Alex Butterworth

Adventures in Zoomland: Transitions in Scale and the Visual Exploration of Historical Knowledge Graphs as Sequential Storytelling

Abstract: This chapter proposes a conceptual framework for the design of graphical interfaces to knowledge graphs that employ the concept of 'zoom', broadly defined to encompass levels of scale, resolution and abstraction – to enable exploratory forms of historical hypothesis formation and testing and of narrative elaboration. It does so through a process of illustrative and speculative reflection on a series of recent or current digital history projects in which the author has been involved that involve the semantic modelling of data representing sociabilities, mobilities, temporalities and identity and which seek to encourage a hermeneutic sensitivity to evidential uncertainty. Reference is included to methods of knowledge graph creation and refinement, both automated and participatory, while the extended concept of 'zoom' is located within its media and aesthetic history.

Keywords: exploratory interfaces, digital hermeneutics, knowledge graph, scalable data, historical narrative

1 Introduction

The increasing abundance of historical data in machine-readable form, the progressive refinement of methods for its analysis, and the popularization of innovative interfaces for its exploration, create the possibility of novel forms of historical narrative within digitally mediated environments. What follows in this chapter is a proposed framework for understanding and utilizing the concept of 'zoom' in the speculative design of graphical interfaces that enable exploratory forms of narrative history-making, which are driven by and embedded in semantically modelled knowledge graphs, grounded in formal ontologies. The framework is considered as a foundational requirement for the subsequent development of a narrative gram-

Note: With thanks to Ed Silverton, Andrew Richardson, Sarah Middle and Duncan Hay for stimulating discussions during the design and development of practical experiments in the application of the approaches to narrative data visualization described in this chapter.

https://doi.org/10.1515/9783111317779-002

mar in which the third meaning latent in transitions between modes of zoom applied to configurations of knowledge can be effectively managed, to controlled and expressive effect.

The concept of the *knowledge graph* will be expanded on in Section 2. Such graphs model historical data in a form which is highly plastic and fungible, enabling the virtuous creative cycle of exploration and contribution that is envisaged. In combination with associated domain specific vocabularies and taxonomies that further organize the data contained in the graph, the interrogation of such knowledge graphs may generate complex historical insights. The filmic technique of Zoom, considered more closely in Section 5, may be broadly understood as a range of operations that vary framing and focalization, here applied analogously to the visualization of data within an explorable graphical interface: the *Zoomland* of our envisaged historical adventures.

The conduct of historical enquiry and production of historical accounts that the Zoomland framework enables resonates with many aspects of the consideration of history-making offered by film theorist Siegfried Kracauer in his posthumously published work, *History: The Last Things Before the Last* (2014*)*. History, he observes, is constrained by a "law of levels", according to which, "the contexts established at each level are valid for that level but do not apply to findings at other levels". Yet, grounded in his insight that "discerning historians aspiring to history in its fullest form favor an interpenetration of macro and micro history", he identifies a productive equivalence with the filmic medium in which, "the big must be looked at from different distances to be understood; its analysis and interpenetration involve a constant movement between the levels of generality". Jolted by this movement across scales, the "historian's imagination, his interpretive designs" are freed from the canalizing effect of data overload, "inviting less committed subjectivity to take over" (Kracauer 2014: 76; 69; 70; 70). From this may emerge, Kracauer believes, "the historical idea" which, Rodowick notes in his critical consideration of Kracauer's thesis, "inaugurates a new terrain in which a wide variety of primary historical material may distribute itself and organize itself, illuminating previously unthought patterns of intelligibility" (Rodowick 1987: 117).

Kracauer's work was written at the dawn of an era in which the potential of filmic zoom to reshape an audience's relationship to the process of fabulation was fully demonstrated by a new generation of directors, including arguably the greatest proponents of the technique, Robert Altman and Roberto Rosselli. Adopting a fluidity of style that has been likened to improvisational jazz, Altman – according to Jay Beck – placed the audience for his 1970 film *MASH* "in an active spectatorial position, being asked to sift through the audiovisual information in order to follow one of the potential narrative paths" (quoted by Hall 2018: 26). A decade earlier, in his film recounting the nineteenth-century campaign for national independence, *Viva Italia*, Rossellini had already applied zoom techniques

to specific history-making effect. Commenting on a scene in which Garibaldi's men fight on a hill, John Belton notes that under Rossellini's direction, "Long shot equals then. Zoom shot equals now. The two shots in tandem are no longer limited to an imitation of event. What we are watching is our own aesthetic and ideological distance from the event" (Belton 1980/81: 24).

In recent years, the notion of the *Macroscope* has become established to describe a form of digital instrument by which the historian can similarly bridge and synthesise the big and the finely curated small data to which the digital turn has given rise. The benefits it may deliver have been variously conceived. For Tim Hitchcock, it makes possible the "radical contextualization" of individuals "by exploring the multiple archival contexts in which they appear or are represented"; for Armitage and Guldi it enables, a "weaving together [of data] into one inter-related fabric of time", while it allows Julia Laite, who quotes the preceding, to "construct a prismatic view of a complex phenomenon, to write a kind of history in the round [. . .] [to] massively expand the synchronic and diachronic connections between people, places, and experiences" (Laite 2020: 965; Armitage and Guldi 2014). The proposed framework seeks to address those desires, reapplying the concept of zoom to an exploratory data environment in which scale is collapsed and "historical ideas" generated. It is a framework that supports a range of manifestations of that negotiated data, from a relatively abstract rendering of data points as graphically arranged nodes to immersive experiences, with the evidential sources, in print or image form, directly accessible.

2 Uncertainty in the work of the narrative historian

Whilst rooted in the observed needs of diverse historians and projects of historical research, the Zoomland framework also foregrounds certain attitudes that I have adopted in my own work as a narrative historian: one whose aim is to present deeply researched historical accounts to a broad and non-specialist audience and readership. Among these is the aspiration to a sophistication in storytelling (or story-building) that equates to both the literary and the cinematic, with their vast but quite distinct toolboxes of tricks and sleights-of-hand and genres: techniques by which the relationship between percipient and story (Propp's *szujet*, at its rawest) is mediated and manipulated into fabular constructions and complex narratives (Propp 1968).

This aspiration could be seen as problematically contrarian for an author of traditional analogue narrative history, though perhaps more congenial in the realm of

digital history, since it involves a distrust of the very voice in which historical narratives are expected to be written for a non-academic readership. The nub of this unease is encapsulated for me in the editorial advice I repeatedly received from publishers to be 'magisterially authoritative'. The implication was less that an author should demonstrate a mastery of their subject, which of course entails a full awareness of the absences and lacunae and hermeneutic challenges, than that they should perform absolute confidence as a presiding narrator of the past. Tone, tense and genre were to be employed as mechanisms of gentle coercion: the use of conditional and subjunctive constructions should be eschewed, polyvocalism tightly managed, the knowable facts wielded as talismans.

Both books I have published were written in fierce but subtle tension with such imperatives, favoring implication over forthright assertion, cherishing ambiguity at their core. One book (*Pompeii: The Living City*) examined the social and political networks of a single, uniquely preserved Roman town as a microcosm for a period of imperial rule and Mediterranean trade and culture, interpreted through shards of epigraphical and material evidence; the other (*The World That Never Was*) looked at a late-nineteenth-century milieu of political conspiracies and subversion, in which the global diaspora that spread anarchist doctrine, and the international coordination of its policing, was discerned through the reports of informants to their secret service paymasters (Butterworth and Laurence 2005; Butterworth 2010). The books set out to prompt readers to engage with interpretive positions which might sometimes be uncomfortable or confusing: even leaving the reader suspended, temporarily, in a state of uncertainty, providing space for reflection on the evidential foundation of the narrative.

Techniques for highlighting uncertainty were important when confronted with subjects that exposed the margins of confident knowability, dependent as my books were upon the reconstruction of fragmentary and unreliable historical and archaeological evidence, and demanding of speculative boldness. These techniques are, I believe, even more essential as touchstones when grappling with the design of digital engagements with the past, which rest on data whose biases and absences require constant attentiveness. This engagement must remain open, exploratory and dialogic. In cases where exploration is heavily channeled by authoritative constraints, these will soon be felt as a deterrent by the user who craves the ability to pursue their own instincts of inquiry, frustrating their desire to probe beyond surface appearances and to test multiple configurations of data. Either that, or those constraints will have been so deeply insinuated as to co-opt the liberating process of enquiry itself to rhetorically manipulative ends. The responsibility in designing such engagements is therefore social and political, as well as scholarly.

The proposed framework conceives of digital narrative history as a mode that embraces dynamism, dialogism and multiplicity. Authorship in this environment inheres in acts of expeditionary trace-making performed by each investigator of the dataset. The *data text* (a state sequence, as discussed in Section 10 onwards) that is produced by the investigator may, as they prefer, be deemed ephemeral and discarded, or else preserved and enmeshed with those of others; it may be supportive of other trace-texts or it may contest them, in part or as a whole. The initiating author only proposes a path and rehearses the methods of its pursuit. These start with enquiry and analysis that is more or less informed by the existing knowledge of the investigator, both of the domain and the dataset, with each gradually accumulating. The investigator progresses to hypothecating and iteratively testing these hypotheses, then to the recording of insights, the presentation of argumentation, and its further elaboration as narrative. The narrative forms emerge from the steps and missteps of the investigation, each of which constellates the salient nodes in the knowledge graph, faceting their relations and elucidating their semantic significance. Any externally conferred authority is complemented by the collective recognition of skill in the dance of meaning-making, and the insights afforded by the traces that have been made. I will return in Section 7 to consider this interaction between the individual investigator and the collective acknowledgement of meaning-making.

3 The projects referenced and their datascapes

The conceptual framework proposed here derives from extensive and long-term investigations of user requirements across diverse historical datasets. It is illustrated by reference to a number of recent digital history projects, primarily three with which I am currently, or have been involved. Those three projects (and their associated datasets) are *The Lysander Flights*, *Tools of Knowledge* and *Crimes in London*. *Crimes in London* was an experiment involving the use of machine learning to explore the narratives of criminal activity and witness accounts in the Old Bailey criminal trials, and is distinct from the *Old Bailey Macroscope* created by colleagues from the Sussex Humanities Lab: a project which is also considered, as a key precedent for work in this area.[1] *The Lysander Flights* is a detailed study and exploratory narrative account of the operations in which the low-flying RAF Lysander aircraft flew secret agents and resistance leaders into and out of occupied France between 1940 and 1944: their conduct, context and consequences.

1 "Old Bailey Voices", accessed February 18, 2023, https://oldbaileyvoices.org/macroscope.php.

The dataset on which it rests is compiled from minute-by-minute operational reports, and records of personnel and strategic planning, the organization of resistance circuits, and the executions of resistance members. *Tools of Knowledge* concerns the creative communities involved in the production and use of scientific instruments in Britain during the four-and-a-half centuries up to 1914, and the instruments themselves as encodings and vectors of both craft and scientific knowledge.[2] At its heart is a large and meticulously curated legacy database, the product of a single curator's long career, which has been semantically remodeled and complemented with artefact data from multiple museum collections, as well as data from diverse sources that broadly contextualizes the entities involved: persons, places, objects, businesses and institutions.

Taken together, these projects present a spectrum of possibilities for complex investigative and narrative methods – covering sociabilities, mobilities, temporalities and identity. As described in Section 3, the datasets on which they rest are semantically modelled, both in order to capture the greatest possible nuance in the data and to enable interoperability with other datasets. Each entity mentioned above is defined as a node within a knowledge graph: each is potentially linked to every other entity type, as the data dictates. These linking edges are themselves additionally typed using controlled vocabularies to describe in detail the nature of the particular relationship, with each relationship characterized as an event. A social relationship between people might be defined as family, or business, or legal, for example, with a taxonomical account of finer grain relations such as husband or brother-in-law (family), partner, apprentice or employee (business), prosecuting lawyer or co-defendant (legal). Each entity node is further characterized by its attributed properties: nationality or gender, for example, in the case of people; type, size or material for an object in a museum collection, with additional relevant and domain specific attributes of *call sign* and engine type for a Lysander aircraft.

Wherever possible, the relations between entities are themselves modelled as events, which may involve actors (animate or inanimate, as individuals or assemblages), locations, activities and times. A crime is modelled as an event, as is a wartime operation, but so too are the atomized constituents of each of these: a button accidentally left at a crime scene, an encounter with anti-aircraft fire over the city of Tours. All of these entity nodes carry properties of their own, which include information about their evidential basis or provenance, and metadata about the reliability of the data. The automated geocoding of a placename against

2 "Tools of Knowledge, Modelling the Creative Communities of the Scientific Instrument Trade, 1550–1914", accessed February 18, 2023, https://toolsofknowledge.org/about/, (AHRC AH/T013400/1).

a gazetteer might have a probability of 78% and may or may not have been human-authenticated, a time might be *circa*, or represented by a span of uncertain start or conclusion ("born before", "died by"), or it might be a spot date standing in for a span (working life: only record, 1756). The qualitative and quantitative characteristics of the data vary, and with these their tractability to more abstract or concrete representation, with the potential for media that actualize the former being linked in to the datasets. This modelling of the historical data will be usefully born in mind in relation to what follows.

4 How to make a knowledge graph

Knowledge graphs are a form of structured knowledge representation, derived from potentially multiple data sets and sources, and modelled at various levels of refinement and authority, according to the interests of those who have contributed to their construction and content. Varied pipelines composed of methods and tools support more or less automated processes to extract the data and then transform and load it (ETL), or extract and load the data and then transform it (ELT): the sequencing of the phases determined by a combination of source type, analytical needs, and storage capacity for processing. The components of the pipelines range from Natural Language Processing methods – including Named Entity Recognition or Cluster Labelling approaches that may leverage word or concept embeddings to capture semantic nuance – through to semi-automated tools such as OpenRefine. At the time of writing the rapid emergence of Generative Pre-Trained Transformers and experimental tools such as GPTIndex promise even higher levels of automation.[3]

At its most refined, the resultant knowledge graph is aligned with a formal schema, which is itself reconciled where possible with published ontologies, including those for source citation: CIDOC-CRM, BIO CRM, VIAF, Fabio, Citeo, etc (potentially using OpenRefine's proprietary near-cousin, OntoRefine, or equivalent tools).[4] Beyond this, specific interest groups that contribute rich and densely modelled data, generated by them during a tightly focused project of research,

3 "Welcome to the GPT Index", accessed February 18, 2023, https://gpt-index.readthedocs.io/en/latest/.

4 "What is the CIDOC-CRM?", accessed February 18, 2023, https://cidoc-crm.org; Jouni Tuominen, Eero Hyvönen and Petri Leskinen, "Bio CRM: A Data Model for Representing Biographical Data for Prosopographical Research", accessed February 18, 2023, http://www.sparontologies.net/ontologies/fabio; https://sparontologies.github.io/fabio/current/fabio.html; https://sparontologies.github.io/fabio/current/fabio.html, accessed February 18, 2023, https://viaf.org.

may either embroider extensions to this core 'authority' schema, or devise novel ontologies for their own immediate interest. As Beretta has explained in relation to the symogih.org project, identifying those points of overlap and concurrence enable interoperability, and allow immediate semantic mappings to be made between the core graph's nodes and those populating its hinterland (Beretta 2021).

Stepping back one degree more in the graph, we find ourselves further distanced from rigid formalizations and edging towards the folksonomical and emergent: bottom-up terms and concepts, linked to the schema only by means of a 'fuzzy ontology' that accommodates them loosely. This material, which could encompass free-text comments contributed by Zoomland explorers or digitized historical texts or transcripts of audiovisual account, may in turn be computationally annotated, extracted, abstracted and modelled, using similar methods to those discussed. These processes can preserve the micro-graphs representing, in simple terms, a set of entities linked by their description of an event. At a higher level of abstraction, they may also capture those semantic vectors drawn together by the strength and density of their connecting edges to reveal concepts that have clustered according to some semantic affinity, within the high-dimensional space generated by the representation of the myriad word associations derived from textual sources. The folksonomical free text source may therefore be conceived of, in graphical terms, as loose constellations of nodes floating around the ever more densely modelled data towards the graph's core, spatially distributed as it would be by a force-directed layout algorithm. Constellations that are barely tethered at the periphery may be more closely integrated to the core, through the step-by-step linkage of nodes by means of newly minted edges. The tasks involved in the identification of these edges – as candidature for semantic interest and contextual relevance – may be performed, in turn, by a combination of machine learning methods and the contribution made by those human explorers of the knowledge graph who are drawn to assist in its development at a more artisanal level.

To most effectively support a virtuous cycle of contribution, the nature of the user's engagement with the graphical interface, by means of which they can explore the semantic content of the graph, should include exploratory reward and remotivation. The process of contribution should be well considered but relatively frictionless; whatever friction is encountered should result from an intentional choice on the part of the explorer to deviate into a more conscious engagement with the data, or else it should be algorithmically-generated as an injunction to heightened criticality, based on an assessment of the validity of the contribution. Within modes and layouts, however, the operational mechanism for informational and cognitive adjustment is here conceptualized by analogy with cinematic 'zoom', as it has variously been deployed.

5 A media archaeology of zoom: Digression and immersion

So, what is meant here by 'zoom' and where does the concept sit in a hermeneutics of data exploration and emergent digital storytelling? In traditional cinematic terms zoom is, at its simplest, a change in the focal length of the lens which, in contrast to a camera move towards a subject, retains the relative size and significance of the subject and background, while narrowing the view-shed. It signals attentive interest, but is only a tentative step towards recentering that interest which a camera move would have advanced a step further, and with a subjective focalization offering the most complete recentering, by means of an edit or *cut*. Whatever the zoom's velocity and extent – a short *whip* or deep *crash*; *buried* and contemplative – its effect is to generate a subliminal sense of provisionality. Like the suspended moment of evidential uncertainty mentioned in Section 3, it offers an instant of self-reflection, in which the viewer wonders where their interest might next be led: whether inward to the soul of the subject, or outward to contextualize its situation, or even displaced onto some other subject which will be viewed relationally to it. Such is the potent effect of Zoom in cinematic form on the spectator.

Fixed in an authored audio-visual or graphical sequence, the speculation engendered by the zoom as by any other composition or camera move, according to the skill of its construction and deployment, propels narrative engagement. For us, in seeking to define the design requirements for an exploratory Zoomland, it might translate into a moment that informs and compels a more active expression of agency in what the next configuration of visual information will be. However, for such progressive choices to be most appropriately informed, the fixed relationship of subject and background that the zoom ensures may benefit from inflections which at least hint at context, or even by previews of those next steps and what they might reveal. The experience here might be conceived as similar to the liminal zone of peripheral vision, where an apprehension of movement or change in light intensity is sufficient instinctually to prompt further investigation. So, for example, the visualization which, on completion of its putative zoom, would switch fully from one faceting of the knowledge graph to another – to a looser temporal granularity, perhaps, or tighter geospatial resolution – here surfaces only those entities algorithmically predicted to be most salient to the explorer. How entities are selected for this proleptic function and how they are evoked will vary by the circumstantial configuration of the interface.

Although the most conventional understanding of zoom – as a tightening of view-shed and loss of peripheral vision – may confer benefits in certain Zoomland

scenarios, it should not preclude other applications analogous to the technique. Nor should the interactive application of zoom be inflexibly reverential to its origins in cinematic or televisual media, where its practical use has been limited and often modish. Rather, the core mechanisms of zoom may be beneficially modified for this novel grammar of Zoomland storytelling in combination with other techniques for the manipulation of the cinematic image and their cognitive effects, such as dollying moves (in or out) that either complement or contradict the zoom. Changes in depth of field can blur or sharpen background or foreground, shifting attention between them or altering their respective resolution to suggest varying relationships or impressionistic affinities. The lenticular effects of pulled focus may also redirect us away from an individual subject to the social or collective: a notion developed, with different intent, in the theory of 'lenticular lenses' applied to the digital analysis of text.[5] By embracing an expanded definition of Zoomland's media archaeology, all these translated affordances of cinematic zoom become available for use in varying permutations, sequenced to produce expressively nuanced results.

Of equal relevance and complementary to such cinematic precedents, though, for current purposes, are insights drawn from the great theorist-practitioners of realist and modernist prose fiction: the potency of digression in Hugo and the layering of social and spatial networks; the densely evoked cognitive realism of Balzac, with its accumulation of persuasive detail; Tolstoy's eagle eye that discovers relevance in and between Napoleon's cold and the chaos of the battle of Borodino, as Carlo Ginzburg has referenced with regard to the macroscopic ambitions of micro-historical practice (Ginzburg 2012, 209). All are suggestive of principles for novel modes of storytelling with the historical data, which will refine the ideal interface design framework envisaged. This is particularly so for the more abstract levels of zoomed visualization at which it is the identification of patterns in the data that approximate to contextualizing story beats. The distribution of pickpocketing in eighteenth century London viewed in relation to that of the watch houses that facilitated its policing, for example; or audits of stock held by an instrument maker, derived from probate auction catalogues, with the dimensions of instruments, individually and collectively reckoned against the likely square footage of the display and storage space in their commercial premises; or the intensity of a Lysander pilot's flight record correlated with pilot-related accidents.

Examples such as these may be evocative as well as analytically revealing, even when encountered through the abstract visualization of relevant data. How-

5 Al Idrissou, Leon van Wissen, and Veruska Zamborlini, "The Lenticular Lens: Addressing Various Aspects of Entity Disambiguation in the Semantic Web" Paper presented at Graphs and Networks in the Humanities, Amsterdam (3–4 February, 2022).

ever, to envisage the framework for an exploratory historical storyworld exclusively in these terms would be negligent, at a moment when the dominant media and genres for storytelling are themselves immersive: in their photo-realistically rendered virtuality, their kinetic preoccupations, and often in their exploratory and constructionist affordances too. Indeed, the digital capture of physical objects and environments as point clouds, whether through photogrammetry of instruments that are representative of those from probate auctions, or laser scans of a crashed Lysander's cockpit, or extruded maps of historical London ground plans, raises the prospect of further negotiations between data models, of quite different scales and semantics. The prospect afforded is one that makes digitally manifest Chekhov's widely paraphrased and abbreviated advice to the short story writer: "Don't tell me the moon is shining; show me the glint of light on broken glass" (Yarmolinsky 2014: 14). In such cases, zoom could serve an important mediating and interpretive function, bridging the concrete and the abstract, to simulate the situated immediacy of lived experience, albeit in a form that must be responsibly hedged around with caveats and correctives. So, whilst the suggested framework does not countenance the integration of historical data exploration and ludic environments, I would argue that it should encompass notions of both "playability" (with dual senses of "play": as rehearsal without formal constraint, and of the scope for minor movement within complex structures or in mechanical motion) and of immersion.

6 The idea of scale – the three principal axes

Before progressing further through the speculative design of frameworks for Zoomland visualizations, it will be useful to consider briefly the most familiar affordance to which principles of zoom are applied: that of scale. At every turn in their work, the historical researcher must confront the question of the most appropriate scales at which to scope their subject, or into which their subject may usefully be decomposed. At what level of scale in the data available to them is it possible to detect change and assess the range of causal effect, and with what confidence, as determined in part by the quality and representative claims that can be made of that data. Following directly from this, and further complicated for historians working in a primarily narrative mode, are questions regarding the methods by which any decompositions can be recombined, and the validity of the insights produced by those recombinations. Zoom, in a straightforward conceptualization, allows experimental movement and comparison between the general and specific, the whole and the component. For the conscientious expert historian

its products will prompt not only further questions and hypotheses about how each informs the other, but also an implicit problematization of those relations and the assumptions that underlie them.

With its remit to rehearse investigative expertise for a broad public, and to cultivate their capacity for such data exploration, the Zoomland interface must go further than merely making visually accessible such straightforward zoom-in-scale. It must additionally surface these processes, reifying the dialectic interplay between them through the presentation of data that is faceted as contextually apposite, at any point. The scales of data involved in the exemplary projects discussed in this chapter, considered in terms of volume and complexity of modelling or detail, exemplifies a broad range of possible historical datasets. As such they also illuminate issues in the management of faceted data visualization and how these mechanisms may serve to bridge between one scale and another: as understood most directly in terms of spatial and temporal data, and within the graph of actor-relations. Furthermore, they should accommodate varied modes of critical interrogation: visual (pattern-matching), speculative (causal hypothesis-forming) and emotional (situated).

To address these requirements, the proposed framework for zoomable exploration is conceived as comprising three principal *axes*, each interacting with a continuous scaling, calibrated around three recognized *scales* of study (micro, meso and macro) and three core *modes* of visual organization (spatial, temporal and network). The three axes which variously characterize the conceptualization of zoom are defined as: (1) Level of Detail, involving the nature and amplitude of information appropriate to the current zoom level; (2) Level of Abstraction, from statistically evaluative to situated and immersive, or critically reflective to emotionally co-opted; (3) Cognitive, by which the depth and faceting of contextual information is focalized. A fourth potential axis, concerning authority and certainty, will be mentioned additionally in Section 12. 'Level' is used here, in relation to all axes, to describe the level of zoom that is set at any moment, along the range of possible settings.

Zooming on the Level of Detail axis, calibrated for informational resolution, is analogous to the illusionistic methods by which the developers of immersive environments switch in and out images or models, as the zoom reaches different pixel or polygon counts, as imperceptibly as is technically possible within a smooth animation, to avoid the lag that may be caused by loading overlarge image files relative to available processing power. It may also be imagined as an extrapolation of the methods employed on zoomable digital maps, where the most zoomed-out view emphasizes topography and transport infrastructure, for the purposes of long-distance route planning; while the most zoomed in exposes infrastructure for pedestrian way-finding; with an intermediate level which highlights resources or points for interest that may localize the interest of the former user, or expand the

horizons (or advertise to the consumerist needs) of the latter. In this respect, it begins to intersect with Cognitive Zoom.

This third axis of zoom, Cognitive Zoom, has the greatest novelty as a principle for visualizing data and also involves a greater complexity of dynamic calculation in transforming the graphed data according to rules customized for individual datasets. It may be understood most straightforwardly as the constellation of context that is deemed most revealing at any moment in relation to the immediate objective of the exploration. Or in presentational terms, the context that the viewer needs to know at this point on their narrative journey to make fullest sense of what follows. Its close antecedents lie in the work of Marian Dörk, his formal conceit of the "information flaneur", and the visualization styles prototyped in his work on monadic exploration. This draws on Tarde's concept of the monad to challenge the distinction between whole and part, while more recent collaborative work progresses to the concept of the 'fold', the latter elaborating the former as a topological model that exists in a continual state of dynamic folding of space, movement and time (Dörk et al. 2011; Brüggemann et al. 2020). For present purposes, though, the concept is animated rather through reference to filmic technique – the dollied zoom – by means of which the central object of interest is simultaneously approached focally and actually, with the foreground held steady while the background is advanced or, in this case, made informationally salient.

Level of Abstraction is probably the most readily understood as being, at one extreme of the axis, photorealistic immersion whose verisimilitude produces a sense of situated presence, potentially involving haptic, tactile and olfactory senses in addition to the visual. It is, or might be conceived, crudely, as the ultimate historical theme park experience. More difficult to envision is the other extreme of the axis where the information contained in the knowledge graph into which the dataset has been modelled is expressed at its rawest: a dense tangle of nodes and edges that exist only in a condition of illegible high-dimensionality. As discussed in Section 12, for whatever purpose this is thought tractable to human interpretation, this multivariant data must be reduced to the simplest form at which it can be visualized, in two or three dimensions (or possibly four, if some means of filtering by time is included). Between these poles of the axis lie innumerable possible calibrates: progressing away from the highest point of abstraction, at each of these the visualized dimensionality of the data is further reduced, by means of the application of algorithmic or graphical filtering. Conversely, we may approach this from the concrete extreme of the Level of Abstraction axis, by which the relations concealed in the data that drive the immersive simulacrum with which the participant is first presented are incrementally decomposed for analysis. Here again, intersection with the axis of Cognitive Zoom is likely to occur, whilst intersection with the Level of Detail axis of zoom is also possible.

7 How authority is conferred, indicated and deployed

For the user, the drift towards Abstraction in this interplay of axes is liable to entail a gradual weakening of those points of intellectual and imaginative purchase derived from immediate human experience, and a drift towards ever higher levels of abstract conceptualization, that invoke and require more scholarly domain knowledge. An immediately comprehensible example of this might relate to time and human memory. The experience of a day or a week or a month or a season is familiar to a contemporary reader or data explorer and may be tacitly invoked to afford points of identification with historical experience. The temporal divisions, rhythms and functions of these time spans are, of course, often radically different according to context, historical and otherwise, as reflected even in their naming: the exigencies or planting and harvest, of economic behaviors dependent on sailing seasons, of *fasti* and *nefasti* or feast days and fast days, of lives lived by daylight or by standardized railway time. Nevertheless, the shared experience of time at these scales can provide imaginative purchase on that difference, which an author may more or less explicitly evoke.

It is this potential to analogize, however approximately, which is drawn on by techniques such as the graphical distortion of an otherwise uniform segmentation of time to communicate more subjective temporal experiences – intensity of activity, the anxiety of waiting – in relation to timeline visualization.[6] Career courses, lifespans, even generational shifts are similarly amenable to an analogizing approach, to produce an immediacy of narrative engagement. In static forms of data visualization, convention insists on regularity in the relationship between any property and its spatial distribution or proportional scaling, with any deviation from this rule considered as irresponsible rhetoric which risks misleading the interpreter. Such techniques may be afforded a different legitimacy, however, where the information space is itself recognized as a dynamically expressive environment. In this case, the relationship between background and foreground can itself be experimentally adjusted by the configuration of zoom settings, on any axis, while comparison between sequential animated states makes apparent any rhetoric devices. As when toggling between maps that represent geographies in Euclidian space according to conventional projections, and cartogram representations that morph according to the geospatial distribution of additional data, the agency of the ex-

6 Alex Butterworth, "On the growth of my own mind: Visualising the creative process behind Wordsworth's autobiographical epic, 'The Prelude', in context", accessed February 18, 2023, https://searchisover.org/posters/butterworth.pdf.

plorer and the perception of narrative development converge to generate imaginative engagement and suggest new interpretations (Dorling 1996).

The effect of such techniques becomes increasingly etiolated, however, as the temporal duration expands. When we reach the level of historical periodization, of longue durées and epochal change, an effective historical hermeneutic is likely to depend on specialist scholarly frameworks, often particular to a field or domain. The application of these collectively tuned knowledge-producing practices may reveal significance, suggest hypotheses and even bring to the surface provisional narratives in the patterning of data which to a layperson will remain mute and perplexing. The design of visualization layouts at such a level of zoom on the Abstract-Concrete axis must, it could be suggested, primarily serve the requirements of these more sophisticated users, even at the risk of presenting a complex and deterrent appearance to the non-expert. A visual form that is as simple as a scatterplot, such as that produced by Hitchcock and Jackson for their *Old Bailey Macroscope*, with its axes representing number of words per trial and sequential trial references numbers, is interpretable, even in the most basic terms, only with implicit knowledge of file naming conventions (and their relationship to chronologies, at micro and meso scale), as well as changes in methods of court recording or relevant legislation. A deeper analysis of change in trial length over a period of centuries might require more "microscopic" knowledge of courtroom architecture and its effects, and/or an instrumentalized understanding of the impact on judicial practices of changing carceral regimes or the imperatives of the political economy.

An analogue author of a narrative history can more easily be inclusive. In the course of ostensible story-telling they may educate the reader almost imperceptibly, skillfully seeding and integrating the necessary information through subtle and contextually apposite digression or allusion. Interfaces to data-driven zoomable analysis can and should seek similarly to mitigate the obstacles to properly informed interpretation by the broadest possible range of users. It is something that Hitchcock and Jackson's macroscope attempts, generously and engagingly, in relation to the example of courtroom architecture mentioned, through the provision of a three-dimensional visualization of the Old Bailey courtroom (styled for one historical arrangement, although potentially in multiple period-accurate manifestations).[7] Within this virtual space, users may position themselves to emulate and gain analytical insight from the experience of witnesses, lawyers, jury, judge or the accused as they were spatially situated in an environment that em-

7 Tim Hitchcock, "Big Data, Small Data and Meaning". Historyonics, 9 November 2014, accessed June 17, 2023, http://historyonics.blogspot.com/2014/11/big-data-small-data-and-meaning_9.html.

bodies and demands an almost theatrical performance of roles from those participating in the courtroom proceedings.

As will already be apparent from these relatively simple examples of the highly varied states into which data can be configured, all of which would conform to the framework for a Zoomland info-graphical environment, there is a high risk that the data explorer will experience both visual and cognitive disorientation, unmoored from reassuringly familiar conceptual models. The simultaneous availability of multiple options and functions for changing position, layout, filters or parameters is a necessity for freedom of enquiry but offers no firm footholds. It is therefore imperative that the explorer be best equipped to make meaningful progress, which here implies a forward movement through sequential meaning-making. Which bring us to the principle of the *subject-in-transit,* introduced here as a necessary complement to the earlier described principles of zoomable axes.

8 The subject-in-transit

The principle of the subject-in-transit proposes a mechanism for the multi-factorial means of ensuring a situated focalization while exploring the Zoomland datascape of an historical knowledge graph. The subject-in-transit comprises a single-entity node – a person, object, place, moment in time – or could involve a set of nodes which constitute and are contained within an event entity. This node provides a persistent object of attention travelling through varying states of data selection (by means of filtering and faceting) in combination with a range of graphical layouts: until a deliberate decision is taken to transfer focus, the node retains its identity through any transition in graphical presentation. Different permutations of selection and layout, which may be adjusted experimentally or with particular intention, enable varying interpretive possibilities to be discovered in the graphed data. Nodes may be laid out according to time or spatial properties, or in a network form, with each subject to rules of reorganization triggered by passing designated points in any of the axes of Zoom.

In simple terms, nodes might stack or aggregate on a Level of Detail Zoom, with the aggregation determined variously by polygonised proximity (map), temporal subdivision (hours, days, months, years, decades), or different gauges of network clustering and affiliation (membership of a guild, apprenticeship to a master; frequency of association with a particular crime or area of London; involvement in a resistance network or military organization). On the Cognitive axis, nodes linked to the subject-in-transit in the graph – directly or indirectly,

though only via edge types which are dynamically determined to reflect contextual semantic relevance to the current layout – may be exposed and constellated to suggest new analytical perspectives for the user, with the degree of allowable graph traversal determined by the adjustable Zoom level that is set. In this way, related people, ideas or objects (spatial/network) or locations and events (spatial/temporal) are displayed, to a greater or lesser extent, and according to further dynamically determined principles of relevance. The operation of Zoom on the Abstract-Concrete axis cross-references the type of data and associated media that are available with the layout type to define the mode of presentation: in spatial layout mode, the zoom might run from organizing nodes into an abstract bubble graph visualization according to type (country when in combination with a low Level of Detail, parish when high) through to the vertical extrusion of a historical map underlay at one extreme of the axis, or at the other, an explorable 3D representation of the street in which the subject-in-transit was located at that moment in historical time. Matrices that define how nodes manifest visually in each permutation or Layout, Axis and Zoom level are pre-configured in the course of data preparation and management.

Such an approach is consistent with a mode of historiographical knowledge production that seeks to induce reflexivity, and which embraces *zoom* as a means to engage and shift affective identification. It enables more nuanced engagement with the polyvocality, uncertainty and ambiguity generated by the tensions across scales of data, while enabling narrative devices that may challenge the user-actant with the unrecognized biases encoded in their exploratory choices. Drawing on Marian Dörk's theorization of the "information flaneur" as an imagined percipient casually constellating their information environment, the principle of the subject-in-transit emphasizes instead the chosen ego node within the knowledge graph as a sustained guarantor of coherence (Dörk et al. 2011).

Hitchcock and Jackson's *The Old Bailey Macroscope* has already sketched an application of the principle, although not explicitly elaborated as such or grounded in a knowledge graph.[8] Within the *Macroscope*, the user is able to switch between layout views while maintaining a persistent focalization on a single subject-entity, in this case the trial. Consequently, the hearing that has been presented in immersive form may also be identified and highlighted in the scatterplot, or vice versa, the 3d courtroom may be accessed via the selection of a single scatterplot glyph. It may also be examined, singly or comparatively, as a sequence of standardized courtroom discourse elements, derived from a dataset created by Magnus Huber and here presented as a sequence of strips on a horizontal bar graph, color-coded

8 "Old Bailey Voices", accessed February 18, 2023, https://oldbaileyvoices.org/macroscope.php.

by discourse type and linked through to the transcript of the trial.[9] Whilst sometimes available in a dashboard form, meaning is inherent more in the counterpoint of layouts, yet the transition between layouts is abrupt rather than fluid. That is, when compared with filmic grammar, it aligns more closely with a "cut" than an animated zoom, although the shifts in scale certainly embody something of the latter effect, albeit with a loss in selectively carried over context. *The Old Bailey Macroscope* therefore offers a highly revealing prototype and illustration of zoom, notably on the abstract-concrete axis, with two distinct explorable visualizations of data that sit respectively towards – but not quite at – its two extremes.

These identifiable lacunae in the *Old Bailey Macroscope* are, however, themselves usefully suggestive of challenges and opportunities. The scatterplot layout positions its visual representation of data on axes that carry human-readable meaning, albeit requiring implicit expert knowledge to interpret, while the 3D courtroom scenario intentionally eschews (even alienates) the user from a sense of immersion: the space is not "skinned" in surface detail, the human figures are near-skeletal maquettes; the voicing of the trial is machine-generated, uninflected and not distinguished by speaker. It is a project that leaves the dramatization to the television fictions that it has informed. However, the extreme of the abstract-concrete axis might allow for further progress in that direction, which in turn could activate a data-orientated negotiation between the levels of evidential certainty (Section 11) and the speculative affordances involved in immersive evocation. Moving from the 3D courtroom, along the axis in the other direction, towards abstraction, it has not yet tested either the territory of meso-level abstract-concrete zoom where, for example, the visual interface might allow for the experimental correlation of trial duration data directly with modification of courtroom design, expressed perhaps in more diagrammatic terms of space syntax and topic-annotated graphs of courtroom discourse. Nor does the *Old Bailey Macroscope* venture to the further extremes of abstraction, of the underlying multi-dimensional knowledge graph, in whose visualization even the guide ropes of expert knowledge must be cast aside in favor of the emergent, if imminently inchoate: one whose analytical potential will be touched upon in Section 12.

9 The data finds its shape within the framework

For the three projects which I have discussed as exemplary case studies for a prospective Zoomland approach, however, the balance of considerations is different.

9 http://hdl.handle.net/11858/00-246C-0000-0023-8CF, accessed February 18, 2023.

In two of these, the scale of the data is somewhat or substantially smaller, in terms of the total sources, indicating the aforementioned focus on meso and micro scales of zoom on most axes, but the detail of the modelling substantially greater, while its semantic grounding amplifies the structure of the graph itself, as a carrier of meaning. In the third case study, coincidentally developed from the Old Bailey dataset, the data in question is derived from textual analysis at scale and is loosely semantically modelled, and for purposes of zoomable analysis, roughly compatible with that of the other two projects. The effect envisaged is to increase the amplitude of analysis along the Abstract-Concrete axis, most notably in combination with Cognitive Zoom, by which the Subject-in-transit can be contextualized.

A recent and highly effective use of Virtual Reality for historical evocation, consonant with the Second World War in the air subject matter of the *Lysander Flights*, was *1943 Berlin Blitz*, produced by the BBC.[10] The experiment is situated in a Lancaster Bomber as it flew in formation through heavy flak to rain incendiary devastation on the German cities below. The experience creates a powerfully immersive visual and physical effect, with its emotional impact intensified by the use of a voice over contemporary report by Wynford Vaughan-Thomas, a radio journalist who accompanied a mission with his sound engineer. For fifteen minutes, in an abbreviated version of the eight-hour journey, the lived experience of the terrified crew is reanimated. Its position on the Abstract-Concrete axis is clear, with the voice of the reporter, equipped with contextualizing information (carefully censored and skewed towards propaganda), fulfilling something of the function of Cognitive Zoom. How though might Cognitive Zoom, deployed at a more abstract point on the axis, generate an equivalent intensity of understanding and even empathy by revealing patterns in the data that evoke the inner world and thought processes of subject's situated experience?

The subject-in-transit here might be, by closest analogy, the pilot of the Lysander aircraft on an equally hazardous operation to infiltrate agents into occupied France. We might be introduced to the interiority of the experience by visualizing data that reveals the intensity of their schedule of sorties, how much recovery time they have enjoyed between, or how much leave they have taken in the last month, from which we might derive the imagined level of stress and fatigue; or the same question might be approached through how many crews' and colleagues' names they have watched wiped from the blackboard in the previous forty-eight hours of 1943 as casualties on similar missions (by way of reference to a common visual trope of combat films).

10 "1943 Berlin Blitz in 360°", accessed February 18, 2023. https://www.bbc.com/historyofthebbc/100-voices/ww2/360berlin.

Or it could be a criminal defendant, or an instrument maker, or even a scientific instrument. What, we might ask, are the chances of conviction for a girl of sixteen brought to the Old Bailey in the 1760s for the theft of a gentleman's handkerchief in a theatre, what is the likely sentence they will receive, and in what language are they most likely to hear themselves described? Or for a young Italian immigrant maker of chronometers who has just won a silver medal in the Greenwich Observatory trials of 1854, coming in behind a prominent London-based company with a three-generation lineage of supplying the admiralty, how many chronometer makers who have been in a similar position to him still have a going business two decades later. In summary, is this a world hospitable to talented incomers and, if so, what are the secrets of success? Perhaps that is a question for which the answer is carried by an object rather than a person: a chronometer that has been in service for decades, was assembled from parts supplied with recognition by half a dozen artisanal enterprises, with the elegant housing made of an unusual alloy of unknown origin detected by X-ray Fluorescent analysis, was stamped by an eminent Edinburgh maker, and which had been regularly repaired by the same workshop twice a year, before its retirement from Admiralty service and acquisition by the National Museums of Scotland, its eminence secured by careful curatorial preservation and thanks to its regular exhibition. Yet how truly representative is it of the hundreds of similar instruments supplied to the Royal Navy's ships throughout the nineteenth century?

10 Narrative exploration as hypothesis forming and testing

For the data to be so revealingly articulate, of course, it must be sifted and sorted, filtered and faceted using compound Boolean arguments, and rendered legible to others as a set or sequence of related insights, graphically presented and, where necessary, commented with explanation. The initial experience of such multiaxial zoom environments is almost certain to be haphazard, even in the hands of an expert operator, and resistant to easy interpretation. Its navigation and the construction of meaning with it entails an iterative process of trial and error, of hypothesis formation and testing, and a responsiveness to unanticipated possibilities and unexpected insights. In terms of a design specification for a supportive graphical interface, it must encourage rather than inhibit speculative forays, with the operator confident of retaining or easily regaining their orientation.

Where a user-authored sequence of states departs from an existing argument or narrative account, with which it shares certain states or state-pairs in common – having started, by definition, with all in common – or where its structure

in development *ex nihilo* is algorithmically identified as correlating with significant states and state-pairs within such an existing account, the points and the extent of agreement or divergence between the new and prior sequences may also be graphically indicated. It is quite possible that such relationships will be discovered between the current authored sequence and *multiple* analogous but somewhat contesting accounts, each of which will offer its distinct interpretation, to produce further graphing of intersection and divergence. In such cases, the comparison that is graphically indicated at any time will be between the active account and a single correlated sequence, but with the option to cycle through the other sequences, for paired comparison. By these means, a dialogic approach to negotiating authority is encouraged whilst the legibility of difference is also ensured, by dint of managing its parameters. The sense of empowerment experienced by the explorer in turn prompts them to contribute additional evidential data in support of their narrative or argument.

The comparison of arguments and narrative accounts requires, of course, a state machine – actively recording states of layout preference and data configuration – in order to ensure that the branching paths of experimentation can be followed, backtracked, retraced and deviated from, in an arborescence of growing complexity but also redundancy. Ideally, this state machine management would include some means to mark those paths most favored, as the argument that they represent becomes more established, whether by authorial annotation or frequency of iteration: tracks through a network of states and edges that the interface may inflect with visual prominence, as in the erosion of desire lines in physical landscapes.

11 Variable authority as a further axis of zoom

To complement our three core axes of zoom, a fourth and distinct type is implied by this capacity for anticipatory probing of the data: one that runs from looseness to certainty, or from a dispersed to a hierarchical authority. The axis of zoom is, in this case, manifest within the graphical environment as a scalar of inclusiveness that can be adjusted – or, in more tangible User Experience terms, 'brushed' – to produce either a more generously broad, or a narrower and more tightly defined focalization of the interpreted graph. This conceptual focalization – which may also be considered as the amplitude allowed by the chosen depth of focus – is technically realized as the degrees of node-traversal in the knowledge graph, filtered for availability by their confidence scores, that that are enabled during the resultant query.

To this end, the (metaphorical) tuning-dial is calibrated to winnow data according to their provenance metadata, against two measures. The first is by type and level of authority, conferred by externally validated status. This might be variously by reference to a relevant publication with a Unique Identifier that has been validated by an authority source, or else by reputation acquired among the contributing community, for example by the collective validation of contributions, of which scored records are maintained internally. The second is conferred by the reliability of the interpretation of the relationship between data represented by that edge: calculated as a combination of the confidence score as assigned by contributors and/or as probabilistically determined in the machine-learning process.

Whilst an aggregation or disaggregation of nodes occurs during a zoom-in-resolution, a zoom-in-certainty is manifest simply as the addition or removal of nodes, as they fall within or without the acceptable range. The node and edges constellated in the micro-graph freeze and preserve user-focused interest at moments of transition between layouts and modes: concepts which are themselves explicitly modelled within the knowledge graph. However, at any time it should be possible for the user to surface – with minimal active enquiry (one or two steps of interaction) – the specific provenance data pertinent to the set of nodes and edges currently filtered, or to any individual node or edge visible, and to access this data either as an annotation or wherever possible, at one further remove, by direct linkage to source. Such information is then presented: displayed as text, perhaps, either in a floating pop-up window associated with graphic elements, or within a dedicated side bar window, or as speech, whether pre-recorded or synthesized using speech-to-text methods.

12 Conclusion: Towards a narrative grammar of zoomable exploration

Behind the framework for managing the fluid visualization of historical data in narrative form that is described in this chapter, lies a process grounded in the modelling and later querying and analysis of the knowledge graph. This involves, fundamentally, the translation of one graph form into another: from a knowledge graph formalized by alignment with formal ontologies that carry rich implicit knowledge, into contextually determined sub-graphs that are extracted and made accessible to the data explorer according to the specific requirements of the moment. These sub-graphs constellate ego networks which are centered on the subject-in-transit, with various allowed affordances of graph traversal and with exploration possible to

varying degrees of amplitude. Neither a general nor particular definition of the algorithmic substrate of such a graph translation system, by which the constraints are dynamically applied, will be attempted here. It will develop through accretion, as a product of the repeated practice of expert exploration and enquiry: rehearsals of human expertise that are manifest as the authoring of narrative sequences out of visualization states. This contribution of candidate edges and the new micro-graphs that they link together will be amplified through the application of machine learning methods, trained on those human-produced accounts, whether narrative or argumentative in form.

In conclusion, it is possible to note how this more sophisticated level of knowledge extraction and narrative modelling depends on abstraction in two axes, combined with concrete specificity in another. The abstraction takes the form of high-dimensionality vector spaces, on the one hand, representing the graph of semantic associations, in which a process of clustering by variable similarity or affinity of vector pairs – or, in more refined instances, of matrices – signify potential relevance. Within the Zoomland graphical environment, these might be inspected for utility and interest through the application of unsupervised non-linear dimensionality-reduction algorithms, such as UMAP or T-SNE, and may even be human-labelled *in situ* to augment the explicit record of the knowledge graph. Meanwhile, and by extreme contrast, Concrete specificity is to be discovered in the character of those more visually immersive states, ranging up to the level of three-dimensional photorealism. As will be apparent, a crucial role is played in this by the second axis, of Abstraction-Concreteness.

The interplay of these two axes may be effectively combined in narrative sequences, with the modes and layouts of each state determined by the specific permutation by the axial configurations. To achieve the desirable coherence of narrative construction, however, will additionally require the application of a narrative grammar of state sequencing, one whose definition might draw on research such as that by Neil Cohn into the 'structure and cognition of sequential images', in which the ordering and interplay of knowledge faceting has the potential to generate 'third meanings' as powerfully as more purely figurative imagery (Cohn 2013). These meanings may be apprehended analytically or in more purely affective terms, with the most skilled exploratory authors orchestrating their arrangements into compelling, informative and persuasive accounts. The anticipated outcome of this next step into Zoomland will be exploratory adventures that may stand comparison with the best linear forms of narrative history.

In the fullest realization of this half-prototyped medium, we can imagine Chekhov's moon seen through the eyes of an agent in the secret war of 1943 – a passenger in the rear cockpit of a Lysander aircraft, exfiltrated from a field in occupied France and huddled over a thermos of coffee laced with rum – as it reflects off the

silvery French waterways below: a living map by which the pilot is tracing a route to safety. That tangible journey, though, is also a trace through contextual knowledge: the spine of a story, dynamically rendered across the myriad synapses of a semantically modelled knowledge graph.

References

Armitage, David and Jo Guldi. *The History Manifesto*. Cambridge University Press, 2014.

Belton, John. "The Bionic Eye: Zoom Esthetics," *Cinéaste*, 11.1 (1980–81): 20–27.

Beretta, Francesco. "A challenge for historical research: Making data FAIR using a collaborative ontology management environment (OntoME)," *Semantic Web 12* (2021) 279–294.

Brüggemann, Victoria, Mark-Jan Bludau and Marian Dörk, "The Fold: Rethinking Interactivity in Data Visualization," *DHQ: Digital Humanities Quarterly*, 14.3 (2020).

Butterworth, Alex and Ray Laurence. *Pompeii: The Living City*. London: Weidenfeld and Nicholson, 2005.

Butterworth, Alex. *The World That Never Was: A True Story of Dreamers, Schemers, Anarchists and Secret Agents*. London: Bodley Head, 2010.

Cohn, Neil. *The Visual Language of Comics: Introduction to the Structure and Cognition of Sequential Images*. London: Bloomsbury Academic, 2013.

Dorling, Daniel. *Area Cartograms: Their Use and Creation*. London: Environmental Publications, 1996.

Dörk, Marian, Sheelagh Carpendale, and Carey Williamson. "The Information Flaneur: A Fresh Look at Information Seeking," *CHI 2011: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM (2011): 1215–1224.

Ginzburg, Carlo. *Threads and Traces. True False, Fictive* trans. Tedeschi, A.C and J. Tedeschi. Berkeley: University of California Press, 2012.

Hall, Nick. *The Zoom: Drama at the Touch of a Lever*. New Jersey: Rutgers University Press, 2018.

Hitchcock, Tim. "Big Data, Small Data and Meaning" Historyonics, 9 November 2014. Accessed June 27, 2023. http://historyonics.blogspot.com/search/label/macroscope

Idrissou, Al, Leon van Wissen, and Veruska Zamborlini. "The Lenticular Lens: Addressing Various Aspects of Entity Disambiguation in the Semantic Web" Paper presented at Graphs and Networks in the Humanities, Amsterdam (3–4 February, 2022).

Kracauer, Siegfried. *History. The Last Things Before The Last*. Princeton: Markus Wiener Publishers, 2014.

Laite, Julia. "The Emmet's Inch: Small History in a Digital Age," *Journal of Social History*, 53.4 (2020): 963–989.

Propp, Vladimir. *Morphology of the Folk Tale*. English trans. Laurence Scott. University of Texas Press, 1968.

Rodowick, D.N. "The Last Things Before the Last, Kracauer and History," *New German Critique* 41 (1987): 109–139.

Tuominen, Jouni, Eero Hyvönen and Petri Leskinen. "Bio CRM: A Data Model for Representing Biographical Data for Prosopographical Research," accessed February 18, 2023. http://www.spar ontologies.net/ontologies/fabio.

Yarmolinsky, Avrahm. *The Unknown Chekhov: Stories and Other Writings Hitherto Untranslated by Anton Chekhov*. New York: Noonday Press: 2014.