

Voynich Manuscript Research: Discoveries & Framework

This manuscript consolidates the documented discoveries and analytical frameworks developed for the Voynich Manuscript.

It is based on original contributions made in the Voymap-STA system and associated datasets, combined with semantic and positional annotations across specific folios. These approaches aim to provide a reproducible foundation for computational and human analysis.

1. Voymap-STA: Mapping into a Single Searchable Alphabet (STA)

- Ingests EVA/v101 transcriptions and normalizes them into a Single Transcription Alphabet.
- Removes inconsistencies between transcription systems, enabling unified search across the corpus.
- Provides per-folio statistics, token concordances, and global frequency analysis.
- Enables large-scale computational text analysis without manual conversion of datasets.

2. Ground Truth Positional and Semantic Mapping

- Created CSV datasets for folios f56v and f57v with positional coordinates for each glyph and illustration.
- Annotated hypothesized phonetic and semantic values for each mapped token.
- Linked glyph clusters to botanical, astronomical, and linguistic referents.
- Structured data format supports both human inspection and machine learning training.

3. Methodological Contributions

Voynich Manuscript Research: Discoveries & Framework

- Combined spatial annotation with transcription normalization for richer context in decoding attempts.
- Designed reproducible mapping templates, making hypotheses testable and shareable.
- Approach bridges paleographic study with computational linguistics, allowing statistical testing of semantic mappings.

4. Potential Applications

- Automated Voynich-to-natural-language alignment models using provided datasets.
- Large-scale pattern mining for glyph co-occurrence and positional dependencies.
- Comparative studies against other medieval cipher manuscripts.
- Public challenge frameworks for competing decoding hypotheses.

Conclusion

These discoveries represent an infrastructure and methodology advance for Voynich studies, laying the groundwork for future decoding and linguistic research. While the full semantic solution to the Voynich Manuscript remains elusive, these tools and datasets enable systematic exploration and verifiable testing of hypotheses.