689. https://doi.org/10.1016/j.gloenvcha.2010.11.006
9. Haque CE (1995) Climatic hazards warning process in Bangladesh: experience of, and lessons from, the 1991 April cyclone. Environ Manag 19(5):719–734. https://doi.org/10.1007/BF02471954
10. Information and Communication Technology Division, Government of Bangladesh (2026) National AI policy Bangladesh 2026–2030, draft V2.0. Government of Bangladesh, Dhaka. https://aipolicy.gov.bd/docs/national-ai-policy-bangladesh-2026-2030-draft-v2.0.pdf. Accessed 21 July 2026
11. Joshi P, Santy S, Budhiraja A, Bali K, Choudhury M (2020) The state and fate of linguistic diversity and inclusion in the NLP world. In: Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, pp 6282–6293. https://doi.org/10.18653/v1/2020.acl-main.560
12. Lemos MC, Kirchhoff CJ, Ramprasad V (2012) Narrowing the climate information usability gap. Nat Clim Change 2(11):789–794. https://doi.org/10.1038/nclimate1614
13. Ministry of Communication and Information Technology, Government of Nepal (2025) National artificial intelligence policy 2025. Government of Nepal, Kathmandu. https://giwmscdnone.gov.np/media/pdf_upload/National%20AI%20Policy-Final_uxc94vg.pdf. Accessed 21 July 2026
14. Ministry of Environment, Government of Sri Lanka (2023) National policy on climate change 2023. Government of Sri Lanka, Colombo. https://env.gov.lk/web/index.php/en/23-downloads. Accessed 21 July 2026
15. Ministry of Environment, Forest and Climate Change, Government of Bangladesh (2022) National adaptation plan of Bangladesh 2023–2050. Government of Bangladesh, Dhaka. https://unfccc.int/documents/637588. Accessed 21 July 2026
16. Ministry of Forests and Environment, Government of Nepal (2021) National adaptation plan 2021–2050. Government of Nepal, Kathmandu. https://unfccc.int/sites/default/files/resource/NAP_Nepal_2021.pdf. Accessed 21 July 2026
17. Mohamed S, Png MT, Isaac W (2020) Decolonial AI: decolonial theory as sociotechnical foresight in artificial intelligence. Philos Technol 33:659–684. https://doi.org/10.1007/s13347-020-00405-8
18. Moser SC (2010) Communicating climate change: history, challenges, process and future directions. WIREs Clim Change 1(1):31–53. https://doi.org/10.1002/wcc.11
19. Nekoto W, Marivate V, Matsila T et al (2020) Participatory research for low-resourced machine translation: a case study in African languages. In: Findings of the Association for Computational Linguistics: EMNLP 2020, pp 2144–2160. https://doi.org/10.18653/v1/2020.findings-emnlp.195

20. Piller I (2016) Linguistic diversity and social justice: an introduction to applied sociolinguistics. Oxford University Press, Oxford. https://doi.org/10.1093/acprof:oso/9780199937240.001.0001
21. Raji ID, Smart A, White RN et al (2020) Closing the AI accountability gap: defining an end-to-end framework for internal algorithmic auditing. In: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, pp 33–44. https://doi.org/10.1145/3351095.3372873
22. Roy C, Sarkar SK, Aberg J, Kovordanyi R (2015) The current cyclone early warning system in Bangladesh: providers' and receivers' views. Int J Disaster Risk Reduct 12:285–299. https://doi.org/10.1016/j.ijdrr.2015.02.004
23. Sambasivan N, Kapania S, Highfill H, Akrong D, Paritosh P, Aroyo LM (2021) Everyone wants to do the model work, not the data work: data cascades in high-stakes AI. In: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, article 39. https://doi.org/10.1145/3411764.3445518
24. United Nations Educational, Scientific and Cultural Organization (2021) Recommendation on the ethics of artificial intelligence. UNESCO, Paris
25. United Nations Office for Disaster Risk Reduction (2015) Sendai framework for disaster risk reduction 2015–2030. UNDRR, Geneva
26. Uekusa S, Matthewman S (2023) Preparing multilingual disaster communication for the crises of tomorrow: a conceptual discussion. Int J Disaster Risk Reduct 87:103589. https://doi.org/10.1016/j.ijdrr.2023.103589
27. Van Horne YO, Alcala CS, Peltier RE et al (2023) Toward language justice in environmental health sciences in the United States. Environ Health Perspect 131(8):085001. https://doi.org/10.1289/EHP12306
28. Vaughan C, Dessai S (2014) Climate services for society: origins, institutional arrangements, and design elements for an evaluation framework. WIREs Clim Change 5(5):587–603. https://doi.org/10.1002/wcc.290
29. World Meteorological Organization (2022) Early warnings for all: executive action plan 2023–2027. WMO, Geneva

Figures



The linguistic adaptation gap appears where recognition or language resources do not travel through the full chain.

Figure 1

Linguistic Adaptation Integration Framework (LAIF). The framework treats multilingual climate-risk communication as a five-stage institutional chain: recognition, language-resource infrastructure, service linkage, participatory validation and accountability. The

linguistic adaptation gap appears when recognition or upstream language resources do not travel through the complete chain. The horizontal bridge represents formal coordination among AI/ICT, language, meteorological, climate-adaptation, disaster-management, telecommunications and local-government bodies.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- HSSCsupplementarycodingmatrix.csv