UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

A Visual Analytics System for Syndromic Surveillance and Pandemic Readiness within a One Health Context

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR of PHILOSOPHY

By

Aseel Basheer

Norman, Oklahoma

2025

A VISUAL ANALYTICS SYSTEM FOR SYNDROMIC SURVEILLANCE AND PANDEMIC READINESS WITHIN A ONE HEALTH CONTEXT

A DISSERTATION APPROVED FOR THE
SCHOOL OF COMPUTER SCIENCE

BY THE COMMITTEE CONSISTING OF

Dr. Dean Hougen, Committee Chair

Dr. David Ebert, Committee Co-Chair

Dr. Chongle Pan, Committee Member

Dr. Charles Nicholson, Graduate College Representative

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my advisors, Dr. David Ebert and Dr. Dean Hougen, for their unwavering support, insightful guidance, and thoughtful mentorship throughout my PhD. Their expertise and vision have been instrumental in shaping the direction of my research, and their patience and encouragement have helped me stay focused and motivated during the most difficult stages of this work.

I would also like to thank the members of my dissertation committee, Dr. Chongle Pan and Dr. Charles Nicholson, for their valuable feedback and thought-provoking questions that strengthened the quality and clarity of this dissertation.

I am thankful to my colleagues and collaborators at the Data Institute for Societal Challenges (DISC) at the University of Oklahoma. I would like to recognize Dr. Wolfgang Jentner, whose encouragement, support, discussions, and technical assistance enriched my work in countless ways.

On a personal note, I would like to thank my family and friends for their unwavering love and support, which have been instrumental in my success. To my mother, my brothers, my spouse, and my lovely children, thank you for your sacrifices, encouragement, and for always reminding me of the bigger picture. Your strength and support have been my foundation. To my father, whose love and guidance shaped who I am. I wish he were here to celebrate this moment.

Table of Contents

List of Tables

List of Figures

Abstract

Emerging infectious diseases continue to pose a significant global threat, as pandemics disrupt societies, economies, and health systems around the world. To effectively anticipate and mitigate such crises, it is essential to gain timely insights from a variety of complex data sources that encompass human, animal, and environmental health. Integrating this data within a One Health framework, supported by advanced visualization and modeling techniques, provides a promising approach to developing more effective preparedness and response strategies. This dissertation advances the use of One Health data, encompassing human, animal, and environmental domains, in conjunction with interactive visual analytics methods to enhance pandemic preparedness and response. A systematic review of One Health dashboards and visualization systems published between 2015 and 2024 has been presented. This review categorizes systems by purpose, diseases addressed, data domains, interactivity, and user groups, revealing both promising trends and significant gaps. While many studies still rely on single-domain datasets, there is a clear movement toward integrated approaches, though fully comprehensive One Health implementations remain rare. Building on these insights, the analysis examines the role of visualization during the COVID-19 pandemic, with a particular focus on dashboards that were widely used to communicate risks, policies, and projections. This analysis maps data types, visualization tasks, and representations, highlighting strengths, limitations, and opportunities for advancing visualization research to better support future pandemic response. To effectively apply these findings, the Predictive Intelligence for Pandemic Prevention (PIPP) dashboard has been introduced. This dashboard integrates various One Health datasets, such as clinical outcomes, animal health records, environmental variables, and wastewater surveillance. Additionally, *the One Health Mixed-Effects Modeling Approach (OH-MEMA)*, which is an interactive tool, is proposed, built on a lag-aware, non-linear mixed-effects modeling framework that captures variability in both spatial and demographic factors. Predictive results are coupled with interactive visualizations, time series comparisons, regional analyses, and analytic provenance trees, enabling users to explore scenarios, evaluate predictors, and iteratively refine models. Together, these contributions demonstrate the value of combining epidemiological modeling with interactive visual analytics to provide timely, interpretable, and actionable insights. By integrating multiple data domains, this work illustrates the potential of a One Health visual analytics approach to support early detection, guide interventions, and strengthen resilience against future pandemics.

Chapter 1

Introduction

The world has experienced and survived over 249 pandemics (Wikipedia 2023). The most recent pandemic was COVID-19, declared a global health emergency by the World Health Organization (WHO) on January 30, 2020, and subsequently declared a pandemic on March 11, 2020. According to World Health Organization (WHO) reports, the SARS-CoV-2 (COVID-19) pandemic has resulted in millions of deaths worldwide (WHO 2022). This pandemic has profoundly altered daily routines and reshaped how societies process and communicate information. During the peak of the COVID-19 pandemic, many governments worldwide provided daily briefings to convey policies using visualizations of collected and computed data, such as daily statistics on fatalities and hospital admissions (Allen et al. 2024).

Beyond COVID-19, global threats such as H5N1 and mpox continue to underscore the need for improved outbreak management and early disease detection (León-Figueroa et al. 2022). Innovative solutions are crucial for addressing the growing challenges surrounding disease surveillance and pandemic response. *One Health* is a global health initiative that recognizes the interconnected health of people, animals, and the environment (World Health Organization 2024). One Health aims to identify, mitigate, and respond to health hazards at the human-animal-environment interface, such as the COVID-19 pandemic (Sinclair 2019; Collignon and McEwen 2019). These proactive measures can be more effective in reducing the scope and impact of an outbreak than reactive actions taken after the pathogen has already spread extensively (Food and Agriculture Organization (FAO) 2024). It has proven to be a promising approach to addressing public health issues at local, regional, and national levels (Queenan et al. 2017). Key organizations, including the World Health Organization (World Health Organization 2024) and the Food and Agriculture Organization (FAO) (Food and Agriculture Organization of the United Nations 2024), emphasize the One Health approach as a transformative method for managing zoonotic diseases, antimicrobial resistance (AMR), vector-borne diseases, and various health challenges. For example, the WHO's One Health framework emphasizes the importance of integrating human, animal, and environmental health to enhance proactive disease surveillance and response, particularly in the aftermath of the COVID-19 pandemic (World Health Organization 2023). The interconnectedness of various health domains underscores the importance of systems that support public health officials, researchers, and policymakers in tasks such as early detection, prediction, planning, mitigation, and monitoring (Degeling et al. 2015).

Effective pandemic preparedness requires the capability to quickly detect, interpret, and respond to various health signals across human, animal, and environmental domains (Kamalrathne et al. 2023). The One Health paradigm emphasizes the interconnectedness of multiple datasets, underscoring the need for integrated approaches to disease surveillance (Singh et al. 2024). However, a significant challenge lies in combining diverse data sources into actionable models. This is particularly important at detailed spatial and temporal scales, where decision-makers must be able to understand and trust the information (Mubareka et al. 2023; Faijue et al. 2024; Mukherjee et al. 2023). The heterogeneity of data types, ranging from clinical case counts and weather indicators to wastewater pathogen concentrations, complicates both temporal modeling and visualization. Differences in spatial and temporal scales, measurement units, and data collection frequencies further hinder integration (Faijue et al. 2024; Mukherjee et al. 2023).

Time-series analysis is essential for identifying trends, seasonality, and anomalies in epidemic surveillance. Moreover, when visualizing such data, combining multiple modalities can lead to visual clutter, making it difficult for analysts to interpret key signals (Thorve et al. 2018).

Current pandemic response systems have several critical limitations (Gotz and Borland 2016; Rind et al. 2013). Health departments work with fragmented data that comes in different formats from various sources, which makes integration challenging and delays response times (Mubareka et al. 2023; Carroll et al. 2014). Existing visualization tools typically show basic charts and maps that focus on population-level trends, missing the subtle, fine-grained patterns that could inform targeted interventions (Faijue et al. 2024; Mukherjee et al. 2023; Bisset et al. 2014). When time is critical, decision-makers often face cognitive overload from cluttered interfaces that don't clearly point out what matters most (Thorve et al. 2018; Betz et al. 2023). Another key limitation is that conflicting forecasts from different modeling approaches make uncertainty assessment difficult (Xu et al. 2021). Traditional modeling workflows struggle with the insufficient integration of spatial and temporal dimensions (Luo et al.), and many tools lack the necessary features for effective decision support (Kim et al. 2024; Rind et al. 2013).

The COVID-19 pandemic has highlighted the potential and limitations of current epidemiological visualization approaches. For tracking and visualizing pandemic data, numerous systems were rapidly developed (Reinert et al. 2020a; Afzal et al. 2020; Comba 2020). While several integrated frameworks have emerged, a universally adopted comprehensive analytical framework that fully integrates diverse data sources and supports real-time decision-making across all scales remains elusive. For instance, the evolution of wastewater-based disease surveillance has shown promising results as an early warning system (Daza-Torres et al. 2024; Peccia et al. 2020); however, fusing these data with clinical and demographic data remains challenging.

Visualization has been crucial not only for communicating the spread of the disease but also for developing and distributing vaccines. It played a key role in sharing results with both the public and the scientific community, as well as in improving the quality of scientific publications (Clark 2023). In response to the severe impact of such pandemics, a more effective strategy for preventing, intervening in, and mitigating the spread of infectious diseases, including endemics, epidemics, and pandemics, is increasingly recognized as the *One Health* approach, which integrates human, animal, and environmental health (CDC 2023).

Given the variety of One Health data sources across different spatial and temporal scales, integrated visual analysis tools are necessary for effective analysis, correlation, predictive and causal analysis, understanding, and decision-making to support effective management and response (Dey et al. 2020).

This dissertation examines how One Health data and visual analytics can be integrated to enhance pandemic preparedness and response. Pandemics and emerging infectious diseases continue to challenge global health systems, highlighting the need for innovative approaches to early detection, monitoring, and decision-making. The One Health framework offers a promising foundation for addressing these challenges. However, current systems for integrating and analyzing such diverse datasets remain fragmented, and visualization practices often fail to provide timely, actionable insights.

To address these gaps, the dissertation begins with a systematic review of existing One Health systems and dashboards, as presented in Chapter 3. This review synthesizes research published over the last decade, categorizing systems by purpose, diseases addressed, data types, user groups, and visualization techniques. The findings reveal a fragmented landscape: while some systems successfully integrate human, animal, and environmental data, many remain siloed or focus on limited domains. This review highlights both the promise of integrated approaches and the urgent need for more comprehensive visual analytics systems to support One Health initiatives.

Then, we examine the role of visualization during global health crises, as illustrated in Chapter 4. Visualization played a central role in communication and decision-making during recent pandemics, but its applications often focused narrowly on data dissemination. By analyzing the literature on visualization in pandemic contexts, this work introduces the concept of *visualization readiness*, which refers to the ability of the visualization community to respond quickly and effectively to emerging crises. The analysis identifies significant challenges, including underutilization of visualization beyond communication, limited infrastructure for complex data, and insufficient evidence of long-term impact. Addressing these issues is crucial for preparing visualization research and practice for future crises.

The dissertation then transitions from analysis to design, presenting the Predictive Intelligence for Pandemic Prevention (PIPP) dashboard in Chapter 5. This interactive system integrates diverse One Health datasets, ranging from human clinical outcomes to wastewater and environmental data, into a unified platform. The dashboard emphasizes usability and interactivity, allowing users to explore data through maps, hierarchical charts, and dynamic time series. By enabling the combination of datasets that would otherwise remain isolated, the system supports early identification of emerging threats and provides accessible insights for both experts and the public.

Finally, we introduce *OH-MEMA*, which is an integrated One Health mixed-effects modeling approach for syndromic surveillance. OH-MEMA is a visual analytics framework that integrates predictive modeling with interactive exploration, as illustrated in Chapter 6. At its core is a lag-aware, mixed-effects statistical modeling engine that accounts for spatial and demographic variability while supporting flexible exploration of time series data. The framework allows domain experts to iteratively build, evaluate, and refine models, aided by an analytic provenance tree that records modeling decisions and outcomes. By unifying human, animal, and environmental datasets in an interpretable modeling system, this framework provides a powerful tool for evidence-based decision-making in public health.

Although Chapters 5 and 6 both focus on predictive analytics within a One Health context, they describe two separate dashboards. Chapter 5 presents the *Predictive Intelligence for Pandemic Prevention (PIPP)* dashboard, which implements single-predictor modeling using classical and decomposition-based time-series methods to capture temporal trends and seasonality from individual data streams such as clinical outcomes, animal health records, environmental variables, or wastewater surveillance within a secured environment. In contrast, Chapter 6 introduces *OH-MEMA*, a distinct, publicly accessible dashboard that implements multi-factor prediction through lag-aware mixed-effects modeling and visualization. This second system serves as a methodological companion to the PIPP dashboard, enabling flexible exploration and forecasting using either user-provided or publicly available datasets. Together, these dashboards demonstrate complementary approaches to integrating predictive modeling and interactive visualization within the One Health framework.

The dissertation roadmap can be summarized in Figure 1.1.

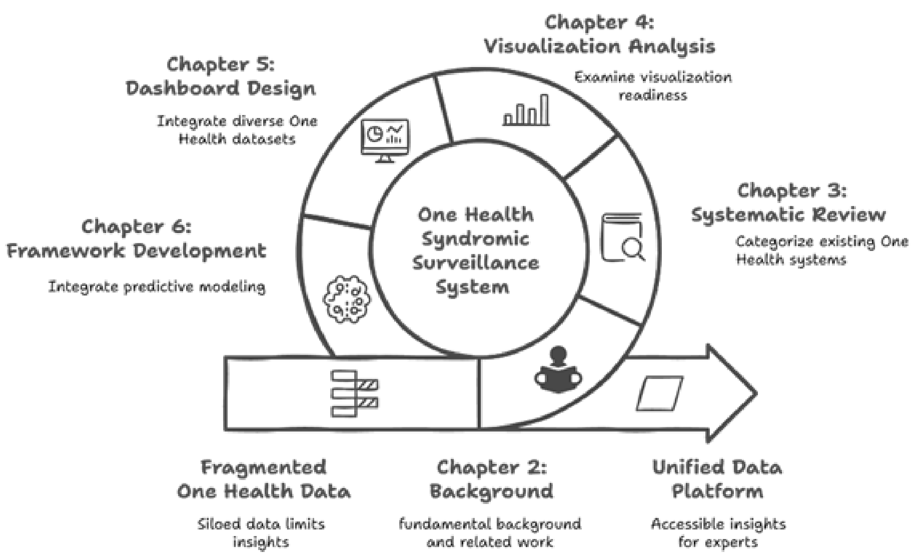


Figure 1.1: The dissertation roadmap.

1.1 Needs and Challenges in One Health Visual Analytics

Fully integrated visual analytics platforms that fuse these heterogeneous datasets with predictive modeling are rare, and existing literature reveals key gaps in:

1. **Heterogeneous data integration:** Few systems combine human, animal, and environmental data streams in a cohesive analytical framework.
2. **Interactive exploration:** Many dashboards prioritize display over exploration, limiting users' ability to test hypotheses or adjust models dynamically.
3. **Predictive modeling:** Outbreak prediction, even with predictive models, requires consideration of complex, non-linear, and lagged relationships between predictors and outcomes.

These needs require a framework that integrates data, predictive analytics, and interactive visualization in a cohesive, human-in-the-loop system.

1.2 Thesis Statement and Contributions

Thesis statement: Interactive exploration, integration, and modeling of multi-domain One Health data can enhance early outbreak detection, improve predictive accuracy, and support evidence-based decision-making in public health.

The contributions of this dissertation are:

1. **Systematic mapping of existing systems:** A meta-analysis of One Health dashboards and visualization systems (2015–2024), mapping data sources, target diseases, analytical tasks, and interactivity levels to reveal integration and capability gaps (Chapter 3). In addition, a systematic mapping of visualization practices for pandemics (2013–2024) is provided (Chapter 4). Together, these surveys benefit both the

visualization research community and health-related domains by providing a comprehensive overview of the state of the art.

2. **Conceptualization of visualization readiness:** An evaluation of how visualization was deployed during COVID-19 (2013–2024) (Chapter 4), leading to the conceptualization of *visualization readiness* as a framework for assessing the preparedness of visualization research and practice to respond to urgent societal needs such as pandemics. This contribution includes identifying strengths, shortcomings, and challenges, as well as proposing opportunities for advancing visual analytics in pandemic readiness.

3. **Framework and prototype implementation:** The design, development, and demonstration of the Predictive Intelligence for Pandemic Prevention (PIPP) dashboard (Chapter 5). The dashboard integrates multiple One Health data sources, including clinical outcomes, animal health records, environmental variables, and wastewater surveillance. It also includes a surveillance tool in the form of an interactive One Health visual analytics platform that integrates heterogeneous datasets, lag-aware mixed-effects predictive models, and analytic provenance tracking to support evidence-based public health decision-making.

4. **Development of a lag-aware mixed-effects visual analytics framework:** The development of *OH-MEMA*, a methodological framework that integrates lag selection, mixed-effects modeling, and interactive visualization for disease forecasting (Chapter 6). This framework demonstrates how statistical modeling and visual analytics can be combined to support the timely collection of pandemic intelligence across heterogeneous One Health datasets.

The ultimate goals of this dissertation are to:

1. Advance understanding of the state of One Health–aligned visual analytics by connecting literature analysis with practical design needs.

2. Provide a reusable framework for integrating predictive models and interactive visualization to support early detection and intervention.

3. Demonstrate the feasibility and utility of such a framework through a COVID-19, Flu, and RSV case studies using human, animal, and environmental datasets.

The rest of this dissertation is organized as follows:

- Chapter 2 provides the fundamental background needed for the rest of the dissertation.
- Chapter 3 discusses in detail the results of the One Health review.
- Chapter 4 discusses in detail the results of the visual representations review.
- Chapter 5 describes the proposed framework for the prediction dashboard.
- Chapter 6 introduces the proposed framework for the mixed-effects modeling framework used within the system.
- Chapter 7 concludes the dissertation by providing a summary and future directions.

Chapter 2

Background and Related Work

This chapter provides an overview of the theoretical foundations, prior research, and existing systems that inform the development of this dissertation. It discusses the One Health framework and its role in pandemic preparedness, examines syndromic surveillance and data integration methods, reviews visual analytics applications in epidemiology, and outlines mixed-effects modeling approaches relevant to disease forecasting. Together, these topics establish the context for the visual analytics framework and modeling strategies proposed in later chapters.

2.1 One Health Framework and Pandemic Preparedness

Current systems that incorporate One Health data often utilize a variety of data sources and visual analytics techniques. However, their purposes, methodologies, and levels of interactivity can differ greatly. As a result, the rapid evolution and growth toward a One Health approach has spawned the creation of various visual and analytical systems to improve public health initiatives in regions across the globe (Rüegg et al. 2018). These new and existing systems, built around One Health, are designed to address various problems and purposes related to the users, datasets, diseases, and tasks that the system performs.

There are numerous survey papers regarding the One Health approach and its applications, such as (Sharan et al. 2023; Rüegg et al. 2018; Bordier et al. 2020; Mukherjee et al. 2023). Sharan et al. (2023) focused specifically on surveillance strategies for zoonotic diseases, assessing global surveillance frameworks, their design, and the core components necessary for establishing surveillance systems at the human-animal-environment interface. In contrast, Rüegg et al. 2018 assessed the effectiveness and integration of One Health initiatives. They employed a structured methodology to assess the impact and coherence of One Health initiatives, with a focus on the evaluation framework rather than the functional features of specific systems.

Bordier et al.(2020) addressed surveillance systems in the One Health context, aiming to characterize their organizational and operational aspects rather than their visual or interactive functionalities. Their work seeks to establish a conceptual framework for understanding the essential organizational and functional characteristics of One Health surveillance systems. They explored the factors that impact their implementation, such as inter-sectoral collaboration, and the specific barriers and levers that influence their performance.

Another review by Mukherjee et al. examines the potential of the One Health framework as a solution for cross-sectoral challenges, particularly in pathogen detection and managing zoonotic viral and bacterial diseases (Mukherjee et al. 2023). This review emphasizes the importance of the One Health approach, which connects the detection and effective treatment of diseases by highlighting the interdependent relationship between human and animal health. The paper underscores the vital role that One Health initiatives play in preventing and controlling infectious diseases.

Although numerous studies explore the concept of One Health and its significant roles, they tend to focus on specific visual methods and disease characteristics. These studies do not provide a comprehensive review of One Health systems across a broad spectrum of infectious diseases and applications. Given the limited knowledge about the systems used to implement One Health initiatives worldwide, we prepared a comprehensive review to evaluate existing systems that utilize One Health datasets for infectious disease detection and management.

Unlike previous surveys, our work systematically reviewed existing systems and dashboards that integrate One Health datasets for infectious disease management. We categorized these systems based on their purpose, the diseases they address, the types of datasets used (human, animal, environmental), the user groups interacting with the system, and the visual analytics techniques employed. We then analyzed these classifications to identify trends and highlight areas for future improvement in the development of comprehensive One Health systems.

2.2 Syndromic Surveillance Systems

The primary objective of our work is to investigate how public health surveillance data can be integrated with other One Health datasets by developing user-friendly data visualization models. This aims to highlight the significance of visual analytics techniques and advocate for the timely identification of potential pandemic threats. Time series prediction and correlation analysis are integral to our system, as they directly support the goals of early detection, understanding complex relationships, and making informed public health decisions, all of which are crucial for pandemic prevention and response.

2.2.1 Dashboards for Pandemics and Health Surveillance

Numerous papers in the field of visual analytics have been published, particularly those focused on analyzing pandemic threats, modeling epidemic dynamics, and assessing public health responses. For instance, Afzal et al. 2011 focus on building an interactive decision support system to study epidemic models and their impacts. Similarly, Maciejewski et al. 2011 have developed a visual analytics toolkit to analyze the impact of decision measures implemented during a simulated pandemic influenza scenario. Our work builds upon previous work to integrate diverse One Health data, including human health information and data related to animals, weather conditions, and Google Trends data. This integration provides a more comprehensive view of disease spread and pandemic dynamics by enabling the identification of correlations and trends that may not be evident when analyzing a single dataset. For example, our approach can reveal how changes in environmental factors or animal health indicators might correlate with emerging human health threats, providing a more comprehensive understanding of potential pandemic risks. Additionally, Afzal et al. 2020 have developed a visual analytics environment that explores the spread of COVID-19 by providing county-level information, including population, demographics, and hospital bed availability. While this approach offers valuable insights into the resources available to combat the pandemic, our dashboard extends these capabilities by incorporating dynamic, interactive features. Users can explore temporal-spatial data through tools such as choropleth maps, sunbursts, and line charts, utilizing brushing techniques for more granular and engaging analysis. This enhances user interaction and provides a detailed, multi-dimensional view of pandemic data, making it easier for experts and the general public to understand and respond to emerging threats.

2.2.2 Time Series Prediction

Time series analysis plays a pivotal role in representing COVID-19 data by capturing temporal trends, handling dependencies, evaluating interventions, forecasting future trends, and detecting anomalies. These capabilities are crucial for understanding the progression of the pandemic and making informed public health decisions. By leveraging time series models, researchers and policymakers can gain valuable insights into the dynamics of COVID-19, leading to more effective response strategies and resource allocation (Betcheva et al. 2024). The pandemic's progression is inherently temporal, with patterns that evolve due to various factors like policy interventions, seasonal effects, and changes in public behavior (Gumel et al. 2021). COVID-19 data exhibits strong temporal dependencies, such as the effect of past infection rates on future case numbers. Time series models are designed to account for these dependencies, making them effective for forecasting and analyzing trends based on historical data (Bamana et al. 2024). Moreover, time series analysis enables the evaluation of the impact of public health interventions by comparing changes in trends before and after interventions are implemented, which is crucial for assessing the effectiveness of measures such as lockdowns, vaccination campaigns, and social distancing (Zhou et al. 2020). Furthermore, time series analysis effectively detects anomalies and trends in COVID-19 and health surveillance data, allowing public health officials to identify unusual spikes or declines in case numbers (Paltiel et al. 2020).

Our study builds upon and extends the existing research by integrating a broader spectrum of One Health data, including human-related factors, animal health data, and environmental conditions. While previous studies, such as those by (Mavragani and Gkillas 2020; Shinde et al. 2020, and Rahimi et al. 2023), have focused on specific aspects like Google Trends data or seasonal time series analysis, our work aims to explore the relationship of these diverse datasets within the context of COVID-19. By doing so, we seek to identify more comprehensive correlations and patterns that can inform public health strategies.

2.2.3 Correlation Analysis

Correlation Analysis helps identify significant associations between the pandemic and various variables, including environmental conditions, demographic factors, and public behavior. For instance, Li et al. 2020 studied the relationship between COVID-19 incidence and the *Air Quality Index (AQI)* as well as five meteorological variables: daily temperature, highest temperature, lowest temperature, temperature difference, and sunshine duration. Their findings revealed a significant correlation between COVID-19 incidence and the AQI, with temperature being the only meteorological parameter that consistently correlated well with COVID-19.

Similarly, Rui et al. 2021 analyzed the state-level correlation between COVID-19 risk and weather/climate factors in the USA, discovering that variables such as maximum and minimum temperatures, humidity, cloud coverage, and atmospheric ozone density were strongly associated with the COVID-19 pandemic in many states. Using Kendall and Spearman-rank correlation tests, they highlighted the impact of these factors on the virus's spread. In a related study, Yuan et al. 2021a assessed the relative risk of meteorological factors in relation to the onset of COVID-19 using *Generalized Additive Models (GAM)* and the *Distributed Lag Nonlinear Model (DLNM)*. They found that the mean temperature, wind speed, and relative humidity were negatively correlated with the daily number of new COVID-19 cases, while the diurnal temperature range was positively correlated with the daily number of new COVID-19 cases. Additionally, research by Lipsitch 2020 predicted that SARS-CoV-2 would transmit more efficiently in winter than in summer, emphasizing the critical role of temperature in COVID-19 transmission.

We aim to use and generate indices based on multiple climate factors and correlate these with public and animal health data.

2.2.4 Causality Assessment

Causality assessment aims to determine whether a specific factor directly influences the outcome of interest, such as the severity of COVID-19. Establishing causality, particularly in the context of COVID-19, is challenging due to the complexity of the factors involved and the pandemic's dynamic nature. Kang et al. (2021) proposed a causal inference framework that combines structural causal models with machine learning techniques to explore potential causal relationships between environmental factors and COVID-19 severity in various Chinese cities. Their study concluded that environmental factors were unlikely to directly contribute to the exacerbation of the COVID-19 pandemic.

Another study, by Chernozhukov et al. 2021, evaluated the dynamic impact of various policies adopted by U.S. states on COVID-19 growth rates and social distancing behavior, utilizing a causal structural model framework. Their analysis highlighted the significant role of policies and behavioral responses in determining the number of COVID-19 cases and deaths. For instance, their counterfactual experiments suggested that early national mask mandates could have substantially reduced the growth rate of cases and deaths, potentially saving thousands of lives.

Based on our search, we can conclude that much research has been done in the field of finding the relationship between various COVID-19 aspects and one or two different factors, but no one has gone so far as to find correlations in terms of comprehensive One Health data, i.e., between COVID-19 aspects with all the other factors—human, animal, weather, and geospatial data along with Google Trends data.

2.3 Visual Analytics in Public Health

Understanding how our framework fits within the broader landscape of visual analytics and epidemiological modeling requires examining several interconnected research areas.

2.3.1 Visual Analytics for Health and Epidemiology

Modern visual analytics has evolved beyond simply creating attractive charts and now focuses on supporting human reasoning throughout the entire analytical process (Andrienko et al. 2018; Sacha et al. 2014). In health applications, this means helping domain experts explore data, build models, and validate findings in ways that leverage their expertise while benefiting from computational power. For instance, Bryan et al. 2015 developed spatiotemporal forecasting systems that integrate predictive analytics with geographic simulation, demonstrating that visual tools can support real-time decision-making during pandemics. However, significant challenges remain. During health emergencies, many visualization tools fail to align with users' actual needs, often contributing to cognitive overload rather than supporting clear decision-making (Rind et al. 2013).

2.3.2 Interactive Systems for Epidemiological Modeling

Modern epidemiological modeling increasingly emphasizes collaboration between human experts and computational tools. Stolper et al. (2014) pioneered progressive visual analytics approaches that enable users to interact with algorithms as they operate, allowing users to adjust the analysis direction based on intermediate results. This concept has yielded valuable insights in health applications, where model assumptions must be rapidly adjusted as new information becomes available. Stolper et al. (2016) extended this work by proposing mixed-initiative systems where both users and algorithms take turns leading the analysis. Their framework includes intelligent data preparation, automatic visual suggestions, and system-guided exploration while maintaining user control. Several successful applications demonstrate the potential of these approaches. Reinert et al. (2020a) developed PanViz 2.0 as a pandemic modeling tool that integrates epidemiological simulation with interactive visualization, demonstrating how public health decision-makers can evaluate intervention strategies through visual interfaces. The collaborative approach is also reflected in systems like TimeFork (Badam et al. 2014) and frameworks based on *opposite models* (Krak et al. 2022), where users guide model construction through interactive visual interfaces. The current study employs this approach to building mixed-effects models.

2.3.3 One Health Data Integration and Visualization

The One Health paradigm presents unique challenges for visual analytics. One major challenge is in integrating human clinical data with environmental and animal surveillance information at appropriate spatial and temporal scales (Mubareka et al. 2023; Faijue et al. 2024; Mukherjee et al. 2023). For instance, Mukherjee et al. 2023 investigated approaches for integrating veterinary syndromic data with human health indicators that show potential for cross-species disease surveillance. However, many existing systems treat different data sources as separate analytical streams rather than truly integrated inputs to unified models, as presented by Wimberly et al. 2022 and Kamalrathne et al. 2023. Our framework addresses this gap by incorporating One Health data sources directly into mixed-effects modeling workflows.

2.3.4 Evaluation and Validation in Health Visual Analytics

Rigorous evaluation remains a critical challenge in health visual analytics (Kim et al. 2024; Rind et al. 2013). Betz et al. (2023) presented the *ESID: Epidemiological Scenarios for Infectious Diseases* tool, which identifies common evaluation gaps, including insufficient attention to real-world usability and limited assessment of decision support effectiveness. Recent work has begun developing more robust evaluation approaches. Rydow et al. (2022) introduced visual sensitivity analysis methods that help researchers understand how parameter variations influence model outputs—directly relevant to our context of mixed-effects modeling. Xu et al. (2021) developed a platform that utilizes advanced visualizations, including textured-tile calendars and multivariate rose charts, to compare model predictions with actual data across space and time. Case studies reveal that the platform can uncover prediction uncertainty error trends in barely populated areas.

2.4 Mixed-Effects Models for Disease Forecasting

Our framework combines traditional statistical approaches with modern computational methods. Visual analytics can enhance every stage of the machine learning pipeline, from data quality assessment to model interpretation, which is particularly important for mixed-effects models where understanding both population-level and group-specific patterns is crucial (Yuan et al. 2021b). Recent technical developments have made mixed-effects modeling more accessible and powerful. Sokhansanj and Rosen (2022) demonstrated how gradient-boosting methods can integrate with random effects to address geographic and temporal biases. Similarly, our approach uses mixed-effects models with month or age group as random effects to account for latent temporal or demographic variations. This allows us to model systematic heterogeneity in the data—such as seasonal effects or age-specific disease dynamics—thereby improving prediction accuracy and generalizability across counties. The integration of feature selection with mixed-effects approaches has also advanced significantly. Visual analytics platforms, such as *Feature Engineering Visualization (FeatureEnVi)* presented by Chatzimparmpas et al. 2022, combine automated feature selection with user-guided exploration. This interactive feature engineering enables domain experts to leverage their knowledge while benefiting from algorithmic optimization. This approach is particularly beneficial in One Health datasets that encompass diverse variables with complex relationships.

Our framework addresses several key gaps identified in these works. While existing visual analytics approaches have demonstrated value in epidemiological contexts, few systems effectively integrate One Health data sources within unified analytical workflows. Current mixed-effects modeling tools often lack the interactive capabilities needed for rapid model development and validation during health emergencies. Most importantly, existing approaches rarely combine the interpretability of traditional statistical methods with the interactive capabilities needed for practical public health decision-making.

Chapter 3

One Health Systems and Visual Strategies

In this chapter [1], we investigated how One Health systems, integrated digital platforms combining human, animal, and environmental health data, are currently designed and implemented for infectious disease detection and management. We identified integration patterns, functional purposes, and user interactivity across 202 reviewed articles published between 2015 and 2024, spanning various databases, including IEEE Xplore, ScienceDirect, PubMed, and Google Scholar, as well as other sources.

3.1 Material and methods

We followed the steps in the *Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)* guidelines (Page et al. 2021a) to conduct this literature review. As shown in Figure 3.1, the systematic review process started with an initial search on IEEEXplore, ScienceDirect, PubMed, and Google Scholar using relevant keywords. After selecting relevant papers from these databases, we followed the *backward citation chasing* approach by

[1] This chapter is based on the published paper in the One Health journal: A. Basheer, M. Tran, B. Khan, W. Jentner, A. Wendelboe, J. Vogel, K. Kuhn, M. Wimberly, and D. Ebert, "Comprehensive Review of One Health Systems for Emerging Infectious Disease Detection and Management," One Health, 2025.

looking at the reference lists of selected papers to find earlier relevant studies (cited articles) (Haddaway et al. 2022). Therefore, we examined the references cited in those papers ("other sources") to identify additional studies. Next, we applied specific criteria to decide which papers could be included in the review. After that, we categorized the selected papers based on various factors, including their purpose, the data they included, and the diseases they addressed. Finally, we concluded the process by analyzing the search statistics to synthesize the results.

3.1.1 Search strategy

A keyword search was conducted using IEEEXplore, ScienceDirect, PubMed, and Google Scholar, with the primary search strategy being "One Health Dashboard Visualization System," "One Health Dashboard," or "One Health System." To refine the results, only papers published between 2015 and 2024 (inclusive) were included. Additionally, relevant citations from review papers identified through this search were also considered, provided they pertained to One Health systems. Papers were also gathered from the Surveillance Outbreak Responses Management and Analysis (SORMAS) website (SORMAS Foundation 2024), specifically those related to the SORMAS system. The focus was primarily on research articles, with review papers being collected to a limited extent.

3.1.2 Inclusion and exclusion criteria

The inclusion criteria for the study were as follows: Systems were required to exhibit a *One Health Focus*, integrating data from at least one of these categories: human health, animal health, and environmental health. An *Infectious Disease Focus* was necessary, targeting the detection, prediction, planning, mitigation, or monitoring of infectious diseases, including the genetic evolution of pathogens, mutations, or species jumping. Systems had to be *Data-Driven*, utilizing datasets such as patient records, veterinary data, or environmental and climate data. In addition, they incorporated *Visualization*, employing techniques like graphs,

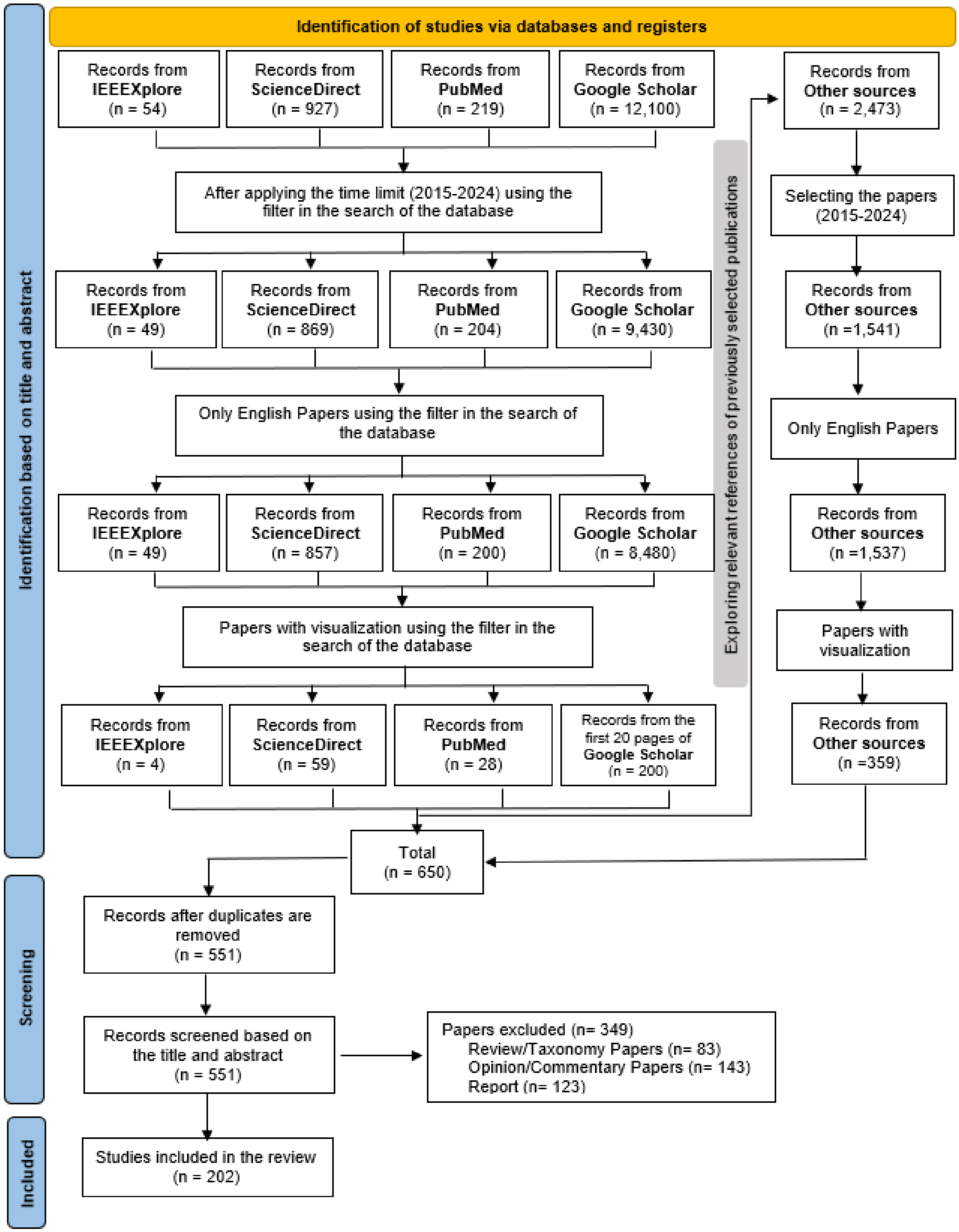


Figure 3.1: The PRISMA flow diagram

charts, geospatial maps, network graphs, time-series plots, and hierarchical charts to convey information. Lastly, the systems offered various levels of *User Interaction*, ranging from high to low interactivity, and were required to be based on research publications from the last 10 years (2015–2024) to ensure current relevance.

The exclusion criteria for this study were applied in multiple stages as illustrated in Figure 3.1. First, we removed papers published outside the targeted timeline (2015–2024). Next, we excluded papers not published in English. Following this, we excluded papers without a visualization component. Lastly, theoretical papers that described frameworks or proposals without demonstrating a functional system or prototype were not included, nor were bibliographic review papers that simply overviewed previous publications.

3.1.3 Classification and Data Extraction

This systematic review comprehensively analyzed existing systems and dashboards that utilize One Health datasets for infectious disease management, as shown in Figure 3.2. The review aimed to identify, categorize, and evaluate key characteristics of these systems, focusing on their purpose, data types, user groups, interactive functionalities, visual representations, analytical techniques, and implementation status.

As shown in Table 3.1, each feature in the review framework was defined as a distinct dimension for characterizing disease surveillance and response systems, while the categories within each feature describe the specific options along that dimension. For example, the *Purpose of the System* feature specifies the overarching goal the system is designed to achieve in the public health domain. Categories under this feature include: *Detection Systems* that emphasize early identification of unusual events or anomalies that may signal an outbreak (Henning 2004; Buehler et al. 2004); *Prediction/Forecasting Systems*, which aim to anticipate disease trends, outbreaks, or pandemics (Shmueli and Koppius 2011); *Planning Systems*, which support decision-making for future preparedness (Turoff et al. 2009); *Mitigation/Response Systems*, which are focused on strategies to reduce the impact or respond to outbreaks

Figure 3.2: We classified the collected papers in terms of the purpose of the system, diseases, dataset type, users, tasks, visual representations, visual analytics techniques, interactive analysis ability, system or dashboard, and availability

(Hollingsworth et al. 2011); and *Monitoring Systems*, which continuously track disease patterns and public health indicators over time to provide situational awareness (World Health Organization 2006; Buehler et al. 2004). While both detection and monitoring involve surveillance, detection systems are distinguished by their emphasis on early warning and anomaly detection, whereas monitoring systems provide ongoing observation of disease dynamics.

Another essential categorization involves the type of *diseases* or range of diseases targeted by the system. This includes a range of infectious and non-infectious diseases, such as emerging and re-emerging diseases, zoonotic diseases, and vector-borne diseases. Identifying the disease focus helps assess the applicability and specificity of a system in addressing different public health challenges (Jones et al. 2008).

The *datasets* utilized by these systems were classified into several types. *Human Data*, includes patient records, health information, and human mobility data; *Animal Data*, encompasses veterinary data, wildlife tracking, and information on zoonotic diseases; *Environmental Data*, consists of climate data, geospatial data, and other environmental factors; and *Integrated Data*, which combines multiple datasets, including Human, Animal, and Environmental data (Who 2017).

The intended *users* of the system include a broad range of stakeholders, such as public health officials, healthcare providers, researchers, veterinarians, policymakers, and the general public. Understanding the user base is critical for evaluating the system's usability and impact on decision-making processes (Petkovic et al. 2023; Alderwick et al. 2021).

Users engage in several *tasks* with the system. These include *Data collection*, which involves gathering information on the spread of diseases or health conditions; *Analysis*, where users perform statistical and epidemiological analyses; *Data exploration*, which includes the use of visual tools to interpret trends and relationships; *Decision-making*, that leverages system outputs to inform public health policies and responses; and *Collaboration*, that enables users to share data, insights, and reports with other stakeholders. While *Purpose of the System* describes the system's ultimate function in the public health domain, *Tasks* capture the concrete activities users undertake within the system. Importantly, the same task (e.g., decision-making) may occur across different purposes, though the context and intended outcomes differ.

Visual Representations play a central role in these systems, facilitating data interpretation and communication. Common types include *Graphs and charts*(e.g., line, bar, and pie charts) for quantitative comparisons; *Geospatial maps*(e.g., heatmaps and choropleth maps) for spatial pattern recognition; *Network graphs* to illustrate the relationships between variables or events; *time-series plots* for trend analysis over time; and *Dashboards* serving as integrated platforms that showcase multiple visual components simultaneously, enabling users to gain insights at a glance. Lastly, *Hierarchical charts*, such as sunburst charts and treemaps, are effective for exploring data organized in a hierarchical structure.

Visual analytics techniques are essential for analyzing and interpreting visual data. Visual analytics is "the science of analytical reasoning facilitated by interactive visual interfaces" (Keim et al. 2008). It supports human decision-making by integrating automated analysis with visualization, enabling the exploration of complex, multivariate data (Al-Hajj 2014). Various methods are employed within this framework, starting with *statistical and analytical approaches* such as *Descriptive statistics* for summarizing data distribution, *Statistical analysis* involves hypothesis testing for validating patterns, *Regression models* for predictive and explanatory purposes, and *Cluster analysis* is used to detect patterns or hotspots of disease incidence.

Spatial and geospatial analysis is another key area that visualizes data across geographic regions to identify hotspots or areas of impact. Techniques such as heatmaps and choropleth maps represent data density or frequency over geographical locations, aiding in the visualization of disease spread. *Geospatial models* incorporate geographical variables to analyze spatial relationships, which are useful for examining disease vectors or environmental factors. *CARTO map* tiles further facilitate interactive exploration of geospatial data.

Data integration and processing techniques are essential for system performance. Simplifying the data collection process helps consolidate information from various sources for real-time analysis.

Automatic data entry enhances the efficiency of this process, while frameworks and indices provide standardized indicators or scores for comparing regions or health statuses. *Data normalization and scaling* ensure consistency across diverse datasets, which ultimately improves the comparability of metrics.

The use of *machine learning and AI* has become invaluable in the realm of visual analytics as well. These systems utilize *Supervised and unsupervised learning* to perform classification, clustering, and anomaly detection in health data. *Deep learning* is employed to process large datasets and complex patterns, such as those used in image recognition in medical diagnostics. *Predictive analytics* helps forecast future disease trends, supports early warning systems, and enhances response planning.

In addition to these techniques, other specialized methods, such as simulation models, optimization algorithms, and interactive dashboards, enhance user engagement and decision-making, ultimately leading to more effective health analytics.

The *interactive analysis capability* refers to the extent to which users can engage with the data. In systems with *high interactivity*, users have full control to manipulate and explore the data through features such as drill-down, zoom, filters, and dynamic updates. *Moderate interactivity* allows for limited exploration, providing some filters and options to adjust views. In contrast, *low interactivity systems* primarily offer static visualizations, giving users minimal control over the data.

Finally, systems are categorized by their implementation status and accessibility. The System/Dashboard category assesses whether the study presents a functional system. A "Yes" indicates a working prototype or operational platform, while a "No" denotes a theoretical or conceptual discussion. The *online availability* category evaluates whether the system or dashboard can be accessed online by external users, which is crucial for reproducibility, transparency, and practical application.

3.2 Results

In this section, we present the findings of our systematic review. We first describe the study selection process (3.2.1) and summarize the characteristics of the included studies (3.2.2). Within the study characteristics, we examine nine key features of the systems under review, including their purposes, targeted diseases, datasets used, intended users, tasks supported, visual representations, analytical techniques, interactivity levels, and overall system availability. We then analyze the relationships within and between these features, highlighting co-occurrence patterns and overlaps that reveal trends and gaps in the current literature. Figures 3.3–3.16 and Table 3.1–3.2 provide visual and tabular summaries of these findings.

3.2.1 Study selection

As shown in Figure 3.1, the systematic review process began with an initial search across IEEEXplore, ScienceDirect, PubMed, and Google Scholar using relevant keywords. After selecting pertinent papers from these databases, we employed a backward citation chasing approach by examining the reference lists of those papers to identify additional relevant studies. These cited papers ("other sources") were then subjected to the same screening and selection process.

Initially, our search yielded approximately 13,300 papers from the four databases. Applying a time filter to include only papers published within the last 10 years (2015–2024) reduced the number to 10,552. Limiting the results to English-language papers further reduced the total to 9,586.

We then filtered for papers containing visualizations, resulting in 91 papers from IEEEXplore, ScienceDirect, and PubMed. From Google Scholar, we reviewed the first 20 result pages (200 papers), giving a total of 291 papers. We examined the reference lists of these 291 papers using backward citation chasing, identifying 2,473 additional papers. After applying the same filtering and selection criteria, this number was narrowed to 359.

Combining the 291 initially selected papers with the 359 identified through reference chasing resulted in 650 papers. After removing duplicates and further screening based on inclusion and exclusion criteria, we finalized a set of 202 papers for the review.

3.2.2 Study characteristics

In this section, we provide an overview of the characteristics of the studies included in our analysis, as illustrated in Figure 3.15. There are nine features, each with sub-categories, and some sub-categories have further divisions, as shown in Table 3.1. In the following sections, we will investigate the relationships within each feature, between features, and among the subcategories of all the features.

Features characteristics

In terms of *Purpose of the System*, we found that the most common purpose of the papers is monitoring, accounting for 56% of the studies, as seen in works such as (Olugasa and Fasunla 2020; Curriero et al. 2021b; O'Brien and Xagoraraki 2019; Gustafsson et al. 2023b; Sun et al. 2024). This is followed by planning, which constituted 38% of the studies, including research (Yadav et al. 2023; Di Lorenzo et al. 2023b; Al-Hemoud et al. 2021a; noa 2024). Detection was noted in 33% of the papers, as highlighted in (Timme et al. 2020; Dalisay et al. 2024; Lam et al. 2024). Mitigation and response accounted for 18% of the studies, including (Tegegne et al. 2023), while prediction and forecasting were less emphasized, appearing in only 6% of the papers (Di Lorenzo et al. 2023b).

The analysis of the system purposes reveals that many papers serve multiple functions, as shown in the Figure 3.3. Monitoring frequently overlapped with detection (18 papers) and planning (17 papers), indicating that systems designed to observe phenomena often also aim to identify issues and support decision-making. Similarly, detection and planning co-occurred in 14 papers, highlighting the close link between identifying events and preparing responses. In contrast, prediction/forecasting and mitigation/response were less often combined with other purposes, suggesting these are more specialized or standalone functions. To illustrate these overlaps, we generated a heatmap of the co-occurrence matrix, which visually highlights the strong associations among monitoring, detection, and planning. Overall, while some systems focus on a single purpose, many integrate multiple objectives to provide comprehensive support for One Health initiatives.

The majority of systems focus on Emerging and Re-emerging Diseases such as COVID-19 52% (Di Lorenzo et al. 2023b; Zyoud 2024; Fasina et al. 2021; Sun et al. 2024; Ormea et al.

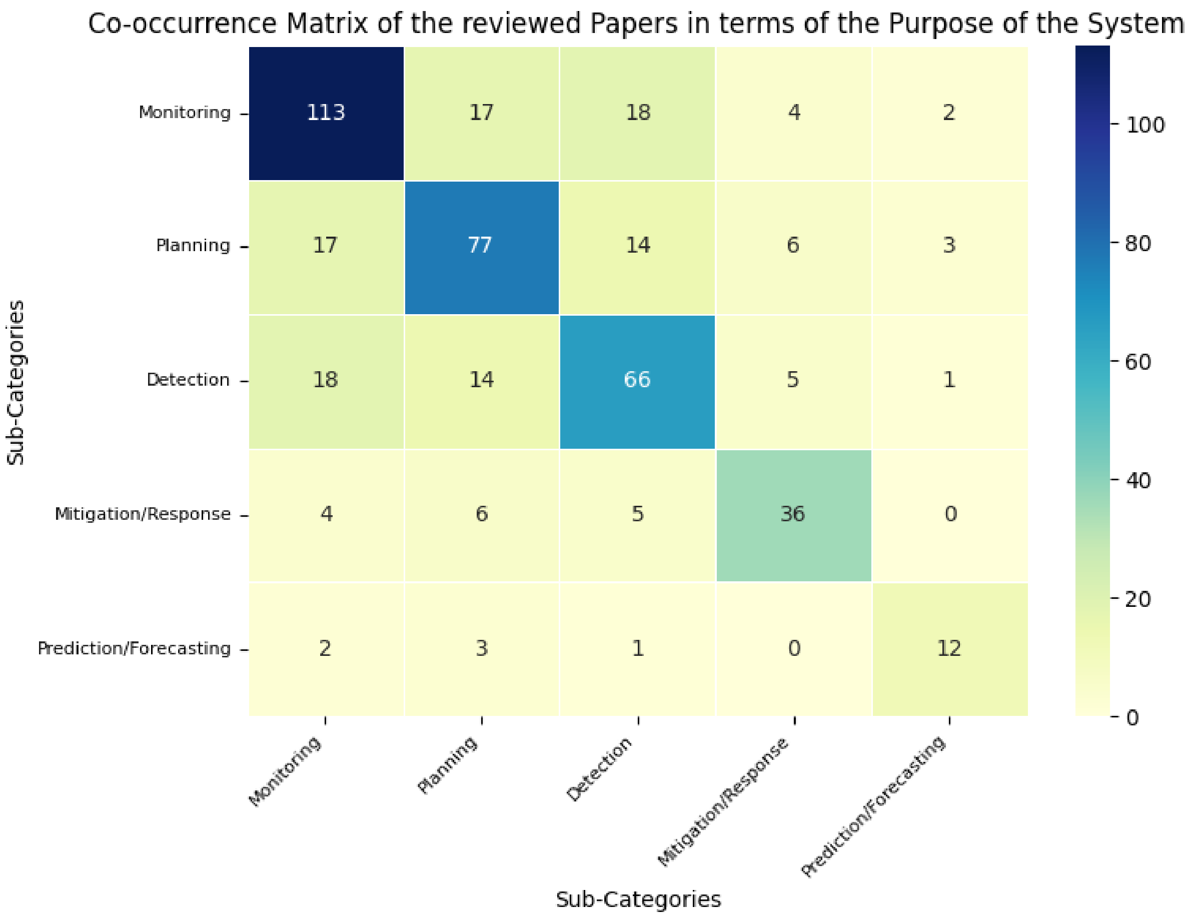


Figure 3.3: Co-occurrence Matrix of the reviewed Papers in terms of the Purpose of the System.

2023). Other significant areas include Zoonotic Diseases 37% (Behravesh et al. 2024; Ho 2022), Vector-borne Diseases 15% (Nyatanyi et al. 2017), Neglected Tropical Diseases 10% (Kahariri et al. 2024), Food Safety and Waterborne Diseases 7% (Osman et al. 2023), and Antimicrobial Resistance (AMR) 11% (Timme et al. 2020).

Analysis of the co-occurrence matrix (Figure 3.4) reveals that Emerging and Re-emerging Diseases often overlapped with Zoonotic Diseases (51 papers) and, to a lesser extent, with Vector-borne Diseases (16 papers) and Neglected Tropical Diseases (11 papers). Similarly, Zoonotic Diseases co-occurred with Vector-borne Diseases (25 papers), underscoring their interconnection through shared vectors and transmission pathways. In contrast, areas such as AMR and food safety/waterborne diseases showed fewer overlaps, indicating that these

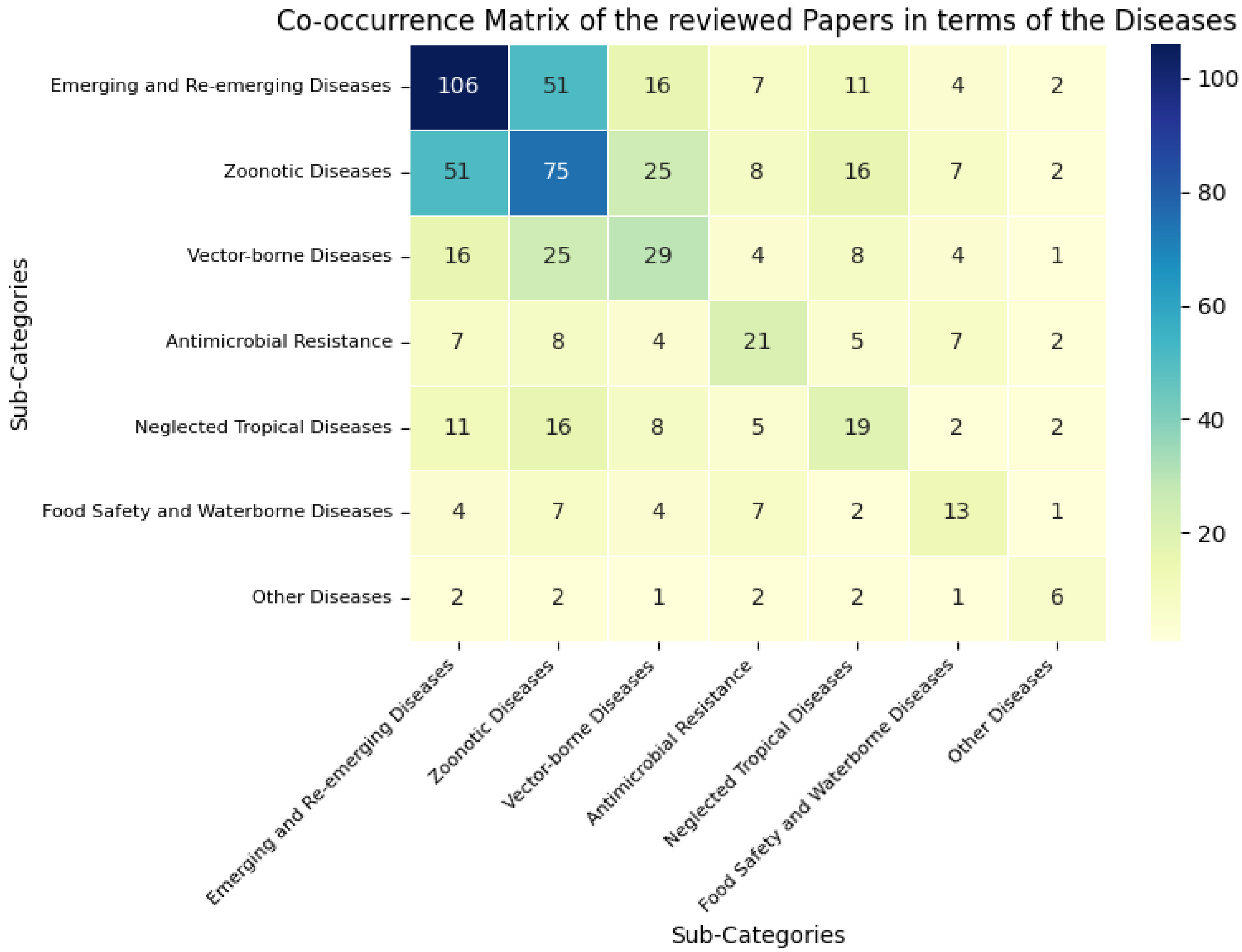


Figure 3.4: Co-occurrence Matrix of the reviewed Papers in terms of the Diseases.

remain more specialized or narrowly scoped within the current body of One Health literature. Also, we visualized the relationships between disease groups using an overlap network (Figure 3.5).

Among the studies analyzed, most systems relied on single-domain data, with Human data used in 20% of the studies, like in (Louw et al. 2022; Kassim et al. 2020; Di Lorenzo et al. 2023a), Animal data in 10% of the studies, such as (Al-Hemoud et al. 2021b; Gustafsson et al. 2023a; Raineri et al. 2024), and Environmental data in 8% of the studies, such as (Neves et al. 2023; Shrestha et al. 2020). Integrated datasets were common with Human–Animal combinations appearing in 28% of the studies, like in (Bridget et al. 2024; European Food Safety Authority (EFSA) et al. 2023; Cossaboom et al. 2022), Human–Environmental 6%,

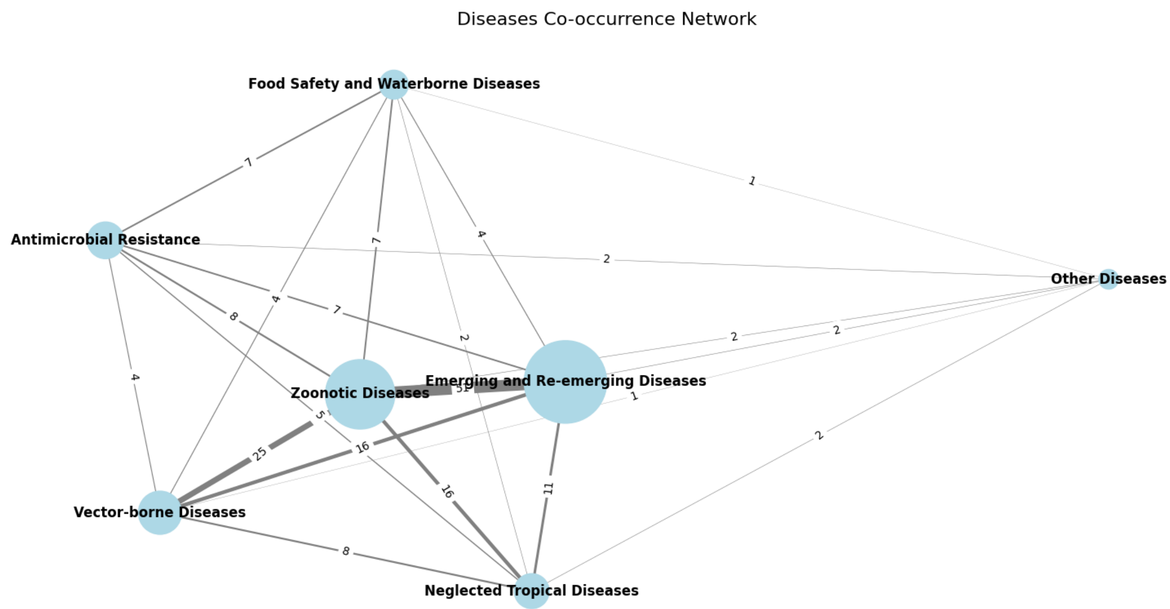


Figure 3.5: The overlap network for the diseases mentioned in the reviewed papers. The nodes represent different disease groups, while the edges indicate the frequency of their co-occurrence in various systems. The thickness of the edges represents the strength of the connections, and the node size indicates the number of papers associated with each node.

like in (Curriero et al. 2021a), Animal–Environmental where just 2%, such as (Matere et al. 2020a), and full integration across Human, Animal, and Environmental domains around 11% of the studies, such as (Das et al. 2024). These counts indicate that while single-domain systems are still prevalent, integration, particularly between Human and Animal datasets, is increasingly adopted. Some of the other papers provide general system descriptions without specifying the types of datasets used. To illustrate the patterns of data usage and integration, we generated a heatmap of the co-occurrence matrix (Figure 3.6), highlighting both partial and full multi-domain overlaps.

The primary users of the system are Public Health Officials, comprising 53% of the user base. Significant studies supporting this group include (Coetzer et al. 2019; Bisson et al. 2015; Matere et al. 2020b; Kassim et al. 2020; Ghai et al. 2022; Kawakyu et al. 2024; Iftikhar et al. 2023).

Additionally, the General Public makes up 27% of users (Silenou et al. 2022; Innes et al.

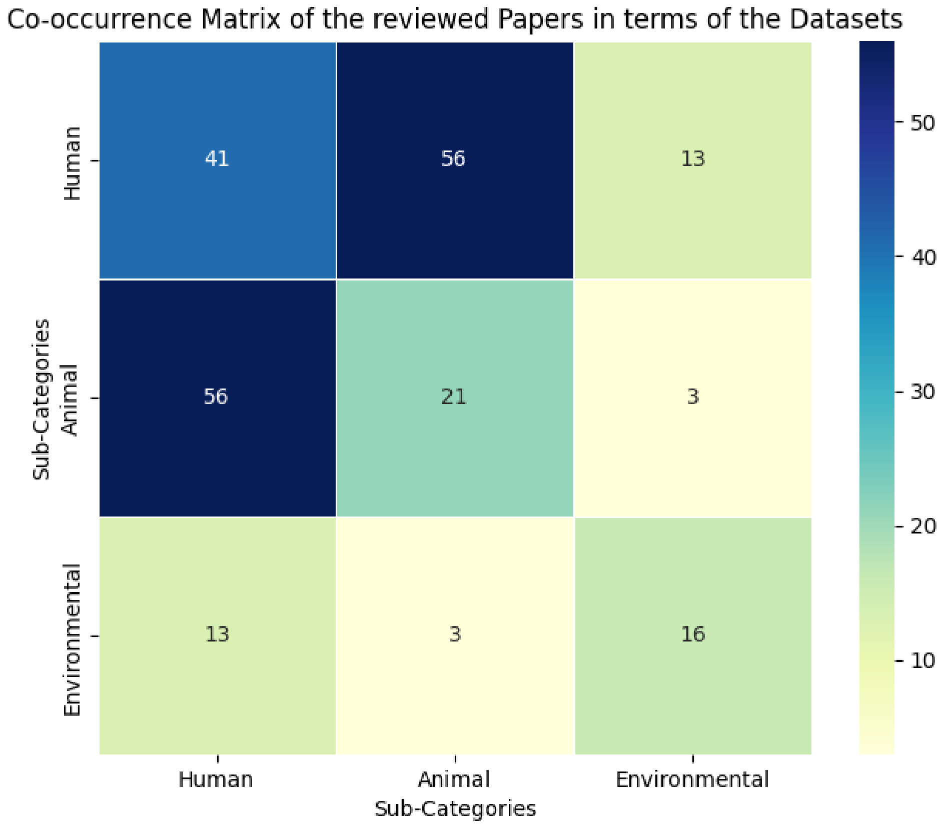


Figure 3.6: Co-occurrence Matrix of the reviewed Papers in terms of the Datasets Type.

2019; Xiao and Zhang 2023; Divi et al. 2024; Nyatanyi et al. 2017; Abd-alrazaq et al. 2023), and Veterinarians constitute 13% (Tan et al. 2023).

The analysis of user groups shows that many systems are designed for multiple audiences rather than a single stakeholder type. Public Health Officials, the largest group, frequently overlap with Policymakers (38 studies) and Researchers (36 studies), underscoring the strong alignment between surveillance, decision-making, and scientific inquiry. They also co-occurred with the General Public in 34 studies, suggesting that systems aimed at official use are often also adapted for public-facing communication. Researchers overlapped with Policymakers in 24 studies, reflecting the role of evidence-based research in guiding policy. Veterinarians, while a smaller group overall, still co-occurred with Public Health Officials (17 studies) and Researchers (13 studies), highlighting their contributions in zoonotic and animal-related contexts. In contrast, overlaps involving the General Public and Veterinarians

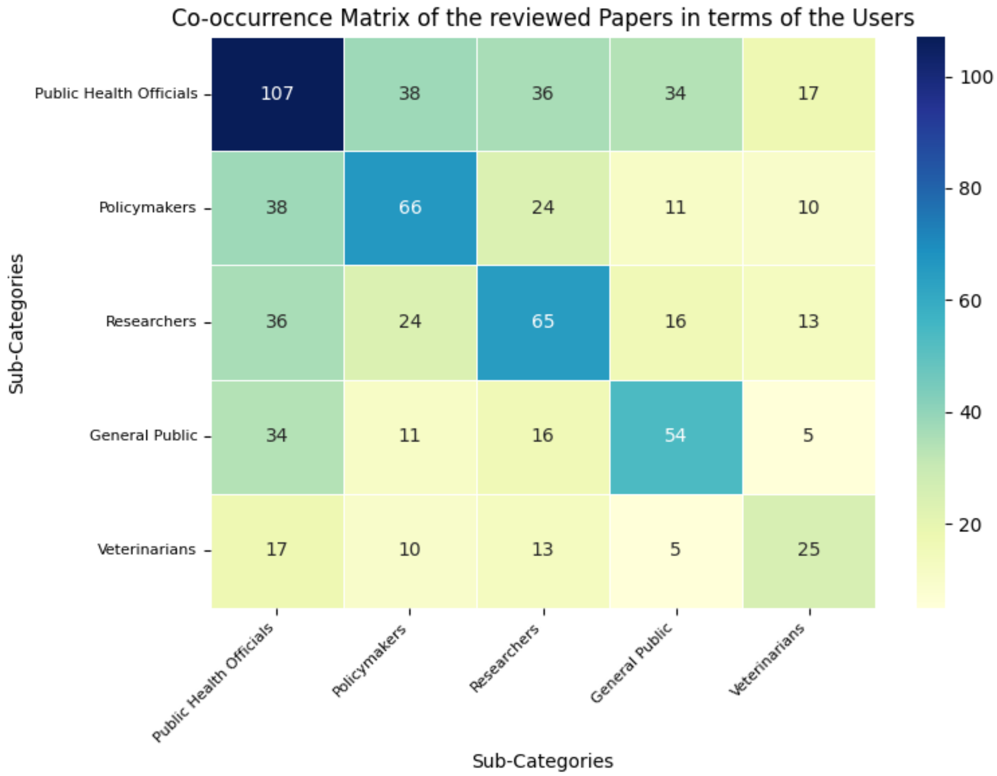


Figure 3.7: Co-occurrence Matrix of the reviewed Papers in terms of the Users.

were relatively rare (5 studies), pointing to a gap in systems that directly connect veterinary insights with lay audiences. To illustrate these relationships, we generated a heatmap of the co-occurrence matrix (Figure 3.7), which highlights the particularly strong ties between Public Health Officials, Policymakers, and Researchers.

Also, we visualized the relationships between user groups using an overlap network (Figure 3.8). In this representation, the nodes represent different user groups, while the edges indicate the frequency of their co-occurrence in various systems. The thickness of the edges illustrates the strength of the connections— for instance, the link between Public Health Officials and Policymakers is strong, with 38 studies highlighting their connection. In contrast, the connection between Veterinarians and the General Public is weaker, supported by only 5 studies.

Analysis is the dominant task 84% as highlighted in (Al-Hemoud et al. 2021a; Curriero et al. 2021b; Chen et al. 2024; Arsenault et al. 2024; Hanifah et al. 2022; Neves et al. 2023;

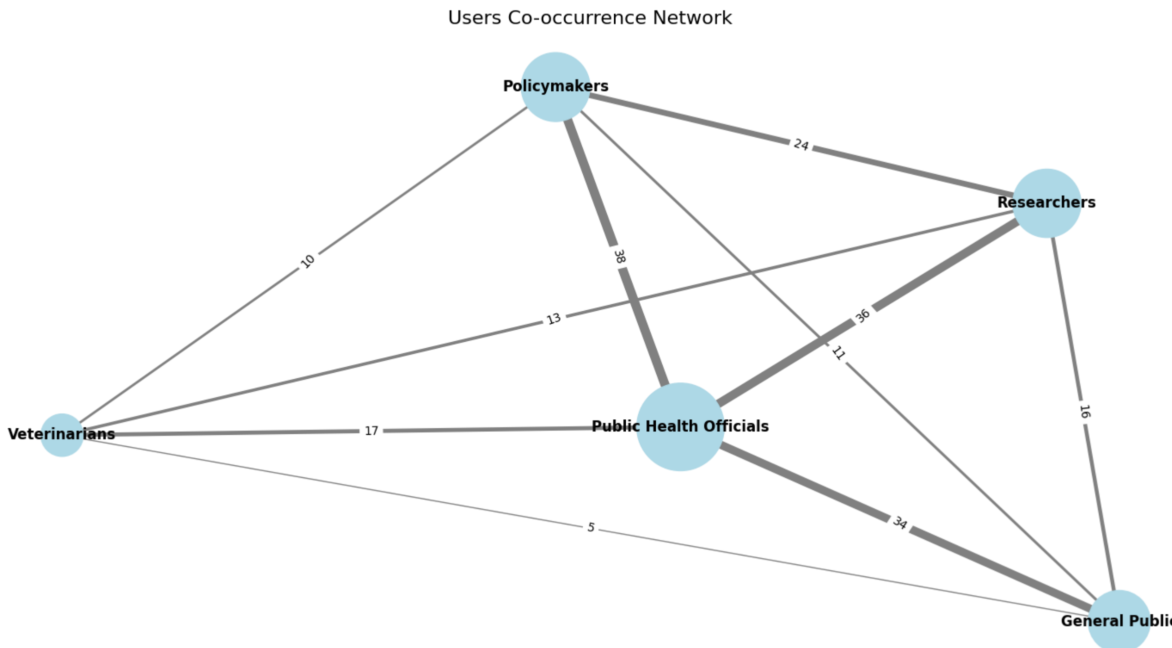


Figure 3.8: The overlap network for the users mentioned in the reviewed papers. The nodes represent different user groups, while the edges indicate the frequency of their co-occurrence in various systems. The thickness of the edges represents the strength of the connections, and the node size indicates the number of papers associated with each node.

Kawakyu et al. 2024; Louw et al. 2022; Silenou et al. 2022; Kassim et al. 2020), with other tasks such as Data Collection 30% (Yadav et al. 2023; Di Lorenzo et al. 2023b; Olugasa and Fasunla 2020), Exploration 22% (Sun et al. 2024; Pedretti et al. 2022; Lushasi et al. 2020), Decision Making 21% (noa 2024; Cossaboom et al. 2022), and Collaboration 18% (Cruz et al. 2024; Kaur et al. 2022) playing supportive roles.

When examining the relationships among the main tasks supported by One Health dashboards, we observe a high degree of overlap between analysis and other tasks, particularly with data collection (54 studies), reflecting how analytical processes often depend on robust and continuous data gathering. Similarly, exploration frequently co-occurs with analysis (36 studies) and data collection (30 studies), suggesting that dashboards often serve as tools to both investigate and contextualize collected data. While decision-making is most commonly linked to analysis (35 studies) and collaboration (28 studies), its overall co-occurrence values are lower, indicating that fewer systems move beyond descriptive and exploratory

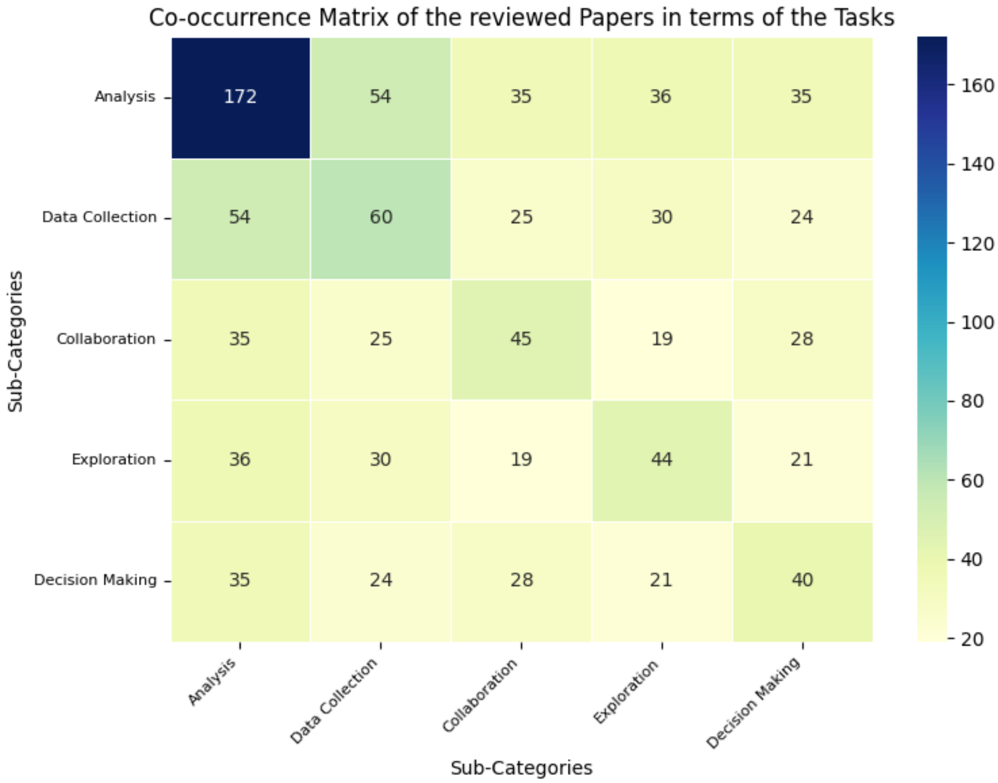


Figure 3.9: Co-occurrence Matrix of the reviewed Papers in terms of the Tasks.

tasks to explicitly support decision-oriented workflows. Interestingly, collaboration emerges in connection with all other tasks, most notably analysis (35 studies) and data collection (25 studies), highlighting that while not always the primary purpose, many systems are designed to facilitate multi-stakeholder engagement alongside technical functionality. These patterns are clearly visible in the Co-occurrence Matrix (Figure 3.9), which quantifies task pairings, and in the overlap network (Figure 3.10), which emphasizes the central role of analysis as a bridging function across dashboard purposes.

The review revealed that Graphs and Charts were the most prevalent visual format, used in approximately 64% of the surveyed studies (Al-Hemoud et al. 2021a; Curriero et al. 2021b; Arsenault et al. 2024; Pedretti et al. 2022). Geospatial maps were also prominent, appearing in 33% of the papers (Matere et al. 2020b; Vesterinen et al. 2019), reflecting the strong role of spatial analysis in One Health applications. Dashboards were identified in 20% of the studies (Yadav et al. 2023; Di Lorenzo et al. 2023b; Curriero et al. 2021b), emphasizing

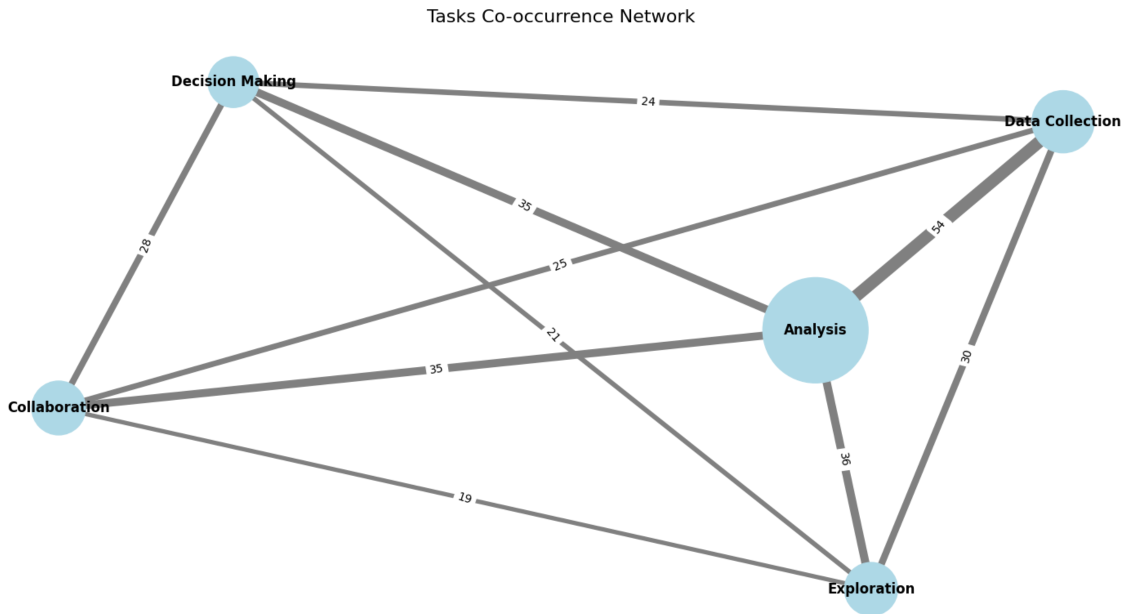


Figure 3.10: The overlap network for the reviewed papers tasks. The nodes represent different task groups, while the edges indicate the frequency of their co-occurrence in various systems. The thickness of the edges represents the strength of the connections, and the node size indicates the number of papers associated with each node.

integrative, interactive visual platforms. Meanwhile, network graphs accounted for 10% of the usage (Desvars-Larrive et al. 2024), and hierarchical charts such as treemaps or sunburst diagrams appeared in 4% of the papers (Lushasi et al. 2020). These findings highlight that while conventional charting and geospatial approaches currently dominate, integrative and structural representations are being increasingly adopted.

The co-occurrence analysis revealed several substantial overlaps between visualization categories, highlighting common combinations that researchers employ in practice (Figure 3.11). For example, Tables and Figures were frequently used together, with more than ninety papers reporting this pairing, underscoring their complementary role in presenting both structured data and explanatory diagrams. Similarly, Tables and Flowcharts (38 papers) and Tables and Dashboards (28 papers) appeared as strong overlaps, indicating that tabular representations often serve as a foundation for more complex or integrative visualizations. Geographic maps also showed notable co-occurrences, particularly in combination

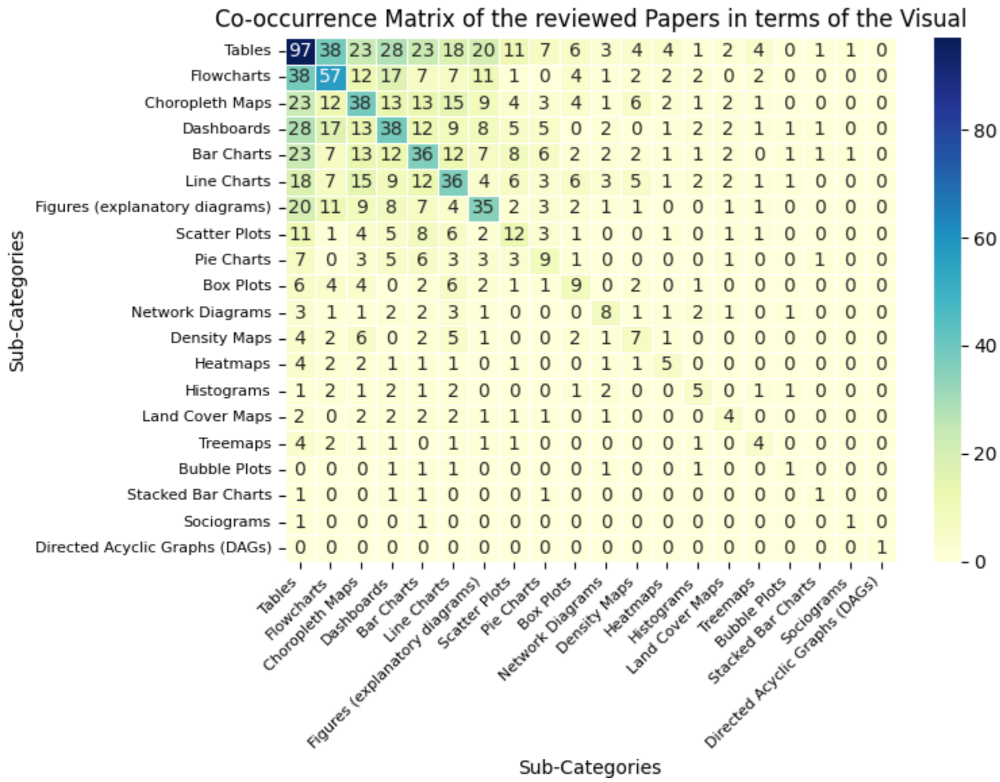


Figure 3.11: Co-occurrence Matrix of the reviewed Papers in terms of the Visual Representations.

with Line Charts and Choropleth Maps (15 papers), which reflects the integration of temporal and spatial perspectives in epidemiological studies. Beyond these dominant clusters, several less frequent but meaningful overlaps emerged, including Network Diagrams with Choropleth Maps and Scatter Plots with Tables, indicating specialized combinations employed in more targeted analytical contexts. These patterns are further illustrated in the chord diagram (Figure 3.12), which highlights not only the dominant clusters of visualization practices but also the nuanced links between less common pairs, showing how diverse visual forms are strategically combined to convey multifaceted insights.

The co-occurrence analysis of visual analytics techniques revealed not only frequently used methods but also meaningful overlaps between them (Figures 3.13 and 3.14). Statistical and Analytical Methods (35%) and Spatial/Geospatial Analysis (32%) (Yadav et al. 2023; Curriero et al. 2021b) were the most commonly applied approaches, often appearing

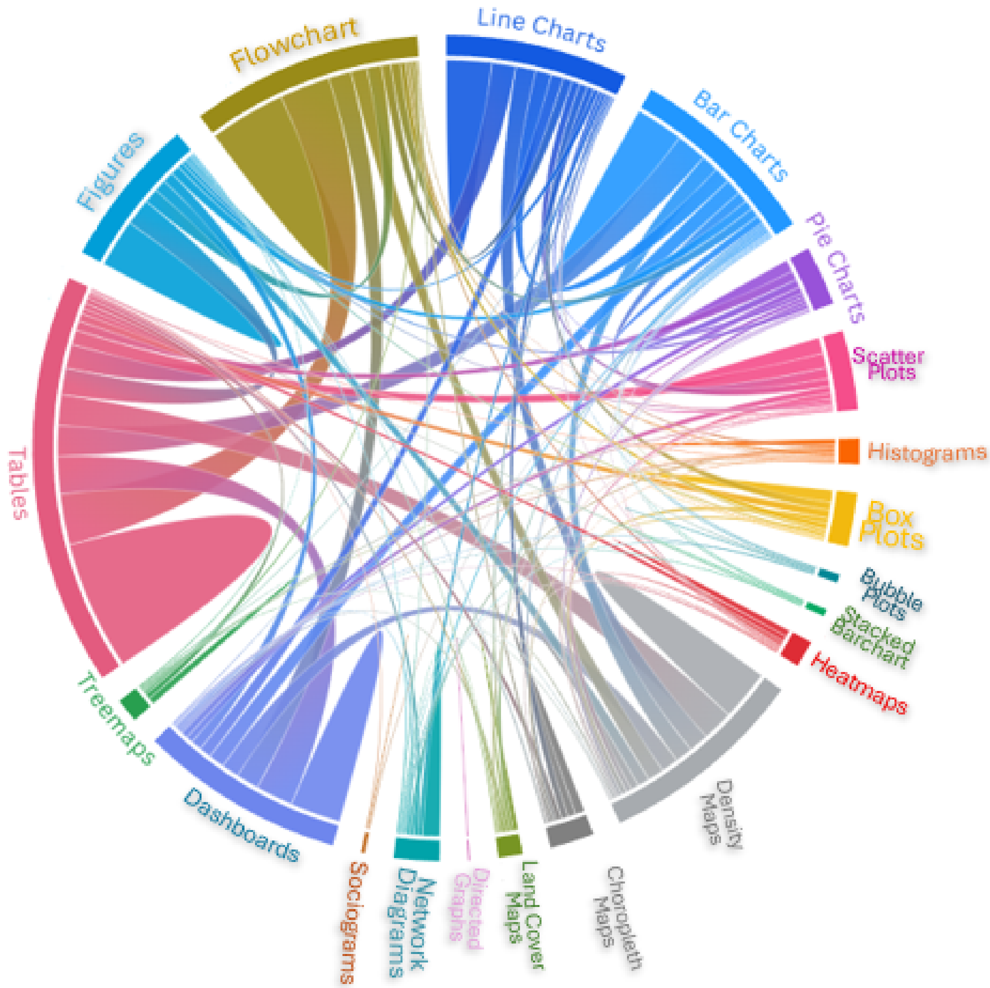


Figure 3.12: The chord diagram of the visual representations for the reviewed papers.

together in the same studies, reflecting the combination of traditional statistical evaluation with spatially explicit visualizations. Machine Learning and AI (16%) (Di Lorenzo et al. 2023b; Wang et al. 2024) also co-occurred with statistical and geospatial methods, indicating that predictive modeling and automated analytics are frequently integrated into broader analytical workflows.

Less dominant but still meaningful overlaps were observed between Statistical and Analytical Methods and Data Integration and Processing, highlighting the importance of structured pipelines for handling heterogeneous datasets. For example, Process Mapping overlaps moderately with geospatial models and indicator frameworks, while Deep Learning co-occurs with supervised and unsupervised learning, suggesting selective adoption in more advanced

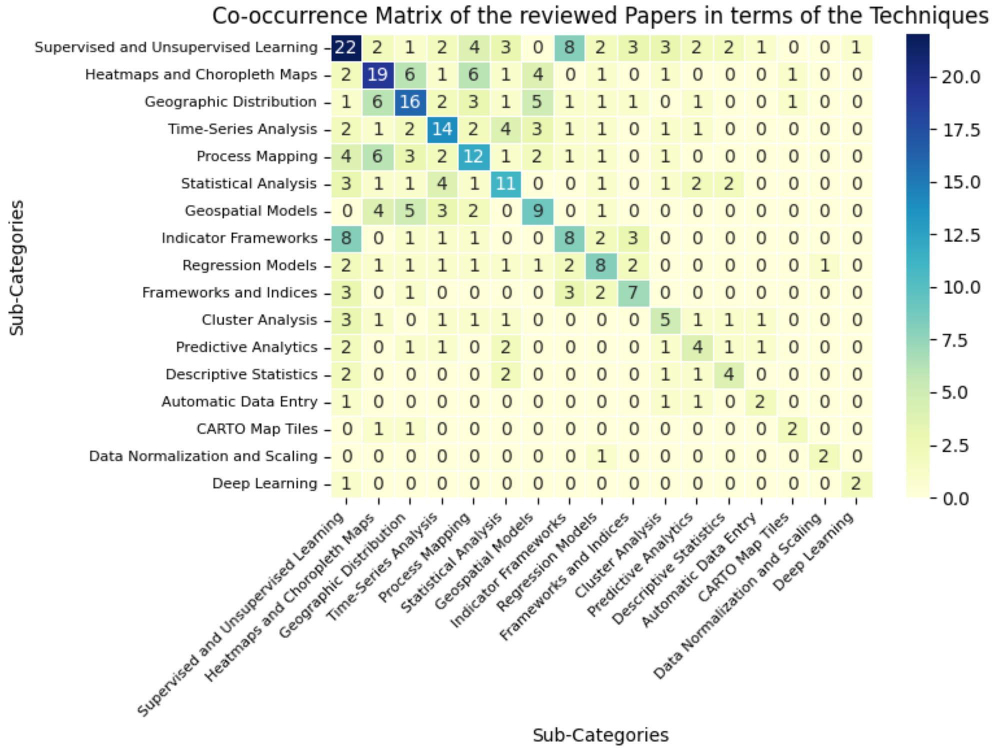


Figure 3.13: Co-occurrence Matrix of the reviewed Papers in terms of the Visual Analytics Techniques.

workflows. Similarly, techniques such as CARTO Map Tiles or Automatic Data Entry are primarily used in studies that emphasize data integration and operational pipelines. Overall, these patterns underscore clusters of methodological practices in which complementary techniques are strategically combined to support robust, multi-dimensional analyses in One Health and epidemiological research.

Most systems have Low to Moderate Interactivity 46% (noa 2024; Faradiba et al. 2024; Desvars-Larrive et al. 2024) and 38% (Coetzer et al. 2019; Hanifah et al. 2022; Pedretti et al. 2022), respectively), with fewer offering High Interactivity 16% (Neves et al. 2023).

Overall, the analysis indicates a strong focus on monitoring COVID-19 using integrated datasets with a wide range of users, emphasizing data analysis through various visual representations and moderate interactivity.

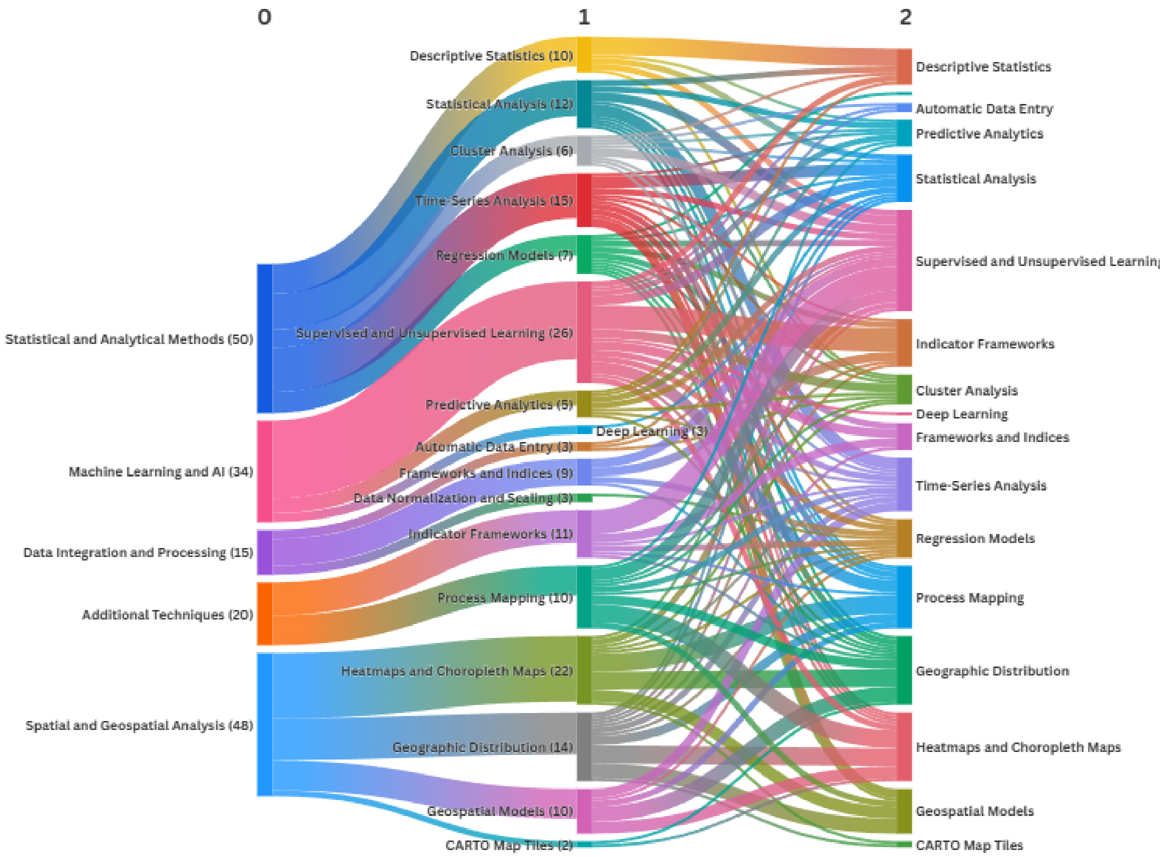


Figure 3.14: The chord diagram of the Visual Analytics Techniques for the reviewed papers.

Relationships between Features

Key observations from the matrix, illustrated in the heatmap in Figure 3.16, reveal clear patterns and potential clusters, as well as opportunities for further development.

- **Purpose of the System: Second-order relationships** (Purpose and Diseases): Detection systems focus heavily on emerging and re-emerging diseases (19 occurrences) and zoonotic diseases (13), while monitoring systems dominate for emerging/re-emerging diseases (42) and zoonotic diseases (33). Planning and mitigation/response systems are more widely distributed across diseases, but still exhibit moderate counts for human-dataset and emerging disease applications (20). Prediction/forecasting systems are rare, with a maximum count of 4. Vector-borne, neglected tropical, and food/waterborne diseases are often overlooked.

Figure 3.15: The categories of the classification are sorted by their respective percentages.

Third-order relationships (Purpose, Diseases, and Users): Detection systems mostly target veterinarians (6) for zoonotic diseases and public health officials (20) for emerg-

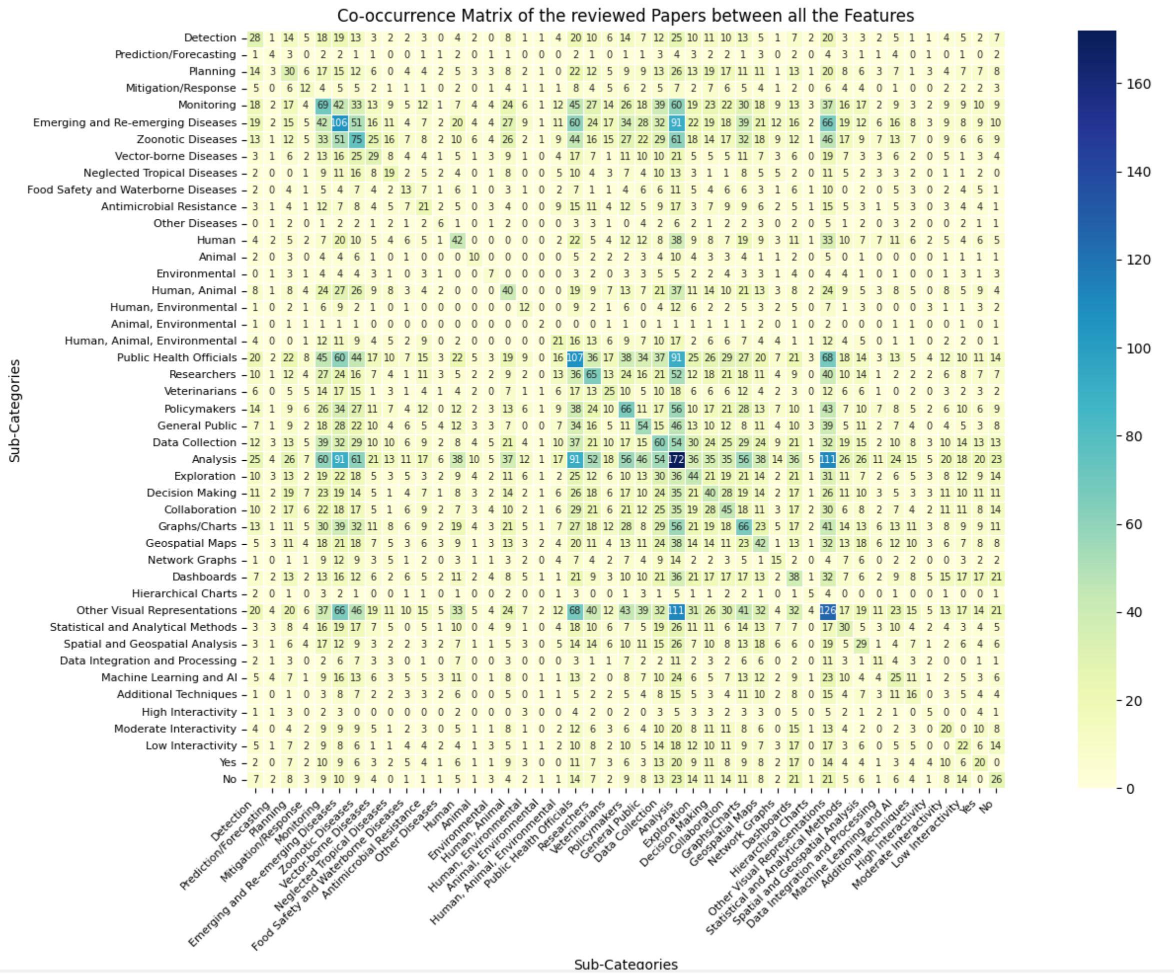


Figure 3.16: Relationships between Features.

ing diseases. Monitoring systems primarily serve public health officials (45), researchers (27), and, to a lesser extent, policymakers (26). Planning systems involve public health officials (22), policymakers (66), and researchers (36) across different disease types. Mitigation/response systems demonstrate moderate engagement with public health officials (8) and policymakers (17), but few connections exist with veterinarians or the general public.

Fourth-order relationships (Purpose, Diseases, Users, and Tasks): Monitoring systems for public health officials focus on analysis (91) and data collection (37), with moderate collaboration (29). Monitoring systems for researchers emphasize analysis (52), exploration (25), and decision-making (26). Detection systems targeting veterinarians prioritize data collection (12) with some analysis (6). Planning systems for policymakers support decision-making (66) and collaboration (35), underscoring their role in facilitating coordinated responses and resource allocation. Mitigation/response systems exhibit a mixed emphasis on analysis (7) and data collection (5), varying by disease type and user role.

Fifth-order relationships (Purpose, Diseases, Users, Tasks, and Visual Representations): Monitoring systems for public health officials utilize dashboards (36), graphs/charts (30), and geospatial maps (24) to support analysis-intensive workflows. Researchers' monitoring systems employ a diverse range of visualizations, including dashboards (21), graphs/charts (18), geospatial maps (20), and other visual representations (40), aligning with exploratory and decision-oriented tasks. Detection systems for veterinarians often use graphs/charts (6) to facilitate data collection. Planning systems for policymakers rely heavily on dashboards (9) and graphs/charts (9) to support decision-making and collaboration.

Sixth-order relationships (Purpose, Diseases, Users, Tasks, Visual Representations, and Interactivity): Public health monitoring systems exhibit moderate to high interactivity (20–25 occurrences), facilitating structured analysis and actionable insights. Research monitoring systems feature higher interactivity (25) to allow exploratory workflows. Detection systems for veterinarians tend to have low interactivity (6–7), reflecting the specialized and task-specific nature of these tools. Planning systems for policymakers also exhibit moderate interactivity (20–22), which is suitable for collaborative decision-making.

textbfSeventh-order and higher relationships (Purpose, Diseases, Users, Tasks, Visual Representations, Interactivity, Visual Analytics Techniques, and Availability): Monitoring systems for public health officials use statistical/analytical methods (91) with dashboards (36) and moderate interactivity (20), often available online (Yes: 20). Researchers combine machine learning/AI (24) with diverse visualizations (other visual representations: 40) and high interactivity (25), facilitating exploration of multi-domain datasets. Detection systems for veterinarians employ statistical methods (6–7) with specialized visualizations and limited interactivity (6–7), often operating offline. Planning systems for policymakers integrate dashboards (36) with analysis and decision-support tasks, available online (Yes: 17–20), supporting collaborative workflows.

- **Diseases:** At the **second-order level** (Disease and Purpose), emerging and re-emerging diseases dominate most systems: detection (19), monitoring (42), and monitoring overall (106 for emerging diseases). Zoonotic diseases are also prominent in detection (13) and monitoring (33), whereas vector-borne (13) and neglected tropical diseases (9) are less frequent. At the **third-order level** (Disease, Purpose, and Users), public health officials mainly interact with emerging disease monitoring systems (60), while veterinarians are involved in zoonotic disease detection (15). At the **fourth-order level** (Disease, Purpose, Users, and Tasks), analysis is dominant for emerging disease monitoring for public health officials (91), and data collection dominates detection for veterinarians (8). At **higher orders**, including visual representation and interactivity, dashboards (36) and graphs/charts (30) are common for emerging disease systems, with moderate to high interactivity, while vector-borne and neglected tropical disease systems use specialized visualizations with lower interactivity.

- **Dataset Type:** At the **second-order level** (Dataset and Purpose), human datasets dominate (42 occurrences), followed by human-animal datasets (40), while animal-only (10), environmental-only (7), and human-animal-environment datasets (12) are less frequent. **Third-order relationships** (Dataset, Purpose, and Users) show public health officials primarily use human datasets (107), researchers use human-animal datasets (36), and veterinarians engage more with animal-only datasets (25). At the **fourth-order level** (Dataset, Purpose, Users, and Tasks), human datasets support analysis-heavy monitoring systems (91 occurrences for analysis), while multi-domain datasets enable exploration (36) and collaborative decision-making (29). **Higher-order relationships** show that dataset type predicts visual representation and interactivity: dashboards and graphs are used for human datasets (66 occurrences for "other visual representations"), while animal datasets often use specialized charts with lower interactivity (6-7 occurrences).

- **Users:** At the **second-order level** (Users and Purpose), public health officials dominate monitoring and analysis (107), researchers are linked to exploratory or predictive tools (65 for analysis), veterinarians are associated with detection and animal-centered systems (25), and policymakers are involved in planning and decision-making (66). **Third-order relationships** (Users, Tasks, and Visual Representations) indicate public health officials focus on analysis (91) using dashboards (36), researchers balance analysis (52) and exploration (25) with diverse visualizations ("other visual representations" 40), and veterinarians' systems rely on data collection (12) with specialized visualizations (5-6). At the **fourth-order level** (Users, Disease, Tasks, and Interactivity), dashboards for public health officials are moderately interactive (20), research systems are highly interactive (25), and veterinary systems are lower in interactivity (7). **Higher-order relationships** show that user type strongly determines system archetype, visual technique, analytic approach, and interactivity.

- **Tasks:** At the **second-order level** (Task and Purpose), analysis dominates (91 occurrences), monitoring (69), and data collection (39) are also prominent, while decision-making (26) and exploration (25) are moderately represented. **Third-order relationships** (Task, Purpose, and Users) highlight that public health officials' monitoring systems emphasize analysis (91), researchers focus on exploration (25) and decision-making (26), and veterinarians' detection systems emphasize data collection (12). At the **fourth-order level** (Task, Users, Visual Representations, and Interactivity), dashboards (36) and graphs (30) are common for analysis tasks, whereas exploratory tasks leverage diverse visualizations with higher interactivity (e.g., other visual representations 40). **Higher-order relationships** indicate task drives visualization design and interactivity: structured analysis and dashboards, exploratory and diverse visualizations, and data collection and specialized charts.

- **Visual Representations:** At the **second-order level** (Visual and Purpose), dashboards (36), graphs/charts (30), and geospatial maps (18) are dominant, while hierarchical charts (3) and network graphs (9) are less frequent. **Third-order relationships** (Visual, Users, and Tasks) show dashboards serve public health officials (36) for analysis (91), graphs/maps support researcher exploration (18-20), and other visualizations support veterinarian data collection (5-6). At the **fourth-order level** (Visual, Tasks, Users, and Interactivity), dashboards and graphs support moderate interactivity (20-30), whereas other representations are lower in interactivity (6-7). **Higher-order relationships** reveal that visual representation is strongly linked with task type, user type, dataset type, and interactivity.

- **Visual Analytics Techniques (Approach):** At the **second-order level** (Technique and Purpose), statistical/analytical methods (91) dominate monitoring and detection, while machine learning/AI (24) and data integration techniques (11) are selectively applied. **Third-order relationships** (Technique, Users, and Tasks) show public health officials rely on statistical methods for analysis (91), researchers combine ML/AI with exploration (25), and veterinarians rely on simple statistical approaches for data collection (12). **Fourth-order level** (Technique, Visual, Task, and Interactivity) highlights ML/AI often accompanies high interactivity and diverse visualizations (25 occurrences), while statistical-method-heavy systems favor dashboards with moderate interactivity (36 occurrences).

- **Interactivity:** At the **second-order level** (Interactivity and Purpose), moderate interactivity is most common (20-25 occurrences), high interactivity occurs in exploratory research systems (25), and low interactivity is tied to detection or veterinary systems (7). **Third-order relationships** (Interactivity, Users, and Tasks) indicate public health dashboards are moderately interactive (20), research exploration is highly interactive (25), and veterinary tools are low-interactivity (7). **Fourth-order level** (Interactivity, Visual, Task, and Dataset) shows high interactivity aligns with dashboards and graphs (30-36 occurrences) for analysis/exploration of human datasets, while low-interactivity visualizations are tied to animal/environmental datasets (6-7 occurrences).

- **Available Online:** At the **second-order level** (Available and Purpose), most systems are available online (Yes: 17-20 occurrences), particularly monitoring and analysis platforms, while offline systems are rarer. **Third-order relationships** (Available, Users, and Tasks) indicate online systems target public health officials and researchers for analysis/decision-making (52-91 occurrences). **Fourth-order level** (Available, Visual, Task, and Interactivity) shows online systems use dashboards and graphs (30-36 occurrences) with moderate to high interactivity, whereas offline systems support specialized tasks with lower interactivity (6-7 occurrences). **Higher-order relationships** suggest online availability correlates with interactivity, user role, analytic complexity, and system purpose.

The heatmap in Figure 3.16 effectively visualizes these relationships, making patterns of frequency and emphasis across features immediately apparent, while Table 3.2 summarizes the high, medium, and low correlations between features, providing both a quantitative and qualitative overview of the observed system archetypes.

3.2.3 Publication Trends and Venue Analysis

The analysis revealed a significant upward trend in One Health-related publications over the past decade, with a notable surge in 2023 and 2024, as shown in Figure 3.18. The bar chart illustrates the number of reviewed papers per year, highlighting the steady increase and recent peak in publications. The journal *One Health* emerges as the dominant publication venue, accounting for 40 papers, underscoring its central role in disseminating research in this field (Figure 3.17). Other frequently appearing venues include *Frontiers* in Public Health, *PLOS One*, and *Frontiers in Veterinary Science*, reflecting the interdisciplinary nature of One Health scholarship. The increasing diversity of journals and platforms in recent years highlights the expanding relevance and recognition of the One Health approach across public health, veterinary science, environmental health, and informatics. This trend suggests a growing commitment within the scientific community to addressing complex health challenges through integrated, cross-sectoral collaboration.

3.3 Discussion

Understanding the landscape of One Health systems through this analysis provides valuable insights that can inform future development and research by identifying gaps, highlighting effective practices, and setting priorities. Based on the analysis, only a small percentage (6%) of systems focus on prediction/forecasting (Di Lorenzo et al. 2023b). This highlights the need for future systems to enhance predictive capabilities and response strategies, which are essential for pandemic preparedness and effective outbreak management. Prior work has emphasized the value of predictive health information systems in preventing and controlling infectious diseases, including epidemics. For example, Myers et al. argue that linking environmental data with disease systems could enable the anticipation of outbreak conditions, thereby improving response efforts and resource allocation (Myers et al. 2000). Similarly, Keyel et al. highlight the disconnect between academic model development and its application in local public health decision-making, calling for standardized, actionable forecasting tools that can support real-time decisions in vector control and public health planning (Keyel et al. 2021). Our findings support these observations and suggest that a stronger integration

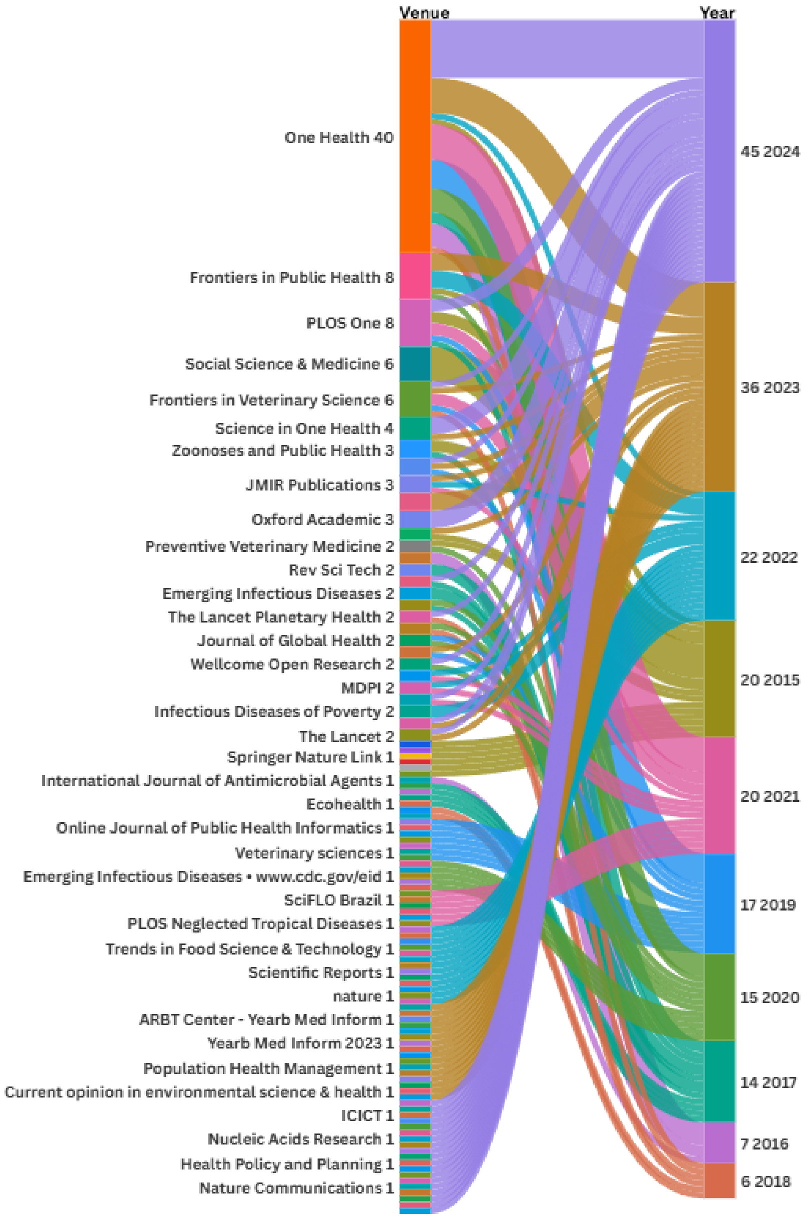


Figure 3.17: The distribution of journals in each year.

of predictive modeling into One Health systems could address a substantial shortcoming in current implementations.

With only a small proportion of systems utilizing predictive analytics (6%) (Chen et al. 2024) and machine learning techniques (16%) (Marty et al. 2024), there is a clear path for growth. Incorporating these advanced methods can improve the ability to forecast disease trends and make data-driven decisions. The presence of spatial and geospatial techniques

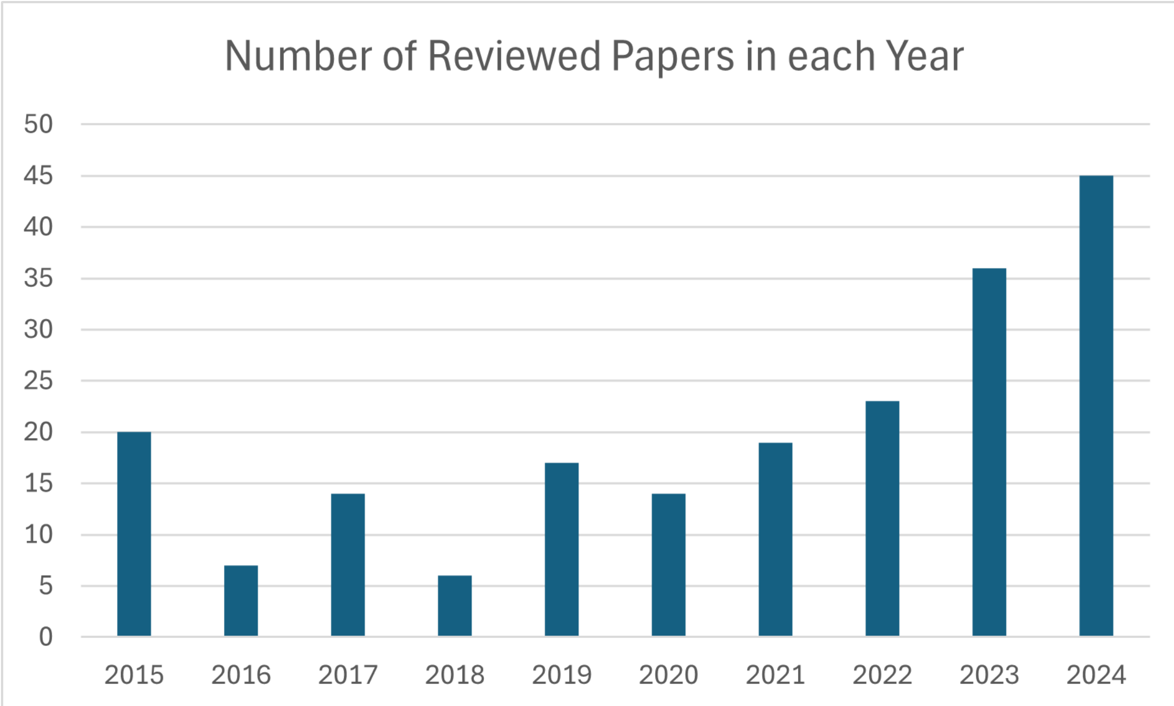


Figure 3.18: Number of the reviewed papers published each year. Venues are listed by publication count, and years are presented in chronological order, from most recent to earliest.

(32%) (Thumbi et al. 2015; Larkins et al. 2024) underscores the importance of understanding geographic distribution and environmental influences. Future systems can utilize geospatial analysis to pinpoint disease hotspots and monitor the spread of emerging infectious diseases.

Regarding diseases, the focus is primarily on COVID-19, while other important areas such as zoonotic diseases, vector-borne diseases, and antimicrobial resistance receive less attention. Research should expand to encompass these emerging threats, aligning with the holistic One Health approach. Regarding data integration, few systems utilize integrated datasets that combine human, animal, and environmental data, which is fundamental to the concept of One Health. This presents an opportunity to develop systems that better integrate diverse data sources to understand complex disease dynamics.

The popularity of visualizations such as graphs/charts, dashboards, and geospatial maps suggests their effectiveness in communicating data insights. Future systems can continue to use and enhance these visual tools, perhaps combining them for more comprehensive visualization capabilities. Additionally, systems with high or moderate interactivity allow users to engage deeply with the data, emphasizing interactive features (e.g., drill-downs, filtering). In future systems, interactive features can improve user experience and data exploration capabilities, enabling stakeholders to make informed decisions.

The analysis reveals different user types, including public health officials, such as healthcare providers, researchers, veterinarians, policymakers, and the general public. This diversity of user groups underscores the need for systems tailored to the specific goals and workflows of different stakeholders, offering capabilities such as data collection and export for researchers, decision-support tools for policymakers, and accessible dashboards for public awareness and engagement. Critically, co-development of systems with end users has proven valuable in ensuring systems meet real-world needs, foster adoption, and support sustained use. For instance, the EPIDEMIA system for malaria forecasting was designed through workshops with computer scientists, modelers, and public health partners, resulting in its successful implementation in Ethiopia's Amhara region (Merkord et al. 2017). Similarly, ArboMAP, a system developed for West Nile virus forecasting, emphasizes automation and usability for decision-makers, allowing public health departments to routinely produce and act on weekly risk predictions (Nekorchuk et al. 2024).

Since most systems emphasize analysis (84%) (Rist et al. 2015; Errecaborde et al. 2017; Thumbi et al. 2015; Zheng et al. 2019), future tools could also enhance support for tasks such as data collection, collaboration, and decision-making. For example, integrating real-time data streams and collaborative features could improve how users work together in pandemic response efforts. With many systems offering only low interactivity, future research can focus on enhancing user interaction with the data. This can be achieved through features such as real-time updates, predictive modeling integration, and dynamic filtering to facilitate deeper analysis. There is an opportunity to adopt or improve techniques like hierarchical charts, network graphs, and flow diagrams, which could offer richer, multidimensional insights into data relationships.

The multidimensional analysis of feature relationships reveals several overarching trends that reflect the current landscape, strengths, and limitations of visualization systems within One Health and pandemic-related research contexts. The patterns identified across second- to higher-order relationships demonstrate not only recurring design choices but also systematic imbalances in how visualization systems are conceptualized, implemented, and targeted toward specific users and diseases.

The findings around disease coverage and user groups can help policymakers identify priority areas for resource allocation. For instance, increased focus on integrating environmental data might better address zoonotic disease outbreaks related to climate change. Enhancing Preparedness and Response Strategies: Insights from monitoring and planning-focused systems can guide the design of tools that not only detect emerging threats but also offer strategic recommendations for outbreak response and long-term planning. We summarized the keys for enhancing One Health systems in Figure 3.20.

The relationships between users, datasets, and purposes underscore a human-dataset bias. Public health officials dominate system design, particularly in monitoring and analysis tasks, often using human-only datasets. Although One Health inherently emphasizes cross-domain integration, animal and environmental datasets remain marginal, and systems addressing all three domains simultaneously are rare. Veterinarian-centered tools exist, but focus narrowly on detection with limited interactivity and specialized visualization types. This pattern highlights a critical gap: despite a conceptual commitment to the One Health paradigm, the operational and technical integration of human, animal, and environmental data remains limited.

Visualization strategies reinforce the analytical orientation of most systems. Dashboards, graphs, and maps are the most common representations, optimized for structured analysis rather than open-ended exploration. High interactivity is largely restricted to research-oriented tools, whereas operational or veterinary systems tend to employ static or low-interactivity designs. This separation between interactive exploration and routine monitoring highlights a persistent divide in usability and accessibility. Systems with moderate to high interactivity, typically online dashboards, offer significant potential for real-time collaboration and decision-making but remain underutilized in domains outside public health.

The strong prevalence of statistical and analytical methods indicates a mature but conservative technological profile. While machine learning and AI methods are increasingly featured in research-oriented systems, they are seldom integrated into operational tools for real-time monitoring or prediction. This suggests a technological readiness gap, where advanced analytics are available but not yet embedded in routine visualization workflows.

The hierarchical relationships across features collectively suggest that system design in the One Health domain follows recognizable archetypes:

- *Monitoring–Analysis–Dashboard* systems for public health officials;
- *Detection–Data Collection–Specialized Visuals* for veterinarians; and
- *Exploration–Diverse Visuals–High Interactivity* systems for researchers.

While each archetype serves its purpose, the lack of hybrid systems bridging these patterns limits integrative insight. Effective One Health visualization platforms should therefore aim for multi-user adaptability, data-type fusion, and flexible visual-analytic workflows that support both structured monitoring and exploratory analysis.

By recognizing trends, addressing limitations, and building upon successful methodologies, future research and development efforts can support the One Health approach more effectively, ultimately improving disease detection, prediction, and prevention strategies across the human-animal-environment interface, as shown in the suggestions summary in Figure 3.19.

Figure 3.19: Some suggestions for Improving One Health Approach

3.4 Summary

This systematic review examined One Health datasets for the management of infectious diseases. Our paper systematically categorizes and evaluates One Health dashboards and systems based on their use of visualization, visual analytics, types of data, and functionality, aiming to understand their effectiveness in infectious disease management tasks, such as prediction, monitoring, and mitigation. Analyzing systems from 2015 to 2024, the review assessed 202 articles identified through specific keywords in various databases. The relationships between the categories revealed through this review highlight that while visualization

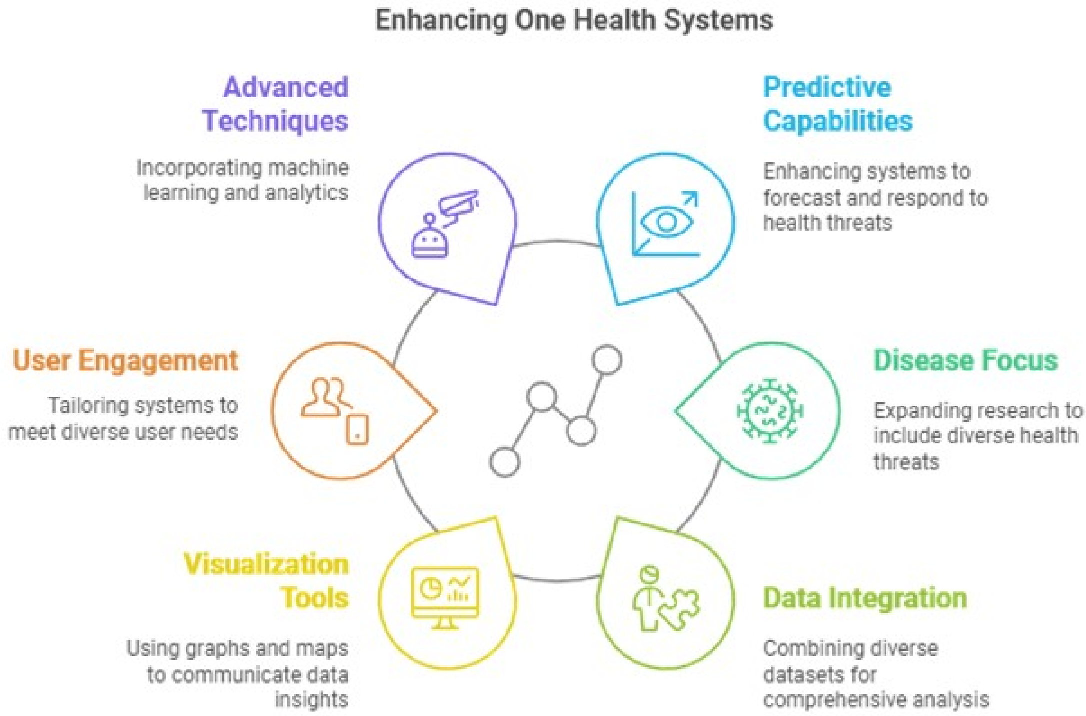


Figure 3.20: Proposed keys for Enhancing One Health Systems

systems in the One Health ecosystem have achieved considerable sophistication in analytical monitoring, they remain fragmented across domains and user roles. The next generation of tools must move beyond descriptive visualization toward integrated, predictive, and participatory analytics—realizing the full potential of One Health as an actionable framework for global disease surveillance and response. While a significant portion of studies still rely on single-domain data, there is a clear trend toward integrating multiple datasets, especially Human and Animal data. Fully integrated systems that combine human, animal, and environmental data accounts are limited, and most of these are commentary pieces rather than functional systems. This indicates a growing recognition of the One Health framework, but also highlights a gap—and an opportunity—for more robust, applied systems that truly integrate data across all three domains. Based on the comprehensive review, we believe that our paper is unique in its focus and scope compared to existing literature on similar topics.

Table 3.1: Infectious Disease System Classification Framework

Feature	Definition / Category
Purpose of the System	Detection: Identify emerging disease events Prediction/Forecasting: Estimate future disease trends Planning: Support preparedness and resource allocation Mitigation/Response: Guide interventions to reduce outbreak impact Monitoring: Continuous tracking of trends over time
Diseases	Emerging and Re-emerging Diseases Zoonotic Diseases Vector-borne Diseases Neglected Tropical Diseases Food Safety and Waterborne Diseases Antimicrobial Resistance
Dataset Types	Human data: patient records, mobility, clinical tests Animal data: veterinary records, wildlife tracking Environmental data: climate, geospatial, ecological variables Integrated data: combination of multiple sources (One Health approach)
Users	Public health officials Researchers Veterinarians Healthcare providers Policymakers General public
Tasks	Data collection: gathering relevant data Analysis: statistical, epidemiological, computational evaluation Exploration: creating and interpreting visual representations Decision making: applying insights to guide interventions Collaboration: sharing data and findings with stakeholders
Visual Represent.	Graphs/Charts: line, bar, pie, scatter Geospatial Maps: heatmaps, choropleth, density maps Network Graphs: relationships between variables or outbreaks Dashboards: integrated multi-component visualization platforms Hierarchical Charts: treemaps, sunbursts Other Visual Representations
Visual Analytics Tech.	Statistical and Analytical Methods Spatial and Geospatial Analysis Data Integration and Processing Machine Learning and AI
Interactive Ability	High interactivity, Moderate interactivity, Low interactivity
System / Dashboard	Yes (functional system/dashboard), No (conceptual/theoretical)
Availability	Yes, No

Table 3.2: Summary of Feature Relationships Across All System Aspects

Feature	**High Correlation**	**Medium Correlation**	**Low Correlation**
Purpose of System	Monitoring, Public Health, Analysis, Dashboards, Moderate/High Interactivity	Planning, Policymakers, Decision-Making, Dashboards, Moderate Interactivity	Mitigation/Response, Mixed Tasks/Users
Diseases	Emerging/Re-emerging, Detection, Public Health, Monitoring, Dashboards, Moderate/High Interactivity	Zoonotic, Monitoring, Human	Vector-borne, Neglected Tropical Diseases, Food/Waterborne, AMR, Other
Dataset Type	Human, Analysis, Dashboards	Human-Animal, Human-Environment, Exploration, Diverse Visualizations	Animal, Environmental, Human-Animal-Environment, Low Interactivity
Users	Public Health Officials, Monitoring, Analysis	Researchers, Exploration, Diverse Visualizations	Veterinarians, Detection, Data Collection, Specialized Visualizations, Dashboards
Tasks	Analysis, Monitoring, Public Health Officials, Dashboards	Data Collection, Detection, Veterinarians, Specialized Visualizations	Decision-Making, Planning, Rare System Types
Visual Representations	Dashboards, Graphs/Charts, Other Visual Representations Monitoring, Analysis	Geospatial Maps, Emerging/Zoonotic Diseases, Research	Network Graphs, Hierarchical Charts, Less Common
Visual Analytics Techniques	Statistical Analytical Methods, Analysis, Monitoring	Machine Learning/AI, Research, Exploration	Additional Techniques, Rarely Used, Specialized Tasks
Interactivity	Moderate/High, Monitoring, Public Health Officials Dashboards	Moderate, Planning, Policymakers Decision-Making	Low, Detection, Veterinarians Specialized Visualizations
Availability Online	Yes, Monitoring, Planning, Public Health/Policymakers		No, Rare or Specialized Systems

Chapter 4

A Review of Visual Representations

During the COVID-19 pandemic, many visualization scientists, researchers, and practitioners collaborated closely with domain experts, gaining comprehensive insights into users, data, and tasks within these contexts. To better understand their experiences, we conducted a three-stage survey that focused on the visual representations used in the COVID-19 pandemic response [1]. Adherent to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Reporting Guidelines (Page et al. 2021b), a widely used framework for conducting and reporting systematic literature reviews, we systematically identified, screened, and analyzed relevant articles on various publications and search platforms, developed a classification scheme, and coded these articles according to the scheme. Given the thousands of COVID-19 visualization papers in VIS and non-VIS venues, direct engagement with the authors was impractical. Instead, we used a systematic literature analysis and classification approach. Dykes et al. (2022) examined VIS researchers in the RAMPVIS project, summarizing 29 design idioms. Although insightful, their method was not scalable for a broader literature analysis.

We aim to thoroughly investigate the role of visualization in pandemic-related research.

[1] This chapter is based on the paper submitted to the Computer Graphics Forum (CGF) in August 2025: A. Basheer, Y. Xing, M. Chen, D. Ebert, K. Gaither, W. Jentner, T. Schreck, and A. Abdul-Rahman, "Pandemic Readiness: Through the Lens of Visual Representations Used in Fighting the COVID-19 Pandemic," CGF, 2025.

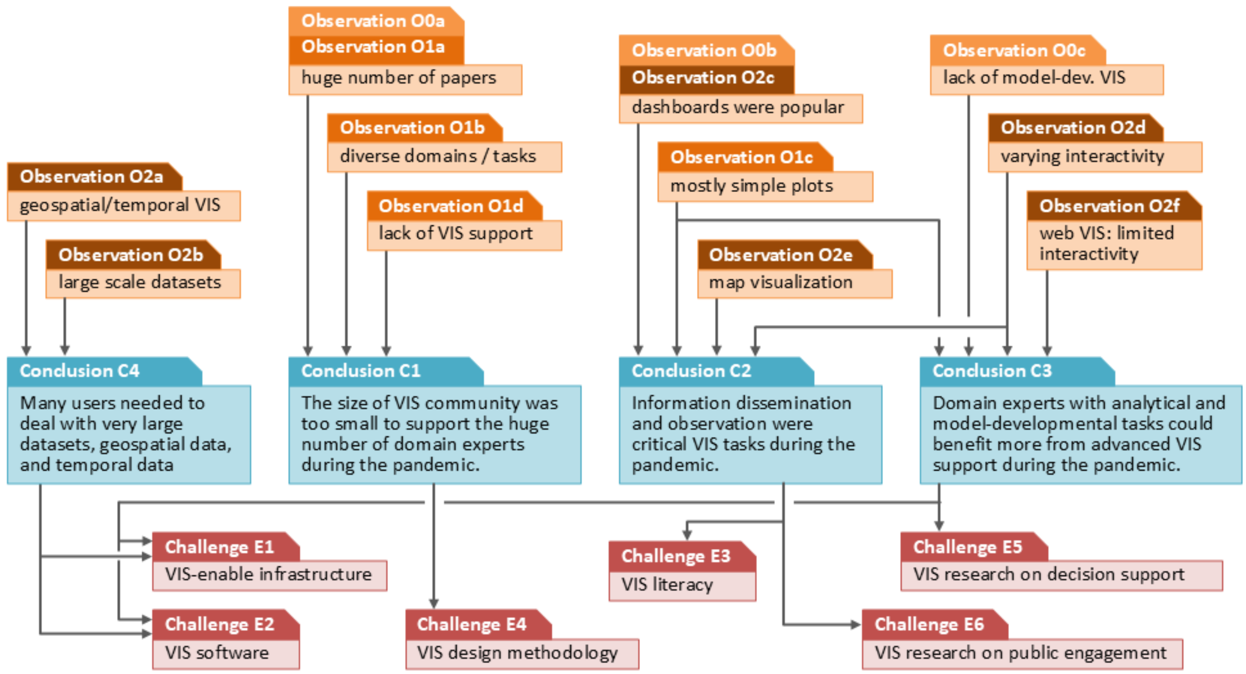


Figure 4.1: Reasoning Diagram of the observations, conclusions, and challenges gathered from the surveys conducted.

This involves progressively broadening our selection criteria to encompass a diverse yet relevant range of work, including visualization techniques, applications, and analytical approaches, within the context of pandemics.

Through the lens of these visual representations and a careful examination of the pertinent literature, we deduced the data types used by domain experts and their corresponding visualization tasks. By connecting the dots among domain subjects, data types, visual representations, and visualization tasks, we assessed the strengths and weaknesses of current visualization capabilities in terms of readiness for future pandemics, identified areas needing improvement, and recommended necessary actions. This is useful for other crises (e.g., natural disasters, critical infrastructure failures, and routine health surveillance – areas that need analysis of data for decision-making, evaluating interdiction/response measures, and communication to implementers and the public. Fig. 4.1 summarizes the reasoning diagram of our observations, conclusions, and challenges gathered from the three-stage survey. Addressing these challenges will better prepare visualization research to respond to future global crises.

Below is a summary of our findings from the three-stage survey:

High Demand for VIS Support: There is a significant need for visualization (VIS) support in pandemic-related research, but resources such as tools, infrastructures, and experts are insufficient.

Underutilization of VIS Beyond Dissemination: While visualization is commonly used for data dissemination, it has untapped potential for exploratory analysis, decision-making, and model development.

Need for Stronger Evidence of VIS Impact: Although VIS played a crucial role in COVID-19 data analysis and decision-making, more empirical research and case studies are needed to establish its long-term value.

Growth in Data Complexity: Pandemic-related data has expanded in size and diversity, particularly in geospatial and temporal domains, requiring more sophisticated visualization tools.

4.1 Methodology and Approach

To thoroughly examine the role of visualization in pandemic-related research, we conducted a three-stage survey, gradually expanding the scope and diversity of our article sources. We aim to capture a wide yet relevant range of work, encompassing visualization techniques, applications, and analytical approaches within the context of pandemics.

Through an initial survey, we developed a foundational collection of visualization articles related to pandemic research. By issuing the broad query *"pandemic AND visualization"* on Google Scholar, we received 350,000 results. After a careful review and analysis of the first 20 pages of the Google Scholar results, we collected 106 papers from 68 different publication venues (Figure 4.3 (b)). This stage, known as **Survey 0**, offered an overview of how visualization has been applied in pandemic research from a broad and general perspective, with

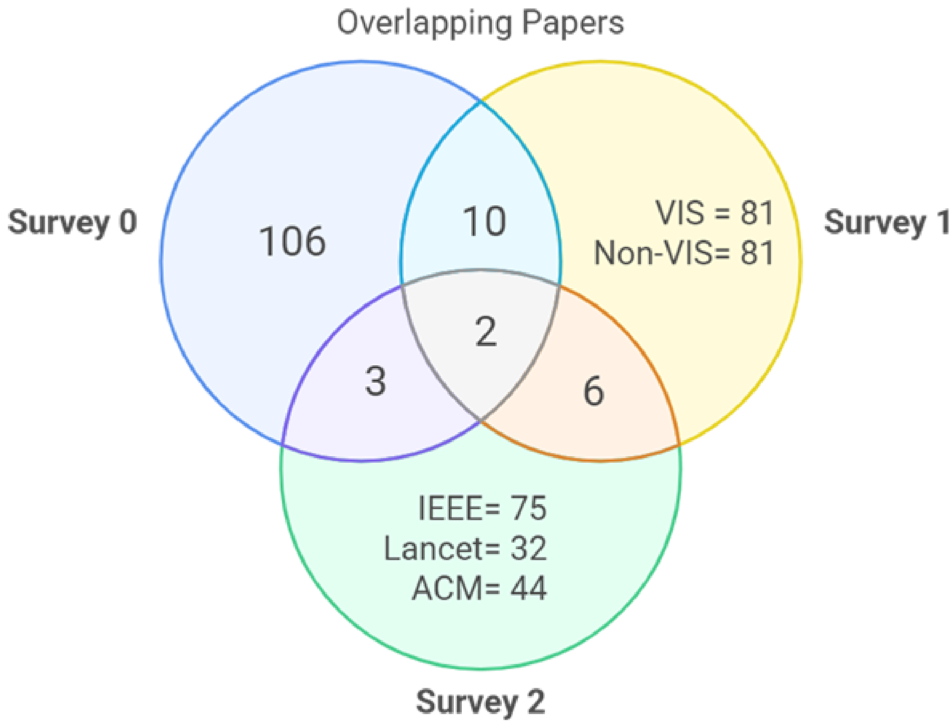


Figure 4.2: Total number in each of the three surveys, **Survey 0**, **Survey 1**, and **Survey 2**, and the number of overlapping papers between them.

many papers published in various non-visualization venues.

To ensure our search reflects pandemic-related research and includes sufficient studies within the visualization community, we expanded the query from *"pandemic AND visualization"* to *"pandemic OR healthcare OR epidemiological modeling OR emergency response AND visualization."* We searched and reviewed articles from Google Scholar as well as well-established visualization and user interface research venues such as IEEE Transactions on Visualization and Computer Graphics (TVCG), IEEE Computer Graphics and Applications (CG&A), Computer Graphics Forum (CGF), ACM Transactions on Interactive Intelligent Systems (ACM TiiS), and the ACM Conference on Intelligent User Interfaces (ACM IUI). By applying specific selection criteria, as detailed in Figure 4.7, we curated a collection of 162 articles, 81 from visualization-focused venues, and 81 from other relevant domains, in 76 different publication venues (Figure 4.3 (c)). We refer to this stage as **Survey 1**, which aimed to balance domain breadth with visualization depth.

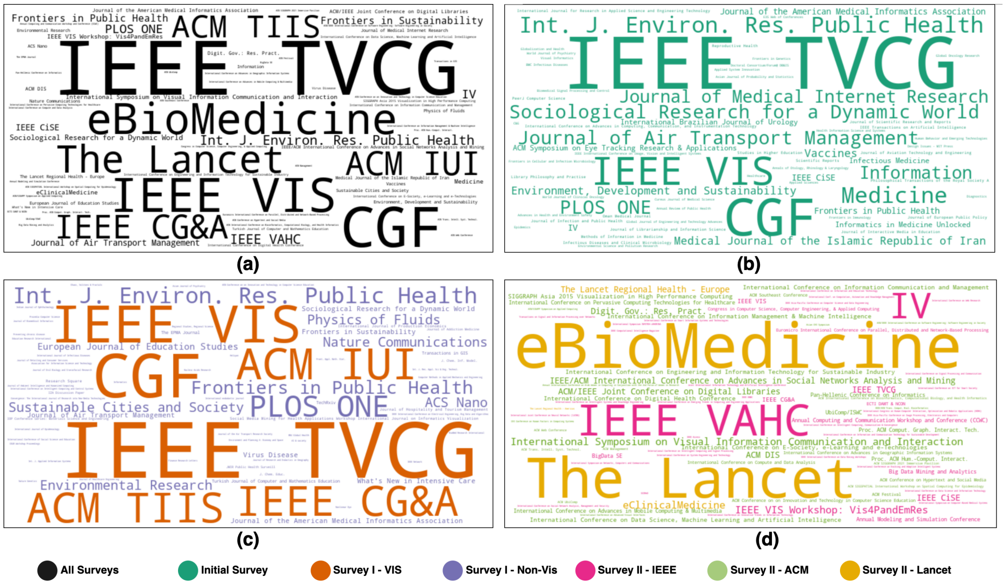


Figure 4.3: The analyzed publication venues are visualized as follows: (a) a word cloud displaying **all** venue names across three surveys; (b) venues from **Survey 0** (green); (c) venues from **Survey 1**, with orange indicating visualization-related venues and purple indicating non-visualization venues; (d) venues from **Survey 2**, with pink representing IEEE venues, light green for ACM venues, and yellow for The Lancet. In all word clouds, font size is proportional to the number of publications within the respective survey.

After completing **Survey 2**, we recognized that The Lancet contains numerous research papers on global health and pandemics that feature visualization and visual analytics that were not included in our Google Scholar search during **Survey 0** and **Survey 1**.

We identified this gap by conducting targeted searches directly on The Lancet website using the same keywords as in our earlier surveys, which retrieved several relevant publications that did not appear within the first 20 pages of Google Scholar results, where we had previously collected articles.

Additionally, while **Survey 1** focused on venues for visualization-related publications, we aimed at a broader investigation that systematically includes visualization applications beyond the dedicated visualization research community. We conducted another round of surveys to address these gaps, strictly adhering to the PRISMA methodology (Page et al. 2021b) for a systematic literature review. To target influential publications in medicine and global health, such as The Lancet, and to ensure the completeness of our review, we systematically searched major digital libraries (ACM Digital Library, IEEE Xplore, and The Lancet) using the same Boolean query employed in **Survey 1**. Applying the paper selection criteria as detailed in Figure 4.7, we finalized a dataset of 151 articles. This stage is referred to as **Survey 2**.

Fig. 4.2 represents the overlap of the articles between three different surveys: **Survey 0**, **Survey 1**, and **Survey 2**. The overlap sections indicate shared papers: 10 papers are common between **Survey 0** and **Survey 1**. 3 papers overlap between **Survey 0** and **Survey 2**. 6 papers are shared between **Survey 1** and **Survey 2**. Only two papers are common across all three surveys. Fig. 4.2 also illustrates the distribution and overlap of articles between surveys, highlighting shared research across different domains.

We chose to analyze the three surveys separately due to fundamental differences in their scope, data sources, publication focus, and intended audiences, which reflect distinct facets of pandemic-related visualization research.

In addition to following PRISMA guidelines, we applied a systematic observational methodology to guide the analysis, interpretation, and synthesis of the surveys. This methodology included:

1. **Scope Definition and Categorization:** Defining dimensions such as visualization type, application domain, data characteristics, target audience, and venue type, and categorizing each paper accordingly.

2. **Feature Extraction and Coding:** Systematically extracting key features from each publication (methods, data types, evaluation approaches, challenges) and recording them in a structured matrix.

3. **Cross-Survey Comparison:** Comparing features across surveys to identify overlaps, gaps, and trends, highlighting differences in focus, audience, and methodological rigor (Figure 4.2).

4. **Pattern Recognition and Synthesis:** Detecting recurring techniques, common chal-

Type	Survey0 (106 papers)	Survey1_Non VIS (81 papers)	Survey1_VIS (81 papers)	Surevy2-IEEE (75 papers)	Surevy2-Lancet (32 papers)	Surevy2-ACM (44 papers)
3d visualization	0.00%	0.00%	0.00%	2.67%	3.13%	9.09%
animation	0.94%	0.00%	1.23%	0.00%	0.00%	4.55%
area plot	4.72%	12.35%	18.52%	8.00%	28.13%	9.09%
bar chart	33.96%	35.80%	80.25%	54.67%	50.00%	43.18%
beeswarm plot	2.83%	0.00%	3.70%	0.00%	0.00%	0.00%
box and whisker plot	7.55%	3.70%	13.58%	2.67%	34.38%	4.55%
bubble map	0.94%	1.23%	2.47%	2.67%	3.13%	6.82%
choropleth map	14.15%	0.00%	18.52%	0.00%	0.00%	18.18%
cluster visualization	15.09%	11.11%	30.86%	0.00%	0.00%	0.00%
dashboards	18.87%	30.86%	55.56%	22.67%	3.13%	22.73%
dendrogram	0.00%	1.23%	1.23%	0.00%	0.00%	0.00%
density visualization map	3.77%	2.47%	7.41%	0.00%	0.00%	0.00%
donut chart	0.00%	0.00%	0.00%	0.00%	0.00%	4.55%
dotplot	2.83%	0.00%	3.70%	0.00%	3.13%	4.55%
flow map	0.94%	0.00%	1.23%	0.00%	0.00%	0.00%
funnel plot	0.94%	1.23%	2.47%	0.00%	0.00%	0.00%
glyph map	1.89%	13.58%	16.05%	0.00%	0.00%	0.00%
heat map	0.94%	0.00%	1.23%	16.00%	25.00%	6.82%
histogram	8.49%	2.47%	13.58%	2.67%	3.13%	0.00%
image	1.89%	0.00%	2.47%	1.33%	12.50%	0.00%
interactive map	8.49%	0.00%	11.11%	45.33%	15.63%	6.82%
list	1.89%	1.23%	3.70%	0.00%	0.00%	0.00%
network visualization map	30.19%	16.05%	55.56%	8.00%	9.38%	9.09%
pie chart	6.60%	6.17%	14.81%	13.33%	0.00%	4.55%
pixel-based visualization	0.94%	1.23%	2.47%	0.00%	0.00%	0.00%
rose tree visualization	0.94%	1.23%	2.47%	0.00%	0.00%	0.00%
scatter map	0.94%	3.70%	4.94%	0.00%	0.00%	0.00%
scatter plot	2.83%	19.75%	23.46%	16.00%	43.75%	2.27%
storytelling	3.77%	1.23%	6.17%	0.00%	0.00%	4.55%
table	1.89%	6.17%	8.64%	5.33%	21.88%	9.09%
themeriver	0.94%	0.00%	1.23%	0.00%	0.00%	0.00%
time series plot	35.85%	54.32%	82.72%	73.33%	56.25%	34.09%
tree map	6.60%	4.94%	13.58%	2.67%	0.00%	2.27%
violin chart	0.94%	1.23%	2.47%	1.33%	12.50%	0.00%
vortex regions	0.00%	1.23%	1.23%	0.00%	0.00%	0.00%
word cloud	5.66%	2.47%	9.88%	6.67%	0.00%	11.36%

Figure 4.4: A summary of the visual representations used in the pandemic surveyed papers.

lenges, and emerging trends using quantitative summaries (word clouds, Figure 4.3) and qualitative synthesis.

5. **Identification of Gaps and Challenges:** Using observed patterns to highlight areas with insufficient visualization support, limited analytical tools, or gaps in decision-making integration.

6. **Validation and Triangulation:** Cross-checking observations across surveys, venues, and search strategies, including targeted searches (e.g., The Lancet in Survey 2), to ensure robustness.

This observational framework provided the foundation for the reasoning behind conclusions, challenges, and recommendations, and it underpins the definition and mapping of Visualization Readiness Levels (VRLs).

We chose to analyze the three surveys separately due to fundamental differences in their scope, data sources, publication focus, and intended audiences, which reflect distinct facets of pandemic-related visualization research.

4.1.1 Survey 0 – Data Collection and Categorization

The Sample of Articles Our methodology systematically searches for relevant literature on pandemic visualization, reviews irrelevant articles, and prioritizes those that use visualizations. To understand the current landscape of pandemic visualization, we analyzed articles published from 2013 to 2024 that incorporated visualizations of pandemic-related topics, including epidemics and public health outbreaks. In this survey, we searched Google Scholar using the terms *"pandemic + visualization"*, which returned 350,000 results. We then removed irrelevant articles, such as not VIS survey articles, studies unrelated to COVID-19 or other pandemics, theoretical articles without visualizations, and bibliometric analysis articles, providing us with 106 articles. We examined the content of the collected articles and noticed a trend in which most visualizations are simple and can be created without advanced visualization knowledge.

Classification Process. Our classification categories build on the design triangle by Miksch and Aigner 2014, extending the space to visual representations and Visual Analytics (VA) techniques. We used the term visualization tasks to follow the four levels of visualization tasks proposed by Chen and Golan 2015. All co-authors collaboratively defined the categories in a brainstorming session, refining criteria and analysis methods. We conducted classification exercises on randomly selected articles to ensure consistency and address ambiguities. Four authors independently coded articles, resolving uncertainties in bi-weekly discussions.

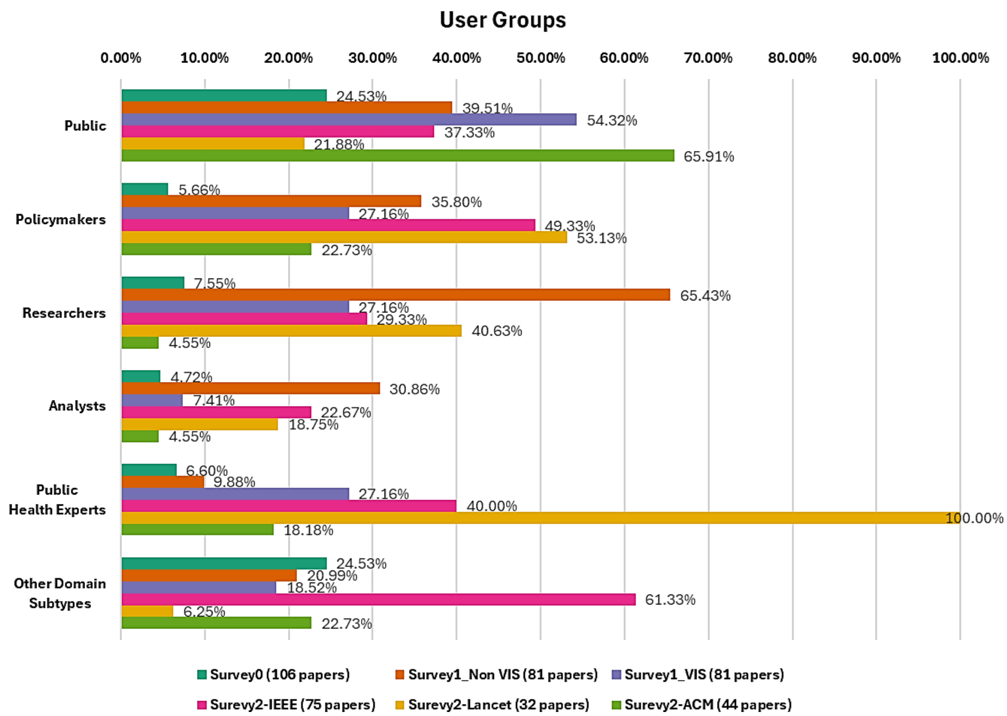


Figure 4.5: A summary of user groups in the pandemic surveyed papers.

In our paper, we define these classifications and their summaries:

Data – sources and types must acknowledge the diverse datasets used in research, analysis, and decision-making processes. Data encompasses information collected, curated, or generated from various sources, including but not limited to empirical observations, digital records, and computational simulations.

Users – individuals or groups interacting with data, visualizations, and information products for various purposes, including analysis, decision-making, communication, and understanding.

Visualization Tasks – represent four categories of visualization tasks:

- Disseminative visualization (VD): In this category, visualization serves as a presentational aid for sharing information or insights with others.
- Observational visualization (VO): Here, visualization acts as an operational aid that facilitates intuitive and/or rapid observation of the captured data.
- Analytical visualization (VA): This category uses visualization as an investigative tool

for examining and understanding complex relationships, such as correlation, association, causality, and contradiction.

- Model-developmental visualization (VM): In this category, visualization functions as a developmental aid aimed at enhancing existing models, methods, algorithms, and systems in terms of the pandemic's spread, as well as creating new ones.

Visual Representations – encompass various methods to visually communicate complex data and insights, such as line graphs, bar charts, and network visualization maps.

Visual Analytics Techniques – can be defined as methods that combine data visualization with analytical approaches to uncover insights and relationships within complex datasets.

In Fig. 4.4, we found that among the most preferred visualization types in this survey were bar charts (33. 96%), e.g. (Mbunge et al. 2021), network visualization maps (30. 19%), e.g. (Michailidis 2022), time series plots (35.85%), e.g., (Wentzel et al. 2023), dashboards (18.87%), e.g., (Tat'Yana et al. 2022), and choropleth maps (14. 15%), e.g., (Comba 2020).

In terms of users, the majority of the papers in this survey target the public (24.53%), e.g., (Padilla et al. 2022). Other domain subtypes like data visualization experts (24.53%), e.g., (Shi et al. 2022), followed by researchers (7.55%), e.g., (Chang et al. 2022), public health experts (6.60%), e.g., (Gonen et al. 2023), policymakers (5.66%), e.g., (Ko et al. 2014), and analysts (4.72%), e.g., (Batch et al. 2023) as we can see in Figure 4.5.

There is a stronger emphasis on visualizations that support analysis and observation rather than solely on dissemination or model development, where analytical visualizations (30.19%), e.g.,(Ryan et al. 2018) are the most common, followed by observational visualizations (16.04%), e.g.,(Mayer et al. 2023), disseminative visualizations (12.26%), e.g.,(Tat'Yana et al. 2022), and model-developmental visualizations (7.55%), e.g.,(Praharaj et al. 2023) as we can see in Fig. 4.6.

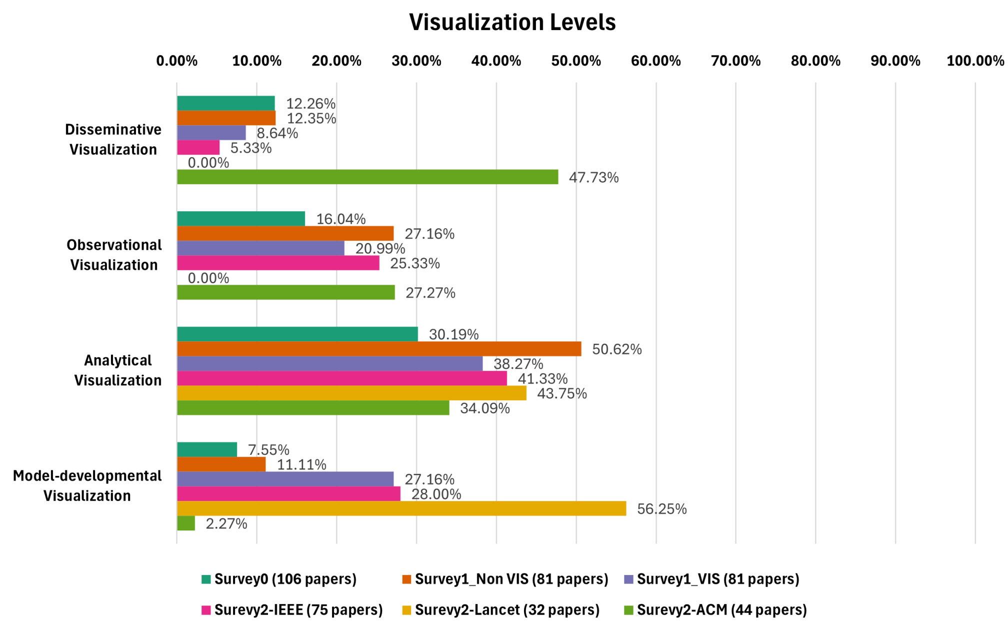


Figure 4.6: A summary of visualization tasks in pandemic-surveyed papers.

Survey 0 – Analysis & Summary

Survey 0 indicates a significant reliance on traditional visualizations and dashboards, yet there is a distinct opportunity to improve them through predictive modeling, interactivity, and advanced visualization techniques.

O0a There are many pandemic-related research papers currently available. A Google Scholar search for *"pandemic"* yields 4,190,000 results, while *"pandemic + visualization"* returns 350,000 results. We reviewed the first 20 pages of Google Scholar results for the *"pandemic + visualization"* query. After carefully reviewing and analyzing the 20 pages, we collected 106 articles in **Survey 0**, as illustrated in Fig. 4.7.

O0b Dashboards are one of the most popular types of visualization, representing 18.87% of the surveyed papers as shown in Fig. 4.4. This suggests that many studies emphasize high-level summary-based data presentations aimed at broad audiences, including policymakers and public health experts. However, further analysis is needed to determine their intended audiences and the motivations behind their design.

O0c Model-developmental visualizations are underrepresented (7.55%), as shown in Fig. 4.6, indicating a gap in visual tools for predictive analytics. Although complex VA methods exist, such as network visualizations, cluster visualizations, and glyph maps, they are not predominant. Most of the visualizations in the survey focus on the observation and analysis of historical data, rather than predicting future trends or revealing hidden relationships. This highlights the need to develop advanced VA techniques that integrate machine learning, temporal modeling, and uncertainty visualization to improve predictive capabilities. Such techniques could better support decision-making in dynamic and evolving public health contexts.

4.1.2 Survey 1 – Data Collection and Categorization

The Sample of Articles. As illustrated in Fig. 4.7, **Survey 1** builds upon the broader scope of **Survey 0** by refining both the search strategy and the target publication sources. While **Survey 0** relied solely on a basic query (*pandemic + visualization*") on Google Scholar to gather a wide and general set of papers across disciplines, **Survey 1** adopted a more focused approach. We first expanded the search terms to include a richer set of pandemic-relevant topics, such as *"healthcare + visualization"*, *"epidemiological modeling + visualization"*, and *"emergency response + visualization"*. We then applied this query across two sources: (1) Google Scholar, to retain coverage of interdisciplinary research, and (2) selected high-impact, well-established visualization and user interface venues—including IEEE TVCG, IEEE CG&A, Computer Graphics Forum (CGF), ACM Transactions on Interactive Intelligent Systems (ACM TiiS), and ACM Intelligent User Interfaces (ACM IUI)—accessed through their respective digital libraries (e.g., IEEE Xplore, ACM Digital Library, Wiley Online Library). The goal was to identify visualization-related contributions both from within the visualization research community and from broader application domains. After applying consistent exclusion criteria, we curated a dataset of 162 articles, comprising 81 papers from visualization-centric venues and 81 from non-VIS but domain-relevant venues. This dual-source strategy allowed us to examine how visualization was used and discussed across both research traditions, enabling a comparative perspective not present in Survey 0.

Classification Process. We used the same definition of the classifications conducted in **Survey 0** in terms of data, users, tasks, visual representations, and visual analytics techniques. We found that these sources of data can include epidemiological studies, e.g.,(Khan et al. 2024), online and social media platforms, e.g.,(Wentzel et al. 2023), government reports, e.g.,(Yu et al. 2021), academic databases, e.g.,(Samet et al. 2020), and research repositories, e.g.,(Padilla et al. 2022). Furthermore, data can be simulated or generated for specific research purposes, e.g.,(Rydow et al. 2022). It serves as the foundation for analysis, research, decision-making, and knowledge creation across various disciplines and domains. As shown in Fig. 4.5, the users range includes non-specialists, the general public (e.g.,(Bernasconi and Grandi 2021), domain experts, as well as policymakers (Pala et al. 2022).

As we can see in Fig. 4.6, in terms of visualization tasks, we found that 12.35% and 8.64% are disseminative visualization, 27.16% and 20.99% are observational visualization, 50.62% and 38.27% are analytical visualization, and 11.11% and 27.15% are model-developmental visualizations, representing the percentages of non-VIS and VIS articles, respectively. We also found that the visualization types most frequently mentioned in the VIS and HCI venues include time series plots (82.72%), e.g.,(Kunjir et al. 2017), bar charts (80.25%), e.g.,(Ra et al. 2020), and network visualization maps (55.56%), e.g.,(Villanustre et al. 2021). In contrast, non-VIS and non-HCI venues feature time series plots (54.32%), bar charts (35.80%), and dashboards (30.86%). These methods are frequently featured in numerous articles, indicating their widespread popularity and effectiveness in visualizing various types of data. Our analysis reveals a range of visualization methods in the papers, from traditional charts to innovative techniques like network visualization maps. This variety suggests that researchers and practitioners are exploring different visual data representations to address specific needs and challenges. We also noted emerging trends where visualization methods, such as choropleth maps and interactive maps (Antweiler et al. 2022), were employed. This indicates a growing interest in geospatial data visualization and interactive exploration tools,

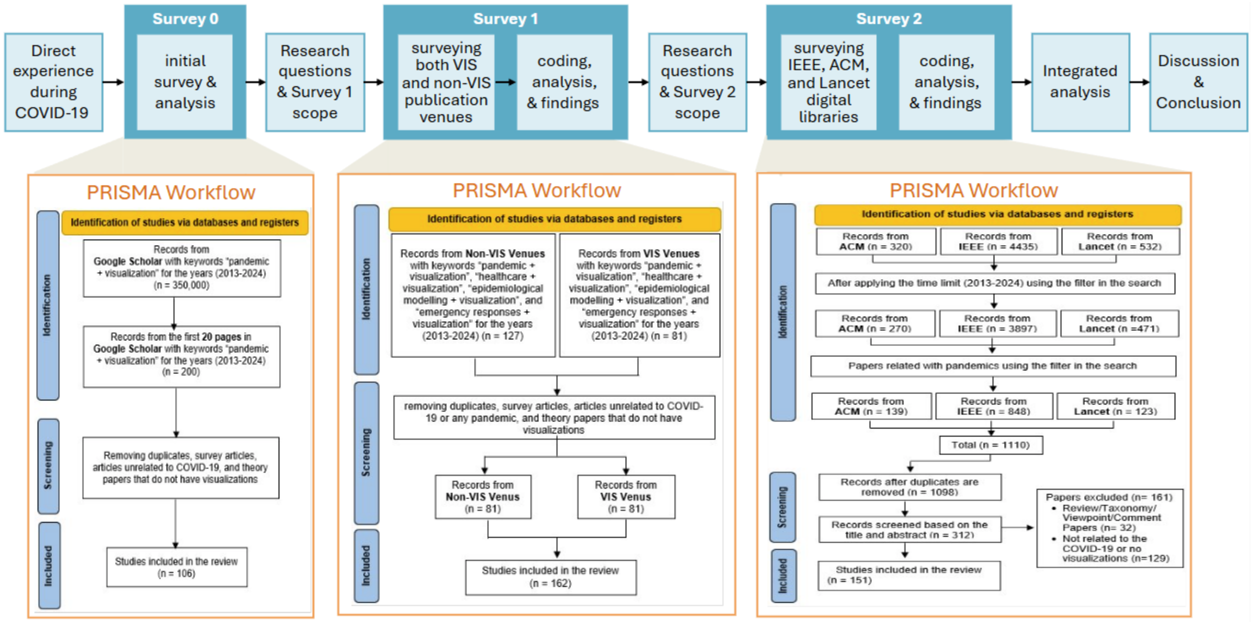


Figure 4.7: An overview of the three-round survey process, illustrating the global research workflow. Each survey follows the PRISMA framework, detailing the systematic approach for paper selection, review, and analysis in **Surveys 0**, **1**, and **2**.

likely driven by the increasing availability of location-based data and the desire to understand spatial patterns more effectively. Additionally, this implies that location-based events can be easily communicated through geospatial representations. In this survey, we found that analytical models, including statistical methods such as clustering, e.g., (Pérez-Campuzano et al. 2022), forecasting, e.g., (Yu et al. 2021), and curve fitting, e.g., (Kurniawan and Kurniawan 2021), along with machine learning algorithms such as neural networks, e.g., (Ahmed et al. 2021), topic modeling, and sentiment analysis, e.g., (Ilyas et al. 2022), are integrated into the visualization process to reveal patterns and derive insights. Interactive features enable users to engage with visualizations, modify parameters, and analyze data from multiple perspectives, promoting iterative analysis and hypothesis testing.

This survey highlights the diverse range of data sources, visualization methods, and analytical techniques employed in pandemic visualization research. Describes evolving trends, emerging tools, and design challenges, providing a comprehensive overview of the field.

Survey 1 – Analysis & Summary

Due to the ease of development and deployment, the findings demonstrate that much of the work during the pandemic focused on creating dashboards to present COVID-19 data. However, notable exceptions exist where more advanced Visual Analytics (VA) methods were used beyond basic data selection and aggregation. These exceptions are summarized below:

Predominance of Dashboards: A significant portion of the work involves the development of dashboards to visualize the COVID-19 data (Bach et al. 2023). These dashboards typically present aggregated data in a user-friendly interface, allowing users to explore trends, patterns, and statistics related to the pandemic.

Limited Sophisticated VA Approaches: While dashboards are common, they are designed to minimize functionality to a "single page" view. Many works focus on basic data selection, aggregation, and visualization techniques without exploring advanced VA workflows.

Exemplary Works: Despite the prevalence of dashboards, some exemplary works incorporate detailed VA workflows. For instance, Baumgartl et al. (Baumgartl et al. 2021) presented an approach using advanced VA techniques to analyze pathogen transmission pathways.

Need for Advancement: Exemplary works show the potential for sophisticated VA approaches in analyzing COVID-19 data. Advancements in VA can uncover hidden patterns and support more effective decision-making in combating the pandemic.

Advanced visual analytics techniques are crucial for deriving actionable insights and informing pandemic responses, thereby enhancing COVID-19 dashboards. From our survey, here are our findings:

O1a This survey used additional search terms, including *"healthcare + visualization," "epidemiological modeling + visualization,"* and *"emergency responses + visualization,"*

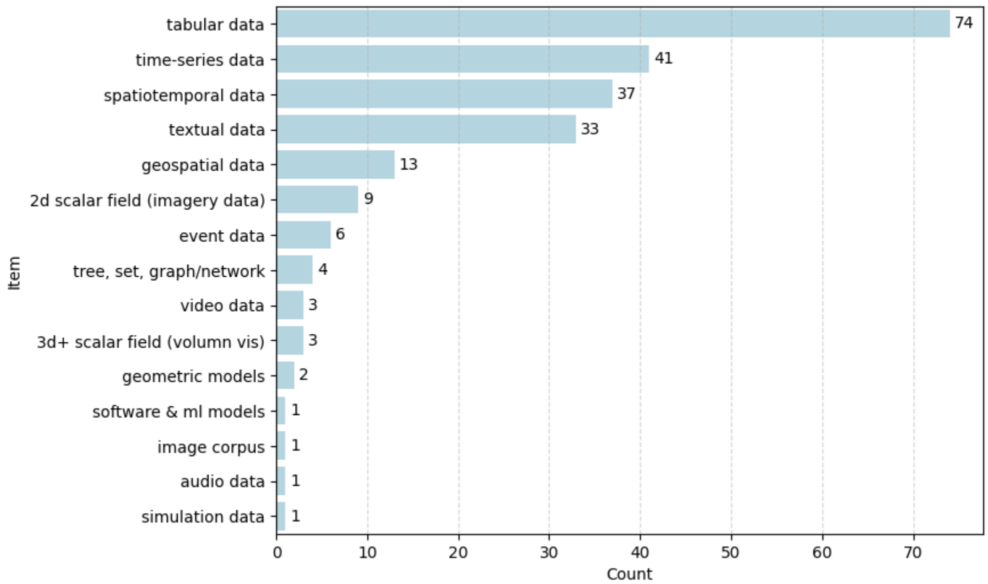


Figure 4.8: Data Type Distribution for **Survey 2**.

that result in 18,200, 18,300, and 20,200 articles in Google Scholar, respectively. We randomly selected 81 papers from non-VIS and non-HCI venues, classified them, and compared them with 81 papers from VIS and HCI venues. See our supplemental material for the full list of papers.

O1b The subject areas are diverse, including epidemiological data analysis, web visualization tools, AI and machine learning for pandemic analysis, healthcare analytics, decision support systems, mobile health analysis, social media sentiment analysis, business analytics, network analysis for pandemic risk, and extended and virtual reality applications.

O1c Our survey found that most visualizations used simple plots like line charts, bar charts, and network maps, as shown in Fig. 4.4.

O1d Many papers do not mention collaborations with VIS researchers or practitioners, potentially leading to a reliance on basic visualization techniques, as illustrated in Fig. 4.5.

4.1.3 Survey 2 – Data Collection and Categorization

Building on the foundation laid by **Survey 1**, which analyzed pandemic-related visualizations in VIS and non-VIS publication venues from the perspectives of: *data source*, *users*, *visualization tasks*, and *visual representations*, **Survey 2** expanded the exploration with broader publication venues. Unlike **Survey 1**, which relied primarily on Google Scholar search engines and selected visualization-related venues, **Survey 2** employed a more structured approach for paper selection, focusing on three major digital libraries: ACM, IEEE, and Lancet.

Survey 2 delved deeper into the data, usage, and developmental aspects of pandemic visualization projects and introduced new codes:

- **Data Size**: As a supplement to the *data type* classification in **Survey 1**, this code evaluates the scale of data visualized in the articles, distinguishing between small and large-scale datasets that require visualizations to aid user comprehension.
- **Use of Advanced Algorithms & Level of Interaction**: These two codes build on the *visual analytics techniques* category in **Survey 1**, offering a more detailed evaluation of the visualization systems. They measure 1) whether advanced algorithms are employed to support data processing and analysis, and 2) the level of user-data interaction the visualization systems provide, assessing the depth of their visual analytical capabilities.
- **Public Engagement**: This code independently evaluates the purpose of the visualizations, focusing on whether they are designed to engage the public or facilitate dissemination beyond specialized audiences.
- **Tools and Platforms**: This code aims to understand which toolkit or software generates the visualizations and in what format the visualization systems are deployed in each paper.

Survey 2 – Analysis & Summary

As shown in Fig. 4.7, we conducted **Survey 2** by collecting papers from ACM, IEEE, and The Lancet using the exact keywords from **Survey 1**. The initial search yielded 320 ACM, 4,435 IEEE, and 532 Lancet papers, reduced to 270, 3,897, and 471 after applying the 2013–2024 time frame. A careful review and analysis for pandemic relevance identified 139 ACM, 848 IEEE, and 123 Lancet papers. After excluding reviews, taxonomies, viewpoints, and papers unrelated to COVID-19 or visualization, 151 remained. **Survey 2** categorized these papers for a comprehensive analysis. The complete coding book for this survey is displayed in Table 4.2.

First, the **data types** discussed in the papers were categorized into various formats, including textual data, tabular data, tree/set/graph or network data, geospatial data, time-series data, and spatiotemporal data. Other classifications included height fields, such as 2D scalar fields for terrain visualization and 2D scalar fields for imagery data. More advanced data types, such as 3D scalar fields (for volume visualization), geometric models, point cloud data, vector fields (for flow visualization), tensor fields (for tensor visualization), and document data (text corpora), were also considered. Furthermore, some papers focused on software and machine learning models (e.g., VIS4ML), image corpora, simulation data, event data, audio data, and video data.

The **data size** of the datasets used in the papers was assessed based on both explicit and implicit indications of data scale as reported in the reviewed papers. We used a four-level coding scale reflecting both stated and inferred characteristics: [0] small size explicitly mentioned by the authors, [1] above small size or big data, [2] observed and estimated as small size (e.g., a one-page table), and [3] observed and estimated as above small size. This coding allowed us to distinguish between explicit author statements and researcher interpretations, while maintaining consistency across diverse publication styles. The grouping into a binary classification (small vs. above small or big data) in the visualization readiness analysis was performed later to simplify comparison and to ensure an adequate sample size per category.

The **user types** targeted by the visualizations were classified into three groups: the public (non-specialists and general audiences), domain experts (including subtypes of expertise specific to various fields), and policymakers (decision-makers and government officials).

We also evaluated whether the papers employed **advanced algorithms**, classifying them as 1) either simple algorithms or 2) using advanced algorithms. Additionally, the types of **visual representations (chart types)** used in the visualizations were analyzed to understand their diversity and usage patterns.

The **level of interaction** provided by the visualizations was categorized into four levels: 1) advanced interaction, which involved interacting directly with algorithms; 2) medium interaction, characterized by linked panels; 3) basic interaction, such as filtering and simple interactivity; and 4) static visualizations that lacked any interactivity.

The **tools and platforms** used in the papers were grouped into several categories, including 1) software tools (e.g., Tableau, Power BI), 2) computational toolkits (e.g., R, Python libraries like matplotlib, plotly, and seaborn), 3) web-based applications, cloud platforms (e.g., AWS, Google Cloud), 4) mobile applications, and 5) immersive technologies such as VR/XR/AR or physicalization.

Finally, we considered the need for **public engagement and dissemination**. Papers were categorized as (1) not requiring dissemination, typically research-focused or expert-oriented papers intended for academic or professional audiences, or (2) mentioning the need for dissemination, where authors explicitly discuss the importance of sharing findings with the public or broader stakeholders.

The data analysis presented in this section is derived from the coded data, which readers can access via `https://tinyurl.com/pandemic-survey2`, including detailed annotations and classifications applied during the systematic review process, serving as the foundation for the observations and insights discussed.

Data Types: The analysis of data types in the surveyed papers reveals notable trends in pandemic-related visualization research. The most common data type observed was tabular data, appearing in 74 papers (49%) (Vashisth et al. 2022; Jayanti and Yulyantari 2021), highlighting its centrality in pandemic analysis. Time-series data followed closely, identified in 41 papers (27%) (Sharma et al. 2020; Talavia et al. 2021), reflecting the temporal dynamics inherent in monitoring and analyzing pandemics. Other significant data types included spatiotemporal data (37 papers, 24%) (Srabanti et al. 2021; Lavanya et al. 2021), textual data (33 papers, 22%) (Tao et al. 2020; Sharma et al. 2023), and geospatial data (13 papers, 9%) (Walker and Sharma 2021; Panyadee et al. 2023). While less frequent, data types such as 2D and 3D scalar fields, trees, sets, graph/network data, video, and geometric models also appeared in specific cases, demonstrating their applicability in niche domains.

Data Size: The papers analyzed datasets of varying sizes. Datasets explicitly mentioned as small size were used in 33 papers (21.9%) (García-Cuesta et al. 2024; Cárdenas et al. 2021), while those identified as above small size or big data appeared in 56 papers (37.1%) (Khan et al. 2022; Reinert et al. 2020b). Small-sized datasets estimated by observation (e.g., one-page tables) were seen in 7 papers (4.6%) (Collins et al. 2022; Wang et al. 2023), and datasets estimated as above small size appeared in 55 papers (36.4%) (Gupta et al. 2024; Swaminathan et al. 2022). In summary, 26.5% of the papers used *small-sized* datasets, while 73.5% focused on analyzing *above small-sized* datasets.

Users: The intended user types were classified into three primary groups: *policymakers*, the *general public*, and *domain experts*. Among the surveyed papers, 59 (39%) focused on policymakers (Faris and Almir 2024; Fangfang et al. 2021), reflecting the importance of supporting decision-making processes with accessible and actionable insights. Similarly, 56 papers (37%) targeted the general public (Wang and Peng 2022; Han et al. 2024), aiming to communicate complex information to non-specialists and broader audiences.

A significant majority of 143 papers (95%) (Papageorgiadis et al. 2022; Yang et al. 2022) included domain experts as users, highlighting that visualization tools developed for this group constitute the predominant focus in pandemic-related research. The four most frequently addressed user groups within this category were public health experts, data scientists, epidemiologists, and healthcare professionals.

It is also worth mentioning that five papers discussed visualization systems explicitly designed for educational purposes (Bovo et al. 2022; Hurtig et al. 2022; Papageorgiadis et al.

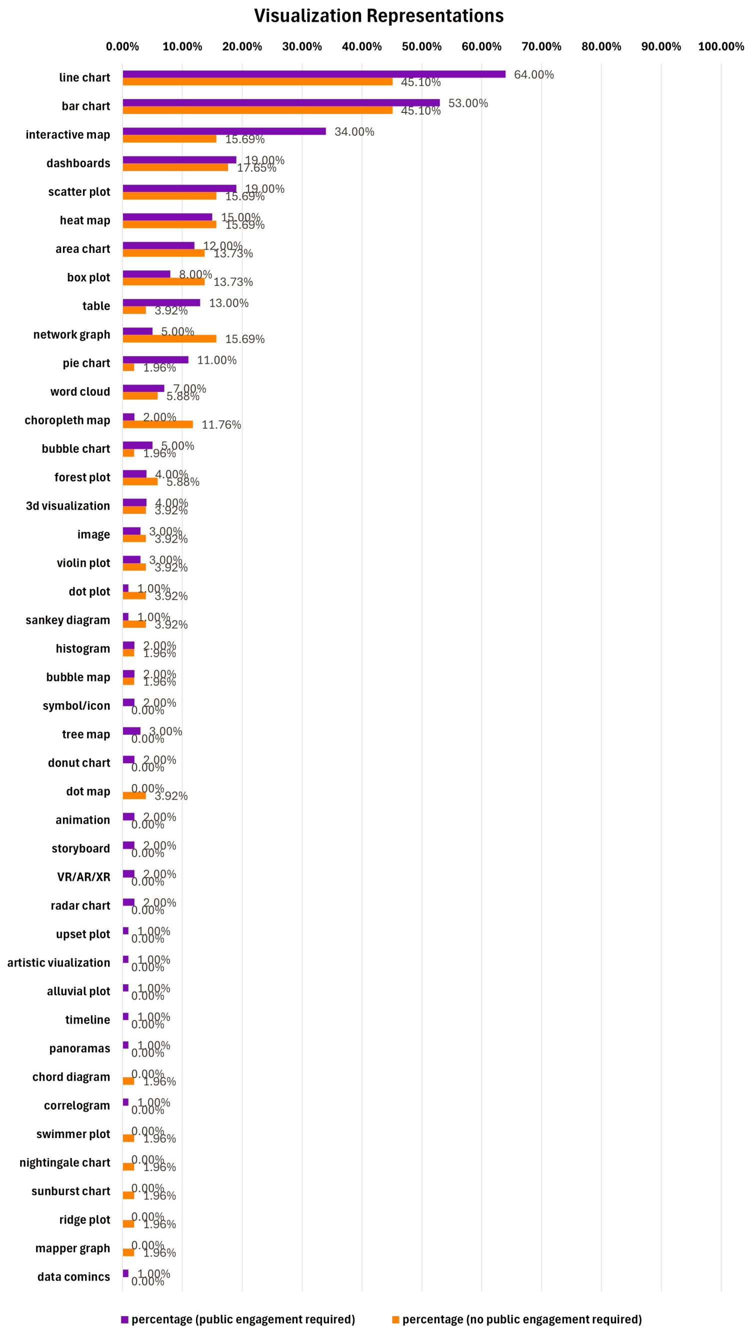


Figure 4.9: Visual Representation Type Percentage in **Survey 2** – categorized by public engagement requirement.

2022; Singla and Pushe 2023; MacLean et al. 2021). This indicates a smaller but meaningful interest in leveraging visualization to enhance teaching and learning in pandemic-related topics or under the restrictions of remote teaching.

Visual Representations: The analysis of visual representation types in Survey 2 reveals

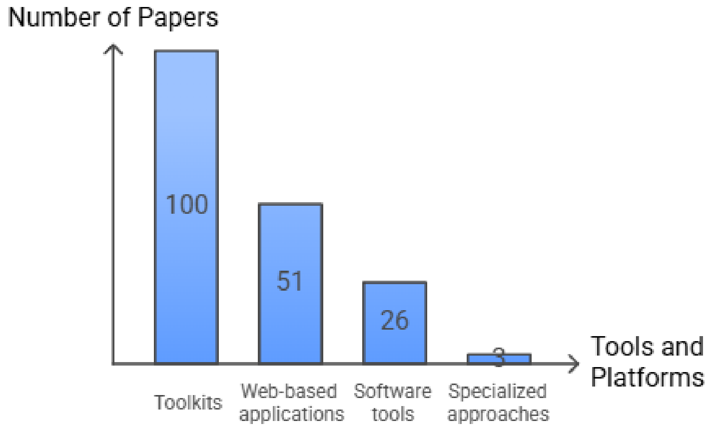


Figure 4.10: Usage of Tools and Platforms in **Survey 2**.

several prevalent choices (Fig. 4.4). The line chart was the most frequently used representation, appearing in 88 papers (56%) (Vashisth et al. 2022; Jayanti and Yulyantari 2021), followed by bar charts in 76 papers (48%) (Talavia et al. 2021; Firiza et al. 2024). Geospatial maps were also widely adopted and featured in 42 papers (27% (Afzal et al. 2020; Tao et al. 2020). Despite their limitations in advanced interaction, dashboards remained a popular representation method used in 28 papers (18%) (Usman et al. 2022; Famadico et al. 2024). Other common visuals included scatter plots (27 papers, 17%) (Wagenhäuser et al. 2024; Friedrich et al. 2024) and heatmaps (23 papers, 15%) (Banerjee et al. 2020; Kolossváry et al. 2023). Visual representations of area charts, tables, box plots, network graphs, pie charts, and word clouds were observed in 10–20 papers. Compared to the findings from **Survey 1**, the overall distribution of visual representation types is notably similar. Basic charts, such as bar and line charts, continue to dominate, reflecting their versatility and ease of use or production in communicating essential data trends and patterns. Map-based visualizations are also popular, particularly for geospatial data requiring spatial representation. Additionally, dashboards continue to be a prominent visualization method in both surveys, reinforcing the argument for the predominance of dashboards as a key tool in pandemic-related visualizations.

Level of Interaction & Advanced Algorithm: The level of interaction provided by the visual analytics tools varied across the surveyed papers. To categorize the tools' interactive features, we divided them into four levels: Advanced, Medium, Basic, and Static. The most common level observed was the basic level of interaction, identified in 74 papers (51%) (Albert et al. 2020; Gaie 2020). This level relies on simple analytical techniques such as hover-over functionality for additional details, zooming in and out on maps, or other basic interactivity features. Medium interaction, observed in 43 papers (29%) (Ambavi et al. 2020; Papageorgiadis et al. 2022), includes functionalities such as linked visual components, search and filter options, and moderate levels of interactivity. Advanced interaction, seen in only 22 papers (15%) (Mee et al. 2022; Sahoo et al. 2021), is characterized by more sophisticated, algorithm-driven analytics or visual analytics tools that allow for algorithm adjustments and refinements, enabling high user-data engagement. Lastly, static visualizations, which lack any form of interactivity - used to present data without user manipulation - were used in 7 papers (5%) (Ahmed et al. 2023; Hartanto et al. 2022).

For algorithmic complexity, the classification into "simple algorithms" and "advanced algorithms" was based on the methodological detail provided in each paper. Simple algorithms included descriptive or rule-based techniques, such as basic statistics, clustering, or regression, whereas advanced algorithms encompassed computationally intensive or adaptive approaches, including machine learning, deep learning, or optimization-based methods. Determinations were made by reviewing the methodological sections of each paper and coding according to the depth and sophistication of the described techniques.

Regarding the use of advanced algorithms, 104 papers (71%) (Jayanti and Yulyantari 2021; Sarkar et al. 2020) incorporated algorithm-driven techniques, reflecting a substantial number of works that aim to enhance analytical capabilities through machine learning, deep learning, or predictive modeling. The remaining 42 papers (29%) (Wang et al. 2022; Kharismawati Kamaludin et al. 2021) did not use advanced algorithms.

Public Engagement: Regarding public engagement, 100 papers (66.2%) (Macrohon and Jeng 2021; Cui and Kong 2021) emphasized the need for dissemination to communicate information to a broader audience. The remaining 51 papers (33.8%) (Tackx et al. 2021; Min et al. 2021) did not prioritize dissemination, likely targeting specialized or expert audiences. To explore differences between papers aimed at public engagement and those not, we compared their visual representations (Fig. 4.9). Map-based visualizations were significantly more prevalent in the engagement group, highlighting the importance of spatial context for general audiences. Semantic and storytelling elements—icons, symbols, data comics, storyboards, and animations—were exclusive to this group, as were immersive approaches (AR, VR, XR) and art-based visualizations. These choices reflect a deliberate effort to enhance accessibility, engagement, and narrative appeal.

Tools and Platforms: In **Survey 1**, we highlighted a key finding: *"Many papers do not mention a collaboration with VIS researchers or practitioners, which may result in a reliance on basic visualization techniques."* This lack of collaboration may contribute to the absence of advanced visualization techniques and a reliance on simple visualization plots. To further investigate, we hypothesized that teams without VIS experts may face limitations in producing and developing visualizations or visual analytical tools. These constraints could indirectly lead to the predominance of simple charts and basic interactions as the main approaches. To explore this, **Survey 2** included a detailed analysis of the VIS artifacts produced and the deployment platforms utilized across the papers.

As shown in Fig. 4.10, the surveyed papers used various tools and platforms. R and Python libraries were the most common, appearing in 100 papers (66%), highlighting their flexibility and accessibility. Web-based applications using technologies like HTML, CSS, JavaScript, and D3.js appeared in 51 papers (34%), emphasizing interactivity. Software tools like Tableau and Power BI were used in 26 papers (17%), though less frequently due to limited customization. Specialized tools, such as VR, XR, AR, and mobile apps, were rare, appearing in only three papers, likely for niche use cases.

This analysis questions whether the tools used for visualizations or visual analytics are related to the level of interaction in the systems. We calculated correlation coefficients be-

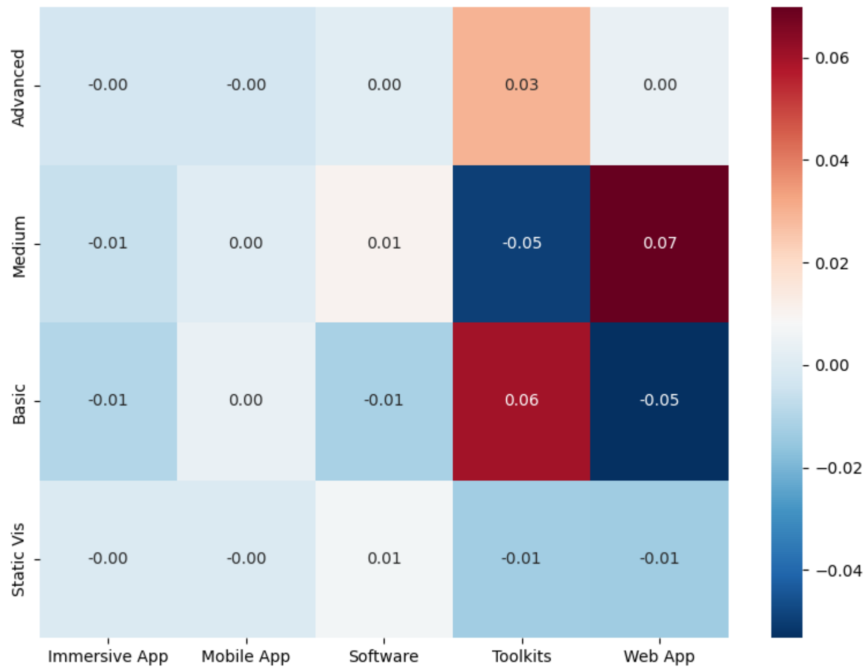


Figure 4.11: The correlation heatmap illustrates the relationship between the tool/platform and the level of interaction in **Survey 2**.

tween the tools/platforms and interaction levels, as shown in Fig. 4.11. The findings suggest potential growth in advanced web-based visual analytics for pandemic research. The usage patterns highlight the need for collaboration with visualization and engineering experts to develop more advanced, interactive systems for complex data analysis and decision-making.

O2a **Survey 2** confirms **Survey 1** findings, showing datasets from diverse sources—empirical observations, digital records, simulations, and social media (Fig. 4.8). The prominence of geospatial and temporal data underscores their role in tracking cases and infection rates, while social media text highlights public sentiment. These data types present opportunities for visualization, requiring intuitive mapping tools for geospatial/temporal data and advanced sentiment analysis and topic modeling for textual data.

O2b The survey highlights a trend toward handling large-scale datasets in pandemic-related visualizations, emphasizing the importance of advanced visualization techniques to make such data comprehensible and actionable.

O2c Dashboards gained significant popularity starting in 2020, coinciding with the COVID-19 outbreak (Fig. 4.12). They were heavily used for pandemic data visualizations due to their ability to integrate, aggregate, and present real-time data concisely. However, this trend diminished as the pandemic eased, partly due to decreased pandemic-related research papers.

O2d Fig. 4.9 shows varying interactivity and algorithmic sophistication in pandemic visualizations. While many papers use advanced algorithms (ML, deep learning, LLMs), their visual tools often feature only basic to medium interactivity. This gap presents an opportunity to enhance the synergy between computation and user interfaces. Future work should improve interactivity to leverage advanced algorithms fully.

O2e Fig. 4.9 shows that map-based visualizations were more common in the public engagement group, emphasizing their importance to general audiences. Storytelling elements like icons, symbols, data comics, storyboards, and animations appeared exclusively in this group, reflecting efforts to enhance accessibility and engagement.

O2f Our survey revealed that web-based applications, while flexible, typically had medium interaction levels (Fig. 4.11). The tools used in the computational toolkits ranged from basic to advanced interactions. Although the VIS community regularly uses D3.js for sophisticated visuals, pandemic research favors computational toolkits for efficiency and simplicity, as shown in Fig. 4.10.

4.2 The Future Roadmap

In **Survey 0**, a search for pandemic-related papers returns a result of 350,000 (out of which we selected 106 papers), indicating a clear demand for more extensive VIS support. This finding is echoed in **Survey 1**, where the 81 papers selected from 127 published in the VIS and HCI venues provide evidence that VIS researchers assisted domain experts in multiple sectors, including health experts, policymakers, and epidemiologists. In addition, the 81 papers published in other venues highlight the urgent need for broader VIS support. These findings were also observed in **Survey 2**, where 44, 75, and 32 papers were collected from 320, 4,435, and 532 articles from ACM, IEEE, and the Lancet, respectively. **Survey 2** also found an increase in the size of the data set and the types of data, specifically geospatial and temporal data.

The findings of the three surveys can be summarized as the following conclusions:

C1 [Observation **O0a**, **O1a**, **O1b**, **O1d**] Many pandemic-related areas and domain experts need VIS support, especially during a pandemic. Still, insufficient resources are available to support them (e.g., infrastructures, tools, and VIS specialists). Figure 4.4 shows that in several key areas, such as epidemiological modeling, healthcare analytics, and decision-support systems, the majority of work appeared in non-VIS venues, with little to no evidence of VIS-specialist collaboration. Although most VIS researchers are engaging with practical applications, the current size of the VIS research community does not scale up with the number of pandemic-related areas and domain experts needing VIS support.

C2 [Observation **O0b**, **O1c**, **O2c**, **O2d** **O2e**] Visualization is commonly perceived as a tool primarily for dissemination, though in practice, many pandemic-related areas and domain experts can benefit from exploratory, observational, analytical, and model-develop visualization (Chen et al. 2022; Dykes et al. 2022).

C3 [Observation **O0c**, **O1c**, **O2d**, **O2f**] There is little doubt about the role of visualization in data surveillance and information dissemination during a pandemic. Although VIS may have supported data analysis and decision-making more directly, more persuasive evidence (e.g., theoretical reasoning, empirical findings, practical case studies) would help make VIS a fixture in these areas. The COVID-19 pandemic highlights the need to enhance the use of VIS in decision-making and public engagement by refining our

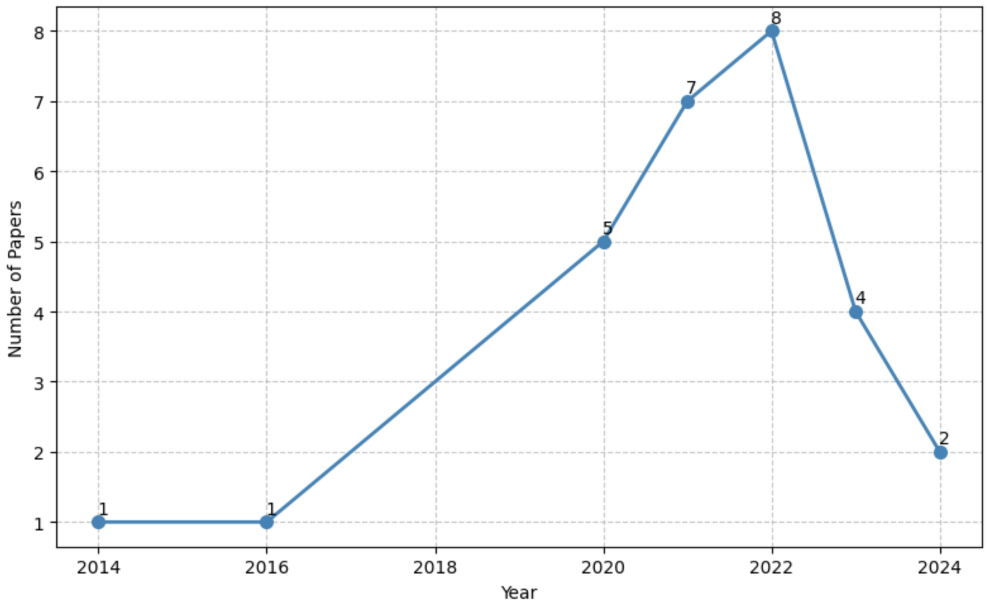


Figure 4.12: A line chart summarizing the dashboard usage per year in **Survey 2**.

theoretical and empirical understanding of VIS techniques and systems. Our analysis suggests that advanced visualization techniques could improve the quality and effectiveness of these works. Collaborating with VIS researchers and practitioners could lead to more insightful and impactful data representations, ensuring technical accuracy, accessibility, and innovative visualizations tailored to research objectives and audience needs.

C4 [Observation **O2a**, **O2b**] During the COVID-19 pandemic, many users had to process and deal with very large datasets, geospatial data, and temporal data. Pandemic-related data sets come from multiple sources, including geospatial and temporal data. These data types are crucial for tracking cases and infection rates, highlighting the need for visualization tools tailored to spatial and temporal analysis. Given the complexity of the data, intuitive geospatial mapping tools and sophisticated text analysis methods are necessary to make the data actionable. In addition, pandemic-related visualizations increasingly involve large datasets, which require advanced visualization techniques to ensure that the data remains interpretable and useful for decision making.

To address these challenges, the VIS community should build on its contributions and the experience gained during the COVID-19 pandemic and make serious efforts to advance VIS as a scientific field and ubiquitous technology. Research in these areas will directly tackle the mentioned challenges.

E1 **VIS-enabled Infrastructures** [Addressing **C3**, **C4**] – COVID-19 highlights the critical need for having the infrastructure to support large-scale data collection and surveillance, data analysis and visualization, and model development and deployment (Chen et al. 2022). Although VIS specialists demonstrated their capabilities to develop large systems during the pandemic (Khan et al. 2022; Jentner et al. 2023), ideally, there would be readily available infrastructure to provide routine VIS services in non-pandemic periods for domain experts in many pandemic-related areas, which could be deployed quickly when a pandemic emerges. The routine VIS services expose domain experts to more advanced VIS techniques while appreciating the broader range of VIS applications, in addition to disseminative visualization.

E2 **VIS Software** [Addressing **C3**, **C4**] – While almost all domain experts can use plotting facilities in some software, most do not have the experience of using more advanced VIS tools. During 2020, there were only a few dashboards for observing COVID-19 data, and the broader deployment of dashboards for different regions occurred more than a year after the pandemic began. It is necessary to broaden the use of software systems with dashboard features, such as Tableau, so VIS specialists and users can rapidly introduce dashboards into pandemic response workflows. In addition, we need to incorporate more advanced VIS techniques into open-access software to support data analysis and model development efforts, and extend the open-source culture in VIS from visual design and plot generation (e.g., D3.js) to software for analytical and model development visualization, as well as VIS systems and infrastructures.

E3 **Visualization Literacy** [Addressing **C2**] – While it is critical for advancing VIS techniques and expanding the provision of VIS, it is also necessary to improve the general knowledge about VIS and, in particular, its uses beyond disseminative visualization. This can help domain experts quickly identify the needs for VIS in various workflows featuring data observation, data analysis, model development, and decision-making, and engage VIS researchers at an earlier stage of workflow development. Visualization literacy has been a research topic in VIS, most commonly around youth education and users' visualization skills (e.g., (Chevalier et al. 2018; Firat et al. 2022)). It is desirable to broaden the research scope to include domain experts' appreciation of when and why VIS is useful, as well as their skills in identifying VIS needs.

E4 VIS Design Methodology [Addressing **C1**] – While it is essential to engage users in selecting or designing visual representations, algorithms, and user interfaces in developing VIS tools and systems, it is also desirable to improve the efficiency of VIS design processes. During the COVID-19 pandemic, many VIS researchers had to design VIS tools and systems with limited time and user engagement. This suggests the need to advance theoretical understanding and develop requirements analysis methodologies that enable VIS researchers to identify requirements rapidly. In the VIS literature, there are several "faster" design methods that complement the traditional user-centered method, including the semi-automated "show-me" method (Mackinlay et al. 2007), the constrained programming method (Moritz et al. 2019), the guideline-based method (Meyer et al. 2015), and the "symptom-cause-remedy-side effect" method (Chen and Ebert 2019). It is highly desirable to make concerted efforts to enhance the efficiency and effectiveness of VIS design processes, particularly for emergency responses.

E5 VIS Research on Decision Support [Addressing **C3**] – While there are many case studies where VIS techniques play an integral role in decision processes, it is necessary to conduct further research to gain an in-depth understanding between the interplay among different components in a decision process, including the human mind, visualization, and algorithms. Ideally, such research will facilitate more theory-informed and systematic inclusion of VIS techniques in decision processes than the largely intuitive and spontaneous approaches.

E6 VIS Research on Public Engagement [Addressing **C2**] – The role of VIS in public engagement has been extensively demonstrated and convincingly affirmed during the COVID-19 pandemic, which also suggests the need for conducting further research into issues such as effectiveness, uncertainty, biases, and privacy in visualization, and misinformation and disinformation through visual imagery. Meanwhile, developing more advanced and deployable VIS techniques for public engagement, such as storytelling and personalized visualization, is highly desirable.

Fig. 4.1 summarizes the reasoning diagram of our observations, conclusions, and challenges gathered from the three-stage survey.

Based on the observations, conclusions, challenges, and recommendations presented above, We define the **Visualization Readiness Levels (VRLs)**, which classify the readiness of visualization practices in terms of key dimensions: infrastructure, software availability, visualization literacy, design methodology, decision-support integration, and public engagement. VRLs provide a structured framework for assessing the maturity and preparedness of visualization techniques, particularly in urgent societal contexts, such as pandemic response.

- **Level 0 (Absent):** No evidence of visualization engagement.
- **Level 1 (Emergent):** Dissemination-oriented visualization exists, but analytical or interactive support is limited.
- **Level 2 (Developing):** Some exploratory and analytical tools available, but not widely accessible or deployable.
- **Level 3 (Mature):** Visualization infrastructure, software, and literacy support for ongoing use.

- **Level 4 (Ready):** Visualization techniques are integrated into workflows, decision-making, and public engagement, and can be rapidly deployed in crises.

Table 4.1 maps the observations (O0a–O2f), conclusions (C1–C4), and recommendations (E1–E6) to the corresponding VRL levels (0–4), illustrating how each stage of readiness corresponds to concrete VIS practices, gaps, and actionable steps. This mapping not only highlights the current state of visualization preparedness but also provides a systematic pathway for advancing readiness from absent or emergent levels to fully integrated, deployable visualization capabilities.

The VRL framework provides not only a retrospective assessment of the visualization community's pandemic response but also a methodological guide for future evaluations. Researchers can apply the framework to other urgent societal domains (e.g., climate crises, cybersecurity, disaster management) by systematically collecting evidence of visualization practices, mapping them to readiness levels, and identifying gaps across the six key dimensions represented by Recommendations E1–E6: VIS-enabled infrastructures (E1), VIS software (E2), visualization literacy (E3), VIS design methodology (E4), VIS research on decision support (E5), and VIS research on public engagement (E6).

While our review corpus did not directly mention funding constraints, we acknowledge that funding availability, especially timely and targeted support, was a significant factor affecting visualization research during the pandemic. Several reports indicate that early COVID-19 research funding was predominantly directed toward vaccine and therapeutic development, with less emphasis on data visualization and the creation of analytical tools. For example, in the UK, major rapid-response grants from UK Research and Innovation (UKRI) and the Medical Research Council (MRC) prioritized vaccine trials, epidemiological modeling, and clinical interventions, as illustrated in the Oxford–AstraZeneca vaccine development story (UK Research and Innovation 2021). Similarly, in the United States, the NIH's RADx and ACTIV initiatives channeled millions into diagnostics, vaccine development, and drug trials, but offered limited immediate calls for visualization-focused projects (National

Table 4.1: Mapping Observations, Conclusions, and Recommendations to Visualization Readiness Levels (VRLs).

VRL	**Definition**	**Observations**	**Conclusions / Challenges**	**Recommendations**
0 (Absent)	No evidence of VIS engagement in the domain.	O0a, O1a, O1b, O1d	C1: Lack of VIS engagement in many pandemic areas	E4: Improve VIS design methodology
1 (Emergent)	Dissemination-oriented visualization exists, but analytical or interactive support is limited.	O0b, O2c, O1c, O2e, O2d	C2: VIS often limited to dissemination	E3: Improve visualization literacy; E6: Improve VIS research on public engagement
2 (Developing)	Some exploratory and analytical tools are available, but not widely accessible or deployable.	O2a, O2b	C3: VIS can support decision-making and analysis but needs stronger evidence; C4: Large and complex datasets require tailored visualization	E1: Develop VIS-enable infrastructure; E2: Broaden VIS software adaptation
3 (Mature)	Visualization infrastructure, software, and literacy support ongoing domain use.	O0c, O2d, O2f	C3: VIS can support decision-making and analysis but needs stronger evidence	E1: Develop VIS-enable infrastructure; E2: Broaden VIS software adaptation; E5: Conduct research on VIS in decision support
4 (Ready)	VIS techniques are integrated into workflows, decision-making, and public engagement, deployable in crises.	–	VIS fully integrated and deployable in urgent contexts	E6: Advance VIS research for public engagement and deployable solutions

Institutes of Health 2021; Collins and Stoffels 2020; Ndwandwe and Wiysonge 2021). This emphasis was understandable given the urgency of containment and prevention, yet it meant that many visualization researchers lacked the rapid funding mechanisms necessary to pivot their work to pandemic-related applications. Without such targeted calls, research groups often had to continue their pre-pandemic projects rather than developing urgently needed visualization tools for public health decision-making.

4.2.1 Implementation of VRL Scoring and Visualization Dashboard

To make the evaluation of the visualization readiness more actionable and reproducible, a supporting scaffold that operationalizes the assessment process has been developed. This includes observable indicators, metrics, or evidence for each dimension at each VRL level, for example, documenting the availability of analytical dashboards (E2), the presence of trained VIS personnel (E3), the integration of visualizations into decision workflows (E5), or the use of public-facing visualization tools (E6). By providing this scaffold, VRL assessments become more transparent, systematic, and comparable, allowing researchers in other domains to adopt or adapt it for their own readiness evaluations.

To operationalize the VRL framework, we implemented a data-driven scoring system using Python, Pandas, and Plotly/Dash. The six dimensions, Data Size, User Group, Advanced Algorithm, Public Engagement, Interaction, and Tools/Platforms, were used to assess readiness for visualization. For each paper, a checklist was completed to record observed tokens and map each dimension to a numeric score according to predefined rules (see Table 4.2 for the full codebook). Multi-label fields (semicolon-separated) were expanded, and the maximum relevant score was used for aggregation within each dimension. A per-paper Visualization Readiness Level (VRL) was then computed as the arithmetic mean of the six dimension scores and mapped to discrete readiness levels using the following thresholds: Absent (<0.5), Emergent (0.5–1.0), Developing (1.0–1.5), Mature (1.5–2.0), and Ready (≥2.0). Each dimension, as well as the overall score, was mapped to a label corresponding to these VRL levels, facilitating an interpretable representation of visualization readiness for each paper.

To support exploration and analysis of the results, an interactive dashboard was developed using Dash and Plotly. The dashboard is publicly available: **`https://vrl.disc.ourcloud.ou.edu/`**. The dashboard includes several key components: **year range slider and publications per year bar chart** (Figure 4.13) showing the number of papers published per year with a selectable year range, enabling users to explore temporal trends;

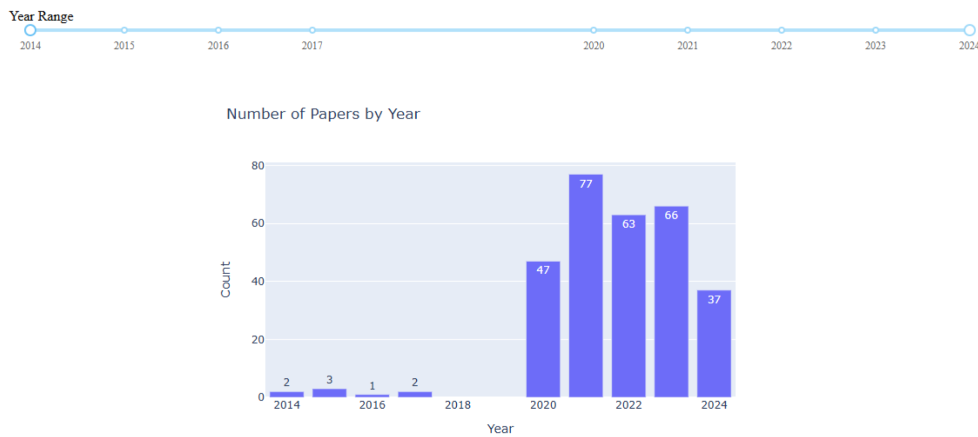


Figure 4.13: Number of papers published per year with a selectable year range, illustrating temporal trends in VRL-related research

bar charts and a radar chart (Figure 4.15) displaying the distribution of scores across the six VRL dimensions and summarizing average scores for an at-a-glance comparison of overall readiness; and **two heatmaps**, including a dynamic heatmap (Figure 4.14) that shows relationships between selected dimensions and a static co-occurrence heatmap illustrating label co-occurrences across all dimensions. Additionally, a **paper-level table** (Figure 4.16) allows filtering, sorting, and inspection of individual papers and their coded attributes for deeper analysis.

The dashboard supports interactive exploration through a year range slider (Figure 4.13) that allows users to dynamically filter the dataset by publication year. Adjusting the slider updates all other visual components in real time. Specifically, the bar chart reflecting the number of papers published per year is refreshed to display only data within the selected range. Simultaneously, the heatmap (Figure 4.14), the VRL dimension bar charts and radar chart (Figure 4.15), and the paper-level table (Figure 4.16) are all updated to reflect papers published during the chosen years. This coordinated filtering enables users to examine temporal trends, relationships among VRL dimensions, and detailed paper-level information within any specific period of interest, thereby supporting focused and comparative analysis across time.

This methodology is operationalized in an interactive dashboard, which complements the

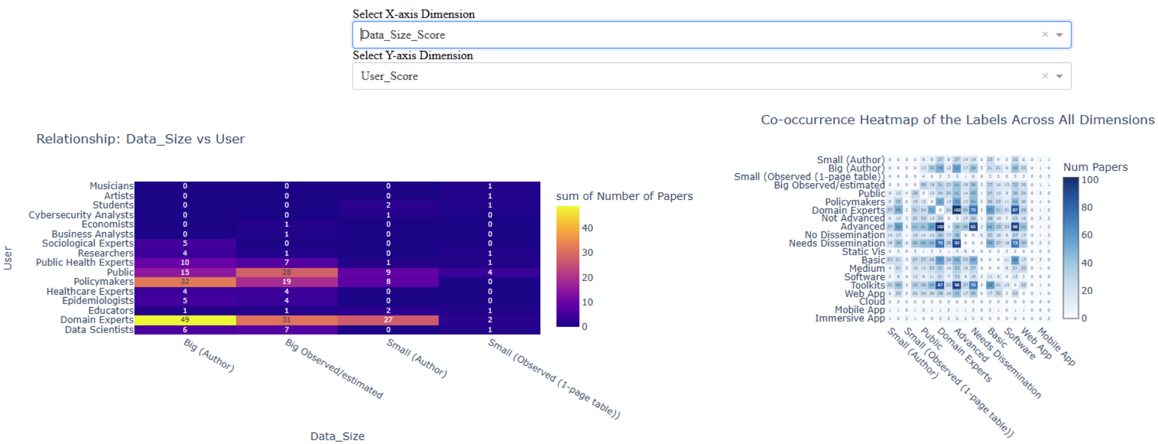


Figure 4.14: This view contains a dropdown menu and two correlation Heatmaps. The right heatmap shows the correlation of all the VRL dimensions, while the left Heatmap focuses on the comparison of the subcategories of the two selected VRL dimensions. Users can choose dimensions from a dropdown menu to explore relationships and correlations between readiness aspects.

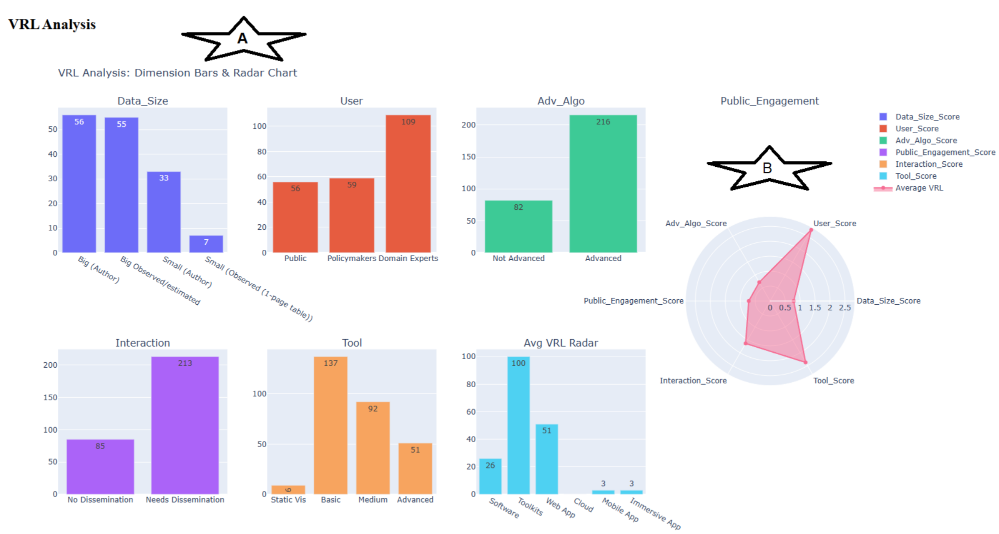


Figure 4.15: A) Distribution of scores across the six VRL dimensions, highlighting variation and patterns in visualization readiness across papers. B) Radar chart summarizing average VRL scores across all six dimensions, providing an at-a-glance comparison of overall readiness levels.

conceptual VRL framework. The numeric scoring and visualization of each dimension directly correspond to the six key aspects outlined in Table 4.2. Moreover, the per-paper and aggregated VRL scores reflect the readiness levels described in Table 4.1, illustrating how observations, conclusions, and recommendations map onto concrete visualization practices.

Paper-Level Details

Paper ID	Paper Title	Publisher	Venue
filter c			
IEEE01	Real-Time Data Visualization to Enhance Situational Awareness of COVID pandemic	IEEE	International Conference on Computational Science and Computational Intelligence (CSCI)
IEEE02	Visualization of Covid-19 Pandemic Data: An Analysis	IEEE	International Conf. on Computation, Automation and Knowledge Management (ICCAKM)
IEEE02	Visualization of Covid-19 Pandemic Data: An Analysis	IEEE	International Conf. on Computation, Automation and Knowledge Management (ICCAKM)
IEEE02	Visualization of Covid-19 Pandemic Data: An Analysis	IEEE	International Conf. on Computation, Automation and Knowledge Management (ICCAKM)
IEEE02	Visualization of Covid-19 Pandemic Data: An Analysis	IEEE	International Conf. on Computation, Automation and Knowledge Management (ICCAKM)
IEEE03	Interactive Data Visualization: Open Data of Pandemic Disease	IEEE	International Conference on Cybernetics and Intelligent System (ICORIS)
IEEE04	Data Visualization Platform and Tourism Resilience in the Post COVID-19 Pandemic in Chiang Mai, Thailand	IEEE	Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electron
IEEE05	Covid-19 Tracker: A data visualization tool for time series data of pandemic in India	IEEE	IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE)
IEEE06	Daily Visualization of Statewide COVID-19 Healthcare Data	IEEE	Workshop on Visual Analytics in Healthcare (VAHC)
IEEE06	Daily Visualization of Statewide COVID-19 Healthcare Data	IEEE	Workshop on Visual Analytics in Healthcare (VAHC)

1 / 30

Figure 4.16: Interactive table displaying detailed information for each of the reviewed papers, including coded attributes and VRL scores, with filtering and sorting functionality for deeper analysis.

By integrating the scoring methodology with interactive visualizations, the dashboard provides a practical tool for assessing visualization readiness, identifying gaps, and monitoring progress over time, making the VRL framework actionable, reproducible, and adaptable for future studies in pandemic response or other urgent societal contexts.

4.3 Summary

Our survey analysis highlights the necessity for more advanced and interactive visualization tools in pandemic-related research. It reveals a gap between computational capabilities and the interactivity of existing visualization tools. The findings emphasize the importance of incorporating diverse data sources, such as geospatial, temporal, and social media data, which require specialized visualization techniques. Although dashboards gained popularity during the COVID-19 pandemic, their effectiveness diminished as the pandemic came to an end. This presents an opportunity to improve the integration of advanced algorithms with more interactive visual tools. The survey advocates for increased collaboration with visualization (VIS) and visual analytics (VA) experts to develop more effective data analysis and decision-making systems. This collaboration will ensure that visualizations effectively communicate information.

The survey also highlights several key challenges: a lack of resources for pandemic-related visualization, insufficient recognition of visualization as an analytical tool, and the need for more substantial empirical evidence of its impact. To address these issues, expanding infrastructure, enhancing software capabilities, and improving visualization literacy are essential, all while promoting interdisciplinary collaboration. Future research should focus on evaluating the role of visualization in decision support, improving design efficiency, and developing rapid prototyping methods for emergency response, thereby ensuring better preparedness for future global crises. Additionally, more research is needed to understand the impact of visualization on decision-making and public engagement and to streamline design processes, especially in emergencies. In summary, we conclude the following challenges:

VIS-Enabled Infrastructures: Establishing permanent VIS infrastructures will ensure readiness for future pandemics and facilitate routine support for domain experts.

Advanced VIS Software: Expanding access to and improving VIS tools, particularly open-source solutions, will enhance real-time pandemic response.

Visualization Literacy: Educating domain experts about the broader applications of VIS beyond dissemination can improve early adoption in workflows.

Efficient VIS Design Methods: Developing faster, user-centered design methodologies will enhance the adaptability of VIS tools during emergencies.

VIS for Decision Support: Research on how VIS interacts with human decision-making and algorithms will lead to more structured, theory-driven applications.

VIS for Public Engagement: Further investigation into VIS effectiveness, biases, and misinformation handling is needed to improve public trust and communication strategies.

Based on the observations, conclusions, challenges, and recommendations presented above, we defined the **Visualization Readiness Levels (VRLs)**, which classify the readiness of visualization practices in terms of key dimensions: infrastructure, software availability, visualization literacy, design methodology, decision-support integration, and public engagement.

VRLs provide a structured framework for assessing the maturity and preparedness of visualization techniques, particularly in urgent societal contexts, such as pandemic response.

In addition, an interactive publicly available dashboard is accessible at **`https://vrl.disc.ourcloud.ou.edu/`**, which was developed to facilitate exploration and analysis of these results. The dashboard allows users to select a year range using an interactive slider, dynamically filtering all visualizations, including the publication bar chart, heatmaps, dimension-specific bar charts, radar chart, and paper-level table. This interactivity enables users to examine trends across time, explore relationships between VRL dimensions, and inspect individual papers in detail. By linking the VRL scoring framework with an interactive visualization interface, this tool provides both a methodological foundation and a practical means for assessing visualization readiness across studies.

Table 4.2: Coding Book for Visualization Readiness Review

Category	Details and Coding Rules
Data Type	Includes types of datasets analyzed or visualized in the paper, such as: Textual data, Tabular data, Tree/Set/Graph/Network data, Geospatial data, Time-series data, Spatiotemporal data, Height fields (2D surface/terrain), 2D Scalar Fields (imagery), 3+D Scalar Fields (volume visualization), Geometric Models, Point Cloud data, Vector/Tensor Fields, Document Data (Text Corpus), Software & ML Models (VIS4ML), Image Corpus, Ensemble Datasets, Event Data, and Video Data.
Data Size	[0] Author explicitly stated small data size. [1] Author stated or implied large/big data. [2] Observed or estimated as small (e.g., one-page table). [3] Observed or estimated as large-scale or complex data.
User Type	Categorized based on the intended or evaluated audience. Each paper may include multiple user types (separated by ";"): [Public] Non-specialists or the general public; [Policymakers] Decision-makers and institutional stakeholders; [Domain Experts] Researchers, scientists, or technical professionals.
Advanced Algorithm Usage	[0] – Visualization does not integrate or depend on advanced algorithms (e.g., purely descriptive or observational); [1] – Visualization incorporates or builds upon advanced computational or machine learning algorithms.
Task Type	Indicates the primary purpose of visualization: [Visualization for observing data] [Visualization for disseminating information] [Visualization for analyzing data] [Visualization for model development].
Visual Representations (Chart Types)	Categories derived from the visualization taxonomy spreadsheet, including bar charts, line charts, scatter plots, networks, maps, timelines, or custom hybrid visualizations. Multiple chart types can be assigned per paper.
Level of Interaction	[0] Static Vis – no user interaction; [1] Basic – limited filtering or selection; [2] Medium – coordinated or linked views; [3] Advanced – interactive algorithms or real-time updates.
Tools / Platforms	Classified into main technological categories: [Software] General visualization software (Tableau, Power BI, etc.); [Toolkits] Computational Toolkits (Python, R, D3.js, etc.); [Web App.] Browser-based visualization systems; [Cloud] Cloud platforms (AWS, Google Cloud); [Mobile App.] Mobile-based tools; [Immersive] VR/AR/XR applications or physicalizations.
Public Engagement	[0] – Visualization not designed for dissemination or public outreach; [1] – Visualization includes dissemination, outreach, or engagement component (e.g., dashboards, public reporting platforms).

Chapter 5

Design of the system Framework

In this chapter [1], we represent the *Predictive Intelligence for Pandemic Prevention (PIPP)* Dashboard, a novel visual analytics platform designed to support early detection and response to emerging infectious diseases through the integration of diverse One Health datasets. By combining human, animal, environmental, and digital surveillance data, the system enables a comprehensive exploration of factors influencing disease dynamics. The dashboard provides interactive and user-friendly visualizations, including dynamic maps, hierarchical sunburst charts, and automatically updating time-series views, allowing users to monitor weekly changes across states and counties.

Through its data-driven and predictive capabilities, PIPP facilitates the timely identification of potential public health threats, promoting evidence-based decision-making among researchers, policymakers, and health officials. By transforming complex multisource data into actionable insights, the system exemplifies the power of visual analytics in pandemic preparedness and prevention.

The novelty of this work lies in several key aspects: a) **Integration of Diverse Datasets:** The project uniquely combines different types of "One Health" data, which includes human, animal, weather, geospatial, and Google Trends data. This comprehensive integration is

[1] This chapter is based on the published paper in HICSS 2025: A. Basheer, W. Jentner, and D. Ebert, "Leveraging Visual Analytics and Diverse Datasets for Proactive Pandemic Surveillance: A One Health Approach," HICSS, 2025, 1499-1508.

crucial for early detection and prompt public health actions during a pandemic. Combining these datasets enables the identification of correlations that would not be apparent from a single dataset alone. For example, a correlation between a spike in Google searches for flu symptoms, a drop in temperatures, and an increase in animal deaths could indicate an emerging zoonotic threat. b) **PIPP Dashboard:** The creation of an interactive dashboard that offers multiple views and visualization options for different types of One Health datasets is innovative. The dashboard enables users to select specific states and counties, view interactive maps, utilize hierarchical data visualizations such as the sunburst chart, and visualize data through dynamic line charts that update automatically to display weekly changes. c) **User-Friendly Data Visualization:** The emphasis on developing intuitive and interactive data visualization models simplifies the presentation of complex information. The interactive visualizations on this dashboard make complex data relationships easy to understand, allowing experts and the general public to quickly grasp important insights. d) **Highlighting Visual Analytics Techniques:** By showcasing the significance of visual analytics, the project turns raw data into actionable insights. This is particularly important for managing the spread of infectious diseases and making informed decisions in public health. Additionally, it helps to highlight new trends and abnormalities, enabling the prompt identification and response to potential pandemic threats. e) **Timely Identification of Pandemic Threats:** The project advocates for the early identification of potential public health threats by leveraging advanced analytics and AI/ML techniques. This proactive approach is essential for mitigating the impact of public health emergencies. f) **Collaboration with Public Health Officials:** The project collaborates with public health officials to develop a comprehensive visual analytics system that integrates traditional and new health surveillance data sources. This collaboration ensures that the system is aligned with real-world needs and practices. g) **Customizable and Scalable System Design:** The system is designed to be flexible, allowing for adding new data sources and analytics methods as needed. This adaptability ensures that the system can evolve and remain relevant as new challenges and data sources emerge. h) **Case Study on COVID-19:** Using forecasting methods to predict COVID-19 deaths and cases as a case study demonstrates the practical application and effectiveness of the developed tools and techniques.

To ensure that each novel aspect of this work was substantively validated, I designed a multi-layered evaluation strategy that combined quantitative, qualitative, and case-based analyses. Specifically:

- **a) Integration of Diverse Datasets:** We empirically demonstrated cross-domain correlations, such as between COVID-19 cases, NET temperature indices, and Google search trends, that became apparent only when datasets were integrated, as visualized through the lag-aware heatmaps (Figure 5.1d).

- **b–d) Visualization and Analytical Design:** Usability and interpretability were validated through qualitative feedback from epidemiologists and public health officials, who confirmed that the system enhanced exploratory analysis and hypothesis generation.

- **e–g) Predictive and Collaborative System Design:** Forecasting performance was validated quantitatively using MAE, RMSE, correlation, and CI metrics (Table 5.1), demonstrating high predictive accuracy and reliability. Collaboration with public-health stakeholders ensured that system requirements and visualization workflows aligned with operational needs.

- **h) Case Study on COVID-19:** The Oklahoma COVID-19 case study served as a holistic validation of system capabilities, confirming that integrated data sources and predictive methods yielded actionable epidemiological insights.

These novel contributions collectively enhance the ability to monitor, understand, and respond to pandemics, making the project a significant advancement in the field of public health and pandemic prevention.

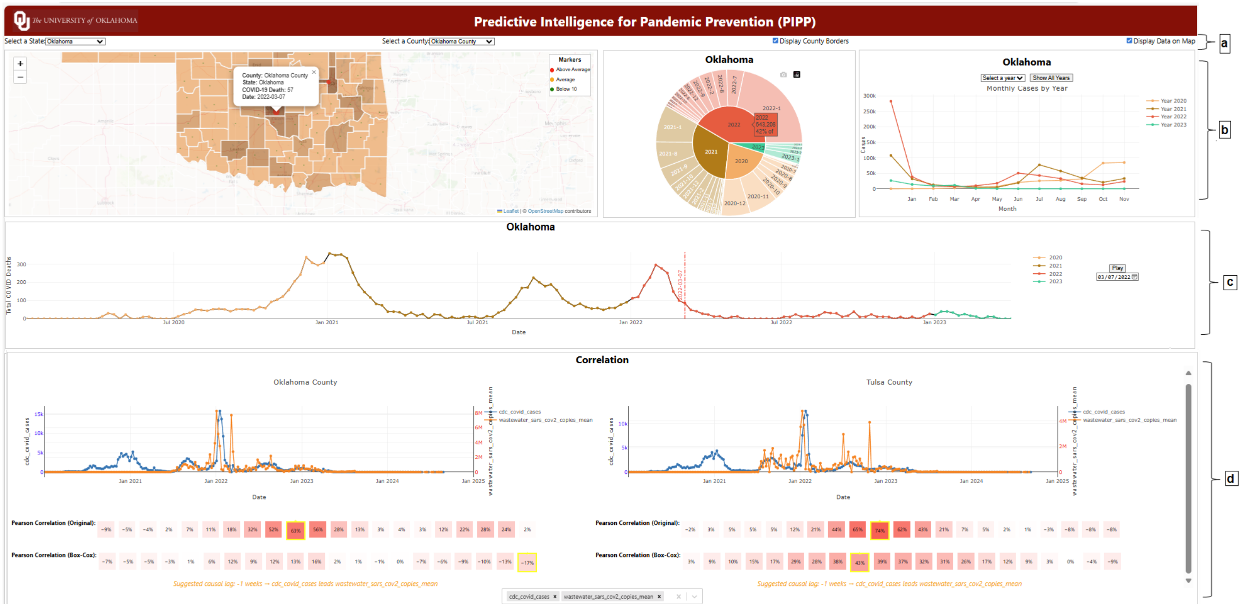


Figure 5.1: Predictive Intelligence for Pandemic Prevention (PIPP) Dashboard. The dashboard has many views, such as a) Options Panel to select a region of interest (state / specific county). b) Three types of interactive visualizations: An interactive map exploring disease spread in the region using a tooltip, Sunburst, and line charts to visualize the data hierarchically. c) A line chart that displays COVID-19 deaths or cases for the whole state with an automatic timer that will go through the line chart to display the information for every week, and at the same time, it changes the details on the interactive map for each county. d) Combined Datasets: Each selected county has a line chart with two different datasets and heatmaps to visualize the correlation.

5.1 System Description

This section provides more details on our developed dashboard and prediction models.

5.1.1 System Goals

The system goals are derived from existing literature, expert views, and the practical needs of public health surveillance, ensuring that the developed dashboard aligns with requirements for effective pandemic prevention, detection, intervention, and management.

Exploring the Relationship Between the Disease Spread and Other One Health Datasets: The primary goal is to investigate the relationship between human health surveillance aspects and other One Health datasets, as well as infection predictive trends at the

Figure 5.2: We have a line chart for each time series used in the PIPP Dashboard.

county and state levels, to enable responses and interventions based on human, animal, weather, and Google Trends data. By analyzing these datasets at the county level, the system enables users to identify potential correlations or patterns that could help understand the outbreak dynamics at each location, as transmission modalities vary based on community demographics (Clarkson 2023).

Developing User-Friendly Visual Analytics System: We develop a user-friendly system that enables users to zoom in, filter by criteria such as time and location, and explore data more effectively. By leveraging visual analytics, we highlight the importance of these techniques in managing One Health issues, enabling the integration of complex datasets to identify patterns, trends, and correlations, ultimately supporting better decision-making.

Advocating for Timely Identification of Potential Pandemic Threats: Our work emphasizes the significance of identifying potential pandemic threats promptly. Proactive surveillance activities and interventions initiated before the transmission of a pathogen be-

Figure 5.3: The stages of the Prediction Process in Our System.

comes widespread can be more effective in reducing the scope and impact of an outbreak than reactive measures taken after the pathogen has spread extensively. Swift and effective public health responses are crucial for minimizing human suffering and mitigating the broader social and economic effects of disease outbreaks.

5.1.2 Task and design approach

The project employs an integrative approach, combining multiple data streams for comprehensive analysis through visualizations such as choropleth maps and multi-series line charts, thereby enhancing insights into pandemic dynamics. Developed as a scalable web framework, it is a robust tool for detailed data exploration in public health research.

Integrative and Layered Approach: The dashboard employs an integrative, layered, and parallel approach across multiple surveillance data streams, including environmental, infectious disease, and open-source intelligence sources. This approach ensures comprehensive data coverage and facilitates a holistic analysis.

Choropleth Map Design: a) A choropleth map visualizes variables normalized by the region's population. This design choice is deliberate because choropleth maps are widely used and easily interpreted, even by untrained users. Users can select which variable to display on the map, which defaults to showing counties in Oklahoma state (Figure 5.1b). b) A tooltip appears on hovering over regions, providing detailed information, which enhances user interaction and understanding.

Tempo-Spatial Data Visualization: The dashboard features a map and line charts to visualize tempo-spatial data. This combination allows users to see both spatial distributions and temporal trends of selected regions and variables of interest (Figure 5.1c).

Temporal Overview with Line Charts: a) A line chart visualizes the selected variable from the map as a time series, aggregated for all regions or a user-selected specific region. This chart provides a comprehensive temporal overview, encompassing the minimum and maximum times of all variables displayed in the dashboard. b) The chart includes a brushing feature, allowing users to select a specific time range, which is then highlighted and applied to all other line charts by re-scaling their x-axis (focus+context technique).

Comparative Analysis with Dual-Axis Line Charts: The panel below the main chart shows two dual-axis line charts for two different selected regions, allowing users to compare variables from arbitrary data sources across regions (Figure 5.1d). To further enhance exploratory analysis, we included lag-aware Heatmap correlation that captures the temporal relationships between two selected data series, as we can see in (Figure 5.1d). The heatmap shows correlation values across a range of lags from -10 to +10 weeks. This design enables users to investigate how shifts in one variable may precede or follow changes in another, which is particularly relevant for early-warning signals in pandemic surveillance. Interactive features enable users to hover over each heatmap cell, displaying the exact correlation value at a given lag and providing immediate feedback on potential lead–lag effects between the two time series.

Multi-Series Line Charts for Detailed Analysis: Each of the four data sources has its own panel (Figure 5.2e), featuring two multi-series line charts for two selected regions. Hovering over these charts displays a vertical line in all other line charts, aligned to the date closest to the hovered date, along with a tooltip detailing the actual numbers.

Enhanced Data Correlation and Insight: By combining diverse datasets into a single visualization platform, the dashboard enables users to gain a more comprehensive understanding of the factors influencing the spread and impact of a pandemic. For example, users can overlay pandemic case data with environmental factors, such as temperature and humidity, to visualize the correlation.

User-Friendly Web Framework: The application is developed as a web framework using Python for the data, correlation, and prediction API, and React + Recharts for displaying

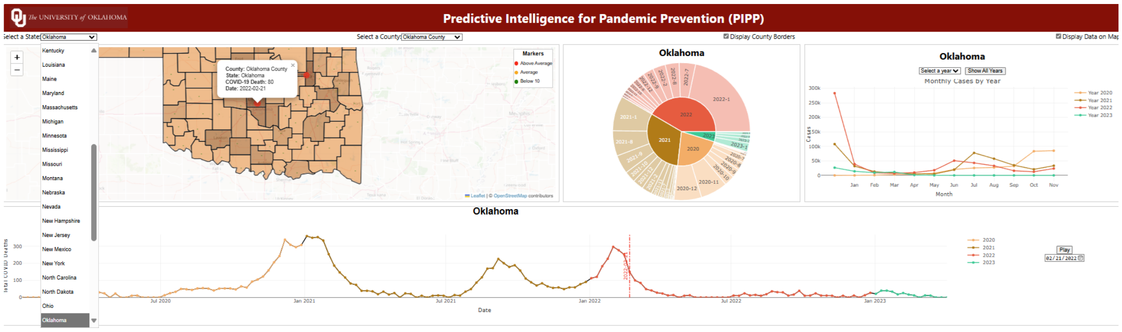


Figure 5.4: A dropdown menu to select a state.

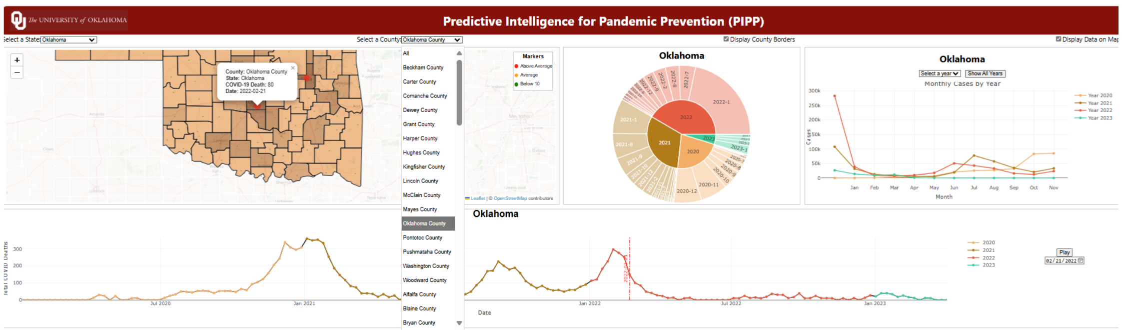


Figure 5.5: A dropdown menu to select a county.

the data as a static website. This choice ensures scalability, accessibility, and ease of use for a broad audience.

By adhering to these design rationales, the dashboard effectively supports the visualization and analysis of complex, multi-dimensional data, aiding in the early detection and management of public health emergencies.

5.1.3 User Interaction and Exploratory Workflow

To complement the system architecture and design rationale, this section illustrates how users interact with the PIPP Dashboard through a series of dynamic and intuitive operations. The dashboard employs a coordinated multiple-view design, where actions taken in one view (e.g., the map) are immediately reflected in other visualizations (e.g., time-series plots and correlation heatmaps). Figures 5.4–5.10 illustrate the step-by-step workflow.

Step 1: Selecting a Region of Interest. Users begin with a toolbar that contains a dropdown menu (Figure 5.4) listing the states, and another dropdown menu (Figure 5.5) dis-

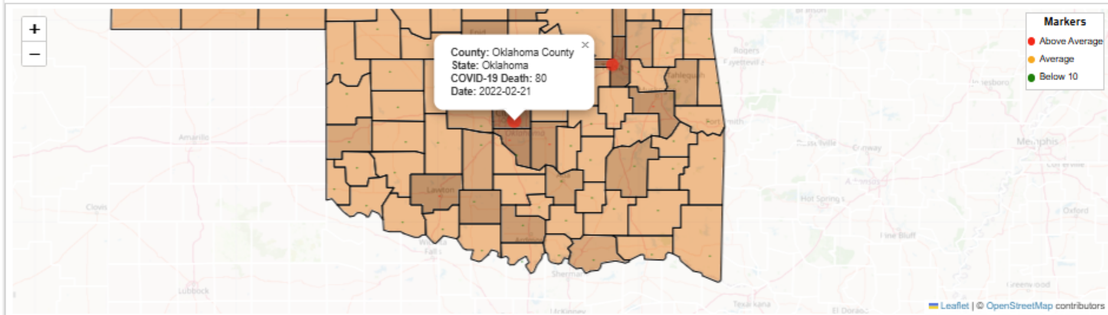


Figure 5.6: A map view for the selected state that shows the county borders.

playing the counties of the selected state. Figure 5.6 displays a choropleth map of Oklahoma, which encodes a selected health variable (e.g., normalized COVID-19 deaths). Hovering over a county displays a tooltip summarizing key indicators for that region. Selecting a county filters the dataset and triggers updates in the linked charts below.

Each county is overlaid with circular markers whose size and color encode the severity of the selected indicator. The marker radius scales with the magnitude of the variable (e.g., the square root of COVID-19 deaths), allowing users to visually compare relative intensity across regions. The color represents categorical thresholds of severity — red indicates values above the state average, orange represents moderate values near the average, and green marks regions with comparatively low values. This dual encoding provides an intuitive visual cue for identifying spatial disparities and high-risk regions at a glance.

Step 2: Exploring Data Hierarchies. There are two linked visualizations, the **Sunburst chart** (Figure 5.7-a) and a **hierarchical line chart** (Figure 5.7-b), to reveal multiscale temporal relationships within the data. These views enable users to explore disease dynamics across nested temporal levels (year, month, and week) for the selected region.

The *Sunburst chart* (Figure 5.7-a) encodes the hierarchical structure of COVID-19 cases, where each ring represents a different temporal resolution. The central node corresponds to the year, and the surrounding arcs correspond to individual months. The size of each arc reflects the aggregated number of new cases for that time period, and the color follows a predefined palette distinguishing years (e.g., orange for 2020, brown for 2021, red for 2022, and green for 2023). This color scheme is harmonized with the line chart to ensure visual

Figure 5.7: Exploring data hierarchies through a) Sunburst Chart and b) Hierarchical line chart.

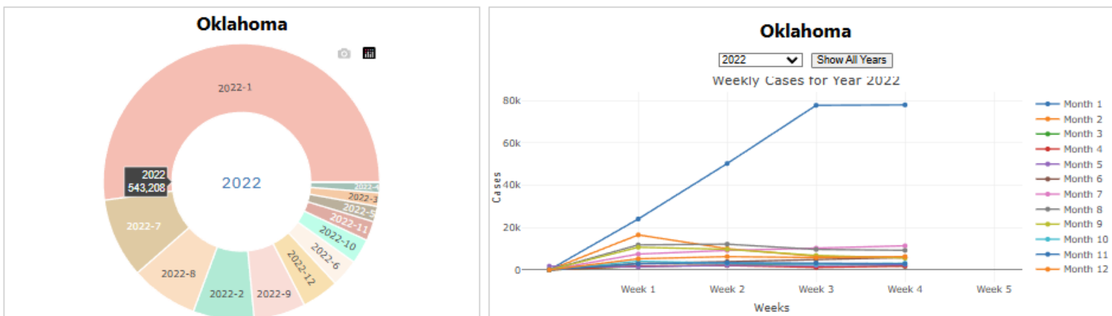


Figure 5.8: Exploring data synchronization between the sunburst and hierarchical line charts.

continuity between the two representations. Users can interactively click a ring segment to drill down into a specific year or month, dynamically updating the linked line chart.

The accompanying *line chart* (Figure 5.7-b) displays the temporal trend corresponding to the selected level of detail. When no specific year is selected, it shows monthly totals across all years, highlighting seasonal variations. When a specific year is selected (either from the Sunburst or directly within the line chart), both visualizations update synchronously: the Sunburst expands to show monthly-level detail for that year, while the line chart transitions to weekly aggregation, plotting weekly case counts per month within that year, as shown in Figure 5.8. Hovering over a line segment reveals the exact week, triggering synchronized updates on the map (Figure 5.6) to highlight the corresponding spatial distribution for that date. This interaction provides a cohesive linkage between hierarchical temporal analysis and spatial visualization, allowing users to fluidly transition between macro and micro views of disease progression.

Step 3: Temporal Playback and Cross-Panel Synchronization. The third panel

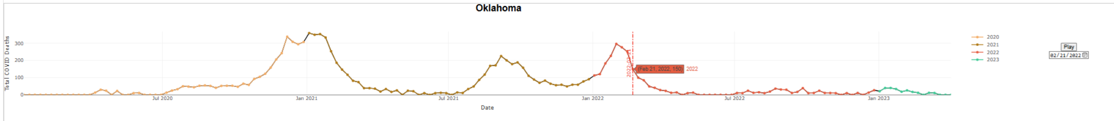


Figure 5.9: A dynamic line chart that shows the temporal evolution of disease spread at the state level.

focuses on the temporal evolution of disease spread at the state level (Figure 5.9). It introduces a **dynamic line chart** synchronized with all other temporal visualizations across the dashboard. This coordination enables users to explore disease dynamics seamlessly over time within a unified temporal context. The line chart displays weekly aggregated COVID-19 cases or deaths, with time represented on the x-axis and counts on the y-axis. Consistent with the color hierarchy established in the earlier panels, each year is represented using the same color scheme: orange for 2020, brown for 2021, red for 2022, and green for 2023, ensuring perceptual continuity across the interface.

A key feature of this panel is the *automatic temporal playback* mechanism, which simulates the progression of the pandemic over time. When activated, the timer iterates through each week chronologically, updating both the chart and the map. As the line chart's indicator moves along the timeline, the corresponding week's data are highlighted, and the choropleth map (Figure 5.6) simultaneously updates to reflect county-level variations for that week. Also, a vertical reference line appears and simultaneously aligns with the corresponding time points on all other line charts within the dashboard in real-time (Figure 5.2). This dynamic linkage enables users to observe spatial diffusion patterns and temporal peaks simultaneously, revealing how outbreaks emerged, spread, and declined across regions.

In addition to automated playback, users can manually pause, skip forward or backward, or hover over a specific point on the line chart to inspect a given week. This manual control offers flexibility for detailed analysis of specific events (e.g., sudden spikes or declines), while maintaining a synchronized state across both panels. Together, this interaction bridges temporal and spatial scales, enhancing users' understanding of the dynamic nature of disease progression across Oklahoma.

Figure 5.10: Comparative analysis of multiple data streams for two selected counties.

Step 4: Comparative Analysis of Multiple Data Streams. The fourth panel supports cross-dataset exploration by providing linked visualizations that enable users to compare multiple data streams for the selected two counties side by side using dual-axis line charts. This panel combines *dual-axis line charts* with a *lag-aware correlation heatmap* to reveal temporal relationships and potential lead–lag effects between variables. (Figure 5.10).

Each dual-axis line chart displays two selected variables from different datasets, such as COVID-19 hospitalizations and SARS-CoV-2 concentrations in wastewater, or disease incidence in animals and environmental measures like temperature or humidity. Users can select which variables to display on each axis, allowing flexible exploration of arbitrary data pairings.

The panel features two lag-aware correlation heatmaps: one for the original Pearson correlations and another for Box-Cox correlations between the two selected time series over a user-defined lag range (e.g., –10 to +10 weeks). Each heatmap cell represents the correlation coefficient at a specific lag, with intensity (opacity) corresponding to the magnitude of the correlation, making stronger correlations more visually prominent. Peak correlations are highlighted, and the visualization provides suggested causal directions, indicating which variable may lead or lag the other. Hovering over a heatmap cell dynamically shifts the second data series along the x-axis to visualize the corresponding lag, allowing for an immediate assessment of lead–lag relationships. This visualization allows users to identify whether changes in one variable precede or follow changes in another, providing early-warning insights and uncovering temporal patterns that may not be immediately apparent in the line charts alone.

Interactions in this panel are fully synchronized with other panels. Hovering over a specific week in either line chart highlights the corresponding time point across all linked line charts, including those in the panels above or under it. This synchronization ensures that users maintain a consistent temporal reference as they explore correlations and compare trends across datasets. Selecting a time window through brushing or playback updates both the line charts and the heatmap, enabling dynamic focus and context analysis.

Through this integrated approach, the dashboard facilitates the identification of meaningful relationships across diverse One Health datasets. Users can investigate potential drivers of disease trends, such as the relationship between environmental conditions and viral spread, or correlations between human, animal, and wastewater indicators, all while maintaining temporal and spatial context within the dashboard.

Step 5: Multi-Source Deep Exploration. The fifth panel presents detailed multi-series line charts for the selected counties, enabling users to compare temporal trends across two regions simultaneously (Figure 5.2). Each county has its own pair of line charts displaying multiple variables, such as COVID-19 cases, hospitalizations, or environmental and animal health indicators.

Hovering over any point in a line chart triggers a vertical guideline that spans all line charts across the other panels. This synchronized marker allows users to immediately see corresponding values for other variables, counties, or hierarchical temporal levels. The choropleth map in Panel a also updates to highlight the spatial distribution corresponding to the hovered time point, maintaining linkage between temporal and spatial views.

The dual-chart setup for each county allows direct comparison of trends between two regions, revealing similarities or divergences in disease dynamics, environmental factors, or other One Health datasets. Users can also interact with the charts by brushing to select a specific time range, dynamically updating all linked charts, and enabling detailed focus and context analysis.

This panel, in combination with Panels 2 and 4, provides a comprehensive interface for exploring temporal, spatial, and multi-variable data. It supports nuanced comparisons between counties, revealing patterns that might indicate shared drivers of disease progression or region-specific anomalies, while maintaining consistent color coding and synchronized interactivity across the dashboard.

5.1.4 System and Analysis Flow

The procedure of our system has several steps (see Figure 5.3).

Data Cleaning: We start our process by cleaning each dataset, i.e., handling missing data. Then, that dataset is split into training (70% of the original data) and test 30% of the original data) subsets.

Forecasting: Different methods have been applied to perform data prediction, such as LSTM (Hochreiter and Schmidhuber 1997) and STL (RB 1990), as well as STL applied to preprocessed data.

Long Short-Term Memory (LSTM) is an advanced machine learning algorithm that demonstrates excellent prediction performance, particularly for time series analysis (Song et al. 2020). LSTM is based on the assumption of a sequential connection between data points, making it suitable for time series data where earlier values are anticipated to impact future values. A particular recurrent neural network (RNN) is designed to avoid long-term dependency problems by remembering historical information.

Seasonal-Trend Decomposition (STL) STL is a universal and robust algorithm for time series decomposition, with high tolerance to outliers and applicability to various types of time series data. STL can decompose data into seasonality, trend, and residual components. STL assumes that there is a seasonal component in the data, which repeats at regular intervals (e.g., weekly in our case).

Data Transformation + STL to the processed data, such as Log, SQRT, Moving Average, Exponential Smoothing, and various scalers, including standard, Min-Max, and Robust scaler.

Normalization: The results from applying the different methods have been normalized to a range of 0 to 1 to facilitate fair comparison.

5.1.5 Performance Evaluation

The performance of the proposed models has been evaluated using various approaches, including the mean absolute error (MAE), the root-mean-square error (RMSE), the Pearson correlation coefficient, and the confidence interval (CI). Both MAE and RMSE are used to evaluate the performance of a model in absolute terms (Willmott and Matsuura 2005). The correlation coefficient indicates the degree to which changes in the value of one variable are associated with changes in the value of the other. A sign preceding the coefficient of correlation (+ or -) indicates the direction of correlation (Udovičić et al. 2007).

Confidence Interval (CI) is the range within which the true value is expected to fall with a certain level of confidence (in this case, we are using a confidence level of 95%) (Hazra 2017).

5.2 Case Study

We are utilizing various datasets and visual analytics techniques to better understand the COVID-19 outbreak and other factors in the State of Oklahoma at the county level. Our main focus is on implementing proactive surveillance and timely public health interventions. The visualization models we are developing aim to facilitate a clearer understanding of the outbreak and its associated factors. Therefore, we are considering One Health data, i.e., human, animal, and weather data that include geospatial data along with Google Trends data

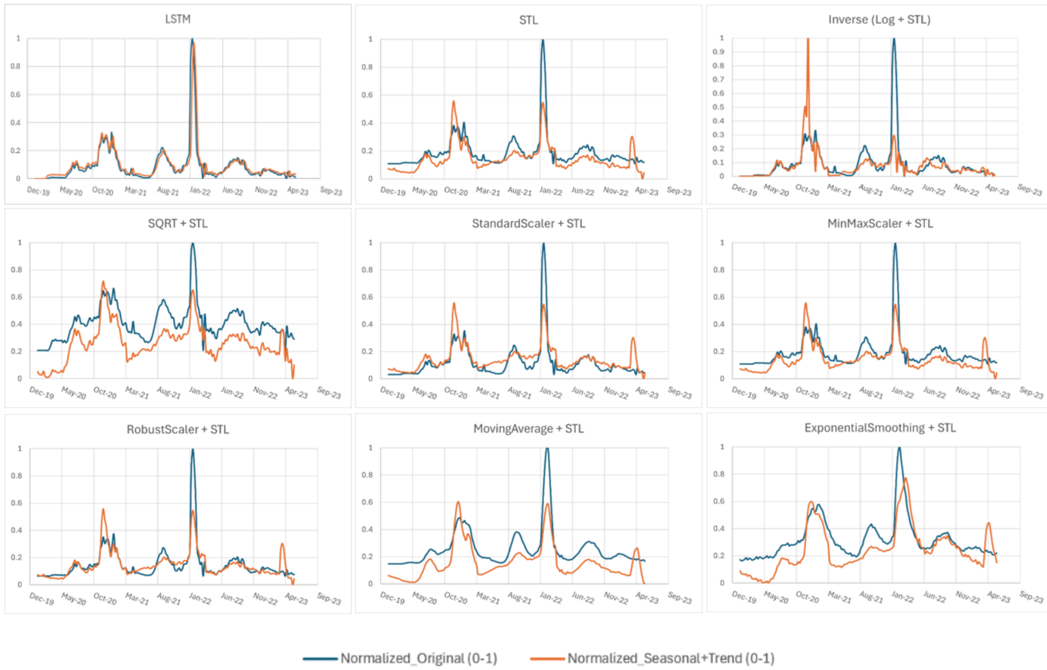


Figure 5.11: Prediction results of applying different methods to the COVID-19 cases in Oklahoma County. The blue line is the original dataset that has been normalized to the scale [0,1], while the orange line is the summation of the Seasonal and Trend decomposition of STL.

on the topic of "COVID-19 Symptoms Search Trends" for the years 2020-2022 in Oklahoma at the county level as a case study to explore the relationship between COVID-19 cases/deaths and other datasets.

Covid-19 deaths and cases datasets from CDC: This aggregate dataset was structured to include cumulative COVID-19 cases and deaths reported to the CDC from states and territories over time (Centers for Disease Control and Prevention, COVID-19 Response 2023).

Wastewater Human Pathogens: This dataset includes various human pathogens, such as "SARS-CoV-2" and "RSV," collected from sampling locations in Oklahoma.

Animal Ailments: This dataset contains animal ailments reported by Banfield Pet Hospital for various locations in Oklahoma. Where the ailments are the diagnoses assigned by a veterinarian during the inpatient visit.

Weather Data (gridMET): "gridMET" is a comprehensive dataset that provides daily, high-resolution gridded information on various surface meteorological variables such as maximum and minimum temperature, wind velocity, and humidity (Climatology Lab 2023).

COVID-19 Symptoms Search Trends from Google: For more than 400 health symptoms, the COVID-19 Symptom Search Trends Dataset contains aggregated, anonymized trends from Google searches. A time series for each region is provided in the dataset to show the search volume per symptom (PAIR-code 2020).

Our visualization models are designed to facilitate a deeper understanding of the outbreak by integrating these datasets. The system provides interactive visualizations that allow users to explore:

Temporal and Spatial Trends: Users can examine how COVID-19 cases and deaths correlate with changes in environmental factors, pathogen levels, and symptom search trends across different counties (Figure 5.10).

Table 5.1: Performance Comparison of Prediction Methods

Method	MAE	RMSE	Correlation Coefficient	Confidence Interval
Log + STL	0.127	0.148	**0.916**	0.016
SQRT + STL	0.152	0.162	0.882	0.020
MovingAverage + STL	0.106	0.121	0.864	0.188
LSTM	**0.033**	0.080	0.848	0.019
ExponentialSmoothing + STL	0.099	0.119	0.829	0.025
STL	0.058	0.082	0.816	**0.014**
StandardScaler + STL	0.047	0.081	0.816	**0.014**
MinMaxScaler + STL	0.058	0.082	0.816	**0.014**
RobustScaler + STL	0.037	**0.073**	0.816	**0.014**

Forecasting Analysis: We apply various forecasting methods, including Log, SQRT, Moving Average, Exponential Smoothing, and normalization techniques, followed by STL decomposition and LSTM models, to predict future COVID-19 cases and deaths (Figure 5.11). This helps in identifying potential future spikes or declines in cases.

Threat Detection: By analyzing trends and correlations among different datasets, the system can highlight areas with increasing infection rates or unusual patterns (Figure 5.10). For example, rising levels of SARS-CoV-2 in wastewater or a sudden increase in symptom search trends may signal emerging outbreaks.

We applied several forecasting methods to predict COVID-19 cases and deaths, utilizing different data transformation approaches, including logarithmic, square root, moving average, and exponential smoothing. Normalization techniques, such as standard, Min-Max, and Robust scalers, were also explored. Then, we applied the STL decomposition method to the processed data and combined the two parts of the STL method output (seasonal and trend) for the prediction results. We then compared the different methods with the LSTM method as shown in Figure 5.11.

We employ a rolling forecasting window of two weeks ahead to predict COVID-19 cases and deaths. At each time point, the models generate predictions for the subsequent two weeks using the most recent available data.

For performance evaluation of the applied methods, we used MAE, RMSE, Pearson correlation coefficient, and confidence interval. Based on the results provided in Table 5.1, we can see that:

- The correlation coefficient values range from 0.816 to 0.916, indicating a relatively strong positive linear relationship between the predicted and actual values across all methods, where "Log + STL" was the most effective method, followed by "SQRT + STL."

- Lower MAE values indicate better predictive performance. The methods "LSTM" and "RobustScaler + STL" have the lowest MAE values, indicating that they provide the most accurate predictions on average.

- Similar to MAE, lower RMSE values indicate better predictive performance. Again, the "LSTM" method has the lowest RMSE value, followed closely by "RobustScaler + STL".

- A smaller confidence interval indicates more precise predictions, making the prediction more reliable and the true value likely closer to the predicted value. STL and the applying scalers have the smallest confidence intervals, making them the best choices for precise predictions.

5.3 Discussion

Throughout our design and development process, we have received continuous feedback from stakeholders, including epidemiologists, public health experts, and veterinarians. We quote the feedback from an epidemiologist: "[...] the One Health approach is simultaneously a challenge and opportunity to access, analyze, and interpret data from these different sources and structures meaningfully. Presenting them in the proposed dashboard is meant to facilitate hypothesis generation. In the present example, the common feature is the x-axis, which represents time. By superimposing these line graphs, we can examine patterns and question if there are correlations." A public health expert stated: "[...] the PIPP dashboard provides unique insight into the transmission specifics of an infectious disease, including possible outbreaks and seasonal patterns. Combined, this information is a crucial public health tool because it enables the generation of more real-time signals with broader population coverage compared to traditional surveillance methods. Ultimately, this facilitates better preparedness and better control and prevention measures targeting."

This feedback validates the system's usability and relevance to real-world public-health practice, supporting the claims related to user-friendly visualization, collaboration with officials, and actionable analytic insights.

5.4 Summary

As of today, the data used in the PIPP dashboard requires at least weekly data points at the county level after preprocessing. We envision a data ingestion assistant using heuristics and LLM capabilities to guide the user. Another design goal is to extend correlation and prediction techniques. One promising approach is multi-variate time-series prediction combined with feature importance techniques, where the user could select a target variable, and the system suggests which datasets and columns are useful for prediction. Another step is to enhance the system to detect anomalies in multivariate time series and regions, and display alerts to the user. A feedback loop will be created, allowing users to label false positives, thereby improving detection capabilities over time.

We developed a dashboard with interactive visualization techniques to explore the relationship between health and environmental-related data. Additionally, we have used some forecasting methods to predict COVID-19 deaths and cases, and we have used evaluation metrics such as MAE, RMSE, correlation coefficient, and confidence interval to evaluate the performance of the applied methods. The quantitative evaluation (MAE, RMSE, correlation, CI) confirms the system's predictive accuracy and supports the novelty aspects related to forecasting and timely threat identification. The stakeholder feedback further validates the system's usability, interpretability, and practical value for decision-making. Collectively, these results provide comprehensive validation for the PIPP dashboard's contributions, including data integration, visualization design, predictive modeling, and collaborative development. The dashboard is well-suited for exploring data sources and generating hypotheses for predictors of various diseases.

Chapter 6

Modeling for Syndromic Surveillance

In this chapter [1], we present *OH-MEMA*, which is an integrated One Health mixed-effects modeling approach for syndromic surveillance. OH-MEMA is a visual analytics framework designed to assist domain experts, such as public health researchers, epidemiologists, and disease surveillance analysts, our primary target users. This framework helps them model and understand disease dynamics using integrated One Health data. The system combines data from human clinical records, environmental monitoring, and wastewater surveillance into a single, unified platform. Central to this system is a mixed-effects statistical modeling engine that enables predictive modeling while accounting for variability across different counties, time periods, and population groups.

Our framework is designed to assist domain experts in iteratively building and evaluating models, exploring trends, and making evidence-based decisions. Unlike many traditional tools, the system offers interactive features that allow users to upload datasets, select predictors, visualize trends, compare regions, and configure modeling parameters without coding. Model performance is automatically evaluated using RMSE, MAE, and correlation metrics, and results are presented through an interactive provenance tree that supports comparison across multiple modeling runs. While we are currently collaborating with public health

[1] This chapter is based on the paper submitted to PLOS One, Oct 2025: A. Basheer, P. M. Khiabani, W. Jentner, A. Wendelboe,J. Vogel, K. Kuhn, M. C. Wimberly, D. Hougen, and D. Ebert, “OH-MEMA: An Integrated One Health Mixed-Effects Modeling Approach for Syndromic Surveillance," PLOS One 2025.

partners in Oklahoma to assess and refine system usage, this work represents a step toward developing flexible, interpretable, and integrated decision-support tools for proactive disease surveillance.

A key innovation of our system is the visual analytics provenance tree, which captures the evolution of modeling configurations and associated performance metrics over time. This tree allows users to track and compare runs based on selected predictors (e.g., wastewater SARS-CoV-2 concentrations, NET), outcome variables (e.g., hospitalizations), and random effects (e.g., county, month, age group). Each node in the tree encodes evaluation metrics such as RMSE, MAE, and correlation through color and size, enabling intuitive visual comparison. This design promotes transparency, supports iterative refinement, and encourages evidence-based model selection.

Beyond this visualization component, our system introduces a set of interactive capabilities that enable rapid model development and exploration. Users can select outcome variables and input features through a guided interface, configure random effects using drop-down menus, and initiate model training with a single click. Model training is handled server-side and optimized for fast computation, returning results within seconds for datasets at a weekly level. Once models are built, users can drill down into specific configurations, view comparative performance, and assess how different modeling choices impact predictive accuracy. These capabilities lower the barrier to entry for domain experts without statistical training, while still supporting advanced analyses for public health researchers and epidemiologists.

To ensure the reliability and usability of the proposed framework, this chapter employed a two-tiered validation approach encompassing both quantitative and qualitative assessments. Quantitatively, the mixed-effects models were validated using established statistical performance metrics, namely Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and correlation, calculated across multiple counties and outcome variables. These metrics provided objective measures of predictive accuracy and robustness under varying data and model configurations. Qualitatively, the system was evaluated through expert feedback sessions involving disease surveillance analysts and epidemiologists, focusing on interpretability, workflow integration, and cognitive workload. Using the NASA-TLX instrument and open-ended interviews, participants assessed the usability and mental demand of the tool's interactive modeling and provenance-tracking features. Together, these validation methods confirm that the system not only produces reliable predictive outputs but also effectively supports user-centered analytical reasoning and model comparison within public health decision-making workflows.

By combining explainable modeling techniques with interactive controls, performance feedback, and provenance tracking, our framework supports a more agile and user-centered modeling workflow. It empowers stakeholders to quickly explore multiple hypotheses, refine models based on real-time feedback, and uncover insights that would be difficult to obtain using traditional static modeling tools. In doing so, our system advances the state of the art in visual analytics for syndromic surveillance and demonstrates the practical value of integrating One Health data for timely public health response.

6.1 System Methodology

Our goal was to enhance pandemic prediction at the county level through a transparent modeling process. This process integrated data-driven feature selection with a non-linear mixed-effects model, taking into account human clinical outcomes, environmental conditions, and wastewater-based epidemiology.

As shown in Figure 6.2, we started by preparing the dataset for modeling the target variable and its predictors. This involved applying Min-Max normalization, which not only standardized the data range but also supported privacy protection through data perturbation (Jain and Bhandare 2011). Following normalization, we used Seasonal-Trend Decomposition using LOESS (STL) (Hafen et al. 2009). The main idea was to apply STL decomposition

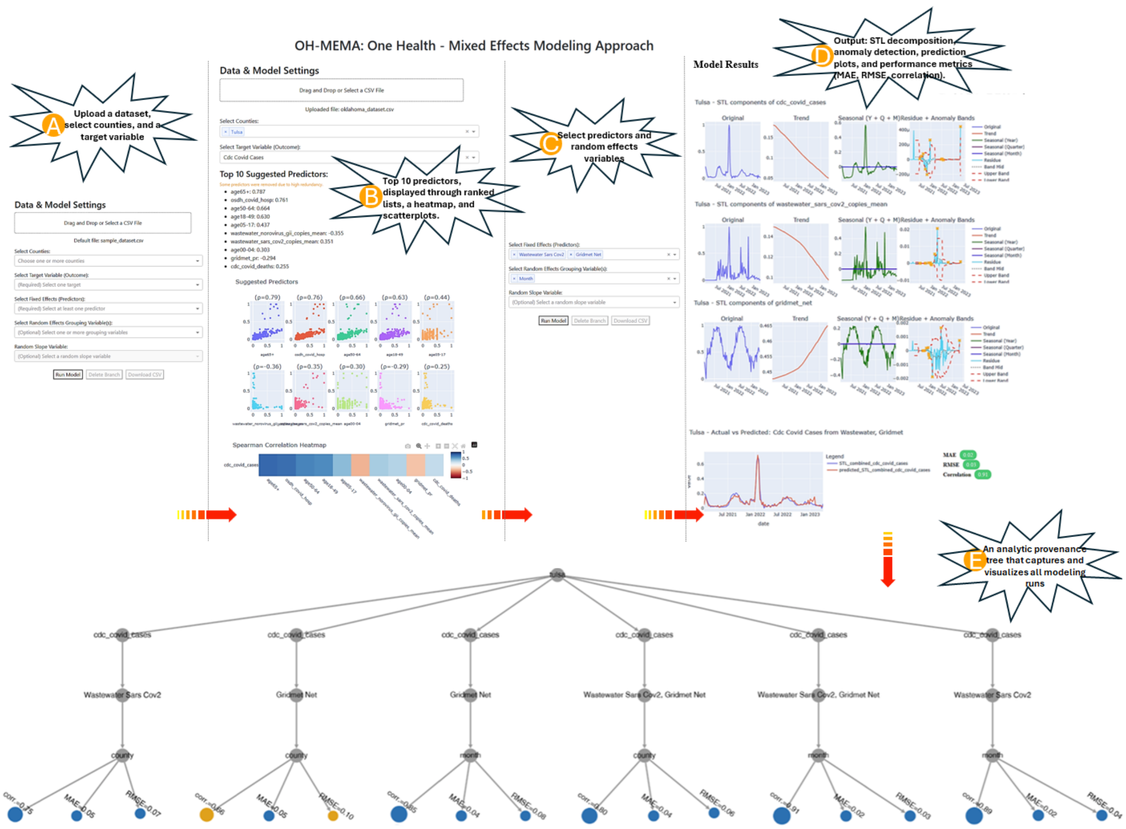


Figure 6.1: Visual Analytics Workflow OH-MEMA. (**A**) After uploading a dataset, users select counties and a target variable. (**B**) The system calculates Spearman correlations between the target and all other variables, displaying the top 10 predictors via ranked lists, a heatmap, and scatter plots. (**C**) Users configure fixed and random effects, optionally specify interaction (slope) variables, and run the model. (**D**) Output includes STL decomposition, anomaly detection, prediction plots, and model performance metrics (MAE, RMSE, correlation). (**E**) An analytic provenance tree that captures and visualizes all modeling runs, allowing users to explore past configurations and performance, and reload any previous model for comparison. (link: `https://mema.disc.ourcloud.ou.edu/`)

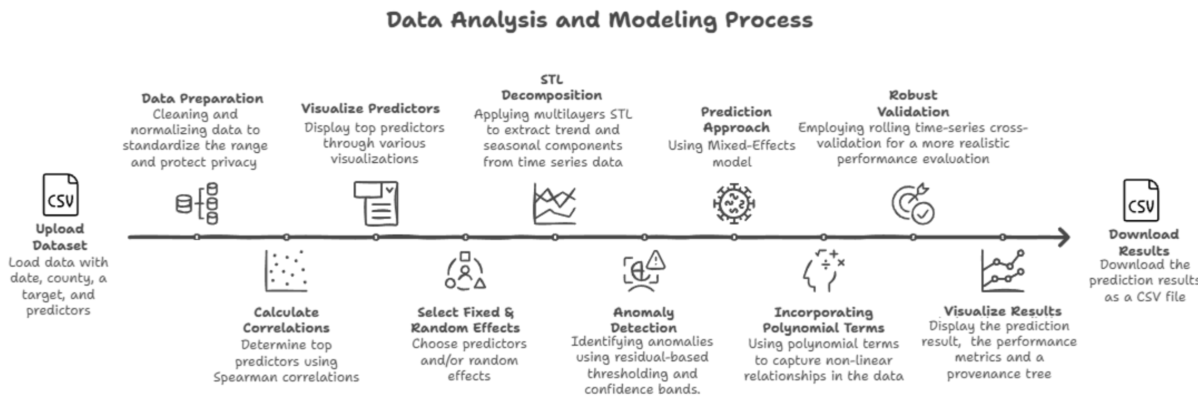


Figure 6.2: Data Analysis and Modeling Process.

to the selected target and predictor variables for each county individually. This approach enabled us to extract smoothed trend and seasonal components, preserving spatial heterogeneity while capturing consistent temporal features.

To enhance the interpretability and modeling capabilities of epidemiological predictors in our visual analytics system, we employed a layered STL to decompose and denoise the county-level time series (Hafen et al. 2009). The process began with a log transformation to stabilize variance and reduce the impact of extreme values, which was particularly relevant for count-based health metrics (Basheer et al. 2025). We then decomposed the target variable and each predictor using STL and compiled a list of cleaned, county-specific decomposition results. This list included the trend, multiple seasonal layers (yearly, quarterly, monthly), and residual components.

To support anomaly detection and interpretation in the visual interface, we employed a residual-based thresholding mechanism grounded in rolling statistics. We computed a centered moving average and standard deviation over a configurable window applied to the final residuals from the multi-layer STL decomposition. Observations that exceeded ± 2 standard deviations were flagged as anomalies, supporting the detection of potential outbreaks or signal shifts. Corresponding upper and lower confidence bands were retained for visual encoding. This banding resembled Bollinger Bands (Bollinger 2002), a widely used technique for time-series anomaly detection, and aligned with established statistical process control practices in public health surveillance (Fricker Jr et al. 2008). The anomaly flags and confidence bands were retained as first-class data attributes and visually encoded to support the identification of atypical patterns and emerging signals, as illustrated in Figure 6.3.

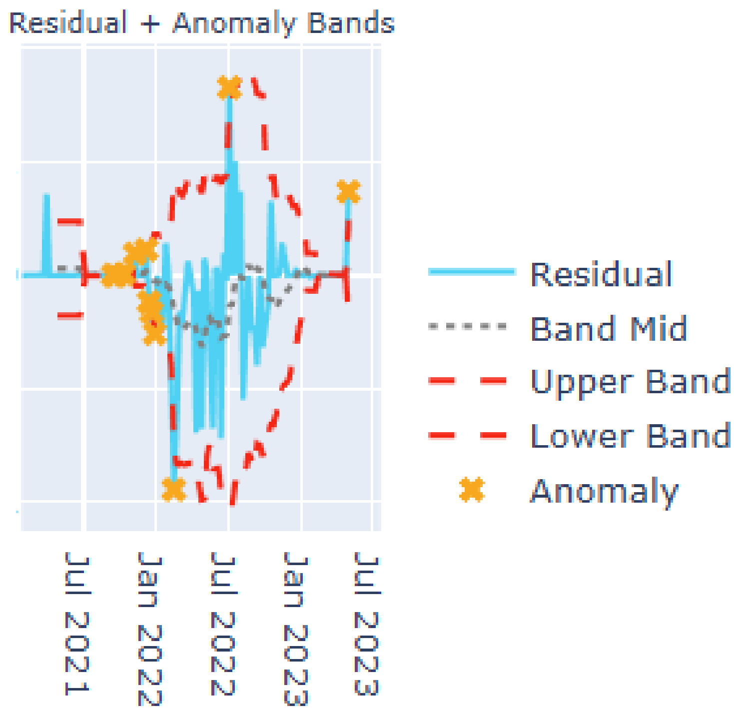


Figure 6.3: Anomaly detection from STL decomposition showing residual, band mid, upper band, lower band, and anomalies.

After that, we updated the model preparation pipeline to conditionally include random effects and random slopes in the final modeling DataFrame. The "month" was excluded from the "date" column to be used as a random effect. This ensured compatibility with models that expect these variables and resolved errors related to categorical variable aggregation and unintended duplication.

The prediction approach builds on the approach of Daza-Torres et al. (Daza-Torres et al. 2024), who also used a mixed-effects model for prediction. However, we extend their methodology in several significant ways. While their research focuses exclusively on COVID-19 hospitalization as the disease outcome, our model is designed to predict a variety of pandemic outcomes, including cases, deaths, and hospitalizations. We have broadened the set of predictors from the original focus on COVID-19-related data, such as wastewater, test positivity rate (TPR), and reported cases, to include additional data sources, including environmental factors like the Net Effective Temperature (NET) index, as well as other relevant predictors. Furthermore, whereas their analysis models the interaction between wastewater and the COVID-19 wave, our approach accommodates interactions between any predictor and other temporal factors.

We also incorporate polynomial terms to capture non-linearity (Vitanov and Dimitrova 2021). For each county, continuous predictors (e.g., weather indicators) are log-transformed and expanded with cubic polynomial terms to capture non-linear relationships. Both continuous and binary anomaly features (e.g., weather anomalies) are shifted across a defined temporal window of lags (from -4 to 0 weeks). This lag structure enables systematic testing of causal lead-lag relationships between predictors and the target variable (Stübinger and Adler 2020). To determine the most relevant lag, we fit a series of mixed-effects models for each predictor and lag combination, using Spearman correlation (De Winter et al. 2016) between fitted and observed values as a selection criterion. The final model includes only those lagged terms that yield the strongest association with the outcome.

Additionally, we have expanded the structure of the random effects from a single grouping by county to a more comprehensive framework that considers variability by age group, sex group ratios, and temporal effects across different months.

Importantly, our approach is interactive, allowing users to choose their outcome, predictors, and grouping structures through a visual tool that highlights key relationships and provides real-time results, as shown in Figure 6.1.

Finally, unlike the original study, which reports results from a single, static model fit, our method utilizes rolling time-series cross-validation for each county. This yields a more robust and realistic evaluation of predictive performance.

6.1.1 Model Formulation

The mixed-effects model used in this study accounts for both fixed and random effects while incorporating lagged and nonlinear (polynomial) predictors. The general model takes the form:

$$y_{it} = \mu + \sum_{k} \beta_k x_{kit}^{(l)} + \sum_{j} \gamma_j z_{jit}^{(l)} + b_{g[i]} + u_{g[i]} x_{rit}^{(l)} + \epsilon_{it} \tag{6.1}$$

where y_{it} denotes the log-transformed target outcome (e.g., COVID-19 hospitalizations) for county i at time t, and μ is the global intercept. The term $x_{kit}^{(l)}$ represents lagged and/or polynomial-transformed fixed-effect predictors (e.g., log-lagged weather or wastewater concentrations), with associated coefficients β_k. Additional variables $z_{jit}^{(l)}$ capture interaction terms or anomaly indicators (such as seasonal anomalies), with corresponding coefficients γ_j.

The model also includes group-specific random effects. The random intercept $b_{g[i]}$ accounts for variation across groups such as month, county, age group, or sex group. Additionally, the random slope $u_{g[i]}$ allows one selected predictor $x_{rit}^{(l)}$ (e.g., weather or wastewater) to vary across the same groups. Finally, ϵ_{it} represents the residual error. Incorporating lagged predictors enables the model to account for known delays between environmental signals and clinical outcomes, with each lag evaluated independently.

6.1.2 Random Effects Handling

To account for variability across spatial and demographic dimensions, we incorporated random effects into our mixed-effects modeling framework. Random effects enable the model to capture unobserved heterogeneity that arises from grouping structures, such as geography (e.g., counties), time (e.g., months extracted from the date), or population characteristics (e.g., age or sex groups). The month was selected as a random effect to capture unobserved temporal heterogeneity that may influence outcomes across different time periods, such as seasonal fluctuations in COVID-19 cases or environmental conditions. By modeling month as a random effect, the model accounts for systematic variation across months without estimating a separate coefficient for each one, which helps prevent overfitting and preserves generalizability to future time periods (Briz-Redón 2021). This approach aligns with previous research that emphasizes the importance of hierarchical structure in public health data modeling (Sokhansanj and Rosen 2022).

6.1.3 Prediction Model Evaluation

To ensure robust performance estimation, we implement rolling time-series cross-validation. For each county:

- The time series for each county is split chronologically, with the earliest 80% of observations used for training.
- The most recent 20% of observations are reserved for testing, simulating a realistic forecasting scenario.
- The model is retrained at each step as the training set expands over time (forward-chaining cross-validation).
- Lagged versions of key predictors (e.g., NET, wastewater) are tested individually. The best lag is selected based on test set correlation and mean absolute error (MAE).
- Model performance is evaluated using standard metrics: mean absolute error (MAE), root mean squared error (RMSE), Spearman correlation, and R^2.

6.2 Visual Analytics System Description

Our methodology is based on three main pillars: selecting features through visual analytics, user-configurable mixed-effects modeling, and providing explainable evaluations of the results, all delivered through an online interactive tool as shown in Figure 6.1. The full system is deployed and freely accessible via a web interface at `https://mema.disc.ourcloud.ou.edu/`. Users can upload their own datasets and view the results quickly and easily.

6.2.1 Feature Selection

Once the user uploads a CSV file containing relevant data, such as county, date, and at least one outcome variable, the system will identify valid predictors and outcomes. When the user selects a target variable (for example, COVID-19 hospitalizations), as shown in Figure (6.1−A), the system calculates Spearman correlations between this target and all other columns in the dataset. To improve model interpretability and reduce redundancy, the system includes a redundancy detection step: predictors that are highly correlated with each other (above a configurable threshold, e.g., 0.95) are automatically filtered, leaving only one representative feature from each correlated group. Users are notified when predictors are removed due to redundancy. The top 10 non-redundant predictors are then presented through a ranked list, interactive scatterplots, and a heatmap of correlation strengths, as shown in Figure (6.1−B). This selection process, summarized in Algorithm 1, guided user involvement by combining statistical associations with visual pattern recognition.

The user can choose one or more predictors to include as fixed effects in the prediction model. Additionally, they can select variables such as county, month, age, or sex group to serve as random effects in the model. The slope variable, which is one of the chosen predictors, can also be selected to enable the user to specify the interaction between the random effects and that predictor. Once the selections are made, the user should click the "Run Model" button to perform the calculations, as shown in Figure (6.1−C).

6.2.2 User Interface

Our system's user interface was built to support transparency, user control, and step-by-step interaction for non-expert users in public health and epidemiology. It consists of four main panels, as shown in Figure 6.1:

- **Panel A (Data Selection):** Users can upload a CSV file or use the sample dataset that is included in the system. The interface automatically validates column names and updates the drop-down menus accordingly. Figure 6.1−A shows the upload panel, where the user selects the target variable and counties of interest.

- **Panel B (Feature Correlation):** The system computes Spearman correlations between the selected target and all predictors. The predictors that are highly correlated with each other (above a configurable threshold, e.g., 0.95) are automatically filtered, leaving only one representative feature from each correlated group. Users are notified when predictors are removed due to redundancy. The top 10 predictors are displayed in a ranked list, interactive scatter plots, and a correlation heatmap as shown in Figure (6.1−B).

- **Panel C (Model Configuration):** Users configure fixed and random effects. Optional controls allow the selection of interaction terms and slope variables. Tooltips explain the statistical meaning of each option, helping non-expert users understand mixed-effects modeling decisions as shown in Figure (6.1−C).

- **Panel D (Results & Interpretation):** Model predictions, STL decomposition results, and detected anomalies are shown with overlaid line charts. Each county-level STL decomposition is rendered as an interactive, multi-panel visualization (Figure 6.4). Users can hover over lines to reveal exact values and timestamps, toggle components (e.g., trend, yearly seasonality, anomalies) through the legend, and zoom into specific temporal windows. Residual anomalies are emphasized using color-coded markers, allowing users to identify outlier periods and visually relate them to potential external events. The interface automatically refreshes this view upon changes to the selected county or variable, creating an immediate feedback loop between model configuration and data interpretation. This interactivity supports transparent reasoning and exploratory sensemaking by enabling users to isolate and interpret different temporal dynamics in the data. Also, each STL decomposition figure supports direct manipulation through Plotly's interactive toolbar. Users can zoom into specific temporal windows by dragging across the plot or clicking the zoom icons; double-clicking resets the view to full. Because all panels share the same time axis, zooming or panning in one panel simultaneously updates all others, preserving temporal context between trend, seasonal, and residual components. Additionally, users can download any figure in high-resolution format via the built-in "Download as PNG" control, allowing seamless export of insights for reports and reproducible documentation. Evaluation metrics (MAE, RMSE, and correlation) are displayed, and results can be downloaded as a CSV file as shown in Figure (6.1−D).

- **Panel E (Version History and Model Traceability):** Every model run is automatically stored with its configuration, selected county, target, predictors, random effects, and its resulting performance metrics as shown in Figure (6.1−E). A visual tree interface allows users to explore and compare past runs, as shown in Figure 6.5. Each node displays a modeling decision (e.g., selected predictors or random effects), while performance metrics (correlation, MAE, and RMSE) are shown as color-coded and size-coded leaf nodes. Users can hover over a node to visually highlight it, enabling quick identification within dense trees. Users can click any node to reload a previous configuration, enabling rapid iteration, insight tracking, and reproducible modeling workflows. Users can also delete the unwanted branches by clicking on the "Delete Branch" button and then selecting the node to delete its dependencies. This visual traceability supports iterative exploration, reproducibility, and transparent reasoning. Altogether, these interactions allow users to explore, compare, and manage model configurations in a transparent, white-box environment.

The interface is implemented using Dash and Plotly for interactive visualization, enabling dynamic updates as users modify their selections. Tooltips, hover info, and dropdown menus provide further guidance during exploration. This interface acts as a white-box system,

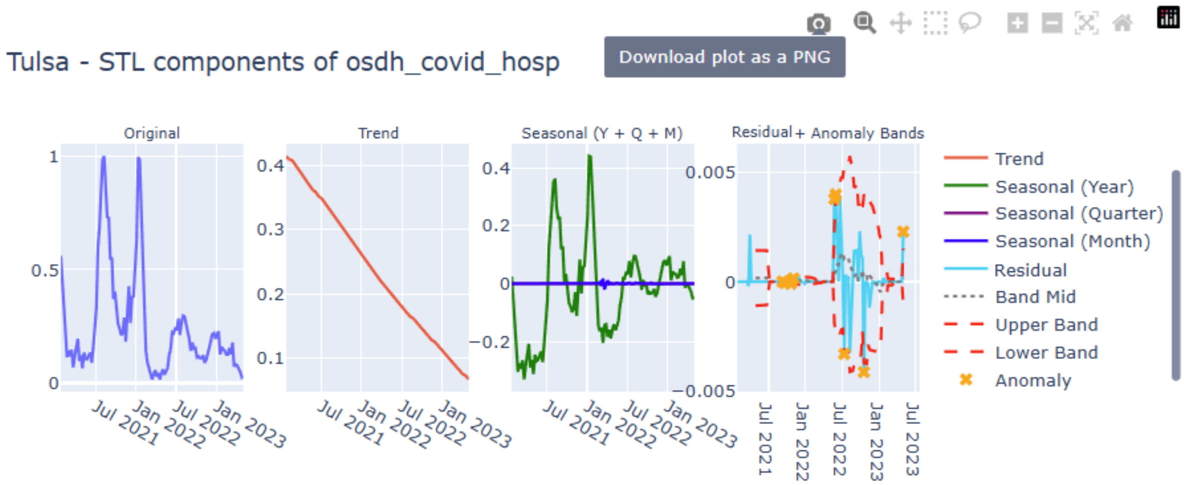


Figure 6.4: STL decomposition of county-level predictions showing trend, seasonal components, and residual with anomalies.

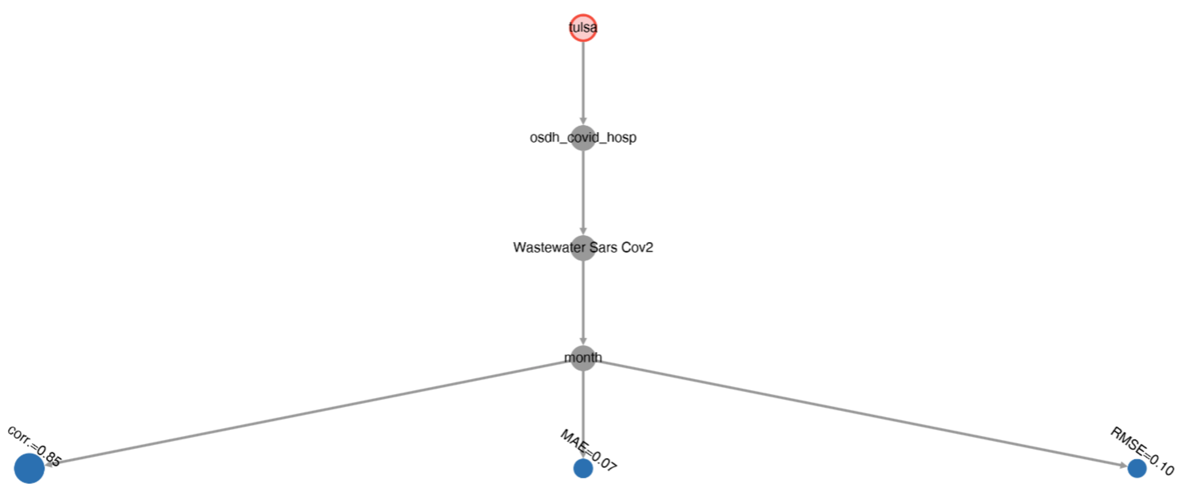


Figure 6.5: Visual tree of past model runs showing decisions, performance metrics, and interactive node operations.

promoting transparency and explainability in model building.

Algorithm 1: Feature Selection and Modeling Workflow

Input: CSV file with weekly county-level data
Output: The prediction results from the mixed-effects model

Step 1: Upload the dataset with columns such as `date`, `county`, target, and potential predictors
Step 2: Select one or more counties to subset the dataset
Step 3: Choose a target variable
- Spearman correlations will be calculated between the target and all other columns to find the top 10 related predictors
- Visualize the top 10 predictors using:
  - Ranked list of correlation coefficients
  - Correlation heatmap
  - Interactive scatterplots

Step 4: Select one or more predictors as fixed effects
Step 5 (Optional): Select one or more random effects (e.g., county, month, age group)
Step 6 (Optional): Choose one fixed-effect predictor to model random slopes
Step 7: Run the model and return:
- STL decomposition of predictors and target
- Anomaly detection
- Model prediction results with performance metrics (MAE, RMSE, correlation)
- Append current model configuration and metrics to an analytics provenance tree for visualization

Step 8 (Optional): Delete a branch from the analytics provenance tree
- The delete button is enabled only after at least one model run has produced a tree
- Selecting a branch allows the user to remove that branch and all its associated model configurations from the tree
- If the selected branch is the only branch in the tree, the system prompts the user for confirmation before deletion

6.2.3 Analytics Provenance for Model Exploration and Comparison

To support iterative model refinement and comparative analysis, we integrated an analytics provenance mechanism (North et al. 2011) into the dashboard, as shown in Figure (6.1−E). Each time a model is run, the system logs the selected configuration—including county, target variable, fixed effect predictors, and random effect groups—along with the resulting evaluation metrics (correlation, MAE, and RMSE). These versions are visualized as a dynamic, interactive tree using Dash Cytoscape in Python, where each branch represents a modeling decision (e.g., target variable, predictors, random effects), and leaf nodes display the corresponding performance metrics.

The use of provenance trees is grounded in extensive research demonstrating their value for tracking analytic workflows and supporting sensemaking in complex data analysis tasks (Huynh et al. 2018; Oliveira et al. 2018). Provenance visualization captures the history and rationale behind analytic decisions, enabling users to retrace steps, compare alternative hypotheses, and understand how changes in model parameters impact outcomes (Madanagopal et al. 2019; Ragan et al. 2015). This explicit representation of analytic provenance is particularly critical in epidemiological modeling, where transparency and reproducibility are necessary for building trust with domain experts and decision-makers (Chen and Plale 2015).

To improve interpretability, leaf nodes were colored using a color-blind–safe palette (Okabe and Ito 2008) to provide quick insight into model quality: blue indicates strong correlation ($\rho > 0.7$) or low error (MAE/RMSE < 0.1), while orange highlights weaker performance. This feature enhances interpretability, supports inclusive design, and empowers users to make informed, feedback-driven modeling decisions. This feature empowers users to make more informed modeling decisions and fosters an exploratory, feedback-driven analytical workflow. Additionally, node size encodes the magnitude of performance; larger nodes correspond to stronger correlations or lower errors, helping users identify optimal models at a glance.

The system also supports interactive branch deletion within the provenance tree, allowing users to dynamically refine or simplify model structures. The **Delete Branch** button remains disabled until at least one model run has been completed and a tree structure exists, as shown in Figure 6.6, ensuring that deletion actions are only possible when meaningful model configurations are present. Once enabled, users can click the button, and the interface prompts them to select a node to remove. After that, users should select the branch that they want to delete by selecting a node from the tree, and then the system will automatically delete the selected node and all of its connected descendants. This design enables flexible model editing without requiring users to rebuild the entire configuration from scratch. If the selected node represents the only branch in the tree (for example, the root node associated with a county), the system displays a confirmation dialog to prevent accidental deletion of the entire structure. This interactive control enhances usability by supporting incremental exploration, model refinement, and correction of configuration errors in a transparent and intuitive way.

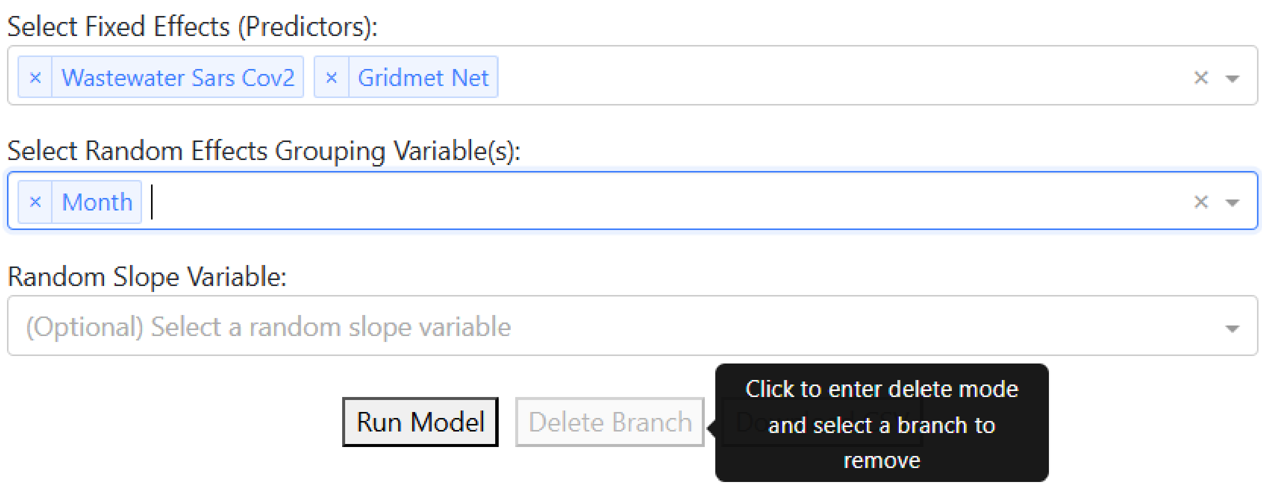


Figure 6.6: Interactive branch deletion in the provenance tree. The delete button remains disabled until at least one model run has been completed and a tree has been created. When enabled, users can delete a selected branch, with a confirmation prompt appearing if it is the only branch in the tree. Tooltips provide guidance on the button's functionality.

The tree structure supports an overview and detailed analytical workflow, enabling users to quickly evaluate model performance across multiple experiments and identify effective configurations, such as those with low MAE or high correlation coefficients. Users can interactively reload previous runs by clicking on any node, facilitating iterative exploration and model refinement. This interactivity aligns with best practices in provenance visualization research, which emphasizes user-driven exploration of analytic histories to support trade-off analysis and informed decision-making (Menin et al. 2021; Chakhchoukh et al. 2022).

6.3 Case Studies: County-Level Surveillance in the US (2020–2024)

To evaluate the effectiveness of our system in supporting proactive public health interventions and outbreak understanding, we apply it to a real-world case study focused on the U.S. states such as Arizona, Arkansas, and Oklahoma. This case study integrates geospatial and temporal datasets across human and environmental health domains, aligning with the One Health paradigm. Our goal is to explore correlations and predictive relationships among heterogeneous data sources at the county level from 2020 to 2024.

- **Human Clinical Data from CDC and OSDH** We incorporate aggregated weekly-level COVID-19 clinical data, including cumulative confirmed cases, hospitalizations, and deaths. These records are sourced from both the U.S. Centers for Disease Control and Prevention (CDC), which collects standardized data across all states and territories (Centers for Disease Control and Prevention, COVID-19 Response 2023), and the Oklahoma State Department of Health (OSDH), which provides more localized reporting. These outcome variables serve as core targets for predictive modeling and anomaly detection.

- **Wastewater Surveillance of Human Pathogens** This dataset includes quantifications of human pathogens, such as SARS-CoV-2, Influenza, and Respiratory Syncytial Virus (RSV), from municipal wastewater samples collected across Oklahoma counties. Wastewater surveillance offers a non-invasive, early-warning signal of community-level pathogen presence, even when clinical testing is underreported or delayed (Oeschger et al. 2021).

- **Environmental and Meteorological Data (gridMET)** We include environmental data from the gridMET dataset, which provides high-resolution gridded daily surface meteorology for the contiguous United States. Variables such as minimum and maximum temperatures, wind speeds, vapor pressure deficits (VPD), and relative humidities are aggregated at the county level (Climatology Lab 2023). To quantify thermal comfort conditions at the county level, we compute the Net Effective Temperature (NET) index, an indicator that accounts for the combined effects of air temperature, relative humidity, and wind speed. The NET index (Zhang et al. 2023), is defined as:

$$\mathrm{NET} = 37 - \frac{37 - T}{0.68 - 0.0014 \cdot RH + \frac{1}{1.76 + 1.4 \cdot v^{0.75}}} - 0.29 \cdot T \cdot (1 - 0.01 \cdot RH) \tag{6.2}$$

where T is the daily mean near-surface air temperature (in °C), RH is the relative humidity (in %), and v is the wind speed (in m/s). NET provides a more holistic representation of thermal stress than temperature alone and has been widely used to evaluate spatial inequality in thermal comfort across regions (Zhang et al. 2023). In our case study, we calculate the weekly mean NET values from the gridMET meteorological data for each Oklahoma county, serving as a key environmental predictor of respiratory-related health outcomes.

For the following case studies, we selected the two most populous counties in each state and integrated environmental, clinical, and wastewater data sources into mixed-effects models to evaluate their predictive power for respiratory-related outcomes. In our case studies, we calculate the weekly mean NET values from the gridMET meteorological data for each county, serving as a key environmental predictor of respiratory-related health outcomes.

Each case study is supported by a table of the results and an analytics provenance tree. Each tree branch represents a unique modeling configuration, including the selected target variable, predictors, random effects, and corresponding model performance metrics. Leaf nodes display correlation (ρ), MAE, and RMSE. Node colors provide quick insight into model quality: green indicates strong correlation ($\rho > 0.7$) or low error (MAE/RMSE < 0.1), while orange highlights weaker performance. Additionally, node size encodes the magnitude of performance; larger nodes correspond to stronger correlations or lower errors, helping users identify optimal models at a glance.

6.3.1 Case Study I: Arizona (Flu and RSV Cases)

For Arizona, we modeled weekly county-level influenza and RSV cases from the Arizona Department of Health Services (ADHS) (of Health Services) using the Net Effective Temperature (NET) as the primary environmental predictor. Models were evaluated both with and without the month included as a random effect to capture temporal variability. When investigating across all Arizona counties, the models showed a consistent pattern: average correlations increased by approximately 10–15% when month was included as a random effect, while MAE and RMSE decreased accordingly. On average, RSV models achieved statewide correlations of approximately 82% with month effects (compared to 76% without), and flu models reached around 79% (versus 71%). For example, Table 6.1 summarizes model performance for Maricopa and Pima counties, while Fig. 6.7 provides a visual provenance view of modeling decisions. Results indicate that incorporating the month as a random effect improved predictive performance across both counties, particularly for RSV. For example, in Maricopa County, RSV predictions achieved a correlation of 89% with month effects compared to 84% without. These results highlight the value of accounting for seasonal variation in respiratory disease dynamics.

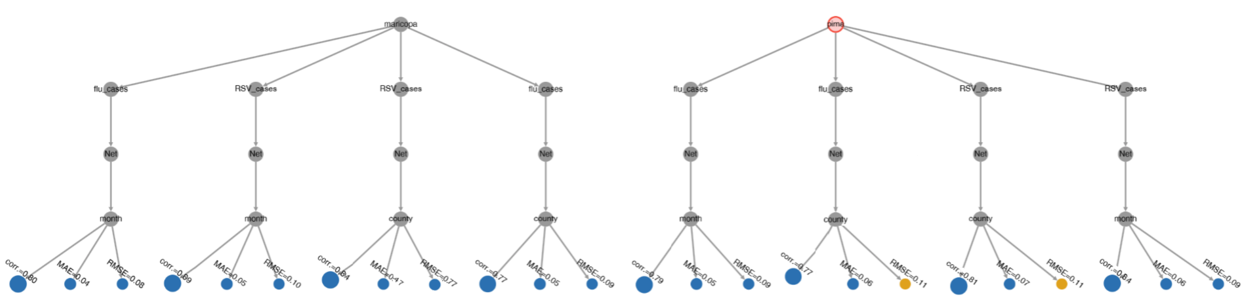


Figure 6.7: Visual summary of model analytics provenance for comparing two counties, the left tree is for Maricopa County, while the right tree is for Pima County, in the State of Arizona. This figure complements Table 6.1 by enabling users to visually trace and compare modeling decisions over time.

This demonstrates that while performance varies between individual counties, the overall

Table 6.1: Model performance for Flu and RSV cases prediction for the State of Arizona using NET with and without random effects (Month).

County	Target	Predictors	Random Effects	Correlation	MAE	RMSE
Maricopa	Flu Cases	NET	Month	80%	0.044	0.081
			County	77%	0.050	0.093
	RSV Cases	NET	Month	89%	0.051	0.052
			County	84%	0.047	0.077
Pima	Flu Cases	NET	Month	79%	0.059	0.092
			County	77%	0.062	0.110
	RSV Cases	NET	Month	84%	0.062	0.094
			County	81%	0.071	0.108
State Average	Flu Cases	NET	Month	79%	0.052	0.082
			County	71%	0.061	0.094
	RSV Cases	NET	Month	82%	0.056	0.092
			County	76%	0.076	0.113

statewide trends confirm the robustness of including temporal structure in the mixed-effects framework.

6.3.2 Case Study II: Arkansas (COVID-19 Cases and Deaths)

For Arkansas, weekly county-level COVID-19 cases and deaths from the U.S. Centers for Disease Control and Prevention (CDC) (Centers for Disease Control and Prevention, COVID-19 Response 2023) were modeled using the Net Effective Temperature (NET) as the primary environmental predictor. Models were evaluated both with and without the month included as a random effect to account for temporal variability.

When averaging across all Arkansas counties, we observed a consistent improvement when incorporating the month as a random effect. On average, statewide correlations for case predictions increased from about 53% without month effects to 80% with them, while death predictions rose from roughly 38% to 72%. Correspondingly, MAE and RMSE values decreased, underscoring the value of including temporal variation, as shown in Table 6.2. These statewide results confirm the patterns observed in Pulaski and Washington counties, where accounting for monthly variability substantially enhanced the predictive accuracy of NET-based models for both COVID-19 cases and deaths.

Table 6.2 presents the results for Pulaski and Washington counties, complemented by the provenance visualization in Fig. 6.8. In Pulaski County, the correlation for COVID-19 cases increased from 43% to 79%, while the death rate improved from 31% to 65%. A similar pattern was observed in Washington County, where case prediction improved from 54% to 78% and death prediction from 36% to 71%. Although death prediction showed weaker correlations compared to cases, reflecting noisier and less frequent reporting, the inclusion of month as a random effect substantially enhanced overall model accuracy.

Table 6.2: Model performance for COVID-19 prediction for the State of Arkansas using NET with and without random effects (Month).

County	**Target**	**Predictors**	**Random Effects**	**Correlation**	**MAE**	**RMSE**
Pulaski	Cases	NET	Month	79%	0.039	0.065
			County	62%	0.044	0.076
	Deaths	NET	Month	62%	0.09	0.12
			County	34%	0.113	0.147
Washington	Cases	NET	Month	78%	0.033	0.058
			County	54%	0.047	0.074
	Deaths	NET	Month	71%	0.079	0.107
			County	36%	0.113	0.136
State Average	Cases	NET	Month	80%	0.033	0.058
			County	53%	0.043	0.070
	Deaths	NET	Month	72%	0.040	0.108
			County	38%	0.110	0.134

6.3.3 Case Study III: Oklahoma (COVID-19 Hospitalizations)

As a case study, we apply our visual analytics framework to explore the correlation between multiple heterogeneous datasets and their predictive power for COVID-19 hospitalizations in Tulsa County, Oklahoma. The goal is to support proactive public health surveillance and intervention by modeling the impact of environmental and wastewater signals on human health

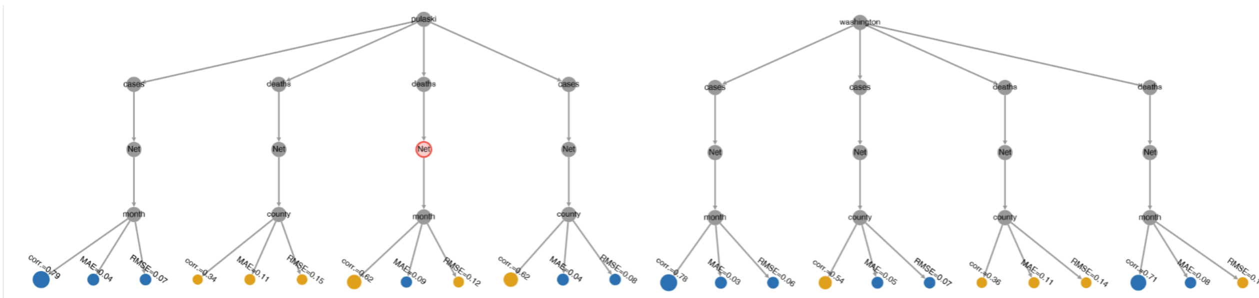


Figure 6.8: Visual summary of model analytics provenance for comparing two counties, the left tree is for Pulaski County, while the right tree is for Washington County, in the State of Arkansas. This figure complements Table 6.2 by enabling users to visually trace and compare modeling decisions over time.

outcomes. In our modeling approach, these datasets are treated as fixed effects—variables believed to have a consistent effect on the target outcome. The outcome of interest is the weekly number of COVID-19 hospitalizations at the county level. To capture hierarchical and contextual variation, we include random effects for county, month, age group, and sex group. These allow the model to account for unobserved heterogeneity across demographic strata and time.

We assessed three predictive models (M1–M3) using the results generated by our interactive visual analytics system, as shown in Table 6.3. This includes model evaluation metrics (e.g., MAE, RMSE) and correlation statistics, all computed on held-out validation data from the most recent 20% of the time series using rolling time-series cross-validation. These metrics reflect model performance on future observations not included during training. Complementing these quantitative results, Figure 6.9 provides a visual summary of model analytics provenance. Each tree branch represents a unique modeling configuration, including selected target variable (COVID-19 hospitalizations), predictors (wastewater or NET), random effects (county, month, sex group, or age group), and their associated performance metrics. Leaf nodes display correlation (ρ), MAE, and RMSE values, with node color and size encoding model quality to facilitate intuitive comparison of modeling decisions over time.

Our findings show that incorporating STL (Seasonal-Trend decomposition using Loess)

Table 6.3: Performance of the COVID-19 hospitalizations for the Tulsa County, State of Oklahoma, prediction models with and without STL decomposition. C: County, M: Month, S: Sex group, A: Age group.

Model	Random Effects	MAE	RMSE	Corr. (%)	MAE (STL)	RMSE (STL)	Corr. (STL) (%)
M1: Wastewater	County	0.1399	0.1977	63.71	0.0899	0.1377	66.71
	Month	0.1196	0.1729	75.92	0.0696	0.1029	85.12
	Sex Group	0.1297	0.1955	65.79	0.0927	0.1365	67.79
	Age Group	0.1213	0.1860	64.91	0.0913	0.1360	67.91
	C+M+S	0.1102	0.1496	78.21	0.0802	0.1178	85.32
	C+M+A	0.0827	0.1073	82.50	0.0627	0.1003	86.50
M2: NET	County	0.1802	0.2478	52.32	0.1002	0.1478	66.32
	Month	0.1362	0.1907	66.60	0.0762	0.1007	86.20
	Sex Group	0.1802	0.2478	52.32	0.1151	0.1489	62.32
	Age Group	0.1802	0.2478	52.32	0.1102	0.1478	62.62
	C+M+S	0.1370	0.1908	66.46	0.0802	0.1196	83.21
	C+M+A	0.1370	0.1908	66.46	0.0802	0.1178	85.32
M3: Wastewater + NET	County	0.1208	0.1802	74.21	0.0808	0.1202	77.21
	Month	0.0856	0.1089	84.35	0.0696	0.0989	88.35
	Sex Group	0.1159	0.1777	75.21	0.0807	0.1201	77.10
	Age Group	0.1084	0.1597	75.03	0.0804	0.1137	79.03
	C+M+S	0.0835	0.1084	85.69	0.0735	0.1084	86.69
	C+M+A	0.0673	0.0930	89.47	**0.0613**	**0.0930**	**89.47**

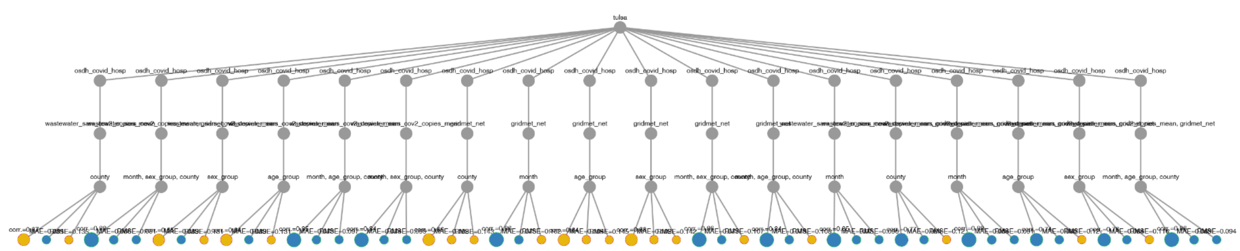


Figure 6.9: Visual summary of model analytics provenance of the COVID-19 hospitalizations for the Tulsa County, State of Oklahoma. Each tree branch represents a unique modeling configuration, including the selected target variable, predictors, random effects, and corresponding model performance metrics. Leaf nodes display correlation (ρ), MAE, and RMSE. Node colors provide quick insight into model quality: blue indicates strong correlation ($\rho > 0.7$) or low error (MAE/RMSE < 0.1), while orange highlights weaker performance. Additionally, node size encodes the magnitude of performance; larger nodes correspond to stronger correlations or lower errors, helping users identify optimal models at a glance. This figure complements Table 6.3 by enabling users to visually trace and compare modeling decisions over time.

significantly improved prediction performance across all models. When comparing the same models with and without STL, those using STL consistently yielded lower MAE and RMSE, as well as higher Spearman correlation coefficients. For instance, model M3, which combines all data sources (wastewater and NET), achieved the best overall performance when the random effects included county, month, and age group. This configuration yielded an MAE of **0.067**, an RMSE of **0.093**, and a correlation of **89.47%**, underscoring the value of multi-source integration.

We also observed that the structure of random effects plays a crucial role in prediction accuracy. While models with only a single grouping variable (e.g., county or month) performed reasonably well, those with multiple random effects—particularly combinations such as county, month, and age group—consistently outperformed simpler models. These findings highlight how combining One Health datasets through an explainable and interactive platform can enhance our understanding of outbreak dynamics.

6.3.4 Case Study IV: Oklahoma (COVID-19 Cases)

To evaluate the robustness of our visual analytics system, we conducted another case study focused on predicting COVID-19 cases for Tulsa County, Oklahoma. We assessed three models (M1–M3), each incorporating different combinations of predictor types: wastewater SARS-CoV-2 and weather NET index. Table 6.4 presents the results, showing model performance with and without STL decomposition. Complementing these quantitative results, Figure 6.10 provides a visual summary of model analytics provenance. Each tree branch represents a unique modeling configuration, displaying the target variable (COVID-19 cases), predictors (wastewater or NET), random effects (county or month), and their associated performance metrics. Leaf nodes show correlation (ρ), MAE, and RMSE, with node colors indicating model quality—blue for strong correlation ($\rho > 0.7$) or low error (MAE/RMSE < 0.1), and red for weaker performance—allowing users to visually trace and compare modeling decisions over time.

Table 6.4: Performance of prediction models for COVID-19 cases prediction with and without STL decomposition.

Model	**Random Effects**	**MAE**	**RMSE**	**Corr. (%)**	**MAE (STL)**	**RMSE (STL)**	**Corr. (STL) (%)**
M1 = Wastewater	County	0.0543	0.1021	64.08	0.0443	0.0721	75.08
	Month	0.0313	0.0531	84.04	0.0213	0.0431	89.04
M2 = NET	County	0.0718	0.1342	48.52	0.0518	0.1042	66.52
	Month	0.0486	0.0891	78.85	0.0426	0.0591	85.15
M3 = Wastewater + NET	County	0.0489	0.0922	70.06	0.0489	0.0622	81.06
	Month	0.0313	0.0509	84.28	**0.0213**	**0.0309**	**91.28**

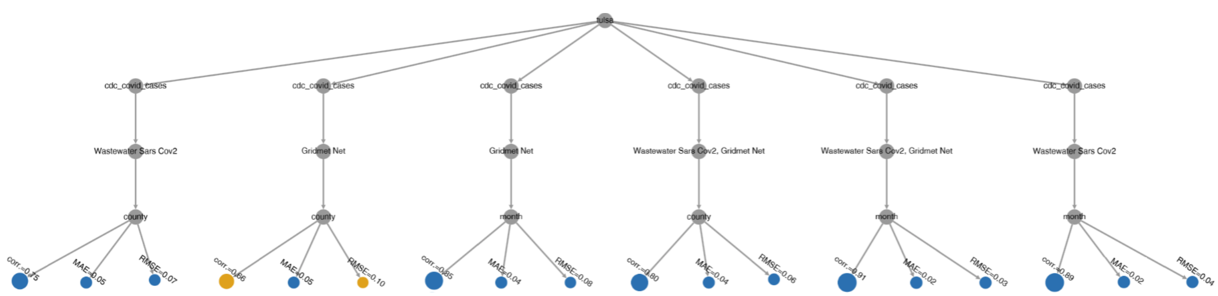


Figure 6.10: Visual summary of model analytics provenance. Each tree branch represents a unique modeling configuration, including the selected target variable, predictors, random effects, and corresponding model performance metrics. Leaf nodes display correlation (ρ), MAE, and RMSE. Node colors provide quick insight into model quality: blue indicates strong correlation ($\rho > 0.7$) or low error (MAE/RMSE < 0.1), while orange highlights weaker performance. This figure complements Table 6.4 by enabling users to visually trace and compare modeling decisions over time.

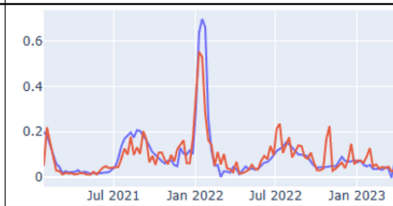
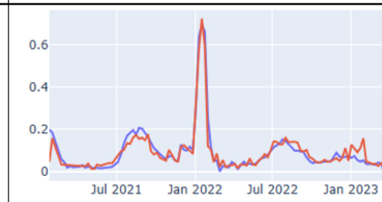
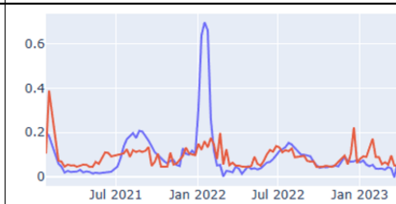
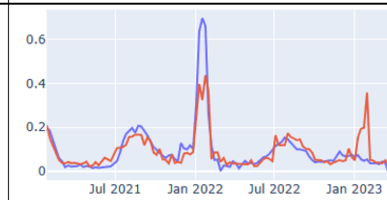
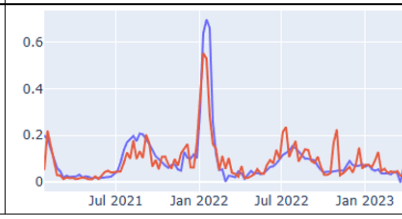
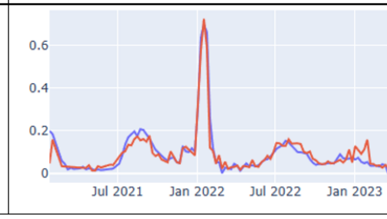
Case Study II: COVID-19 cases prediction using STL		
Model\Random Effects	**County**	**Month**
M1= wastewater		
M2=NET		
M3= wastewater + NET		

Figure 6.11: Predicted (red) vs. actual (blue) COVID-19 case counts across selected models. This figure complements Table 6.4 by visually showing model performance with county or month as random effects.

Overall, models using STL decomposition consistently outperformed their non-STL counterparts, achieving lower MAE and RMSE and higher correlation. The most accurate models included multiple predictors and random effects, which were based on the month. For example, the full model (M3), which includes wastewater and NET data, achieved the best performance with STL preprocessing: an MAE of **0.021**, an RMSE of **0.030**, and a Spearman correlation of **91.28%**. By contrast, the same model without STL achieved only **84.28%** correlation.

Interestingly, month-based random effects generally led to stronger model performance than county-based ones, suggesting consistent seasonal and behavioral patterns across regions. However, this improvement might also stem from the relative coarseness of the grouping: since there are fewer months than counties, more data are available per group, leading to more stable random effect estimates and potentially lower error (Gelman and Hill 2007). Among single-source models, wastewater data alone (M1) performed best, followed by NET

(M2).

These findings reinforce the benefit of integrating heterogeneous predictors and leveraging STL decomposition to isolate trend and anomaly signals in time-series data. They also demonstrate the utility of the system in facilitating comparative exploration and hypothesis testing across modeling scenarios.

The results from the two case studies demonstrate the potential of our visual analytics framework to support explainable, multivariate, and data-driven public health modeling. By enabling interactive model configuration and displaying clear performance comparisons, the system helps users understand the contribution of different predictors and random effect groupings, promoting transparency and supporting proactive public health decision-making.

6.4 User Evaluation of the OH-MEMA Tool

6.4.1 Purpose and Study Design

To evaluate the usability, perceived workload, and effectiveness of the **OH-MEMA** system, a qualitative user evaluation was conducted. The purpose of this study was to examine how domain experts interact with the visual analytics interface, interpret model outputs, and integrate mixed-effects modeling into their analytical workflows for syndromic surveillance.

Participants were invited to complete a short ∼30–40 minute online session conducted via Zoom. Each session involved hands-on interaction with the OH-MEMA dashboard, where users explored model configuration options, applied mixed-effects modeling, and interpreted resulting visualizations.

6.4.2 Participants and Confidentiality

Participants included professionals from relevant One Health domains, such as public health researchers, epidemiologists, and disease surveillance analysts. To ensure confidentiality, each participant was assigned an anonymized identifier (e.g., `U1`, `U2`). Contact information used for scheduling was stored separately and deleted after the coordination was complete. All feedback and workload ratings were analyzed in aggregate form to maintain anonymity.

6.4.3 Procedure and Tasks

Each participant was asked to complete a series of guided tasks designed to simulate realistic analytical use of the tool:

1. **Select Parameters:** Choose a county, a target variable to predict (e.g., COVID-19 hospitalizations), and one or more predictors (e.g., NET).
2. **Apply Modeling:** Run a mixed-effects model using either *county* or *month* as a random effect.
3. **Explore Results:**
    - Compare model performance across different predictor combinations.
    - Interpret results using visual performance metrics.
    - Examine temporal patterns and interdependencies in syndromic surveillance data.
    - Trace and Compare Model Versions:
        - Use the interactive provenance tree to review and compare previously generated models.
        - Identify optimal configurations by inspecting color and size cues that encode model accuracy.
        - Restore prior configurations directly from the tree to continue exploration or refine parameter choices.

After completing the analytical tasks, participants filled out the **NASA Task Load Index (NASA-TLX)** (Hart 2006) questionnaire to rate perceived workload across six dimensions: mental demand, physical demand, temporal demand, performance, effort, and frustration. Each dimension was rated on a 10-point scale (1 = very low, 10 = very high).

6.4.4 NASA-TLX Evaluation

The NASA-TLX instrument was selected as a validated and efficient tool for assessing perceived cognitive and physical workload in interactive analytics environments. Participants provided quantitative ratings on each workload dimension, offering insights into how demanding users found the modeling and visualization processes within OH-MEMA. These measures complemented qualitative feedback to provide a comprehensive and balanced assessment of the user experience.

6.4.5 Open-Ended Feedback

In addition to the NASA-TLX ratings, participants were invited to provide open-ended feedback regarding:

- The perceived effectiveness of the models they created;
- Features they found confusing or particularly useful;
- The potential integration of the tool into their professional workflow; and
- Suggestions for improving usability, clarity, and visualization design.

Responses were qualitatively coded to identify recurring themes related to interpretability, user experience, and analytical efficiency. This analysis provided deeper insight into how users perceived the system's support for model exploration and decision-making.

6.4.6 User Evaluation Results and Discussion

NASA-TLX Workload Ratings

Four participants representing public health and disease surveillance domains completed the evaluation: two epidemiologists and two disease surveillance analysts. The NASA Task Load Index (NASA-TLX) scores across six workload dimensions are summarized in Figure 6.12.

Overall, participants reported low cognitive and physical workload while using the OH-MEMA tool. Average mental demand (3.0), physical demand (2.5), and temporal demand (2.8) were all low, indicating that the interface did not require excessive effort or time pressure. The performance dimension averaged 8.8, suggesting that users felt highly successful in completing the assigned modeling tasks. Similarly, low average effort (2.5) and frustration (2.3) scores suggest that users found the interface intuitive and efficient.

As shown in Figure 6.12, the visual summary of NASA-TLX results highlights a clear contrast between the high perceived performance and low workload dimensions. This pattern reinforces that participants were able to perform complex modeling operations with minimal cognitive strain. The one participant reporting higher workload scores (7 across mental, physical, and temporal demand) noted difficulty interpreting predictor variables, suggesting that the higher effort was related to data interpretation rather than the usability of the interface itself.

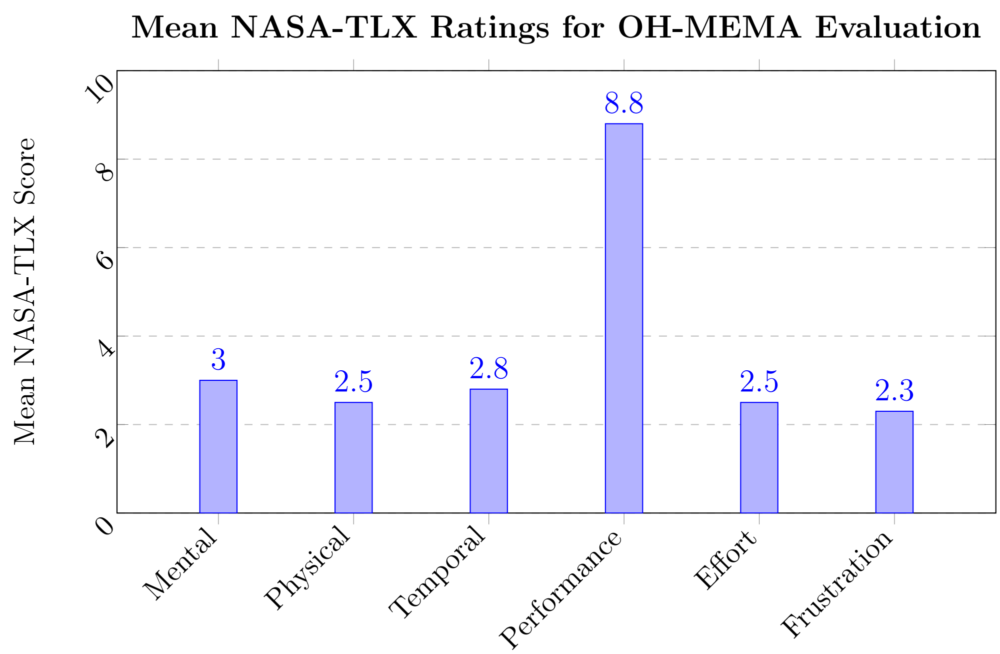


Figure 6.12: Average perceived workload ratings from four participants using the OH-MEMA tool. Low demand and frustration scores indicate low cognitive load, while high performance scores reflect strong task success.

Qualitative Feedback

Open-ended responses provided detailed insights into perceived effectiveness, usability, and areas for improvement. All participants agreed that the models produced by OH-MEMA were effective and informative. Users emphasized the value of being able to explore the relationship between One Health variables and outcomes through visual and interactive modeling.

Positive feedback and usability highlights. Participants appreciated the interactive design and visual feedback mechanisms. Features such as dual-axis charts for county comparison, hover highlights, tooltips, and the ability to redownload model output files were frequently mentioned as useful and intuitive. Users also highlighted the simplicity of transitioning from univariate to multivariate models and the visualization of model history as strengths that supported transparency and exploration.

Areas for improvement. Despite the overall positive experience, participants offered constructive suggestions:

- Improve the clarity of predictor variable names and descriptions to enhance interpretability.
- Provide an inline user guide or quick tutorial explaining model output and terminology.
- Incorporate additional analytical features such as confidence intervals or population density as potential factors.
- Allow deletion or management of historical model trees to prevent interface clutter during extended use.

Workflow integration. All participants indicated that OH-MEMA would make sense within their analytical workflow, particularly for data exploration, model interpretation, and hypothesis generation in surveillance contexts. One epidemiologist noted that the tool supports both “the exploration phase of understanding the data" and “assessing which One Health factors have a stronger impact on outcomes of interest."

The user evaluation provided valuable insights into both the cognitive workload associated with interactive modeling and the practical usability of the OH-MEMA system. The findings guided iterative refinements to the dashboard's layout, labeling, and analytical workflow, ensuring that the tool effectively supports interdisciplinary One Health users in syndromic surveillance and predictive modeling. The NASA-TLX results and qualitative feedback collectively indicate that the OH-MEMA tool is usable, efficient, and cognitively lightweight for expert users. Participants achieved high task performance with low perceived workload, demonstrating that the system design effectively supports mixed-effects modeling and comparative analysis tasks. Suggestions related to documentation, feature labeling, and optional advanced metrics will guide future interface refinements. Overall, these findings validate OH-MEMA's potential as a practical visual analytics system for One Health syndromic surveillance.

6.5 Discussion

Our study highlights the advantages of integrating multiple One Health data sources—human clinical data, wastewater pathogen concentrations, and weather variables—to improve public health syndromic surveillance and epidemic/pandemic detection, prediction, and management at the county level. The combination of these heterogeneous datasets, facilitated by our visual analytics system, enables a more nuanced understanding of the complex drivers behind disease dynamics, thereby improving predictive accuracy.

The significant improvements in prediction performance achieved by applying STL decomposition indicate that separating the seasonal and trend components from the raw time series data helps reveal underlying patterns and reduces noise, thereby enhancing model robustness. This finding aligns with prior research emphasizing the importance of seasonality and trend adjustment in epidemiological forecasting (Basheer et al. 2025).

Moreover, our exploration of different random effect groupings revealed that incorporating multiple hierarchical factors—specifically county, month, and age group—provides a better fit than models with simpler random structures. This suggests that spatial, temporal, and demographic heterogeneities are key contributors to variation in COVID-19 hospitalization rates. Including these random effects not only improves prediction accuracy but also supports a more realistic representation of the underlying data-generating process.

Our visual analytics system is essential for enabling domain experts to explore and compare multiple model configurations interactively. By providing transparent access to model parameters, random effect choices, and performance metrics, it promotes explainability and facilitates informed decision-making. This transparency is critical for public health applications, where trust and understanding of predictive models are essential. Also, integrating provenance visualization into our visual analytics framework not only enhances transparency and reproducibility but also empowers domain experts to make data-driven decisions with confidence, which is essential in the high-stakes context of pandemic response and One Health surveillance.

In addition to the technical evaluation, a qualitative user study was conducted to assess the usability, workload, and effectiveness of the OH-MEMA system. Domain experts, including epidemiologists and disease surveillance analysts, completed hands-on tasks using the dashboard and provided both quantitative ratings via the NASA-TLX instrument and qualitative feedback. The results indicated low perceived cognitive and physical workload, with high performance scores, suggesting that participants were able to successfully navigate complex modeling tasks with minimal strain. Open-ended feedback highlighted the interactive visualizations, provenance tree, and tooltip features as particularly useful for understanding model outcomes and comparing predictor configurations. Participants also appreciated the ability to manage the analytics provenance tree, including the optional deletion of branches, which helped prevent interface clutter and maintained clarity during iterative modeling. Collectively, these findings reinforce that OH-MEMA effectively supports expert users in exploratory mixed-effects modeling, offering a balance between analytical rigor, interpretability, and workflow integration in One Health surveillance contexts.

However, several limitations remain. The temporal resolution and data completeness vary across datasets, which may impact model stability. In addition, while our system performs well with the current scope of county-level data across several variables, its scalability under larger or real-time data scenarios, such as national-level syndromic surveillance, has not yet been fully evaluated.

To address this, the system has been designed using modular data processing and model execution pipelines that can be adapted to parallel and distributed computing environments. Preliminary tests suggest that pre-processing tasks, such as STL decomposition and mixed-effects modeling, can be batch-executed across counties or demographic groups, allowing for future scalability. However, future work will focus on formal benchmarking, database optimization, and integration of real-time data ingestion to support public health operations at scale.

Future work could also explore the integration of more granular spatial and demographic data, including mobility patterns and vaccination rates, to further enhance predictive capabilities.

Overall, our approach demonstrates the promise of combining One Health data streams with advanced statistical modeling and interactive visualization to support timely and actionable public health surveillance and intervention. By offering an explainable and scalable decision-support platform, this work contributes toward building more resilient and responsive systems for future epidemic and pandemic response.

6.6 Summary

In this chapter, we introduced **OH-MEMA**, the *One Health Mixed-Effects Modeling Approach*, a visual analytics system designed to enable transparent and interactive modeling of pandemic-related outcomes using heterogeneous One Health data. The framework integrates mixed-effects models with seasonal-trend decomposition and visual tools for predictor selection, model configuration, and interpretability. By modeling diverse factors such as wastewater viral concentrations and weather conditions as fixed effects, and incorporating demographic and geographic groupings as random effects, OH-MEMA facilitates the exploration of complex, multiscale relationships that influence public health outcomes. The system is publicly accessible at `https://mema.disc.ourcloud.ou.edu/` to promote broader adoption and reproducibility.

To support iterative exploration and model refinement, OH-MEMA includes a visual analytics provenance tree that records each modeling configuration, including the target, predictors, random effects, and evaluation metrics such as correlation, MAE, and RMSE. This feature enhances transparency, reproducibility, and user-guided decision-making throughout the modeling process.

We demonstrated the effectiveness of OH-MEMA through four case studies: predicting influenza and RSV cases in Arizona, COVID-19 cases and deaths in Arkansas, and COVID-19 cases and hospitalizations in Oklahoma. The findings highlight the advantages of integrating multiple data sources to enhance forecast accuracy and capture contextual variations through the use of hierarchical random effects.

Combining predictors from human and environmental domains significantly improved performance, especially when accounting for seasonality via STL decomposition. OH-MEMA enables users not only to identify the most relevant predictors but also to understand how random effects shape model performance and interpretability.

To further assess the system's usability and perceived value, we conducted a qualitative user study with domain experts from public health and epidemiology. Participants engaged in guided exploratory sessions using OH-MEMA to configure models, compare predictors, and interpret results. Feedback from these interviews indicated that users found the system intuitive, informative, and supportive of data-driven reasoning. Experts particularly valued the provenance tracking, hierarchical modeling interface, and transparent visualization of fixed and random effects. The study also revealed opportunities to enhance accessibility and extend model interpretability features for broader audiences, informing future design iterations of the framework.

By uniting statistical rigor with visual interactivity, OH-MEMA fosters a white-box modeling environment that supports hypothesis generation, policy analysis, and early warning applications for public health officials and researchers. This work underscores the potential of integrated, explainable modeling platforms to advance data-driven decision-making in complex and evolving One Health contexts.

Chapter 7

Conclusion

This dissertation examined how integrating One Health data, encompassing human, animal, and environmental domains, with visual analytics can enhance pandemic preparedness and response. Across four studies, it examined the current state of One Health systems, the role of visualization in pandemics and the levels of visualization readiness, the design and implementation of the PIPP dashboard, and the development of a mixed-effects modeling framework for interactive, explainable analysis. Together, these works contribute both conceptual insights and practical systems toward building more resilient and data-informed approaches to emerging health threats.

7.1 Main Contributions

The dissertation makes the following contributions:

- **Systematic mapping of One Health dashboards (2015–2024):** The review categorized systems by purpose, data domains, visualization techniques, and interactivity. It revealed a strong interest in integrated approaches but found that fully realized One Health implementations remain rare. This provides a baseline for future research and identifies specific gaps that require attention.

- **Conceptualization of visualization readiness:** The survey analysis of visualization in pandemic contexts introduced the idea of "visualization readiness", the ability of VIS/VA communities to rapidly support crisis response. The study identified opportunities for VIS beyond communication, emphasizing infrastructure, literacy, and theory-driven design for decision support and public engagement.

- **Design of the Predictive Intelligence for Pandemic Prevention (PIPP) dashboard:** The system integrates diverse datasets (human health, animal health, environmental, and wastewater) with interactive visualization tools. It provides stakeholders with a flexible environment for exploration, hypothesis generation, and preliminary forecasting.

- **Development of a lag-aware mixed-effects visual analytics framework:** This modeling system combines statistical rigor with interactive transparency, enabling users to explore predictors, evaluate model structures, and compare outcomes. Features such as STL decomposition and provenance tracking support interpretability, reproducibility, and user-guided analysis.

Collectively, these contributions advance the integration of heterogeneous data, epidemiological modeling, and interactive visualization into a coherent One Health visual analytics framework.

7.2 Future Directions

Several avenues remain for extending this work:

- **Scalability and infrastructure:** While demonstrated at the county level, the framework should be stress-tested and optimized for larger scales (state, national, or global) and for streaming data environments.

- **Visualization for decision support and communication:** More empirical studies are needed to understand how visual analytics impacts expert workflows, public trust, and policy-making. This includes developing rapid prototyping methods for emergencies and strategies for managing misinformation.

- **Interdisciplinary collaboration:** Advancing One Health visual analytics requires sustained collaboration across epidemiology, veterinary science, environmental science, visualization research, and public health practice. Institutionalizing such collaboration will be crucial for achieving a long-term impact.

- **User-selectable forecasting horizon:** Enhance the PIPP dashboard transparency and usability by allowing users to view and select the forecasting window for predictions. This would enable users to customize the temporal scope of the forecasts and more effectively interpret the results for informed decision-making.

- **Benchmarking against existing surveillance systems:** Compare OH-MEMA's predictive and analytical performance with established public health surveillance systems such as ESSENCE, using open-source or publicly available data. Such benchmarking would help contextualize the framework's strengths and limitations and highlight opportunities for methodological refinement.

- **Temporal aggregation and model performance:** Examine how the mixed-effects model performs when using monthly data instead of weekly data. Assessing the impact of temporal aggregation on predictive accuracy and model stability would provide additional insight into the robustness and generalizability of the modeling approach.

7.3 Final Thoughts

By situating dashboards, visualization readiness, system design, and explainable modeling within a unified framework, this dissertation illustrates how One Health visual analytics can help bridge gaps between data, models, and decision-making. Its contributions lay the foundation for future systems that are not only more integrated and predictive but also more transparent, interpretable, and actionable—qualities essential for building resilience against the pandemics of tomorrow.

Bibliography

Empowering Global AMR Research Community: Interactive GIS ..., May 2024. URL `https://wellcomeopenresearch.org/articles/9-234`.

Alaa Abd-alrazaq, Najeeb Al-Shorbaji, Kheira Lakhdari, Ahmed Elbokl, Hani Farouk, Hoda Wahba, Naema Elgasser, Rania Mohell, Tamer Emara, Haleh Ayatollahi, Sharareh R. Niakan Kalhori, Reza Rabiei, Mahmood Tara, Laila Akhu-Zaheya, Raeda Al-Qutob, Sadam Alabed Alrazak, Eiman Al-Jafar, Dari Alhuwail, Hassan Ghazal, Zineb El Otmani Dehbi, Najib Al Idrissi, Ouafaa Fassi Fihri, Lahcen Belyamani, Tariq Shahzad, Zakiuddin Ahmed, Arfan Ahmad, Zubair Shah, Eman Abu Hamra, Adel Taweel, Abdulqadir J. Nashwan, Mowafa Househ, Mounir Hamdi, Dena Al-Thani, Tanvir Alam, Abdulwahhab Alshammari, Sana Alnafrani, and Haitham Alali. Middle East and North African Health Informatics Association (MENAHIA): Technological initiatives for 'One Health'. *Yearbook of Medical Informatics*, 32:320–329, December 2023. ISSN 0943-4747. doi: 10.1055/s-0043-1768740. URL `https://www.thieme-connect.de/products/ejournals/abstract/10.1055/s-0043-1768740`. Publisher: Georg Thieme Verlag KG.

Shehzad Afzal, Ross Maciejewski, and David S. Ebert. Visual analytics decision support environment for epidemic modeling and response evaluation. In *6th IEEE Conference on Visual Analytics Science and Technology, IEEE VAST 2011, Providence, RI, USA, October 23-28, 2011*, pages 191–200. IEEE Computer Society, 2011. doi: 10.1109/VAST.2011.6102457. URL `https://doi.org/10.1109/VAST.2011.6102457`.

Shehzad Afzal, Sohaib Ghani, Hank C. Jenkins-Smith, David S. Ebert, Markus Hadwiger, and Ibrahim Hoteit. A visual analytics based decision making environment for COVID-19 modeling and visualization. In *31st IEEE Visualization Conference, IEEE VIS 2020 - Short Papers, Virtual Event, USA, October 25-30, 2020*, pages 86–90. IEEE, 2020. doi: 10.1109/VIS47514.2020.00024. URL `https://doi.org/10.1109/VIS47514.2020.00024`.

Imran Ahmed, Misbah Ahmad, et al. A framework for pandemic prediction using big data analytics. *Big Data Research*, 25, 2021.

Imran Ahmed, Abdellah Chehri, and Gwanggil Jeon. Artificial intelligence and blockchain enabled smart healthcare system for monitoring and detection of covid-19 in biomedical images. *Trans. Comput. Bio. & Bioinformatics*, 2023.

Samar Al-Hajj. Collaborative visual analytics for public health: facilitating problem solving and supporting decision-making. 2014.

Ali Al-Hemoud, Manar AlSaraf, Mariam Malak, Musab Al-Shatti, Meshael Al-Jarba, Ahmad Othman, Hanadi Al-Shammari, and Alya Al-Shatti. Analytical and Early Detection System of Infectious Diseases and Animal Health Status in Kuwait. *Frontiers in Veterinary Science*, 8, July 2021a. ISSN 2297-1769. doi: 10.3389/fvets.2021.676661. URL https://www.frontiersin.org/journals/veterinary-science/articles/10.3389/fvets.2021.676661/full. Publisher: Frontiers.

Ali Al-Hemoud, Manar AlSaraf, Mariam Malak, Musab Al-Shatti, Meshael Al-Jarba, Ahmad Othman, Hanadi Al-Shammari, and Alya Al-Shatti. Analytical and early detection system of infectious diseases and animal health status in kuwait. *Frontiers in Veterinary Science*, 8:676661, 2021b.

Justin Amadeus Albert et al. Will you be my quarantine: A computer vision and inertial sensor based home exercise system. In *Conf. Perv. Comp. Tech. Healthcare*, pages 380–383, 2020.

Hugh Alderwick, Andrew Hutchings, Adam Briggs, and Nicholas Mays. The impacts of collaboration between local health care and non-health care organizations and factors shaping how they work: a systematic review of reviews. *BMC public health*, 21(1):753, 2021.

William L. Allen et al. Next slide please: The politics of visualization during Covid-19 press briefings. *J. Euro. Pub. Pol.*, 31(3):729–755, 2024.

Heer Ambavi, Kavita Vaishnaw, et al. CovidExplorer: A multi-faceted AI-based search and visualization engine for COVID-19 information. In *Proc. 29th ACM Int. Conf. Inf. & Knowledge Man.*, pages 3365–3368, 2020.

Natalia Andrienko, Tim Lammarsch, Gennady Andrienko, Georg Fuchs, Daniel Keim, Silvia Miksch, and Andrea Rind. Viewing visual analytics as model building. In *Computer graphics forum*, volume 37, pages 275–299. Wiley Online Library, 2018.

Dario Antweiler et al. Uncovering chains of infections through spatio-temporal and visual analysis of covid-19 contact traces. *Comp. & Graph.*, 106:1–8, 2022.

Catherine Arsenault, Anagaw Derseh Mebratie, Solomon Kassahun Gelaw, and Dessalegn Shamebo. Routine data in a primary care performance dashboard, Ethiopia. *Bulletin of the World Health Organization*, 102(07):465–475A, July 2024. ISSN 00429686, 15640604. doi: 10.2471/BLT.23.291122. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC11197649/pdf/BLT.23.291122.pdf.

Benjamin Bach, Euan Freeman, et al. Dashboard design patterns. *IEEE Trans. Vis. Comput. Graph.*, 29(1):342–352, 2023. URL https://doi.org/10.1109/TVCG.2022.3209448.

Sriram Karthik Badam, Jieqiong Zhao, Niklas Elmqvist, and David S Ebert. Timefork: Mixed-initiative time-series prediction. In *2014 IEEE Conference on Visual Analytics Science and Technology (VAST)*, pages 223–224. IEEE, 2014.

Apollinaire Batoure Bamana, Mahdi Shafiee Kamalabad, and Daniel L Oberski. A systematic literature review of time series methods applied to epidemic prediction. *Informatics in Medicine Unlocked*, page 101571, 2024.

Amitava Banerjee, Laura Pasea, Steve Harris, et al. Estimating excess 1-year mortality associated with the COVID-19 pandemic according to underlying conditions and age: a population-based cohort study. *The Lancet*, 395(10238):1715–1725, 2020.

Aseel Basheer, Wolfgang Jentner, and David Ebert. Leveraging visual analytics and diverse datasets for proactive pandemic surveillance: A one health approach. 2025.

Andrea Batch, Peter WS Butcher, Panagiotis D Ritsos, and Niklas Elmqvist. Wizualization: A "hard magic" visualization system for immersive and ubiquitous analytics. *IEEE Trans. Vis. Comp. Graph.*, 30(1):507–517, 2023.

Tom Baumgartl, Markus Petzold, Marcel Wunderlich, Markus Höhn, , et al. In search of patient zero: Visual analytics of pathogen transmission pathways in hospitals. *IEEE Trans. Vis. Comput. Graph.*, 27(2):711–721, 2021. URL `https://doi.org/10.1109/TVCG.2020.3030437`.

Casey Barton Behravesh, Dominique F. Charron, Amanda Liew, Natalia Cediel Becerra, Catherine Machalaba, David T.S. Hayman, Janice R. Ciacci Zanella, Elmoubasher Farag, Abhishek Chaudhary, Hayley Belles, Wiku B. Adisasmito, Salama Almuhairi, Pépé Bilivogui, Salome A. Bukachi, Natalia Casas, Andrew A. Cunningham, Nitish Debnath, Osman Dar, Baptiste Dungu, George F. Gao, Margaret Khaitsa, Marion P.G. Koopmans, John S. Mackenzie, Serge Morand, Vyacheslav Smolenskiy, Lei Zhou, Wanda Markotter, and Thomas C. Mettenleiter. An integrated inventory of One Health tools: Mapping and analysis of globally available tools to advance One Health. *CABI One Health*, 3(1), June 2024. doi: 10.1079/cabionehealth.2024.0017. URL `https://www.cabidigitallibrary.org/doi/10.1079/cabionehealth.2024.0017`. Publisher: CABI.

Anna Bernasconi and Silvia Grandi. A conceptual model for geo-online exploratory data visualization: The case of the COVID-19 pandemic. *Information*, 12(2), 2021. ISSN 2078-2489. URL `https://www.mdpi.com/2078-2489/12/2/69`.

Lidia Betcheva, Feryal Erhun, Antoine Feylessoufi, Peter Fryers, Paulo Gonçalves, Houyuan Jiang, Paul Kattuman, Tom Pape, Anees Pari, Stefan Scholtes, et al. An adaptive research approach to covid-19 forecasting for regional health systems in england. *INFORMS Journal on Applied Analytics*, 2024.

Pawandeep Kaur Betz, Julien Stoll, Valerie Grappendorf, Jonas Gilg, Moritz Zeumer, Margrit Klitz, Luca Spataro, Anna Klein, Lena Rothenhäusler, Hartmut Bohnacker, et al. Esid: A visual analytics tool to epidemiological emergencies. *CoRR*, 2023.

Keith R Bisset, Jiangzhuo Chen, Suruchi Deodhar, Xizhou Feng, Yifei Ma, and Madhav V Marathe. Indemics: An interactive high-performance computing framework for data-intensive epidemic modeling. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 24(1):1–32, 2014.

Isabelle-Anne Bisson, Benard J. Ssebide, and Peter P. Marra. Early Detection of Emerging Zoonotic Diseases with Animal Morbidity and Mortality Monitoring. *EcoHealth*, 12(1): 98–103, March 2015. ISSN 1612-9202, 1612-9210. doi: 10.1007/s10393-014-0988-x. URL http://link.springer.com/10.1007/s10393-014-0988-x.

John Bollinger. *Bollinger on Bollinger bands*. McGraw-Hill New York, 2002.

Marion Bordier, Theethawat Uea-Anuwong, Aurélie Binot, Pascal Hendrikx, and Flavie L. Goutard. Characteristics of One Health surveillance systems: A systematic literature review. *Preventive Veterinary Medicine*, 181:104560, August 2020. ISSN 01675877. doi: 10.1016/j.prevetmed.2018.10.005. URL https://linkinghub.elsevier.com/retrieve/pii/S0167587718303659.

Riccardo Bovo et al. Cone of vision as a behavioural cue for vr collaboration. *Proc. ACM Hum.-Comput. Interact.*, 6, November 2022. URL https://doi.org/10.1145/3555615.

Magoba Bridget, Gebrekrstos Negash Gebru, George S Odongo, Calle Hedberg, Adel Hussein Elduma, Joseph Sam Kanu, James Bangura, James Sylvester Squire, and Monique A Foster. Digitalizing disease surveillance: experience from Sierra Leone. *Health Policy and Planning*, page czae039, May 2024. ISSN 1460-2237. doi: 10.1093/heapol/czae039. URL https://academic.oup.com/heapol/advance-article/doi/10.1093/heapol/czae039/7685088.

Álvaro Briz-Redón. The impact of modelling choices on modelling outcomes: a spatio-temporal study of the association between covid-19 spread and environmental conditions in catalonia (spain). *Stochastic Environmental Research and Risk Assessment*, 35(8):1701–1713, 2021.

Chris Bryan, Xue Wu, Susan Mniszewski, and Kwan-Liu Ma. Integrating predictive analytics into a spatiotemporal epidemic simulation. In *2015 IEEE Conference on Visual Analytics Science and Technology (VAST)*, pages 17–24. IEEE, 2015.

James W Buehler, Richard S Hopkins, J Marc Overhage, Daniel M Sosin, and Vincent; CDC Working Group Tong. Framework for evaluating public health surveillance systems for early detection of outbreaks: recommendations from the cdc working group. *MMWR Recommendations and Reports*, 53:1–11, 2004.

Román Cárdenas et al. A modeling and simulation platform for space-based compartmental modeling of pandemic spread. In *Ann. Modeling & Sim. Conf.*, pages 1–12, 2021.

Lauren N Carroll, Alan P Au, Landon Todd Detwiler, Tsung-chieh Fu, Ian S Painter, and Neil F Abernethy. Visualization and analytics tools for infectious disease epidemiology: a systematic review. *Journal of biomedical informatics*, 51:287–298, 2014.

CDC. One health basics. https://www.cdc.gov/onehealth/basics/index.html, 2023. Accessed:2024-03-24.

Centers for Disease Control and Prevention, COVID-19 Response. Weekly United States COVID-19 Cases and Deaths by County. https://data.cdc.gov/dataset/Weekly-United-States-COVID-19-Cases-and-Deaths-by-/yviw-z6j5/about_data, 2023. Version date: July 10,2023.

Mehdi Chakhchoukh, Nadia Boukhelifa, and Anastasia Bezerianos. Understanding how in-visualization provenance can support trade-off analysis. *IEEE Transactions on Visualization and Computer Graphics*, 29(9):3758–3774, 2022.

Yung-Chun Chang, Chih-Hao Ku, and Duy-Duc Le Nguyen. Predicting aspect-based sentiment using deep learning and information visualization: The impact of covid-19 on the airline industry. *Inf. & Man.*, 59(2):103587, 2022.

Angelos Chatzimparmpas, Rafael M Martins, Kostiantyn Kucher, and Andreas Kerren. Featureenvi: Visual analytics for feature engineering using stepwise selection and semi-automatic extraction approaches. *IEEE Transactions on Visualization and Computer Graphics*, 28(4):1773–1791, 2022.

Dongyu Chen, Yingwei Liu, Yang Liu, Kejun Zhao, Tianhan Zhang, Yangyang Gao, Qi Wang, Baoan Song, and Gefei Hao. ChemFREE: a one-stop comprehensive platform for ecological and environmental risk evaluation of chemicals in one health world. *Nucleic Acids Research*, 52(W1):W450–W460, July 2024. ISSN 0305-1048, 1362-4962. doi: 10.1093/nar/gkae446. URL https://academic.oup.com/nar/article/52/W1/W450/7687434.

M. Chen and D. S. Ebert. An ontological framework for supporting the design and evaluation of visual analytics systems. *Computer Graphics Forum*, 38(3):131–144, 2019.

M. Chen et al. RAMPVIS: Answering the challenges of building visualisation capabilities for large-scale emergency responses. *Epidemics*, 39:100569, 2022.

Min Chen and Amos Golan. What may visualization processes optimize? *IEEE Trans. Vis. Comp. Graph.*, 22(12):2619–2632, 2015.

Peng Chen and Beth A Plale. Big data provenance analysis and visualization. In *2015 15th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing*, pages 797–800. IEEE, 2015.

Victor Chernozhukov, Hiroyuki Kasahara, and Paul Schrimpf. Causal impact of masks, policies, behavior on early covid-19 pandemic in the us. *Journal of econometrics*, 220(1):23–62, 2021.

Fanny Chevalier et al. Observations and reflections on visualization literacy in elementary school. *IEEE Comp. Graph. & App.*, 38(3):21–29, 2018.

Jocalyn Clark. How Covid-19 bolstered an already perverse publishing system. *BMJ*, 380, 2023. doi: 10.1136/bmj.p689.

Melissa D Clarkson. Web-based covid-19 dashboards and trackers in the united states: Survey study. *JMIR Human Factors*, 10:e43819, 2023.

Climatology Lab. GridMet: Daily Meteorological and Climatological Data. https://www.climatologylab.org/gridmet.html, 2023. Accessed on July 13, 2023.

Andre Coetzer, Terence P. Scott, Khadija Noor, Lambert F. Gwenhure, and Louis H. Nel. A Novel Integrated and Labile eHealth System for Monitoring Dog Rabies Vaccination Campaigns. *Vaccines*, 7(3):108, September 2019. ISSN 2076-393X. doi: 10.3390/vaccines7030108. URL https://www.mdpi.com/2076-393X/7/3/108.

Peter J Collignon and Scott A McEwen. One health—its importance in helping to better control antimicrobial resistance. *Tropical medicine and infectious disease*, 4(1):22, 2019.

Christopher Collins et al. Covid connect: Chat-driven anonymous story-sharing for peer support. In *Proc. ACM Designing Interactive Sys. Conf.*, pages 301–318, 2022.

Francis S Collins and Paul Stoffels. Accelerating covid-19 therapeutic interventions and vaccines (activ): an unprecedented partnership for unprecedented times. *Jama*, 323(24): 2455–2457, 2020.

Joao LD Comba. Data visualization for the understanding of covid-19. *Computing in Science & Engineering*, 22(6):81–86, 2020.

Caitlin M. Cossaboom, Natalie M. Wendling, Nathaniel M. Lewis, Hannah Rettler, Robert R. Harvey, Brian R. Amman, Jonathan S. Towner, Jessica R. Spengler, Robert Erickson, Cindy Burnett, Erin L. Young, Kelly Oakeson, Ann Carpenter, Markus H. Kainulainen, Payel Chatterjee, Mike Flint, Anna Uehara, Yan Li, Jing Zhang, Anna Kelleher, Brian Lynch, Adam C. Retchless, Suxiang Tong, Ausaf Ahmad, Paige Bunkley, Claire Godino, Owen Herzegh, Jan Drobeniuc, Jane Rooney, Dean Taylor, and Casey Barton Behravesh. One Health Investigation of SARS-CoV-2 in People and Animals on Multiple Mink Farms in Utah. *Viruses*, 15(1):96, December 2022. ISSN 1999-4915. doi: 10.3390/v15010096. URL https://www.mdpi.com/1999-4915/15/1/96.

Robert Jay C. Cruz, Rene Victor A. San Juan, Ralph Christian S. Cristobal, Lee Ann D. Lo, and Leonila B. Valdez. Mosquinator: A Community-Based Mobile and Web Application Framework for combating Dengue and providing Awareness for Residents of Pasig City. In *2024 7th International Conference on Information and Computer Technologies (ICICT)*, pages 326–331, March 2024. doi: 10.1109/ICICT62343.2024.00059. URL https://ieeexplore.ieee.org/document/10541723. ISSN: 2769-4542.

Liuyuan Cui and Weiwei Kong. Visualization analysis of spatiotemporal data of covid-19. In *Int. Conf. Int. Sys. & Know. Eng.*, pages 565–571, 2021.

Frank C Curriero, Cara Wychgram, Alison W Rebman, Anne E Corrigan, Anton Kvit, Timothy Shields, and John N Aucott. The lyme and tickborne disease dashboard: A map-based resource to promote public health awareness and research collaboration. *PLoS One*, 16(12):e0260122, 2021a.

Frank C. Curriero, Cara Wychgram, Alison W. Rebman, Anne E. Corrigan, Anton Kvit, Timothy Shields, and John N. Aucott. The Lyme and Tickborne Disease Dashboard: A map-based resource to promote public health awareness and research collaboration. *PLOS ONE*, 16(12):e0260122, December 2021b. ISSN 1932-6203. doi: 10.1371/journal.pone.0260122. URL https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0260122. Publisher: Public Library of Science.

Soledad Natalia M. Dalisay, Carlo R. Lumangaya, Lorenzo Maria C. De Guzman, Robert Neil F. Leong, Taggart G. Siao, Juan Alfonso Leonardia, Chiqui De Verya, and Vicente Y. Belizario Jr. A qualitative analysis of the implementation of the water, sanitation, and hygiene in schools program in the Philippines using the One Health lens. *International Journal of One Health*, pages 1–11, January 2024. ISSN 24558931, 24555673. doi: 10.14202/IJOH.2024.1-11. URL https://www.onehealthjournal.org/Vol.10/No.1/1.html.

Madhuchhanda Das, Venencia Albert, Samaresh Das, Karma Gurmey Dolma, Tapan Majumdar, Pranjal Jyoti Baruah, Suranjana Chaliha Hazarika, Basumoti Apum, and Thandavarayan Ramamurthy. An integrated FoodNet in North East India: fostering one health approach to fortify public health. *BMC Public Health*, 24(1):451, February 2024. ISSN 1471-2458. doi: 10.1186/s12889-024-18007-w. URL https://bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-024-18007-w.

Maria L Daza-Torres, J Cricelio Montesinos-López, Heather N Bischel, Colleen C Naughton, Angel N Desai, Marlene K Wolfe, Alexandria B Boehm, and Miriam Nuño. A mixed-effects model to predict covid-19 hospitalizations using wastewater surveillance. *Journal of Environmental Chemical Engineering*, 12(2):112485, 2024.

Joost CF De Winter, Samuel D Gosling, and Jeff Potter. Comparing the pearson and spearman correlation coefficients across distributions and sample sizes: A tutorial using simulations and empirical data. *Psychological methods*, 21(3):273, 2016.

Chris Degeling, Jane Johnson, Ian Kerridge, Andrew Wilson, Michael Ward, Cameron Stewart, and Gwendolyn Gilbert. Implementing a one health approach to emerging infectious disease: reflections on the socio-political, ethical and legal dimensions. *BMC Public Health*, 15:1–11, 2015.

Amélie Desvars-Larrive, Anna Elisabeth Vogl, Gavrila Amadea Puspitarani, Liuhuaying Yang, Anja Joachim, and Annemarie Käsbohrer. A One Health framework for exploring zoonotic interactions demonstrated through a case study. *Nature Communications*, 15(1):5650, July 2024. ISSN 2041-1723. doi: 10.1038/s41467-024-49967-7. URL https://www.nature.com/articles/s41467-024-49967-7.

Samrat K Dey, Md Mahbubur Rahman, Umme R Siddiqi, and Arpita Howlader. Analyzing the epidemiological outbreak of covid-19: A visual exploratory data analysis approach. *Journal of medical virology*, 92(6):632–638, 2020.

Alessio Di Lorenzo, Iolanda Mangone, Patrizia Colangeli, Daniela Cioci, Valentina Curini, Giacomo Vincifori, Maria Teresa Mercante, Adriano Di Pasquale, Nicolas Radomski, and Simona Iannetti. One health system supporting surveillance during covid-19 epidemic in abruzzo region, southern italy. *One Health*, 16:100471, 2023a.

Alessio Di Lorenzo, Iolanda Mangone, Patrizia Colangeli, Daniela Cioci, Valentina Curini, Giacomo Vincifori, Maria Teresa Mercante, Adriano Di Pasquale, Nicolas Radomski, and Simona Iannetti. One health system supporting surveillance during COVID-19 epidemic in Abruzzo region, southern Italy. *One Health*, 16:100471, June 2023b. ISSN 2352-7714. doi: 10.1016/j.onehlt.2022.100471. URL `https://www.sciencedirect.com/science/article/pii/S2352771422001033`.

Nomita Divi, Jaś Mantero, Marlo Libel, Onicio Leal Neto, Marinanicole Schultheiss, Kara Sewalk, John Brownstein, and Mark Smolinski. Using EpiCore to Enable Rapid Verification of Potential Health Threats: Illustrated Use Cases and Summary Statistics. *JMIR Public Health and Surveillance*, 10(1):e52093, March 2024. doi: 10.2196/52093. URL `https://publichealth.jmir.org/2024/1/e52093`. Company: JMIR Public Health and Surveillance Distributor: JMIR Public Health and Surveillance Institution: JMIR Public Health and Surveillance Label: JMIR Public Health and Surveillance Publisher: JMIR Publications Inc., Toronto, Canada.

J. Dykes et al. Visualization for epidemiological modelling: challenges, solutions, reflections & recommendations. *Phil. Trans. Royal Society A*, 380(2233), 2022.

Kaylee Myhre Errecaborde, Katharine M Pelican, Heidi Kassenborg, Ong-Orn Prasarnphanich, Linda Valeri, Erinaldi Yuuzar, Rama Prima Syahti Fauzi, Nyoman Sri Budayanti, Agus Suwandono, Wayan T Artama, et al. Piloting the one health systems mapping and analysis resource toolkit in indonesia. *Ecohealth*, 14:178–181, 2017.

European Food Safety Authority (EFSA), John Berezowski, Katinka de Balogh, Fernanda C Dórea, Simon Rüegg, Alessandro Broglia, Andrea Gervelmeyer, and Lisa Kohnle. Prioritisation of zoonotic diseases for coordinated surveillance systems under the One Health approach for cross-border pathogens that threaten the Union. *EFSA Journal*, 21(3), March 2023. ISSN 18314732, 18314732. doi: `10.2903/j.efsa.2023.7853`. URL `https://data.europa.eu/doi/10.2903/j.efsa.2023.7853`.

Darlington David Faijue, Allison Osorio Segui, Kalpita Shringarpure, Ahmed Razavi, Nadeem Hasan, Osman Dar, and Logan Manikam. Constructing a one health governance architecture: a systematic review and analysis of governance mechanisms for one health. *European journal of public health*, 34(6):1086–1094, 2024.

Nichol John F Famadico et al. A covid-19 epidemiological analysis and forecasting dashboard for hospitals using time-series analysis and deep learning. In *IEEE 3rd Int. Conf. Comp. & Machine Int.*, pages 1–5, 2024.

Liu Fangfang et al. Research on visualization of emergency response to covid-19:——social network analysis based on microblog data. In *6th Int. Conf. Int. Comp. & Signal Processing*, pages 1205–1208, 2021.

Faradiba Faradiba, Satibi Satibi, Lutfan Lazuardi, and Anselma Ivanawati. Development of web-based application of medicine management dashboard for pharmacists in primary health care using a user-centered design: evidence from Indonesia. *JURNAL ILMU KEFARMASIAN INDONESIA*, 22(1):122–136, April 2024. ISSN 2614-6495. doi: 10.35814/jifi.v22i1.1574. URL http://jifi.farmasi.univpancasila.ac.id/index.php/jifi/article/view/1574. Number: 1.

Mehmedović Faris and Karabegović Almir. Data science-driven system design with lstm neural networks: A covid-19 perspective. In *INFOTEH*, pages 1–6, 2024.

Folorunso O. Fasina, Olubunmi G. Fasanmi, Yilma J. Makonnen, Charles Bebay, Bernard Bett, and Kristina Roesel. The one health landscape in Sub-Saharan African countries. *One Health*, 13:100325, December 2021. ISSN 2352-7714. doi: 10.1016/j.onehlt.2021.100325. URL https://www.sciencedirect.com/science/article/pii/S2352771421001154.

Elif E. Firat, Alark Joshi, and Robert S. Laramee. VisLitE: Visualization literacy and evaluation. *IEEE Comp. Graph. & App.*, 42(3):99–107, 2022.

Stefania-Mihaela Firiza et al. Visualizing the covid-19 pandemic: Retrospective insights and implications for public health strategies. In *IEEE 18th Int. Symp. Appl. Comp. Int. & Inf.*, pages 000461–000466, 2024.

Food and Agriculture Organization (FAO). One Health Initiative. https://www.fao.org/one-health/en/, 2024. Accessed on April 24, 2024.

Food and Agriculture Organization of the United Nations. One Health, 2024. URL https://www.fao.org/one-health/en/. Accessed: 2024-11-04.

Ronald D Fricker Jr, Benjamin L Hegler, and David A Dunfee. Comparing syndromic surveillance detection methods: Ears'versus a cusum-based methodology. *Statistics in Medicine*, 27(17):3407–3429, 2008.

Vincent D Friedrich, Peter Pennitz, Emanuel Wyler, Julia M Adler, et al. Neural network-assisted humanisation of covid-19 hamster transcriptomic data reveals matching severity states in human disease. *EBioMedicine*, 108, 2024.

Christophe Gaie. Providing detailed information on national policies to cope with the covid-19 pandemic. *Dig. Gov.: Res. & Prac.*, 2(1):1–11, 2020.

Esteban García-Cuesta et al. Explainable risk prediction on streaming data: a LSTMs application to SARS-CoV-2 patients. In *IEEE Int. Conf. Evo. & Adap. Int. Sys.*, pages 1–9, 2024.

Andrew Gelman and Jennifer Hill. *Data analysis using regression and multilevel/hierarchical models*. Cambridge university press, 2007.

Ria R. Ghai, Ryan M. Wallace, James C. Kile, Trevor R. Shoemaker, Antonio R. Vieira, Maria E. Negron, Sean V. Shadomy, Julie R. Sinclair, Grace W. Goryoka, Stephanie J. Salyer, and Casey Barton Behravesh. A generalizable one health framework for the control of zoonotic diseases. *Scientific Reports*, 12(1):8588, May 2022. ISSN 2045-2322. doi: 10.1038/s41598-022-12619-1. URL `https://www.nature.com/articles/s41598-022-12619-1`.

Tuna Gonen, Yiwen Xing, Cagatay Turkay, Alfie Abdul-Rahman, Radu Jianu, Hui Fang, Euan Freeman, Franck P Vidal, and Min Chen. Visual analytics based search-analyze-forecast framework for epidemiological time-series data. In *2023 IEEE VIS Workshop on Visualization for Pandemic and Emergency Responses (Vis4PandEmRes)*, pages 1–7. IEEE, 2023.

David Gotz and David Borland. Data-driven healthcare: challenges and opportunities for interactive visualization. *IEEE computer graphics and applications*, 36(3):90–96, 2016.

Abba B Gumel, Enahoro A Iboi, Calistus N Ngonghala, and Elamin H Elbasha. A primer on using mathematics to understand covid-19 dynamics: Modeling, analysis and simulations. *Infectious Disease Modelling*, 6:148–168, 2021.

Mousumi Gupta, Madhab Nirola, Arpan Sharma, Prasanna Dhungel, et al. Geospatial analysis of contagious infection growth and cross-boundary transmission in non-vaccinated districts of north-east indian states during the covid-19 pandemic. *The Lancet Regional Health-Southeast Asia*, 28, 2024.

Wiktor Gustafsson, Fernanda C Dórea, Stefan Widgren, Jenny Frössling, Gema Vidal, Hyeyoung Kim, Wonhee Cha, Arianna Comin, Ivana Rodriguez Ewerlöf, and Thomas Rosendal. Data workflows and visualization in support of surveillance practice. *Frontiers in Veterinary Science*, 10:1129863, 2023a.

Wiktor Gustafsson, Fernanda C. Dórea, Stefan Widgren, Jenny Frössling, Gema Vidal, Hyeyoung Kim, Wonhee Cha, Arianna Comin, Ivana Rodriguez Ewerlöf, and Thomas Rosendal. Data workflows and visualization in support of surveillance practice. *Frontiers in Veterinary Science*, 10, February 2023b. ISSN 2297-1769. doi: 10.3389/fvets.2023.1129863. URL `https://www.frontiersin.org/journals/veterinary-science/articles/10.3389/fvets.2023.1129863/full`. Publisher: Frontiers.

Neal R Haddaway, Matthew J Grainger, and Charles T Gray. Citationchaser: A tool for transparent and efficient forward and backward citation chasing in systematic searching. *Research synthesis methods*, 13(4):533–545, 2022.

Ryan P Hafen, David E Anderson, William S Cleveland, Ross Maciejewski, David S Ebert, Ahmad Abusalah, Mohamed Yakout, Mourad Ouzzani, and Shaun J Grannis. Syndromic surveillance: Stl for modeling, visualizing, and monitoring disease counts. *BMC Medical Informatics and Decision Making*, 9:1–11, 2009.

Su Yeon Han, Joon-Seok Kim, , et al. CyberGIS-Vis for democratizing access to scalable spatiotemporal geovisual analytics: A case study of COVID-19. In *Proc. 5th ACM Int. W. Spa. Comp. Epi.*, pages 19–22, 2024.

Ni'mah Hanifah, Guardian Yoki Sanjaya, Nuryati Nuryati, Aprisa Chrysantina, Niko Tesni Saputro, and Mardiansyah Mardiansyah. Using District Health Information System (DHIS2) for Health Data Integration in Special Region of Yogyakarta. *Jurnal Pengabdian kepada Masyarakat (Indonesian Journal of Community Engagement)*, 8(1):48, March 2022. ISSN 2541-5883, 2460-9447. doi: 10.22146/jpkm.40379. URL https://jurnal.ugm.ac.id/jpkm/article/view/40379.

Sandra G Hart. Nasa-task load index (nasa-tlx); 20 years later. In *Proceedings of the human factors and ergonomics society annual meeting*, volume 50, pages 904–908. Sage publications Sage CA: Los Angeles, CA, 2006.

Chandra Hartanto et al. Multivariate sales forecast model towards trend shifting during covid-19 pandemic: A case study in global beauty industry. In *Proc. Int. Conf. Eng. & Inf. Tech. Sust. Ind.*, pages 1–7, 2022.

A. Hazra. Using the confidence interval confidently. *Journal of Thoracic Disease*, 9(10): 4125–4129, 2017. doi: 10.21037/jtd.2017.09.14.

Kelly J Henning. What is syndromic surveillance? *MMWR: Morbidity & Mortality Weekly Report*, 53, 2004.

Calvin Wai-Loon Ho. Operationalizing "One Health" as "One Digital Health" Through a Global Framework That Emphasizes Fair and Equitable Sharing of Benefits From the Use of Artificial Intelligence and Related Digital Technologies. *Frontiers in Public Health*, 10: 768977, May 2022. ISSN 2296-2565. doi: 10.3389/fpubh.2022.768977. URL https://www.frontiersin.org/articles/10.3389/fpubh.2022.768977/full.

Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural Comput.*, 9(8): 1735–1780, 1997. doi: 10.1162/NECO.1997.9.8.1735. URL https://doi.org/10.1162/neco.1997.9.8.1735.

T Déirdre Hollingsworth, Don Klinkenberg, Hans Heesterbeek, and Roy M Anderson. Mitigation strategies for pandemic influenza a: balancing conflicting policy objectives. *PLoS computational biology*, 7(2):e1001076, 2011.

Nathan Hurtig et al. Network visualization and assessment of student reasoning about conditionals. In *Inn. & Tech. Comp. Sci. Edu. Vol. 1*, page 255–261, 2022. ISBN 9781450392013.

Trung Dong Huynh, Mark Ebden, Joel Fischer, Stephen Roberts, and Luc Moreau. Provenance network analytics: an approach to data analytics using data provenance. *Data Mining and Knowledge Discovery*, 32(3):708–735, 2018.

Sana Iftikhar, S.M. Mursalin, Syeda QurratulAin, and Carrie S. McNeil. Control of Dengue Epidemic in Lahore: A fortuitous application of One Health Approach. *One Health Cases*, page ohcs20230014, May 2023. ISSN 2958-4345. doi: 10.1079/onehealthcases.2023.0014. URL `http://www.cabidigitallibrary.org/doi/10.1079/onehealthcases.2023.0014`.

Haider Ilyas et al. Analysis and visualization of Covid-19 discourse on Twitter using data science: A case study of the USA, the UK and India. *Global Knowledge, Memory & Communication*, 71(3):140–154, 2022.

Elisabeth A. Innes, Clare Hamilton, Joao L. Garcia, Andreas Chryssafidis, and David Smith. A one health approach to vaccines against Toxoplasma gondii. *Food and Waterborne Parasitology*, 15:e00053, June 2019. ISSN 24056766. doi: 10.1016/j.fawpar.2019.e00053. URL `https://linkinghub.elsevier.com/retrieve/pii/S2405676619300022`.

Y Kumar Jain and Santosh Kumar Bhandare. Min max normalization based data perturbation method for privacy protection. *International Journal of Computer & Communication Technology*, 2(8):45–50, 2011.

Ni Ketut Dewi Ari Jayanti and Luh Made Yulyantari. Interactive data visualization: open data of pandemic disease. In *3rd Int. Conf. Cybernetics & Int. Sys.*, pages 1–5, 2021.

Wolfgang Jentner et al. Visualization of covid-19 incidences and treatment capacities with coronavis. In *IEEE VIS Workshop Vis. Pandemic & Emer. Responses*, 2023.

Kate E Jones, Nikkita G Patel, Marc A Levy, Adam Storeygard, Deborah Balk, John L Gittleman, and Peter Daszak. Global trends in emerging infectious diseases. *Nature*, 451 (7181):990–993, 2008.

Samuel Kahariri, S. M. Thumbi, Bernard Bett, Marianne W. Mureithi, Nazaria Nyaga, Allan Ogendo, Mathew Muturi, and Lian Francesca Thomas. The evolution of Kenya's animal health surveillance system and its potential for efficient detection of zoonoses. *Frontiers in Veterinary Science*, 11:1379907, June 2024. ISSN 2297-1769. doi: 10.3389/fvets.2024.1379907. URL `https://www.frontiersin.org/articles/10.3389/fvets.2024.1379907/full`.

Thushara Kamalrathne, Dilanthi Amaratunga, Richard Haigh, and Lahiru Kodituwakku. Need for effective detection and early warnings for epidemic and pandemic preparedness planning in the context of multi-hazards: Lessons from the covid-19 pandemic. *International journal of disaster risk reduction*, 92:103724, 2023.

Qiao Kang, Xing Song, Xiaying Xin, Bing Chen, Yuanzhu Chen, Xudong Ye, and Baiyu Zhang. Machine learning-aided causal inference framework for environmental data analysis: a covid-19 case study. *Environmental Science & Technology*, 55(19):13400–13410, 2021.

I Kassim, C Arinze, D Tom-Aba, O Adeoye, C Ihekweazu, T D McHugh, I Abubakar, G Krause, S Mwakasungula, H Masanja, and R W Aldridge. Mobile-based and open-source infectious disease surveillance and outbreak management in Tanzania. *European*

Journal of Public Health*, 30(Supplement_5):ckaa166.1347, September 2020. ISSN 1101-1262, 1464-360X. doi: 10.1093/eurpub/ckaa166.1347. URL `https://academic.oup.com/eurpub/article/doi/10.1093/eurpub/ckaa166.1347/5915340`.

Jasleen Kaur, Jasmine Kaur, Ajay Singh Dhama, Shelja Jindal, Kamini Walia, and Harpreet Singh. Strengthening the Surveillance of Antimicrobial Resistance in India Using Integrative Technologies. *Frontiers in Public Health*, 10:861888, April 2022. ISSN 2296-2565. doi: 10.3389/fpubh.2022.861888. URL `https://www.frontiersin.org/articles/10.3389/fpubh.2022.861888/full`.

Nami Kawakyu, Allison Osterman, Jimi Michel, Dominic Mutai, Edith Jepleting, Grace Njenga, Parysa Oskouipour, and Jessica C Shearer. Developing and refining the COVID-19 to Routine Immunization Information System Transferability Assessment (CRIISTA) tool: a decision support tool to leverage COVID-19 immunization information system investments for routine immunization. *Oxford Open Digital Health*, 2(Supplement_1):i75–i85, May 2024. ISSN 2754-4591. doi: 10.1093/oodh/oqae006. URL `https://academic.oup.com/oodh/article/2/Supplement_1/i75/7663966`.

Daniel A Keim, Florian Mansmann, Jörn Schneidewind, Jim Thomas, and Hartmut Ziegler. Visual analytics: Scope and challenges. In *Visual data mining: Theory, techniques and tools for visual analytics*, pages 76–90. Springer, 2008.

Alexander C Keyel, Morgan E Gorris, Ilia Rochlin, Johnny A Uelmen, Luis F Chaves, Gabriel L Hamer, Imelda K Moise, Marta Shocket, A Marm Kilpatrick, Nicholas B DeFelice, et al. A proposed framework for the development and qualitative evaluation of west nile virus models and their application to local public health decision-making. *PLoS neglected tropical diseases*, 15(9):e0009653, 2021.

Saiful Khan, Scott Jones, et al. Feature-action design patterns for storytelling visualizations with time series data. *arXiv:2402.03116*, 2024.

Saiful Khan et al. Rapid Development of a Data Visualization Service in an Emergency Response. *IEEE Trans. Services Comp.*, 15(2):1251–1264, 2022.

Nurul Kharismawati Kamaludin et al. Analysis strategy of tourism development during the pandemic in Yogyakarta. In *Proc. Conf. Inf. Comm. & Management*, pages 142–148, 2021.

Denisse Kim, Bernardo Cánovas-Segura, Manuel Campos, and Jose M Juarez. Visualization of spatial–temporal epidemiological data: A scoping review. *Technologies*, 12(3):31, 2024.

Sungahn Ko, Jieqiong Zhao, et al. Vasa: Interactive computational steering of large asynchronous simulation pipelines for societal infrastructure. *IEEE Trans. Vis. Comp. Graph.*, 20(12):1853–1862, 2014.

Márton Kolossváry, Chris deFilippi, et al. Identification of pre-infection markers and differential plasma protein expression following SARS-CoV-2 infection in people living with HIV. *EBioMedicine*, 90, 2023.

Iurii Krak, Olexander Barmak, and Eduard Manziuk. Using visual analytics to develop human and machine-centric models: A review of approaches and proposed information technology. *Computational Intelligence*, 38(3):921–946, 2022.

Ajinkya Kunjir et al. Data mining and visualization for prediction of multiple diseases in healthcare. In *Conf. Big Data Analy. & Comp. Int.*, pages 329–334, 2017.

Arif Kurniawan and Fachrul Kurniawan. Time series forecasting for the spread of covid-19 in Indonesia using curve fitting. In *3rd East Indonesia Conf. Computer & Inf. Technology*, pages 45–48, 2021.

Steven Lam, Vivian Hoffmann, Bernard Bett, Eric M. Fèvre, Arshnee Moodley, Chadag Vishnumurthy Mohan, Javier Mateo-Sagasta, and Hung Nguyen-Viet. Navigating One Health in research-for-development: Reflections on the design and implementation of the CGIAR Initiative on One Health. *One Health*, 18:100710, June 2024. ISSN 23527714. doi: 10.1016/j.onehlt.2024.100710. URL `https://linkinghub.elsevier.com/retrieve/pii/S2352771424000363`.

Andrew Larkins, Soulasack Vannamahaxay, Vannaphone Puttana, Malavanh Chittavong, Fongsamouth Southammavong, Mayfong Mayxay, Davina Boyd, Mieghan Bruce, and Amanda Ash. Scaling up one health: A network analysis in lao PDR. 18:100661, 2024. ISSN 23527714. doi: 10.1016/j.onehlt.2023.100661. URL `https://linkinghub.elsevier.com/retrieve/pii/S2352771423001817`.

Addepalli Lavanya et al. A real-time human mobility visualization of covid-19 spread from east asian countries. In *Conf. Soc. Net. Analy., Man. & Sec.*, pages 1–8, 2021.

Darwin A León-Figueroa, D Katterine Bonilla-Aldana, Monica Pachar, Luccio Romaní, Hortencia M Saldaña-Cumpa, Claudia Anchay-Zuloeta, Milagros Diaz-Torres, Carlos Franco-Paredes, José Antonio Suárez, Juan David Ramirez, et al. The never-ending global emergence of viral zoonoses after covid-19? the rising concern of monkeypox in europe, north america and beyond. *Travel medicine and infectious disease*, 49:102362, 2022.

He Li, Xiao-Long Xu, Da-Wei Dai, Zhen-Yu Huang, Zhuang Ma, and Yan-Jun Guan. Air pollution and temperature are associated with increased covid-19 incidence: a time series study. *International journal of infectious diseases*, 97:278–282, 2020.

M Lipsitch. Will covid-19 go away on its own in warmer weather. *Center for Communicable Disease Dynamics (CCDD) at the Harvard TH Chan School of Public Health*, 2020.

Candice Louw, Rafael Paffenholz, Carolin Verset, and Gérard Krause. Global Good Open Source Software Development in Response to the COVID-19 Pandemic – Perspectives from SORMAS Implementation in Europe. In Brigitte Séroussi, Patrick Weber, Ferdinand Dhombres, Cyril Grouin, Jan-David Liebe, Sylvia Pelayo, Andrea Pinna, Bastien Rance, Lucia Sacchi, Adrien Ugon, Arriel Benis, and Parisis Gallos, editors, *Studies in Health Technology and Informatics*. IOS Press, May 2022. ISBN 978-1-64368-284-6 978-1-64368-285-3. doi: 10.3233/SHTI220553. URL `https://ebooks.iospress.nl/doi/10.3233/SHTI220553`.

Dan Luo, Yadong Wu, Kai Liu, and Weihan Zhang. Visual analysis of epidemic epidemiological investigation track data.

Kennedy Lushasi, Rachel Steenson, Jubilate Bernard, Joel Jackson Changalucha, Nicodem James Govella, Daniel T. Haydon, Husna Hoffu, Felix Lankester, Frank Magoti, Emmanuel Abraham Mpolya, Zacharia Mtema, Hesron Nonga, and Katie Hampson. One Health in Practice: Using Integrated Bite Case Management to Increase Detection of Rabid Animals in Tanzania. *Frontiers in Public Health*, 8:13, February 2020. ISSN 2296-2565. doi: 10.3389/fpubh.2020.00013. URL `https://www.frontiersin.org/article/10.3389/fpubh.2020.00013/full`.

Ross Maciejewski, Philip Livengood, Stephen Rudolph, Timothy F. Collins, David S. Ebert, Robert T. Brigantic, Courtney D. Corley, George A. Muller, and Stephen W. Sanders. A pandemic influenza modeling and visualization tool. *J. Vis. Lang. Comput.*, 22(4):268–278, 2011. doi: 10.1016/J.JVLC.2011.04.002. URL `https://doi.org/10.1016/j.jvlc.2011.04.002`.

Jock Mackinlay et al. Show Me: Automatic presentation for visual analysis. *IEEE Trans. Vis. Comput. Graph.*, 13(6):1137–1144, 2007.

Eli MacLean et al. Web-based 3d visualization system for anatomy online instruction. In *Proc. ACM Southeast Conf.*, page 194–198, 2021. ISBN 9781450380683. URL `https://doi.org/10.1145/3409334.3452080`.

Julio Jerison E Macrohon and Jyh-Horng Jeng. A real-time COVID-19 data visualization and information repository in the Philippines. In *9th Int. Conf. Information & Education Technology (ICIET)*, pages 443–447, 2021.

Karthic Madanagopal, Eric D Ragan, and Perakath Benjamin. Analytic provenance in practice: The role of provenance in real-world visualization and data analysis environments. *IEEE Computer Graphics and Applications*, 39(6):30–45, 2019.

Stephen Makonin, Daniel McVeigh, Wolfgang Stuerzlinger, Khoa Tran, and Fred Popowich. Mixed-initiative for big data: The intersection of human+ visual analytics+ prediction. In *2016 49th Hawaii international conference on system sciences (HICSS)*, pages 1427–1436. IEEE, 2016.

Aileen M. Marty, Christian K. Beÿ, and Kristi L. Koenig. 2024 mpox outbreak: A rapidly evolving public health emergency of international concern: Introduction of an updated mpox identify-isolate-inform (3i) tool. 19:100927, 2024. ISSN 23527714. doi: 10.1016/j.onehlt.2024.100927. URL `https://linkinghub.elsevier.com/retrieve/pii/S2352771424002532`.

Joseph Matere, Piers Simpkin, Jay Angerer, Emmanuella Olesambu, Selvaraju Ramasamy, and Folorunso Fasina. Predictive livestock early warning system (plews): Monitoring forage condition and implications for animal production in kenya. *Weather and Climate Extremes*, 27:100209, 2020a.

Joseph Matere, Piers Simpkin, Jay Angerer, Emmanuella Olesambu, Selvaraju Ramasamy, and Folorunso Fasina. Predictive Livestock Early Warning System (PLEWS): Monitoring forage condition and implications for animal production in Kenya. *Weather and Climate Extremes*, 27:100209, March 2020b. ISSN 22120947. doi: 10.1016/j.wace.2019.100209. URL `https://linkinghub.elsevier.com/retrieve/pii/S2212094718300343`.

Amaryllis Mavragani and Konstantinos Gkillas. Covid-19 predictability in the united states using google trends time series. *Scientific reports*, 10(1):20693, 2020.

Benedikt Mayer et al. A characterization of interactive visual data stories with a spatio-temporal context. *Comp. Graph. Forum*, 42(6):e14922, 2023.

Elliot Mbunge et al. A critical review of emerging technologies for tackling covid-19 pandemic. *Human Beh. & Emerging Tech.*, 3(1):25–39, 2021.

Paul Mee et al. Tracking the emergence of disparities in the subnational spread of COVID-19 in Brazil using an online application for real-time data visualisation: A longitudinal analysis. *The Lancet Regional Health–Americas*, 5, 2022.

Aline Menin, Ricardo Cava, Carla Maria Dal Sasso Freitas, Olivier Corby, and Marco Winckler. Towards a visual approach for representing analytical provenance in exploration processes. In *2021 25th International Conference Information Visualisation (IV)*, pages 21–28. IEEE, 2021.

Christopher L. Merkord, Yi Liu, Abere Mihretie, Teklehaymanot Gebrehiwot, Worku Awoke, Estifanos Bayabil, Geoffrey M. Henebry, Gebeyaw T. Kassa, Mastewal Lake, and Michael C. Wimberly. Integrating malaria surveillance with climate data for outbreak detection and forecasting: the EPIDEMIA system. 16(1):89, 2017. ISSN 1475-2875. doi: 10.1186/s12936-017-1735-x. URL `https://malariajournal.biomedcentral.com/articles/10.1186/s12936-017-1735-x`.

Miriah Meyer, Michael Sedlmair, P. Samuel Quinan, and Tamara Munzner. The nested blocks and guidelines model. *J. Inf. Visualization*, 14(3):234–249, 2015.

Panagiotis D Michailidis. Visualizing social media research in the age of covid-19. *Information*, 13(8):372, 2022.

Silvia Miksch and Wolfgang Aigner. A matter of time: Applying a data–users–tasks design triangle to visual analytics of time-oriented data. *Computers & Graphics*, 38:286–290, 2014. ISSN 0097-8493. URL `https://www.sciencedirect.com/science/article/pii/S0097849313001817`.

Sitao Min et al. Crowdmap: Spatiotemporal visualization of anonymous occupancy data for pandemic response. In *Proc. Int. Conf. Adv. Geo. Inf. Syst.*, pages 630–633, 2021.

Dominik Moritz, Chenglong Wang, Greg L. Nelson, et al. Formalizing visualization design knowledge as constraints: Actionable and extensible models in Draco. *IEEE Trans. Vis. Comput. Graph.*, 25(1):438–448, 2019.

Samira Mubareka, John Amuasi, Arinjay Banerjee, Hélène Carabin, Joe Copper Jack, Claire Jardine, Bogdan Jaroszewicz, Greg Keefe, Jonathon Kotwa, Susan Kutz, et al. Strengthening a one health approach to emerging zoonoses, 2023.

Riya Mukherjee, Jasmina Vidic, Sandrine Auger, Hsiao-Chuan Wen, Ramendra Pati Pandey, and Chung-Ming Chang. Exploring disease management and control through pathogen diagnostics and one health initiative: A concise review. *Antibiotics*, 13(1):17, 2023.

Monica F Myers, DJ Rogers, J Cox, Antoine Flahault, and Simon I Hay. Forecasting disease risk for increased epidemic preparedness in public health. *Advances in parasitology*, 47: 309–330, 2000.

National Institutes of Health. Nih radx initiative expands covid-19 testing innovation: Additional types of rapid tests, 10 2021. URL `https://www.nih.gov/news-events/news-releases/nih-radx-initiative-expands-covid-19-testing-innovation-additional-types-rapid-tests`. Accessed: 2025-08-12.

Duduzile Ndwandwe and Charles S Wiysonge. Covid-19 vaccines. *Current opinion in immunology*, 71:111–116, 2021.

Dawn M Nekorchuk, Anita Bharadwaja, Sean Simonson, Emma Ortega, Caio MB França, Emily Dinh, Rebecca Reik, Rachel Burkholder, and Michael C Wimberly. The arbovirus mapping and prediction (arbomap) system for west nile virus forecasting. *JAMIA open*, 7(1):ooad110, 2024.

Aitana Neves, Daniel Walther, Trinidad Martin-Campos, Valerie Barbie, Claire Bertelli, Dominique Blanc, Gérard Bouchet, Frédéric Erard, Gilbert Greub, Hans H. Hirsch, Michael Huber, Laurent Kaiser, Stephen L. Leib, Karoline Leuzinger, Vladimir Lazarevic, Mirjam Mäusezahl, Jorge Molina, Richard A. Neher, Vincent Perreten, Alban Ramette, Tim Roloff, Jacques Schrenzel, Helena M. B. Seth-Smith, Roger Stephan, Dillenn Terumalai, Fanny Wegner, and Adrian Egli. The Swiss Pathogen Surveillance Platform – towards a nation-wide One Health data exchange platform for bacterial, viral and fungal genomics and associated metadata. *Microbial Genomics*, 9(5), May 2023. ISSN 2057-5858. doi: 10.1099/mgen.0.001001. URL `https://www.microbiologyresearch.org/content/journal/mgen/10.1099/mgen.0.001001`.

Chris North, Remco Chang, Alex Endert, Wenwen Dou, Richard May, Bill Pike, and Glenn Fink. Analytic provenance: process+ interaction+ insight. In *CHI'11 Extended Abstracts on Human Factors in Computing Systems*, pages 33–36. 2011.

Thierry Nyatanyi, Michael Wilkes, Haley McDermott, Serge Nzietchueng, Isidore Gafarasi, Antoine Mudakikwa, Jean Felix Kinani, Joseph Rukelibuga, Jared Omolo, Denise Mupfasoni, Adeline Kabeja, Jose Nyamusore, Julius Nziza, Jean Leonard Hakizimana, Julius Kamugisha, Richard Nkunda, Robert Kibuuka, Etienne Rugigana, Paul Farmer, Philip Cotton, and Agnes Binagwaho. Implementing One Health as an integrated approach to health in Rwanda. *BMJ Global Health*, 2(1):e000121, February 2017. ISSN 2059-7908. doi: 10.1136/bmjgh-2016-000121. URL `https://gh.bmj.com/lookup/doi/10.1136/bmjgh-2016-000121`.

Evan O'Brien and Irene Xagoraraki. Understanding temporal and spatial variations of viral disease in the US: The need for a one-health-based data collection and analysis approach. *One Health*, 8:100105, December 2019. ISSN 2352-7714. doi: 10.1016/j.onehlt.2019.100105. URL https://www.sciencedirect.com/science/article/pii/S2352771419300588.

Taylor M Oeschger, Duncan S McCloskey, Rose M Buchmann, Aakash M Choubal, Juan M Boza, Saurabh Mehta, and David Erickson. Early warning diagnostics for emerging infectious diseases in developing into late-stage pandemics. *Accounts of Chemical Research*, 54 (19):3656–3666, 2021.

Arizona Department of Health Services. Respiratory illness summary dashboard. https://www.azdhs.gov/preparedness/epidemiology-disease-control/infectious-disease-epidemiology/respiratory-illness/dashboards/index.php#respiratory-summary. Accessed: 9/8/2025.

Masataka Okabe and Kei Ito. Color universal design (cud)-how to make figures and presentations that are friendly to colorblind people. *Retrieved April*, 21:2017, 2008.

Wellington Oliveira, Daniel De Oliveira, and Vanessa Braganholo. Provenance analytics for workflow-based computational experiments: A survey. *ACM Computing Surveys (CSUR)*, 51(3):1–25, 2018.

Babasola Oluseyi Olugasa and Ayotunde James Fasunla. One health data audit: a spatio-temporal approach to cultivating sustainable multi-disciplinary collaboration and communication in zoonoses surveillance, control and stepwise elimination. *PAMJ-One Health*, 1(5), April 2020. ISSN 2707-2800. doi: 10.11604/pamj-oh.2020.1.5.22350. URL https://www.one-health.panafrican-med-journal.com/content/article/1/5/full. Number: 5.

Verónica Ormea, Roberto Evaristo, Salomon Durand, Andrew Lover, and Simon R. Rüegg. COVID-19 exposes weaknesses in public health in the Peruvian Amazon and highlights opportunities for a One Health approach. *CABI One Health*, page ohcs20230006, April 2023. ISSN 2791-223X. doi: 10.1079/cabionehealth.2023.0006. URL http://www.cabidigitallibrary.org/doi/10.1079/cabionehealth.2023.0006.

Abdinasir Yusuf Osman, Halima Mohamed, Farah I. Mumin, Heba Mahrous, Asma Saidouni, Sharifo Ali Elmi, Amira Khalif Adawe, Abdikani Abdullahi Mo'allim, Mutaawe Lubogo, Sk Md Mamunur Rahman Malik, Athman Mwatondo, Tajudeen Raji, Abdifatah Dirie Ahmed, Alimuddin Zumla, Osman Dar, Richard Kock, and Siobhan M. Mor. Prioritization of zoonoses for multisectoral, One Health collaboration in Somalia, 2023. *One Health*, 17: 100634, December 2023. ISSN 23527714. doi: 10.1016/j.onehlt.2023.100634. URL https://linkinghub.elsevier.com/retrieve/pii/S2352771423001544.

Lace Padilla et al. Multiple forecast visualizations (mfvs): Trade-offs in trust and performance in multiple covid-19 forecast visualizations. *IEEE Trans. Vis. Comput. Graph.*, 29 (1):12–22, 2022.

Matthew J Page, Joanne E McKenzie, Patrick M Bossuyt, Isabelle Boutron, Tammy C Hoffmann, Cynthia D Mulrow, Larissa Shamseer, Jennifer M Tetzlaff, Elie A Akl, Sue E Brennan, et al. The prisma 2020 statement: an updated guideline for reporting systematic reviews. *bmj*, 372, 2021a.

Matthew J Page et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, 372, 2021b. URL https://www.bmj.com/content/372/bmj.n71.

PAIR-code. COVID-19 Symptom Dataset Map. https://pair-code.github.io/covid19_symptom_dataset/map/?date=2020-10-26, 2020. Accessed on September 12, 2023.

Daniele Pala et al. A new interactive tool to visualize and analyze covid-19 data: The periscope atlas. *Int. J. Env. Res. & Pub. Hea.*, 19(15):9136, 2022.

A David Paltiel, Amy Zheng, and Rochelle P Walensky. Assessment of sars-cov-2 screening strategies to permit the safe reopening of college campuses in the united states. *JAMA network open*, 3(7):e2016818–e2016818, 2020.

Chalermchai Panyadee, Adisorn Kuntharose, et al. Data visualization platform and tourism resilience in the post COVID-19 pandemic in Chiang Mai, Thailand. In *J. Int. Conf. Dig. Arts, Media & Tech. ECTI N. Sec. Conf. Elec., Elect., Comp. and Tele. Eng.*, pages 215–218, 2023.

Konstantinos Papageorgiadis, Konstantinos Georgiou, et al. COVID-Vis: Visualizing knowledge exchange on scientific software development in the COVID-19 era. In *Proc. 26th Pan-Hellenic Conf. Informatics*, pages 367–372, 2022.

Jordan Peccia, Alessandro Zulli, Doug E Brackney, Nathan D Grubaugh, Edward H Kaplan, Arnau Casanovas-Massana, Albert I Ko, Amyn A Malik, Dennis Wang, Mike Wang, et al. Measurement of sars-cov-2 rna in wastewater tracks community infection dynamics. *Nature biotechnology*, 38(10):1164–1167, 2020.

Ana Soledad Pedretti, Santiago Márquez Fosser, Roberto José Mercau Cossio, Jorge Ariel Esteban, Paz Rodriguez, Bernardo Julio Martinez, Santiago Andres Frid, Daniel Luna, Adriana Iannicelli, Fernando Plazzotta, and Maria Florencia Grande Ratti. Evaluation of Telehealth Service for COVID-19 Outpatients: A Dashboard to Measure Healthcare Quality and Safety. In Paula Otero, Philip Scott, Susan Z. Martin, and Elaine Huesing, editors, *Studies in Health Technology and Informatics*. IOS Press, June 2022. ISBN 978-1-64368-264-8 978-1-64368-265-5. doi: 10.3233/SHTI220098. URL https://ebooks.iospress.nl/doi/10.3233/SHTI220098.

Darío Pérez-Campuzano, Luis Rubio Andrada, et al. Visualizing the historical covid-19 shock in the us airline industry: A data mining approach for dynamic market surveillance. *J. Air Trans. Man.*, 101:102194, 2022.

Jennifer Petkovic, Olivia Magwood, Lyubov Lytvyn, Joanne Khabsa, Thomas W Concannon, Vivian Welch, Alex Todhunter-Brown, Marisha E Palm, Elie A Akl, Lawrence Mbuagbaw, et al. Key issues for stakeholder engagement in the development of health and healthcare guidelines. *Research Involvement and Engagement*, 9(1):27, 2023.

Sarbeswar Praharaj, Patricia Solis, and Elizabeth A Wentz. Deploying geospatial visualization dashboards to combat the socioeconomic impacts of Covid-19. *Env. & Plan. B: Urban Analytics & City Science*, 50(5):1262–1279, 2023.

Kevin Queenan, Julie Garnier, L Rosenbaum Nielsen, Sandra Buttigieg, D de Meneghi, Martin Holmberg, Jakob Zinsstag, Simon Rüegg, Barbara Häsler, and Richard Kock. Roadmap to a one health agenda 2030. *CABI Reviews*, (2017):1–17, 2017.

Muthusami Ra et al. COVID-19 outbreak: Tweet based analysis and visualization towards the influence of coronavirus in the world. *Gedrag. Organ. Rev*, 33:8–9, 2020.

Eric D Ragan, Alex Endert, Jibonananda Sanyal, and Jian Chen. Characterizing provenance in visualization and data analysis: an organizational framework of provenance types and purposes. *IEEE transactions on visualization and computer graphics*, 22(1):31–40, 2015.

Iman Rahimi, Fang Chen, and Amir H. Gandomi. A review on COVID-19 forecasting models. *Neural Comput. Appl.*, 35(33):23671–23681, 2023.

Raissa R Raineri, Chelsey B Shivley, Paul J Plummer, and Amanda J Kreuder. Assessing stakeholders' opinions on sharing antimicrobial susceptibility testing data from animals is key to developing a centralized database and dashboard tool. *Journal of the American Veterinary Medical Association*, 262(2):216–225, 2024.

CLEVELAND RB. Stl: A seasonal-trend decomposition procedure based on loess. *J Off Stat*, 6:3–73, 1990.

Audrey Reinert, Luke S Snyder, Jieqiong Zhao, Andrew S Fox, Dean F Hougen, Charles Nicholson, and David S Ebert. Visual analytics for decision-making during pandemics. *Computing in Science & Engineering*, 22(6):48–59, 2020a.

Audrey Reinert, Luke S. Snyder, et al. Visual analytics for decision-making during pandemics. *Comput. Sci. Eng.*, 22(6):48–59, 2020b.

Alexander Rind, Taowei David Wang, Wolfgang Aigner, Silvia Miksch, Krist Wongsuphasawat, Catherine Plaisant, Ben Shneiderman, et al. Interactive information visualization to explore and query electronic health records. *Foundations and Trends® in Human–Computer Interaction*, 5(3):207–298, 2013.

Cassidy Logan Rist, Carmen Sofia Arriola, and Carol Rubin. Prioritizing zoonoses: a proposed one health tool for collaborative decision-making. *PloS one*, 9(10):e109986, 2015.

Rongxiang Rui, Maozai Tian, Man-Lai Tang, George To-Sum Ho, and Chun-Ho Wu. Analysis of the spread of covid-19 in the usa with a spatio-temporal multivariate time series model. *International Journal of Environmental Research and Public Health*, 18(2):774, 2021.

Gabriel Ryan et al. At a glance: Pixel approximate entropy as a measure of line chart complexity. *IEEE Trans. Vis. Comp. Graph.*, 25(1):872–881, 2018.

Erik Rydow, Rita Borgo, Hui Fang, Thomas Torsney-Weir, Ben Swallow, Thibaud Porphyre, Cagatay Turkay, and Min Chen. Development and evaluation of two approaches of visual sensitivity analysis to support epidemiological modeling. *IEEE Transactions on Visualization and Computer Graphics*, 29(1):1255–1265, 2022.

Simon R. Rüegg, Liza Rosenbaum Nielsen, Sandra C. Buttigieg, Mijalche Santa, Maurizio Aragrande, Massimo Canali, Timothy Ehlinger, Ilias Chantziaras, Elena Boriani, Miroslav Radeski, Mieghan Bruce, Kevin Queenan, and Barbara Häsler. A Systems Approach to Evaluate One Health Initiatives. *Frontiers in Veterinary Science*, 5:23, March 2018. ISSN 2297-1769. doi: 10.3389/fvets.2018.00023. URL `http://journal.frontiersin.org/article/10.3389/fvets.2018.00023/full`.

Dominik Sacha, Andreas Stoffel, Florian Stoffel, Bum Chul Kwon, Geoffrey Ellis, and Daniel A Keim. Knowledge generation model for visual analytics. *IEEE transactions on visualization and computer graphics*, 20(12):1604–1613, 2014.

Debashis Sahoo, Gajanan D Katkar, Soni Khandelwal, et al. AI-guided discovery of the invariant host response to viral pandemics. *EBioMedicine*, 68, 2021.

Hanan Samet, Yunheng Han, John Kastner, and Hong Wei. Using animation to visualize spatio-temporal varying COVID-19 data. In *Proc. ACM SIGSPATIAL Int. W. Model. & Understanding Spread COVID-19*, pages 53–62, 2020.

Ovi Sarkar, Md Faysal Ahamed, and Pallab Chowdhury. Forecasting & severity analysis of covid-19 using machine learning approach with advanced data visualization. In *Int. Conf. Comp. & Inf. Tech.*, pages 1–6, 2020.

Manjeet Sharan, Deepthi Vijay, Jay Prakash Yadav, Jasbir Singh Bedi, and Pankaj Dhaka. Surveillance and response strategies for zoonotic diseases: a comprehensive review. *Science in One Health*, 2:100050, 2023. ISSN 29497043. doi: 10.1016/j.soh.2023.100050. URL `https://linkinghub.elsevier.com/retrieve/pii/S2949704323000446`.

Garima Sharma et al. A scalable architecture for rapid development of trackers amid pandemics like COVID-19. In *Int. Conf. Comp. Int. & Sus. Eng. Sol.*, pages 592–598, 2023.

Sharad Sharma et al. Real-time data visualization to enhance situational awareness of covid pandemic. In *Int. Conf. Comp. Sci. & Comp. Int.*, pages 352–357, 2020.

Yang Shi et al. Breaking the fourth wall of data stories through interaction. *IEEE Trans. Vis. Comput. Graph.*, 29(1):972–982, 2022.

Gitanjali Rahul Shinde, Asmita B. Kalamkar, Parikshit N. Mahalle, Nilanjan Dey, Jyotismita Chaki, and Aboul Ella Hassanien. Forecasting models for coronavirus disease (COVID-19): A survey of the state-of-the-art. *SN Comput. Sci.*, 1(4):197, 2020.

Galit Shmueli and Otto R Koppius. Predictive analytics in information systems research. *MIS quarterly*, pages 553–572, 2011.

Nistha Shrestha, Muhammad Yousaf Shad, Osman Ulvi, Modasser Hossain Khan, Ajlina Karamehic-Muratovic, Uyen-Sa DT Nguyen, Mahdi Baghbanzadeh, Robert Wardrup, Nasrin Aghamohammadi, Diana Cervantes, et al. The impact of covid-19 on globalization. *One Health*, 11:100180, 2020.

Bernard C Silenou, Carolin Verset, Basil B Kaburi, Olivier Leuci, Stéphane Ghozzi, Cédric Duboudin, and Gérard Krause. A Novel Tool for Real-time Estimation of Epidemiological Parameters of Communicable Diseases Using Contact-Tracing Data: Development and Deployment. *JMIR Public Health and Surveillance*, 8(5):e34438, May 2022. ISSN 2369-2960. doi: 10.2196/34438. URL `https://publichealth.jmir.org/2022/5/e34438`.

Julie R Sinclair. Importance of a one health approach in advancing global health security and the sustainable development goals. *Rev Sci Tech*, 38(1):145–154, 2019.

Samradhi Singh, Poonam Sharma, Namrata Pal, Devojit Kumar Sarma, Rajnarayan Tiwari, and Manoj Kumar. Holistic one health surveillance framework: synergizing environmental, animal, and human determinants for enhanced infectious disease management. *ACS Infectious Diseases*, 10(3):808–826, 2024.

Anshu Singla and Vandana Pushe. Text and audio framework for an effective mathematical symbols' visualization for blind students. In *Proc. 4th Int. Conf. Inf. Management & Machine Int.*, 2023. ISBN 9781450399937. URL `https://doi.org/10.1145/3590837.3590933`.

Bahrad A Sokhansanj and Gail L Rosen. Predicting covid-19 disease severity from sars-cov-2 spike protein sequence by mixed effects machine learning. *Computers in biology and medicine*, 149:105969, 2022.

Xuanyi Song, Yuetian Liu, Liang Xue, Jun Wang, Jingzhe Zhang, Junqiang Wang, Long Jiang, and Ziyan Cheng. Time-series well performance prediction based on long short-term memory (lstm) neural network model. *Journal of Petroleum Science and Engineering*, 186: 106682, 2020.

SORMAS Foundation. SORMAS: Surveillance Outbreak Response Management and Analysis System, 2024. URL `https://sormas.org/`. Accessed: 2024-10-07.

Sanjana Srabanti et al. Covid-19 ensemblevis: Visual analysis of county-level ensemble forecast models. In *IEEE Vis. Anal. Health.*, pages 1–5, 2021.

Charles D Stolper, Adam Perer, and David Gotz. Progressive visual analytics: User-driven visual exploration of in-progress analytics. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1653–1662, 2014.

Johannes Stübinger and Katharina Adler. How to identify varying lead–lag effects in time series data: Implementation, validation, and application of the generalized causality algorithm. *Algorithms*, 13(4):95, 2020.

Zhi-Shan Sun, Er-Ya Wan, Yannick Luther Agbana, Han-Qing Zhao, Jing-Xian Yin, Tian-Ge Jiang, Qin Li, Si-Wei Fei, Logan Blair Wu, Xin-Chen Li, Qi-Yu Zhang, Jing-Shu Liu, Si-Yu Gu, Zhao-Yu Guo, Jing-Bo Xue, Le-Fei Han, Xiao-Xi Zhang, Shang Xia, Susan Christina Welburn, Moses Okpeku, Xiao-Kui Guo, Zhao-Jun Wang, Jun-Hu Chen, Yi Zhang, Xiao-Nong Zhou, and Kokouvi Kassegne. Global One Health index for zoonoses: A performance assessment in 160 countries and territories. *iScience*, 27(4), April 2024. ISSN 2589-0042. doi: 10.1016/j.isci.2024.109297. URL `https://www.cell.com/iscience/abstract/S2589-0042(24)00518-2`. Publisher: Elsevier.

Aravind Swaminathan et al. Analysis and prognosis of covid-19 using machine learning algorithms and visualization using tableau. In *7th Int. Conf. Comm. and Elec. Sys.*, pages 1096–1103, 2022.

Raphael Tackx et al. Quantified us: a group-in-the-loop approach to team network reconstruction. In *Adj. Proc. ACM Int. J. Conf. Per. & Ubi. Comp. & Proc. ACM Int. Symp. Wear. Comp.*, pages 502–507, 2021.

Yaksh Talavia, Priyanka Singh, and Sathiamoorthy Manoharan. Covid-19 tracker: A data visualization tool for time series data of pandemic in India. In *IEEE Asia-Pacific Conf. Comp. Sci. & Data Eng.*, pages 1–5, 2021.

Alwyn Tan, Mo Salman, Bruce Wagner, and Brian McCluskey. The Role of Animal Health Components in a Biosurveillance System: Concept and Demonstration. *Agriculture*, 13(2): 457, February 2023. ISSN 2077-0472. doi: 10.3390/agriculture13020457. URL `https://www.mdpi.com/2077-0472/13/2/457`. Number: 2 Publisher: Multidisciplinary Digital Publishing Institute.

Grace Tao, Yuan Miao, and Sebastian Ng. Covid-19 topic modeling and visualization. In *24th Int. Conf. Inf. Vis. (IV)*, pages 734–739, 2020.

A Kenigsberg Tat'Yana, Anne M Hause, et al. Dashboard development for near real-time visualization of COVID-19 vaccine safety surveillance data in the vaccine safety datalink. *Vaccine*, 40(22):3064–3071, 2022.

Henok Ayalew Tegegne, Carlijn Bogaardt, Lucie Collineau, Géraldine Cazeau, Renaud Lailler, Johana Reinhardt, Frederick T. A. Freeth, Emma Taylor, Joaquin M. Prada, and Viviane Hénaux. OH-EpiCap: a semi-quantitative tool for the evaluation of One Health epidemiological surveillance capacities and capabilities. *Frontiers in Public Health*, 11, May 2023. ISSN 2296-2565. doi: 10.3389/fpubh.2023.1053986. URL `https://www.frontiersin.org/journals/public-health/articles/10.3389/fpubh.2023.1053986/full`. Publisher: Frontiers.

Swapna Thorve, Mandy L Wilson, Bryan L Lewis, Samarth Swarup, Anil Kumar S Vullikanti, and Madhav V Marathe. Epiviewer: an epidemiological application for exploring time series data. *BMC bioinformatics*, 19:1–10, 2018.

S. M. Thumbi, M. Kariuki Njenga, Thomas L. Marsh, Susan Noh, Elkanah Otiang, Peninah Munyua, Linus Ochieng, Eric Ogola, Jonathan Yoder, Allan Audi, Joel M. Montgomery, Godfrey Bigogo, Robert F. Breiman, Guy H. Palmer, and Terry F. McElwain. Linking human health and livestock health: A "one-health" platform for integrated analysis of human health, livestock health, and economic welfare in livestock dependent communities. 10(3):e0120761, 2015. ISSN 1932-6203. doi: 10.1371/journal.pone.0120761. URL https://dx.plos.org/10.1371/journal.pone.0120761.

Ruth E. Timme, William J. Wolfgang, Maria Balkey, Sai Laxmi Gubbala Venkata, Robyn Randolph, Marc Allard, and Errol Strain. Optimizing open data to support one health: best practices to ensure interoperability of genomic data from bacterial pathogens. *One Health Outlook*, 2(1):20, October 2020. ISSN 2524-4655. doi: 10.1186/s42522-020-00026-3. URL https://doi.org/10.1186/s42522-020-00026-3.

Murray Turoff, Starr Roxanne Hiltz, Connie White, Linda Plotnick, Art Hendela, and Xiang Yoa. The past as the future of emergency preparedness and management. *International Journal of Information Systems for Crisis Response and Management (IJISCRAM)*, 1(1): 12–28, 2009.

Martina Udovičić, Ksenija Baždarić, Lidija Bilić-Zulle, Mladen Petrovečki, et al. What we need to know when calculating the coefficient of correlation? *Biochemia Medica*, 17(1): 10–15, 2007.

UK Research and Innovation. The story behind the oxford–astrazeneca covid-19 vaccine success, 2021. URL https://www.ukri.org/who-we-are/how-we-are-doing/research-outcomes-and-impact/mrc/the-story-behind-the-oxford-astrazeneca-covid-19-vaccine-success/. Accessed: 2025-08-12.

Muhammad Usman et al. A multiscale geospatial dataset and an interactive visualization dashboard for computational epidemiology and open scientific research. *IEEE Comp. Graph. & App.*, 43(1):39–52, 2022.

Rashmi Vashisth et al. Visualization of covid-19 pandemic data: an analysis. In *3rd Int. Conf. Comp., Auto. and Know. Man.*, pages 1–6, 2022.

Heidi M. Vesterinen, Tracey V. Dutcher, Kaylee M. Errecaborde, Michael W. Mahero, Katelyn W. Macy, Ong-Orn Prasarnphanich, Heidi Kassenborg, Erinaldi Yulizar, Rama P. Fauzi, Nyoman S. Budayanti, Agus Suwandono, Wayan T. Artama, Linda Valeri, and Katharine M. Pelican. Strengthening multi-sectoral collaboration on critical health issues: One Health Systems Mapping and Analysis Resource Toolkit (OH-SMART) for operationalizing One Health. *PLOS ONE*, 14(7):e0219197, July 2019. ISSN 1932-6203. doi: 10.1371/journal.pone.0219197. URL https://dx.plos.org/10.1371/journal.pone.0219197.

Flavio Villanustre et al. Modeling & tracking Covid-19 cases using big data analytics on HPCC system platform. *J. Big Data*, 8:1–24, 2021.

Nikolay K Vitanov and Zlatinka I Dimitrova. Simple equations method and non-linear differential equations with non-polynomial non-linearity. *Entropy*, 23(12):1624, 2021.

Isabell Wagenhäuser, Kerstin Knies, Tamara Pscheidl, et al. SARS-CoV-2 antigen rapid detection tests: test performance during the COVID-19 pandemic and the impact of COVID-19 vaccination. *EBioMedicine*, 109, 2024.

Sean Walker and Sharad Sharma. Data visualization tool for covid-19 and crime data. In *Int. Conf. Comp. Sci. & Comp. Int.*, pages 238–242, 2021.

Bingyuan Wang et al. Simonstown: An AI-facilitated interactive story of love, life, and pandemic. In *Symp. Vis. Inf. Comm. & Inter.*, pages 1–7, 2023.

Fang Wang, Leilei Xiang, Kelvin Sze-Yin Leung, Martin Elsner, Ying Zhang, Yuming Guo, Bo Pan, Hongwen Sun, Taicheng An, Guangguo Ying, Bryan W. Brooks, Deyi Hou, Damian E. Helbling, Jianqiang Sun, Hao Qiu, Timothy M. Vogel, Wei Zhang, Yanzheng Gao, Myrna J. Simpson, Yi Luo, Scott X. Chang, Guanyong Su, Bryan M. Wong, Tzung-May Fu, Dong Zhu, Karl J. Jobst, Chengjun Ge, Frederic Coulon, Jean Damascene Harindintwali, Xiankui Zeng, Haijun Wang, Yuhao Fu, Zhong Wei, Rainer Lohmann, Changer Chen, Yang Song, Concepcion Sanchez-Cid, Yu Wang, Ali El-Naggar, Yiming Yao, Yanran Huang, Japhet Cheuk-Fung Law, Chenggang Gu, Huizhong Shen, Yanpeng Gao, Chao Qin, Hao Li, Tong Zhang, Natàlia Corcoll, Min Liu, Daniel S. Alessi, Hui Li, Kristian K. Brandt, Yolanda Pico, Cheng Gu, Jianhua Guo, Jianqiang Su, Philippe Corvini, Mao Ye, Teresa Rocha-Santos, Huan He, Yi Yang, Meiping Tong, Weina Zhang, Fidèle Suanon, Ferdi Brahushi, Zhenyu Wang, Syed A. Hashsham, Marko Virta, Qingbin Yuan, Gaofei Jiang, Louis A. Tremblay, Qingwei Bu, Jichun Wu, Willie Peijnenburg, Edward Topp, Xinde Cao, Xin Jiang, Minghui Zheng, Taolin Zhang, Yongming Luo, Lizhong Zhu, Xiangdong Li, Damià Barceló, Jianmin Chen, Baoshan Xing, Wulf Amelung, Zongwei Cai, Ravi Naidu, Qirong Shen, Janusz Pawliszyn, Yong-guan Zhu, Andreas Schaeffer, Matthias C. Rillig, Fengchang Wu, Gang Yu, and James M. Tiedje. Emerging contaminants: A One Health perspective. *The Innovation*, 5(4): 100612, July 2024. ISSN 26666758. doi: 10.1016/j.xinn.2024.100612. URL `https://linkinghub.elsevier.com/retrieve/pii/S266667582400050X`.

Han Wang et al. Software engineers' response to public crisis: lessons learnt from spontaneously building an informative covid-19 dashboard. In *Proc. 44th Int. Conf. Soft. Eng.: Soft. Eng. Soc.*, pages 149–160, 2022.

Mengmei Wang and Shuguang Peng. Covid-19 visualization platform based on population density propagation model. In *IEEE 2nd Int. Conf. Data Science & Comp. App.*, pages 406–411, 2022.

Andrew Wentzel et al. A lens to pandemic stay at home attitudes. In *IEEE VIS Workshop Vis. Pandemic & Emer. Responses*, pages 15–19, 2023.

WHO. 14.9 million excess deaths associated with the covid-19 pandemic in 2020 and 2021. https://www.who.int/news/item/05-05-2022-14.9-million-excess-deaths-were-associated-with-the-covid-19-pandemic-in-2020-and-2021, 2022. Accessed:2024-03-24.

O Who. One health. *World Health Organization*, 736, 2017.

Wikipedia. List of epidemics and pandemics. https://en.wikipedia.org/wiki/List_of_epidemics_and_pandemics, 2023. Accessed on 12 April 2024.

Cort J Willmott and Kenji Matsuura. Advantages of the mean absolute error (mae) over the root mean square error (rmse) in assessing average model performance. *Climate research*, 30(1):79–82, 2005.

Michael C Wimberly, Justin K Davis, Michael B Hildreth, and Joshua L Clayton. Integrated forecasts based on public health surveillance and meteorological data predict west nile virus in a high-risk region of north america. *Environmental Health Perspectives*, 130(8):087006, 2022.

World Health Organization. *Communicable disease surveillance and response systems: guide to monitoring and evaluating*. World Health Organization, 2006.

World Health Organization. Quadripartite Call to Action for One Health for a Safer World, 2023. URL https://www.who.int/news/item/27-03-2023-quadripartite-call-to-action-for-one-health-for-a-safer-world. Accessed: 2024-11-04.

World Health Organization. One Health, 2024. URL https://www.who.int/health-topics/one-health#tab=tab_1. Accessed: 2024-10-28.

Kangpeng Xiao and Longxian Zhang. Wastewater pathogen surveillance based on One Health approach. *The Lancet Microbe*, 4(5):e297, May 2023. ISSN 26665247. doi: 10.1016/S2666-5247(23)00039-3. URL https://linkinghub.elsevier.com/retrieve/pii/S2666524723000393.

Haowen Xu, Andy Berres, Gautam Thakur, Jibonananda Sanyal, and Supriya Chinthavali. Episemblevis: A geo-visual analysis and comparison of the prediction ensembles of multiple covid-19 models. *Journal of biomedical informatics*, 124:103941, 2021.

Santosh Kumar Yadav, Lilee Shrestha, Jyoti Acharya, Tulsi Ram Gompo, Sharmila Chapagain, and Runa Jha. Integrative Digital Tools to Strengthen Data Management for Antimicrobial Resistance Surveillance in the "One Health" Domain in Nepal. *Tropical Medicine and Infectious Disease*, 8(6):291, June 2023. ISSN 2414-6366. doi: 10.3390/tropicalmed8060291. URL https://www.mdpi.com/2414-6366/8/6/291. Number: 6 Publisher: Multidisciplinary Digital Publishing Institute.

Tao Yang et al. Cvas: An interactive visual analytics system for exploring covid-19 information on the web. In *Proc. Int. Conf. Comp. & Data Analy.*, pages 116–121, 2022.

Cheng-Sheng Yu et al. A COVID-19 pandemic artificial intelligence-based system with deep learning forecasting and automatic statistical data acquisition: Development and implementation study. *J. Medical Int. Res.*, 23(5):e27806, 2021.

Jie Yuan, Yu Wu, Wenzhan Jing, Jue Liu, Min Du, Yaping Wang, and Min Liu. Association between meteorological factors and daily new cases of covid-19 in 188 countries: A time series analysis. *Science of The Total Environment*, 780:146538, 2021a.

Jun Yuan, Changjian Chen, Weikai Yang, Mengchen Liu, Jiazhi Xia, and Shixia Liu. A survey of visual analytics techniques for machine learning. *Computational Visual Media*, 7(1):3–36, 2021b.

Jintao Zhang, Qinglong You, Guoyu Ren, Safi Ullah, Inom Normatov, and Deliang Chen. Inequality of global thermal comfort conditions changes in a warmer world. *Earth's Future*, 11(2):e2022EF003109, 2023.

Zhe Zheng, Yi Lu, Kirsty R. Short, and Jiahai Lu. One health insights to prevent the next HxNy viral outbreak: learning from the epidemiology of h7n9. 19(1):138, 2019. ISSN 1471-2334. doi: 10.1186/s12879-019-3752-6. URL `https://bmcinfectdis.biomedcentral.com/articles/10.1186/s12879-019-3752-6`.

Fei Zhou, Ting Yu, Ronghui Du, Guohui Fan, Ying Liu, Zhibo Liu, Jie Xiang, Yeming Wang, Bin Song, Xiaoying Gu, et al. Clinical course and risk factors for mortality of adult inpatients with covid-19 in wuhan, china: a retrospective cohort study. *The lancet*, 395 (10229):1054–1062, 2020.

Shaher Zyoud. Global Mapping and Visualization Analysis of One Health Knowledge in the COVID-19 Context. *Environmental Health Insights*, 18:11786302241236017, January 2024. ISSN 1178-6302. doi: 10.1177/11786302241236017. URL `https://doi.org/10.1177/11786302241236017`. Publisher: SAGE Publications Ltd STM.