Author Accepted Manuscript version of chapter 35 in Lewi, H. (Ed.), Smith, W. (Ed.), vom Lehn, D. (Ed.), Cooke, S. (Ed.). (2019). *The Routledge International Handbook of New Digital Practices in Galleries, Libraries, Archives, Museums and Heritage Sites*. London: Routledge, https://doi.org/10.4324/9780429506765

Succession: A Generative Approach to Digital Collections

Mitchell Whitelaw

Digitisation is turning out corpora that rapidly exceed our human capacity for interpretation. In 2014, the Internet Archive published some 2.6 million machine-extracted images to the Flickr Commons (Miller, 2014); this collection has since grown to 5.2 million items. The National Library of Australia's Trove newspaper database currently records some 218 million automatically transcribed articles (National Library of Australia, 2018a). Europeana's aggregated digital collection numbers over 55 million items (Europeana, 2018). While these massive collections are significant and celebrated digital GLAM projects, they also pose significant challenges for both collection holders and collection users. How might a curator work with millions, or tens of millions, of digital objects? How might a visitor understand the contents and contexts of such collections? Computation has enabled the emergence of these collections; as I have argued elsewhere, computation can also support improved representation and access to these collections (Whitelaw, 2015). This case study shows that computational techniques can also support new forms of creative and cultural practice, new ways of using and reusing collections, and new modes of interpretation and meaning-making.

Succession is an experimental digital GLAM project developed in 2014, while I was a visiting researcher at Newcastle University, UK. It draws on a set of around two thousand image records related to the city of Newcastle-Upon-Tyne and its surrounds. These were harvested from the Flickr Commons, where they were contributed by institutions including the Tyne and Wear Archives and Museums, the British Library, the Internet Archive and others. The project arose out of my encounters with the city, and the thoughts and feelings that the place provoked. I was struck by its densely layered quality; Newcastle was a key Roman settlement, later one of the crucibles of the Industrial Revolution, and is now

finding its way in a 'de-industrial' Britain (Chakrabortty, 2011). As well as its intrinsic interest, this city seemed to speak to a range of very current concerns: energy, urbanisation, industrialisation and its legacies. My aim in developing the work was to find a way to use digital collections to respond to or articulate the complexity of this place, its layered histories and its wider resonances; but also, to use these traces of the past as materials in a generative process that could prompt reconsideration and create new meaning. The work's title draws a term from ecology, alluding to continuous and ongoing change and adaptation. *Succession* aims to mine the city's industrial past to fuel consideration of our possible futures.

Succession was also influenced by an emerging strand of digital GLAM practice that takes a poetic and playful approach to collections, offering serendipitous samples and chinks of algorithmic insight. Tim Sherratt's (2013) *Trove News Bot* tweets archival news articles based on daily headlines; the British Library's *Mechanical Curator* posts random images from the library's digitised books (Baker, 2013). Sherratt's *Eyes on the Past* (2014) harvests faces from digitised newspapers, and has their eyes peer out through the interface, inviting investigation. These approaches reflect an emerging interest in collections as active sites of meaning-making, and experimentation with how we might encounter such collections in an everyday digital environment.

In practice, *Succession* is a web application that draws on a corpus of some two thousand images and combines these elements into new visual composites (or 'fossils') (see Figures 1 and 2). Each fossil is composed of five randomly selected source images—arranged, composited and potentially repeated to make a new image. Each source image is cited in full, with a thumbnail, title and home collection. These citations link back to the item's Flickr Commons page, so that while these composites often radically obscure, transform or juxtapose their elements, the sources and their attendant contexts remain navigable and intact. Composites may be saved, acquiring a permanent URL to become a new citable online object. The generative process of composition is performed live, in the browser, so each viewer will encounter a series of unique composites. The system allows for around 2.5×10^{15} combinations of elements (ignoring spatial and blending variations). At a rate of one per second, it would take around eight million years to show all permutations.

Figure 1. *Succession* composite 1508308880101
http://mtchl.net/succession/#/saved/1508308880101

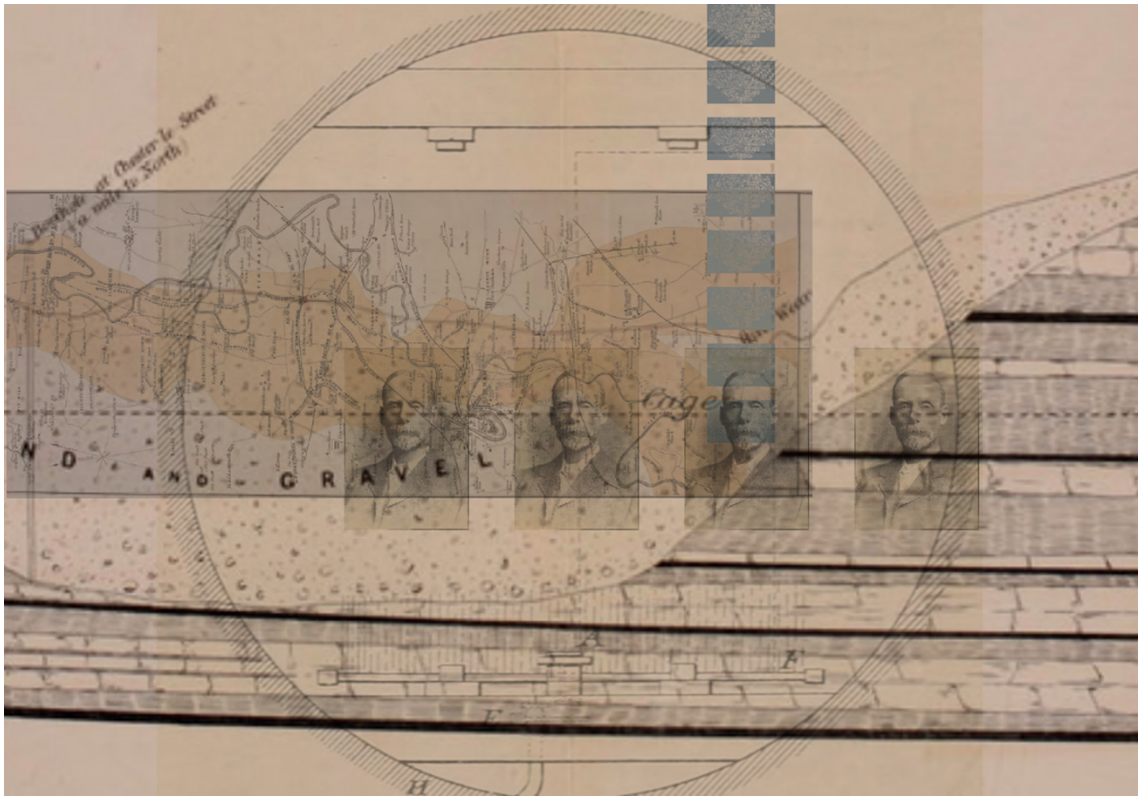

Figure 2. *Succession* composite 1478230078096
http://mtchl.net/succession/#/saved/1478230078096

Authored Spaces and Speculative Heritage

Succession's generative techniques are an instance of what Drucker and Nowviskie (2004) term speculative computing: 'speculative approaches to digital humanities engage subjective and intuitive tools' while enlisting computation for augmentation, rather than simply automation. A speculative approach pursues computation that is 'dynamic and constitutive', 'creating programs that have emergent properties'.

As Drucker and Nowviskie (2004) show, generative procedures have long been used in the arts to augment and extend thought, creating aesthetic and conceptual provocations. *Succession* uses combinatorics, the procedural combination of specific formal elements. Permutation gives combinatorics a form of quantitative leverage; small sets of elements proliferate into vast numbers of possible outcomes. When the permuted elements are symbolic or textual, combinatorics becomes a machine for knowledge or meaning. Ramon Lull's *Ars Magna* (1305) permuted divine principles into sets of theological assertions; these conceptual machines marked out domains of knowledge for investigation (see Gardner, 1958). In the 1960s, poets Brion Gysin and Raymond Queneau used combinatorics to pursue unforeseen or inconceivable meaning using fixed and finite textual means. Bill Seaman's 'recombinant poetics' seeks 'emergent meaning' within digital arrays of textual and audiovisual elements (2001). Ross Gibson and Kate Richards's (2006) *Life after Wartime* applies combinatorics to an archive of 1940s crime-scene photographs, showing how a generative 'story engine' can prompt speculative interpretations of digital heritage.

Bill Seaman (2001) describes his combinatoric generative system as an 'authored electronic space', emphasising that it is expansive, but not arbitrary (p. 426). Similarly, *Succession's* space is authored, in part through its corpus; sources were selected for content relevance and visual potential. The content-base grew and was pruned around the conceptual focus in a slow process of subjective evaluation, exploratory search and tangential investigation. In Seaman's (2002) words, this authorship seeks out a 'resonant unfixity'.

In *Succession*, the rules for compositing elements are also authored, tailored around the idiosyncrasies of the sources and the poetic aims of the work. Layered composition is a formal machine here—a way to combine and arrange visual sources into new artefacts. However, it is also a metaphor; to address a city built on coal, it seemed necessary to

combine and compress things, to obscure them while at the same time hopefully intensifying some of the energy latent in those sources, just as fossil fuels store concentrated solar energy gathered millions of years in the past. To evoke this feeling, the image blending modes are biased to overlay dark elements; this treatment works particularly well with the engraved illustrations in the Internet Archive and British Library collections. Thus, the generative algorithm is not an indifferent automated process; it has been tuned and shaped by the aesthetic potentials of the collection material, and by the (human) concepts and metaphors at play in the work.

Figure 3. *Succession* composite 1413513552860
http://mtchl.net/succession/#/saved/1413513552860

This saved composite in Figure 3 shows how the generated artefacts can also operate as what Drucker and Nowviskie (2004) call 'aesthetic provocations' while enlisting the contexts and referents of their source elements in speculative juxtapositions. This composite is dominated by a 1993 photograph of Wearmouth Colliery in its final week of

operation: a poignant image of the last days of Newcastle coal. But a spectral waterbird (from Gould's 1837 *Birds of Europe*) seems about to splash down in those desolate puddles. Faintly in the background is the bustling River Tyne itself, circa 1880; on the left of frame is a carriage destined for Newcastle's Metro system, under development in the mid-1970s. Almost imperceptible at bottom left is the *HMS Opal*, a torpedo destroyer, under construction at the Sunderland shipyards in 1915. Thus, this composite encompasses not only 150 years of urban history, but a latent portrait of 20th-century capital, the rise and (UK) fall of extractive industry, war, urbanism, and pre-Industrial naturalism and the non-human lives it records.

This example shows how formal and visual transformations inflect narrative or historical interpretations. Thus, these generative artefacts are not simply bundles of citations, but speculative visual propositions. Layering emphasises simultaneity and atemporal juxtaposition, rather than chronology: Gould's duck, about to dive into the colliery puddles; or perhaps swimming on the 1880s Tyne? Faded traces evoke the presence of the lost; visual collisions prompt an interpretive search for coherence, patterns of connection in the authored space of the system.

Thus, generative systems like *Succession* can be both prompts for interpretation and humanistic ways of speaking in themselves. In addressing complex topics (or 'wicked problems') such as extractive industry, capital and urban change, it seems necessary to respond in kind. In this project, generative techniques provide a way to speak complex multitudes, as well as an engine for unforeseeable combinations, using digital cultural collections as the seeds of something new. Seaman (2001) terms this 'unfixity'; Gibson (2006) calls it a 'restlessness' that prompts imagination through 'artful imbalances and implied possibilities for completion or patterning'. This active imagination is essential, Gibson argues, in understanding and potentially altering 'the continuous tendencies that are making us as they persist out of the past into the present'.

Succession and Digital GLAM Practice

As well as demonstrating some of the creative potential of generative techniques applied to digital heritage objects, *Succession* illuminates a number of current and emerging issues with wider relevance for the digital GLAM sector.

The promise and challenge of large-scale digital collections is a central theme here. *Succession's* approach exploits—and depends on—both the scale of these collections and their distinctively uncurated or computational quality. The Internet Archive's Book Images collection on Flickr is a key source. This collection is a fundamentally computational corpus, generated through an algorithm that identified and extracted images from the Archive's digitised books collection, appending existing book-level metadata and automatically extracted text (Miller, 2014). Automation maintains many of the valuable features of conventional collections, such as descriptive and structured metadata, while radically amplifying their scale; the collection currently numbers some 5.2 million images. By comparison, the Tyne and Wear Archives and Museums Flickr collection, another key source, has around 2,200 images. This large scale is both challenge and opportunity. It offers a wealth of potential resources, but it challenges traditional means of access and interpretation based on human curatorship. The Archive provides no curatorial narrative, no guiding themes or choice selections. In the duality that Lev Manovich identifies of narrative and database, this collection sides entirely with the database (Manovich, 2002, p. 225). It is readily accessible, indeed highly and usefully structured. Flickr's description and tags have been populated with consistent metadata, enabling browsing the collection by subject, publisher or other features. Notably, in the absence of human curatorial framing, all representations of the collection are themselves computational: search or browse listings generated by database queries. The emergence of large-scale digital collections will continue to challenge traditional approaches to managing, interpreting and using heritage objects.

The Internet Archive's broad scale, curator-free approach to generating a collection also shapes its content. Many of the images in the stream would never be deemed worthy of the human effort of digitisation and description. They are often oblique, obscure, fragmentary and generally marginal. Often, they must be read in the context of their

original (book) source to make any kind of sense. In the corpus harvested for *Succession*, a characteristic mixture of diagrams, illustrations and advertisements are drawn from publications such as the *Transactions of the Institution of Mining Engineers*—a primary source documenting the literal machinery of coal-powered industrialisation. These fragments are powerfully evocative but only available as resources thanks to the uncurated, algorithmic collection-making pursued by the Internet Archive. It is also notable that the Archive itself is not a traditional collecting institution; the Book Images collection is as much a product of an ambitious digital-first institutional approach, as it is of a specific algorithmic intervention.

Succession is also enabled by aggregation, another key emerging feature of digital collections practice. In this case, Flickr Commons aggregates content from a wide variety of sources and makes it readily accessible through a single point of entry. As well as the Internet Archive and Tyne and Wear Archives and Museums, *Succession* draws on collections including the British Library, the UK National Archives, the Imperial War Museum and as far afield as the Library of Congress and the Powerhouse Museum in Sydney. Aggregation enables a broad harvest that cuts across different institutional contexts and collecting practices, once again marginalising traditional curated interpretations, and enabling a complex new thematic aggregate to be formed through a single search query.

Flickr's application programming interface (API) works in tandem with aggregation. The API provides machine-readable collection data, meaning that users can write software to query and harvest that data at scale; *Succession* is completely dependent on the Flickr API to identify and gather its sources. Thus, it demonstrates the utility of GLAM APIs, but equally shows that as computational approaches to engaging with heritage develop, APIs and other forms of data sharing will condition and constrain it. APIs from large aggregators such as Trove and Europeana support significant re-use and experimentation (Europeana Pro, 2018; National Library of Australia, 2018b), while collections outside these aggregators, absent APIs or other forms of data sharing, will not benefit from this engagement.

Finally, *Succession* demonstrates some of the novel forms of creative re-use that are enabled by digital collections and their computational tractability. As well as combining sources to form new composites with emergent visual and semantic properties, *Succession* shows how re-use can draw attention to its sources. This work deliberately

includes thumbnails and descriptions of sources alongside each composite fossil. This cues the audience to the constituent parts of each composite, as well as inviting engagement. Each thumbnail links to the original source on Flickr, where it can be further investigated. Cryptic diagrams from the Internet Archive collection can be traced back to their source publications, which in turn provide rich new layers of information—on mine-shaft engineering or geological surveys. The looming hulls of ships at dock lead into the rich history of Newcastle's shipyards and their international trade. An array of speckled eggs will bring you to William Hewitson's 1833 volume *British Oology*, published, like many similar volumes, in Newcastle-Upon-Tyne. Each artefact that *Succession* generates is an opening into an expansive rabbit warren of provenance and context. Digital re-use has been a focus of engagement projects such as the DPLA's *GIF IT UP* (Digital Public Library of America, 2017); here too, sources are carefully attributed in the transformed content. A key difference is that *Succession* uses computational techniques to recombine multiple sources, multiplying the attributions in turn. Like projects such as Sherratt's *Eyes on the Past* (2014), *Succession* shows how computational techniques can blur the lines between creative re-use and collection interface.

The large-scale digitisation of cultural heritage is a transformative process. It unlocks a wide range of potential outcomes yet to be explored. *Succession* demonstrates a set of generative techniques that transform digital heritage artefacts into unforeseeable new composites. These offer prompts for speculative interpretation and imaginative insight, as well as sparks of aesthetic engagement. They also, importantly provide an interface for investigation of the source materials and their contexts. As well as demonstrating some of the generative potential of digital heritage, *Succession* illustrates emerging issues in the field as computational techniques transform the production and use of heritage materials. Large-scale collections, aggregation and machine-readable data are all key to supporting the generative approach that this work demonstrates.

References

Baker, J. (2013). *The Mechanical Curator*. Retrieved from http://britishlibrary.typepad.co.uk/digital-scholarship/2013/09/the-mechanical-curator.html

Berry, D. (2011, 14 January). Digital humanities: First, second and third wave [Blog post]. Retrieved from http://stunlaw.blogspot.com.au/2011/01/digital-humanities-first-second-and.html

Chakrabortty, A. (2011, 17 November). Why doesn't Britain make things any more? *The Guardian*. Retrieved from http://www.theguardian.com/business/2011/nov/16/why-britain-doesnt-make-things-manufacturing?CMP=twt_gu

Digital Public Library of America. (2017). *GIF IT UP.* Retrieved from https://pro.dp.la/events/gif-it-up

Drucker, J. & Nowviskie, B. (2004). Speculative computing: Aesthetic provocations in humanities computing. In S. Schreibman, R. Siemens & J. Unsworth (Eds.), *A companion to digital humanities* (pp. 431–447). Oxford, UK: Blackwell. Retrieved from http://digitalhumanities.org:3030/companion/view?docId=blackwell/9781405103213/9781405103213.xml&chunk.id=ss1-4-10&toc.depth=1&toc.id=ss1-4-10&brand=9781405103213_brand

Europeana. (2018). *Europeana collections*. Retrieved from https://www.europeana.eu/portal/en

Europeana Pro. (2018). *Apps*. Retrieved from https://pro.europeana.eu/resources/apps

Gardner, M. (1958). *Logic machines and diagrams*. New York, NY: McGraw-Hill. Retrieved from http://archive.org/details/logicmachinesdia227gard

Gibson, R. (2006). Places past disappearance. *Transformations 13*. Retrieved from http://www.transformationsjournal.org/journal/issue_13/article_01.shtml

Manovich, L. (2002). *The language of new media*, Boston, MA.: MIT Press.

Miller, R. (2014, 29 August). Millions of historic images posted to Flickr [Blog post]. Retrieved from http://blog.archive.org/2014/08/29/millions-of-historic-images-posted-to-flickr/

National Library of Australia. (2018a). Current work counts—Trove. Retrieved from https://trove.nla.gov.au/system/counts

National Library of Australia. (2018b). *Trove application gallery.* Retrieved from http://help.nla.gov.au/trove/building-with-trove/application-gallery

Richards, K. (2006). Life after wartime. *Canadian Journal of Communication 31*(2). Retrieved from http://cjc-online.ca/index.php/journal/article/view/1729

Seaman, B. (2001). Oulipo | vs | Recombinant Poetics. *Leonardo 34*(5), 423–430.

Seaman, B. (2002). Interview with Yvonne Spielmann. Retrieved from http://www.fondation-langlois.org/html/e/page.php?NumPage=386

Sherratt, T. (2013). *Conversations with collections*. Retrieved from http://discontents.com.au/conversations-with-collections/

Sherratt, T. (2014). *Eyes on the past*. Retrieved from http://eyespast.herokuapp.com/

Whitelaw, M. (2015). Generous interfaces for cultural collections. *Digital Humanities Quarterly 9*(1). Retrieved from http://www.digitalhumanities.org/dhq/vol/9/1/000205/000205.html

Biography

Mitchell Whitelaw is an academic, writer and practitioner with interests in digital art, design and culture, especially generative systems, data visualisation and digital cultural collections. His work has appeared in journals including *Leonardo*, *Digital Humanities Quarterly*, *Fibreculture* and *Senses and Society*. His current work spans materiality, data and culture, with a practical focus on creating 'generous interfaces' for digital and environmental heritage. He has worked with institutions including the State Library of NSW, the National Archives and the National Gallery of Australia, developing innovative interfaces to their digital collections. Mitchell is currently an Associate Professor in the School of Art and Design at the Australian National University.