2020 IEEE Visualization Conference (VIS) | 978-1-7281-8014-4/20/$31.00 ©2020 IEEE | DOI: 10.1109/VIS47514.2020.00018

COVIs: Supporting Temporal Visual Analysis of Covid-19 Events Usable in Data-Driven Journalism

Roger A. Leite* — TU Wien
Victor Schetinger† — TU Wien
Davide Ceneda‡ — TU Wien
Bernardo Henz§ — IFFar
Silvia Miksch¶ — TU Wien

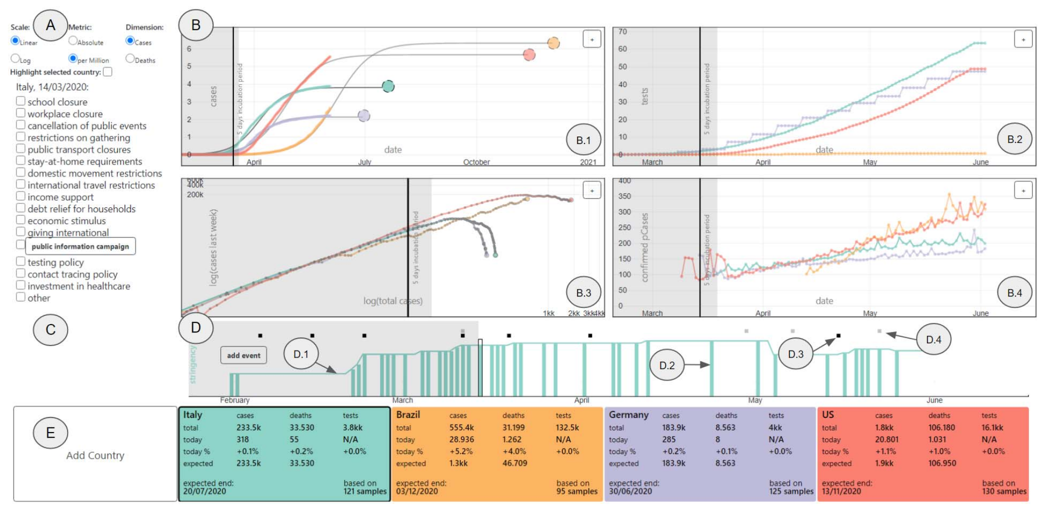

Figure 1: *COVis*'s coordinated multiple view environment: (A) Control Panel: allows the user to change the graphs scale (linear / log), metric (absolute / per million cohabitants), and dimensions (cases / deaths). (B) Line charts: present four different line charts that are coordinated to support exploration of multiple narratives. Respectively, the charts display the relation between: (B1) time x cases/deaths, (B2) time x tests, (B3) total cases/deaths x last week cases/deaths, and (B4) time x cases/deaths projection length. (C) Events Panel: displays information and source references concerning main events occurred in certain time periods. (D) Events Time Chart: chart presenting the policy changes of a country over time. (E) Country Cards: show information concerning the analysed group of countries, allowing the exclusion and inclusion of different countries into the analysis.

Abstract

Caused by a newly discovered coronavirus, COVID-19 is an infectious disease easily transmitted between people through close contacts that had exponential global growth in 2020 and became, in a very short time, a major health, and economic global issue. Real-world data concerning the spread of the disease was quickly made available by different global institutions and resulted in many works involving data visualizations and prediction models. In this paper, (1) we discuss the problem, data aspects, and challenges of COVID-19 data analysis; (2) We propose a Visual Analytics approach (called *COVis*) combining different temporal aspects of COVID-19 data with the output of a predictive model. This combination supports the estimation of the spread of the disease in different scenarios and allows correlating and monitoring the virus development in relation to different government response events; (3) We evaluate the approach with two domain experts to support the understanding of how our system can facilitate journalistic investigation tasks and (4) we discuss future works and a possible generalization of our solution.

Index Terms: COVID-19—Visual Analytics—Data Visualization—Prediction ModelEventsTime-oriented Analysis;

*e-mail: roger.leite@tuwien.ac.at
†e-mail: victor.schetinger@tuwien.ac.at
‡e-mail: davide.ceneda@tuwien.ac.at
§e-mail: bernardo.henz@iffarroupilha.edu.br
¶e-mail: silvia.miksch@tuwien.ac.at

1 Introduction

The recent COVID-19 epidemic has triggered a global crisis that tests our modern civilization's potential for coordination and cooperation. While scientists, epidemiologists, and health specialists have a central role in understanding the disease itself, its social, political and economical aspects can be much harder to analyze [5]. Every country responded to the crisis in a different way and in its own time. Data about contagion rates and their growth has been the main source for public discussion on the subject. However, relating such data to global and local events within each country is a complex puzzle and many specialists still rely primarily on their Microsoft Excel [25] skills to perform data-driven analysis (see Section 6). Journalists, social scientists and economists may be the ones most suited for the task of piecing this puzzle together, but they need the proper tools to visualize the bigger picture in a sea of data and news that evolves on a daily basis.

One of the most important aspects to be understood is the relation between events – political or otherwise – and the spread of the virus in each local context. The Oxford University [11] provides a dataset of public measures for each individual country, along with an index estimating the level of closure of a determinate country. This data can be combined with the global rates of infection and predictive models to construct a rich picture of the development of the crisis. In this paper, we propose a Visual Analytics (VA) approach (called *COVis*) to aid the temporal analysis and exploration of narratives under an investigative framework, focusing on the identification of significant events and time frames. This is a process of storytelling, for which we provide means to be grounded as much as possible in available data and evidence, so that communicators need not rely solely on speculation, exaggeration, and interpretation.

DOI 10.1109/VIS47514.2020.00018

Visualizing time in COVID-19 data is challenging because the daily, linear time dimension provided hides different temporal phenomena occurring simultaneously. We point to four key challenges considered in this work: (Time-1) The virus has a variable incubation period of 5-6 days on average [28], which means actions and events have a variable delay before their noticed impact. (Time-2) The rate of sampling is different and varies for each country, generally increasing in response to the spread of the virus and perceived urgency. These two aspects generate a lot of temporal uncertainty, as they imply that new data records will always be an under-representation of reality, making it harder to support real-time analysis. (Time-3) Each country has its own timeline, usually starting with the identification of the first infection case. All these timelines have different starting points and developments, which cannot be treated equally or independently. Countries infected earlier and their neighbors had to deal with a different scenario both in terms of available information and population behavior than latter ones. As the public awareness of the virus increased and global mobility decreased, countries could react more effectively, generating different behaviors for their timelines. (Time-4) there is a large amount of possible future data from predictive models. While these projections weight heavily on public discussion, it is not clear how to best visualize them and form a reliable perspective, dangerously facilitating misrepresentation [8, 31]. In this work, we tackle these challenges from a data journalism point-of-view, supporting their way of thinking [22] and storytelling tasks. Our main contributions are:

- Identify the COVID-19 pandemic data aspects and their problem challenges.
- Propose a VA approach that incorporates a growth predictive model, allowing the correlation and monitoring of the virus population development in relation with different government response events, as well as the identification of key dates.
- Evaluate our VA solution with two domain experts to support the understanding of our approach.
- Elaborate future works and generalization of our solution.

2 Related Work

The graveness and the ubiquitous impact of the COVID-19 crisis has led to a spike in research about the subject, with a great deal of visualizations sprouting from governments, organizations, academia, and general public. Many countries provided their own visualizations [4, 24] with different features as means to communicate their situation to the public. As data became more available, standardized composites and comparative approaches became popular. This increased data availability and allowed the application of prediction techniques, ranging from simple curve fittings to advanced machine learning and simulation models. However, there was not enough time for all available works being peer-reviewed. A great deal of work remains unpublished, or available through repository sources.

CoronaVIS [3] is an online tool that presents an overview of the hospital‘s bed capacities as well as standard information about the disease development throughout Germany. The main goal is to geographically identify potential hospital occupancy bottlenecks and optimise the distribution of patients. Using different maps visualization, Zhou et al. [32] focus in the relation of geographic aspects of the data, such as regional transmission and spatial segmentation of the epidemic risk. This work also discuss the important role of rapid aggregation of multi-source big data into the understanding of the COVID-19 phenomenon. Different from these works [3, 32], we did not focus on geographical features however, we also used multi-source big data aggregation.

Kevorkian et al. [19] developed a visualization to track the epidemic states and compare the different countries trajectories trough time. In this visualization the authors do not explicitly plot time however, since each dot represents a given day, it is possible to track the growing behavior of the disease. To do so, the authors used a technique equivalent to phase space plots (commonly used to represent dynamical systems in Physics). We present similar technique in one of our views (see Figure 1: B.3), while helping the identification of interesting dates for investigation through the coordinated view, which anchors the real time frame.

COVis provides an estimate of the disease population growth based on Machine-Learning's (ML) curve-fitting, offering a future time-frame. This is a rather simple approach, as our focus is understanding how to construct a narrative from the composite frames of past, present, and future. Many authors have employed more sophisticated ML and predictive models with different goals on this subject. Muthusami et al. [29] used sentiment analysis based on tweeter data to analyse and visualize the influence of COVID-19 in social media. Kai et al. [18] defined two theoretical models to understand the positive impact of face mask wearing on the COVID-19 pandemic. This simulation-based approach considers compares different location scenarios by using line charts and pixel-based matrix view. Different from our approach that is purely data driven, in this work causal models were constructed based on the classic SEIR model of infectious disease [23].

We consider our work to fall within the context of visualization for the digital humanities [1], since our targeted users are journalists. Data-driven journalism shares a close relationship with VA, as it is centered around finding and exposing insights in data. Even so, there are varying degrees of acceptance of technical solutions, and journalists have to rely on their own digital literacy to navigate in the modern data deluge [15]. We support journalists through supporting storytelling, inspired in Chen's [7] *story synthesis framework*, which uses *story slices* as nuclear aggregates of findings that are building blocks of a narrative. Our approach, however, organically creates time slices, which will rely on a journalist to give them meaning.

3 Methodology

To design our approach, we combine the "Nested Model" approach by Munzner [27] with the "Design Triangle" by Miksch and Aigner [26]. Based on Munzner [27], we characterized tasks and data vocabulary of the problem domain, abstracted the problem into operation and data types, designed visual encoding and interaction techniques, and created algorithms for efficiently executing the techniques. To achieve a more accurate tool, based on Miksch and Aigner [26], we define and target the following data, user, and tasks:

Data. Our data is the combination of two datasets, plus our generated predictions. The first dataset [16] consists of the core statistics of COVID-19 development such as confirmed cases, amount of tests, and death toll for each country. The second dataset [11] is from the Oxford University and contains the dates and news sources for public measures, such as movement restrictions or school closures, together with stringency levels (a metric that tries to capture how closed a country is). It provides the event-based dimension to our approach, and allows a direct link to source investigation. This data is integrated with our projections using a logistic regression model, which is discussed in the next section.

User. Our main target public are journalists or investigative reporters. However, we consider that also common public can take advantage of *COVis* to explore and reason over COVID-19 data.

Tasks. This work's main task is story prospecting. From a journalistic point of view, this can be evaluated on the classic principle of the "Five Ws" (Who, What, Where, When, Why) [14]. We aim to support the answering of these five questions around a subject, as the building block for storytelling, with a special emphasis on the causal and temporal aspects (when and why). Using the "five Ws" as guideline, we defined the **system requirements** according to the described data, user, and tasks: Identification of different countries behavior **R1** (Who? and Where?), identification of events influence **R2** (What?), perception of variables change over time **R3** (When?), reasoning concerning changes of data pattern **R4** (Why?).

4 Predictive Model

The incorporation of a predictive method allows a glimpse of future scenarios, helping the understanding of current variables as they are exacerbated and highlighted by their projection through time. There are two prominent approaches to model epidemics, like the COVID-19: Causal Models [18, 20] and Data-Driven Models [9]. We opt for using a data-driven approach due to two factors: (1) the diversity of trustful online data available by official governmental institutions and (2) the comparison potential between countries provided by results generated by the same metrics. Causal models are more complex and flexible, but require specific fine-tuning of the system for each country. In this sense, the local causalities integration result in less comparable results to different countries.

Our predictive ML model results from curve-fitting a logistic function over the number of cases/deaths in relation to time, to determine the date when growth would cease. Logistic curves are widely used to represent population growth [2, 12], which is the case of the cases/deaths. The function parameters are fit using least-squares optimization. As new data become available, the predictive model can be instantly updated. Due to the behavior of the logistic function, the influence of new samples is diluted, which is a problem since more recent data tends to be more accurate and reliable. To account for this, a lower sigma-weight is given to the last quarter of samples, forcing the latter and more accurate data to have a larger influence on the prediction. Based on the i-th position of the sample, with N being the number of samples, the sigma-weights are computed following:

$$\sigma_i = \begin{cases} 1, & \text{if } i \leq \dfrac{3N}{4} \\ 1 - \dfrac{i}{N} + 0.6, & \text{otherwise.} \end{cases} \tag{1}$$

It is important to note that the focus of our work is not the prediction model. Any different model or data from outsourced predictions could be used. The main role of prediction data is to explore visualization of (probable) future time, which guides public discussion and reaction, most notably when it changes quickly, as can be analyzed in our bottom-right graph (see Figure 1: B.4).

5 *COVis*: an Interactive VA Environment

COVis's layout is composed by five coordinated views that are interactively combined to support the gaining of data insights (see Figure 1). Next, we describe the different views and derive possible story cues from them. The main technologies used to develop the prototype were Angular and D3.js.

(A) Control Panel: This panel allows the user to switch between scales (linear / log), metric (absolute numbers / per million inhabitants), and dimension (active cases / deaths) to be displayed (R2). Comparing absolute and per million inhabitants data gives different perspectives on the planning and potential for mobilization of each country. Death data, on the other hand, might be a more accurate measure of spread than active cases, as it does not rely on testing rates, assuming a constant mortality rate after infection.

(B) Line Charts: Each of the line graphs shows a distinct facet of time, telling different aspects of each countries' development during the crisis (R1-R4). In the first graph (top left) we can visualize the overall growth of cases or deaths, fitted to a logistic curve that projects a possible point of inflection, or stability. Past (the line itself), present (the end of the line), and future (the idealistic projection) are superimposed in it, allowing us to visually correlate them. The four charts present interactive features such as: zoom, context dragging, mouse over data nodes, and view-size maximization.

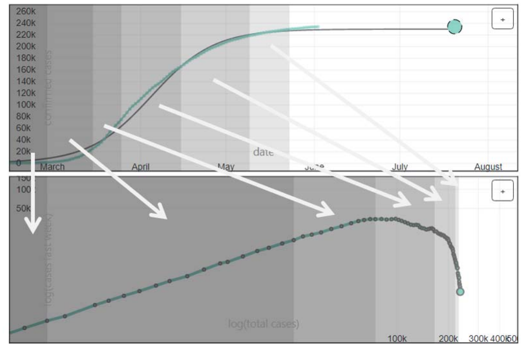

Figure 2: Example of gradient formed from multiple *temporal shades* by selecting all "Stay at Home" events for Italy. Each grayscale represents an epoch with the same events in effect. As brighter, as more events are influencing the curve. The top graph depicts cases over time (see Figure 1: B1), while bottom graph illustrated last week cases over total cases (see Figure 1: B.3). Here, time can be perceived in two different ways: both explicitly (top), and implicitly (bottom). The arrows indicate representations of same time periods. The disruption in linear behavior can be perceived just as the fourth epoch kicks in, suggesting the presence of significant ongoing events.

The top left graph (see Figure 1: B1) shows the amount of confirmed cases, which can be interpreted as the sampling rate of our signal. The first derivative (slope) of this graph shows the increase of tests, indicating each countries response. Italy, as the protagonist of the epidemic in the west, has been the most desperate to react and contain the epidemic, having the most consistent increase in tests, and performing the most tests per million inhabitants. The US, while initially delayed in responding, was able to quickly scale its test count to over 10 million, the highest by a large margin, attesting to the country's economic potential and ability to mobilize resources, which can be contrasted with Brazil. Both countries showed a similar critical growth in cases, but Brazil was not able (or willing) to scale its testing accordingly. An informed journalist could observe that the country is ongoing an acute political crisis, and be able to work these visible patterns into a story. Germany's testing behavior can be seen as strikingly methodical, with weekly increments in reporting.

The bottom left graph (see Figure 1: B3) shows the log of the amount of cases (or deaths) in the last week, over the log of the total cumulative amount. While there is not an **explicit** time dimension, time is encoded **implicitly** through the accumulation of total cases on the X axis. Since countries achieved different counts at different times, every position in X corresponds to a different date in each line. This can be noted when using the *temporal shading* feature, where the date will be centered according to the day the selected country achieved that count. In this graph, time gets compressed on the X axis as the growth rate decreases for the country, since the weekly increase cannot keep up with the accumulated amount. Since the spread of the virus is exponential in nature, drops in the line can be understood as deviations from this expectation, and therefore successful (or situational) braking of spread.

Finally, the bottom right graph (see Figure 1: B4) displays the deviation over time for the predicted date of cessation of growth. It is the combined perspective of the three other graphs, allowing us to visualize the stability of both growth and testing for each country. Elevation in this line is caused by the estimated end date being pushed forward, because new measurements are causing the model to fit the growth rate to a different curve, with a different behavior. This graph can be used to identify dates that generated instability, allowing journalists to further investigate for relevant events.

(C) Events Panel: The events panel can display up to 17 event-types [11] concerning containment and closure, economic response, and health systems. Using the checkbox, all events from the respective type are highlighted in the timeline, and in the line chart views (B) by the *temporal shading* effect (Section 5: (B)). By selecting a date in the events time chart (D), events that occurred in that date are highlighted, providing details and source of the event (R2, R4). On the other hand, by selecting an event type in the events panel (C), its frequency is highlighted in the event time chart (D).

(D) Events Time Chart: The chart is colored according to the selected country in the country cards view (E)(R3,R4). This interactive chart displays four main features: the line (see Figure 1: D.1) shows the stringency level [11] through time, which indicates how strict are local policies, while vertical bars mark governmental events (see Figure 1: D.2). World events that are relevant for the epidemic, such as one million global cases, are represented by black squares (see Figure 1: D.3), and create a consistent time anchor that supports comparison between countries. Furthermore, the user can add new events, represented by gray squares (see Figure 1: D.4), which allows journalists to use their knowledge to enrich analysis and tailor narratives. By passing the mouse over bars (D.2) or squares (D.3 and D.4), a summary of the event is displayed. By selecting the bars, the user can investigate the different types of events of that specific date (marked in the events panel (C)).

By selecting an event date, a *temporal shade* is cast on all graphs. This simple feature aids in tackling our four temporal challenges (Time-1 to Time-4). The shade darkens the time frame where the event could **not** have an accountable influence, defined as its past time, plus five days of incubation. This helps the user understand possible associations and causality (Time-4), while accounting for the incubation delay (Time-1).

When multiple dates are selected, for example when selecting all events of a certain type from the events panel (C), multiple overlapping *temporal shades* are cast (Fig. 2). This forms a gradient-like visual pattern with multiple grayscale levels, where each level represents a stable "epoch" in relation to the events selected (that is, influenced by the same events). This provides insightful details, allowing one to better understand the temporal structure of each country (3), where event frequencies can be quickly perceived. Since the *temporal shades* are consistent along all graphs, they generate a new time scale that can tie the different timescales on each graph through visual association (see Arrows in 2). For example, *temporal shades* on the performed tests graph (Fig.1, B.2) allows one to visualize correlations between different testing rates and identify sampling patterns between epochs (Time-2).

(E) Country Cards: This panel displays raw data(R2) about total, and last day, amounts of cases deaths and tests, as well as the ending predicted date (see Section 4). The country cards also allows the selection of a specific country(R1), highlighting its respective events-time and curve; and the addition of additional countries with available data into the visualization. We chose a qualitative seven colors palette [13] to encode different countries trough the views, thus, the *COVis* allow the comparison of maximum seven countries.

6 Evaluation: User Study

In order to estimate the usability of *COVis* and its value for journalists, we conducted a qualitative evaluation with two domain experts. Qualitative studies are a useful means to generate insights with relatively few study participants [17, 21]. The evaluation was divided in four parts: (1) we asked about the participant background, (2) we presented the system, (3) the participants could freely interact with the system, investigating questions of their interest, and, (4) we asked the participants about their discoveries and impressions of *COVis*. The two individual interviews were performed and recorded via a virtual conference platform (zoom). In average, every evaluation took two hours to be completed.

Participants background interview. The first participant is a journalism university professor (P1) while the second participant is an independent journalist (P2). Both declared having experience working with data investigation during journalistic story development and highlighted Microsoft Excel [25] as the common tool. P1 highlighted the common lack of mathematical knowledge in the journalism community, pointing out the difficulties that log-scale graphs interpretations bring to the journalists‘ final report.

***COVis*'s exploration.** During a non-oriented system exploration we asked the participants to "think aloud" about potential stories inspired by the data and graphs they were experimenting. This way, we could validate the "five Ws" task/requirements usage and their natural positive and negative reaction to the system. As a result, both participants developed and completed stories using the defined requirements. Being able to visually couple personalized events with existing data events was the feature that added most to the story development by both participants. P1 drew a narrative that tried to contextualize the countries different testing rates, which we summarized when explaining Figure 1: B3 in Section 5B. P2 suggested an analysis of the negative impact of Brazilian President Bolsonaro appalling statements, such as "it is just a little flu" [10], or "we are all gonna die one day" , which could be achieved by treating them as events and adding them to our timeline.

Feedback to *COVis*. P1 appreciated the different charts combination and integration, and highlighted the dimension and scale flexibility as a good support to understand and create stories. P2 made several positive comments concerning the feasibility of reasoning about events in time trough the different dimensions perspectives (multiple views) and coordinated interactions. In comparison, P2 commented that analysis of temporal data is usually a cumbersome task for journalists, specially when involving many events, with Microsoft Excel [25] spreadsheets still being a standard. In this sense, the visual aid was extremely appreciated.

Without specific task instruction the participants naturally performed story explorations that followed their particular investigative styles and preferences. P2 highlighted the ad-hoc nature of journalistic workflows, which are defined on a personal or institutional basis, and was interested on the possibilities for developing a workflow. P1, as a journalism professor, suggested to evaluate the tool by giving an assignment to his students to write stories using it. This would allow us to see different workflows that could emerge, and grading could provide a form of quantitative evaluation. Both participants gave positive feedback on the fruitfulness of possible writing material and suggested future improvements (see Section 7).

7 Conclusion and Future Work

We presented *COVis* to support journalistic story telling through temporal analysis of epidemic data and governmental events. Our approach tackles important challenges in temporal visualization of COVID-19 data, providing an accessible environment for information discovery. We received positive feedback from experts in the field of journalism, which highlighted the lack of technical solutions that can bridge the gap between data analysis and story synthesis. Based on current limitations, we derive three research challenges to inspire future works: (1) Improve the learning-curve and the analysis power of journalists, guidance could be incorporated [6], as well as didactic devices for data interpretation [30]. (2) Perform more thorough evaluation to understand the potential of our approach in assisting data-driven journalism, and its impact on the story writing process. (3) Use data from more sophisticated predictive models, allowing both the comparison of different future scenarios and exploration/visualization of uncertainty in their projections.

Acknowledgments

This work was supported and funded by the Austrian Science Fund (FWF) with the grant #P31419-N31 (KnoVA).

REFERENCES

[1] Visualization and the Digital Humanities:. 38. doi: 10.1109/MCG.2018.2878900

[2] L. J. Allen. Some discrete-time si, sir, and sis epidemic models. *Mathematical biosciences*, 124(1):83–105, 1994.

[3] T. D. Analysis and V. G. D. of Daniel A. Keim at the University of Konstanz. Coronavis. https://coronavis.dbvis.de/en/ (accessed: 07/06/2020).

[4] Austria. Gesundheitsministerium data web page. https://info.gesundheitsministerium.at/?l=en, see Epidemiological Data Tab (accessed: 07/06/2020).

[5] S. Callaghan. Covid-19 is a data science issue. *Patterns*, 1(2):100022, 2020. doi: 10.1016/j.patter.2020.100022

[6] D. Ceneda, T. Gschwandtner, T. May, S. Miksch, H.-J. Schulz, M. Streit, and C. Tominski. Characterizing guidance in visual analytics. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):111–120, 2017. doi: 10.1109/TVCG.2016.2598468

[7] S. Chen, J. Li, G. Andrienko, N. Andrienko, Y. Wang, P. H. Nguyen, and C. Turkay. Supporting Story Synthesis: Bridging the Gap between Visual Analytics and Storytelling. *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–1, 2018. doi: 10.1109/TVCG.2018.2889054

[8] J. Drucker. Data as capta. *Digital Humanities Quarterly*, 5(1):1–21, 2011.

[9] S. P. Ellner, D. Z. Childs, M. Rees, et al. Data-driven modelling of structured populations. *A practical guide to the Integral Projection Model. Cham: Springer*, 2016.

[10] T. Guardian. "brazil's jair bolsonaro says coronavirus crisis is a media trick". https://www.theguardian.com/world/2020/mar/23/brazils-jair-bolsonaro-says-coronavirus-crisis-is-a-media-trick (accessed: 24/08/2020).

[11] T. Hale, A. Petherick, T. Phillips, and S. Webster. Variation in government responses to covid-19. *Blavatnik School of Government Working Paper*, 31, 2020.

[12] M. Haque, F. Ahmed, S. Anam, and R. Kabir. Future population projection of bangladesh by growth rate modeling using logistic population model. *Annals of Pure and Applied Mathematics*, 1(2):192–202, 2012.

[13] M. Harrower and C. A. Brewer. Colorbrewer. org: an online tool for selecting colour schemes for maps. *The Cartographic Journal*, 40(1):27–37, 2003.

[14] G. Hart. The five W's: An old tool for the new task of task analysis. *Technical communication*, 43(2):139–145, 1996.

[15] A. Hermida and M. L. Young. Finding the Data Unicorn: A hierarchy of hybridity in data and computational journalism. *Digital Journalism*, 5(2):159–176, Feb. 2017. doi: 10.1080/21670811.2016.1162663

[16] O. W. in Data. *Data on COVID-19 by Our World in Data*, 2020 (accessed Jun 3, 2020). https://github.com/owid/covid-19-data/tree/master/public/data.

[17] T. Isenberg, P. Isenberg, J. Chen, M. Sedlmair, and T. Möller. A systematic review on the practice of evaluating visualization. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2818–2827, 2013.

[18] D. Kai, G.-P. Goldstein, A. Morgunov, V. Nangalia, and A. Rotkirch. Universal masking is urgent in the covid-19 pandemic: Seir and agent based models, empirical validation, policy recommendations. *arXiv preprint arXiv:2004.13553*, 2020.

[19] A. Kevorkian, T. Grenet, and H. Gallee. Tracking the covid-19 pandemic: Simple visualization of the epidemic states and trajectories of select european countries & assessing the effects of delays in official response. *medRxiv*, 2020.

[20] J. S. Koopman and J. W. Lynch. Individual causal models and population system models in epidemiology. *American Journal of Public Health*, 89(8):1170–1174, 1999.

[21] S. Kriglstein and M. Pohl. Choosing the Right Sample? Experiences of Selecting Participants for Visualization Evaluation. In W. Aigner, P. Rosenthal, and C. Scheidegger, eds., *EuroVis Workshop on Reproducibility, Verification, and Validation in Visualization (EuroRV3)*. The Eurographics Association, 2015. doi: 10.2312/eurorv3.20151146

[22] S. C. Lewis and O. Westlund. Big Data and Journalism: Epistemology, expertise, economics, and ethics. *Digital Journalism*, 3(3):447–466, May 2015. doi: 10.1080/21670811.2014.976418

[23] M. Y. Li and L. Wang. Global stability in some SEIR epidemic models. In *Mathematical approaches for emerging and reemerging infectious diseases: models, methods, and theory*, pp. 295–311. Springer, 2002.

[24] W. Meters. Covid-19 coronavirus pandemic. https://www.worldometers.info/coronavirus/ (accessed: 07/06/2020).

[25] Microsoft. Excel. office.microsoft.com/en-us/excel/ (accessed: 07/06/2020).

[26] S. Miksch and W. 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.

[27] T. Munzner. A nested model for visualization design and validation. *IEEE transactions on visualization and computer graphics*, 15(6):921–928, 2009.

[28] W. H. Organization et al. Coronavirus disease 2019 (covid-19): situation report, 73. 2020.

[29] M. Ra, B. Ab, and S. Kc. Covid-19 outbreak: Tweet based analysis and visualization towards the influence of coronavirus in the world.

[30] Z. Wang, L. Sundin, D. Murray-Rust, and B. Bach. Exploring Cheat Sheets for Data Visualization Techniques. p. 2.

[31] C. Xiong, J. Shapiro, J. Hullman, and S. Franconeri. Illusion of Causality in Visualized Data. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):853–862, Jan. 2020. doi: 10.1109/TVCG.2019.2934399

[32] C. Zhou, F. Su, T. Pei, A. Zhang, Y. Du, B. Luo, Z. Cao, J. Wang, W. Yuan, Y. Zhu, et al. Covid-19: Challenges to GIS with big data. *Geography and Sustainability*, 2020.