Bachelor thesis

Crisis Cartography: An AI-Driven Dashboard for Real-Time Outbreak and Trend Detection via ActivityPub

Panayiotis Ioakim

University of Cyprus
Department of Computer Science

Department of Computer Science

May 2025

UNIVERSITY OF CYPRUS

Faculty of Pure and Applied Sciences

Department of Computer Science

Crisis Cartography: An AI-Driven Dashboard for Real-Time Outbreak and Trend Detection via ActivityPub

Panayiotis Ioakim

Advisor:

Dr. Panayiotis Kolios

The Thesis was submitted in partial fulfillment of the requirements for obtaining the Computer Science degree of the Department of Computer Science of the University of Cyprus

TABLE OF CONTENTS

List of Abbreviations

BERT	Bidirectional Encoder Representations from Transformers
CDC	Centers for Disease Control and Prevention
DID	Decentralized Identity
ECDC	European Centre for Disease Prevention and Control
HTML	HyperText Markup Language
IEEE	Institute of Electrical and Electronics Engineers
IPFS	InterPlanetary File System
JSON	JavaScript Object Notation
LDA	Latent Dirichlet Allocation
LLaMA	Large Language Model Meta AI
ML	Machine Learning
NER	Named Entity Recognition
NLP	Natural Language Processing
NOAA	National Oceanic and Atmospheric Administration
ORM	Object-Relational Mapper
OSINT	Open-Source Intelligence
RSS	Really Simple Syndication
SQL	Structured Query Language
SSL	Secure Sockets Layer
VADER	Valence Aware Dictionary and sEntiment Reasoner
W3C	World Wide Web Consortium
WHO	World Health Organization
XMPP	Extensible Messaging and Presence Protocol

1 Abstract

This thesis develops and evaluates a decentralized AI driven system for real–time health monitoring using the ActivityPub protocol. Public posts from federated platforms like Mastodon, PeerTube, Lemmy, and Misskey are collected and processed through a Django + Celery pipeline. Latent Dirichlet Allocation uncovers key health topics, while a LLaMA based model classifies relevance and sentiment. Grafana dashboards display live trends.

Our analysis reveals that shifts in public emotion and urgency often precede official outbreak reports. Topic modeling uncovers emerging concerns around symptoms and preventive measures. Geomapped data pinpoints new hotspots before case counts rise. User engagement metrics – replies, boosts and favourites, highlight how communities spread and endorse health advice.

These findings demonstrate that federated social networks can yield timely, actionable insights for public health. By combining open, distributed data sources with machine learning and real-time visualization, our system augments traditional surveillance. Future work should expand platform coverage, enhance model accuracy, and explore privacy preserving analytics.

2 Introduction

Social media has changed how we talk about health. It has also changed how we track it. Every day, people share symptoms, health concerns, and local events online. These posts can signal larger patterns, such as flu outbreaks, new variants of COVID, or mental health trends. But until recently, most of this data remained inaccessible inside the platforms that owned it.

That is starting to shift. Federated networks like Mastodon, PeerTube, Lemmy, and Misskey offer a new model. They use the ActivityPub protocol to connect independent servers. No single company controls the data. This makes it possible to collect information in a more open and decentralized manner, shifting away from reliance on centralized social media APIs. However, it does not necessarily simplify the process of data collection. Instead, it enables access to communities that are part of the federated network and have opted into data sharing. In essence, ActivityPub changes not just where the data comes from, but who you can reach and what you can build from those connections.

This thesis explores how federated networks and artificial intelligence can work together for public health. It focuses on building a system that collects health-related posts in real time, analyzes them using natural language processing (NLP) and machine learning (ML), and turns them into useful insights. These insights help spot early warning signs of outbreaks and changes in peoples sentiment. Everything is visualized through Grafana, giving health professionals a clear view of whats happening, where, and when.

The research concentrates on one core idea: public health data does not have to come from official reports alone. It can come from everyday conversations too; especially when we have the tools to find meaning in the noise. This system aims to be that tool.

2.1 Objectives & Contributions

Objectives

- **Federated Data Integration**: Develop a Django application that fetches public health posts from Mastodon, PeerTube, Lemmy, and Misskey, overcoming heterogeneous APIs and rate limits.
- **Automated Health Analysis**: Implement LLaMA based NLP alongside LDA topic modeling and sentiment analysis to classify content and surface emerging themes.
- **RealTime Visualization**: Create Grafana dashboards that update live, revealing outbreak signals, sentiment shifts, and geospatial hotspots.
- **Scalable Architecture**: Orchestrate data fetching and processing with Celery and Redis under Docker, ensuring fault tolerant, autonomous operation.

Contributions

1. **End to End Federated Pipeline**: Unifies data collection, AI–based content processing, and interactive dashboards.
2. **Hybrid ML Framework**: Blends transformer inference (LLaMA) with classical topic models for nuanced health-content tagging.
3. **Live Dashboards**: Offers public health professionals immediate visibility into user sentiment, symptom spikes, and location-based alerts.
4. **Robust Deployment**: A containerized setup that scales seamlessly across instances and recovers automatically from failures.
5. **Critical Decentralization Analysis**: Examines privacy, interoperability, and scalability challenges within the ActivityPub ecosystem.

2.2 Research Questions

This study is driven by four central questions:

1. In what ways can the ActivityPub protocol support the collection of decentralized public health data as it emerges in real time?
2. Which machine learning methods can be used effectively for identifying and classifying health-related content from federated social media networks?
3. How can live visualizations make public health data more useful and accessible to policymakers and public health professionals?
4. What technical obstacles and ethical concerns arise when collecting and analyzing federated health data?

3 Literature Review

3.1 ActivityPub Protocol Overview

ActivityPub is a decentralized social networking protocol standardized by the World Wide Web Consortium (W3C) in January 2018. Building upon the ActivityStreams 2.0 data format, it facilitates both client-to-server and server-to-server interactions, enabling the creation, updating, and deletion of content, as well as the delivery of notifications and content across federated networks. This protocol underpins various platforms within the "Fediverse," such as Mastodon, PeerTube, and Lemmy, promoting interoperability and user autonomy across diverse social applications [1].

3.2 Geolocalized Emotion and Public Health

This study by Feng and Kirkley (2021)[2] investigates how emotions expressed online, particularly on Twitter, relate to local policies and offline mobility during the early months of the COVID-19 pandemic in the United States. Using around 13 million geotagged tweets across 49 major U.S. cities, the authors constructed a measure called online geolocalized emotion (OGE), which quantifies daily sentiment around COVID-19-related topics like distancing, masking, economy, China, and the Trump administration.

The study finds strong consistency in emotional trends across cities, especially in response to national events. However, the emotional reaction to policies especially local ones ,varies significantly between cities. This variation, not the overall trend, reveals deeper regional differences in how populations emotionally responded to policy decisions.

Interestingly, online emotional responses are found to correlate strongly with physical mobility (driving and walking levels), but weakly with case and death counts. Using time series and Granger causality analysis, the authors show that sentiment changes can predict shifts in mobility up to two weeks in advance. Cities with more engagement online also tended to show stronger feedback between sentiment and movement.

The study makes clear that federal policies like Trump's national emergency declaration had some emotional effect, but local measures (e.g., shelter-in-place orders) sparked stronger emotional responses. This supports the idea that people respond more emotionally to policies that directly affect their daily life. It also suggests that policy communication strategies might benefit from paying closer attention to local sentiment, especially when aiming for behavioral compliance.

Despite strong overall trends in negative sentiment, the subtopics reveal more nuance. For example, "distancing" saw an increase in positivity over time, possibly indicating growing acceptance. The study also highlights how sentiment around Trump and China became more negative in response to federal action, reinforcing the role of political framing in emotional response.

The authors argue that integrating social media sentiment with mobility data offers a way to inform and possibly predict public behavior during crises. Still, they caution against

assuming that online emotions reflect the scientific reality of the pandemic, noting a weak link between emotion and actual infection data. Ethical concerns, such as user privacy and digital bias, are also raised.

In summary, the study reveals a feedback loop: policy influences emotion, which influences behavior. That loop, if properly understood, could be used to design more effective, spatially targeted interventions in future crises.

3.3 Social Media and Public Attitudes toward Vaccination

The Chandrasekaran et al. (2022)[3] study captures public attitudes toward COVID-19 vaccination by analyzing 2.94 million tweets from individuals between January and April 2021. Using CorEx topic modeling and VADER sentiment analysis, the authors identified 16 distinct topics grouped into six themes: vaccination experiences, pharmaceutical industry, policy debates, rollout logistics, attitudes toward vaccination, and expressions of gratitude to healthcare workers.

The most discussed issue by far was vaccine regulation, especially around mandates versus personal choice. Tweets reflected sharp divisions. Some called for mandatory vaccines in workplaces and schools, while others warned of overreach and loss of bodily autonomy. Sentiment was mixed but trended slightly positive overall on this topic.

The second most discussed issue was vaccine hesitancy where the sentiment was clearly negative. Concerns ranged from the speed of development, potential side effects, and mistrust in institutions, to fears about long-term consequences. These worries were persistent, and despite targeted campaigns, they remained largely unresolved during the study period.

Interestingly, tweets discussing post-vaccination experiences, especially side effects, were also mostly negative in sentiment but this negativity was nuanced. Many posts balanced personal discomfort with a sense of collective responsibility, often ending with affirmations about safety or community protection.

Other topics, such as vaccine efficacy, trial approvals, appointment logistics, and public endorsements, trended positively throughout. This suggests that while skepticism remained high in pockets, there was broader trust in the science and process at least among the Twitter-using public.

This work contributes to ongoing public health infoveillance by offering fine-grained temporal tracking of sentiment across multiple facets of vaccination discourse. Its use of user-only tweet data (excluding media outlets) strengthens its ability to reflect actual public emotion, not just media framing.These approaches underscore the need for real-time, topic-specific emotional monitoring at scale. Federated systems that support multilingual, community-driven, and privacy-respecting data collection would be well positioned to replicate this kind of insight, especially when integrated into public health dashboards or decision-making workflows.

3.4 Social Media's Role During the COVID-19 Pandemic

This scoping review[4] analyzes 81 peer-reviewed empirical studies on the role of social media during the first wave of the COVID-19 pandemic (Nov 2019-Nov 2020). The studies were grouped into six public health themes: infodemics, public attitudes, mental health, detection or prediction of cases, government responses, and quality of health information in videos. The dominant platform studied was Twitter, followed by China's Sina Weibo. Most studies relied on content and sentiment analysis, with relatively few employing machine learning.

A major focus across studies was the spread of misinformation termed "infodemics" by the WHO. Roughly one-fourth of analyzed tweets contained false or unverifiable content. While false posts were more frequent, posts based on scientific information had higher engagement and trust. WhatsApp and Facebook were also identified as major hubs for misinformation, with WhatsApp showing the highest prevalence of false content among all platforms examined. Interestingly, many conspiracy theories were found to originate from regular users, not bots or influencers, though bots often amplified hate speech or polarization.

Sentiment analysis revealed sharp public mood swings though the descriptive analyses do not formally test or explicitly state the causal reasons underlying these shifts. Most people started with fear, then moved to anger, fatigue, or even humor. Many tweets expressed frustration toward government policies or praised healthcare workers. Some studies noted a gendered dimension to social media expression, with women tweeting more about caregiving and health safety, while men focused more on policy or sports cancellations.

Several studies confirmed that search engine queries and social media mentions of symptoms like dry cough or fever predicted case spikes up to 812 days in advance. Social media, therefore, held real predictive power for health surveillance but was largely underused in real-time applications. Only six studies addressed outbreak prediction, and none developed sustained surveillance infrastructure.

On mental health, few studies tackled the issue rigorously. Those that did found a strong correlation between heavy social media use and anxiety, depression, or social risk sensitivity. This reinforces the concern that public health crises magnify psychological strain via the same platforms used for information access.

One key gap identified was the limited application of machine learning and natural language processing for real-time decision-making. While studies deployed tools like LDA, BERT, or random forests for post-hoc analysis, there was almost no work on integrating these methods into live policy systems.

Although the Review does not explicitly propose alternative platform architectures, its finding that Twitter was the leading social media sourcefollowed closely by Sina Weibo underscores the empirical rationale for exploring federated, privacy-respecting, open alternatives . An ActivityPub-based approach if paired with machine-learning pipelinescould help build a real-time, distributed public-health observatory capable of tracking sentiment shifts, detecting early outbreaks, and flagging misinformation without relying on centralized social-media monopolies. The work surveyed here thus provides a strong empirical foundation for why such alternatives merit development.[4]

3.5 The Use of Social Networking Sites in Public Health Practice

This systematic review Capurro et al. (2014)[5] analyzes how social networking sites (SNSs) have been used in public health practice and research. Covering 73 peer-reviewed articles published up to early 2012, the review maps trends, identifies gaps, and evaluates how SNSs like Facebook, Twitter, and MySpace have supported surveillance, engagement, communication, and research. Most of these studies were recent: over two-thirds were published in the final two years of the review period reflecting the rapid rise of SNSs.

The majority of studies were observational and descriptive while few of them used experimental methods. Only five randomized controlled trials were found. This shows that while SNSs were being widely explored for data collection or monitoring, they had not yet been fully tapped for testing interventions or two-way communication. Most studies (86%) simply described user behavior or platform usage. Only 3% used SNSs for interactive communication or engagement. This underuse of SNS's core strength-multidirectional interaction is a missed opportunity.

Despite this, SNSs proved valuable for reaching populations traditionally considered hard to reach, such as adolescents, LGBTQ+ communities, or people at risk of STDs and substance abuse. In fact, nearly 60% of the reviewed studies targeted such groups. This shows the potential of SNSs for health outreach, particularly when anonymity and ease of access are critical.

A few articles stood out for their innovation. For example, the platform "PatientsLikeMe" allowed researchers to gather patient-reported outcomes on off-label drug use. Other studies used SNSs for recruitment into longitudinal or behavioral health studies, maintaining contact with participants, and promoting testing for HIV.

Only one study was conducted in a low-income country. This highlights a serious equity gap, especially as mobile-first SNS access is growing in many such regions. For future global public health frameworks especially decentralized onesthis gap matters. Federated platforms like those using the ActivityPub protocol could bridge these gaps, especially when they empower local ownership, anonymity, and ethical participation in data exchange.

The review ends with a call for more active and interactive uses of SNSs in public health, including experimentation, longitudinal tracking, and full use of platform features. It also notes the importance of keeping pace with the fast-changing SNS landscape, as older platforms like MySpace quickly became obsolete.

For projects building decentralized systems to monitor health sentiment or outbreak indicators such as those implementing ActivityPub this study offers essential insight. It shows where current SNS-based approaches fall short and points to a future where open, federated, real-time systems could carry forward the work more ethically, equitably, and sustainably.

3.6 Digital Coping with Malady

Stephen A. Rains' *Cope with Malady Digitally* (2018) [6] examines how communicating technologies, including social mass media, blogs, and online communities, have changed the way individuals deal with malady. The book synthesizes a broad range of empirical research to illustrate how these technologies ease health information acquisition, enable patients to share experiences, and provide platforms for seeking and offering social support. Rains highlights that communicating technologies play a relevant role in reducing isolation, nurturing public engagement, and normalizing the experiences of illness. The book introduces the concept of digital coping by categorizing the affordances of communicating technologies such as visibility, accessibility, and control. These affordances allow users to make their experiences visible to others, connect across physical and temporal boundaries, and maintain control over the sharing of sensitive health-related information.

Rains discusses how patients employ digital tools like online communities and blogs to share updates or seek advice without the constraints of real-time interactions. The author also highlights the mental and social dimensions of malady, arguing that communicating technologies help patients navigate their experiences and relationships. By enabling interactions with others who have similar conditions, these platforms offer empathy, practical advice, and validation. Notwithstanding, Rains acknowledges potential challenges such as overexposure, negative interactions, anxiety triggered by others' struggles, and the risk of dysfunctional coping behaviors. In exploring patient-to-patient interactions, the book discusses the evolving nature of patient-provider relationships in the digital age, accenting the potential for these technologies to empower patients and facilitate a dynamic approach to health management.

3.7 Preventing the Next Pandemic

Mark Smolinski's article "Preventing the Next Pandemic" (2022) [7] examines why novel pathogens will continue to emerge and what strategies can halt them early. He argues that speed of detection is critical and that social innovation underpins effective surveillance.

Smolinski outlines five priority actions. First, engaging the public through participatory surveillance systems such as Flu Near You and Outbreaks Near Meenables communities to report symptoms in real time. This approach can reveal clusters before patients even seek clinical care. Second, a One Health framework integrates human, animal, and environmental data. Thailand's PODD system, for example, uses veterinary reports to spot livestock outbreaks and has prevented costly animal-to-human spillovers. Third, expanding epidemic intelligence via tools like EIOS and crowdsourced networks such as EpiCore accelerates early warning and verification of outbreak signals. Fourth, regional collaboration embodied by networks like CORDS builds trust across borders and allows shared resources to contain threats swiftly. Finally, standardized timeliness metrics (time to detect, verify, respond) help countries benchmark their readiness and identify persistent gaps.

Together, these measures form a cohesive model for pandemic prevention. Speed, community involvement, cross-sector collaboration, and measurable benchmarks work in concert to keep outbreaks local and avoid global spread.

3.8 Detecting Natural Disasters Using Social Media

Àgata Lapedriza and colleagues (2023) designed a deep-learning computer vision system that automatically detects natural disasters from images shared on social media platforms [8].

They constructed INCIDENTS1M, a dataset of 1,787,154 photographs labeled across 43 incident categories (e.g., avalanches, floods, earthquakes) and 49 place categories.Out of these, 977,088 images carried at least one positive incident label and 810,066 were negative; for place labels, 764,124 were positive and 1,023,030 were negative.

Using a multi-task learning paradigm with a convolutional neural network, the model learned to avoid false positives for example, distinguishing a benign fireplace from an active blaze. Evaluation on real-world social media streams (Flickr, Twitter) confirmed its ability to identify documented events such as the 2015 Nepal earthquake and Chilean tremors.

This tool offers humanitarian agencies a low-latency data source to understand disaster progression and aftermath. Future work could focus on quantifying incident severity over time and combining image analysis with accompanying text to improve classification accuracy.

3.9 Related Work

The integration of artificial intelligence (AI) in public health data analysis has garnered significant attention in recent years. This section reviews existing research and applications pertinent to this domain while highlighting the novelty of our approach.

3.9.1 Federated Learning in Healthcare

Federated Learning (FL) enables collaborative model training across decentralized data sources, preserving data privacya critical aspect in healthcare. Dayan et al. demonstrated the application of FL in predicting clinical outcomes for COVID-19 patients across multiple institutions, highlighting its effectiveness in sensitive data environments [9].

3.9.2 AI-Driven Public Health Monitoring

AI has been instrumental in enhancing public health monitoring. Yang et al. explored the role of AI in improving public health through predictive analytics and personalized interventions [10]. While their work emphasizes predictive models trained on historical healthcare data, our approach leverages real-time data streams from diverse social media platforms to capture evolving health discussions. This allows for immediate identification of public concerns, misinformation trends, and emerging outbreaks before they are reflected in official reports.

3.9.3 Social Network Analysis in Epidemiology

Analyzing social networks provides insights into disease transmission dynamics. Christakis and Fowler examined how social networks influence health behaviors and the spread of infectious diseases, suggesting that network analysis can inform public health interventions [11]. Unlike traditional approaches that analyze static datasets, our system processes dynamically changing user-generated content, continuously updating risk assessments and public health metrics. This real-time adaptability offers a more responsive and proactive approach to health monitoring.

3.9.4 BlueDot

BlueDot is a Toronto based infectious disease intelligence provider that uses AI, natural language processing, and epidemiological modelling to detect, monitor, and forecast outbreaks in near real time [12,13]. Founded in 2008 and commercialized in 2013, its Insights platform serves publichealth agencies, enterprises, and travel stakeholders worldwide.

Data Acquisition BlueDot aggregates signals from a rich ecosystem of open-source and proprietary feeds, including:

- **Official public health bulletins & reports**: WHO Disease Outbreak News, CDC MMWR, ECDC rapidrisk feeds, PHAC bulletins, OIE WAHIS notifications.
- **Expert moderated forums & syndromic trackers**: ProMED-mail RSS, HealthMap API, FluTrackers discussion forums.
- **Global news media**: Custom scrapers ingesting ∼300000 articles/day from 35000+ outlets (Reuters, AFP, BBC, Xinhua, local dailies) across 65 languages.
- **Peer reviewed literature & preprints**: PubMed RSS for emerging pathogen papers; medRxiv/bioRxiv preprints.
- **Airtravel network data**: IATA/OAG ticket itinerary feeds for $> 4\,000$ airports, driving importation/exportation risk scores .
- **Environmental & vectorsuitability models**: Climate metrics (NOAA/NASA), remotesensing of vector habitats.
- **Demographic & mobility proxies**: WorldPop population grids; aggregated, anonymized mobilephone mobility traces.

Analytical Framework The analytical framework employed involves sophisticated NLP pipelines that process ingested text from extensive open-source intelligence (OSINT) and traditional sources. These pipelines systematically identify critical event metadata, including pathogen type, specific geographic locations, and relevant dates. This structured identification is pivotal for creating high-resolution epidemiological datasets that include both traditional disease data-such as case numbers and hospitalization rates and supplementary data on environmental conditions, historical prevalence, and socio-economic factors that influence disease spread.

Machine learning classifiers then apply rigorous signal-to-noise separation techniques, en-

hancing the accuracy of outbreak detection by filtering out irrelevant or misleading data. This stage leverages predictive analytics, integrating supplementary information such as international travel patterns, local demographic details, and climate variability, which significantly enhances the precision of outbreak forecasting.

Subsequent to AI-driven alert generation, expert epidemiologists perform an essential validation role, meticulously assessing each alert against biological plausibility and contextual coherence. This dual-layer validation ensures that only alerts meeting stringent criteria of accuracy and reliability are escalated, thereby minimizing false positives and enhancing actionable intelligence for public health response and pharmaceutical strategic planning.

Client Access & Outputs Subscribers access BlueDots intelligence via:

- The Insights web portal-with interactive dashboards, personalized briefs, and Year-in-Review reports [14].
- Data-as-a-Service REST APIs exposing alerts, time-series, and probabilistic forecasts for direct programmatic integration [15].
- Configurable email alerts, webhooks, and BI connectors for real-time integration into enterprise workflows [15, 16].

Real World Use Cases BlueDots platform anticipated the Zika spread six months in advance and issued one of the first alerts for COVID-19 on December 31, 2019 nine days before the WHO notice [12, 16]. Its AI-powered biothreat intelligence has also guided supply-chain resilience planning and pharma R&D site selection [15, 16].

Relation to This Research This work parallels BlueDots approach by combining automated signal ingestion with AI-driven analysis and human validation. However, instead of proprietary feeds, it leverages decentralized ActivityPub networks to source community-generated health signals. A hybrid pipeline of topic modeling and transformer-based classification is used to surface emerging trends, which are then visualized in real-time dashboards under an open, reproducible framework. Beyond text posts, the system also processes extended media sources such as transcribed PeerTube videos, enriching the signal pool with diverse forms of user-generated content. By tapping into federated social platforms such as Mastodon, PeerTube, and Lemmy, this system broadens the scope of early detection to include grassroots-level reports that may be overlooked by centralized systems. The emphasis on transparency, modularity, and interoperability not only enhances trust but also facilitates independent verification and community-driven improvements.

3.9.5 Novelty of Our Approach

While previous studies have explored federated learning in institutional healthcare settings, AI in predictive public health, and traditional social network epidemiology, our work introduces a novel integration of these domains with a focus on real-time adaptability. By continuously analyzing user-generated unfiltered conversations, first hand reports and multimedia contributions as they happen from decentralized platforms, our system captures emerging health concerns. This contrasts with existing models that rely on static or retrospective data, making our approach uniquely suited for real-time crisis response and public health intelligence.

4 Research Methodology

This research adopts a design-oriented approach, focusing on the development and implementation of an ActivityPub-based web application to enhance public health data management. The primary objective is to explore how the integration of federated networks, machine learning models, and data visualization tools can improve the real-time collection, analysis, and presentation of statistical insights that support public health decision-making. The Web application, developed using Python and the Django framework, leverages ActivityPub APIs to collect posts and media data from decentralized platforms such as Mastodon, PeerTube, Lemmy, and Misskey. To supplement the data and test the processing pipeline, posts are also fetched from Bluesky, a non-federated platform that provides access to larger volumes of data. The project incorporates a robust data processing pipeline designed to efficiently manage posts and employ machine learning models to extract health-related insights. Furthermore, Grafana is integrated to visualize these insights in real time, offering a concise and dynamic platform for tracking and analyzing public health data as it evolves.

4.1 Materials and Tools

This research utilized a variety of programming languages, frameworks, libraries, and tools to design, implement, and analyze the web application and its functionality. Below is a categorized breakdown of these materials and tools:

4.1.1 Programming Languages

- **Python 3.12.5:** Used for web application development, data processing, and machine learning pipelines. [17]

4.1.2 Frameworks

- **Django:** The primary web framework for server-side logic, database interaction, and API communication. [18]
- **Celery:** Used for task parallelization and asynchronous job execution. [19]
- **Django-cron:** Employed for scheduling periodic data-fetching and processing tasks. [20]
- **Redis:** Used as a message broker for Celery, enabling efficient task queue management and caching. [21]

4.1.3 Libraries and Tools

4.1.3.1 Machine Learning and NLP

- **Transformers:** A library from Hugging Face, used for machine learning tasks such as text classification and summarization.[22]
- **Llama Model:** Specifically, the Llama 3.2 3B and 1B instruct model for health-related data analysis. [23]
- **Whisper:** Used for transcribing audio content from PeerTube videos to text for analysis. [24]
- **NLTK (Natural Language Toolkit):** Utilized for text preprocessing, including stopword removal, tokenization, and lemmatization. [25]
- **spaCy:** Used for natural language processing tasks, such as named entity recognition (NER). [26]
- **Gensim:** Used for LDA (Latent Dirichlet Allocation) topic modeling and Phrases for n-gram generation. [27]

4.1.3.2 Data Collection and Processing

- **BeautifulSoup:** Used for web scraping and HTML content parsing. [28]
- **Geopy:** Integrated for geolocation services to extract locations from text and assign coordinates. [29]
- **YouTube-DL:** Utilized for downloading video content from PeerTube for transcription and analysis. [30]

4.1.3.3 Database and Storage

- **PostgreSQL:** A relational database system for storing and managing all collected data and results. [31]
- **Django ORM:** Used for interacting with the PostgreSQL database. [18]

4.1.3.4 Visualization

- **Grafana:** Integrated as the visualization platform for real-time dashboards displaying health-related insights and trends. [32]

4.1.3.5 Task and Dependency Management

- **Docker:** Used for containerizing the web application and its services, ensuring consistent development and deployment environments. [33]
- **Hugging Face Hub:** Accessed for downloading pre-trained models and caching dependencies. [34]

4.1.3.6 Other Utilities

- **SSL:** Managed secure HTTP requests when interacting with APIs.
- **Logging:** Integrated for error tracking and debugging during development and deployment.
- **Unicodedata:** Used for normalizing text data.
- **Subprocess:** Utilized for handling system-level operations like video format conversions.
- **shutil:** Employed for file operations, such as moving or deleting files.
- **Counter:** Used for counting and analyzing word frequencies in textual data.

4.2 Data Collection

The data collection process involved fetching posts and media data from multiple platforms to ensure comprehensive coverage of public health discussions. The platforms included both federated and non-federated systems:

4.2.1 Federated Platforms (ActivityPub-based)

- **Mastodon:** Posts were retrieved using hashtags and keywords relevant to health topics.
- **PeerTube:** Videos were downloaded and transcribed for text analysis, focusing on health-related discussions and trends.
- **Lemmy:** Posts and comments were fetched to gather broader community discussions on health issues.
- **Misskey:** Content was collected to provide additional insights from users in the federated network.

4.2.2 Non-Federated Platform

- **Bluesky:** Bluesky does not follow the ActivityPub protocol but was included in the data collection process to test the processing pipeline and supplement the dataset with a higher volume of posts. This platform provided additional insights into public health-related conversations outside the federated ecosystem.

4.2.3 Supplementary Dataset from IEEE DataPort

- **GeoCoV19 Dataset:** To further evaluate and test the data processing pipeline, the GeoCoV19 dataset [35] was used. This dataset contains hundreds of millions of multilingual COVID-19-related uploads, many of which include location information. It aggregates content from various platforms, including federated sources such as blogs and Facebook posts. Since real-time data collection from ActivityPub-based platforms is limited in volume and scope, this external dataset served as a valuable supplement for training, testing, and benchmarking the system.

4.2.4 Challenges Addressed in Data Collection

The collection process addressed several challenges:

1. **API Limitations and Rate Limits:** Platforms impose restrictions on the number of API calls within a specific timeframe. To address this, efficient rate-limiting strategies, such as batching requests and implementing retry mechanisms, were applied.
2. **Authentication and Access Restrictions:** Some platforms required OAuth authentication or API keys for accessing data. This included handling private instances of federated networks where additional permissions or credentials were necessary.
3. **Data Duplication Across Instances:** In federated platforms like Mastodon, posts often appear across multiple instances. A deduplication mechanism was implemented using unique post IDs or hashes to avoid redundant data storage.
4. **Inconsistent APIs Across Platforms:** Different platforms exposed varying API structures and data formats, requiring custom parsers and adapters for consistent data ingestion into the system.
5. **Handling Non-Standard Data (Bluesky):** Since Bluesky does not follow the ActivityPub protocol, additional effort was required to create separate logic for fetching and processing its data, ensuring compatibility with the existing pipeline.
6. **Language and Text Variability:** Posts were often written in multiple languages or included informal text with abbreviations, slang, and emojis. Natural language processing (NLP) techniques were employed to normalize and process this variability.
7. **Geotagging and Location Extraction:** Many posts lacked explicit geotags, requiring text-based location extraction using geolocation tools such as Geopy to associate content with geographic regions.
8. **Incomplete or Noisy Data:** Posts frequently contained incomplete information, spam, or irrelevant content. Preprocessing steps were applied to filter and clean the data for meaningful analysis.
9. **Video and Media Content:** Platforms like PeerTube hosted video content requiring transcription (using Whisper) before analysis. The process included handling video downloads, format conversions, and transcription errors.
10. **Health-Related Relevance Filtering:** A filtering mechanism was applied to identify posts containing health-related keywords, hashtags, or context. Machine learning models, including Llama, were used to determine relevance.
11. **Time Zone and Timestamp Normalization:** Posts collected from different platforms used varied timestamp formats and time zones. A normalization process ensured consistency for time-based analyses.
12. **Ethical Considerations:** To ensure compliance with ethical guidelines, sensitive or personal information was anonymized or excluded from the dataset during the collection process.
13. **Scalability:** The system needed to handle increasing amounts of data as more instances and platforms were added. Efficient database storage and caching strategies were implemented to manage scalability.

14. **Real-Time Data Retrieval:** For platforms requiring real-time updates, such as Mastodon and Misskey, a streaming-like approach was incorporated to fetch data continuously while balancing system performance.

4.3 Data Analysis

The data analysis pipeline was designed to process, analyze, and extract actionable insights from the collected data. The pipeline consisted of the following steps:

4.3.1 Natural Language Processing (NLP)

Textual data underwent pre–processing using tools such as **NLTK** and **spaCy** to ensure it was prepared for analysis. The pre–processing steps included:

- **Removing stopwords:** To focus on meaningful words.
- **Tokenization, lemmatization, and stemming:** To normalize text and prepare it for analysis.
- **Cleaning data:** To remove noise, such as punctuation, HTML tags, and emojis.

These preprocessing steps ensured a clean and consistent dataset for subsequent machine learning tasks.

4.3.2 Machine Learning Analysis

Relevance Assessment with the Llama Model was central to the data analysis process. It was employed to:

- Assess the relevance of posts for health-related analysis by identifying key terms, context, and overall alignment with public health themes.
- Summarize posts to extract concise and actionable information, particularly for long-form content or transcriptions from PeerTube videos.
- Extract specific details such as health symptoms, locations, and sentiment, enabling further categorization and contextual understanding.
- Validate outputs from other analysis methods, such as LDA, ensuring high accuracy and consistency.

Health Topic Classification

- **LDA (Latent Dirichlet Allocation):** Topic modeling was used to classify posts into general health-related themes based on word distributions and patterns. LDA provided insights into broader categorization, such as mental health discussions, disease outbreaks, or lifestyle-related advice.

Sentiment Analysis The sentiment of posts was analyzed using the **Llama model** to identify emotional tones, such as anxiety, positivity, or negativity, related to health topics. This helped gauge public perception and sentiment trends over time.

4.3.3 Visualization

The processed data was visualized using **Grafana**, which was configured to connect with the **PostgreSQL** database. The following features were implemented:

- **Dynamic Dashboards:** Real-time dashboards allowed public health officials to dynamically track trends, including health topic spikes, symptom frequency, and potential hot spots.
- **Health Insights:** Key metrics, such as the most mentioned symptoms, geographic patterns, and sentiment trends, were visualized to aid in quick decision-making.
- **Trend Monitoring:** Data was aggregated and displayed over specific time frames, enabling the identification of emerging trends, such as sudden increases in discussions around specific diseases or health concerns.

5 System Overview and Architecture

This chapter provides a comprehensive overview of the system designed and implemented for discovering, analyzing, and summarizing health-related content across multiple federated platforms, including Mastodon, Lemmy, Misskey, PeerTube, and Bluesky. We leverage a combination of Docker Compose for containerization, Django for the web framework, Celery for task scheduling and background processing, Redis as a message broker, and PostgreSQL for data storage. Additionally, we integrate a local LLaMA model running locally on GPU for advanced text classification and analysis.

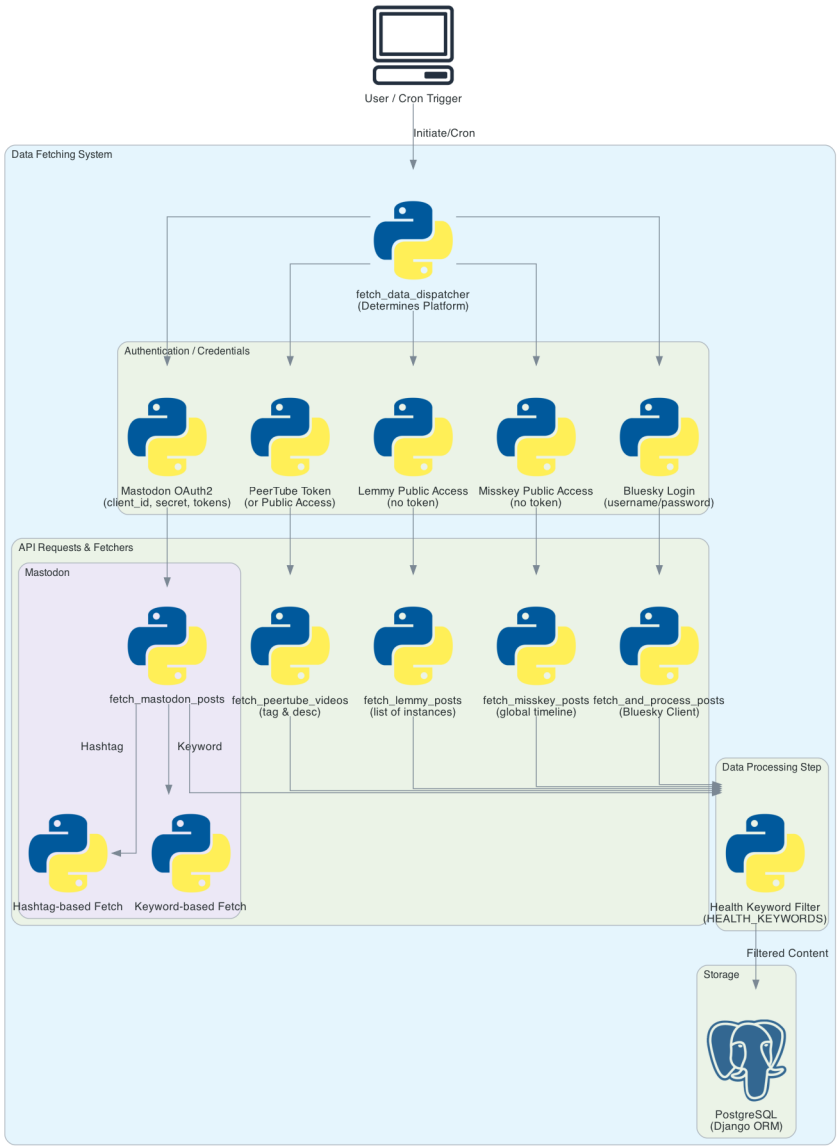
5.1 Data Flow

At a high level, the system aims to:

1. **Fetch** data from a variety of federated platforms (Mastodon, Lemmy, Misskey, PeerTube, Bluesky) using platform-specific APIs or public endpoints.
2. **Store** both raw and partially processed content in a PostgreSQL database via Djangos ORM.
3. **Process** content in background tasks, including:
    - Summarizing lengthy text.
    - Validating whether the post is relevant and suitable for further analysis based on content quality and health-related context.
    - Running classification or advanced analysis on the text (e.g., health categories, sentiment, etc.).
    - Extracting additional metadata such as locations, topics, or LDA-based keywords.
4. **Visualize** final results in a Grafana dashboard that pulls data directly from the database.

Federated Platforms Data Fetching Flow

Because we ingest content from multiple platforms, each with its own API or endpoint, we implement a specialized Fetching Flow. Figure 5.1 shows how the system collects data from Mastodon, Lemmy, Misskey, PeerTube, and Bluesky, stores preliminary results, and performs basic filtering:



Federated Platforms Data Fetching and Processing (Multi-Language)

Figure 5.1: Federated Platforms Data Fetching Flow.

Data from the above flow is then inserted into the database, at which point further processing tasks can be triggered. This approach ensures each platform is treated as a black box whose data ends up uniformly in the system.

5.2 Docker-Based Deployment

To ensure a consistent and maintainable environment, we containerize all services with **Docker Compose**. Listing 5.1 shows an excerpt of the `docker-compose.yaml` configuration that defines the following services:

- `db` (PostgreSQL): Holds schema, tables, and indexes for all data.
- `web` (Django): Runs the main application logic, including endpoints for admin and Celery configuration.
- `redis`: In-memory data store for queuing tasks and caching results.
- `celery` and `celery-beat`: Worker for tasks and a scheduler for periodic tasks, respectively.
- `grafana`: Visualization layer, pointed to the same `db` for data queries.

Listing 5.1: Excerpt from the Docker Compose File

```
services:
  db:
    image: postgres:17
    # ...
  web:
    build: .
    # ...
  redis:
    image: redis:7.2.5-alpine3.19
    # ...
  celery:
    build: .
    # ...
  celery-beat:
    build: .
    # ...
  grafana:
    image: grafana/grafana:latest
    # ...
volumes:
  postgres_data:
  grafana_data:
  llama-model:
```

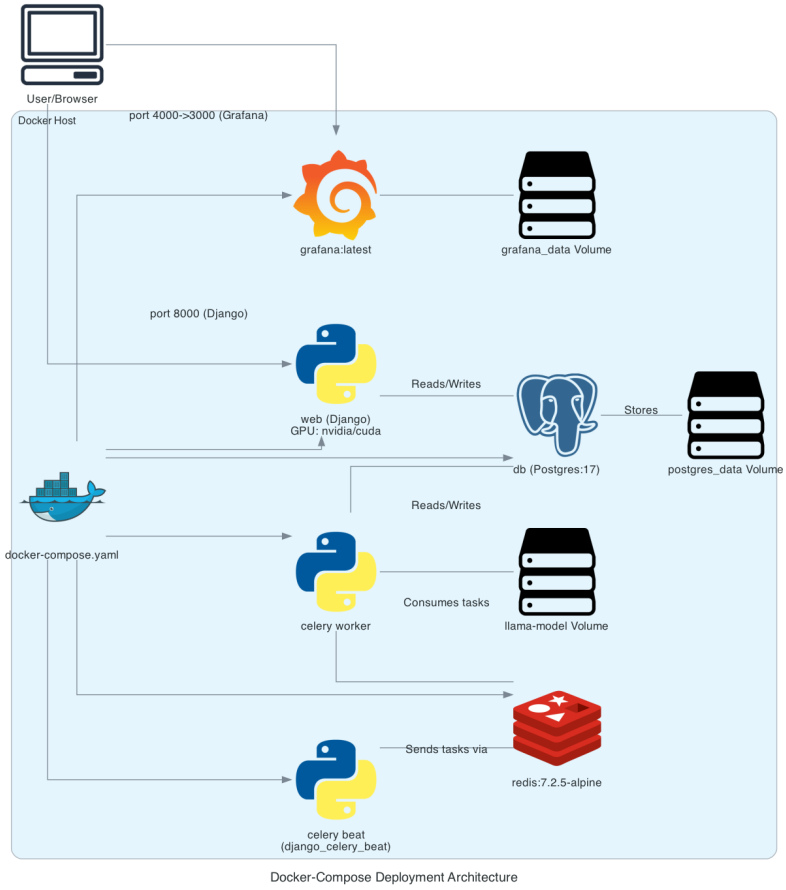


Figure 5.2: Docker Compose Deployment Diagram Showing Containers and Volumes.

Figure 5.2 visually explains how these containers interact internally via the Docker network. The `llama-model` volume is particularly important if you mount your local LLaMA model files for the GPU-enabled container.

5.3 Celery Architecture and Task Flow

In this architecture, **Celery** orchestrates a wide array of tasks related to:

- *Fetching content*: from Mastodon, Lemmy, Misskey, PeerTube, and Bluesky.
- *Scheduling*: controlled by `celery-beat`, which uses `django-celery-beat` for reading from the database.

A user or admin can either:

- *Trigger a job manually*: e.g., Fetch from Mastodon now.
- *Rely on periodic tasks*: e.g., Every 6 hours, fetch from Lemmy.

All tasks are queued in `redis` (the broker). Celery workers pick them up, interact with external APIs, and then write or update records in PostgreSQL.

Figure 5.3 outlines how Celery workers collaborate with Redis and Django to schedule or execute tasks. This makes the system scalable and fault-tolerant.

5.4 LLaMA Pipeline and Processing Diagram

5.4.1 Overview

One of the critical aspects of our system is the ability to perform advanced text analysis, classification, and summarization using a **local LLaMA model** on GPU. This allows deeper health-related extraction: disease categories, user sentiment, potential misinformation indicators, or outbreak detection signals.

5.4.2 Processing Diagram

After data is fetched, a separate Processing pipeline refines and enriches the posts or video transcriptions. Figure 5.4 demonstrates how tasks flow, from pre-processing or summarization decisions to LLaMA-based classification and advanced NLP steps like location extraction and LDA topic modeling.

The main steps are:

1. *Preprocess Text.* The Celery worker cleans the raw text (HTML removal, tokenization).
2. *Decision: Summarize?* If the text exceeds a certain length or meets specific conditions, it is routed to a **summarizer** service or library.
3. *Call LLaMA.* Summarized or raw text is sent to the GPU container that runs the LLaMA model, returning classification JSON with health categories, sentiments, or specialized tags (e.g., flu outbreak, mentalhealth, etc.).

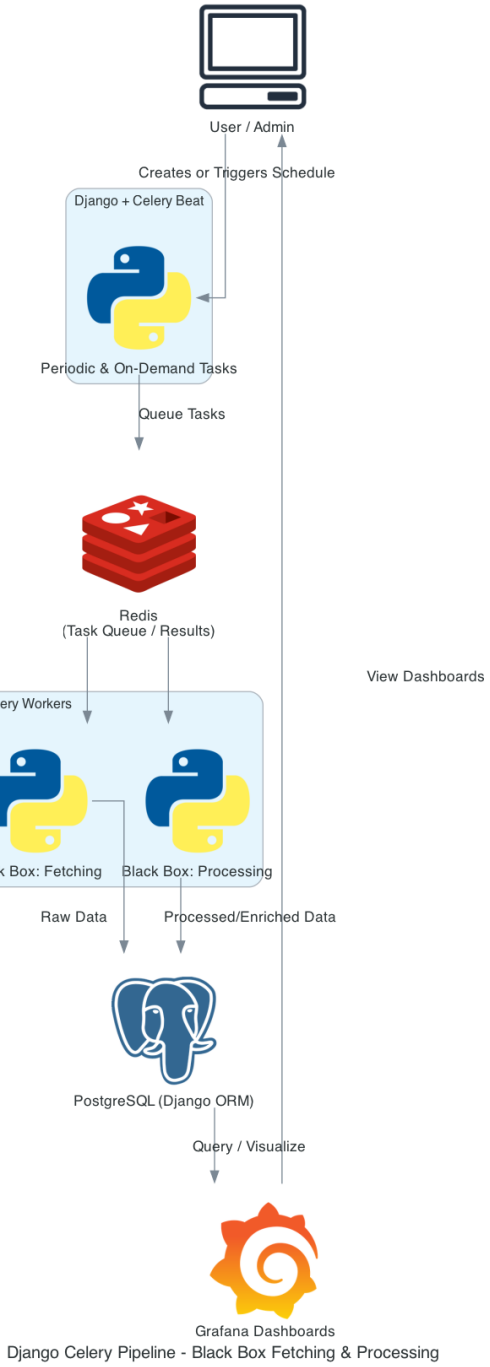


Figure 5.3: Celery Architecture for Background Task Execution and Coordination.

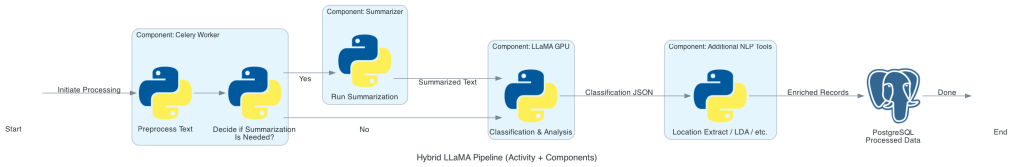


Figure 5.4: Data Processing Diagram Featuring LLaMA Integration and NLP Steps.

4. *Additional NLP.* Additional steps like location extraction or LDA-based topic modeling further enrich the data.
5. *Store in DB.* The final annotated or classified record is persisted in the `PostgreSQL` database.

5.5 Workflow: End-to-End Overview

To consolidate all phases (fetching, processing, storing, and final dashboard reporting), we present an **Workflow Diagram** in Figure 5.5. This single view demonstrates how user triggers or scheduled tasks lead to data ingestion, advanced text analysis, and ultimately feed the Grafana dashboards:

Users can see real-time or historical statistics via Grafana, accessing the same PostgreSQL database where the Celery tasks have stored or updated records with classification results, summaries, and location details.

5.6 Grafana Visualization

Grafana is configured to connect directly to the PostgreSQL instance, enabling real-time and historical analysis of health-related content. The dashboards built in Grafana aggregate and visualize a wide range of metrics, including:

1. **Urgency Tracking for Health Emergencies:** Tracks the number of urgent posts over time, segmented by location. This visualization helps identify areas with high volumes of urgent health-related posts, thereby guiding emergency response efforts and resource allocation.
2. **Lifestyle Impact Analysis:** Displays the reported impacts on lifestyle due to health issues over time. This chart provides insights into how health incidents affect daily life by tracking mentions of lifestyle changes or disruptions.
3. **Preventative Measures and Health Advice Impact:** Analyzes the frequency of posts sharing preventative measures or health advice (e.g., vaccinations, hygiene practices, social distancing). This dashboard helps assess the effectiveness and reach of public health advisories.

4. **Symptom Frequency and Outbreak Detection:** Monitors the occurrence of specific health symptoms, segmented by time and location. This visualization acts as an early warning system by highlighting trends that may indicate potential outbreaks.
5. **Health Sentiment Tracking:** Captures daily sentiment metrics from health-related posts. The dashboard categorizes sentiments (positive, neutral, negative, hopeful, anxious) to provide an overall gauge of public mood towards health issues.
6. **Covid Hashtag Spikes:** Shows daily counts of posts tagged with #covid, indicating spikes in COVID-19-related discussions. This visualization reflects changes in public concern and awareness over time.
7. **LDA Topics with Location:** Maps topics generated from LDA (Latent Dirichlet Allocation) based on post locations. This dashboard helps identify geographic trends in health discussions by topic.
8. **Top Hashtags:** Displays the most frequently used hashtags across posts, capturing trending health topics and keywords in real-time.
9. **Interaction with Health Posts:** Summarizes engagement metrics (replies, reblogs, favorites) for health-related posts. This visualization measures the level of public interaction and interest in various health topics.
10. **Topics and Post Counts (LDA):** Visualizes the distribution of posts by topics generated through LDA, aiding in the analysis of which health topics are most frequently discussed.

Since Grafana queries the database directly, it minimizes additional overhead on the Django application. Users can log into the Grafana UI (exposed by the `grafana` container) to interact with these dashboards, customize visualizations, and gain actionable insights into public health trends.

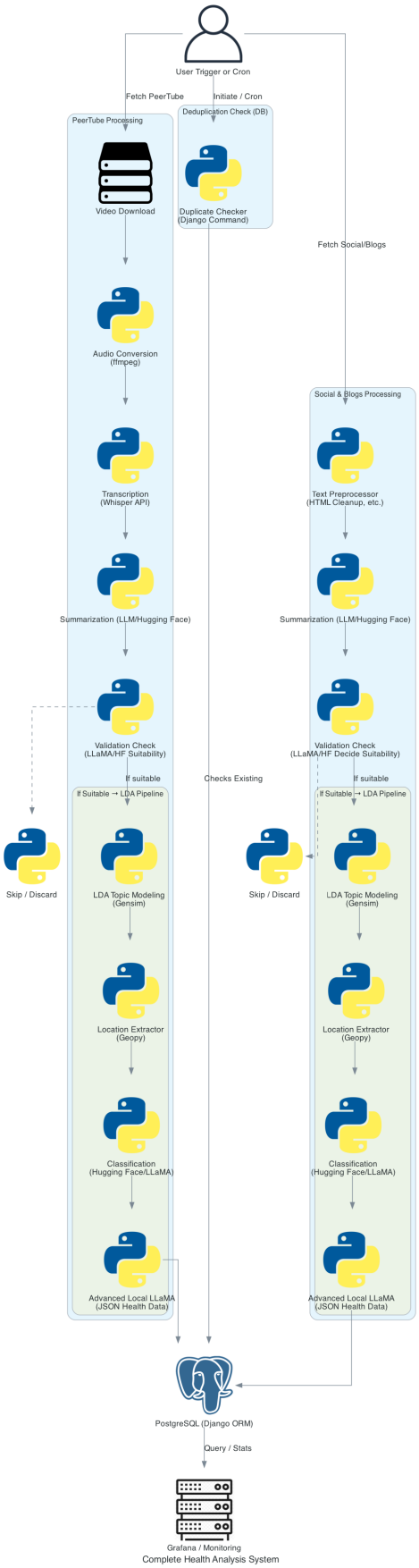


Figure 5.5: Ultimate Workflow: Fetching, Processing, and Displaying with Grafana.

5.7 Code Snippets

This section presents key code snippets used in the implementation of the system. These code snippets illustrate various functionalities such as data fetching, processing, classification, and visualization.

5.7.1 Fetching Data from Mastodon

Mastodon posts are fetched using its public API. The following code snippet demonstrates how data is collected from multiple Mastodon instances and stored in the database:

Listing 5.2: Fetching posts from Mastodon using API

```
import requests
from core.models import MastodonPost

MASTODON_INSTANCES = [
    'https://mastodon.social',
    'https://mastodon.online',
    'https://fosstodon.org'
]

def fetch_mastodon_posts(instance_url, hashtag, access_token):
    headers = {'Authorization': f'Bearer {access_token}'}
    params = {'limit': 40}
    url = f'{instance_url}/api/v1/timelines/tag/{hashtag}'

    response = requests.get(url, headers=headers, params=params)
    if response.status_code == 200:
        posts = response.json()
        for post in posts:
            MastodonPost.objects.create(
                post_id=post['id'],
                content=post['content'],
                published=post['created_at'],
                account_username=post['account']['username'],
                url=post['url']
            )
```

5.7.2 Transcribing PeerTube Videos

PeerTube videos are processed using Whisper AI for transcription. The following code snippet demonstrates how videos are downloaded and transcribed:

Listing 5.3: Transcribing PeerTube videos with Whisper AI

```
import youtube_dl
import whisper

```

```
def transcribe_video(video_url):
    options = {'format': 'best', 'outtmpl': 'video.mp4'}
    with youtube_dl.YoutubeDL(options) as ydl:
        ydl.download([video_url])

    model = whisper.load_model("base")
    transcript = model.transcribe("video.mp4")
    return transcript["text"]
```

5.7.3 Validating Posts Using AI

To determine whether a post contains relevant health information, a LLaMA model is used for classification:

Listing 5.4: Validating posts for health-related content

```
import torch
from transformers import pipeline

def validate_post_content(post_content):
    model_id = "meta-llama/Llama-3.2-3B-Instruct"
    device = "cuda" if torch.cuda.is_available() else "cpu"
    generator = pipeline("text-generation", model=model_id, device=device)

    prompt = f"Determine if the following post is relevant to health
        topics:\n{post_content}\nAnswer 'yes' or 'no'."
    response = generator(prompt)[0]['generated_text']
    return response.lower().strip() == "yes"
```

5.7.4 LDA-Based Topic Modeling

To classify posts into various health-related topics, Latent Dirichlet Allocation (LDA) is applied:

Listing 5.5: LDA topic modeling for health-related posts

```
from gensim import corpora, models

def train_lda(posts):
    processed_posts = [post.content.split() for post in posts]
    dictionary = corpora.Dictionary(processed_posts)
    corpus = [dictionary.doc2bow(text) for text in processed_posts]

    lda_model = models.LdaModel(corpus, num_topics=10, id2word=dictionary,
        passes=15)
    return lda_model
```

5.7.5 Detailed Timeseries Panel Configuration

The following JSON excerpt defines the Health Sentiment Tracking panel, which visualizes weekly counts of positive, neutral, and negative posts.

Listing 5.6: Grafana panel JSON for Health Sentiment Tracking

```
{
  "id": 8,
  "type": "timeseries",
  "title": "Health Sentiment Tracking",
  "gridPos": {"h": 8,"w": 12,"x": 0,"y": 40},
  "fieldConfig": {
    "defaults": {
      "color": {"mode": "palette-classic" },
      "custom": {
        "axisBorderShow": false,
        "axisCenteredZero": false,
        "axisColorMode": "text",
        "axisLabel": "Number of Posts",
        "axisPlacement": "auto",
        "drawStyle": "line",
        "fillOpacity": 62,
        "lineInterpolation": "linear",
        "lineWidth": 1,
        "pointSize": 5,
        "thresholdsStyle": {"mode": "area" }
      },
      "thresholds": {
        "mode": "absolute",
        "steps": [
          { "color": "green",  "value": null },
          { "color": "dark-red", "value": 5000}
        ]
      }
    },
    "overrides": []
  },
  "options": {
    "legend": { "showLegend": true, "placement": "bottom" },
    "tooltip": {"mode": "single",   "sort": "none"    }
  },
  "targets": [
    {
      "datasource": {"type": "grafana-postgresql-datasource", "uid":
          "edzvx3q2em9kwa" },
      "format": "table",
      "rawQuery": true,
      "rawSql": "SELECT\n"
               "  DATE_TRUNC('week', published) AS week,\n"
```

```
                "  COUNT(*) FILTER (WHERE sentiment ILIKE 'Positive') AS
                    positive,\n"
                "  COUNT(*) FILTER (WHERE sentiment ILIKE 'Neutral') AS
                    neutral,\n"
                "  COUNT(*) FILTER (WHERE sentiment ILIKE 'Negative') AS
                    negative\n"
                "FROM (\n"
                "  SELECT sentiment, published FROM core_mastodonpost\n"
                "  UNION ALL SELECT sentiment, published FROM core_blueskypost\n"
                "  UNION ALL SELECT sentiment, published FROM
                    core_peertubevideo\n"
                "  UNION ALL SELECT sentiment, published FROM core_misskeypost\n"
                "  UNION ALL SELECT sentiment, published FROM core_lemmypost\n"
                "  UNION ALL SELECT sentiment, published FROM
                    core_external_blogs\n"
                ") AS combined_posts\n"
                "WHERE sentiment IS NOT NULL\n"
                "GROUP BY week\n"
                "ORDER BY week;\n",
      "refId": "A"
    }
  ]
}
```

These code snippets represent essential components of the system, ensuring efficient data fetching, processing, and classification while integrating machine learning models for enhanced analysis.

5.8 Conclusion

In this chapter, we presented the complete **System Overview and Architecture**, outlining:

1. **Federated Platforms Data Fetching** flow (Figure 5.1), showing how different APIs funnel content into our system.
2. **Docker Compose** setup (Figure 5.2) that coordinates PostgreSQL, Redis, Django, Celery workers, and Grafana.
3. **Celery-based approach** (Figure 5.3) for scheduling and executing tasks in the background.
4. A dedicated **Processing Diagram** featuring LLaMA (Figure 5.4), illustrating text classification, summarization, location extraction, and more.
5. The **Workflow** (Figure 5.5) tying everything together, from ingestion to final visualization.

These design choices aim to ensure scalability (through Docker and Celery workers), reproducibility (Docker images and pinned requirements), and powerful analysis (local GPU-based LLaMA). The next chapter will present experimental results, benchmarks for the LLaMA inference performance, and a qualitative discussion of the systems real-world feasibility.

6 Results

This chapter presents the analytical findings derived from the federated public health monitoring system. Using data collected from various ActivityPub-based platforms (Mastodon, PeerTube, Lemmy, and Misskey) and Dataset from IEEE Dataport including external platforms

6.1 Topic Modeling with LDA

Figure 6.1 shows the top LDA topics identified in the dataset, with the most dominant topic containing keywords like *virus, outbreak, epidemic, infection, disease* occurring over 29,000 times.

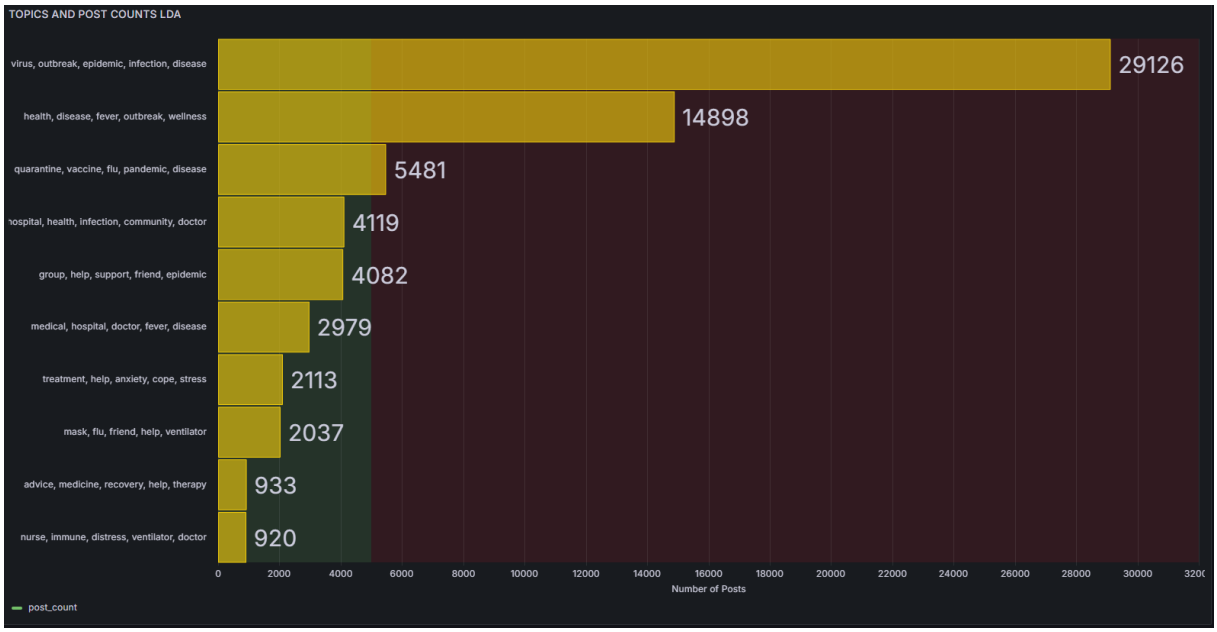


Figure 6.1: Topics and Post Counts Using LDA Topic Modeling

6.2 Emotion Distribution

Figure 6.2 summarizes emotion categories from user posts. *Anxious* and *hopeful* dominate the distribution, showing contrasting public sentiments during health crises.

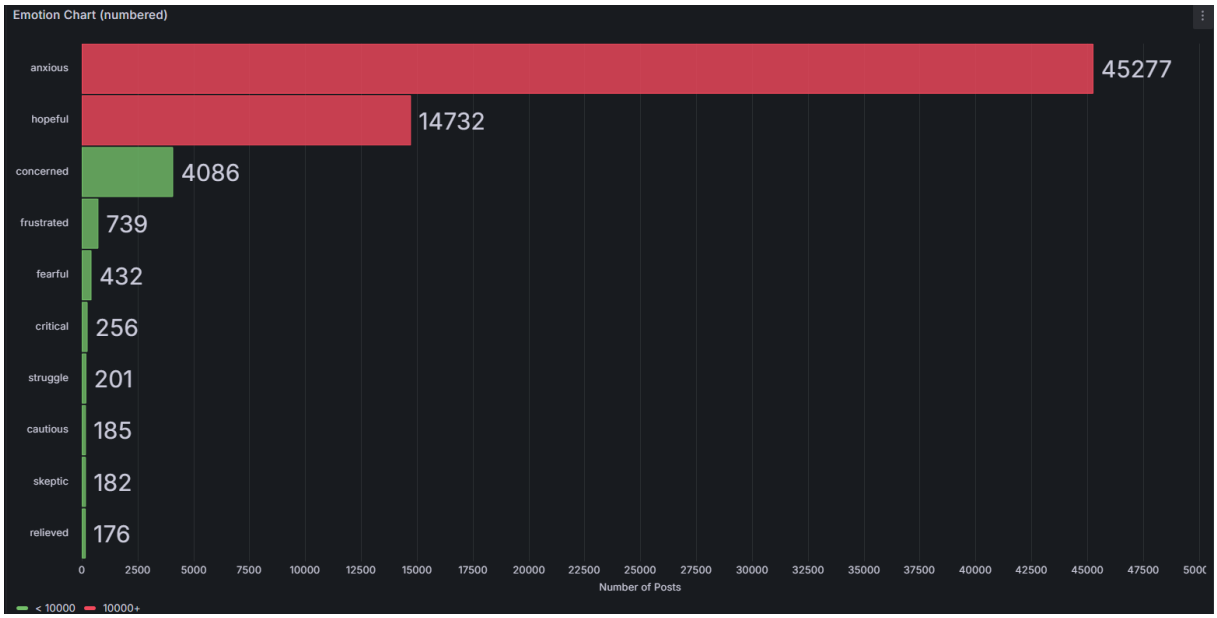


Figure 6.2: Emotion Chart Categorized by Post Count

6.3 Urgency Tracking

Urgency trends over time, depicted in Figure 6.3, indicate a spike in urgent health-related posts between late January and early February.

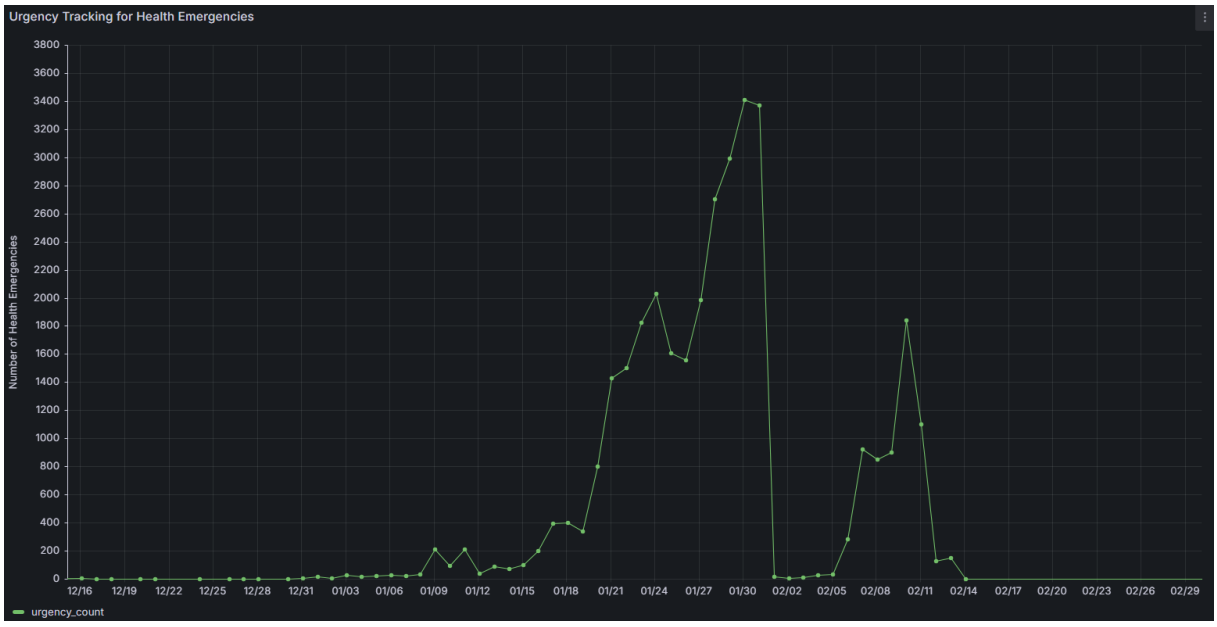


Figure 6.3: Urgency Tracking for Health Emergencies Over Time

6.4 Lifestyle Impact

The lifestyle disruptions caused by health events are visualized in Figure 6.4, where *disruption* emerges as the most affected category.

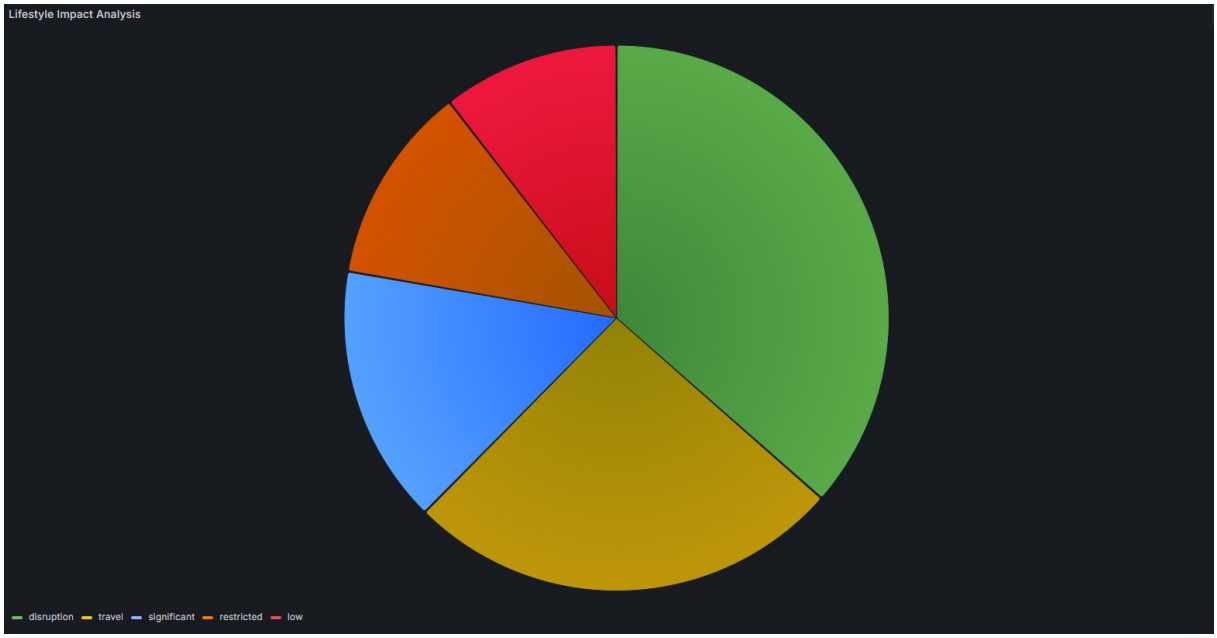


Figure 6.4: Lifestyle Impact Analysis

6.5 Preventative Measures and Health Advice

Figure 6.5 displays the volume of posts advocating various preventative actions, with *vaccination* and *masking* dominating the discourse.

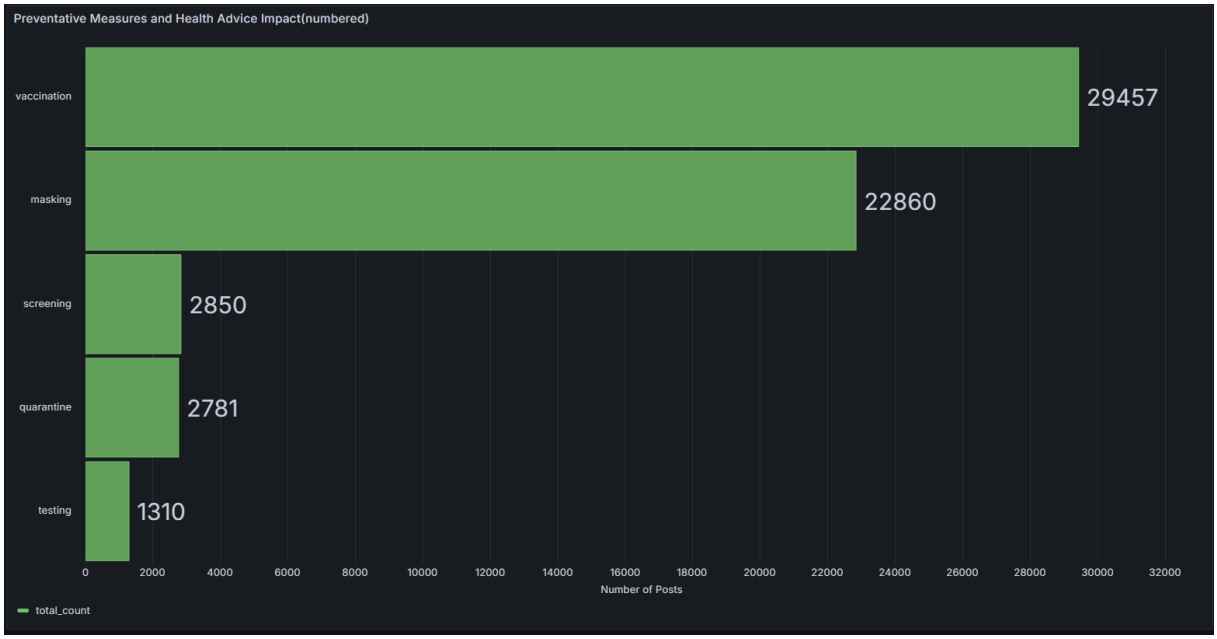


Figure 6.5: Preventative Measures and Health Advice Impact

6.6 Symptom Frequency and Outbreak Detection

As shown in Figure 6.6, symptom reporting experienced a sharp increase during late January, surpassing the outbreak detection threshold.

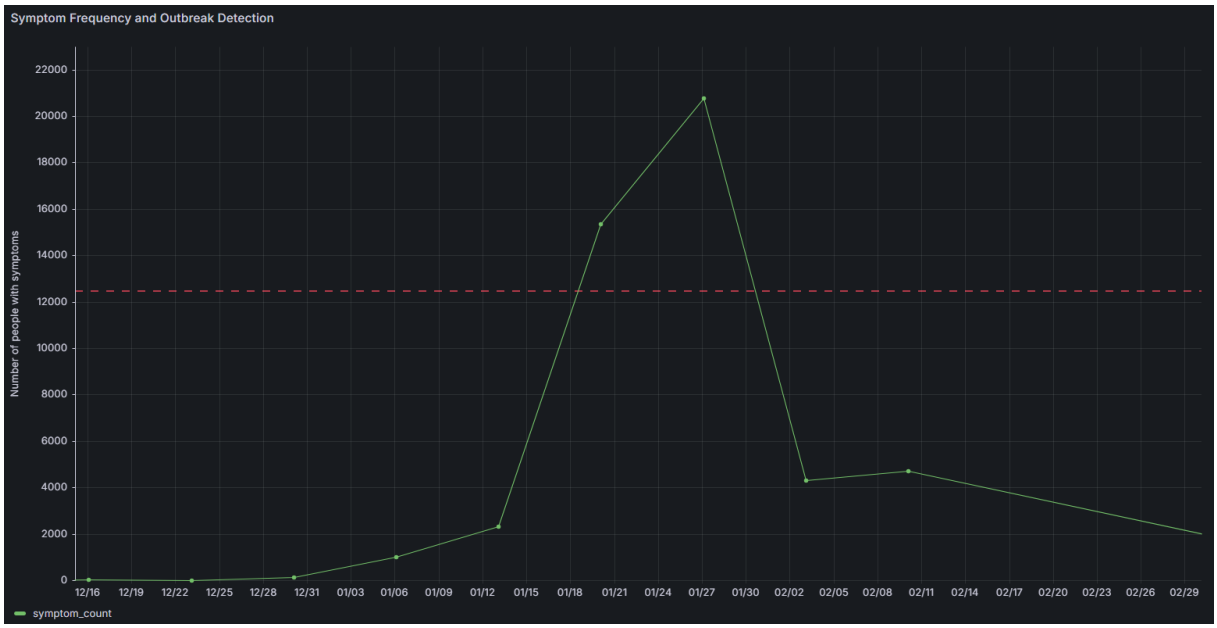


Figure 6.6: Symptom Frequency and Outbreak Detection

6.7 Health Sentiment Tracking

Health-related sentiment trends (positive, neutral, negative) over time are shown in Figure 6.7, reflecting heightened public response in late January.

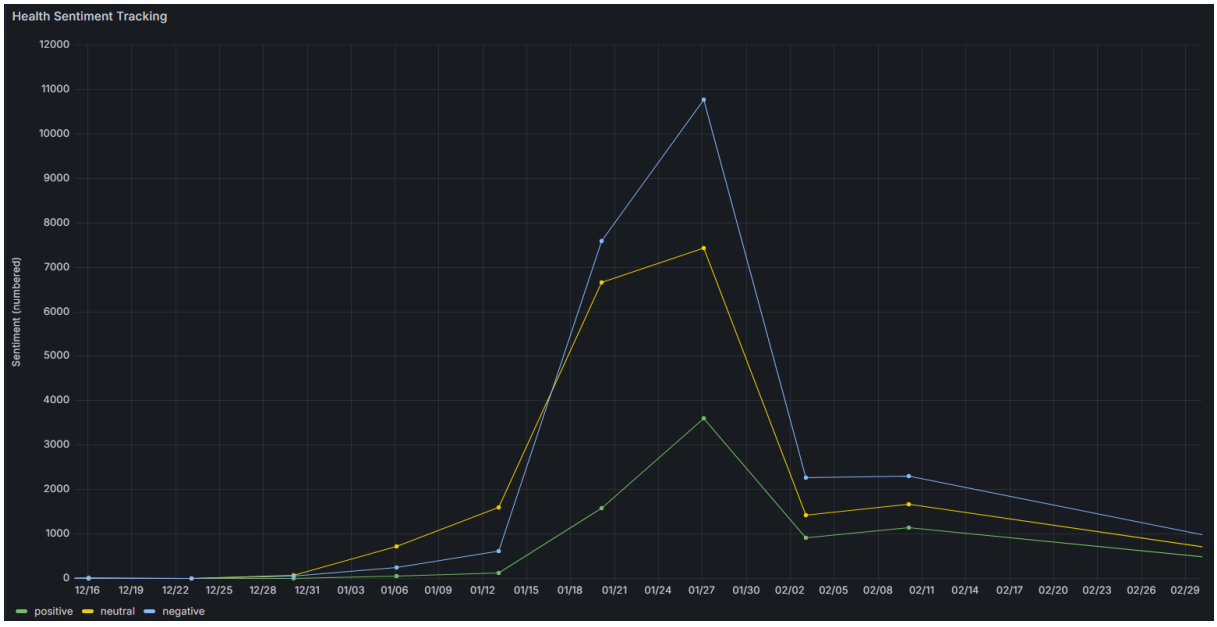


Figure 6.7: Health Sentiment Tracking

6.8 COVID-Related Hashtag Spikes

A separate spike in COVID-tagged content is visible during late 2022 and early 2023, as shown in Figure 6.8.

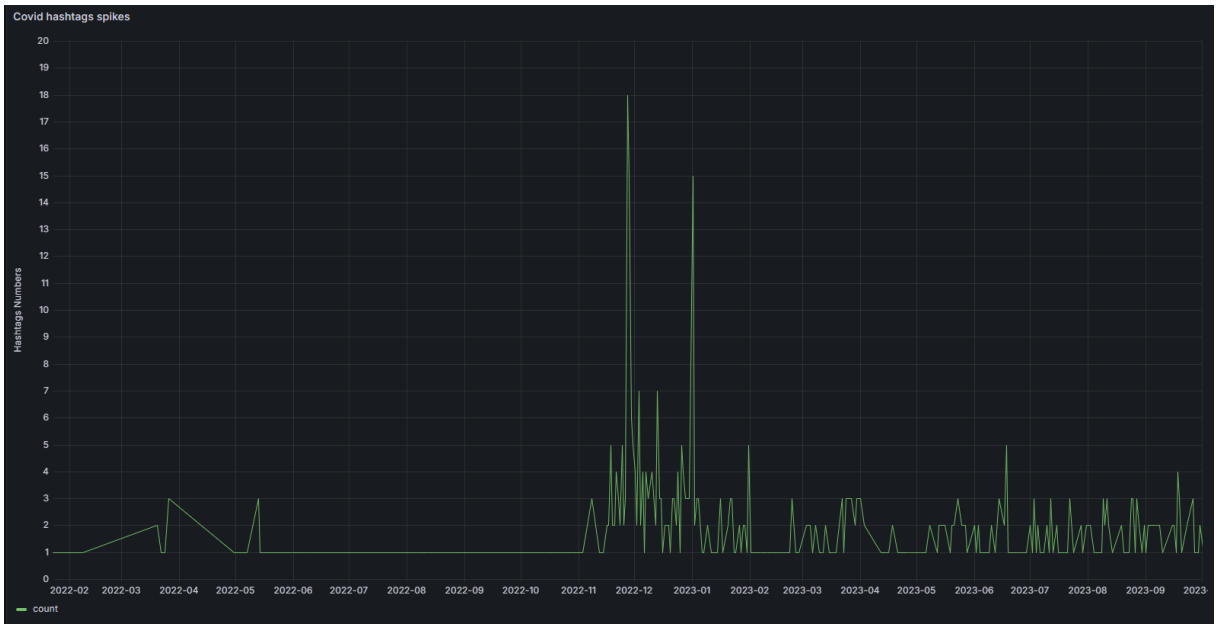


Figure 6.8: COVID-Related Hashtag Spikes)

6.9 Topic Distribution by Location

Figure 6.9 links LDA topics with geographic locations, identifying hotspots of topic engagement and post volume per region.

LDA TOPIC PER WITH LOCATION

latitude	longitude	topic_name	post_count ↓
35.0	105	virus, outbreak, epidemic, infection, disease	4736
30.6	114	virus, outbreak, epidemic, infection, disease	2186
35.0	105	health, disease, fever, outbreak, wellness	2155
30.6	114	health, disease, fever, outbreak, wellness	952
35.0	105	quarantine, vaccine, flu, pandemic, disease	677
35.0	105	hospital, health, infection, community, doctor	640
39.8	-100	virus, outbreak, epidemic, infection, disease	611
40.2	116	virus, outbreak, epidemic, infection, disease	605
35.0	105	group, help, support, friend, epidemic	603
35.0	105	medical, hospital, doctor, fever, disease	427
22.4	114	virus, outbreak, epidemic, infection, disease	318
30.6	114	quarantine, vaccine, flu, pandemic, disease	307

Figure 6.9: LDA Topics and Post Count by Location

6.10 Top Health Hashtags

As visualized in Figure 6.10, the most frequently used hashtags in health-related posts include #health, #mentalhealth, and #healthcare. These reflect major areas of public concern and discourse.

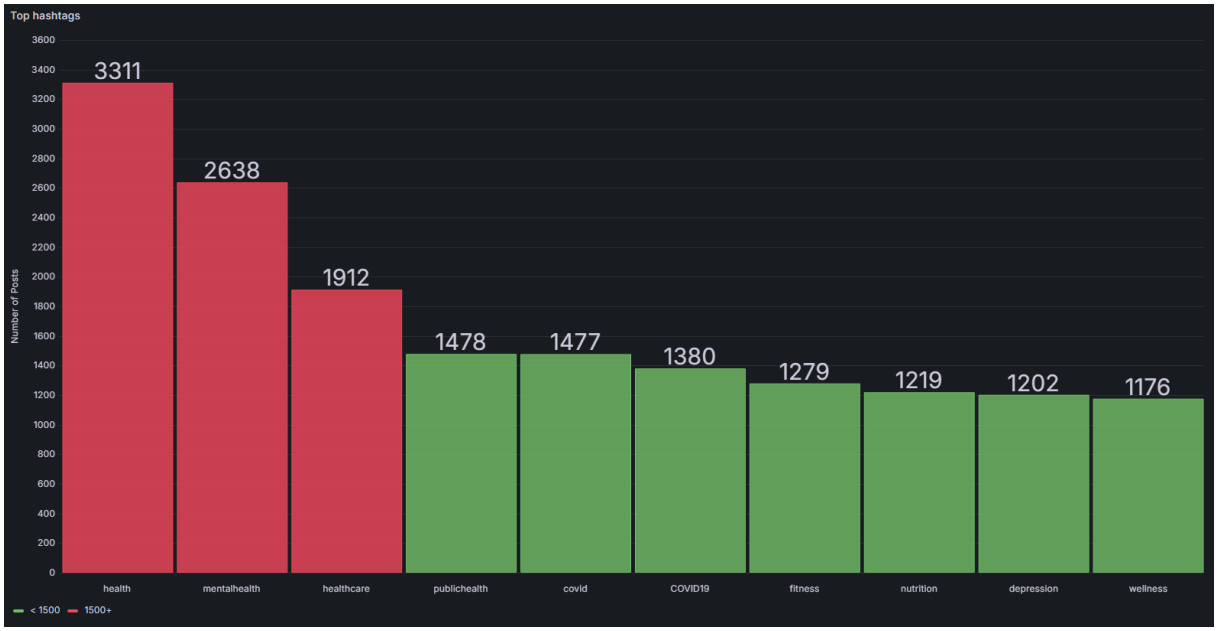


Figure 6.10: Top Health-Related Hashtags by Post Volume

6.11 Location-Based Content Distribution

Figure 6.11 displays the geographic spread of posts based on their content. North America and Europe dominate the activity, with scattered contributions from Asia and Africa.

Figure 6.11: Global Heatmap of Health-Related Posts

6.12 Interaction with Health Posts - Replies

Figure 6.12 presents the volume of replies to health-related posts. Peak interaction periods coincide with spikes in urgency and symptom trends.

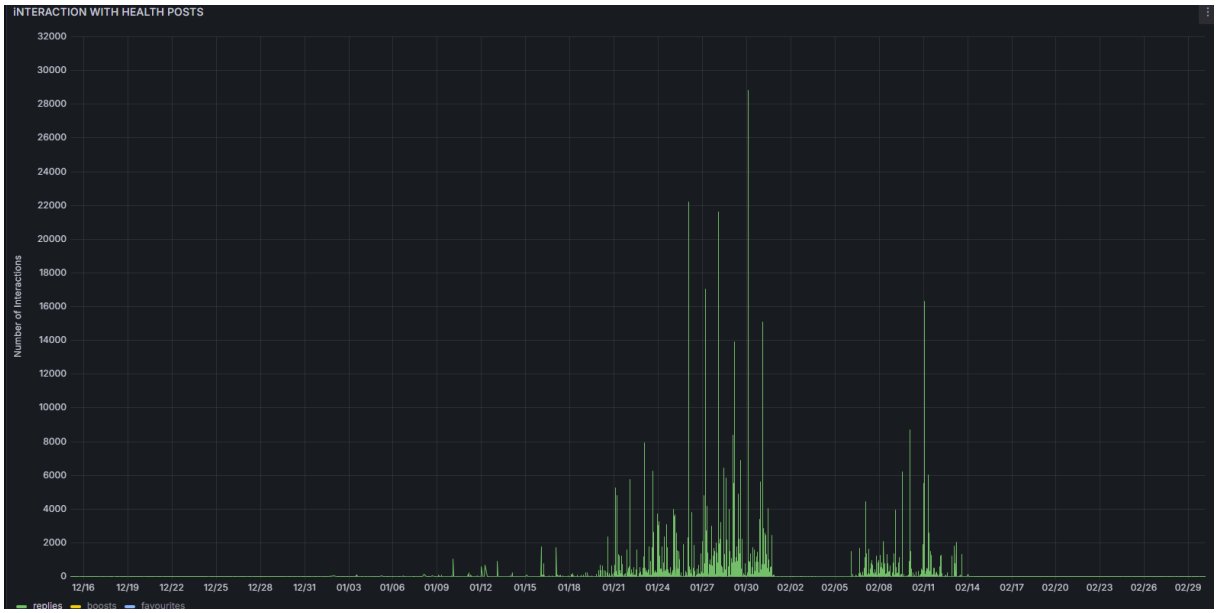


Figure 6.12: Volume of Replies to Health-Related Posts

6.13 Interaction with Health Posts - Boosts

Boosts (analogous to retweets or shares) offer insight into content amplification. Figure 6.13 shows a strong amplification pattern in late January and early February.

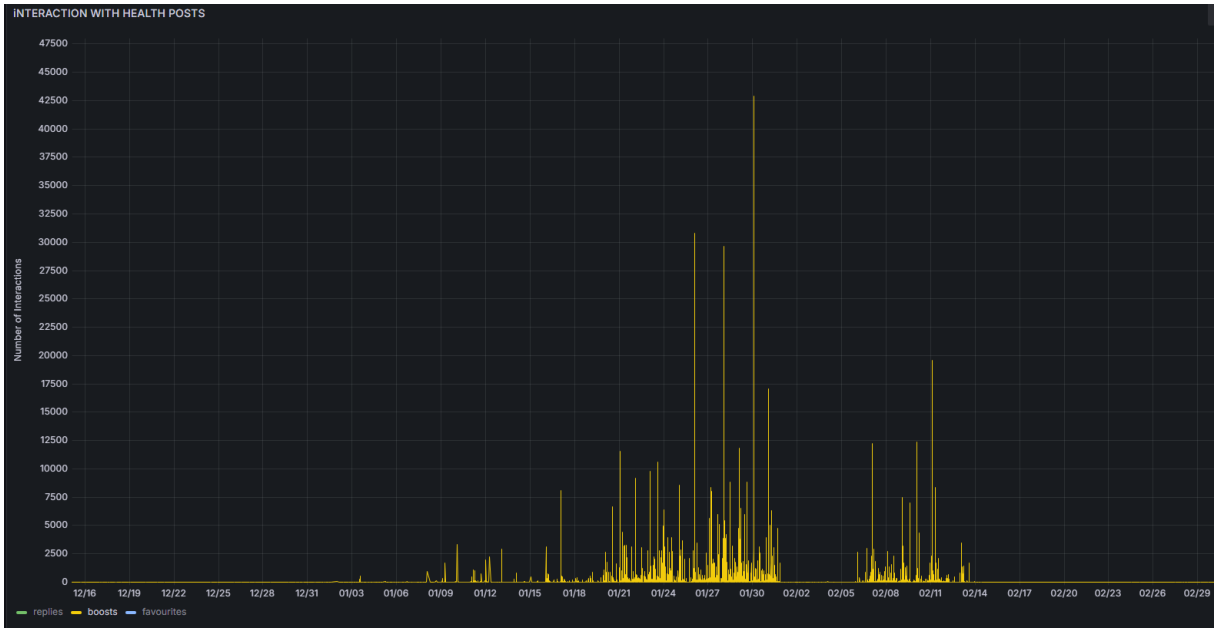


Figure 6.13: Boosts of Health-Related Posts Over Time

6.14 Interaction with Health Posts - Favourites

Favourites reflect positive user endorsement of health content. As seen in Figure 6.14, the engagement trend follows similar temporal patterns as replies and boosts.

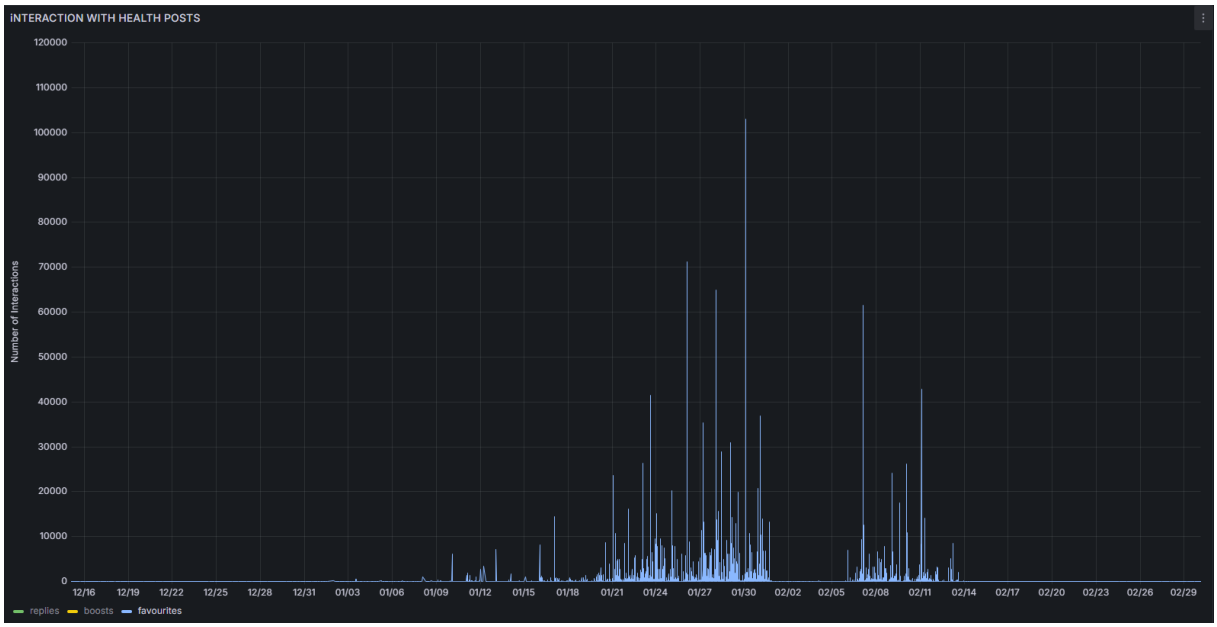


Figure 6.14: Favourites of Health-Related Posts Over Time

7 Discussion

This chapter revisits our four research questions (Section 2.2) and explains how the systems design-grounded in ActivityPub, machine learning, and realtime visualization fulfills each one.

7.1 RQ1: ActivityPub for Decentralized Data Collection

By leveraging the ActivityPub protocol, our application seamlessly integrates multiple independent servers without a single point of control. This architecture allows the system to gather posts from diverse communities and platforms in a uniform way. The federated model not only scales naturally as new instances join but also preserves local autonomy, ensuring broad coverage without sacrificing resilience.

7.2 RQ2: Effectiveness of Machine Learning Methods

Our hybrid ML pipeline combines topic modeling with transformer based classification to identify and tag health related content on the fly. Latent Dirichlet Allocation uncovers thematic clusters, while the LLaMA model refines relevance and sentiment judgments. Together, these methods provide a reliable, automated mechanism for filtering noise and surfacing meaningful patterns in unstructured social media text.

7.3 RQ3: Utility of RealTime Visualization

Connecting live data streams to Grafana transforms raw posts into interactive dashboards that update continuously. Public health users can monitor emerging discussions, sentiment shifts, and geographic patterns as they occur, rather than waiting for batch reports. This immediacy supports faster situational awareness, enabling stakeholders to respond proactively to rising concerns.

7.4 RQ4: Technical and Ethical Challenges

Implementing a federated, AI driven system introduces several challenges:

- **API Diversity**: Each platform exposes unique interfaces and usage limits. We addressed this with modular adapters and adaptive rate control, ensuring uninterrupted ingestion.
- **Model Generalization**: Generic transformer models may misclassify domain specific jargon. Prompt design and lightweight finetuning help mitigate these gaps, though specialized training remains a future enhancement.

- **Privacy and Bias**: Aggregating public posts risks exposing sensitive context and amplifying existing biases. We enforce anonymization, ethical data handling practices, and transparent model auditing to uphold user privacy and fairness.

7.5 Summary

In sum, our federated dashboard demonstrates that:

- ActivityPub enables robust, distributed data collection across heterogeneous servers.
- A combined LDA and LLaMA approach effectively identifies and categorizes health content in real time.
- Live Grafana visualizations turn continuous data flows into actionable insights.
- Careful engineering and ethical safeguards are essential to manage technical heterogeneity, model limitations, and privacy concerns.

Together, these elements validate the viability of a decentralized, AI powered platform for realtime public health monitoring and lay the groundwork for future enhancements in scalability, accuracy, and privacy preservation.

8 Evaluation of Alternative Approaches

In this section, we discuss the alternative model implementations we tested for classifying posts as Sickness Report or not. Specifically, we experimented with a fine-tuned Flan-T5 model and a pre-trained BERT model. The subsections that follow describe the challenges and limitations we encountered with each approach.

8.0.1 Flan-T5 Model for Sickness Report Classification

The Flan-T5 model, a T5-based text-to-text model fine-tuned for better instruction-following, was utilized to classify posts by framing the task as a question-answering problem (e.g., “Is this a sickness report?”). However, several issues were observed:

1. **Not Trained Specifically for Classification Tasks:** Flan-T5 is designed for general instruction-following, resulting in diverse, free-form responses that are less consistent for classification.
2. **Ambiguity in Responses:** The model can generate varied answers such as "I think so", "Yes, it might be", or "Maybe", making it challenging to definitively categorize a post.
3. **Generative Nature of the Model:** Instead of producing discrete labels, the model outputs full sentences, complicating the mapping to clear "yes" or "no" responses.
4. **Lack of Fine-Tuning on the Specific Task:** The small variant (“google/flan-t5-small”) was not fine-tuned on sickness-related data, limiting its ability to accurately classify health-related posts.
5. **Limited Input Context:** The simple prompt used might not provide sufficient context for the model to make accurate judgments.

8.0.2 BERT Model for Sickness Report Classification

A pre-trained BERT model was also employed for this task. This approach involved loading a pre-trained model and tokenizer, pre–processing the input text, and making predictions using PyTorch. The following challenges were identified:

1. **Insufficient or Poor Quality Training Data:** The model was not fine-tuned on a sufficient quantity of high-quality, labeled data specifically for sickness reports, hindering its generalization.
2. **Data Mismatch:** The characteristics of the inference data (e.g., language style, topics) differed significantly from the training data, leading to decreased performance.
3. **Class Imbalance:** An imbalance in the training dataset with many more examples of ’Not a Sickness Report’ than ’Sickness Report’ biased the model toward the dominant class.
4. **Limited Fine-Tuning Epochs:** The model may have been underfitted if trained for too few epochs or overfitted if trained for too many, compromising its ability to generalize.

5. **Model Complexity:** The high number of parameters in BERT makes it sensitive to overfitting, particularly when fine-tuned on a small or non-diverse dataset.
6. **Domain-Specific Challenges:** The use of colloquial language, slang, and ambiguous expressions in posts presents difficulties for a general-purpose BERT model not adapted to medical or health-related content.
7. **Training on General BERT without Adaptation:** Without sufficient domain-specific fine-tuning, the model lacks the specialized knowledge necessary to effectively classify health-related content.

This evaluation of alternative approaches highlights the limitations encountered with both models, reinforcing the need for domain-specific fine-tuning and a tailored model architecture for reliable sickness report classification.

9 Limitations

This chapter discusses the challenges encountered during the development and evaluation of the federated web application for public health data management using the ActivityPub protocol. While the system has demonstrated promising capabilities in aggregating diverse data sources and generating actionable insights, several limitations stemming from both inherent architectural factors and practical constraints remain. These limitations are discussed in the sections below.

9.1 Data Collection and API Integration

The federated nature of the ActivityPub ecosystem means that each platform (e.g., Mastodon, PeerTube, Lemmy, Misskey) operates with its own API structure, data format, and access policy. In practice, many instances are configured with restrictive privacy settings or adopt walled garden approaches that hinder comprehensive data access. As a result, only partial data may be retrievable from platforms that deliberately limit full fetching. The variability in API designs and rate-limiting strategies across instances necessitates ongoing adjustments in the data retrieval workflow, making it challenging to ensure consistency and completeness across the diverse federated landscape.

9.2 Machine Learning Analysis

The project integrates the LLaMA model (in both 3B and 1B parameter configurations) to extract health-related insights from unstructured social media content. Although we invested significant effort into smart prompt engineering refining the input prompts through extensive trial and error to optimize the model's performancethe limitations encountered are inherent to the models themselves. In particular, the relatively low parameter counts and the fact that these models were not specifically trained on health-related data constrain their ability to capture the full complexity and nuance of domain-specific language. Thus, even with careful and creative prompt engineering, the overall accuracy and depth of the extracted insights remain limited by the inherent characteristics of the models employed.

9.3 Decentralized Nature of ActivityPub

A core strength of ActivityPub is its federated design, which empowers independent servers to operate under their own policies and share data without a central authority. However, this decentralization also introduces intrinsic limitations. The absence of standardized API behavior and uniform data retention policies results in fragmented data availability, as each instance enforces its own rules regarding content sharing and privacy. Such fragmentation not only makes real-time aggregation and uniform processing challenging but also requires developers to continuously adapt their systems to accommodate

heterogeneous implementations. This limitation is not due to shortcomings in the implementation but is an inherent consequence of the decentralized modela trade-off between local autonomy and the ease of centralized data aggregation.

9.4 Local Hardware Constraints

The performance of the machine learning analysis is heavily influenced by local hardware capabilities. During development, the system was run on an RTX 3070 with 8 GB of memory a configuration that proved marginal for processing the large volumes of federated data required by the LLaMA model. Although software optimizations and containerization strategies have been applied, the relatively modest GPU capacity has resulted in slower inference times and limited scalability. This constraint affects not only the throughput of data analysis but also the overall responsiveness of the system. Future iterations of the project would benefit from upgrading to more robust hardware solutions, such as GPUs with higher memory and processing power, to better support real-time analytics and scalable data processing.

In summary, while the developed system successfully demonstrates the integration of federated health data and advanced machine learning analysis, the limitations highlighted in data collection, model capabilities, decentralized variability, and local hardware constraints point to key areas for future improvement. Addressing these challenges will be crucial for enhancing the system's robustness, scalability, and overall performance.

10 Future Work

This chapter offers a forward-looking perspective on how the insights from this thesis can be extended and refined. Rather than revisiting identified issues, the focus here is on strategies, enhancements, and new research directions that can further strengthen decentralized social networks built on ActivityPub. The following sections discuss these considerations in terms of protocol extensibility, user-centric security and privacy, scalability, user experience, advanced processing methods, long-term viability, and interdisciplinary collaboration.

10.1 Protocol Extensibility

A promising avenue for future work lies in extending the core ActivityPub protocol to accommodate more diverse and sophisticated forms of interaction. Although the existing specification supports a range of content types, integrating richer media such as voice, video, and real-time streams can broaden the networks overall utility. These enhancements might be introduced through well-defined extensions rather than altering the foundational protocol, allowing developers to adopt new features in a modular fashion. Additionally, bridging ActivityPub with other decentralized protocols such as Matrix, XMPP, or distributed storage solutions like IPFS can foster a more interconnected ecosystem. By focusing on standardized frameworks for adding or modifying features, researchers and implementers can experiment with novel ideas while minimizing fragmentation.

10.2 User-Centric Security and Privacy

Maintaining robust security and privacy remains a central objective for decentralized networks. Future endeavors could build on end-to-end encryption methods to make private communications both secure and user-friendly. Integrating emerging decentralized identity (DID) standards and verifiable credentials may also help users establish portable profiles, which would simplify moving between instances and bolster trust across the network. Another point of focus could be designing more granular access controls that give individuals greater flexibility in determining which audiences can view specific content. By enabling users to customize their online presence, decentralized platforms can encourage nuanced social interactions while ensuring personal data remains under user control.

10.3 Scalability and Performance

As decentralized networks grow, exploring techniques for managing high volumes of traffic and data will be critical. Load-balanced federation architectures, potentially supported by container orchestration tools like Kubernetes, can spread activity across multiple servers, thereby maintaining performance during peak usage. Adaptive rate control mechanisms, which adjust data flow according to real-time network conditions, can further optimize

resource allocation. In tandem with these approaches, developers can refine data storage strategies perhaps through advanced caching systems, graph databases, or distributed file solutions to suit the dynamic demands of social media environments. By cataloging best practices for scalability, researchers can pave the way for smoother adoption of decentralized platforms at large scale.

10.4 User Experience and Adoption

Decentralized platforms must also remain accessible to both newcomers and experienced users alike. Future research can investigate streamlined processes for onboarding, ensuring that account creation and instance selection pose minimal barriers to entry. More effective, privacy-centric discovery mechanisms could make it easier for individuals to locate relevant communities, topics, and people across disparate servers. Equally significant is the continuous refinement of user interfaces to accommodate linguistic diversity and accessibility considerations. By emphasizing inclusive design and iterative user research, the Fediverse can appeal to broader demographics and sustain growth over time.

10.5 Advanced Processing Methods

As the volume and complexity of content grow, so too does the need for more sophisticated data processing, moderation, and analysis. New machine learning models and frameworks, such as *DeepSeek*, have emerged since the inception of this thesis and offer promising capabilities for tasks like content classification, spam detection, and sentiment analysis. Future work could incorporate fine-tuned models building on large-language-model paradigms or specialized neural architectures to automatically detect and filter inappropriate or harmful content. These advances may also help reduce moderation burden on instance operators by providing assistive tools that flag anomalies or potential policy violations. Integrating such technologies, however, demands careful consideration of bias, transparency, and privacy. Designing model interpretability mechanisms and clear user notification systems will help ensure that AI-driven moderation aligns with the ethos of user autonomy that underpins decentralized networks.

10.6 Automated Graph Creation via Conversational Interfaces

A promising direction for future work is the integration of chat-based interfaces that allow users to request dynamic graph creation through natural language prompts. For example, a user could ask, "Show the monthly trend of COVID-related posts in Europe," and the system would automatically generate the corresponding Grafana visualization without manual setup.

This capability would lower barriers for non-technical users, streamline data exploration, and make decentralized monitoring platforms more interactive. While initially focused on

health data, the approach could extend to other fields such as environmental monitoring, education, or public infrastructure analytics, broadening the systems impact.

Implementing this feature would involve combining large language models with secure query templates, enabling safe translation of user intent into internal code. Future research could refine parsing methods, ensure responsible query execution, and develop intuitive feedback loops, making decentralized data analysis more accessible and adaptable across domains.

10.7 Long-Term Studies and Real-World Deployments

To gain deeper insight into how ActivityPub-based ecosystems evolve over time, it is important to conduct ongoing observations and targeted case studies. Multi-year research could shed light on patterns of governance, content diversity, and community resilience, capturing how federation practices adapt as user communities expand. Investigations into real-world deployments whether in educational settings, advocacy groups, or public service contextswould help document the unique hurdles and successes encountered when implementing decentralized social platforms. In gathering this empirical data, researchers can provide tangible guidance for future adopters while contributing to broader conversations on the impact of decentralized systems at societal, organizational, and policy levels.

10.8 Interdisciplinary Collaboration and Funding Models

There remains considerable scope for interdisciplinary collaboration that goes beyond purely technical considerations. Scholars in fields like sociology, law, design, or economics can offer fresh perspectives on the social structures and ethical dimensions of decentralized platforms, informing more holistic development strategies and governance models. These collaborations may reveal frameworks for content moderation that align with democratic values, highlight regulatory implications of cross-border federations, or propose novel approaches to interface design grounded in user psychology. In parallel, open-source communities can serve as innovation hubs for testing new ideas and sharing best practices. Finally, exploring sustainable funding mechanisms such as cooperatives, crowdfunding, or subscription models can ensure that smaller instance operators have the resources they need to remain online, moderate content, and foster community growth.

10.9 Conclusion

Taken together, these directions reveal a rich landscape for future work that goes beyond the initial scope of this thesis. Focusing on protocol extensibility and scalable architectures will ensure that the system can grow with user needs, while user-centric security measures and advanced processing methods will strengthen trust and functionality. Inclusive design practices and longitudinal studies can help us understand and accommodate diverse communities over time, and interdisciplinary cooperation will bridge technical, social, and policy perspectives. By pursuing these avenues, developers and researchers can help ActivityPub-based platforms fulfill their promise of enhanced user agency, democratic governance, and truly global, inclusive communication.

BIBLIOGRAPHY

[1] C. Lemmer-Webber and J. Tallon, "Activitypub," https://www.w3.org/TR/2018/REC-activitypub-20180123/, 2018, world Wide Web Consortium (W3C) Recommendation.

[2] S. Feng and A. Kirkley, "Integrating online and offline data for crisis management: Online geolocalized emotion, policy response, and local mobility during the COVID crisis," *Scientific Reports*, vol. 11, p. 8514, 2021. [Online]. Available: https://doi.org/10.1038/s41598-021-88010-3

[3] R. Chandrasekaran, R. Desai, H. Shah, V. Kumar, and E. Moustakas, "Examining public sentiments and attitudes toward COVID-19 vaccination: Infoveillance study using twitter posts," *JMIR Infodemiology*, vol. 2, no. 1, p. e33909, 2022. [Online]. Available: https://infodemiology.jmir.org/2022/1/e33909

[4] S.-F. Tsao, H. Chen, T. Tisseverasinghe, Y. Yang, L. Li, and Z. A. Butt, "What social media told us in the time of covid-19: a scoping review," *The Lancet Digital Health*, vol. 3, no. 3, pp. e175–e194, 2021. [Online]. Available: https://doi.org/10.1016/S2589-7500(20)30315-0

[5] D. Capurro, K. Cole, M. I. Echavarría, J. Joe, T. Neogi, and A. M. Turner, "The use of social networking sites for public health practice and research: A systematic review," *J Med Internet Res*, vol. 16, no. 3, p. e79, Mar 2014. [Online]. Available: http://www.jmir.org/2014/3/e79/

[6] S. A. Rains, *Cope with Malady Digitally.* City, Country: Publisher Name, 2018, hOW TO CITE THIS BOOK.

[7] M. Smolinski, "Preventing the next pandemic," *Stanford Social Innovation Review*, pp. 55–60, Winter 2022, pDF file.

[8] E. Weber, D. Papadopoulos, . Lapedriza, F. Ofli, M. Imran, and A. Torralba, "Incidents1m: A large-scale dataset of images with natural disasters, damage, and incidents," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 4, pp. 4768–4781, 2023. [Online]. Available: https://doi.org/10.1109/TPAMI.2022.3191996

[9] I. Dayan *et al.*, "Federated learning for predicting clinical outcomes in patients with covid-19," *Nature Medicine*, vol. 27, no. 11, pp. 1735–1743, 2021.

[10] G. Yang *et al.*, "Using artificial intelligence to improve public health: a narrative review," *Journal of Medical Internet Research*, vol. 22, no. 8, p. e18965, 2020.

[11] N. A. Christakis and J. H. Fowler, *Connected: The Surprising Power of Our Social Networks and How They Shape Our Lives.* Little, Brown and Co., 2009.

[12] B. . TRADECRAFT, "An osint approach to event-based surveillance for epidemic intelligence," https://bluedot.global/an-osint-approach-to-event-based-surveillance-for-epidemic-intelligence/, 2022, whitepaper, Sept. 8, 2022.

[13] BlueDot, "The ai revolution in pharma," https://bluedot.global/the-ai-revolution-in-pharma/, 2022, industry Report.

[14] ——, "2021 year-in-review of infectious diseases: Part 1 emerging infectious diseases & pathogens," https://bluedot.global/focus-report-year-in-review-1-emerging-infectious-diseases-pathogens/, 2022, focus Report, Feb. 14, 2022.

[15] ——, "How ai-powered biothreat intelligence improves pharma sales & supply chain performance," https://bluedot.global/ai-powered-infectious-disease-intelligence-in-pharmaceutical-organization/, 2022, industry Report, July 2022.

[16] ——, "How to build a resilient supply chain in the face of ongoing biothreats," https://bluedot.global/how-to-build-a-resilient-supply-chain-in-the-face-of-ongoing-biothreats/, 2022, whitepaper.

[17] Python Software Foundation, *Python 3.12.5 Documentation*, 2024. [Online]. Available: https://docs.python.org/3.12/

[18] Django Software Foundation, *Django Documentation*, 2024. [Online]. Available: https://docs.djangoproject.com/

[19] Celery Team, *Celery Documentation*, 2024. [Online]. Available: https://docs.celeryq.dev/

[20] K. Vyas, "django-cron: Django cron jobs in minutes," https://pypi.org/project/django-cron/, 2024.

[21] Redis, *Redis Documentation*, 2024. [Online]. Available: https://redis.io/documentation

[22] Hugging Face, "Transformers Documentation," https://huggingface.co/docs/transformers/en/index, 2025.

[23] Meta AI, "Llama 3: Open foundation and fine-tuned chat models," https://ai.meta.com/resources/models-and-libraries/llama/, 2024.

[24] OpenAI, "Whisper: Robust Speech Recognition," https://github.com/openai/whisper, 2022.

[25] S. Bird, E. Klein, and E. Loper, *Natural Language Processing with Python*. O'Reilly Media, Inc., 2009. [Online]. Available: https://www.nltk.org/book/

[26] Explosion AI, "spacy 3 documentation," https://spacy.io/usage, 2023.

[27] R. ehek and P. Sojka, "Gensimtopic modelling for humans," https://radimrehurek.com/gensim/, 2010.

[28] L. Richardson, *Beautiful Soup Documentation*, 2023. [Online]. Available: https://www.crummy.com/software/BeautifulSoup/bs4/doc/

[29] Geopy contributors, "Geopy: Python geocoding toolbox," https://geopy.readthedocs.io/, 2021.

[30] youtube-dl contributors, “youtube-dl: Download videos from youtube and more,” https://ytdl-org.github.io/youtube-dl/, 2006.

[31] PostgreSQL Global Development Group, *PostgreSQL 15 Documentation*, 2024. [Online]. Available: https://www.postgresql.org/docs/15/

[32] Grafana Labs, *Grafana Documentation*, 2024. [Online]. Available: https://grafana.com/docs/

[33] Docker, Inc., *Docker Documentation*, 2024. [Online]. Available: https://docs.docker.com/

[34] Hugging Face, “huggingface/hub,” https://huggingface.co/, 2024.

[35] U. Qazi, M. Imran, and F. Ofli, “GeoCoV19: A Dataset of Hundreds of Millions of Multilingual COVID-19 Tweets with Location Information,” https://dx.doi.org/10.21227/et8d-w881, June 2020, iEEE Dataport.