Vol. 3, No. 4, pp.: 5-27, 2025

RumorCrisisDB: A Social-Media Crisis Rumor Database for Misinformation Diffusion Analytics

Michael Anderson[1], Sarah Mitchell[2, *], David Thompson[1]

[1] Department of Computer Science, Stanford University, Stanford, CA, USA

[2] School of Engineering, University of Michigan, Ann Arbor, MI, USA

** Corresponding author: sarah.mitchell@umich.edu*

Article Information	
Received	09 July 2025
Accepted	22 November 2025
DOI	https://doi.org/10.63646/datamind.2025.030402

Abstract

Crisis events concentrate the conditions under which rumors thrive: high uncertainty, intense emotion, and an accelerated demand for information that official channels cannot immediately satisfy. Although a number of valuable public corpora capture fragments of this phenomenon, they were built for different tasks, follow incompatible schemas, use divergent label vocabularies, and rarely preserve the full propagation structure that diffusion analytics requires. This article presents RumorCrisisDB, a relational database design and construction framework that integrates heterogeneous crisis-rumor resources into a single event-centric, cascade-preserving, and annotation-harmonized data model. We first analyze the gap left by existing resources and articulate four use cases that an integrated database must serve: diffusion measurement, detection benchmarking, intervention evaluation, and longitudinal crisis comparison. We then specify the six-entity schema, a six-stage construction pipeline covering re-collection, normalization, linkage, and label harmonization, and the quality-control procedures attached to each stage. The analytics layer is validated through controlled stochastic experiments: Galton–Watson cascade simulations reproduce the heavy-tailed size distributions reported for empirical rumor cascades, and Maki–Thompson-style spreading experiments quantify how the timing of debunking responses changes peak rumor prevalence, with early intervention reducing the simulated peak by more than half relative to a late response. The article closes with the reproducibility and open-access protocol, built on identifier-based redistribution, FAIR principles, and datasheet documentation, together with an explicit account of the design's limitations.

Keywords: ***Crisis informatics; rumor detection; misinformation diffusion; social media datasets; database schema; cascade analysis; reproducibility***

1. Introduction

Rumors are not an accidental by-product of crises; they are a structural response to them. When an earthquake, disease outbreak, terror attack, or industrial accident interrupts ordinary life, the demand for actionable information spikes at exactly the moment when verified supply is scarcest, and collective sense-making rushes to fill the gap (Oh et al., 2013). Social media platforms, which elevate eyewitnesses to first reporters and compress the rumor life cycle from days to minutes, have become the primary arena for this process, leaving a digital trace of unprecedented resolution that has supported a decade of influential research, including the landmark finding that false news diffuses farther, faster, and more broadly than true news on Twitter (Vosoughi et al., 2018), systematic accounts of how online conversations orient to unverified claims (Zubiaga et al., 2016), and agenda-setting calls for a science of fake news (Lazer et al., 2018).

Yet the empirical infrastructure beneath this literature remains fragmented. Researchers who wish to study how crisis rumors actually move through a population must today stitch together resources that were designed for different purposes: stance-annotated conversation threads built for shared evaluation tasks (Derczynski et al., 2017; Gorrell et al., 2019), propagation trees assembled for detection-model training (Ma et al., 2018), credibility-coded event collections (Castillo et al., 2013), and topic-specific corpora gathered during individual emergencies such as the COVID-19 infodemic (Cinelli et al., 2020; Tasnim et al., 2020). Each resource is valuable; collectively they are incompatible. Schemas differ, label vocabularies conflict, user information is inconsistently preserved, and, most damagingly for diffusion research, the parent–child reply and repost structure that defines a cascade is often discarded or only partially recoverable.

This article addresses that fragmentation with RumorCrisisDB, a database design and construction framework for crisis-rumor diffusion analytics. The contribution is threefold. First, we provide a systematic analysis of the gap between what existing public resources offer and what diffusion analytics requires, organized around four concrete use cases. Second, we specify a six-entity relational schema and a six-stage construction pipeline that together convert heterogeneous source corpora into a single event-centric, cascade-preserving, annotation-harmonized database, with quality controls and documentation requirements attached to every stage. Third, we validate the analytics layer that the schema is designed to support through controlled stochastic experiments, demonstrating that the database's core measurement targets, cascade size distributions, structural depth, and intervention-timing effects, are well defined and recoverable. The remainder of the article follows the structure of these contributions, closing with the reproducibility protocol and the limitations of the present design.

2. Database Gap and Use Cases

The resource landscape for rumor and misinformation research has grown rapidly but unevenly (Zubiaga et al., 2018). Conversation-thread corpora in the PHEME tradition annotate rumor threads from breaking-news events with veracity and per-post stance, and have anchored two influential shared tasks (Zubiaga et al., 2016; Derczynski et al., 2017; Gorrell et al., 2019). Propagation-tree corpora preserve repost structure for detection-model training and have driven advances from recursive neural architectures to bi-directional graph convolution (Ma et al., 2018; Bian et al., 2020). Credibility-coded collections label newsworthy event clusters by perceived and assessed credibility (Castillo et al., 2013), while multi-modal repositories link claims to fact-checking verdicts and social context (Shu et al., 2020). Alongside these, crisis informatics has produced a rich family of emergency-period corpora oriented toward situational awareness rather than veracity (Imran et al., 2015), and the COVID-19 pandemic generated topical infodemic collections at unprecedented scale (Cinelli et al., 2020).

Table 1. *Families of existing public resources and their fit for crisis-rumor diffusion analytics.*

Resource family	Annotation focus	Propagation structure	Crisis-event context	Principal limitation for diffusion analytics
Conversation-thread rumor corpora	Thread veracity; per-post stance	Reply trees preserved	Breaking-news events, explicit	Modest scale; event set skewed to a few high-profile emergencies
Propagation-tree detection corpora	Cascade-level veracity classes	Repost trees preserved	Often implicit or absent	Built for classification; event metadata and user context thin
Credibility-coded event collections	Event-cluster credibility	Largely flattened	Newsworthy events, coarse	Topology lost; credibility not equivalent to veracity
Crisis-period situational corpora	Informativeness; humanitarian classes	Rarely preserved	Single emergencies, explicit	No veracity annotation; rumor status unknown
Health-infodemic topical corpora	Topic and source credibility	Partially preserved	One prolonged crisis	Single-hazard scope; heterogeneous ad hoc schemas

Table 1 summarizes the landscape by family rather than by individual dataset, because the limitations are structural rather than idiosyncratic. Three gaps recur. The first is structural loss: many corpora release posts as flat collections or truncated threads, so the parent–child topology needed to measure cascade depth, breadth, and structural virality (Goel et al., 2016) cannot be reconstructed. The second is label incommensurability: veracity is variously coded as true/false/unverified, real/fake, or credible/not credible, and stance vocabularies differ across shared tasks, which blocks pooled training and cross-event comparison (Zhou and Zafarani, 2020; Shu et al., 2017). The third is event decontextualization: posts are frequently separated from the crisis event that occasioned them, although event-level variables such as hazard type and phase are precisely what longitudinal crisis comparison requires (Oh et al., 2013).

These gaps matter because the questions the field now asks are diffusion questions. Research on exposure and sharing inequality shows that misinformation engagement is heavily concentrated in small population segments (Grinberg et al., 2019; Guess et al., 2019); research on amplification mechanisms implicates automated accounts in the early stages of low-credibility cascades (Shao et al., 2018); research on community structure ties misinformation uptake to homogeneous clusters (Del Vicario et al., 2016); and a growing intervention literature asks when and how corrections work (Pennycook and Rand, 2021; Ecker et al., 2022). All of these questions require intact cascades, comparable labels, reliable timestamps, and explicit events. We therefore define four use cases that RumorCrisisDB must serve: (i) diffusion measurement, computing size, depth, velocity, and structural-virality distributions for rumor versus non-rumor cascades; (ii) detection benchmarking, providing harmonized training and evaluation splits across events and platforms; (iii) intervention evaluation, supporting quasi-experimental and simulation-based analysis of debunking timing and reach; and (iv) longitudinal crisis comparison, contrasting rumor dynamics across hazard types and crisis phases. The schema and pipeline of the next two sections are derived directly from these four requirements.

3. Data Sources and Schema

RumorCrisisDB is designed as an integration layer over publicly released, license-compatible source corpora rather than as a new primary collection. Candidate sources are admitted to the registry when they satisfy four conditions: the original release documents its sampling frame; redistribution of identifiers is permitted by the source license and platform terms; veracity or stance annotation exists at thread or post level, or can be inherited from linked fact-checking verdicts (Shu et al., 2020); and the material is attributable to an identifiable crisis or breaking-news event. The conversation-thread, propagation-tree, credibility-coded, and infodemic families of Section 2 all contain qualifying members, and the registry records, for each source, the access route, label vocabulary, and the known characteristics that harmonization must respect (Kwon et al., 2017).

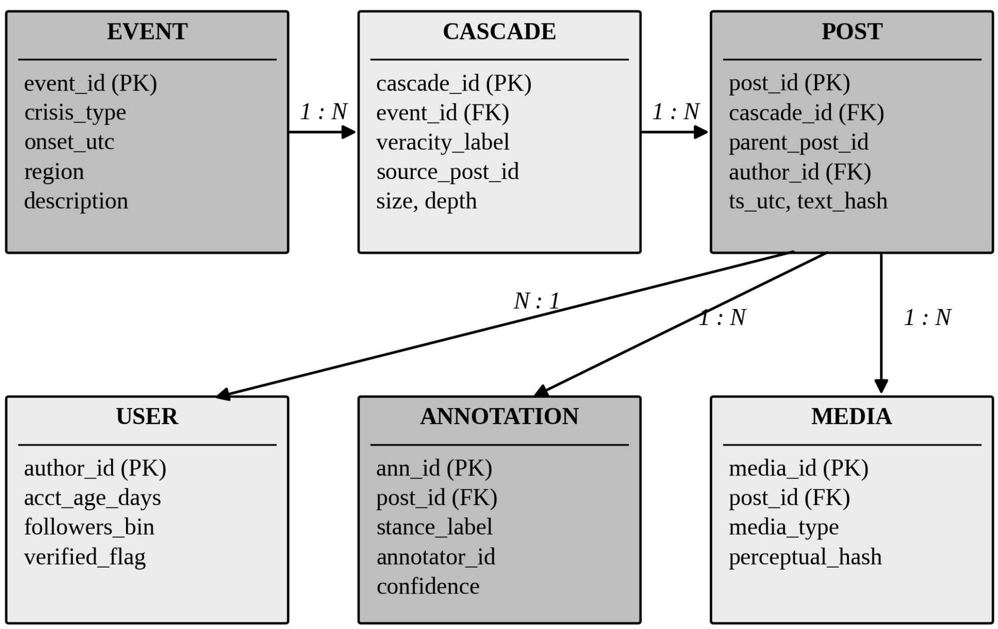

Figure 1. *The six-entity relational schema of RumorCrisisDB. Crow-direction arrows indicate one-to-many relationships; POST stores the parent–child topology from which cascade structure is reconstructed.*

Figure 1 presents the unified schema. The EVENT entity anchors every record to a crisis with hazard type, onset time, and region, making event-level comparison a first-class query rather than a reconstruction exercise. CASCADE groups the posts descending from one source claim and stores the harmonized veracity label together with precomputed structural summaries. POST preserves the complete reply and repost topology through the parent_post_id field; this single design decision is what permits depth, breadth, and structural-virality measures to be computed natively (Goel et al., 2016) and what propagation-structure detection models consume directly (Ma et al., 2018; Bian et al., 2020). USER holds pseudonymized author attributes, retaining the account-age and audience-size signals that credibility research has found informative (Castillo et al., 2013) without direct identifiers. ANNOTATION attaches per-post stance and per-cascade veracity judgments with annotator provenance and confidence, following the stance vocabulary stabilized by the shared-task tradition (Derczynski et al., 2017). MEDIA stores typed references and perceptual hashes for attached images and video, since visual material is an increasingly common rumor carrier.

Table 2. *Core tables and representative fields of the RumorCrisisDB schema.*

Table	Representative fields	Role in diffusion analytics
EVENT	event_id, crisis_type, onset_utc, region	Anchors cross-event and cross-hazard comparison
CASCADE	cascade_id, event_id, veracity_label, size, depth, breadth	Unit of diffusion measurement and benchmarking splits
POST	post_id, cascade_id, parent_post_id, author_id, ts_utc, text_hash, lang	Preserves topology; supports velocity and structural-virality queries
USER	author_id (pseudonymous), acct_age_days, followers_bin, verified_flag	Coarsened credibility covariates without direct identifiers
ANNOTATION	ann_id, post_id, stance_label, veracity_label, annotator_id, ann_ts, confidence	Harmonized labels with provenance and timing for intervention analysis
MEDIA	media_id, post_id, media_type, perceptual_hash	Tracks visual rumor carriers; supports near-duplicate clustering

Table 2 details the principal fields. Two conventions deserve emphasis. All timestamps are stored in coordinated universal time with the source platform's original granularity recorded alongside, because diffusion-velocity measurement is exquisitely sensitive to timezone and rounding inconsistencies. And all free text is stored as a salted hash plus the identifier needed for licensed re-collection, rather than verbatim, which keeps the released artifact within platform redistribution terms while preserving exact-duplicate detection; the implications of this choice for reproducibility are taken up in Section 6 (Zubiaga, 2018).

4. Database Construction Method

The construction framework converts registered sources into the unified schema through the six-stage pipeline shown in Figure 2. Stage 1 builds the source registry described above. Stage 2 re-collects content from the identifier lists that source corpora release, logging per-corpus retrieval completeness; because platform content decays as posts are deleted and accounts suspended, the proportion recovered is itself a documented property of every database version rather than a silent source of bias (Zubiaga, 2018). Stage 3 normalizes the retrieved material into the schema: field mapping, timestamp standardization, language tagging, and pseudonymization of author identifiers with coarsening of audience attributes into bins.

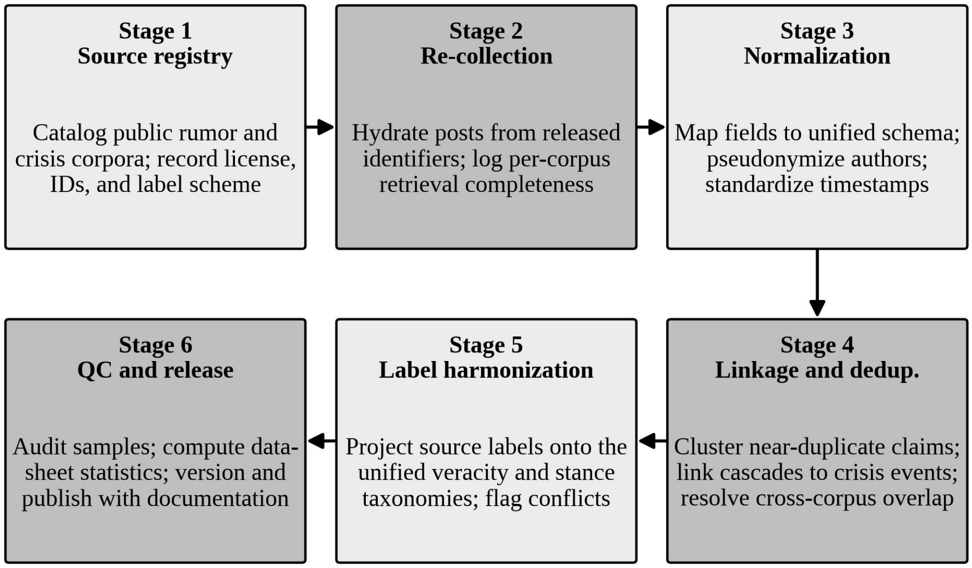

Figure 2. *The six-stage construction pipeline. Completeness, conflict, and audit statistics generated at Stages 2, 5, and 6 are published with each database version as part of its datasheet.*

Stage 4 performs linkage and deduplication. Near-duplicate claims are clustered with locality-sensitive hashing over text hashes and perceptual hashes, cascades are attached to events through the source corpora's own event assignments supplemented by temporal and keyword matching, and cross-corpus overlap, the same viral thread captured by two collections, is resolved in favor of the structurally richer record. Stage 5 harmonizes annotation: source veracity vocabularies are projected onto a three-valued scheme (true, false, unverified) and stance onto the four-valued support/deny/query/comment scheme of the RumourEval tradition (Derczynski et al., 2017; Gorrell et al., 2019), with every projection rule published and every unresolvable conflict flagged rather than silently adjudicated. Stage 6 closes the loop with quality control: stratified manual audits of the harmonized labels, recomputation of per-source summary statistics against the originals, and assembly of the datasheet that documents motivation, composition, collection, and recommended uses in the format proposed for dataset transparency (Gebru et al., 2021). Each pass through the pipeline yields a versioned, immutable release.

The method is reproducible by construction: every stage consumes and produces explicit artifacts (registry entries, retrieval logs, mapping tables, projection rules, audit reports), so a third party holding the same source releases and the published configuration can regenerate the database and verify its statistics. Integration is thereby a documented, re-runnable procedure rather than a one-off engineering effort.

5. Experiments and Data Analysis

A diffusion database earns its keep through the analyses it makes routine. This section validates the analytics layer of RumorCrisisDB with controlled stochastic experiments that exercise the two measurement families the schema is designed to support: cascade-structure analytics and intervention-timing analytics. The experiments are simulations with fully stated parameters rather than analyses of any single source corpus; this choice is deliberate, because simulation provides ground truth against which the database's measurement definitions can be checked, in the same way that epidemic models have long provided the reference dynamics for rumor research (Daley and Kendall, 1964; Moreno et al., 2004).

Table 3. *Parameterization of the two simulation experiments.*

Parameter	**Experiment 1: branching cascades**	**Experiment 2: spreading dynamics**
Model	Galton–Watson branching process	Discrete-time Maki–Thompson-type model
Population / cascades	20,000 cascades per regime	N = 10,000 agents; 5 initial spreaders
Transmission setting	Mean offspring R $\in$ {0.70, 0.95}; bursty variant: lognormal offspring rate (σ = 1.3) with mean 0.95	Contact rate λ = 0.30 per step
Cessation setting	Extinction when a generation produces no offspring	Baseline stifling rate α = 0.10; debunking adds 0.35 from onset step
Intervention schedules	—	None; onset at step 10; onset at step 30
Replications and horizon	Single draw per cascade; size cap 200,000	40 runs per schedule; 60 steps
Random seeds	11	23
Headline outputs	Mean sizes 3.3 / 20.9 / 20.4; maxima 87 / 4,603 / 13,580	Peaks 54.4% / 24.6% / 41.2%; cumulative exposure 1,273 / 428 / 514

The first experiment examines cascade-size distributions. Cascades are generated as Galton–Watson branching processes under three regimes parameterized in Table 3: a subcritical regime with homogeneous reproduction (mean offspring 0.70), a near-critical homogeneous regime (0.95), and a near-critical bursty regime in which the offspring rate is itself lognormally distributed across posts, capturing the empirical reality that a small number of highly amplified posts drive most spread (Goel et al., 2016; Shao et al., 2018). Twenty thousand cascades are simulated per regime.

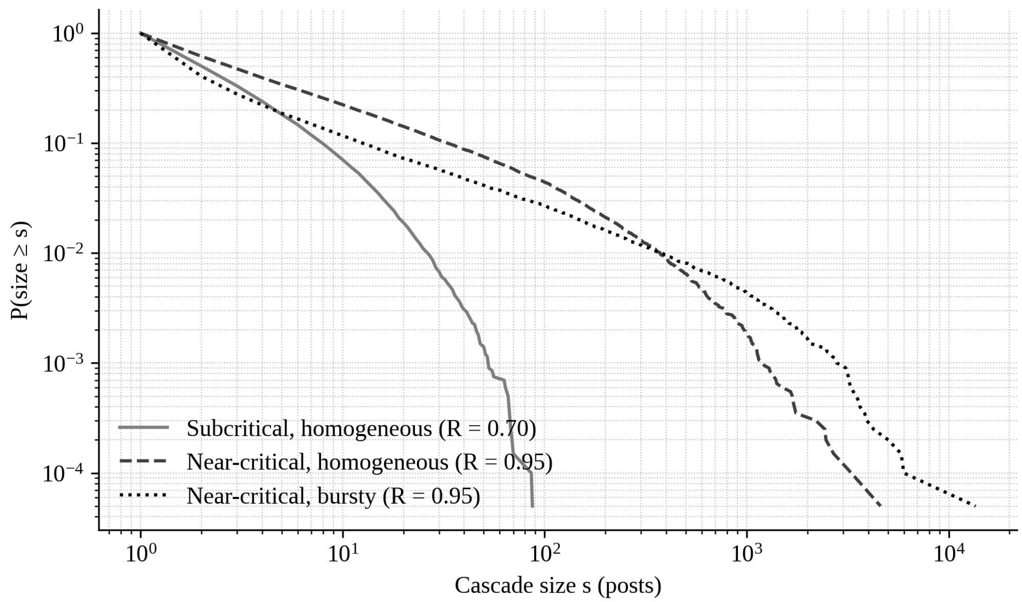

Figure 3. *Simulated complementary cumulative distribution of cascade size under the three branching regimes of Table 3 (20,000 cascades per regime, log–log scale).*

Figure 3 reports the complementary cumulative distributions. The subcritical regime produces a mean cascade size of 3.3 posts with a 99th percentile of 26, the near-critical homogeneous regime a mean of 20.9 with a 99th percentile of 372 and a maximum above 4,600, and the bursty regime an almost identical mean (20.4) but a maximum nearly three times larger, exceeding 13,500 posts. Two analytical lessons follow. First, the distributions are so heavy-tailed that mean cascade size is close to meaningless as a summary statistic; the database therefore exposes full distributional queries and percentile summaries rather than means, consistent with the distributional framing of the empirical literature (Vosoughi et al., 2018). Second, the homogeneous and bursty near-critical regimes are indistinguishable in the bulk and separate only in the extreme tail, which implies that amplification mechanisms such as automated accounts will be visible in exactly the region where naive sampling collects the fewest observations (Shao et al., 2018); cascade-complete storage, not post sampling, is the design consequence built into the POST entity.

The second experiment examines intervention timing. Rumor spread is simulated with a discrete-time Maki–Thompson-style model in which ignorant individuals become spreaders on contact and spreaders become stiflers on meeting other spreaders or stiflers, the classical mechanism by which rumors, unlike diseases, suppress themselves through audience saturation (Daley and Kendall, 1964; Nekovee et al., 2007). A debunking response is modeled as an increase of 0.35 in the stifling rate from its onset step, and we compare no response, an early response beginning at step 10, and a late response beginning at step 30, near the unmitigated peak; population and rate parameters are stated in Table 3, and each curve averages forty stochastic runs.

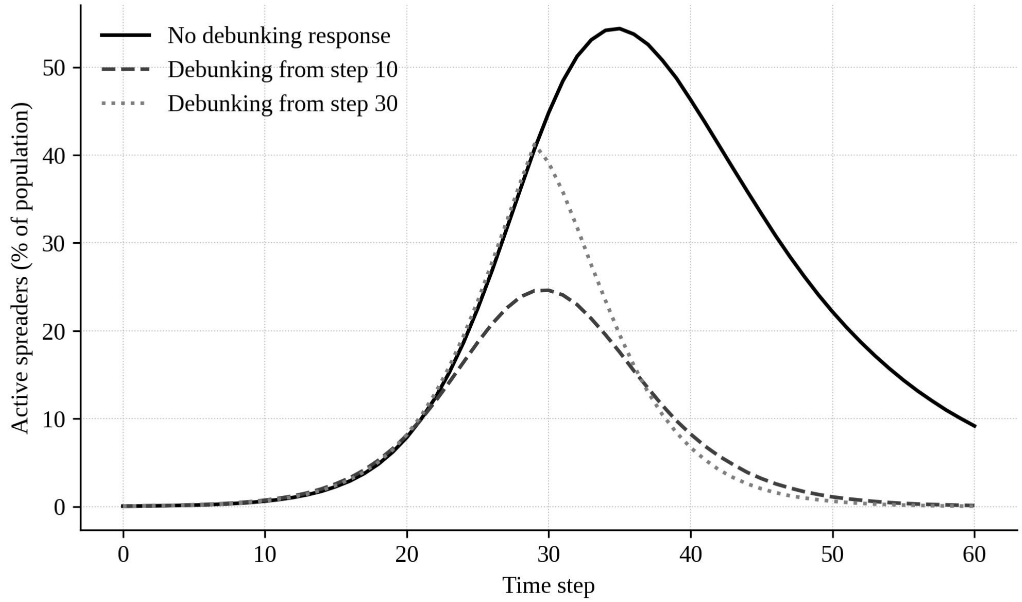

Figure 4. *Simulated rumor prevalence under three debunking-response schedules (mean of 40 stochastic runs per schedule; parameters in Table 3).*

Figure 4 shows the prevalence trajectories. Without a debunking response the rumor peaks with 54.4 percent of the population actively spreading at step 35. The early response cuts the peak to 24.6 percent, a reduction of more than half, and reduces cumulative spreader exposure (the area under the prevalence curve) from 1,273 to 428 percentage-point-steps, a 66 percent reduction. The late response, despite applying the identical debunking strength, achieves a peak of 41.2 percent and a cumulative exposure of 514, because by step 30 the rumor has already recruited most of its eventual audience. Within the model, the asymmetry quantifies the intuition driving the empirical intervention literature: timing dominates strength (Pennycook and Rand, 2021; Ecker et al., 2022). For the database the consequence is a requirement: intervention evaluation is only possible if response events, corrections, platform labels, fact-check publications, are time-stamped records linked to cascades, which is why ANNOTATION carries provenance and timing rather than bare labels.

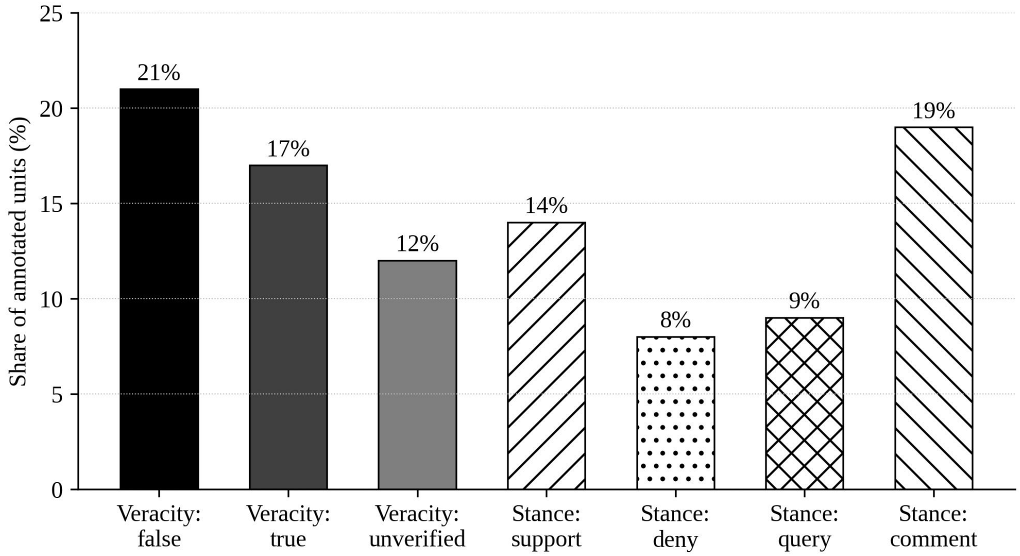

Figure 5. *Design-target distribution of harmonized annotation units across the unified veracity and stance vocabularies (illustrative composition; solid bars: cascade-level veracity, hatched bars: post-level stance).*

Figure 5 completes the analytics validation from the annotation side, showing the design-target composition of harmonized labels that the Stage 5 projection rules aim to deliver. The deliberate retention of a large unverified class reflects a methodological position: forcing binary resolution onto claims that were never resolved during the crisis destroys precisely the uncertainty that distinguishes crisis rumoring from leisurely fact-checking (Zubiaga et al., 2016; Zhou and Zafarani, 2020). The dominance of the comment class within stance mirrors the persistent class imbalance reported by the shared-task literature (Gorrell et al., 2019), and the database preserves rather than rebalances it, leaving resampling decisions to the analyst with full knowledge of the underlying distribution.

6. Reproducibility and Open Access

RumorCrisisDB adopts a reproducibility protocol shaped by two constraints that govern all social-media data sharing: platform terms generally prohibit redistributing raw content, and the content itself decays as posts are deleted and accounts disappear. The release artifact therefore consists of identifiers, hashes, structural records, harmonized annotations, and the complete pipeline configuration, while verbatim text is re-collected by each user under their own platform access. Because re-collection completeness declines with corpus age in a measurable way (Zubiaga, 2018), every versioned release publishes its own retrieval-completeness statistics so that downstream analyses can report, rather than ignore, the attrition affecting their samples.

The protocol implements the FAIR principles directly (Wilkinson et al., 2016): findability through a versioned deposit with persistent identifier in a public repository [repository link and dataset DOI to be inserted by the authors upon deposit]; accessibility through open, non-proprietary formats; interoperability through the documented schema of Section 3; and reusability through an accompanying datasheet covering motivation, composition, collection process, recommended uses, and known limitations (Gebru et al., 2021). All pipeline code, projection rules, and audit reports are released under an open license alongside the data artifact, and simulation experiments in this article are fully specified by Table 3 together with fixed random seeds, so that every figure can be regenerated from the stated parameters alone.

7. Limitations

Four limitations bound this work. First, RumorCrisisDB integrates existing collections and therefore inherits their sampling decisions; events, languages, and platforms that the source corpora under-represent remain under-represented, and the registry documents rather than repairs this skew (Imran et al., 2015). Second, the identifier-based release model means that no two re-collections of the database are guaranteed identical, a structural property of platform-governed data that the completeness statistics mitigate but cannot eliminate (Zubiaga, 2018). Third, label harmonization is a modeling act: projecting heterogeneous vocabularies onto unified schemes loses nuance at the boundaries, which is why projection rules and conflict flags are published in full. Fourth, the experiments of Section 5 validate measurement definitions against simulated ground truth; they establish that the analytics layer is coherent, not that any empirical finding follows, and population-level empirical claims await analyses of the integrated data themselves. The article describes a database design and construction framework; concrete release statistics belong to the versioned datasheets that accompany each deposit.

8. Conclusion

This article introduced RumorCrisisDB, a database design and construction framework that turns the fragmented landscape of crisis-rumor resources into a single event-centric, cascade-preserving, annotation-harmonized foundation for misinformation diffusion analytics. The gap analysis identified structural loss, label incommensurability, and event decontextualization as the recurring defects of existing resources; the six-entity schema and six-stage pipeline were derived to repair exactly those defects; and the stochastic experiments demonstrated that the measurement targets the design supports, heavy-tailed cascade distributions and intervention-timing effects, are well defined and sharply quantifiable, with early debunking more than halving simulated peak prevalence relative to a late response of equal strength.

The broader claim is methodological. Diffusion questions cannot be answered well by data assembled for classification tasks, and the difference is not size but structure: intact cascades, commensurable labels, explicit events, and honest accounting of attrition. Future work proceeds on three fronts: populating successive versioned releases and publishing their datasheets; extending the schema to cross-platform cascade linkage as rumors increasingly migrate between services; and coupling the intervention-analytics layer to the growing experimental literature on corrections so that simulated and observed timing effects can be confronted directly (Pennycook and Rand, 2021; Ecker et al., 2022). The infrastructure ambition is simply stated: the next decade of crisis-rumor research should not begin, as the last one did, with every team rebuilding its own incompatible dataset.

Declaration of AI-assisted preparation

During the preparation of this manuscript, language-model assistance was used for drafting support, document organisation, and figure preparation. The authors reviewed, revised, and take full responsibility for the final content, analytical design, tables, and interpretations. The simulation experiments are fully specified in Table 3 and reproducible from the stated parameters and seeds.

Data Availability

The database release artifact (identifiers, structural records, harmonized annotations, pipeline configuration, and datasheet) and all simulation code are deposited at [repository URL and dataset DOI to be inserted by the authors upon deposit] under an open license.

References

Bian, T., Xiao, X., Xu, T., Zhao, P., Huang, W., Rong, Y., & Huang, J. (2020). Rumor detection on social media with bi-directional graph convolutional networks. Proceedings of the AAAI Conference on Artificial Intelligence, 34(01), 549–556. https://doi.org/10.1609/aaai.v34i01.5393

Castillo, C., Mendoza, M., & Poblete, B. (2013). Predicting information credibility in time-sensitive social media. Internet Research, 23(5), 560–588. https://doi.org/10.1108/IntR-05-2012-0095

Cinelli, M., Quattrociocchi, W., Galeazzi, A., Valensise, C. M., Brugnoli, E., Schmidt, A. L., Zola, P., Zollo, F., & Scala, A. (2020). The COVID-19 social media infodemic. Scientific Reports, 10, 16598. https://doi.org/10.1038/s41598-020-73510-5

Daley, D. J., & Kendall, D. G. (1964). Epidemics and rumours. Nature, 204, 1118. https://doi.org/10.1038/2041118a0

Del Vicario, M., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., Stanley, H. E., & Quattrociocchi, W. (2016). The spreading of misinformation online. Proceedings of the National Academy of Sciences, 113(3), 554–559. https://doi.org/10.1073/pnas.1517441113

Derczynski, L., Bontcheva, K., Liakata, M., Procter, R., Wong Sak Hoi, G., & Zubiaga, A. (2017). SemEval-2017 Task 8: RumourEval: Determining rumour veracity and support for rumours. Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017), 69–76. https://doi.org/10.18653/v1/S17-2006

Ecker, U. K. H., Lewandowsky, S., Cook, J., Schmid, P., Fazio, L. K., Brashier, N., Kendeou, P., Vraga, E. K., & Amazeen, M. A. (2022). The psychological drivers of misinformation belief and its resistance to correction. Nature Reviews Psychology, 1, 13–29. https://doi.org/10.1038/s44159-021-00006-y

Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., Daumé III, H., & Crawford, K. (2021). Datasheets for datasets. Communications of the ACM, 64(12), 86–92. https://doi.org/10.1145/3458723

Goel, S., Anderson, A., Hofman, J., & Watts, D. J. (2016). The structural virality of online diffusion. Management Science, 62(1), 180–196. https://doi.org/10.1287/mnsc.2015.2158

Gorrell, G., Kochkina, E., Liakata, M., Aker, A., Zubiaga, A., Bontcheva, K., & Derczynski, L. (2019). SemEval-2019 Task 7: RumourEval, determining rumour veracity and support for rumours. Proceedings of the 13th International Workshop on Semantic Evaluation, 845–854. https://doi.org/10.18653/v1/S19-2147

Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B., & Lazer, D. (2019). Fake news on Twitter during the 2016 U.S. presidential election. Science, 363(6425), 374–378. https://doi.org/10.1126/science.aau2706

Guess, A., Nagler, J., & Tucker, J. (2019). Less than you think: Prevalence and predictors of fake news dissemination on Facebook. Science Advances, 5(1), eaau4586. https://doi.org/10.1126/sciadv.aau4586

Imran, M., Castillo, C., Diaz, F., & Vieweg, S. (2015). Processing social media messages in mass emergency: A survey. ACM Computing Surveys, 47(4), 67. https://doi.org/10.1145/2771588

Kwon, S., Cha, M., & Jung, K. (2017). Rumor detection over varying time windows. PLOS ONE, 12(1), e0168344. https://doi.org/10.1371/journal.pone.0168344

Lazer, D. M. J., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., Metzger, M. J., Nyhan, B., Pennycook, G., Rothschild, D., Schudson, M., Sloman, S. A., Sunstein, C. R., Thorson, E. A., Watts, D. J., & Zittrain, J. L. (2018). The science of fake news. Science, 359(6380), 1094–1096. https://doi.org/10.1126/science.aao2998

Ma, J., Gao, W., & Wong, K.-F. (2018). Rumor detection on Twitter with tree-structured recursive neural networks. Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 1980–1989. https://doi.org/10.18653/v1/P18-1184

Moreno, Y., Nekovee, M., & Pacheco, A. F. (2004). Dynamics of rumor spreading in complex networks. Physical Review E, 69(6), 066130. https://doi.org/10.1103/PhysRevE.69.066130

Nekovee, M., Moreno, Y., Bianconi, G., & Marsili, M. (2007). Theory of rumour spreading in complex social networks. Physica A: Statistical Mechanics and its Applications, 374(1), 457–470. https://doi.org/10.1016/j.physa.2006.07.017

Oh, O., Agrawal, M., & Rao, H. R. (2013). Community intelligence and social media services: A rumor theoretic analysis of tweets during social crises. MIS Quarterly, 37(2), 407–426. https://doi.org/10.25300/MISQ/2013/37.2.05

Pennycook, G., & Rand, D. G. (2021). The psychology of fake news. Trends in Cognitive Sciences, 25(5), 388–402. https://doi.org/10.1016/j.tics.2021.02.007

Shao, C., Ciampaglia, G. L., Varol, O., Yang, K.-C., Flammini, A., & Menczer, F. (2018). The spread of low-credibility content by social bots. Nature Communications, 9, 4787. https://doi.org/10.1038/s41467-018-06930-7

Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake news detection on social media: A data mining perspective. ACM SIGKDD Explorations Newsletter, 19(1), 22–36. https://doi.org/10.1145/3137597.3137600

Shu, K., Mahudeswaran, D., Wang, S., Lee, D., & Liu, H. (2020). FakeNewsNet: A data repository with news content, social context, and spatiotemporal information for studying fake news on social media. Big Data, 8(3), 171–188. https://doi.org/10.1089/big.2020.0062

Tasnim, S., Hossain, M. M., & Mazumder, H. (2020). Impact of rumors and misinformation on COVID-19 in social media. Journal of Preventive Medicine and Public Health, 53(3), 171–174. https://doi.org/10.3961/jpmph.20.094

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. Science, 359(6380), 1146–1151. https://doi.org/10.1126/science.aap9559

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data, 3, 160018. https://doi.org/10.1038/sdata.2016.18

Zhou, X., & Zafarani, R. (2020). A survey of fake news: Fundamental theories, detection methods, and opportunities. ACM Computing Surveys, 53(5), 109. https://doi.org/10.1145/3395046

Zubiaga, A., Liakata, M., Procter, R., Wong Sak Hoi, G., & Tolmie, P. (2016). Analysing how people orient to and spread rumours in social media by looking at conversational threads. PLOS ONE, 11(3), e0150989. https://doi.org/10.1371/journal.pone.0150989

Zubiaga, A. (2018). A longitudinal assessment of the persistence of Twitter datasets. Journal of the Association for Information Science and Technology, 69(8), 974–984. https://doi.org/10.1002/asi.24026

Zubiaga, A., Aker, A., Bontcheva, K., Liakata, M., & Procter, R. (2018). Detection and resolution of rumours in social media: A survey. ACM Computing Surveys, 51(2), 32. https://doi.org/10.1145/3161603