Using Data Visualisation to tell Stories about Collections

Stephen Boyd Davis
Royal College of Art
London, SW7 2EU
stephen.boyd-davis@rca.ac.uk

Olivia Vane
Royal College of Art
London, SW7 2EU
olivia.fletcher-vane@network.rca.ac.uk

Florian Kräutli
Royal College of Art
London, SW7 2EU
florian.krautli@network.rca.ac.uk

The paper explores visualisation of “big data” from digitised museum collections and archives, focusing on the relationship between data, visualisation and narrative. A contrast is presented between visualisations that show “just the data” and those that present the information in such a way as to tell a story using visual rhetorical devices; such devices have historically included trees, streams, chains, geometric shapes and other forms. The contrast is explored through historical examples and a survey of current practice. A discussion centred on visualising datasets from the British Library, Science Museum and Wellcome Library is used to outline key research questions.


1. INTRODUCTION

The work presented here is concerned with the visualisation of historic data, especially that held by museums and archives. It builds on our previous investigations of the history of data visualisation (particularly eighteenth century innovation), and our explorations of the potential and difficulties of creating new digital visualisations of history. We are interested in time as a graphic organising principle. We aim to develop interactive visualisations that are sufficiently sophisticated for use in investigating research questions, not simply in presenting history to a wider public (Boyd Davis et al. 2013, Boyd Davis & Kräutli 2015). We believe that some sophistication in representing history is called for even when designing for wider audiences (Kräutli & Boyd Davis 2013).

2. VISUALISING HISTORY

The task of visualising history can be conceptualised in several ways. Even a handwritten or typographic list is a form of visualisation (Twyman 1986); very early forms of writing include Sumerian king lists from the third millennium BCE (Corfield 2007:81). Figurative metaphors may be used to evoke the nature of history: trees, roots, streams, animals and statues were repeatedly used to ‘make sense’ of the past. As Rosenberg and Grafton (2010) have comprehensively documented, these two traditions, textual and figurative-metaphorical, have a long history. In the eighteenth century, they were joined by a new approach which treated historical time more consciously as a graphic dimension (Boyd Davis 2012). We have argued elsewhere that this took much of its inspiration from geography – the mapping of terrestrial space became the model for mapping historical time (Boyd Davis 2015a, 2015b). Events were now mechanically distributed in a space that was often of uniform scale, exemplified by the Chart of Biography (Figure 1) of Joseph Priestley (1733–1804).

The ancient figurative-metaphorical tradition was by no means overwhelmed by this emergence of the arithmetic-mechanical. The very neutrality of mechanical approaches seemed to some chronographers to diminish the ability to ‘tell a story’. They rejected such abstemious plotting of events in favour of richer visual forms. As Rosenberg (2007: 83) highlights, Strass (1766–1845) specifically objected to the flatness and neutrality of Priestley’s view, favouring a highly authored, hand-drawn grouping and linking of currents and tributaries in the stream of time (Figure 2). The later Chronological Chart of Sebastian Adams (1825–1898), a famous diagram issued in many editions even down to the present day, likewise exploits streams, along with other rhetorical graphics, in order to narrate key strands of history (Adams 1871).
Emma Willard (1787–1870) like Strass objected to the bare geometry of events distributed in time:

Mere straight lines not wrought into a picture, and presenting no form or comeliness to the eye, are unattractive. The young (and the old too) do not feel any wish to look at them, and thus they carry away no distinct impression. They are like a succession of monotonous sounds, which no one remembers; while the arrangement of sounds in tunes, or lines in pictures, are attended to with pleasure, and easily remembered. (Willard 1849:12n)

Willard offered four solutions. Her perspectival Picture of Nations (Willard 1836a) shows broad, serpentine bands representing kingdoms, empires and peoples which originate in the distance in 4004BC, and come towards the viewer’s own time in the foreground (detail, Figure 3). Her Progressive Geography of the World is a series of maps showing an ever-increasing area of territory bathed in light, from a small area of the Middle East in 1921BC to most of Africa, Asia, Europe and the edge of the Americas in 1492 (though always ringed by dark clouds), culminating in a cloudless complete world map on the final double-page (Willard 1836b). Her Chronological Plan of the History of the United States takes the form of a venerable tree (Willard 1844). Finally her Temple of Time (Willard 1846) is a perspectival view of a classical temple with a floor like the earlier Picture of Nations, columns to the right side representing rulers, and a ceiling divided into five bands containing Statesmen, Philosophers, Discoverers, Theologians, Poets, Painters, Warriors and others.

The use of light and shade augments the interpretation through drawing already exploited by Strass and others.

What should we think about such interpretation: what work does it do and what virtues were claimed for it? The advantages asserted by Willard are above all pleasure and memorability – the same advantages advanced by the authors of far more neutral graphic layouts such as Lenglet du Fresnoy (1674–1755), Barbeau de la Bruyère (1710–1781), Barbeu-Dubourg (1709–1779) and Priestley (1733–1804) discussed elsewhere (Boyd Davis 2015a, 2015b, 2016). Willard suggests that the distinctive perspectival arrangement of the Picture of Nations and Temple of Time assists the viewer in grasping our distance from past events (Willard 1849: 21). She discusses her own use of lighter shades to represent metaphorical enlightenment: ‘This broad light space, in which is placed the STAR [for Christ], represents the Roman Empire. ... Christianity, by the appointment of Providence, is now the religion of the most enlightened and civilized part of mankind’ (p.27). Though she does not explicitly say so, her interpretation constructs an image of a single unified nation (Schulten 2006): its ancestry lies in England (‘our mother country’) and Rome (‘our grandmother’), though she also acknowledges that ‘our grandfather-tribes’ include the Angles and Saxons who inhabit the dark regions of her diagrams (Willard 1849: 29).

Figure 1: Joseph Priestley’s 1765 Chart of Biography (detail). Time is mechanically mapped to a spatial dimension. Collection: Chetham’s Library, Manchester, used with permission. Photo: Stephen Boyd Davis.

Figure 2: Strass’s Stream of Time (1849 edition), detail. Mechanical neutrality is replaced by the rhetoric of hand-drawn rivers of time. Collection and photo: Stephen Boyd Davis.

3. VISUAL RHETORIC

A special feature of Willard’s work is her rhetorical use of light and shade. In the Picture of Nations, Lombards, Goths, Vandals and others lurk in the shadow of the Roman Empire. A starburst of light represents the time and place of Christ. Significant though lesser radiance is assigned to Columbus, the Reformation, and to American Independence.

Of course even Priestley’s abstemious dots, lines and labels are not neutral, something he is happy to admit; his Explanation of a Chart of Biography (Priestley 1764) is open about the kinds of choices he made (and was obliged to make for pragmatic reasons) in selecting 2000 names from all of
history. But the use of visual rhetoric in the charts of Strass, Willard and others of their school is – to use an appropriate phrase – another story. Instead of distributing data across a surface and leaving the viewer to ‘make sense’ of it, these authors tell the viewer formative tales, designed to influence their thinking along particular lines. At the same time, they still allow the viewer to see history ‘from a distance’, gaining an overview or perspective.

3. VISUAL RHETORIC, DIGITAL NARRATIVES

Computing mechanises knowledge. Typically when computing is applied to any domain its use promotes those formats that are most easily mechanised, while what is hard to mechanise is excluded. The arithmetic-mechanical designs of the eighteenth-century pioneers like Priestley lend themselves to computational methods, and many modern digital ‘timelines’ more or less resemble that same uniform mechanical visual style. It is the basis of almost all scientific visualisation. But are there also benefits in narrative-oriented visualisations? If so, how might they be constructed?

It is probably obvious that we do not want the kind of narratives offered by Willard. Whole peoples are expunged or marginalised; they are cast into darkness. One culture’s landmarks are taken to comprise history, the rest relegated to the status of background. But within historical data, narratives may be available at multiple levels, not just at the level of the presenting author. History contains narratives, such as those told by protagonists and by contemporary witnesses. If we could chart the particular, idiosyncratic viewpoint of Pepys’ experience of seventeenth century London through a more neutral landscape of events, what benefits might that yield? In the following century, if we were to trace the thread of evidential material suggesting that Harrison’s solution to the Longitude problem was the ‘right answer’ (as popular histories suggest), how would that compare with an alternative view that assigned status to the documents supporting the approach of ‘rival’ Maskelyne? And aside from these white male heroes, what other protagonists’ stories can be told? Would they enable viewers to negotiate alternative routes through the material, increasing their insight into the contentious and disputed nature of history? For a museum collection or an archive, what multiple, competing and conflicting narratives can be discovered – or created – by historians, researchers, archivists or non-experts? And what should such visualisations look like?

3.1 Some approaches to multiple narratives

Our questions echo the preoccupations of previous initiatives. In the 1980s and early 90’s there was a burst of interest in using digital technologies to dethrone unitary narratives (e.g. Kaplan & Moulthrop 1991), but this was principally focused on hypertext techniques and deliberately avoided making the connections between narrative chunks visually explicit. Virtual museum projects (e.g. Apple Computer 1992) imitated the appearance of physical galleries through which the viewer could wander, interrogating objects. Such immersion generally militates against a sense of perspective or overview; most immersive systems have also avoided the representation of competing narratives. Apple’s Guides project (Oren, Salomon, Kreitman and Don 1990), by contrast, was designed to present imaginary historical characters (preacher, miner, slave, etc.) who each led the user on a different route through material structured using HyperCard; each guide offered a distinctive viewpoint. Such approaches to digitally representing the complexity and contested nature of history are sensitively discussed by Beardon and Worden (1997). More recently, Nowviskie and Drucker (2003), Nowviskie at al. (2011) have championed subjectively inflected mappings of time and place.

The simple fact that any human-authored document represents an individual and inherently fragmentary point of view from within events, rather than an objective record from a presumed external stance, necessarily suggests that our counter- assumptions are integral to humanistic inquiry. Our challenge is to create a framework for graphically representing such elusive and subjective seeming phenomena. (Nowviskie and Drucker 2003:1)

This led to the Neatline project which ‘allows scholars, students, and curators to tell stories with maps and timelines ... it opens new possibilities for hand-crafted, interactive spatial and temporal
interpretation’ (Neatline.org). These tools have not been designed to make sense of ‘big data’, and the published examples emphasise place and geography over time and chronography. We discuss Neatline below.

4. DIGITAL STORYTELLING IN MUSEUMS

Digital storytelling is currently a popular term in museum practice (Wong 2015); however it is rarely used to refer to storytelling through data visualisation. Typical ‘digital stories’ from the museum sector, such as Outside the spacecraft from The National Air and Space Museum of the Smithsonian Institution, Discovering Literature: Romantics and Victorians from the British Library, and Monet and the Birth of Impressionism from the Städel Museum resemble long-form articles with video, photography and graphics built in and around the text.

4.1 Museum storytelling with timelines

Where a timeline is used to structure digital media (such as on the Mucha Foundation and Anne Frank House websites) the viewer’s experience is similar to that from museum ‘digital stories’. Size of images and of other content may be used to differentiate between the immediate story and the broader background, and the timeline may be broken into periods such as “Breakthrough in Paris, 1895–1903” (Alphonse Mucha Timeline). The digital media content is, however, laid out in a more or less horizontal linear way. There are no narrative threads made visual, and there is no possibility to step back from the details, to zoom out and view a history in overview.

A departure from this model is MOMA’s Mapping Artists’ Lives (Figure 4). This interactive map with timeline below allows the user to choose one or more artists and then watch on a time-lapse as the places connected to that artist by home, work, exhibitions, publications, and travel are highlighted with coloured circles. A user can observe as circles appear over time, marking significant locations, and can see where the lives of two artists intersected by watching to see if the differently coloured circles overlap in time. Additionally, the user is able to pause the time-lapse at any point and to adjust the timeline’s lower and upper date markers to see the accumulation of highlighted locations within a chosen timespan. In order to access greater detail, clicking on a coloured circle brings up a separate chronology of events connected to that place to the left of the map. Coordinated displays over time and place were earlier used to represent the Itinerary of King John, built with the MIT Simile timeline tool and Google maps API.

While Mapping Artist’s Lives better enables investigation of the data than the linear timeline stories, there are shortcomings to its power as an analytical and exploratory tool. For instance, it is not possible to easily compare the locations of artists’ activities in different timespans – there is only ever one view visible at any time.

Neatline (Nowviskie et al. 2012) has similar usability limitations for making comparisons. While the user can review the content in any order they choose, it is not possible to see different map views side-by-side or overlaid. These tools construct narrative with sequence and ordered content but with the downside that it is difficult for the user to spot general patterns and trends, or interesting gaps in the data.

4.2 Timelines overviewing narratives

Tate Modern’s large-scale interactive wall Timeline of Modern Art achieves a sense of overview. It is an updated interactive version of the static Modern Artists’ Timeline (Sara Fanelli for Tate Modern, Figure 5), which has similarities with Priestley’s Chart of Biography and The Genealogy of Pop/Rock Music (Tufte 1997).

Figure 4: MOMA’s 2015 Mapping Artists’ Lives. Screen capture, used with permission. Credit Second Story, part of SapientNitro.

Figure 5: Sara Fanelli’s 2006 Modern Artists’ Timeline for Tate Modern, detail. Used with permission.

The interactive Timeline of Modern Art is a large-scale touch-screen display (Figure 6). A stable horizontal line with date labels passes through the centre; around this, examples of artists and artistic movements float out from the centre line, appearing...
and disappearing automatically at intervals. If the user selects a visible name by touching it, further information appears as text and images.

Neither the static nor interactive versions use strict mapping of time for artists or art movements – names float in clusters along a horizontal time axis and there is no suggestion of birth or death dates, or of length of lifetime. For the purpose of these very accessible timelines that seems appropriate. They are visually attractive, and the educational value of being able to see where individuals and movements fit in the wider history is not diminished without this level of detail. In addition, artists' names are positioned to emphasise periods of related artistic activity and publicity – fuzzy time-spans at best. Perhaps for telling a story, detail in dates can be an unnecessary distraction?

![Figure 6: Tate 2015 Timeline of Modern Art. Image courtesy of帧store, Tate.](Image)

While the interactive Timeline of Modern Art is excellent for exploring and discovering artists represented by the collection, the lack of control over what data appears makes it difficult to make a focused investigation into the collection. The artists and art movements that appear seem at random and, rather than serving to bring out a narrative, the animated floating coloured dots are better suited to attracting spectators, and illustrating the scale of the collection. The topography of the collection over time is clear, but not its topology or any sense of connection and causality.

The Cleveland Museum of Art also has a large-scale interactive display (Alexander et al. 2013). The Collection Wall makes visible all the artworks on display at the museum (Figure 7). Similarly to the Tate’s, this interface performs attractive transitions/animations of the collection data. The organic shapes recall the visual metaphors employed in the historical examples discussed earlier.

While transitions offer different views and perspectives on the collection (including curated groups of objects organised around themes such as ‘Love and Lust’), time is not used as an ordering principle. Nevertheless, the Collection Wall suggests how organic transitions/animations might be employed in future interactive displays to tell stories about collection data.

![Figure 7: Cleveland Museum of Art 2013 Collection Wall. Photo: Cleveland Museum of Art. Used with permission.](Image)

5. INTERACTIVE NARRATIVE VISUALISATIONS – DIGITAL JOURNALISM

The scope for narrative visualisation in data journalism has been reviewed by Segel & Heer (2010), Hullman & Diakopoulos (2011), Hullman et al. (2013). While static visualisations have a history of supporting narrative in associated text, an “emerging class of visualizations attempts to combine narratives with interactive graphics” (Segel & Heer 2010: 1). This is the class of contemporary data visualisations where visual rhetorical devices are most observable, and it is this category, largely absent from the museum sector, that seems to offer most potential when applied to digitised cultural collections. These interactive narrative visualisations can represent multiple narratives “inviting verification, new questions, and alternative explanations” (ibid.) in a way that the linear progression of digital material in museum storytelling timelines or the single view of static timelines cannot. The following examples have been chosen because they employ the type of visual rhetoric used in traditional history visualisation.

5.1 Where We Came From and Where We Went, State by State

Where We Came From and Where We Went, State by State from the New York Times is a set of interactive time-wise visualisations showing the migration into and diaspora out of each US state 1900–2012. The y-axis maps the state population percentage with time running along the x-axis. The strands in the diagram represent the proportion of residents who have come from (or gone to) other states/regions. The strands are labelled and colour-coded by region. When the proportion of residents from a state/region is particularly small, the
designers chose to aggregate them across regions to improve legibility.

This visualisation allows the reader to perceive migration/diaspora flow patterns by state, and, when compared to the other graphs, discern differences in these patterns between states. The reader can get an impression of which states are more cosmopolitan and which more regionally insular. Each state has its own history, but epic trends are also made visible: European immigration in the late 19th and early 20th centuries; the mass migration of African-Americans to the North during WWII.

The organic visual language in these graphics makes them more approachable than conventional stacked area graphs. While the braiding is visually impressive and attention-catching, however, there is no clear logic or meaning to when and why the wavy strands cross. In addition while there are trends made visible in the data, they are by no means obvious just from the graphic; the accompanying text is essential to recognise the history represented by the snaking ribbons.

5.2 Over the Decades, How States Have Shifted

A final example also from the New York Times shows how American states have shifted their voting allegiance 1952–2012. Each box represents a state sized by the number of electoral votes. Curving streams follow the shifting path of a state’s voting allegiance as it swings between political left and right over time (vertical axis). Colour is used to highlight shifting allegiance to the political parties.

To track the history of a particular state, the user hovers over that box to highlight its full-length curve (though it is difficult to do this scrolling down the webpage). When no single state is highlighted, restricted use of colour draws attention to broad voting shifts, such as Republican voters for George Bush Sr. in 1988 shifting to Democrat candidate Clinton in 1992. As with Where We Came From… this is a visualisation of quantitative data and the narrative relies a lot on accompanying text – how might a designer build similar visualisations using collection data? And how might a designer adopt this type of interactive graphics to visualise qualitative trends over time?

A time-wise data visualisation by Florian Kräutli at the RCA employs a similar organic visual language. This visualisation automatically draws connectors between two datasets on separate timelines: poets whose writing Benjamin Britten set to music, and the resulting musical works (Figure 8). It offers relational information, topology overlaid on topography. Simple serpentine arcs appear in clusters as the user selects song-cycles or poets, allowing new insights into the pattern of Britten’s tastes.

6. DISCUSSION

We are currently working on data from the British Library, on Wellcome Library data including Medical Officer of Health reports, and on the Babbage Archive at the Science Museum. Our design criteria are based on several years’ research in this field. We intend our timelines to be readily understandable with little or no instruction, but they should not sacrifice sophistication in the process. We wish to represent the subtlety and indeed the problems of historiography, which we have shown elsewhere extend even to the dating mechanisms routinely used by museums and archives (Boyd Davis and Kräutli 2015). Doubt, error, controversy and other forms of uncertainty are endemic to history-making, and we have no desire to conceal these characteristics, which in our view are fundamental and indeed part of the interest of historical investigation. As indicated earlier, our chronographic designs should support inquiry as much as presentation. As we turn our attention to narrative form in relation to data, several questions arise:

(i) What form(s) should we adopt – what is the ‘vocabulary’ of graphics best suited to the task? What role can animation and interaction play?

(ii) To what degree can a story be brought out using computation? Do narrative visualisations always require a high level of authoring?
(iii) How can we support rapid apprehension from uncluttered displays, but still provide depth of information where it is needed?
(iv) What literary narrative devices can be translated into visual terms, and what does narrative theory have to offer?
(v) What forms of inquiry are best framed in narrative terms?
(vi) Who narrates? Can there be a balance between museums encouraging users to construct their own narratives and explore, and offering stories about the collection?

These questions will be pursued over the coming years. We welcome debate on our objectives and the means of achieving them.

7. ACKNOWLEDGEMENTS

This research builds on work by Florian Kräutli (EPSRC EP/J502169/1 and System Simulation Ltd) and Sam Cottrell (AHRC AH/M004775/1 and The National Archives). Olivia Vane’s research is supported by AHRC AH/L503782/1. We gratefully acknowledge the organisations that have given permission for images to be used in this paper.

8. REFERENCES


Priestley, J. (1764) A Description of a Chart of Biography. Warrington.
Using Data Visualisation to tell Stories about Collections
Stephen Boyd Davis, Olivia Vane & Florian Kräutli

228


9. VISUALISATIONS DISCUSSED IN THE TEXT