INTERNATIONAL JOURNAL OF CREATIVE MEDIA RESEARCH

Vol.12, April/October 2025

Telling a Data Story well: Communicating Covid-19 Variants to a wide audience through a Generative Graphic System

DOI: https://doi.org/10.24135/ijcmr.v12i1.159

Simone Gumtau, *University of Portsmouth*

Abstract

This paper will discuss the ethical issues involved with designing scientific data to tell an accessible story, while aiming to stay true to scientific credibility. At the height of the Covid-19 pandemic, at the end of 2020, MSM and social media discourses were concerned with the "new" or "mutant" variant of Covid-19 (Boyd, 2020). This story about variants suddenly emerging added to the fears and anxieties around the pandemic (Luo et al, 2021; Mach et al, 2021). The UK government in secret Whatsapp messages as part of "Project Fear" appears to have planned to utilise these fears to ensure compliance with government-imposed restrictions (Badshah, 2023). Understanding the science, i.e. the evolving nature of genetic sequences (viruses) and how they are transported across communities, I argue here, could enlighten us to see these processes are ever present – thereby alleviating anxieties of the 'one' variant, and also changing behaviours by understanding the importance of keeping to hygiene concepts, for example.

This paper will discuss the ethical issues involved with designing scientific data to tell an accessible story, while aiming to stay true to scientific credibility. At the height of the Covid-19 pandemic, at the end of 2020, MSM and social media discourses were concerned with the "new" or "mutant" variant of Covid-19 (Boyd, 2020). This story about variants suddenly emerging added to the fears and anxieties around the pandemic (Luo et al, 2021; Mach et al, 2021). The UK government in secret Whatsapp messages as part of "Project Fear" appears to have planned to utilise these fears to ensure compliance with government-imposed restrictions (Badshah, 2023). Understanding the science, i.e. the evolving nature of genetic sequences (viruses) and how they are transported across communities, I argue here, could enlighten us to see these processes are ever present – thereby

alleviating anxieties of the 'one' variant, and also changing behaviours by understanding the importance of keeping to hygiene concepts, for example.

A data visualisation design was developed based on genetic sequencing data of the Covid-19 spread through the South East UK, which allowed the identification of the evolution of different virus strains in the region over days. The animation allows the audience to focus on the rapidly appearing evolution of genetic sequences throughout the days of March-May in 2021. This visualisation makes a very strong statement by focusing on the process of mutation in relation to time and space. It is designed to be accessible to the less data literate and therefore target precisely the audience that may otherwise be tempted by misrepresentation, conspiracy theories and fake news. It is also explicitly a rhetorical device (and not claiming to be objective) and is intended to evoke and facilitate further conversations about the nature of the virus spread, as well as about the truth value in data visualisation. It aims to visualise the essence of the process of mutation, which aims to communicate the relationship and behaviour of virus strains in their environment. It should therefore be seen as the visualisation of a system, rather than a data set.

The essence, the behaviour of genetic systems, was prototyped as the data story here. 2 versions were created, one as a vector-based animation, which very quickly communicates the sense of the variations spread across a map of Hampshire. The second version was scratched into wax coatings revealing a shiny surface. This very tactile activity involved a close and slowed down interaction with the data for the designer, and it was intended to communicate the nuances of complexity through analogue means, attempting to strike the balance between scientific information and emotional impact.

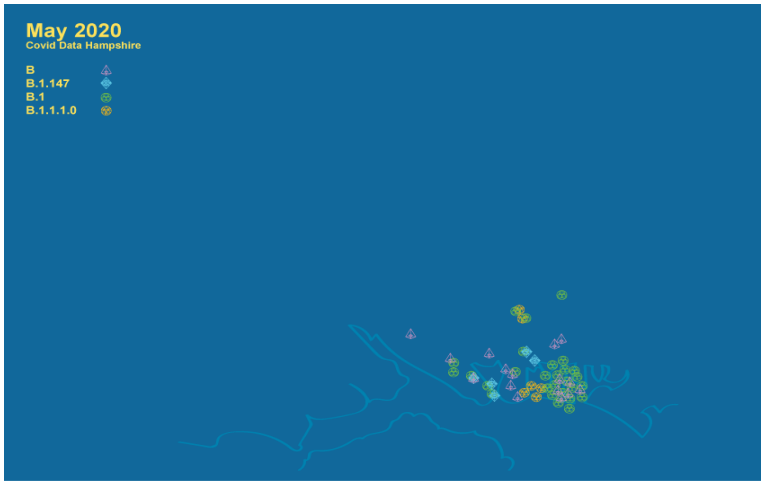

Figure 01: Version 1 of the Covid Data Visualisation, showing virus spread in Hampshire, South coast of the UK in March 2020

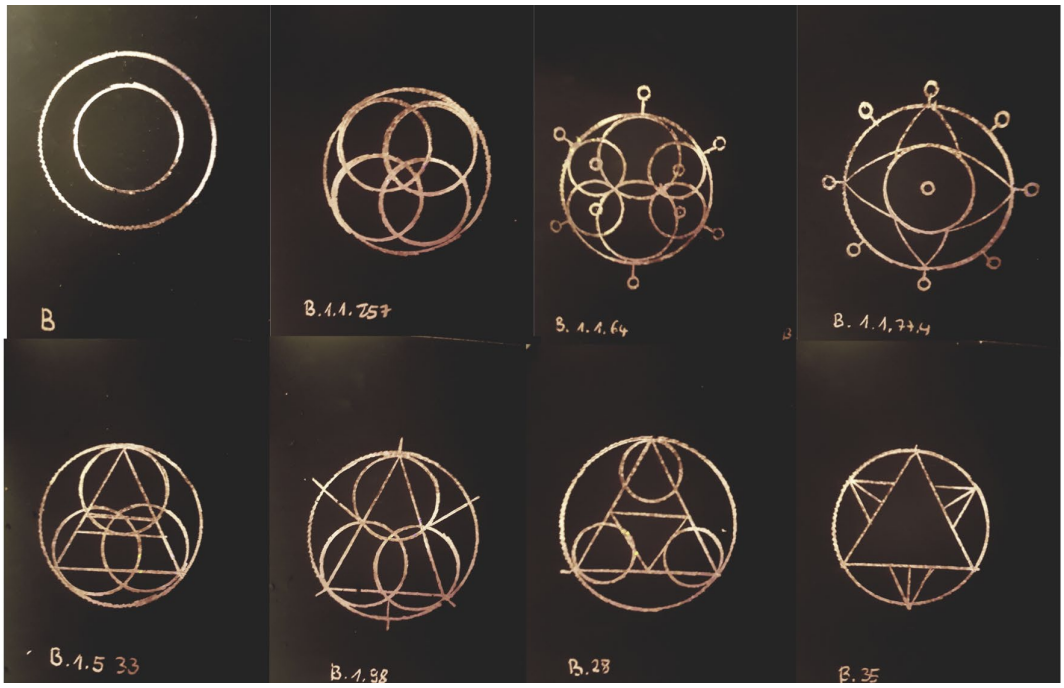

Figure 02: Version 2 of the Covid Data Visualisation, showing the virus lineage communicated as evolving geometric forms - scratched into a wax coating

Figure 03: Snapshot of Version 2 animation, showing the date 27/03/2020

Figure 04: Snapshot of Version 2 animation, showing the following day, 28/03/2020

This design approach is situated in the theory of complexity and systems and builds on the practice of systems and generative art (Galanter, 2012). A system of graphic symbols was devised to allow for generativity and evolution, similar to virus strains. Its aim was to make the information more user-friendly, while maintaining a sense of complexity. The animated geometric forms adapt and evolve into more complex shapes, and it is argued here that this may be a useful design method in data driven design for visualising virus strains. Taking a systems, non-mechanistic worldview (the 'whole is greater than the sum of its parts') the design focuses on the behaviour of the system and relationships rather than discrete objects (Capra, 1996; Latour, 2005). Networks, relationships and environments become transmitters of meaning. In this context, creative practitioners can design systems that can have emergent outcomes, i.e. evolving shapes into unpredictable forms and patterns. Research has shown that generative design systems of large scale data sets can help designers better comprehend inter-relationships between variable parameters, constraints and outcomes and optimise their design process (Chaszar et al, 2016; Matejka et al, 2018, Curralo, 2015). Following Peeters' (2016) suggestion to use Roland Barthes' thoughts on the "Work" and "Text", it could be argued that the audience of a generative design will make their own reading and therefore contribute to the emergence of meaning.

The prototype design is inspired by William Latham's work "MUTATOR 1 and 2" (2013), who painstakingly drew mutating shapes by hand, before translating it into projections generated by evolutionary software art. For the design presented here, the generation of form was created manually, but according to laws of geometry and evolution, manifested through algorithmic coding. Through dialogue with the genetic scientists who had produced the data sets used, the "main message", or the story of the data was distilled for clarity. However, within any process of filtering, cleaning, shaping the data, inevitably we face issues of inbuilt bias. All scientific visualisation methods can suffer from this, as they are also always products of a certain time and culture – although they might be seducing us to think they represent perfect truth and objectivity (D'Ignazio & Klein, 2020; Engebretsen & Kennedy, 2020).

Furthermore, we should consider the positionality and ethics of the designer in relation to data visualisation, and critically engage with notions of truth. There is always a balancing act between aiming for simplicity and too much distortion; as argued by data visualisation experts such as Edward Tufte who referred to too much information as "chartjunk" (1990), the ethics of design (Roberts, 2006), and philosophical discussions of the notions of truth and science (Malpass, 1992; Latour 1987). Most usefully, Albert Cairo provides us with the notion of "truth continuum" (2016), implying that there always must be some compromise.

Birgit Schneider (2018) provides a comprehensive overview of how the practice of empirical data science of the Enlightenment, in the sense of "Humboldtian Science" (after Alexander Humboldt), generates observable and measurable evidence in the form of numbers, which lead to the need for visio-spatial arrangement of statistics in graphic forms. Over time, the tools and conventions of data visualisation have come to stand semiotically for the idea of enlightened, modernist science itself: the bar chart, the line graph, the pie chart familiar to everyday culture. Philosopher Thomas Kuhn (1962) posed the idea that there is always a life cycle to what is considered to be "good science" and that this can enter a period of crisis, to emerge with a shifted paradigm, a new intellectual framework for investigation and scientific endeavour.

Twenty-five years later, Bruno Latour (1987) made waves with his statements around how to relativize scientific certainty. He recounted examples of different stages of genomic research – from Rosalind Franklin's photograph of a DNA string, to the model of a double helix by Francis Cricket and James Watson, which has become the predominant association. At this point, the uncertainty, personalities and environments that lead to the development of this knowledge are forgotten about, and the accepted model is no longer questioned.

Data Visualisation and its sister discipline Information Design have the often-stated aim and ideal of 'presenting information accurately, precisely, and truthfully'. The discipline's values align with ideas of rationality and objectivity – favouring reason over emotion, and the graphic equivalent of this is usually a minimalist approach – trying to communicate only what is necessary, in order to

broadcast the intended message. Now, data visualisation has become an integral part of everyday news cycles; in fact, it has inspired a whole new discipline of 'data journalism'. It straddles data science, design and journalism and is mainly aimed at general audiences. In contrast, data visualisation created for scientists often relies on very complex conventions with a high learning curve, such as the box and whisker plot (Salmond, 2007). This is necessary to communicate complexities of the research with some detail, precision, but also the notion of uncertainty. The data visualisations published on news sites and posted on social media want to grab viewers' attention, and be immediately understood, and have enough impact to be shared – by a much wider audience with mixed levels of data literacy. This creates new challenges and debates around the notion of "truth" in the visualisation of information. Jeff Malpas proposes: "In the modern era […] truth is unproblematic insofar as the possibility of truth is unproblematic. The understanding of the world on the model of the scientific, calculative understanding does in fact presuppose the possibility of truth" (1992, p.287). He explains how we presuppose any claims of truth to be based on an understanding of the world that can be scientifically explored and managed and understood through mathematics. However, as has become clear over time, data and statistics is prone to bias (Noble, 2018), prone to gaps (Criao-Perez, 2020) and in general has become conflated with the problem of misinformation and fake news. For all its virtue, it has to be acknowledged that it is entirely possible to create misleading data stories – either accidentally or on purpose.

Engebretsen and Kennedy (2020) make the case that data visualisation can be employed to facilitate understanding or persuasion, among other things – they are by no means the clean, value free, certain solutions that we tend to see in them. They tend to communicate complexity in a way that presents order, but the order is shaped by the people making the data visualisation according to the social conventions of the time. For data visualisation pioneer Edward Tufte, "every single pixel should testify directly to content" (PBS Offbook, 2013) which implies his aversion to anything that is merely decorative with no communicative intention, something he calls "chartjunk" (Tufte, 1983). He talks about the concepts of "goodness" and "truth" of information, which "beauty" might be a by-product of, but which cannot make up for weak, or even error prone content. Also discussing the tension between the modernist "nobility" of clear, minimalist information and overly decorative communication designed to make an emotional impact, Graphic Design writer Lucienne Roberts says the boundaries are fluid: "even the serious information-based publicity of many charitable organisations has succumbed to sugar-coated enticement" (2006, p.29).

Albert Cairo has engaged with these debates extensively in his writing on data visualisation, and he proposes the "Truth Continuum" as a way of thinking about these problems between truth and representation. He puts forward the argument that "Any visualisation is a model" (2016, p. 93) – in any visualisation, he states, some features and complexity of the data are necessarily removed, so that our human brains can cope with the information. We might think back to the example of the data flood produced by the observational practices of the Enlightenment. It is unrealistic to present

all data points at the same time. For Cairo, a better model for a visualisation means "more truthful, accurate, informative, and understandable" (Cairo, 2016, p.94). So, in this sense of offering a model, how can we understand the Covid-19 datavisualization presented here in this paper?

We will have to ask – does the visualisation sufficiently communicate scientific information? Where does it sit on the 'truth continuum'? Is it on the wrong side of sugar coated? First of all, for the more traditional data scientist, this may not be classed as a data visualisation at all, more like a pretty animation. In this particular prototype, there is no way to dig deeper into the data itself, which is often presented as one way of accommodating complex data and staying truthful: to present different views of the data and allow readers to explore the data set themselves. Close collaboration with the scientists to understand the data, as well as an awareness of the media discourse at the time, enabled the design thinking approach of "feeling" your way into the essence of the story to be told here. In this sense, the story was not set from the beginning, but was generated by understanding and developing empathy with the data, the context, the researchers and potential audiences.

References

Badshah, N. (2023). Matt Hancock wanted to 'frighten everyone' into following Covid rules. The Guardian. 5th March

Boyd, C. (2020). BioNTech boss warns mutated Covid-19 variant could make achieving herd immunity more difficult with vaccines - as he promises jab could be re-made within six weeks in unlikely scenario it doesn't work against the strain. Daily Mail. 22/12/2022. https://www.dailymail.co.uk/news/article-9078579/Is-mutant-Covid-ripping-UK-REALLY-70-infectious-Britain-just-testing-more.html

Cairo, A. (2016). The Truthful Art: Data, Charts, Maps for Communication. New Riders

Engebretsen, M. & Kennedy, H. (2020). Data Visualization in Society. Amsterdam University Press

Capra, F. (1996). The Web of Life. A New Scientific Understanding of Living Systems. New York: DoubleDay.

Chaszar, A., von Buelow, P., & Turrin, M. (2016). Multivariate Interactive Visualization of Data in Generative Design. In R. Attar, A. Chronis, S. Hanna, & M. Turrin (Eds.), Proceedings of the Symposium on Simulation for Architecture and Urban Design (simAUD 2016) (pp. 223-230).

Luo, C., Chen, A., Cui, B. & Liao, W. (2021). Exploring public perceptions of the COVID-19 vaccine online from a cultural perspective: Semantic network analysis of two social media platforms in the United States and China. Telematics and Informatics, Vol 65

Curralo, A. (2015). Generative Design and Information Visualization. ARTECH 2015. 7th International Conference on Digital Arts – Creating Digital e-Motions. Óbidos, Portugal

Criado-Perez, C. (2020). Invisible Women: Exposing Data Bias in a World Designed for Men. Vintage

D'Ignazio, C. & Klein, L. (2020). Data Feminism. MIT Press

Galanter, P. (2012) Computational Aesthetic Evaluation: Past and Future. In: McCormack, J. & D'Inverno, M. (Eds) Computers and Creativity. Springer, Berlin, Heidelberg.

Kuhn, T (1962) The Structure of Scientific Revolution. University of Chicago Press

Latham, William and Gollifer, Susan. 2013. Mutator 1 + 2: Evolutionary Art. In: "Brighton Design Festival", Brighton, 7 Sept - 13 Oct 2013. [Show/Exhibition]

Latour, B. (1987). Science in Action. Harvard University Press

Noble. S. (2018) Algorithms of Oppression: How Search Engines Reinforce Racism. New York University Press

Mach, K.J., Salas Reyes, R., Pentz, B. et al. News media coverage of COVID-19 public health and policy information. Humanit Soc Sci Commun 8, 220 (2021).

Malpas, J. (1992). Retrieving Truth – Modernism, Post-modernism and The Problem of Truth. In Soundings: An Interdisciplinary Journal, 75(2/3), p. 287-306, Penn State University Press

Mateijka, J., Glueck, M., Bradner, E., Hashemi, A., Grossman, T. & Fritzmaurice, G. (2018). DreamLens: Exploration and Visualization of Large Scale Generative Design Datasets. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI). Association for Computing Machinery, New York, NY, USA, Paper 369, 1-12.

PBS Off Book (2013). The Art of Data Visualization. https://www.youtube.com/watch?v=AdSZJzb-aX8

Peeters, M. (2016). Designing in liquid times: Generative graphic design in an age of uncertainty. Plots, 3, 20-33.

Roberts, L. (2006). Good: An Introduction to Ethics in Graphic Design. AVA Academia

Salmond, S. (2007) Taking the Mystery out of Research: Box and Whisker Plots. Orthopaedic Nursing, 26(1), p 33

Schneider, B. (2018). Klimabilder: Eine Genealogie Globaler Bildpolitiken von Klima und Klimawandel. Mattes & Seitz

Tufte, E. (1983). The Visual Display of Quantitative Information. Graphics Press