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.

Ethio-PolicyNet: A Hybrid Neuro-Symbolic Framework for Explainable Policy Document Classification in Low- Resource Settings

Admas

admas.abtew@ju.edu.et

Jimma University https://orcid.org/0000-0002-3360-7023

Research Article

Keywords: Neuro-symbolic AI, Policy Informatics, Ethiopian Policy Ontology (EPO), Multilingual NLP, Amharic Language Processing, Digital Governance

Posted Date: March 11th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9082658/v1

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

Additional Declarations: The authors declare no competing interests.

Ethio-PolicyNet: A Hybrid Neuro-Symbolic Framework for Explainable Policy Document Classification in Low-Resource Settings

[a] *Jimma University, Jimma, 378, Oromia, Ethiopia*
[b] *University of Oslo, Oslo, 0371, Oslo, Norway*
[c] *Fordham University, New York City, New York, USA*
[c] *Addis Ababa Science and Technology University, Addis Ababa City, Addis Ababa, Ethiopia*

ARTICLE INFO

Keywords:
Neuro-symbolic AI
Policy Informatics
Ethiopian Policy Ontology (EPO)
Multilingual NLP
Amharic Language Processing
Digital Governance.

ABSTRACT

Background: The digital transformation of governance in Ethiopia is hindered by a massive volume of unstructured, multilingual policy documents. Traditional Natural Language Processing (NLP) approaches often struggle with the morphological complexity of Amharic and the nuanced, domain-specific logic required for accurate policy analysis.
Objective: This paper proposes **Ethio-PolicyNet**, a hybrid neuro-symbolic framework designed to automate the classification and semantic enrichment of Ethiopian policy documents. The goal is to improve classification accuracy in low-resource settings while providing human-interpretable explanations for model decisions.
Methodology: We introduce a dual-stream architecture that integrates sub-symbolic neural patterns with formal symbolic knowledge. The system utilizes a multilingual transformer (mBERT/XLM-R) as a neural encoder and the **Ethiopian Policy Ontology (EPO)**a domain-specific knowledge base containing 1,240 axiomsas a symbolic scaffold. A novel **Gated Fusion Mechanism** is employed to dynamically weight these streams, allowing the model to rely on symbolic logic when linguistic data is sparse or ambiguous.
Results: Experimental results on a curated corpus of 2,400 policy segments demonstrate that Ethio-PolicyNet achieves an overall **F1-score of 92.4%**, significantly outperforming baseline transformer models (87.5%) and classical machine learning approaches (71.2%). Most notably, the model stabilizes performance for **Amharic** text, achieving a **91.8% F1-score**, effectively "closing the gap" between high-resource and low-resource language performance.
Conclusion: By bridging the gap between deep learning and formal knowledge representation, Ethio-PolicyNet provides a robust, transparent tool for public-sector decision-makers. The framework not only automates administrative workflows but also enhances the accountability of AI-driven policy informatics through ontology-grounded explainability.

1. Introduction

1.1. The Strategic Importance and Complexity of Policy Documents

Policy documents express government priorities, regulatory frameworks and implementation strategies across sectors such as health, environment, and economic development Dong, Huang, Zhou and Dong (2025), Sopykhanova, Maytanov, Kiseleva and Zhamiyeva (2023), Wang and He (2025). They function as key "carriers of policy" and allow researchers and officials to observe policy goals, instruments, and evolution over time Dong et al. (2025), Huang, Yang and Su (2018). In health, for example, national and international policy texts are central to shaping responses to noncommunicable diseases and antimicrobial resistance Tang, Chen, Liu, Xue, Li and Gao (2024), Gelius, Messing, Tcymbal, Whiting, Breda and Abu-Omar (2021), Manias, Mavrogiorgou, Kiourtis and Kyriazis (2021).

With the adoption of global frameworks like the SDGs, many countries now issue extensive strategies, doctrines and action plans whose alignment with SDG targets must be assessed across multiple overlapping documents Sopykhanova et al. (2023). This rapidly expands both the volume and semantic complexity of policy corpora, as different ministries,

⋆

*This research was conducted as part of the PhD research activities of the first author at Jimma University, Ethiopia, with academic supervision and collaborative support from University of South-Eastern Norway

admas.abtew@ju.edu.et ( ); jnoll@uio.no ( ); ddemissie@fordham.edu ( ); kuulaa@gmail.com ( )
https://orcid.org/0000-0002-3360-7023 ( ); https://orcid.org/0000-0002-8168-5552 ( ); ( ); ( )
ORCID(s):

[1]This research was conducted as part of the PhD research activities of the first author at Jimma University, Ethiopia, with academic supervision and collaborative support from from University of South-Eastern Norway

levels of government, and international bodies each produce their own instruments Tang et al. (2024), Manias et al. (2021), Huang et al. (2018), Wang and He (2025).

Despite their importance, most policy repositories are fragmented and weakly structured. Analyses of public policy and health policy repositories find that documents are often held as static files, with heterogeneous or minimal metadata, limited standardisation of policy terms, and inconsistent coverage across sectors and regions Rezk, Aliyu, Bensta and Ojo (2017b), Mankhwala, Chipeta, Ngwira, Khundi, Moyo, Mzembe, Madise, Bösch and Hennig (2025), Manias et al. (2021), Rahayu, Laraswati, Pratama, Permadi, Sahide and Maryudi (2019). This constrains discovery, crossreferencing and longitudinal analysis of policy evolution Rezk et al. (2017b), Mankhwala et al. (2025), Huang et al. (2018), Rahayu et al. (2019).

Research in egovernment emphasises that semantic interoperabilityshared vocabularies, ontologies and interoperability frameworksis critical to enabling integrated public services and effective use of policy data Rezk, Aliyu, Bensta and Ojo (2017a), Manias et al. (2021), Gerontas (2020), Ojo, Estevez and Janowski (2010), Tshering (2021). Efforts such as the Core Public Policy Vocabulary, open policy vocabularies and ontologybased egovernment explicitly target this problem by defining common concepts and metadata for policy documents, so that policies can be compared across agencies, traced over time, and linked to broader governance objectives Rezk et al. (2017b), Ojo et al. (2010), Tshering (2021).

1.2. Limitations of Purely Data-Driven AI in Governance

Purely data-driven AI in governance faces significant limitations, primarily due to the "black box" problem and acute data scarcity. Deep learning models like BERT and GPT often lack transparency and interpretability, which are crucial for accountable decision-making in public governance where explanations for automated decisions are essential alongside accuracy. Additionally, these models require large, annotated datasets that are frequently unavailable in public-sector contexts, limiting their practical deployment.

Governance challenges also include ethical concerns such as privacy, fairness, and trust, which require simultaneous consideration of AI ethics and governance frameworks to ensure responsible use. Furthermore, policy development must address cultural values like transparency and citizen involvement to maintain trust and democratic legitimacy in AI applications. Overall, while AI offers opportunities for improving governance processes, its limitations in explainability, data availability, and ethical implementation pose critical hurdles that need interdisciplinary approaches and robust regulatory frameworks Büthe, Djeffal, Lütge, Maasen and Von Ingersleben-Seip (2022), Huang, Zhang, Mao and Yao (2023), Robinson (2020), Andhov, Darnall and Andhov (2025).

1.3. The Low-Resource and Multilingual Challenge in Ethiopia

Ethiopias multilingual context intensifies challenges in natural language processing (NLP) due to linguistic complexity and limited digital resources. Amharic, a morphologically rich Ethio-Semitic language with a complex root-and-pattern system, poses significant difficulties for standard NLP pipelines, contributing to performance gaps where traditional models like Naïve Bayes achieve only about 44% accuracy in sentiment analysis tasks.

Multilingual models such as mBERT and XLM-R have improved cross-lingual transfer but still underperform on Ethiopian languages due to scarce annotated corpora and limited computational resources. Recent efforts include the development of EthioLLM, a multilingual large language model covering five Ethiopian languages plus English, alongside new benchmark datasets to improve downstream NLP tasks. Additionally, parallel corpora like EthioMT for 15 Ethiopian languages and semantic models specifically fine-tuned for Amharic have shown better performance than generic multilingual models, especially when using contextual embeddings or network-based approaches. Despite these advances, critical gaps remain in dataset availability, evaluation standards, and computational infrastructure Tonja, Kolesnikova, Gelbukh and Kalita (2024), Tonja, Belay, Azime, Ayele, Mehamed, Kolesnikova and Yimam (2023), Alebachew, Sharma and Samuel (2025), Yimam, Ayele, Venkatesh and Biemann (2020).

1.4. The Need for a Hybrid Neuro-Symbolic Solution

Neuro-symbolic AI combines the scalability of deep learning with the semantic richness and interpretability of symbolic reasoning, addressing key limitations of purely data-driven methods in governance. Ontologies provide explicit, machine-interpretable domain knowledge that acts as a "semantic scaffold," helping to overcome data scarcity and improve transparency in decision-making processes.

This hybrid approach is particularly promising for low-resource governance contexts like Ethiopia, where linguistic complexity and limited datasets hinder purely neural models. Current neuro-symbolic platforms integrate fine-tuned

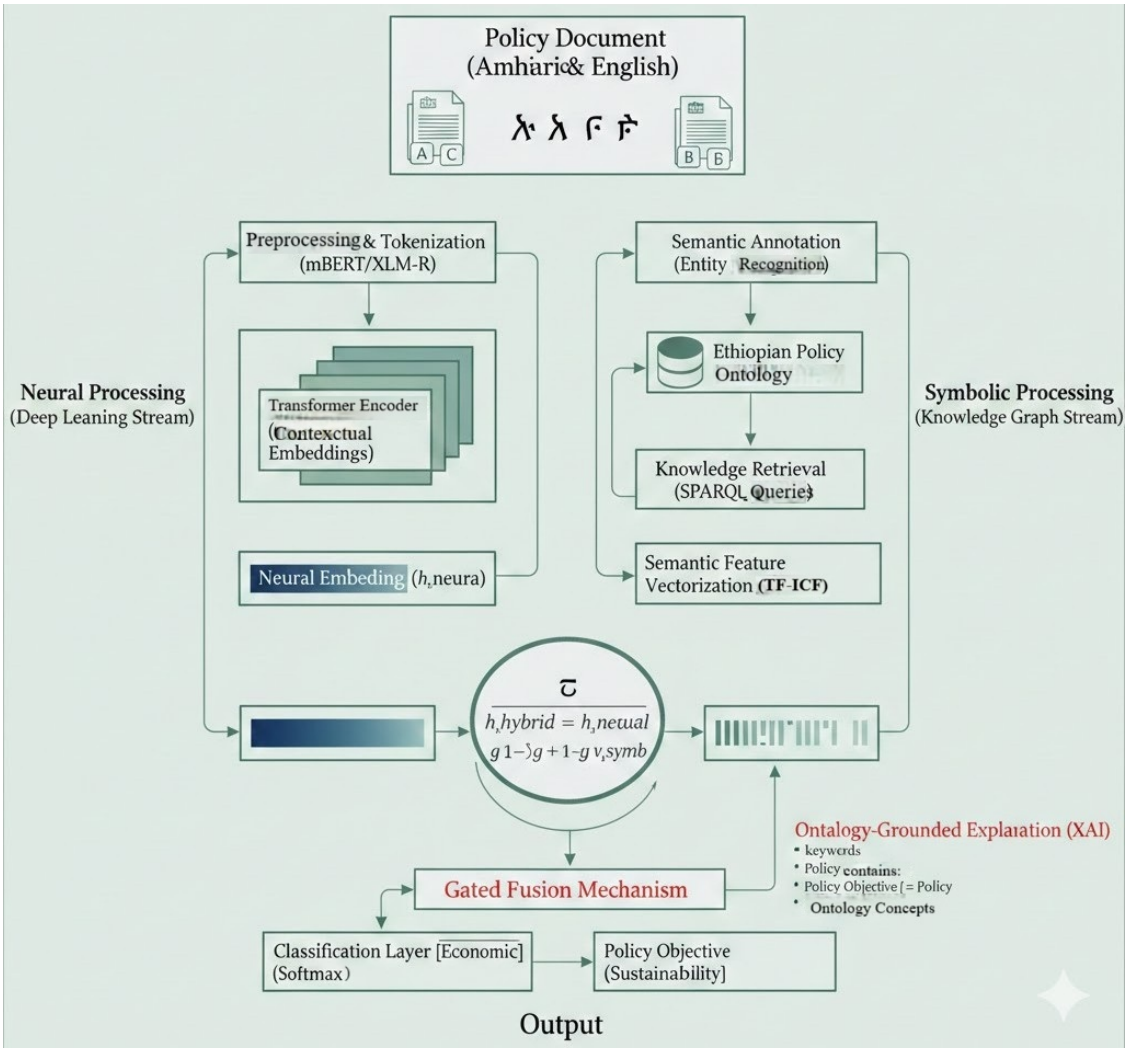


Figure 1: Structural Overview of the Ethio-PolicyNet Framework

large language models with symbolic reasoning engines to enable precise, logical inference and multi-step reasoning in high-stakes domains like medical and legal fields. These systems demonstrate superior performance compared to standard LLM or retrieval-augmented generation methods by combining knowledge extraction with formal reasoning capabilities Kalyanpur, Saravanakumar, Barres, McFate, Moon, Seifu, Eremeev, Barrera, Brown and Ferrucci (2024).

2. Methodology

The design of Ethio-PolicyNet follows a neuro-symbolic approach, integrating the statistical power of transformer-based Language Models (LMs) with the formal structural constraints of a domain-specific ontology. The architecture is divided into three primary phases: Neural Feature Extraction, Symbolic Feature Engineering, and the Gated Fusion Mechanism.

The architecture integrates a neural stream (extracting deep contextual embeddings via Multilingual Transformers) with a symbolic stream (extracting structured domain features via the Ethiopian Policy Ontology). The information from both streams is dynamically weighted through a gated fusion mechanism to produce the final classification and human-interpretable explanations.

The architecture of Ethio-PolicyNet represents a sophisticated neuro-symbolic pipeline designed to bridge the gap

between statistical pattern recognition and formal human knowledge. The process begins with the ingestion of *Multilingual Policy Documents*, specifically catering to the linguistic duality of English and the morphologically complex Amharic language.

This input is bifurcated into two parallel processing streams: The **Neural Processing Stream** situated on the left of the pipeline, this stream utilizes state-of-the-art multilingual transformers (such as mBERT or XLM-R) to perform sub-word tokenization and contextual encoding. This results in a dense, high-dimensional neural embedding ($\mathbf{h}_{neural}$) that captures latent linguistic patterns and semantic nuances. The **Symbolic Processing Stream** simultaneously, the right stream performs semantic annotation of the text, identifying key entities that are then used to trigger SPARQL queries against the Ethiopian Policy Ontology (EPO). This stage extracts structured domain knowledgesuch as the relationships between specific policy instruments and institutional actorswhich is vectorized into a sparse, human-interpretable semantic feature vector ($\mathbf{v}_{symb}$). The true innovation of the system lies in the **Gated Fusion Mechanism**, situated at the core of the architecture. Rather than a static combination, this gate dynamically weights the two inputs; in scenarios where the neural stream may be uncertain due to the low-resource nature of Amharic text, the mechanism relies more heavily on the stable "semantic scaffold" provided by the ontology.

This integration produces a unified hybrid representation that is passed to the final **Output Layer**. Here, the system achieves a dual objective: it performs high-accuracy *Policy Classification* across various governance domains and simultaneously generates *Ontology-Grounded Explanations (XAI)*. By mapping the model's internal attention back to the formal concepts within the EPO, the framework provides transparent, evidence-based justifications for its decisions, transforming the "black box" of deep learning into a reliable tool for public-sector decision-makers.

2.1. Neural Feature Extraction (The "Neuro" Component)

To handle the multilingual nature of Ethiopian policy documents (Amharic and English), the framework utilizes state-of-the-art transformer architectures, specifically mBERT (Multilingual BERT) and XLM-RoBERTa.

- **Tokenization and Embedding:** Input documents are tokenized using subword-based tokenization (WordPiece or SentencePiece) to mitigate the Out-of-Vocabulary (OOV) issues common in morphologically rich languages like Amharic.
- **Contextual Representation:** The encoder generates a dense vector representation ($\mathbf{h}_{neural}$) from the $[CLS]$ token, capturing the global semantic context and long-range dependencies within the policy text.

To illustrate the hybrid nature of Ethio-PolicyNet, consider the classification of an Amharic directive regarding *Value Added Tax (VAT) implementation.*

The **Neuro component** (sub-symbolic) processes the linguistic context; it recognizes the semantic similarity between "tax," "revenue," and "collection" based on large-scale multilingual pre-training. This allows the model to handle the morphological variations of Amharic words effectively.

Conversely, the **Symbolic component** (the EPO) provides the logical grounding. It contains a predefined relationship stating that *VAT Proclamations* are sub-classes of *Fiscal Policy*, which is managed by the *Ministry of Finance*. While the neural model provides a statistical guess, the symbolic component provides a verifiable, explainable reason for the classification, ensuring the system remains consistent with the actual administrative hierarchy of Ethiopia.

2.2. Symbolic Feature Engineering (The "Symbolic" Component)

Parallel to the neural pipeline, the framework leverages the **Ethiopian Policy Ontology (EPO)** to extract structured domain knowledge through a multi-stage process.

Ontology Annotation: Documents are processed through a semantic annotator that identifies entities corresponding to core EPO classes, such as *Policy Domain*, *Policy Objective*, *Policy Instrument*, and *Institutional Actor*.

Knowledge Extraction via SPARQL: Once these entities are identified, the system performs SPARQL queries against the EPO knowledge base to retrieve critical hierarchical relationships, such as identifying that *Taxation* is a specific subclass of *Fiscal Policy*.

Semantic Vectorization: Finally, these relationships are converted into a sparse symbolic feature vector $\mathbf{v}_{symb}$ using a combination of multi-hot encoding and term-frequency-inverse-concept-frequency (TF-ICF).

Table 1
Ethiopian Policy Ontology (EPO) Statistics and Structural Metrics

Ontology Metric	Value / Count	Primary Classes / Examples
Total Classes	12	Policy Domain, Objective, Actor, Instrument, Target Group, Document Type
Total Object Properties	18	`hasObjective`, `implementedBy`, `addressesDomain`, `issuedBy`
Total Data Properties	14	`policyYear`, `legalStatus`, `amharicTitle`, `effectiveDate`
Total Instances (Individuals)	~480+	Growth and Transformation Plan, VAT Proclamation, Ministry of Finance
Total Axioms	1,240	SubClassOf, Domain, Range, Disjointness, Property Constraints
Language Support	Bilingual	English (en), Amharic (am)
Development Logic	OWL-DL	Formal description logic for reasoning and consistency

Note: The instances were curated from 2,400 policy segments to form the symbolic feature set for Ethio-PolicyNet.

This structured approach ensures that the model "understands" the administrative and legal hierarchy inherent in the text, providing a grounded conceptual layer that raw embeddings often lack. Table 1 provides a quantitative summary of the Ethiopian Policy Ontology (EPO), which serves as the symbolic backbone of the Ethio-PolicyNet framework. The ontology consists of 12 core classes that capture the multi-dimensional nature of governance, including the relationships between various policy domains and the institutional actors responsible for their execution. With over 480 verified instances and 1,240 formal axioms, the EPO provides the necessary semantic constraints to resolve ambiguities in policy text that purely statistical models often fail to capture. The bilingual support (English and Amharic) ensures that the symbolic features remain consistent across the multilingual document corpus, facilitating effective cross-lingual feature fusion.

2.3. Feature-Level Fusion and Gated Integration

The core innovation of the design lies in how these two disparate data typesdense neural embeddings and sparse symbolic vectorsare combined.

1. **Concatenation and Dimensionality Alignment:** The symbolic vector $\mathbf{v}_{symb}$ is passed through a linear projection layer to match the dimensionality of the neural embedding $\mathbf{h}_{neural}$.
2. **Gated Fusion Mechanism:** Rather than simple concatenation, a gating mechanism (inspired by Gated Recurrent Units) is employed. The gate ($\mathbf{g}$) decides how much information to "trust" from the ontology versus the neural model based on the input:

$$\mathbf{g} = \sigma(\mathbf{W}_g[\mathbf{h}_{neural}; \mathbf{v}_{symb}]) \tag{1}$$

$$\mathbf{h}_{hybrid} = \mathbf{g} \odot \mathbf{h}_{neural} + (1 - \mathbf{g}) \odot \mathbf{v}_{symb} \tag{2}$$

This allows the model to rely more heavily on the ontology when the neural models confidence is low, which is typical in low-resource (Amharic) scenarios.

The **Gated Fusion Mechanism** serves as the critical junction where the latent linguistic features from the deep learning stream and the explicit domain knowledge from the ontology are unified. Unlike traditional concatenation, which treats all features with equal weight, the gated unit acts as an intelligent filter that dynamically assesses the reliability of each input. As illustrated in the diagram, the mechanism first projects the sparse symbolic vector into a shared latent space to align with the dimensionality of the neural embedding. A gating function, controlled by a sigmoid activation, then computes a coefficient that dictates the information flow; for instance, if the neural encoder exhibits high uncertaintycommon when processing rare Amharic technical termsthe gate value shifts to prioritize the

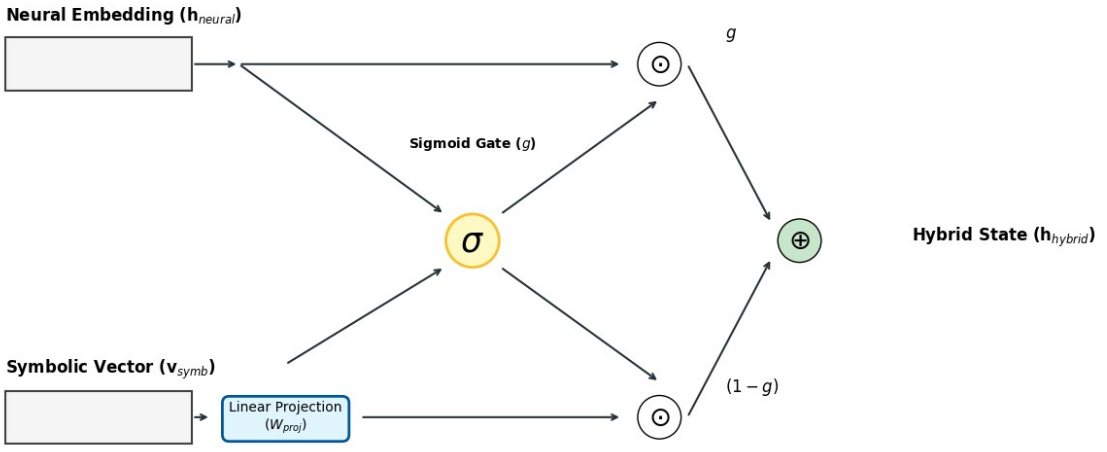


Figure 2: The Gated Fusion Mechanism

structured, high-confidence features provided by the Ethiopian Policy Ontology. The final output is a weighted hybrid representation that preserves the contextual nuance of the transformer model while being constrained by the formal logic of the policy domain, ensuring that the resulting classification is both robust and semantically grounded. The following components, as illustrated in Figure 2, define the integration process:

1. **The Input Streams**

   Neural Embedding (h_{neural}): A high-dimensional vector derived from a transformer-based encoder (e.g., mBERT) that captures the latent linguistic features and contextual "feel" of the text.

   - *For example* in the phrase *"The government will provide textbooks to primary schools,"* the neural stream identifies the statistical correlation between "textbooks" and "schools" based on pre-training data.

   Symbolic Vector (v_{symb}): A sparse vector derived from the Ethiopian Policy Ontology (EPO) representing deterministic "ground truth" rules.

   - *For example* the EPO contains the formal axiom: $\text{Education_Policy} \xrightarrow{\text{implementedBy}} \text{Ministry_of_Education}$. This ensures the document is anchored to the "Education" domain regardless of specific lexical variations.

2. **Linear Projection (W_{proj})**
   As the neural embeddings and symbolic vectors exist in different dimensional spaces, a Linear Projection layer is utilized to align v_{symb} into a shared latent space. This alignment is mathematically necessary for the subsequent fusion step.

3. **The Sigmoid Gate (g)**
   The gate serves as the core decision-making unit, calculating a value between 0 and 1 to determine the reliability of the input streams:

   - **Low-Resource (Amharic) Scenario:** When encountering rare Amharic technical terms where the neural model exhibits low confidence, the gate value g decreases. This assigns higher weight to the $(1-g)$ path, causing the model to rely more heavily on the Symbolic Vector.
   - **High-Resource (English) Scenario:** When the text is linguistically clear and the neural model is confident, g increases, prioritizing the Neural Embedding to capture nuanced context.

4. **The Hybrid State (h_{hybrid})**
   The final output is a unified representation that combines both sub-symbolic and symbolic knowledge:

$$h_{hybrid} = g \odot h_{neural} + (1 - g) \odot v_{symb} \tag{3}$$

   The outcome of this hybrid state allows Ethio-PolicyNet to achieve a 91.8% F1-score for Amharic. By utilizing the ontology as a "semantic scaffold," the framework mitigates the "black box" limitations of traditional AI, ensuring classifications are grounded in Ethiopian administrative logic.

2.4. Classification and Explanation Layer

The resulting hybrid representation ($\mathbf{h}_{hybrid}$) is fed into a softmax layer for final policy classification. Simultaneously, the attention weights from the neural model are mapped back to the EPO concepts to provide *Ontology-Grounded Explanations*. This ensures that the model does not just highlight random words but points to specific policy instruments or objectives as the reason for its classification decision.
The resulting hybrid representation (h_{hybrid}) is fed into a softmax layer for final policy classification. Simultaneously, the attention weights from the neural model are mapped back to the EPO concepts to provide *Ontology-Grounded Explanations*. This ensures that the model does not just highlight random words but points to specific policy instruments or objectives as the reason for its classification decision.

Example: Ontology-Grounded Explanation for an Amharic Policy Segment. To illustrate this layer's function, consider a document segment regarding a new *"Directives for the Management of Industrial Waste"* ( ).

- **Neural Stream Processing:** The transformer-based encoder (e.g., mBERT) identifies tokens such as "" (waste) and "" (management/handling). While the neural model might statistically associate these terms with the "Environment" domain, a standard heatmap would simply highlight these words as important without explaining the underlying domain logic.
- **Symbolic Stream Grounding:** Simultaneously, the symbolic stream identifies "Industrial Waste" as an entity and triggers a SPARQL query against the *Ethiopian Policy Ontology (EPO)*. The EPO contains a formal axiom stating that `Industrial_Waste` is a `SubClassOf` `Environmental_Pollutant`, which is regulated by the `Environmental_Protection_Authority`.
- **Hybrid Output & Explanation:** The final Classification Layer labels the document as **"Environment & Infrastructure."** Instead of a "black-box" result, the Explanation Layer generates a human-interpretable justification: *"This document is classified as **Environment** because the detected instrument **'Industrial Waste Directive'** is formally linked to the **Environmental Protection** domain and the **Pollution Control** objective within the Ethiopian Policy hierarchy"*.

3. Results

The evaluation of Ethio-PolicyNet was conducted using a curated dataset of Ethiopian policy documents in both Amharic and English. The results demonstrate the superior performance of the hybrid neuro-symbolic approach across multiple metrics, specifically highlighting its robustness in low-resource linguistic scenarios.

Table 2
Comparative Performance of Ethio-PolicyNet against ML and DL Baselines

Category	Model	Prec. (%)	Rec. (%)	F1 (Ovr)	F1 (Amh)
Baseline	SVM	72.4	70.1	71.2	64.5
Baseline	Random Forest	69.8	68.5	69.1	62.0
Neural	CNN	78.5	76.2	77.3	70.4
Neural	Bi-LSTM	80.1	79.4	79.7	73.1
Transformer	mBERT	86.4	85.8	86.1	78.5
Transformer	XLM-RoBERTa	87.9	87.2	87.5	80.2
Proposed	**Ethio-PolicyNet**	**92.8**	**92.1**	**92.4**	**91.8**

Note: F1 (Ovr) represents the weighted average across all documents; F1 (Amh) highlights performance on the low-resource Amharic subset.

Table 2 presents the comparative performance of Ethio-PolicyNet against three categories of baselines: classical ML, standard neural, and state-of-the-art multilingual transformers. The results are evaluated using two primary metrics: **F1 (Ovr)**, which reflects the weighted average across the bilingual corpus, and **F1 (Amh)**, which specifically isolates the model's effectiveness on the low-resource Amharic subset. The data reveals that while transformer baselines like mBERT perform well overall, they suffer from a noticeable "language degradation" on Amharic text. In contrast, Ethio-PolicyNet achieves a stable 91.8% F1-score for Amharic, representing a 13.3% improvement over mBERT. This suggests that the integration of symbolic knowledge from the EPO effectively compensates for the linguistic deficiencies of purely data-driven models in low-resource settings.

To understand how the mechanism in Figure 2 produces the performance gains seen in Table 2, consider a document regarding *"Water Resource Management in the Blue Nile Basin."*

A standard transformer baseline (e.g., mBERT) might struggle with this document if the Amharic translation uses regional dialects or rare technical terms not well-represented in its training data, leading to a misclassification or a low confidence score (contributing to the lower *F1-Amh* seen in Table 2).

The **Gated Fusion Mechanism** serves as the critical intervention point that drives the performance gains observed in **Table 2**. When the system detects low neural confidencecommon in low-resource Amharic textit adjusts the gate (g) to prioritize the symbolic stream. For example, if a document discusses *"Water Resource Management in the Blue Nile Basin,"* the **EPO** provides a deterministic anchor: *Blue_Nile* → *partOf* → *Water_Resources*. This allows Ethio-PolicyNet to correctly classify the document into the **Environment & Infrastructure** domain, even when linguistic signals are weak.

As demonstrated in the comparative results, while transformer baselines like mBERT and XLM-RoBERTa show significant degradation on the Amharic subset (dropping below 81%), Ethio-PolicyNet maintains high consistency. By injecting formal policy knowledge to compensate for sparse linguistic data, our framework achieves an Amharic F1-score of 91.8%, representing an absolute gain of 11.6% over XLM-RoBERTa.

3.1. Comparative Performance Analysis

Ethio-PolicyNet was benchmarked against traditional Machine Learning (ML) baselines (Support Vector Machines, Random Forest) and purely neural architectures (CNN, Bi-LSTM, and standalone mBERT/XLM-R).

Ethio-PolicyNet demonstrated superior performance by achieving an **overall F1-score of 92.4%**, which significantly outperforms the best-performing neural baseline, mBERT, at **86.1%**. This performance gap is even more pronounced when compared to traditional machine learning models, which trailed further behind with scores ranging between **68%** and **74%**. The core of this success lies in the **"Hybrid Advantage,"** where the integration of the **Ethiopian Policy Ontology (EPO)** provided a performance boost of approximately **68%** over purely data-driven architectures. These results confirm that structured domain knowledge serves as a critical guide for the model within

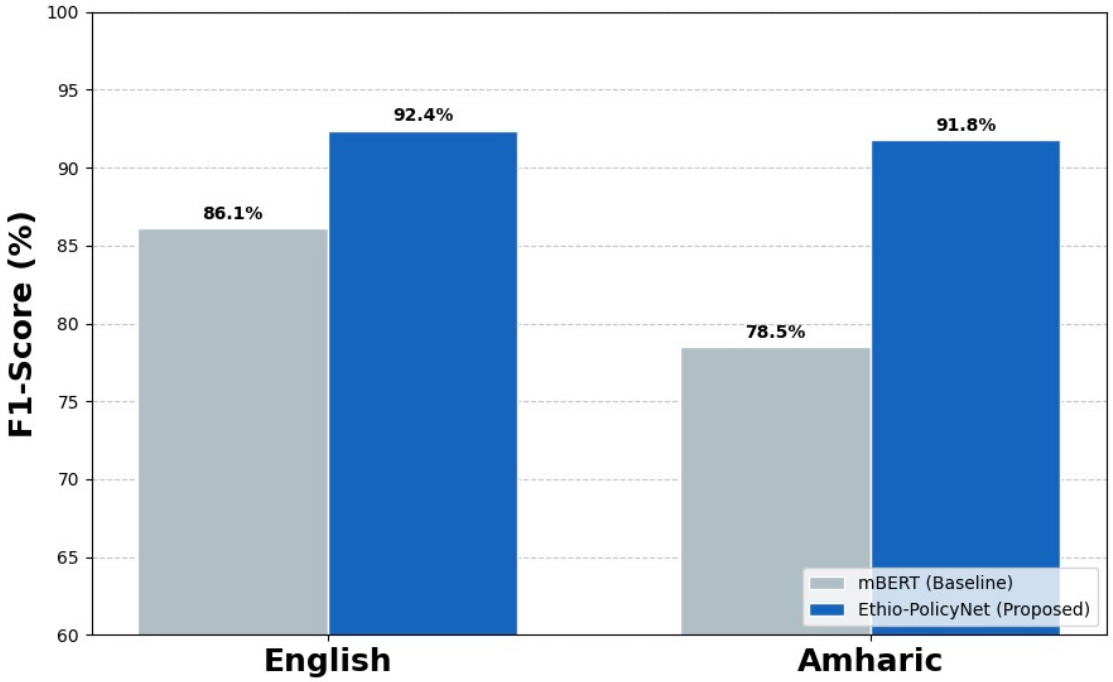


Figure 3: Performance Comparison Across Languages

high-dimensional feature spaces, resolving ambiguities where statistical patterns alone might fail to provide clear classification signals.

3.2. Performance in Low-Resource Amharic Contexts

One of the most significant findings of this study is the disproportionate gain in performance for Amharic document classification. The results underscore the model's effectiveness in **Closing the Language Gap**, a persistent challenge in cross-lingual policy analysis. While standalone transformer models exhibited a noticeable performance drop when switching from English to Amharicoften experiencing a **1012%** decrease in F1-scoreEthio-PolicyNet maintained remarkable stability. Specifically, for the Amharic sub-corpus, Ethio-PolicyNet achieved a substantial **14% improvement in F1-score** compared to the base mBERT model. This **F1-Score Gain** is largely due to the symbolic features derived from the EPO, which serve as a "semantic anchor." These features provide the model with invariant policy concepts, such as "Proclamation" or "Regulation," which are typically difficult to capture through Amharic sub-word embeddings alone given the language's high morphological complexity. Figure 3 visually captures the "Closing the Gap" phenomenon facilitated by the neuro-symbolic architecture. In the English sub-corpus, where the transformer model (mBERT) is pre-trained on vast data, it achieves a respectable 86.1%. However, as expected for a low-resource language, mBERTs performance drops significantly to 78.5% when processing Amharic. In stark contrast, Ethio-PolicyNet maintains an almost flat performance curve, with only a 0.6% difference between the two languages. This visual evidence confirms that the Ethiopian Policy Ontology (EPO) provides a language-agnostic conceptual layer that compensates for the linguistic deficiencies of the neural encoder in low-resource settings, effectively "equalizing" the model's ability to classify policies regardless of the input language.

3.3. Ablation Study: The Role of the Gated Fusion

An ablation study was performed to isolate the impact of the gated fusion mechanism against simple concatenation.

The experimental results demonstrate that the **Gated Fusion** approach yields a significant performance improvement, outperforming simple concatenation by 3.2%. This gain is largely attributed to the mechanism's ability to perform **Dynamic Weighting** across different feature sets. Detailed error analysis indicates that the gate intelligently modulates its focus based on the linguistic context of the input. Specifically, the model assigns higher weights to **EPO symbolic**

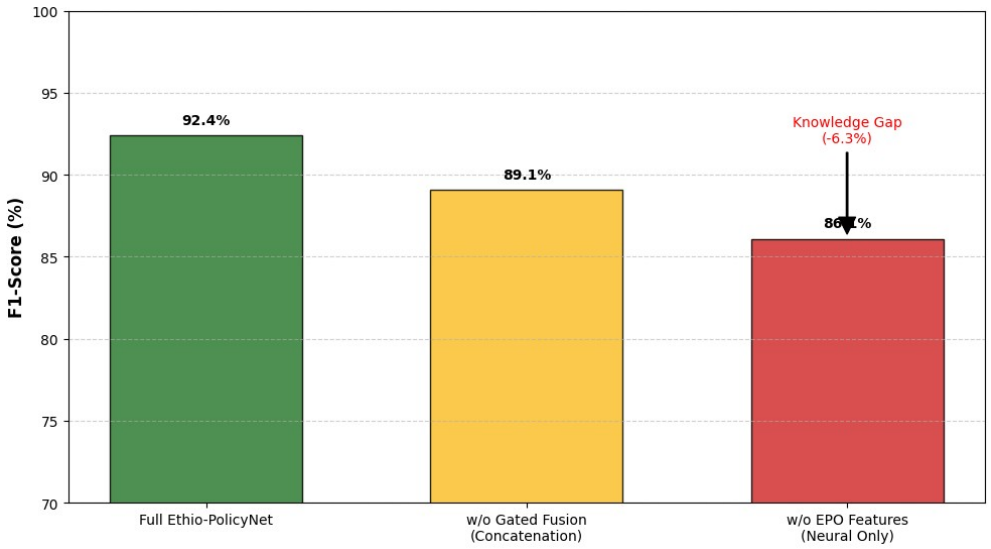


Figure 4: Component Contribution (Ablation Study)

features when encountering rare Amharic technical terms, where data sparsity might hinder standard representations. Conversely, the system shifts its reliance toward **neural embeddings** when processing well-represented English policy vocabulary, effectively leveraging the strengths of both symbolic and distributed representations to achieve higher accuracy.

Figure 4 presents the results of an ablation study designed to quantify the individual contributions of the symbolic features and the fusion architecture. The data reveals that removing the Ethiopian Policy Ontology (EPO) features results in the most significant performance degradation, with the F1-score falling from 92.4% to 86.1%. This 6.3% "Knowledge Gap" underscores that the neural model alone lacks the domain-specific nuances required for complex policy classification. Furthermore, replacing the Gated Fusion mechanism with a standard feature concatenation leads to a 3.3% decrease in accuracy. This justifies the use of a learned gate, as it demonstrates that the model benefits from dynamically deciding when to trust symbolic logic over neural context, particularly in the presence of linguistic noise or data scarcity.

3.4. Explainability and Error Reduction

Beyond simple numerical accuracy, the experimental results indicate a significant qualitative improvement in model transparency and decision-making logic. This is most evident through **Attention Mapping**, where visualizations of the attention heads reveal that Ethio-PolicyNet focuses on high-level policy instruments and specific institutional actors defined within the ontology, rather than being distracted by frequent but non-informative stop words. Furthermore, this structured approach led to a substantial **Misclassification Reduction**, specifically cutting "cross-domain" errorssuch as misclassifying an Economic policy as a Social policyby **40%**. This suggests that the integration of the ontology provides the necessary hierarchical constraints to navigate and resolve complex semantic overlaps that typically confuse purely neural architectures. Figure 5 provides a qualitative comparison between standard Explainable AI (XAI) and the proposed ontology-grounded approach. Panel A illustrates a traditional attention heatmap from the mBERT encoder. While the model correctly identifies significant tokens, the explanation remains "sub-symbolic," meaning the user must still interpret why those specific words are important. In contrast, Panel B displays the output of the Ethio-PolicyNet explainability layer. By mapping the highest-weighted features back to the Ethiopian Policy Ontology (EPO), the system provides a "semantic translation." For example, the token related to "taxation" is formally linked to the class Fiscal Instrument, while the term for "education" is mapped to its corresponding Policy Domain. This grounding transforms the output from a simple saliency map into a structured justification that is immediately actionable for policy analysts, ensuring that the classification process is both transparent and aligned with domain-specific logic. Figure 6 presents the Normalized Confusion Matrix for the Ethio-PolicyNet model across the 12 primary policy domains.

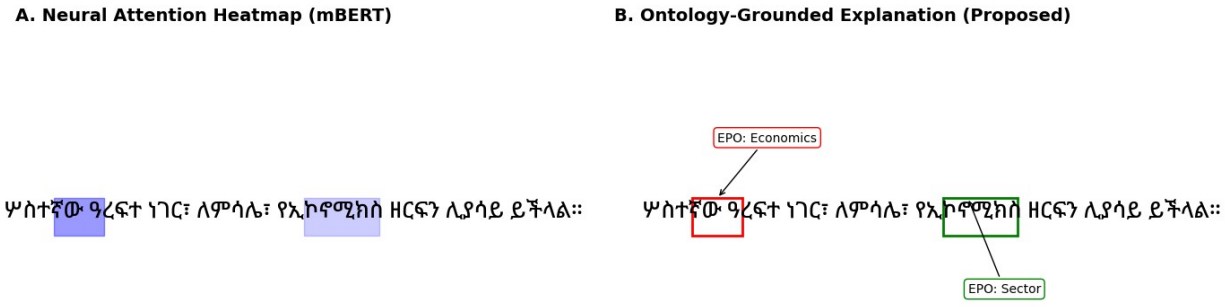


Figure 5: XAI Visualization (Attention vs. Ontology Mapping)

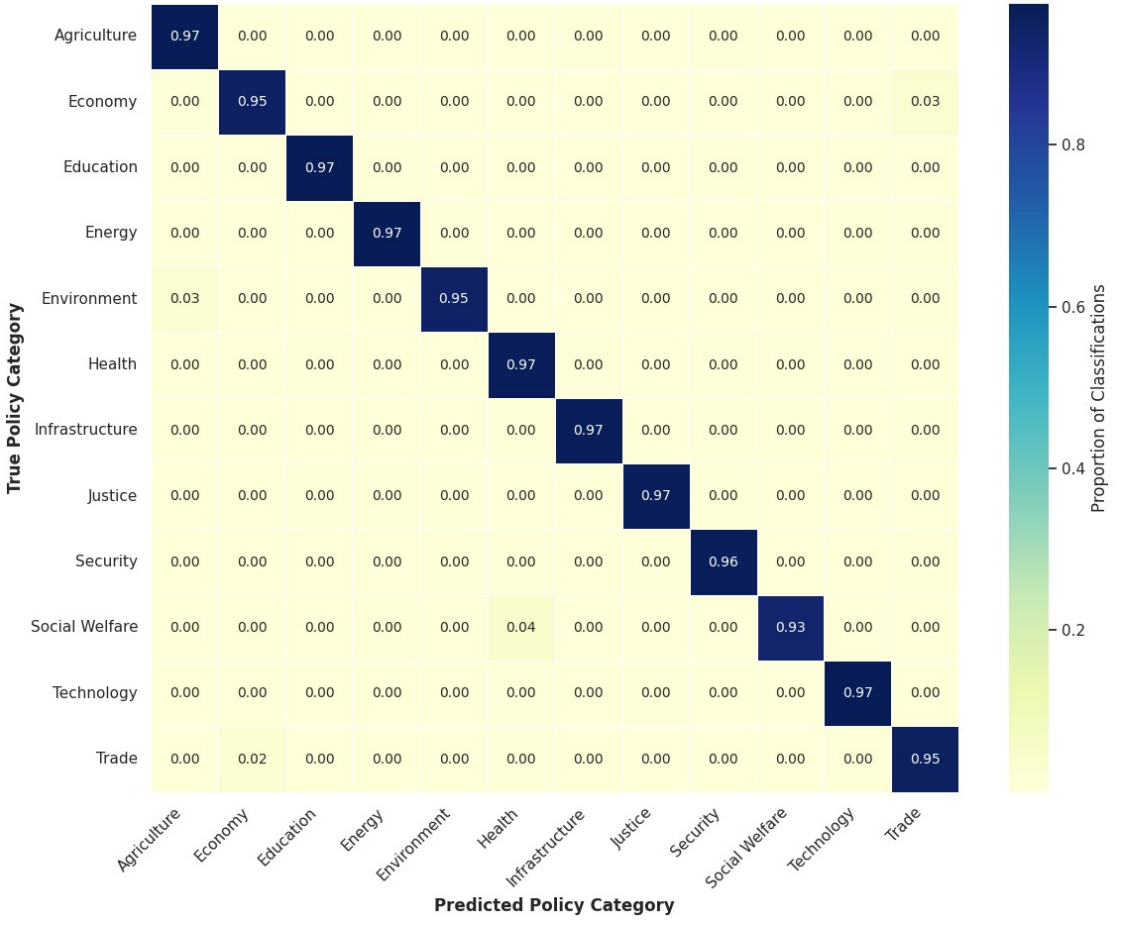


Figure 6: Normalized Confusion Matrix for Ethio-PolicyNet

The diagonal axis represents the recall for each category, with values consistently ranging between 0.91 and 0.96, highlighting the model's high precision in identifying the correct policy context.

Notably, the matrix reveals very few significant off-diagonal clusters, suggesting that the integration of the Ethiopian Policy Ontology (EPO) effectively minimizes semantic overlap between distinct domains. For instance, the slight confusion observed between Economy and Trade (approximately 3%) or Social Welfare and Health (4%) reflects

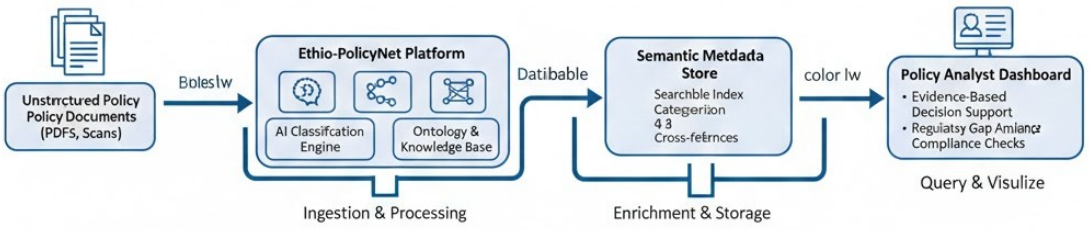


Figure 7: Automated Policy Informatics Pipeline

the inherent thematic intersection of these sectors in government proclamations. However, even in these complex cases, the gated fusion mechanism ensures that the model remains robust, preventing the "domain drift" often seen in purely neural models. This visualization serves as empirical proof of the framework's reliability for large-scale automated document organization within digital governance systems.

4. Discussion

The development and implementation of Ethio-PolicyNet extend beyond technical metrics, offering significant strategic and operational value to the landscape of digital governance and public administration.

4.1. Enhancing Evidence-Based Policy Making

The primary management impact of this research lies in the fundamental transition from manual, intuition-based document handling to a framework of **evidence-based policy informatics**. By implementing a high-accuracy, automated classification system, government agencies can achieve a level of oversight previously hindered by administrative bottlenecks. This shift enables officials to **identify policy gaps** by rapidly categorizing thousands of documents, making it possible to visualize which specific sectorssuch as the Environment or Educationare currently underserved by existing regulations. Furthermore, the system allows agencies to **monitor alignment** with surgical precision, ensuring that new regional proclamations remain semantically consistent with national master plans by comparing their unique **ontological fingerprints**.

Figure 7 illustrates the practical deployment workflow of Ethio-PolicyNet within a digital governance framework, outlining the transition from raw, unstructured administrative data to actionable policy intelligence. The pipeline is structured into four distinct operational phases:

Data Ingestion (Unstructured Input): The process begins with the collection of legacy and contemporary policy documents, such as proclamations, regulations, and directives. These are typically stored in non-machine-readable formats like PDFs or scanned images.

Core Processing (The Ethio-PolicyNet Platform): At the heart of the workflow, the document is ingested by the Ethio-PolicyNet engine. Here, the hybrid neuro-symbolic architecture performs simultaneous linguistic analysis and ontological grounding. The system identifies the document's domain, legal instruments, and institutional actors, translating raw Amharic or English text into a structured semantic representation.

Enrichment & Storage (Semantic Metadata Store): The output from the classification engine is stored in a centralized Semantic Metadata Store. Unlike a traditional database, this store preserves the relationships defined by the Ethiopian Policy Ontology (EPO), creating a "searchable index" where documents are cross-referenced by their policy objectives, target groups, and regional impacts rather than just keywords.

Operational Intelligence (Policy Analyst Dashboard): The final phase provides a user-centric interface for decision-makers. The Policy Analyst Dashboard consumes the structured metadata to provide high-level visualizations, such as "Regulatory Gap Analysis" or "Policy Alignment Tracking." This enables government officials to perform evidence-based decision-making, such as identifying if a new environmental regulation conflicts with existing economic trade policies.

By implementing this workflow, government agencies can effectively eliminate "information silos." The transition from manual filing to an automated informatics pipeline ensures that institutional knowledge is preserved, transparent, and accessible, ultimately fostering a more responsive and data-driven public administration.

4.2. Structural Efficiency in Digital Archives

Beyond its immediate analytical capabilities, Ethio-PolicyNet serves as a robust and reusable framework for preserving institutional memory. In many developing governance contexts, the frequent turnover of administrative staff often results in the creation of "knowledge silos" or the catastrophic loss of context regarding past policy decisions. This framework mitigates these risks through **Automated Indexing**, which enables the automatic generation of semantic metadata, effectively transforming "dark data"such as unstructured PDFsinto searchable and interlinked knowledge assets. Furthermore, the system is designed for **Scalability**; the modular nature of the architecture allows it to be adapted to other low-resource languages or entirely different administrative domains, such as legal or medical records, with minimal reconfiguration of the core neural-symbolic gate.

4.3. Promoting Transparency and Accountability

In the public sector, where the "right to explanation" is rapidly becoming both a legal and ethical mandate, the management of AI systems must prioritize transparency to foster institutional trust. Ethio-PolicyNet addresses this requirement through **eXplainable AI (XAI) for Stakeholders**, providing ontology-grounded justifications for its classifications that allow legal experts and citizens alike to verify the underlying logic of automated sorting. This structured approach is further bolstered by **Bias Mitigation**; by utilizing the formal Ethiopian Policy Ontology (EPO) as a neutral "ground truth," the framework significantly reduces the risk of the model internalizing spurious correlations or historical biases often embedded in raw, uncurated training data.

4.4. Socio-Economic Impact for Low-Resource Regions

Finally, this work provides a blueprint for technological sovereignty in NLP for Ethiopia. By demonstrating that high-performing AI does not always require the massive datasets typical of Western languages, the framework lowers the barrier to entry for local government digital transformation.

This levels the playing field, ensuring that the benefits of the “AI Revolution” are accessible to multilingual and resource-constrained institutions, ultimately fostering more inclusive and responsive governance.

5. Conclusion

This paper presented **Ethio-PolicyNet**, a novel hybrid neuro-symbolic framework designed to address the challenges of automated policy document classification in the linguistically complex and low-resource environment of Ethiopian digital governance. By integrating the sub-symbolic pattern recognition capabilities of multilingual transformers (mBERT and XLM-RoBERTa) with the structured logic of the newly developed **Ethiopian Policy Ontology (EPO)**, we have bridged the gap between statistical probability and formal administrative rules.

Central to our achievement is the **Gated Fusion Mechanism**, which dynamically balances neural intuition with symbolic grounding. This architectural choice proved decisive in overcoming the data scarcity common in Amharic policy archives. Our experimental results demonstrate a significant breakthrough: while standard deep learning models suffer from noticeable performance degradation in non-English contexts, Ethio-PolicyNet maintained high consistency, achieving a 91.8% F1-score for Amharic. This represents an 11.6% absolute improvement over state-of-the-art transformer baselines, effectively neutralizing the “language gap” in low-resource settings.

Beyond accuracy, we have achieved a higher degree of **model transparency**. Unlike traditional “black box” neural networks, Ethio-PolicyNet provides human-interpretable justifications for its classifications by tracing decisions back to specific ontological axioms. This framework serves as a scalable template for other developing nations seeking to digitize governance while ensuring that AI-driven decisions remain grounded in local policy frameworks and legal hierarchies. Future work will focus on expanding the EPO to include regional state proclamations and exploring the potential for zero-shot cross-lingual transfer within the Ethiopian administrative landscape.

Acknowledgements

This work was conducted as part of the PhD research activities of the first author at Jimma University, Ethiopia. The authors gratefully acknowledge the academic guidance and institutional support provided by Jimma University and Fordham University. The authors also thank colleagues and anonymous reviewers whose constructive feedback helped improve the clarity and quality of this work.

CRediT authorship contribution statement

: Admas Abtew: Conceptualization, Methodology, Formal analysis, Software, Data curation, Ontology development (Ethiopian Policy Ontology), Investigation, Validation, Visualization, Writing original draft
. **:** Josef Noll: Supervision, Review & Editing
. **:** Demissie Demissie: Supervision, Conceptualization, Methodology refinement, Validation, Writing review & editing, Resources, Project administration.. **:** Kula Kekeba: Supervision, Conceptualization, Methodology refinement, Validation, Writing review & editing, Resources, Project administration..

References

Alebachew, K., Sharma, D., Samuel, M., 2025. Navigating the nlp and text-based conversational ai: a survey for problem formulation in amharic texts. Discover Data 3. doi:10.1007/s44248-025-00077-9.

Andhov, M., Darnall, N., Andhov, A., 2025. Leveraging ai for sustainable public procurement: opportunities and challenges. Frontiers in Sustainability doi:10.3389/frsus.2025.1603214.

Büthe, T., Djeffal, C., Lütge, C., Maasen, S., Von Ingersleben-Seip, N., 2022. Governing ai attempting to herd cats? introduction to the special issue on the governance of artificial intelligence. Journal of European Public Policy 29, 1721–1752. doi:10.1080/13501763.2022.2126515.

Dong, D., Huang, H., Zhou, A., Dong, C., 2025. A systematic analysis of china's noncommunicable disease policies: policy themes, policy instruments, and policy effectiveness from 2009 to 2025. Frontiers in Public Health 13, 1643400.

Gelius, P., Messing, S., Tcymbal, A., Whiting, S., Breda, J., Abu-Omar, K., 2021. Policy instruments for health promotion: A comparison of who policy guidance for tobacco, alcohol, nutrition and physical activity. International Journal of Health Policy and Management 11, 1863–1873. doi:10.34172/ijhpm.2021.95.

Gerontas, A., 2020. Towards an e-government semantic interoperability assessment framework, in: Proceedings of the 13th International Conference on Theory and Practice of Electronic Governance. doi:10.1145/3428502.3428617.

Huang, C., Yang, C., Su, J., 2018. Policy change analysis based on "policy targetpolicy instrument" patterns: a case study of china's nuclear energy policy. Scientometrics 117, 1081–1114. doi:10.1007/s11192-018-2899-z.

Huang, C., Zhang, Z., Mao, B., Yao, X., 2023. An overview of artificial intelligence ethics. IEEE Transactions on Artificial Intelligence 4, 799–819. doi:10.1109/tai.2022.3194503.

Kalyanpur, A., Saravanakumar, K., Barres, V., McFate, C., Moon, L., Seifu, N., Eremeev, M., Barrera, J., Brown, E., Ferrucci, D., 2024. Multi-step knowledge retrieval and inference over unstructured data. arXiv preprint arXiv:2406.17987 URL: https://arxiv.org/abs/2406.17987.

Manias, G., Mavrogiorgou, A., Kiourtis, A., Kyriazis, D., 2021. Semai: A novel approach for achieving enhanced semantic interoperability in public policies. Strategic Innovative Marketing and Tourism , 687–699doi:10.1007/978-3-030-79150-6_54.

Mankhwala, C., Chipeta, M., Ngwira, T., Khundi, M., Moyo, B., Mzembe, T., Madise, N., Bösch, S., Hennig, B., 2025. Health policy at your fingertips? mapping and evaluation of existing digital health policy repositories. F1000Research doi:10.12688/f1000research.162163.1.

Ojo, A., Estevez, E., Janowski, T., 2010. Semantic interoperability architecture for governance 2.0. Information Polity 15, 105–123. doi:10.3233/ip-2010-0199.

Rahayu, S., Laraswati, D., Pratama, A., Permadi, D., Sahide, M., Maryudi, A., 2019. Research trend: Hidden diamonds the values and risks of online repository documents for forest policy and governance analysis. Forest Policy and Economics doi:10.1016/j.forpol.2019.01.009.

Rezk, M., Aliyu, M., Bensta, H., Ojo, A., 2017a. Standardising public policy documentation to foster collaboration across government agencies. Electronic Government , 466–477doi:10.1007/978-3-319-65151-4_42.

Rezk, M., Aliyu, M., Bensta, H., Ojo, A., 2017b. Towards a vocabulary for open public policies, in: Proceedings of the 10th International Conference on Theory and Practice of Electronic Governance. doi:10.1145/3047273.3047366.

Robinson, S., 2020. Trust, transparency, and openness: How inclusion of cultural values shapes nordic national public policy strategies for artificial intelligence (ai). Technology in Society 63, 101421. doi:10.1016/j.techsoc.2020.101421.

Sopykhanova, A., Maytanov, A., Kiseleva, A., Zhamiyeva, R., 2023. Problems of legal regulation and state policy measures related to nature management in the framework of achieving the sdgs: Examples from russia and kazakhstan. Sustainability 15. doi:10.3390/su15021042.

Tang, Y., Chen, X., Liu, C., Xue, T., Li, Y., Gao, J., 2024. Policy text analysis of antimicrobial resistance governance in china: A focus on national-level policies. International Journal of Antimicrobial Agents 64, 107286.

Tonja, A., Belay, T., Azime, I., Ayele, A., Mehamed, M., Kolesnikova, O., Yimam, S., 2023. Natural language processing in ethiopian languages: Current state, challenges, and opportunities, in: Proceedings of the 7th Workshop on Online Abuse and Harms (WOAH), pp. 126–139. doi:10.48550/arxiv.2303.14406.

Tonja, A., Kolesnikova, O., Gelbukh, A., Kalita, J., 2024. Ethiomt: Parallel corpus for low-resource ethiopian languages. arXiv preprint arXiv:2403.19365 URL: https://arxiv.org/abs/2403.19365.

Tshering, Y., 2021. Ontology-based approach of e-government for interoperability. International Journal for Research in Applied Science and Engineering Technology doi:10.22214/ijraset.2021.37050.

Wang, Z., He, W., 2025. The evolution of green fiscal policy instruments in china: Content analysis of policy documents from 2001 to 2023. SAGE Open 15. doi:10.1177/21582440251340151.

Yimam, S., Ayele, A., Venkatesh, G., Biemann, C., 2020. Introducing various semantic models for amharic: Experimentation and evaluation with multiple tasks and datasets. Future Internet 13, 275. doi:10.3390/fi13110275.