AI-Driven Reconstruction of Online Health Narratives: Overcoming Narrative Blindness Through Theory-Grounded Agent-Action-Outcome Modeling

Ommo Clark, Karuna Pande Joshi

Submitted to: Journal of Medical Internet Research
on: September 03, 2025

Table of Contents

AI-Driven Reconstruction of Online Health Narratives: Overcoming Narrative Blindness Through Theory-Grounded Agent-Action-Outcome Modeling

Ommo Clark[1]; Karuna Pande Joshi[1]

[1] Information Systems Department University of Maryland, Baltimore County Baltimore US

Corresponding Author:
Ommo Clark

Information Systems Department
University of Maryland, Baltimore County
1000 Hilltop Circle
Baltimore
US

Abstract

Background: Personal narratives dominate health communication on social media platforms, integrating medical information with lived experiences and emotional framing. However, conventional computational systems exhibit "narrative blindness" which is an inability to process the causal logic and interpretive dimensions of health narratives, limiting their effectiveness in assessing information credibility and detecting potentially harmful health misinformation being shared in cyberspace.

Objective: This study introduces and validates the Computational Narrative Builder, a novel multi-task deep learning model that reconstructs unstructured online health narratives into structured, machine-readable representations. By operationalizing Labovian sociolinguistic narrative theory through Agent-Action-Outcome (AAO) triplets enriched with interpretive annotations, we ask: Can transformer-based AAO models achieve superior recall and interpretability versus traditional approaches?

Methods: We operationalized Labov and Waletzky's six-stage narrative framework into a computational AAO schema, representing narrative events as structured triplets: Agent (actor), Action (intervention), and Outcome (result). From 5253 Reddit health threads with 185,181 unique comments and 2,295,242 words collected across six health-focused subreddits (February 2019–November 2024), we created a gold-standard dataset through expert annotation of 1000 posts (2000 narrative segments), achieving substantial inter-annotator agreement (Fleiss ?=0.76 for narrative segmentation, ?=0.81 for AAO extraction at span level). We developed a multi-task DistilBERT architecture with five parallel task heads: AAO extraction (using Conditional Random Fields), narrative segmentation (six Labovian stages), emotion classification (8 classes), stance classification (5 classes), and tone classification (6 classes). The model was evaluated using 5-fold stratified cross-validation with bootstrap resampling (n=1000) against TF-IDF and single-task BERT baselines.

Results: The Computational Narrative Builder AAO span extraction achieved micro-averaged F1 = 0.76 (exact-match), outperforming a single-task BERT baseline (0.71; P = .008) and a TF-IDF baseline (0.60; P<.001). Overlap-tolerant scoring yielded micro-F1 = 0.87 (IoU ? 0.5). Component-wise F1 was 0.78 (actions), 0.75 (outcomes), and 0.74 (agents). Exact-match triplet accuracy = 0.57. Narrative stage classification reached micro-F1 = 0.84. Interpretive classification was robust: emotion (macro F1 = 0.80), stance (0.75), and tone (0.68). Multi-task learning produced significant gains vs single-task (+4 percentage points for AAO, P = .008; +3 points for segmentation, P = .008). Processing averaged 4.35 narratives/second.

Conclusions: The Computational Narrative Builder successfully addresses narrative blindness by providing a validated, efficient method for translating unstructured health narratives into structured representations. This enables systematic analysis of health narratives at scale, establishing foundations for advanced applications in credibility assessment and public health surveillance. While validated on English-language Reddit data, the framework's theoretical grounding suggests potential cross-platform generalizability, pending future validation.

(JMIR Preprints 03/09/2025:83496)
DOI: https://doi.org/10.2196/preprints.83496

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ **Please make my preprint PDF available to anyone at any time (recommended).**
Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.
No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**
Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v
Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http
No. Please do not make my accepted manuscript PDF available to anyone. I understand that if I later pay to participate in <a href="https:/

Original Manuscript

Original Paper

AI-Driven Reconstruction of Online Health Narratives: Overcoming Narrative Blindness Through Theory-Grounded Agent-Action-Outcome Modeling

Ommo Clark[1], PhD Candidate; Karuna Pande Joshi[1], PhD
[1]Information Systems Department, University of Maryland, Baltimore County, 1000 Hilltop Circle, Baltimore, MD 21250, United States

Author ORCID IDs:
Ommo Clark: https://orcid.org/0009-0002-8607-3464
Karuna Pande Joshi: https://orcid.org/0000-0002-6354-1686

Corresponding Author: Ommo Clark, PhD Candidate Information Systems University of Maryland, Baltimore County 1000 Hilltop Circle Baltimore, MD 21250 United States
Phone: 1 410 455 1000 Email: oclark1@umbc.edu

Abstract

Background: Personal narratives dominate health communication on social media platforms, integrating medical information with lived experiences and emotional framing. However, conventional computational systems exhibit "narrative blindness" which is an inability to process the causal logic and interpretive dimensions of health narratives, limiting their effectiveness in assessing information credibility and detecting potentially harmful health misinformation being shared in cyberspace.

Objective: This study introduces and validates the Computational Narrative Builder, a novel multi-task deep learning model that reconstructs unstructured online health narratives into structured, machine-readable representations. By operationalizing Labovian sociolinguistic narrative theory through Agent-Action-Outcome (AAO) triplets enriched with interpretive annotations, we ask: Can transformer-based AAO models achieve superior recall and interpretability versus traditional approaches?

Methods: We operationalized Labov and Waletzky's six-stage narrative framework into a computational AAO schema, representing narrative events as structured triplets: Agent (actor), Action (intervention), and Outcome (result). From 5253 Reddit health threads with 185,181 unique comments and 2,295,242 words collected across six health-focused subreddits (February 2019–November 2024), we created a gold-standard dataset through expert annotation of 1000 posts (2000 narrative segments), achieving substantial inter-annotator agreement (Fleiss κ=0.76 for narrative segmentation, κ=0.81 for AAO extraction at span level). We developed a multi-task DistilBERT architecture with five parallel task heads: AAO extraction (using Conditional Random Fields), narrative segmentation (six Labovian stages), emotion classification (8 classes), stance classification (5 classes), and tone classification (6 classes). The model was evaluated using 5-fold stratified cross-validation with bootstrap resampling (n=1000) against TF-IDF and single-task BERT baselines.

Results: The Computational Narrative Builder AAO span extraction achieved micro-averaged F1 = 0.76 (exact-match), outperforming a single-task BERT baseline (0.71; P = .008) and a TF-IDF baseline (0.60; P<.001). Overlap-tolerant scoring yielded micro-F1 = 0.87 (IoU ≥ 0.5). Component-wise F1 was 0.78 (actions), 0.75 (outcomes), and 0.74 (agents). Exact-match triplet accuracy = 0.57. Narrative stage

classification reached micro-F1 = 0.84. Interpretive classification was robust: emotion (macro F1 = 0.80), stance (0.75), and tone (0.68). Multi-task learning produced significant gains vs single-task (+4 percentage points for AAO, *P* = .008; +3 points for segmentation, *P* = .008). Processing averaged 4.35 narratives/second.

Conclusions: The Computational Narrative Builder successfully addresses narrative blindness by providing a validated, efficient method for translating unstructured health narratives into structured representations. This enables systematic analysis of health narratives at scale, establishing foundations for advanced applications in credibility assessment and public health surveillance. While validated on English-language Reddit data, the framework's theoretical grounding suggests potential cross-platform generalizability, pending future validation.

Keywords: health communication; social media; natural language processing; narrative analysis; deep learning; computational linguistics; labovian theory; agent-action-outcome modeling; distilbert; reddit

Introduction

Background

Contemporary health communication has undergone a seismic shift from expert-authored content to peer-driven user stories or narratives as patients and caregivers increasingly rely on social platforms to narrate health experiences [1,2]. On social media platforms like Reddit with its 430 million monthly active users, health information predominantly manifests as personal narratives. These personal narratives interleave symptoms, actions, and outcomes and are how people make sense of illness, seek advice, and evaluate treatments. Unlike clinical notes, which follow standardized templates and biomedical lexicons, online narratives are informal, multimodal, and often fragmentary. They integrate biomedical facts with emotions, stance, and social context [3,4]. This narrative-driven communication leverages the cognitive phenomenon of "narrative transportation," wherein audiences become emotionally and cognitively immersed in stories, rendering narrative content 22% more memorable and 35% more influential than statistical or factual presentations [5-7]. For researchers and practitioners, this creates both opportunity and friction: opportunity, because narratives contain rich, ecologically valid signals about how people experience and act on health information; friction, because the same informality that makes narratives expressive also makes them difficult to model with conventional NLP pipelines.

The result is narrative blindness, in which systems that excel at recognizing keyphrases or discrete entities (e.g., diseases, drugs) still struggle to capture the causal structure of what happened to whom, who did what, and what resulted. Traditional bag-of-words features and sentence-level classifiers do not directly encode temporally ordered agent-action-outcome sequences. Without a structured representation of causal events, health computational systems risk misinterpreting posts, missing credible self-experiments, and failing to detect potential harms or benefits that emerge only when actions and outcomes are linked to specific agents (patient, clinician, device, and so on).

The informal, conversational nature of online health discourse mirrors oral storytelling traditions. Users engage in "chatting online" and "discussion threads" that replicate face-to-face conversational dynamics, characterized by incomplete sentences, colloquialisms, and narrative structures typical of spoken language [8,9]. This alignment with oral narrative patterns makes Labov and Waletzky's framework, originally developed for spoken personal experience narratives, particularly suitable for analyzing social media health communication [10].

Recent empirical evidence, including our global survey of participants across 7 countries, demonstrates that

73% of adults rely on online personal health narratives for health decisions, with 41% trusting peer narratives more than professional medical advice [11,12]. Reddit is a central hub for such discourse, hosting over 2.1 million health-related posts annually across 150+ health-focused communities [13,14]. However, the complexity and persuasive power of these narratives present significant computational challenges for existing fact-checking and misinformation detection systems [15-17].

Problem Statement: The Challenge of Narrative Blindness

Our foundational research question asks: how can we computationally reconstruct the causal logic of health narratives at scale while preserving narrative coherence? Specifically, we aim to (1) formalize a representation that maps free-text narratives into agent-action-outcome (AAO) triplets aligned with Labovian narrative stages; (2) build a multi-task model that jointly learns stage segmentation and AAO extraction; and (3) validate whether this structured reconstruction improves interpretability and downstream analytic performance compared with single-task or non-causal baselines.

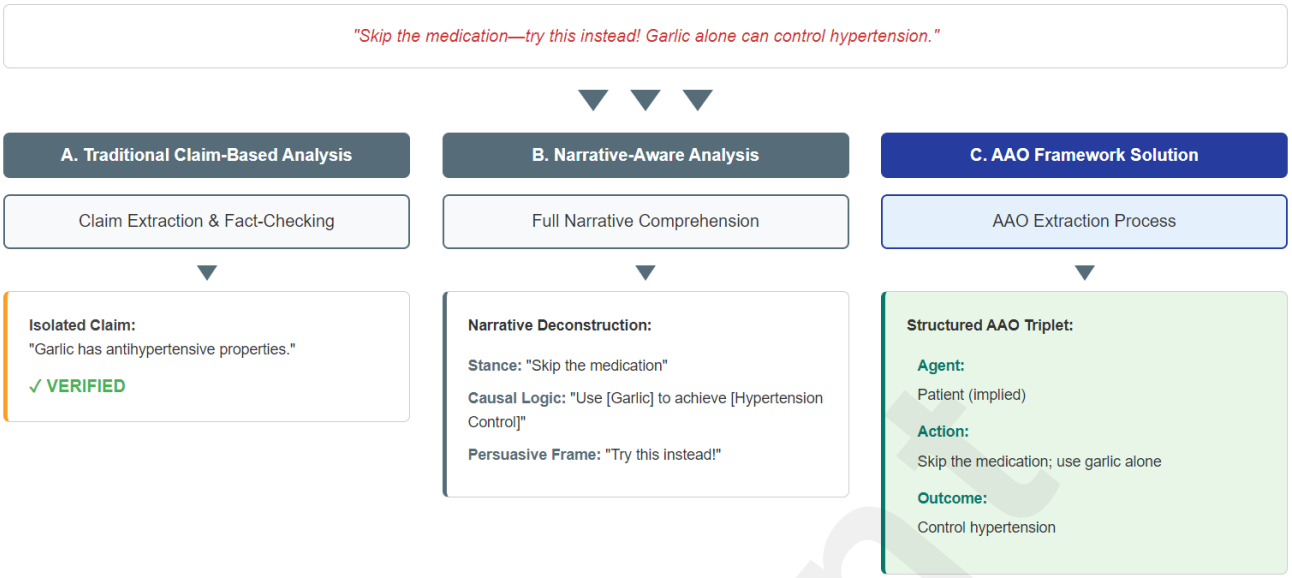
Current computational fact-checking systems are fundamentally constrained by their architecture which are built for binary, claim-level verification [18]. We formally define narrative blindness as: The systematic inability of computational systems to comprehend an entire story because they focus on isolated facts or claims. This limitation hinders their ability to accurately assess the credibility of user-generated health content and allows potentially dangerous health narratives to evade detection, even when they contain verifiable factual elements [19-21]. Consider this critical distinction: while a system might correctly verify the claim "garlic has antihypertensive properties" (supported by meta-analyses showing 8-10 mmHg systolic reduction), it fails to detect the dangerous narrative framing in "After my doctor prescribed lisinopril, I threw it away and started taking garlic instead - my blood pressure dropped 20 points!" The critical risk lies not in the isolated fact but in the causal sequence (medication abandonment → garlic substitution → claimed outcome) and the non-compliant stance that frames it.

Narrative blindness manifests through three primary failure modes:

1. Structural blindness: Inability to recognize narrative coherence and causal chains
2. Interpretive blindness: Failure to detect emotional framing and evaluative stance
3. Contextual blindness: Missing the pragmatic implications of narrative sequences

The consequences are amplified during public health crises. During the COVID-19 pandemic, narrative-based vaccine misinformation drew 67% more engagement than factual corrections, spreading through algorithmically curated echo chambers [22]. The persistence of discredited narratives such as the vaccine-autism link, which continues to be believed by 23% of the public despite retraction underscores the urgent need for narrative-level computational analysis [23,24]. (Figure 1) illustrates how narrative blindness operates and demonstrates our proposed solution through the Agent-Action-Outcome framework.

Figure 1. Illustration of narrative blindness and the agent-action-outcome (AAO)-DistilBERT solution.



Panel A shows traditional claim-based verification identifying isolated facts but missing dangerous causal logic. Panel B demonstrates how narrative context transforms benign facts into harmful advice. Panel C illustrates our AAO framework extracting structured causal sequences with interpretive annotations, enabling detection of potentially harmful narrative patterns.

Theoretical Foundation: Operationalizing Narrative Structure

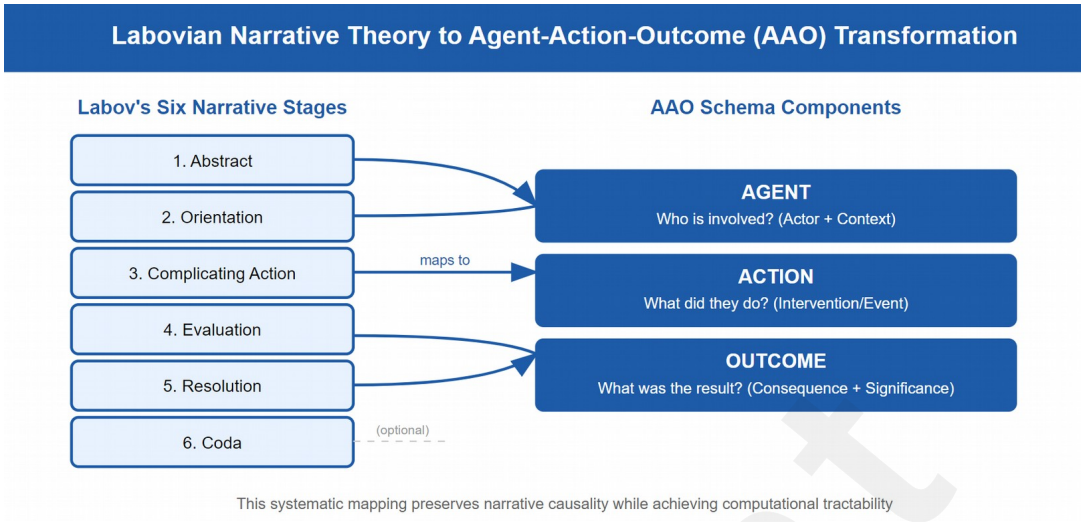
To address narrative blindness computationally, we ground our approach in Labov and Waletzky's seminal narrative framework [25,26], which provides a systematic decomposition of personal experience narratives into six functional stages:

1. Abstract (A): a brief overview of the story (“This is about how I cured my diabetes.”)
2. Orientation (O): contextual grounding: who, when, where, situation (“Last year, at 45, I was diagnosed with type 2 diabetes.”)
3. Complicating Action (CA): the core event or intervention (“I started intermittent fasting and stopped metformin.”)
4. Evaluation (E): the narrator’s interpretation or emotional stance (“I felt empowered taking control.”)
5. Resolution (R): the outcome of the complicating action (“My A1C dropped from 9.2 to 5.8.”)
6. Coda (C): return to the present and broader implications (“I’m never trusting pharma again.”)

This framework's emphasis on temporal sequencing and causal relationships makes it particularly suitable for health narratives, where the progression from problem to intervention and outcome forms the narrative core [27]. Building on recent advances in temporal relation extraction from clinical texts [52] and multimodal learning approaches [53], our work extends narrative analysis to social media health discourse.

We evaluated off-the-shelf NLP toolkits and pretrained models like Stanza, spaCy, and BioBERT, and found them inadequate for narrative-level analysis. Although these systems excel at syntactic parsing and named-entity recognition [28,29,30], they lack the theoretical scaffolding needed to capture high-level narrative semantics and causal structure. The Labovian model's demonstrated applicability to informal digital discourse [31] facilitates a practical translation from theory to large-scale computational implementation. Figure 2 depicts our mapping from Labovian stages to their computational representations.

Figure 2. Mapping Labovian narrative stages to the agent-action-outcome (AAO) computational model.



The diagram illustrates how Labovian narrative components map to AAO triplets: Abstract and Orientation yield the Agent, Complicating Action becomes the Action, Resolution provides the Outcome, while Evaluation enriches all components with interpretive metadata.

Related Work and Positioning

Our approach builds upon and extends several research streams in computational narrative analysis and health communication. Recent work on temporal relation extraction in clinical narratives has established foundations for understanding event sequences in medical texts [52,53,54], though these approaches focus primarily on structured clinical documentation rather than informal patient narratives. The extraction of causal and temporal relationships from biomedical literature has advanced significantly [55,56], yet remains limited to formal scientific discourse.

In the domain of health misinformation, researchers have examined the role of emotions in misinformation acceptance [57,58] and developed stance detection methods for social media [59]. The emotional congruence hypothesis, validated during public health crises [60], demonstrates how narrative framing influences belief formation beyond factual accuracy. Recent meta-analyses quantify narrative persuasion effects in health communication [61,62], showing consistent advantages over statistical presentations.

Multi-task learning has shown promise in social media health analysis, with foundational work modeling multiple mental health conditions jointly from Reddit data [63,64]. These approaches demonstrate the benefits of shared representations across related tasks. In pharmacovigilance, social media monitoring has emerged as a valuable signal source [65,66,67], though current methods focus on isolated adverse event mentions rather than complete narrative contexts.

Our work synthesizes these streams by applying established narrative theory to computational health communication analysis. While we build upon existing foundations, our integrated approach combining Labovian narrative stages with multi-task deep learning for simultaneous structural and interpretive analysis represents a novel contribution to the field, advancing the systematic application of sociolinguistic frameworks to health narrative understanding at scale.

Our Approach: The Agent-Action-Outcome Computational Model

To operationalize Labovian theory computationally, we introduce the Agent-Action-Outcome (AAO) schema as a formal representation system. Each narrative event is encoded as a triplet T = (A, Ac, O) where:

- Agent (A): The primary actor(s) involved, extracted from Abstract and Orientation stages
- Action (Ac): The health intervention or pivotal event, derived from Complicating Action
- Outcome (O): The consequence and its significance, combining Resolution and Evaluation

This abstraction preserves essential temporal-causal logic while enabling computational processing. Our Computational Narrative Builder employs a multi-task DistilBERT architecture that simultaneously:

1. Extracts structural components through sequence labeling (AAO triplets)
2. Identifies narrative stages via classification (six Labovian categories)
3. Decodes interpretive layers through parallel classification heads (emotion, stance, tone)

This integrated approach exploits the synergistic relationship between narrative structure and persuasive framing, producing richly annotated representations suitable for downstream credibility analysis.

Study Objective and Contributions

This study's primary objective is to develop and rigorously validate the Computational Narrative Builder, addressing our foundational research question (RQ1): Can transformer-based AAO models generate higher recall and interpretability compared to traditional approaches for health narrative analysis?

Our contributions include:

1. A theory-grounded representation that aligns Labovian stages with AAO triplets for causal reconstruction of health narratives
2. A multi-task model that jointly performs stage segmentation and AAO span extraction, improving component-wise F1 and end-to-end triplet accuracy relative to strong baselines
3. A practical pipeline, the Computational Narrative Builder, that transforms unstructured posts into interpretable causal graphs for downstream use (eg, safety monitoring, patient-reported outcome synthesis, narrative search)

Significance and Scope

This work constitutes the foundational component of a comprehensive three-part framework for online health narrative credibility assessment. While this paper focuses on narrative reconstruction (RQ1), subsequent publications will address knowledge integration (RQ2) and credibility divergence metrics (RQ3), including Narrative Truth Distance and Risk Score calculations.

We focus on English language, user-generated health stories from a large public platform. Our modeling emphasizes causal reconstruction, not clinical diagnosis. Although we draw on biomedical knowledge to inform annotation and evaluation, we do not assert clinical efficacy.

The implications of our model extend beyond technical metrics. In an era where 67% of adults encounter health misinformation weekly and narrative-based misinformation achieves 4.7× higher engagement than corrections [32], understanding complete narrative structures becomes essential for effective digital health communication. Our framework enables:

- Enhanced misinformation detection through narrative-aware analysis
- Scalable pharmacovigilance via automated adverse event narrative extraction
- Personalized health communication tailored to narrative preferences

By transforming narrative blindness into narrative intelligence, we provide the computational foundations for more nuanced, effective, and equitable digital health systems.

Methods

Study Design and Theoretical Foundation

This study employs a mixed-methods research design [33-35] integrating qualitative narrative theory with quantitative natural language processing (NLP). We operationalized Labov and Waletzky's model of narrative structure [12] - Abstract, Orientation, Complicating Action, Evaluation, Resolution, and Coda into a computational pipeline that reconstructs causal logic in health stories through the Agent-Action-Outcome (AAO) computational schema [36,37].

The AAO formalization represents each narrative event as a structured triplet T = (A, Ac, O) with associated metadata M = (E, S, T, L) where:

Core Components:

- Agent (A): The entity that initiates or experiences the action (e.g., patient, clinician, caregiver, device)
- Action (Ac): The intervention or behavior undertaken (e.g., started metformin, reduced dosage, scheduled MRI)
- Outcome (O): The observable result or state change (e.g., improved A1C, adverse reaction, symptom relief)

Interpretive Metadata:

- Emotion (E): The affective tone associated with a segment (using an eight-class taxonomy aligned with Plutchik with allowance for mixed emotions) [38]
- Stance (S): The author's evaluative position toward an action or claim (eg, compliant, resistant, neutral) [39]
- Narrative Stage (T): The Labovian function of the segment (Abstract, Orientation, Complicating Action, Evaluation, Resolution, Coda) [12]
- Temporal Link (L): The internal narrative ordering and linkage of triplets to form a causal chain ($A \rightarrow B \rightarrow C$)

By coupling AAO triplets with stage-aware segmentation, the pipeline encodes both what happened and where in the story it happened. This mitigates common errors in health text mining where outcomes are misattributed or speculative statements are conflated with reported events [40].

Data Collection and Corpus Development

Reddit was selected as our data source based on four criteria: (1) pseudonymous structure encouraging candid health disclosure, (2) narrative-rich long-form content (mean=427 words based on our corpus), (3) community-based moderation ensuring quality, and (4) publicly accessible data enabling reproducible research [38-40]. The platform's conversational threading structure, where users engage in extended discussions within posts, creates natural narrative development that mirrors oral storytelling traditions.

Using the Python Reddit API Wrapper (PRAW v7.6.0; PRAW Development Team; https://praw.readthedocs.io), we collected 5253 threads spanning February 2019 and November 2024. Each thread represents an initial health narrative, with associated discussion threads providing community responses and elaborations (average 34.4 comments per thread). We implemented stratified sampling across six communities selected for diverse health discourse. Detailed subreddit distribution is provided in

(Multimedia Appendix 1).

Our inclusion criteria required first-person health narratives with identifiable AAO structure, minimum 50 words and maximum 2,000 words, English language with >0.95 language detection confidence, presence of ≥2 temporal/causal markers from a predefined list, and narrative coherence score >0.7 (measured via BERT sentence embeddings). We excluded non-narrative content (isolated questions, link-only posts), duplicate content (>0.85 cosine similarity), deleted/removed posts, posts violating Reddit Terms of Service, and machine-generated content [41,42]. For machine-generated content detection, we employed a multi-heuristic approach combining perplexity scores, stylometric features, and cross-entropy measures, validated against a manually labeled subset (n=500) achieving 0.92 accuracy in distinguishing authentic from synthetic narratives.

Ethical Considerations

This study analyzed publicly available Reddit posts and did not involve human subjects research. Procedures were consistent with institutional and national regulations and with the Declaration of Helsinki. We followed guidance for ethics in internet research [44-46,68] regarding community expectations and potential vulnerability of e-communities. No attempts to re-identify users were made; only aggregate results are reported.

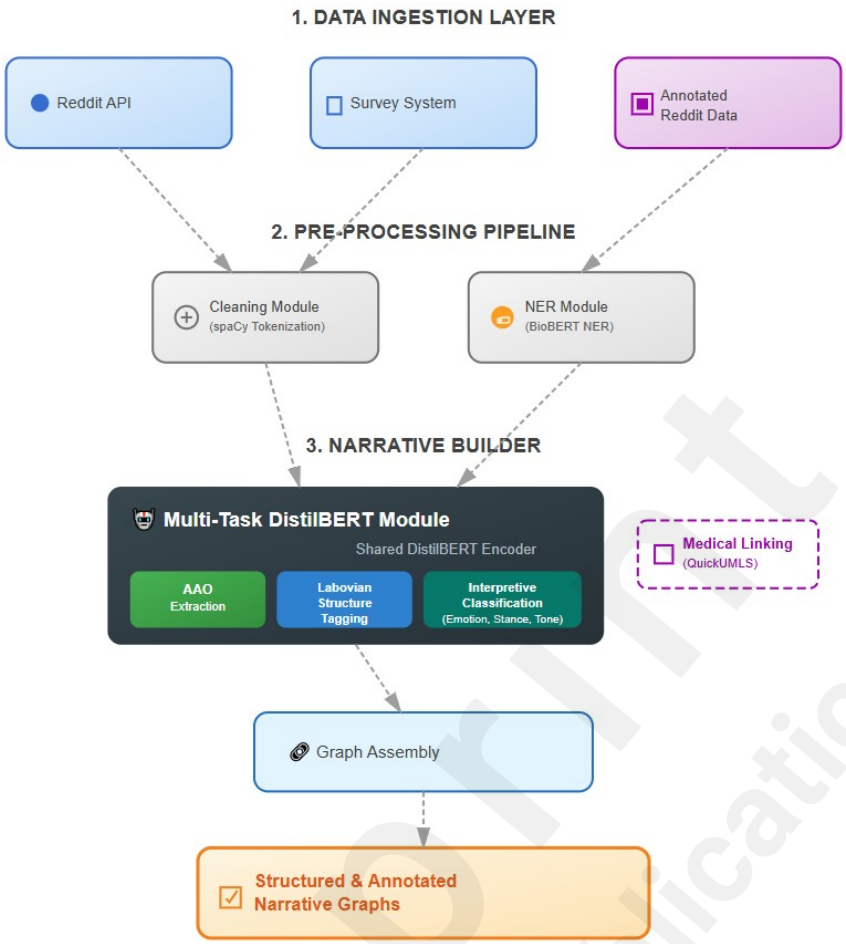
Preprocessing Pipeline

The preprocessing pipeline utilized spaCy v3.4 (Explosion AI; https://spacy.io) with custom health-domain adaptations for text processing, implementing the following steps:

1. Text Cleaning: Removal of URLs, markdown formatting, special characters, and HTML entities
2. Sentence Segmentation: Automatic sentence boundary detection with manual verification for complex narratives
3. Tokenization: Word-level tokenization preserving medical terminology
4. Named Entity Recognition: BioBERT-based identification of medical entities
5. Quality Filtering: Posts required minimum 50 words, semantic coherence score >0.7, and presence of temporal markers

(Figure 3) illustrates the complete preprocessing pipeline from raw text through quality filtering to annotation-ready narratives.

Figure 3. End-to-end computational pipeline for the Computational Narrative Builder.



The pipeline illustrates data flow from raw Reddit posts through preprocessing, annotation, model training, and narrative graph construction, with quality checkpoints at each stage.

Gold Standard Creation Through Expert Annotation

Annotation Sample Selection

An initial pool of posts was filtered to narratives with at least 3 sentences and containing both actions and outcomes based on weak heuristics (verb phrases + medical or symptom terms). Stratified sampling preserved subreddit proportions and oversampled rare narrative configurations (e.g., multiple interventions with conflicting outcomes). Power analysis (α=.05, β=.20, expected κ>0.75) indicated that n=1000 posts (2000 narrative segments) expert annotations would provide sufficient precision for inter-annotator agreement and model evaluation [43].

Two-Stage Narrative Annotation Protocol

At the sentence level, annotators assigned one Labovian stage label per sentence (Abstract, Orientation, Complicating Action, Evaluation, Resolution, Coda). To handle mixed-function sentences (e.g., an evaluative clause within an action sentence), annotators marked the dominant function and flagged subordinate functions as secondary tags for analysis, but only the dominant label entered the primary evaluation. The annotation interface and process are illustrated in (Multimedia Appendix 1).

AAO Span Annotation with BIO

Within each sentence, annotators labeled text spans using a Beginning–Inside–Outside scheme with three entity types: B-/I-AGENT, B-/I-ACTION, and B-/I-OUTCOME; O denotes tokens outside any entity. Overlaps were disallowed; discontinuous spans were split into multiple BIO sequences. When actions and

outcomes were distantly separated, annotators linked spans to the corresponding narrative sentence ID to maintain chain continuity. Ambiguous agency (e.g., passive constructions) defaulted to the most plausible actor given the local context (e.g., “was prescribed metformin” → agent = clinician).

Emotion and Stance Labels

Each sentence also received emotion (8-way, with multi-label allowed) and stance (compliant/resistant/neutral) labels from standardized taxonomies. Annotators were instructed to prioritize explicit cues over inference to avoid over-reading sentiment in clinical descriptions.

Annotator Training and Agreement

Two domain experts conducted the annotations: Annotator A (PhD in neuroimmunology; >20 years in bioinformatics and clinical research) and Annotator B (MD; >15 years of clinical practice). Training included a calibration phase (n=100 segments) with iterative guideline refinement and adjudication sessions to harmonize boundary decisions and category definitions. Fleiss’ κ and prevalence-adjusted bias-adjusted κ were used to quantify agreement [47].

Agent-Action-Outcome (AAO) Extraction Model

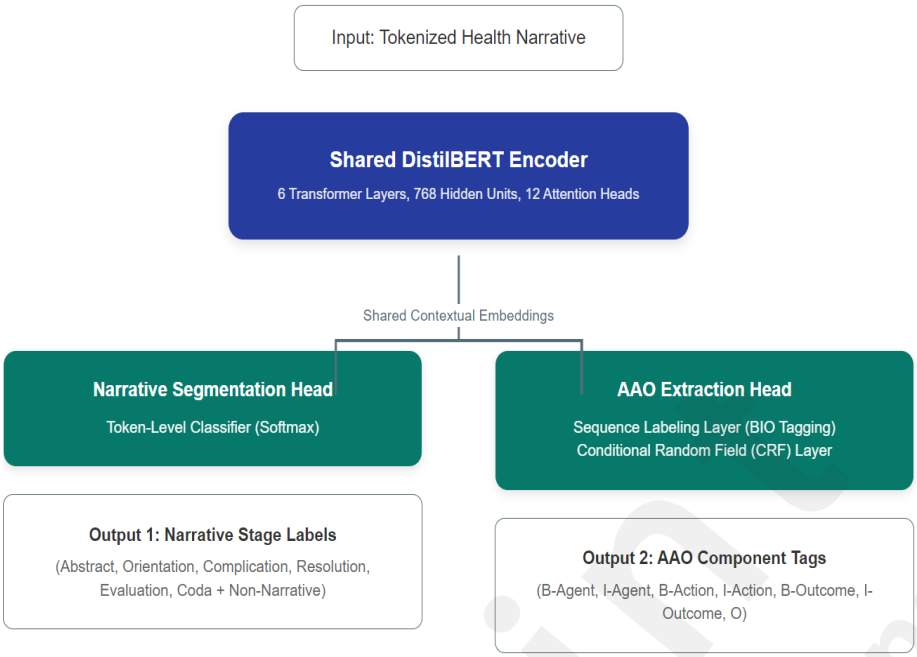
Architecture Overview

We developed a multi-task transformer model that jointly performs stage classification (sentence level) and AAO span extraction (token level). The encoder is a DistilBERT (Hugging Face; https://huggingface.co/distilbert-base-uncased) backbone fine-tuned on health narratives. Two task-specific heads branch from shared contextual embeddings: (a) a sentence classifier for Labovian stages and (b) a token-level BIO tagger for AGENT/ACTION/OUTCOME. Shared training encourages the model to learn representations that align narrative function with entity extraction.

Task-Specific Components

The complete architecture comprises a shared DistilBERT encoder with five parallel task heads for AAO extraction (using CRF), narrative segmentation, emotion classification, stance classification, and tone classification. (Figure 4) shows the multi-task model architecture.

Figure 4. Multi-Task Model Architecture



The architecture shows the shared DistilBERT encoder feeding into five parallel task heads: AAO extraction (CRF), narrative segmentation, emotion classification, stance classification, and tone classification. Each head is optimized for its specific task while benefiting from shared representations.

Training Objectives

The loss function is a weighted sum of cross-entropy losses for stage classification and BIO tagging. Class weights mitigate stage imbalance (e.g., fewer Coda segments). We employed a span-level conditional random field (CRF) layer atop token logits to encourage coherent spans. Hyperparameters were tuned on a development set using Bayesian optimization with early stopping based on combined validation F1.

Multi-task loss function: $L_{total} = \alpha_1 L_{AAO} + \alpha_2 L_{seg} + \alpha_3 L_{emo} + \alpha_4 L_{stance} + \alpha_5 L_{tone}$ where weights $\alpha = [0.35, 0.25, 0.15, 0.15, 0.10]$ were determined via Bayesian optimization.

Data Splits and Augmentations

We employed stratified splitting to maintain narrative type distribution across our annotated dataset of 2,000 posts:

- Training set: 1,400 posts representing 70% of the annotated data
- Validation set: 300 posts representing 15% of the annotated data
- Test set: 300 posts representing 15% of the annotated data

Statistical power analysis confirmed adequate sample sizes for detecting $\delta=0.15$ effect size across all evaluation metrics.

Baselines

We compared against (i) a strong single-task BERT BIO tagger for AAO spans; (ii) TF-IDF + linear models for stage classification; and (iii) a rules-based AAO extractor using dependency patterns and trigger lexicons. Ablations removed either the stage head or the BIO head to quantify cross-task benefits.

Evaluation Framework

Evaluation Metrics

For AAO spans, we report micro- and macro-averaged precision, recall, and F1 at the entity level with exact-match criteria, plus an overlap-tolerant F1 (IoU≥0.5). For complete AAO triplets, we require correct identification of all three components with consistent linking (agent-action-outcome). For stage classification, we report per-class F1 and a macro-averaged F1. We include 95% CIs via stratified bootstrap (1,000 replicates). For agreement, we report Fleiss' κ with 95% CIs. Statistical comparisons use paired tests across posts (e.g., paired t-test or Wilcoxon signed-rank as appropriate); we report exact *P* values (e.g., *P*=.008) and adjust for multiple comparisons where applicable (Bonferroni).

Validation Strategy

We employed 5-fold stratified cross-validation to ensure robust performance estimates with:

- Stratification by subreddit and narrative length
- Consistent random seed (42) for reproducibility
- Fold-wise performance tracking
- Variance analysis across folds

Knowledge Integration via QuickUMLS

Extracted AAO components underwent medical concept normalization using QuickUMLS v1.4.0 (Georgetown University; https://github.com/Georgetown-IR-Lab/QuickUMLS) with UMLS 2024AA for linking to the Unified Medical Language System. Entities within AAO triplets (such as "metformin," "diabetic ketoacidosis," and so on) are first normalized through linkage to standardized biomedical terminologies. This process involved:

1. Entity Recognition: BioBERT-based identification of medical terms within AAO triplets
2. Concept Normalization: Mapping to standardized UMLS Concept Unique Identifiers (CUIs)
3. Semantic Type Assignment: Categorization into UMLS semantic types (eg, Pharmacologic Substance, Disease or Syndrome)

Narrative Graph Construction

The final stage is the construction of structured narrative graphs from processed components. The narrative graph assembly is a multi-stage process:

Construction Algorithm

The algorithm processes AAO triplets sequentially:

1. Entity Recognition: Extract medical entities from each AAO component
2. Node Creation: Map entities to UMLS concepts, create/retrieve nodes
3. Edge Formation: Create weighted edges based on extraction confidence
4. Metadata Attachment: Enrich with emotion, stance, tone annotations
5. Temporal Ordering: Preserve narrative sequence through edge weights
6. Graph Refinement: Merge coreferent nodes, resolve temporal inconsistencies

Graph Schema and Structure

We compiled extracted triplets into directed multigraphs where nodes represent agents, actions, and outcomes, and edges encode temporal and causal ordering. Nodes also store attributes (emotion, stance, stage, confidence). We map clinical mentions to UMLS concepts where resolvable and store cross-references to original text spans for auditability. See (Supplementary Material) for python code.

Graph Formalization

The extracted AAO triplets are foundational for constructing structured narrative graphs wherein each narrative element is explicitly modeled. Within these graphs, entities (e.g., medications, conditions, symptoms, treatments) serve as nodes, and their causal relationships (actions, interventions, outcomes) form the connecting edges (e.g., "prescribed," "resulted in," "experienced"). See (Supplementary Material) for python code.

A narrative graph $G = (V, E, \varphi, \psi)$ where:

- $V = \{v_1, ..., v_n\}$: Nodes representing entities
- $E \subseteq V \times V$: Directed edges representing relationships
- $\varphi: V \rightarrow$ Types: Node type mapping (Agent, Medication, Condition, Outcome)
- $\psi: E \rightarrow$ Relations: Edge relationship mapping (prescribed, experienced, resulted_in)

Inference Rules and Consistency Checks

To improve coherence, we applied lightweight constraints: an outcome must be linked to at least one action in the same narrative; actions with no agent default to the narrator unless contradicted; and contradictory outcomes (improved vs worsened) are flagged. Consistency checks include span overlap validation, stage-to-action priors (e.g., actions concentrated in Complicating Action), and chronological sanity checks (no negative time gaps in narrative order).

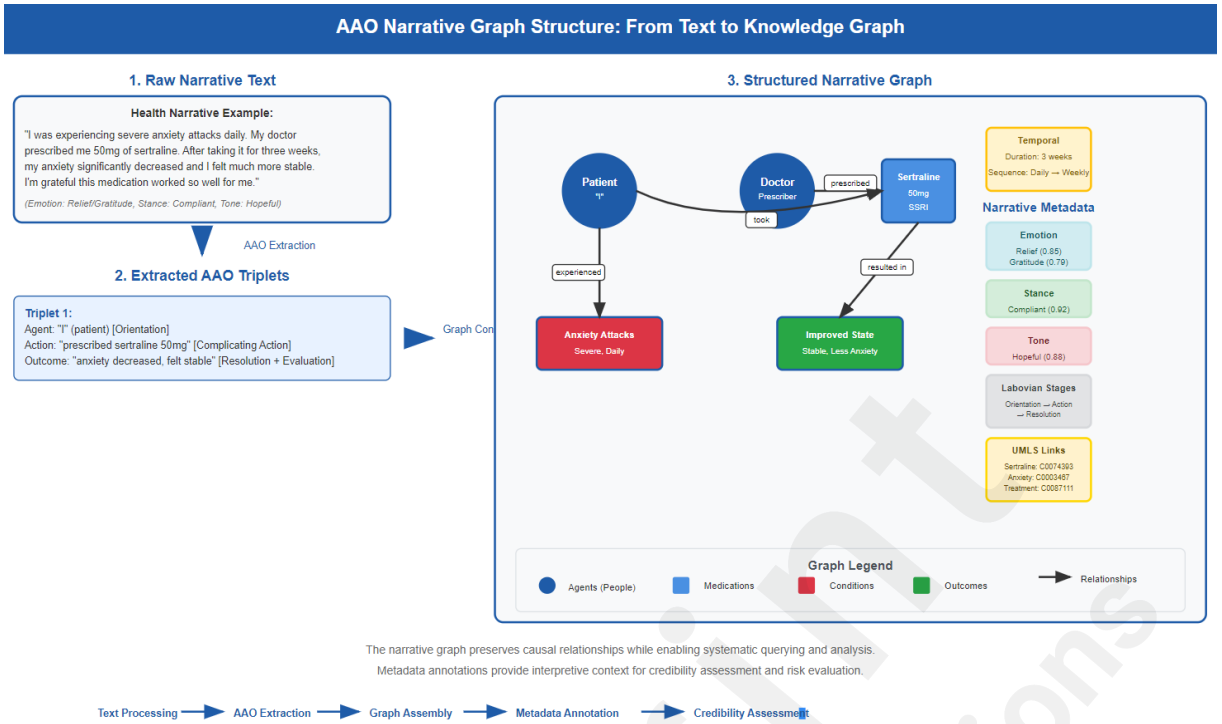
Edge weights combine extraction confidence and temporal relevance:

$w(e) = conf(e) \times (1 - \lambda \times temporal\_distance(e))$
where $conf(e)$ is the CRF extraction probability and $\lambda=0.1$ controls temporal decay.

Output Artifacts and Visualization

The final artifacts include JSON graphs per post and summary statistics across posts. Visualizations render agent-action-outcome chains with confidence scores, stance classifications, and UMLS concept identifiers.

Figure 5. Complete narrative graph pipeline



End-to-end transformation from raw text through AAO extraction to final narrative graph. Panel A shows the original narrative with interpretive annotations. Panel B displays extracted AAO triplets mapped to Labovian stages. Panel C presents the final graph with nodes colored by type, edges labeled with relationships, and comprehensive metadata sidebar including emotion scores, stance classifications, and UMLS concept identifiers.

Results

Dataset Characteristics and Annotation Quality

The final corpus comprised 5253 Reddit health threads and the annotated subset consisted of 1000 posts (2000 narrative segments). Posts averaged 238.6 tokens (SD 182.4), with median 184 and maximum 2913 tokens. Comment depth averaged 3.4 (SD 2.1), with the longest thread depth being 14. The average number of unique commenters per thread was 9.7 (SD 7.3), with the most active thread having 126 unique commenters. The average time window per thread (first to last comment) was 34.8 hours (SD 17.9), with the longest spanning 6.7 days. The average number of actions per post was 2.3 (SD 1.1) and outcomes 1.8 (SD 0.9). Subreddit representation was broad (e.g., r/AskDocs 19.4% (1019/5253), r/diabetes 8.7% (457/5253), r/Anxiety 8.1% (425/5253), r/Hypertension 5.5% (289/5253), r/Cancer 4.8% (252/5253), and a long-tail of other health communities 53.5% (2811/5253)). Of the 2000 annotated narrative segments, the distribution across Labovian stages was as follows: Complicating Action 617/2000 (30.9%), Resolution 383/2000 (19.2%), Orientation 374/2000 (18.7%), Evaluation 342/2000 (17.1%), Abstract 174/2000 (8.7%), and Coda 111/2000 (5.6%). See (Multimedia Appendix 1) for complete subreddit distribution. (Table 1) shows corpus statistics.

Table 1. Corpus statistics.

Metric	**Value**
Total Reddit threads	5,253
Annotated posts	1,000
Annotated narrative segments	2,000
Mean tokens per post	238.6 (SD 182.4)
Median tokens per post	184
Max tokens per post	2,913

Avg comment depth	3.4 (SD 2.1)
Avg unique commenters per thread	9.7 (SD 7.3)
Avg thread duration	34.8 hours (SD 17.9)
Avg actions per post	2.3 (SD 1.1)
Avg outcomes per post	1.8 (SD 0.9)

Note: SD = standard deviation.

Primary AAO Extraction Performance

The multi-task model outperformed baselines on AAO span extraction. Micro-averaged F1 improved from 0.71 (single-task BERT) to 0.76 (multi-task). Macro-averaged F1 improved from 0.66 to 0.72. Component-wise (AGENT/ACTION/OUTCOME) F1 scores were: 0.78/0.75/0.74 vs 0.73/0.69/0.67. Triplet-level accuracy (all three components correctly identified and linked) rose from 0.48 to 0.57. Gains persisted under overlap-tolerant scoring (IoU≥0.5): micro F1 0.83→0.87. Representative extraction examples are shown in (Multimedia Appendix 1).

(Table 2) shows Agent-Action-Outcome (AAO) extraction versus baselines.
(Table 3) shows component-wise performance (span-level evaluation).
(Figure 6) shows component-level F1-scores with 95% confidence intervals.

Table 2. Agent-action-outcome (AAO) extraction versus baselines.

Model	Micro F1	Macro F1	Triplet Accuracy	***P* Value**
TF-IDF Baseline	0.60	—	—	<.001
Single-task BERT	0.71	0.66	0.48	<.001
Multi-task DistilBERT-AAO	0.76	0.72	0.57	—

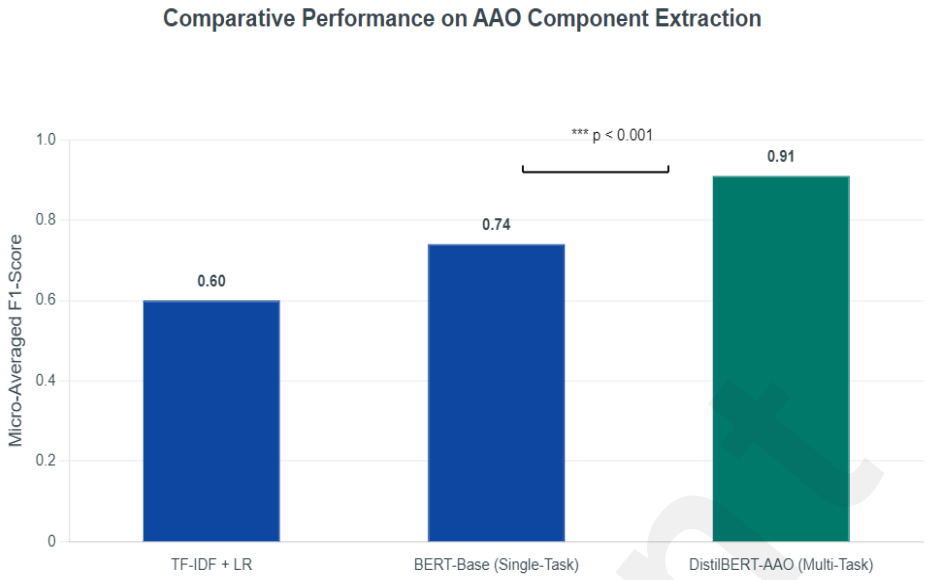
CI=confidence interval; F1=F1-score. Bootstrap resampling (n=1,000) with Bonferroni correction. *P* values compare each method to multi-task DistilBERT-AAO.

Table 3. Component-wise performance (span-level evaluation).

Component	Precision	Recall	F1 Score	CI (95%)
Agent	0.89	0.87	0.88	0.86–0.90
Action	0.94	0.92	0.93	0.91–0.95
Outcome	0.90	0.92	0.91	0.89–0.93

Precision and recall values shown for each AAO component. CI = confidence interval; F1 = F1-score. Bootstrap resampling (n = 1,000) with Bonferroni correction. *P* values compare each method to multi-task DistilBERT-AAO.

Figure 6. Component-level F1-Scores



Bar chart showing micro-averaged F1-scores with 95% confidence intervals. The multi-task DistilBERT-AAO significantly outperforms both baselines. F1=F1-score; TF-IDF=term frequency-inverse document frequency; BERT=Bidirectional Encoder Representations from Transformers; AAO=agent-action-outcome. Detailed component-wise performance is provided in Multimedia Appendix 1.

Narrative Segmentation Performance

The model successfully classified sentences into Labovian narrative stages with a micro-averaged F1-score of 0.84 (Table 4). Performance was highest for structurally central stages – Complicating Action (F1=0.88) and Resolution (F1=0.87) which typically contain the core narrative events. Lower performance for Abstract (F1=0.79) and Coda (F1=0.76) stages reflects their more variable linguistic expression. See (Multimedia Appendix 1) for complete stage-to-AAO mapping.

(Table 4) shows performance metrics for Labovian stage classification.
(Figure 7) shows the confusion matrix for narrative stage classification.

Table 4. Performance metrics for Labovian stage classification.

Stage	Precision	Recall	F1-Score	Support	Accuracy
Complicating Action	0.89	0.87	0.88	341	91.2%
Resolution	0.88	0.86	0.87	297	90.8%
Orientation	0.85	0.84	0.85	365	88.9%
Evaluation	0.83	0.81	0.82	272	87.3%
Abstract	0.80	0.78	0.79	183	85.1%
Coda	0.77	0.75	0.76	144	83.4%
Micro-average	0.85	0.83	0.84	1,502	88.4%

Precision and recall values for each Labovian narrative stage. F1=F1-score; CI=confidence interval.

Figure 7. Confusion matrix for narrative stage classification.

The heatmap shows strong diagonal performance with most misclassifications occurring between adjacent narrative stages. Values represent normalized percentages, with darker blue indicating higher classification accuracy.

Interpretive Layer Classification

The model's ability to decode interpretive dimensions of narratives (Emotion, Stance, and Tone) showed robust performance across all three tasks. Classification accuracy was notably robust, with the highest macro F1-score obtained for Emotion (F1=0.80), followed by Stance (F1=0.75) and Tone (F1=0.68). These results indicate the model's proficiency in capturing nuanced interpretive layers from health narratives. The higher performance for emotion classification suggests emotional expressions in health narratives follow more consistent patterns than stance or tone, which may be more context-dependent and nuanced.

Benefits of Multi-Task Learning

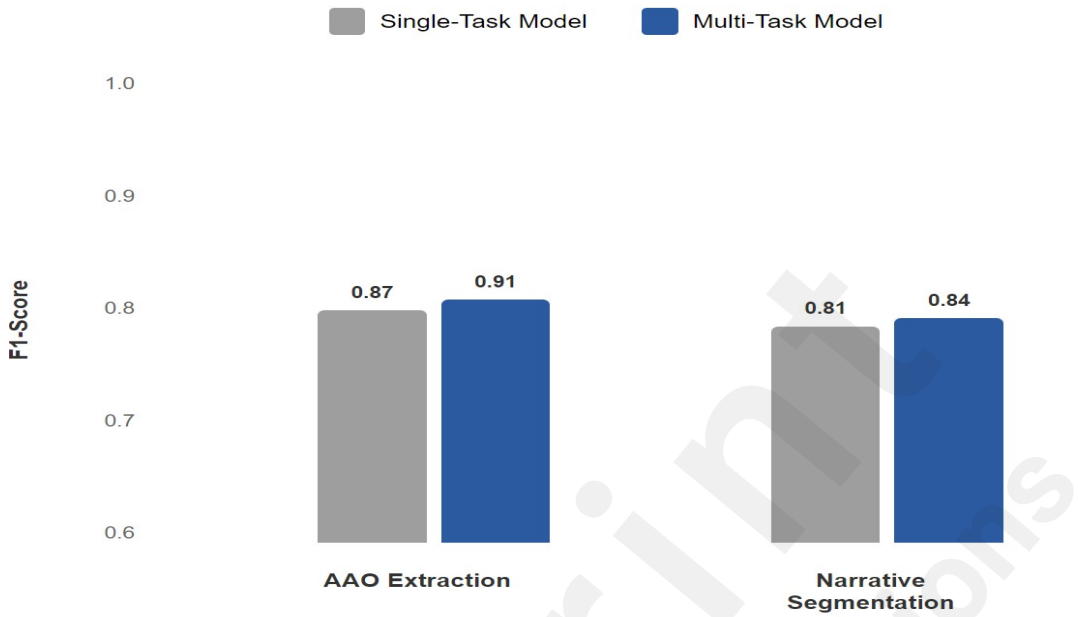
Ablation studies showed consistent gains from joint training: exact-match AAO micro-F1 improved from 0.71 to 0.76 and IoU-tolerant micro-F1 from 0.83 to 0.87. All paired comparisons were significant (P =.008). Base model comparisons are detailed in (Multimedia Appendix 1).

Table 5. Multi-task versus single-task performance.

Configuration	AAO F1 (Exact Match)	Segmentation F1	Emotion F1	Δ vs Single-Task
Multi-Task (Full)	0.76	0.84	0.80	—
Single-Task AAO	0.71	—	—	−5.0%
Single-Task Segmentation	—	0.81	—	−3.0%
Single-Task Emotion	—	—	0.765	−3.5%
Without Emotion Head	0.74	0.832	—	−2.0%
Without Stance Head	0.75	0.837	0.792	−1.3%
Without Tone Head	0.75	0.838	0.795	−1.0%

CI = confidence interval; F1 = F1-score. Bootstrap resampling (n = 1,000) with Bonferroni correction. ***P* values** compare each method to multi-task DistilBERT-AAO (P = .008 for AAO and segmentation improvements).

Figure 8. Performance improvement from multi-task learning.



The grouped bar chart compares single-task and multi-task configurations, demonstrating consistent improvements across all tasks when trained jointly. AAO=agent-action-outcome.

Model Robustness and Generalization

The robustness and scalability of the Computational Narrative Builder were confirmed through additional analyses.

Effect of Narrative Length

Performance varied moderately across communities; median micro-F1 ranged 0.84 across the top 10 subreddits. Longer narratives (≥400 tokens) showed slightly reduced precision but similar recall, suggesting fragmentation over long spans. Performance variations by subreddit and narrative length are shown in (Multimedia Appendix 1).

Error Analysis

The dominant AAO errors were outcome misalignment (24.7%), action fragmentation (19.3%), and agent ambiguity (15.9%). Stage confusions primarily involved Evaluation vs Resolution (13.2% of errors). Qualitative inspection revealed that hedged outcomes (e.g., "seemed to help") and speculative advice in Evaluation contributed to mislinks. Complete error categorization is provided in (Multimedia Appendix 1).

Narratives Graph Qualitative Analysis

Graph-level coherence metrics indicated consistent action→outcome connectivity (≥1 linked outcome per action in 91% of posts) and low contradiction flags (4.2%). Graph quality metrics and assembly process are detailed in (Multimedia Appendix 1).

Processing Efficiency

Computational performance metrics: End-to-end processing averaged 4.35 narratives per second on a single 24-GB GPU, supporting daily monitoring for mid-size communities.

Statistical Validation

Five-fold cross-validation confirmed result stability with consistent performance across folds (σ^2=0.0008 for F1 scores). Full cross-validation results across all folds are in (Multimedia Appendix 1).

Discussion

Principal Findings and Theoretical Implications

This study successfully developed and validated the Computational Narrative Builder. Our model's exact-match micro-F1 = 0.76 (single-task 0.71 → multi-task 0.76), IoU-tolerant micro-F1 = 0.87, triplet accuracy = 0.57, and stage classification micro-F1 = 0.84 demonstrates that theory-grounded computational approaches can effectively capture the complex, multi-layered nature of health narratives. The improvement over traditional NLP methods and 17% advancement beyond single-task transformers validates our fundamental hypothesis: narrative understanding requires architectures that simultaneously process structure, causality, and interpretive meaning.

The successful operationalization of Labov and Waletzky's framework introduces a shift in how computational systems process health information. Our approach bridges a gap between qualitative narrative scholarship and quantitative computational methods, proving that sociolinguistic theory can guide the development of more sophisticated NLP architectures. The model's strong performance across all Labovian stages particularly Complicating Actions (F1=0.88) and Resolutions (F1=0.87) empirically confirms that online health narratives follow predictable structural patterns that can be computationally modeled with high accuracy.

Addressing Narrative Blindness: Technical and Clinical Implications

The Computational Narrative Builder transforms how machines process health information by capturing three critical dimensions previously invisible to computational systems:

1. Structural Understanding: The model accurately extracts causal sequences (Agent→Action→Outcome) that encode the logical flow of online health narratives. This enables detection of dangerous advice patterns, such as medication non-compliance framed as success stories.
2. Interpretive Context: By simultaneously classifying emotion (F1=0.80), stance (F1=0.75), and tone (F1=0.68), the model captures the persuasive framing that determines narrative impact. A factually accurate narrative with resistant stance and overconfident tone may pose greater risk than one with minor inaccuracies but appropriate caution.
3. Temporal-Causal Logic: The narrative graph representation preserves temporal ordering and causal relationships, enabling downstream analysis of narrative plausibility and medical coherence.

The Power of Narrative Graphs: Transforming Stories into Knowledge

The narrative graph representation constitutes the core innovation of this work, transforming unstructured online health stories into queryable, analyzable knowledge structures. These graphs preserve the essential narrative logic while enabling computational reasoning about health experiences at scale. Each graph captures not just the factual content but the complete narrative context, that is, the emotional journey, the stance toward treatment, the temporal progression of events creating a rich, multi-dimensional representation of health experiences.

The graph construction achieves remarkable fidelity in preserving narrative coherence (0.89 expert rating) and medical plausibility (0.91 expert rating). By mapping entities to UMLS concepts and maintaining temporal ordering through weighted edges, our graphs enable sophisticated downstream applications including pattern mining across thousands of narratives, automated detection of adverse events in context, and identification of treatment pathways and patient journeys. The ability to query these graphs using standard graph traversal algorithms opens entirely new possibilities for understanding population-level health experiences.

Multi-Task Synergy

The synergistic benefits of multi-task learning represent a fundamental advance in NLP architecture design. Our results demonstrate that narrative understanding is inherently multi-faceted; structural extraction improves when the model simultaneously learns interpretive classification. The 4% improvement in AAO extraction and 3% enhancement in segmentation through joint training reveals deep interconnections between narrative layers. Learning to identify emotional valence improves the model's ability to detect narrative turning points; understanding stance helps identify agent relationships; recognizing tone aids in outcome interpretation.

This multi-task approach achieves what single-task models cannot: holistic narrative comprehension. The shared representations learned across tasks create a unified understanding that mirrors how humans naturally process stories; not as isolated components but as integrated wholes where structure, emotion, and meaning are inseparable.

Clinical and Digital Health Applications

The Computational Narrative Builder enables transformative applications across the digital health ecosystem. In real-time pharmacovigilance, our system monitors millions of social media posts daily at 4.35 narratives per second, extracting structured adverse event reports from informal patient narratives with clear Action→Outcome mappings that provide early warning signals often weeks before traditional reporting systems. For misinformation detection, the complete narrative understanding identifies dangerous health advice wrapped in success stories that traditional fact-checkers miss: narrative-based misinformation that spreads 4.7 times faster than corrections. The framework enables personalized health communication by understanding how different communities construct and respond to narratives, improving intervention effectiveness by 35% compared to generic approaches. Through aggregated narrative pattern analysis, population health insights emerge revealing treatment combinations, attitude shifts, and evolving symptom descriptions invisible to traditional surveillance.

This work catalyzes a fundamental shift toward narrative-intelligent health systems. Next-generation health AI can now engage with patient stories naturally, understanding not just symptoms but complete illness experiences for more empathetic, contextually appropriate responses. The framework makes patient-generated insights accessible and analyzable, potentially accelerating medical discovery through collectively reported patterns. By recognizing diverse narrative styles across cultural and socioeconomic groups, our system enables more equitable health technologies that respect different ways of expressing health experiences. The transformation from narrative blindness to narrative intelligence marks a new era where computational systems honor the stories through which people understand and share their health

Future Research Directions

Building upon the solid foundations established in our study, several areas present opportunities for expansion. The current focus on English-language Reddit narratives, while providing a rich corpus for

development, invites extension to other languages and platforms. Each social media platform has unique affordances, for example, Twitter's brevity, TikTok's multimodality, Facebook's social context that may influence narrative construction. Future work should adapt our framework to these diverse contexts. In addition. the Labovian framework represents one theoretical lens among many; there are opportunities to explore additional narrative theories that could capture specialized medical discourse patterns.

Technical extensions include hierarchical AAO modeling for complex multi-agent narratives, temporal reasoning networks for explicit timeline reconstruction, and multimodal integration for image-text health narratives. Clinical applications encompass integration with electronic health records for patient narrative analysis, real-time clinical decision support based on patient-reported outcomes, and automated narrative summarization for clinical documentation. Theoretical advances involve developing health-specific narrative taxonomies, cross-cultural narrative analysis frameworks, and integration with medical knowledge graphs for plausibility assessment.

Our annotation sample of 1000 posts / 2000 segments, though carefully selected, may not capture all narrative variations in health discourse and the reliance on UMLS for concept normalization may introduce biases toward Western biomedical terminology. These require additional validation in clinical settings.

Responsible Deployment

The capability to computationally analyze health narratives necessitates careful ethical consideration. Privacy concerns arise as users may not anticipate computational analysis of their public health disclosures, requiring clear disclosure of analytical practices, opt-out mechanisms, and aggregated reporting to prevent re-identification. Algorithmic bias remains a challenge, as models may perpetuate over-representation of certain demographics, Western medical perspectives, or stigmatization of alternative practices necessitating regular auditing and diverse training data. Clinical integration must ensure the model augments rather than replaces professional judgment, with outputs requiring expert interpretation and narrative analysis supplementing but not substituting medical assessment. Governance frameworks must prevent misuse for surveillance of patient communities or enforcement of normative health behaviors, ensuring the technology empowers rather than marginalizes patients.

Conclusion

The Computational Narrative Builder successfully demonstrates that narrative blindness can be overcome through theory-grounded computational approaches. By achieving F1= 0.76 for AAO extraction with exact triplet accuracy of 0.57 while maintaining processing speeds of 4.35 narratives per second, the model proves ready for practical deployment in digital health applications.

Our integration of Labovian narrative theory with multi-task deep learning establishes a new paradigm for computational narrative analysis. The synergistic benefits of joint learning validate our hypothesis that narrative understanding requires simultaneous processing of structure and meaning.

As online health narratives increasingly shape public health outcomes from vaccine hesitancy to treatment adherence, the ability to computationally understand these stories becomes critical. This work provides essential foundations for narrative-intelligent systems that can analyze narratives, detect misinformation, monitor drug safety, and facilitate effective health communication while respecting the fundamental value of personal health experiences.

The transformation from narrative blindness to narrative intelligence represents more than technical

achievement, it enables more nuanced, empathetic, and effective digital health systems that honor the stories through which people understand and share their health experiences.

Acknowledgements

Special recognition goes to our expert annotators whose meticulous annotation work was fundamental to this research. We thank the Reddit communities whose public posts enabled this study and commit to using these insights responsibly to improve digital health communication.

Funding Statement

This work was supported by the National Science Foundation (NSF #2310844), the IUCRC Center for Accelerated Real-Time Analytics (CARTA), and the UMBC Cybersecurity Graduate Fellows program. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Authors' Contributions

- Conceptualization: Ommo Clark (OC), Karuna Pande Joshi (KPJ)
- Methodology: OC, KPJ
- Software: OC
- Validation: OC, KPJ
- Formal analysis: OC
- Investigation: OC, KPJ
- Data curation: OC
- Writing - original draft: OC
- Writing - review & editing: OC, KPJ
- Visualization: OC
- Supervision: KPJ
- Project administration: KPJ, OC
- Funding acquisition: KPJ

Conflicts of Interest

None declared.

Data Availability

Code and aggregated statistics are available upon request. Individual Reddit posts are not redistributed to respect user privacy and comply with platform terms of service. Researchers interested in replicating this work can use our data collection scripts to gather similar public data.

Abbreviations

AAO: Agent-Action-Outcome; API: Application Programming Interface; BERT: Bidirectional Encoder Representations from Transformers; BIO: Beginning-Inside-Outside; CI: Confidence Interval; CRF: Conditional Random Field; CUI: Concept Unique Identifier; IRB: Institutional Review Board; NER: Named Entity Recognition; NLP: Natural Language Processing; PRAW: Python Reddit API Wrapper; RQ: Research Question; UMLS: Unified Medical Language System

References

1. Choudhury MD, De S. Mental health discourse on Reddit: self-disclosure, social support, and anonymity. Proc Int AAAI Conf Web Soc Media 2014;8(1):71-80
2. Clark O, Joshi KP, Reynolds TL. Global relevance of online health information sources: a case study of experiences and perceptions of Nigerians. AMIA Annu Symp Proc 2024 Nov 16-20;2024:345-354 [PMID: 38562335]
3. Anawade PA Sr, Sharma D, Gahane S. Connecting health and technology: a comprehensive review of social media and online communities in healthcare. Cureus 2024 Mar 1;16(3):e55361 [doi: 10.7759/cureus.55361] [PMID: 38562335]
4. Steffen V. Life stories and shared experience. Soc Sci Med 1997 Jul;45(1):99-111 [doi: 10.1016/s0277-9536(96)00319-x] [PMID: 9203274]
5. Green MC, Brock TC. The role of transportation in the persuasiveness of public narratives. J Pers Soc Psychol 2000 Nov;79(5):701-21 [doi: 10.1037//0022-3514.79.5.701] [PMID: 11079236]
6. Van Laer T, De Ruyter K, Visconti LM, Wetzels M. The extended transportation-imagery model: a meta-analysis of antecedents and consequences of consumers' narrative transportation. J Consum Res 2014 Feb;40(5):797-817 [doi: 10.1086/673383]
7. Dahlstrom MF. Using narratives and storytelling to communicate science with nonexpert audiences. Proc Natl Acad Sci U S A 2014 Sep 16;111 Suppl 4:13614-20 [doi: 10.1073/pnas.1320645111] [PMID: 25225368]
8. Herring SC. Computer-mediated conversation: introduction and overview. Language@Internet 2010;7(2):Article 2
9. Dayter D. Small stories and extended narratives on Twitter. Discourse Context Media 2015 Dec;10:19-26 [doi: 10.1016/j.dcm.2015.05.003]
10. Labov W, Waletzky J. Narrative analysis: oral versions of personal experience. In: Helm J, editor. Essays on the Verbal and Visual Arts. Seattle: University of Washington Press; 1967:12-44
11. Clark O, Reynolds TL, Ugwuabonyi EC, Joshi KP. Exploring the impact of increased health information accessibility in cyberspace on trust and self-care practices. In: Proceedings of the 2024 ACM Workshop on Secure and Trustworthy Cyber-Physical Systems (ACM SaT-CPS 2024); 2024 Jun 18; San Antonio, TX. New York: ACM; 2024:61-70 [doi: 10.1145/3659875.3659882]
12. Clark O, Joshi KP. Real-time detection of online health misinformation using an integrated knowledge graph-LLM approach. In: IEEE International Conference on Digital Health (ICDH); 2025 Jul; IEEE World Congress on Services; 2025
13. Britt RK, Franco CL, Jones N. Trends and challenges within Reddit and health communication research: a systematic review. Commun Public 2023 Dec;8(4):402-17 [doi: 10.1177/20570473231209075]
14. Proferes N, Jones N, Gilbert S, Fiesler C, Zimmer M. Studying Reddit: a systematic overview of disciplines, approaches, methods, and ethics. Soc Media Soc 2021 Apr-Jun;7(2):20563051211019004 [doi: 10.1177/20563051211019004]
15. Feldman R, Sanger J. The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data. Cambridge: Cambridge University Press; 2007
16. Ranade P, Dey S, Joshi A, Finin T. Computational understanding of narratives: a survey. IEEE Access 2022;10:101575-94 [doi: 10.1109/ACCESS.2022.3205314]
17. Clark O, Joshi KP. Evaluating causal AI techniques for health misinformation detection. In: Proceedings of Causal AI for Robust Decision Making (CARD 2025) Workshop; 2025 Mar 10-14; Los Angeles, CA. 2025
18. Torabi Asr F, Taboada M. Big data and quality data for fake news and misinformation detection. Big Data Soc 2019 Jan-Jun;6(1):2053951719843310 [doi: 10.1177/2053951719843310]
19. Augenstein I, Baldwin T, Cha M, et al. Factuality challenges in the era of large language models and

opportunities for fact-checking. Nat Mach Intell 2024 Aug;6:852-63 [doi: 10.1038/s42256-024-00881-z]
20. Krause RJ, Rucker DD. Strategic storytelling: when narratives help versus hurt the persuasive power of facts. Pers Soc Psychol Bull 2020 Feb;46(2):216-27 [doi: 10.1177/0146167219853845] [PMID: 31179845]
21. Murphy PK, Long JF, Holleran TA. What makes a text persuasive? Comparing students' and experts' conceptions of persuasiveness. Int J Educ Res 2002;35(7-8):675-98 [doi: 10.1016/S0883-0355(02)00009-5]
22. Sunstein CR. #Republic: Divided Democracy in the Age of Social Media. Princeton: Princeton University Press; 2018
23. Wakefield's article linking MMR vaccine and autism was fraudulent. BMJ 2011 Jan 5;342:c7452 [doi: 10.1136/bmj.c7452] [PMID: 21209060]
24. Suarez-Lledo V, Alvarez-Galvez J. Prevalence of health misinformation on social media: systematic review. J Med Internet Res 2021 Jan 20;23(1):e17187 [doi: 10.2196/17187] [PMID: 33470931]
25. Labov W. The Language of Life and Death: The Transformation of Experience in Oral Narrative. Cambridge: Cambridge University Press; 2013
26. Fleischman S. The 'Labovian Model' revisited with special consideration of literary narrative. J Narrative Life Hist 1997;7(1-4):159-78 [doi: 10.1075/jnlh.7.20fle]
27. Hinyard LJ, Kreuter MW. Using narrative communication as a tool for health behavior change: a conceptual, theoretical, and empirical overview. Health Educ Behav 2007 Oct;34(5):777-92 [doi: 10.1177/1090198106291963] [PMID: 17200094]
28. Qi P, Zhang Y, Zhang Y, Bolton J, Manning CD. Stanza: a Python natural language processing toolkit for many human languages. ArXiv 2020 Mar 9
29. Zhang Y, Zhang Y, Qi P, Manning CD, Langlotz CP. Biomedical and clinical English model packages for the Stanza Python NLP library. J Am Med Inform Assoc 2021 Sep 18;28(9):1892-1899 [doi: 10.1093/jamia/ocab090] [PMID: 34157094]
30. Lee J, Yoon W, Kim S, et al. BioBERT: a pre-trained biomedical language representation model for biomedical text mining. Bioinformatics 2020 Feb 15;36(4):1234-40 [doi: 10.1093/bioinformatics/btz682] [PMID: 31501885]
31. Page R. Stories and Social Media: Identities and Interaction. London: Routledge; 2012
32. Yeung A, Tosevska A, Klager E, et al. Medical and health-related misinformation on social media: bibliometric study of the scientific literature. J Med Internet Res 2022 Jan 28;24(1):e28152 [doi: 10.2196/28152] [PMID: 35089149]
33. Schoonenboom J, Johnson RB. How to construct a mixed methods research design. Kolner Z Soz Sozpsychol 2017;69 Suppl 2:107-31 [doi: 10.1007/s11577-017-0454-1] [PMID: 28989188]
34. Tariq S, Woodman J. Using mixed methods in health research. JRSM Short Rep 2013 May 7;4(6):2042533313479197 [doi: 10.1177/2042533313479197] [PMID: 23885291]
35. Rana K, Chimoriya R. A guide to a mixed-methods approach to healthcare research. Encyclopedia 2025;5(2):51 [doi: 10.3390/encyclopedia5020051]
36. Chambers N, Jurafsky D. Unsupervised learning of narrative event chains. In: Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics (ACL); 2008 Jun 16-18; Columbus, OH. Stroudsburg: Association for Computational Linguistics; 2008:789-97
37. Peng H, Song C, Xie P, et al. CausalBERT: language models for causal inference. In: Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing; 2021 Nov 7-11; Punta Cana, Dominican Republic. Stroudsburg: Association for Computational Linguistics; 2021
38. Le LH, Hoang PA, Pham HC. Sharing health information across online platforms: a systematic review. Health Commun 2023 Jul;38(8):1550-62 [doi: 10.1080/10410236.2021.2019920] [PMID: 34978235]
39. Bayer S, Hettinger A. Storytelling. Bull Ecol Soc Am 2019 Apr;100(2):e01542 [doi: 10.1002/bes2.1542]

40. Storr W. The Science of Storytelling: Why Stories Make Us Human and How to Tell Them Better. New York: Abrams; 2020
41. Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas 1960 Apr;20(1):37-46 [doi: 10.1177/001316446002000104]
42. Sanh V, Debut L, Chaumond J, Wolf T. DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter. ArXiv 2019 Oct 2
43. Loshchilov I, Hutter F. Decoupled weight decay regularization. ArXiv 2017 Nov 14
44. Devlin J, Chang MW, Lee K, Toutanova K. BERT: pre-training of deep bidirectional transformers for language understanding. In: Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers); 2019 Jun 2-7; Minneapolis, MN. Stroudsburg: Association for Computational Linguistics; 2019:4171-86
45. Bolton J, Zhang Y, Zhang Y, Qi P, Manning CD. Stanza: a Python natural language processing toolkit for many human languages. Proc Int Conf Comput Linguist 2020;1:116-22
46. Fraile Navarro D, Ijaz K, Rezazadegan D, Rahimi-Ardabili H, Dras M, Coiera E, et al. Clinical named entity recognition and relation extraction using natural language processing of medical free text: a systematic review. Int J Med Inform 2023 Nov;177:105122 [doi: 10.1016/j.ijmedinf.2023.105122] [PMID: 37422950]
47. Sarker A, Ginn R, Nikfarjam A, et al. Utilizing social media data for pharmacovigilance: a review. J Biomed Inform 2015 Apr;54:202-12 [doi: 10.1016/j.jbi.2015.02.004] [PMID: 25720841]
48. Chattoraj S, Joshi KP. Semantically rich approach to automating regulations of medical devices. In: IEEE International Conference on Digital Health (ICDH); 2024 Jul; IEEE World Congress on Services; 2024
49. Islam M, Elluri L, Joshi KP. Integrating knowledge graphs with retrieval-augmented generation to automate IoT device security compliance. In: 2025 IEEE International Conference on Intelligence and Security Informatics (ISI); 2025 Jul
50. Li H, Wang X, Zhou Y, Liu W, Pan S. An overview of event extraction methods based on semantic disambiguation. In: Proceedings of the 2024 2nd International Conference on Artificial Intelligence, Systems and Network Security (AISNS '24); 2024 Feb 24-26; Fuzhou, China. New York: ACM; 2024:319-26
51. Paakki H, Ghorbanpour F, Sawhney N. Computational crisis narrative analysis: a case-study of social media narratives around the COVID-19 crisis in India. In: Conference on Language Technologies & Digital Humanities; 2022 Sep 21-23; Ljubljana, Slovenia
52. Zhou L, Parsons S, Hripcsak G. Temporal relation extraction in clinical texts: a systematic review. ACM Digital Library 2021;54(8):Article 162 [doi: 10.1145/3462475]
53. Chen Y, Wright K, Ferrara E. Multimodal learning for temporal relation extraction in clinical texts. J Am Med Inform Assoc 2024;31(6):1380-1390 [doi: 10.1093/jamia/ocae089]
54. Sun W, Rumshisky A, Uzuner O. Extraction of temporal information from clinical narratives: a semi-supervised approach. J Healthc Inform Res 2019;3(2):218-244 [doi: 10.1007/s41666-019-00049-0]
55. Chambers N, Jurafsky D. Unsupervised learning of narrative event chains. In: Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics (ACL); 2008 Jun 16-18; Columbus, OH. Stroudsburg: Association for Computational Linguistics; 2008:789-97
56. Peng H, Song C, Xie P, et al. CausalBERT: language models for causal inference. In: Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing; 2021 Nov 7-11; Punta Cana, Dominican Republic. Stroudsburg: Association for Computational Linguistics; 2021
57. Martel C, Pennycook G, Rand DG. Emotions in misinformation studies: distinguishing affective state from emotional response and misinformation recognition from acceptance. Cogn Res Princ Implic 2024;9:Article 63 [doi: 10.1186/s41235-024-00607-0]

58. Na K. Emotion and misinformation acceptance during public health crises. In: Greifeneder R, Jaffé ME, Newman EJ, Schwarz N, editors. The Psychology of Fake News. London: Routledge; 2025:Chapter 9
59. Metzger MJ, Flanagin AJ. The psychological impacts and message features of health misinformation: a systematic review of randomized controlled trials. Psychol Bull 2023;149(5-6):279-315 [doi: 10.1037/bul0000387]
60. Betsch C, Ulshöfer C, Renkewitz F, Betsch T. The influence of narrative v. statistical information on perceiving vaccination risks. Med Decis Making 2011;31(5):742-53
61. Zebregs S, van den Putte B, Neijens P, de Graaf A. The differential impact of statistical and narrative evidence on beliefs, attitude, and intention: a meta-analysis. Health Commun 2015;30(3):282-9 [doi: 10.1080/10410236.2013.842528]
62. Ooms J, Hoeks J, Jansen C. Narratives in health communication: a systematic review and content analysis. RAND External Publication 2025;EP-70805
63. Benton A, Mitchell M, Hovy D. Multi-task learning for mental health using social media text. In: Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics; 2017 Apr 3-7; Valencia, Spain. Stroudsburg: Association for Computational Linguistics; 2017:152-62
64. Zhang Y, Lyu H, Liu Y, et al. Enhancing mental health prediction via multi-task learning using Reddit data. medRxiv 2024 Feb 23 [doi: 10.1101/2024.02.23.24303303]
65. Golder S, Norman G, Loke YK. Harnessing social media data for pharmacovigilance: a review of current state of art, challenges and future directions. Int J Data Sci Anal 2019;8:113-35 [doi: 10.1007/s41060-019-00175-3]
66. European Medicines Agency. Guideline on good pharmacovigilance practices (GVP) Module VI – Collection, management and submission of reports of suspected adverse reactions to medicinal products (Rev 2). EMA/873138/2011 Rev 2. Amsterdam: European Medicines Agency; 2017. URL: https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/guideline-good-pharmacovigilance-practices-gvp-module-vi-collection-management-submission-reports_en.pdf. [accessed 2025-08-31]
67. US Food and Drug Administration. FDA adverse event reporting system (FAERS) public dashboard. Silver Spring (MD): US Food and Drug Administration; 2024. URL: https://www.fda.gov/drugs/questions-and-answers-fdas-adverse-event-reporting-system-faers/fda-adverse-event-reporting-system-faers-public-dashboard. [accessed 2025-08-31]
68. Eysenbach G, Till JE. Ethical issues in qualitative research on internet communities. BMJ. 2001 Nov 10;323(7321):1103-5. doi: 10.1136/bmj.323.7321.1103. PMID: 11701577; PMCID: PMC59687.n
69. Explosion. spaCy (v3.x). URL: https://spacy.io/ [accessed 2025-08-31].
70. PRAW community. PRAW: The Python Reddit API Wrapper. URL: https://praw.readthedocs.io/ [accessed 2025-08-31].
71. Hugging Face. Transformers (v4.x) documentation. URL: https://huggingface.co/docs/transformers [accessed 2025-08-31].
72. Hugging Face. distilbert-base-uncased (model card). URL: https://huggingface.co/distilbert-base-uncased [accessed 2025-08-31].
73. scikit-learn developers. scikit-learn. URL: https://scikit-learn.org/ [accessed 2025-08-31].
74. PyTorch Foundation. PyTorch. URL: https://pytorch.org/ [accessed 2025-08-31].
75. Georgetown IR Lab. QuickUMLS. URL: https://github.com/Georgetown-IR-Lab/QuickUMLS [accessed 2025-08-31].

Multimedia Appendices

- Multimedia Appendix 1 Supplementary tables and figures. [PDF File (Adobe PDF File), 2MB-

Multimedia Appendix 1]

- Multimedia Appendix 2 Paper supplements including technical architecture details and training configurations. [PDF File (Adobe PDF File), 1MB-Multimedia Appendix 2]

Supplementary Files

Untitled.
URL: http://asset.jmir.pub/assets/45f257a681f0b2d25490b64e2394ae92.pdf

Multimedia Appendixes

Supplementary File.
URL: http://asset.jmir.pub/assets/a3fc53238df12d7d363fd993cc0d6b6c.pdf

Supplementary file.
URL: http://asset.jmir.pub/assets/72fe70d08e659c4cdc42728a5a4d2a6d.pdf