System Design Document – Adaptive Taxonomy Mapper

1. System Overview

The Adaptive Taxonomy Mapper is a rule-based inference system designed to map noisy user-provided tags and short story descriptions to a predefined internal genre taxonomy. The primary objective of the system is to infer fine-grained sub-genres while avoiding forced or hallucinated classifications.

The system prioritizes story context over user tags, applies strict taxonomy validation, and explicitly supports an UNMAPPED outcome when confidence is insufficient. This design ensures precision, explainability, and reliability.

2. High-Level Architecture

The system follows a deterministic, modular pipeline:

1. **Input Layer**
   - User tags (optional, noisy)
   - Story description (primary signal)

2. **Preprocessing**
   - Text normalization (lowercasing, punctuation removal)
   - Noise reduction for consistent matching

3. **Signal Extraction**
   - Rule-based keyword and phrase detection
   - Signals are manually curated for each sub-genre

4. **Candidate Selection**
   - Matching sub-genres are collected based on detected signals
   - Context Wins Rule is applied (story > tags)

5. **Taxonomy Validation**
   - Candidate sub-genre is validated against the approved taxonomy
   - Prevents generation of non-existent or hallucinated categories

6. **Final Decision**
   - Valid match → sub-genre assigned
   - No strong or valid match → UNMAPPED

7. **Explainability Layer**
   - A human-readable reasoning statement is generated for each decision

3. Key Design Decisions

3.1 Rule-Based Classification over End-to-End LLM Usage

The system uses explicit, hand-crafted rules instead of relying on a large language model for final classification. This ensures:

- Deterministic behavior
- Reproducible outputs
- Full transparency of decision logic

LLMs, if used in future extensions, would only assist with signal extraction, not decision making.

3.2 Context Wins Rule

When there is a conflict between user-provided tags and story content, the system prioritizes story context. This avoids misclassification caused by vague or misleading tags such as "Action" or "Scary."

3.3 Use of UNMAPPED as a Valid Outcome

The UNMAPPED label is a deliberate and essential design choice. It is returned when:

- The story lacks sufficiently specific genre signals
- Signals are ambiguous or cross-genre
- Mapping would require guessing

This improves system trustworthiness and prevents overfitting or taxonomy misuse.

4. Scalability Considerations

To scale the system to thousands of sub-genres:

- Replace linear keyword scans with an inverted index mapping keywords to sub-genres
- Use hierarchical filtering (Genre → Sub-genre) to reduce search space
- Store signals in external configuration files instead of hard-coded dictionaries
- Cache frequent classification results for repeated patterns

The modular design allows new genres or signals to be added without modifying core logic.

5. Cost Optimization Strategy (High-Volume Usage)

If processing millions of stories per month:

- Apply rule-based classification as the first pass
- Invoke LLMs only for edge cases where rule confidence is low

- Cache LLM-derived signals to avoid repeated calls

- Batch process stories to reduce overhead

This hybrid approach significantly reduces inference cost while maintaining accuracy.

6. Hallucination Prevention Mechanisms

The system prevents hallucination through:

- Strict validation against a predefined taxonomy

- No free-text generation of genre labels

- Explicit fallback to UNMAPPED when validation fails

- No model-generated category creation

These constraints ensure that outputs remain controlled and trustworthy.

7. Limitations and Future Improvements

Current Limitations

- Rule-based signals may miss subtle thematic cues

- Ambiguous stories may be conservatively classified as UNMAPPED

- Manual signal curation requires maintenance

Future Enhancements

- Introduce confidence scoring for partial matches

- Use LLMs for semantic signal extraction (not classification)

- Learn keyword weights from labeled data

- Support multi-label classification for hybrid genres

8. Conclusion

The Adaptive Taxonomy Mapper demonstrates a responsible, explainable, and scalable approach to genre inference. By prioritizing context, enforcing taxonomy constraints, and handling uncertainty explicitly, the system balances precision with flexibility and avoids the risks associated with uncontrolled AI outputs.