

Cristina Linclau¹

¹ Solomon R. Guggenheim Museum, 1071 Fifth Avenue, New, York, NY, United States

clinclau@guggenheim.org

The evanescent flow of the World Wide Web has ushered in a greater tolerance for transience in cultural phenomena, evident in both contemporary artifacts and in the assembling and disassembling of dialogues between them. In such a distributed network that facilitates gleaning information from data, systems are increasingly designed to favor temporal, chance discovery across many sources over prolonged, explicit inquiry within one. This paper investigates serendipity as an emergent digital strategy – once a wonder, now a methodology – and situates the efforts of museums and other cultural institutes to optimally coordinate collections metadata with the saccadic movements of the web within a theory of simulated serendipity. Drawing a line from sociologist Robert K. Merton’s attempt to hammer out “The Serendipity Pattern” in empirical research in the 1940s to the use of the term within Google Arts & Culture own “Experiments,” serendipitous information behaviors can either open new pathways of intuition, or leave a world that was “meaning rich, data poor” in the hapless state of being “data rich, meaning poor” (Boorstein 2004). As unmistakable momentum builds behind Linked Open Data initiatives in museums, this paper speculates how to leverage the precedence of chance encounters to find and foster meaning in a data-driven Web.

Keywords: Museums, Serendipity, Linked Open Data

1. Introduction

For the past decade, cultural collections have at once expanded to include increasingly intangible and obsolescing elements¹ while relying more and more on the cacophony of the World Wide Web to retain relevance with their audiences. In so doing, they are reconciling their own historical practices of authority and permanence with decentralized models of network theory. At a glance, it would seem that transposing the museum experience to the Internet would be a natural evolution, since surfing the web is akin to the casual engagement that Didier Maleuvre attributed to walking museum galleries with “freedom to loiter and tarry, to indulge the long double-take, the retracing of steps, the dreamy pause, the regress and ingress of reverie, the winding progress that is engagement” (Genoways 2006). However, two distinct obstacles complicate the arrival of the much-discussed digital museum, with its endless virtual galleries displaying the entire history of human existence. The first obstacle is an audience that has embraced new media, been empowered by social platforms, and become accustomed to engagement at a lightening pace with endless scope. The second is somewhat less exhilarating: the variability of content across museum collections has engendered organizational systems that inherently lack the web-ready interoperability of more uniform description like that found in, say, library catalogs or jukeboxes. These two problems appear unrelated, but they inform one another in the realization of an online museum that supports collections of events and relationships, as well as objects.

Half a century before the Internet made digital networks a familiar substrate for human interaction, the librarian Suzanne Briet wrote that “what is important are the social networks and cultural forms that construct the meaning and value of documents” (Day 2006). Her statement makes the primacy of context within and among

¹ In 2007 the International Council of Museums (ICOM) updated its definition of “museum,” removing the criteria for them to act on “material evidence” and replacing that evidence with “the tangible and intangible heritage of humanity and its environment.”

repositories apparent and reveals that, as practices and technologies change, so, too, do the meanings made available. With an outlook specifically attuned to web culture, Andrew Hoskins echoes Briet in his 2009 essay, “Digital Network Memory:” “contemporary memory is principally constituted neither through retrieval nor through the representation of some content of the past in the present. Rather, it is embedded in and distributed through our sociotechnical practices” (Hoskins 2009). Today these practices emphasize interaction in escalating degrees, from hyperlinks to virtual reality. This preference for more interactive experiences has ushered in a higher threshold for impermanence in cultural exchange. Historian Walter Ong uses the word *evanescence* to describe the transience of speech in contrast to the permanence of inscription. In his 1982 book *Literacy and Orality*, evanescence is a primary characteristic of the maturation of the literate mind into a state of “secondary orality.” Unsurprisingly, evanescence also perfectly describes the temporality and responsiveness of the entire digital interface of the web. While built on a substrate of inscriptions – computer code – the constant dialogue of hyperlinks, user comments, pop-up ads tailored to your previous browsing history, and ubiquitous embedded media sweep away two fundamental precepts of the literate mentality: authority and permanence. These precepts have historically shaped museum practices and justified their actions. The tensions introduced by the World Wide Web and its secondary oral mode are having a profound effect on the role of museums in cultural production and social memory, demanding new methods that accommodate impermanence in everything from digital content to conservation to cataloging. As the Web ingests existing cultural collections through the surrogacy of their metadata, it is worthwhile to examine the sociotechnical practices that are activating that information for global audiences in order to set the course for institutional relevance and establish a sustainable networked museum environment.

2. Simulating Serendipity

Before joining various endeavors in augmented reality and artificial intelligence under the banner of “Experiments with Google,” the creative coding that was underway at Google Arts & Culture used tags to group its projects by theme (Google Cultural Institute 2017). These tags seemed as far from the lexicon of art history as could be imagined: “Machine Learning,” “Data Visualization,” “Real-Time 3D,” and “Serendipity.” This last tag seems out of place among its bleeding edge neighbors. This technical application of Serendipity is the latest transposition of a word that originated centuries ago as a loosely mystical place, Serendip, which was home to three princes that relayed startling skills of inference in regards to a blind camel. Without indulging in the details, it is enough to project this prototypical combination of cunning and receptivity – wherein a protagonist takes reasonable chances on plausible circumstances – forward through literary history. The plots of mystery and detective novels rely heavily, on the delight of unexpected outcomes and the observant minds that discern them. In 1747, when Horace Walpole coined the term, it was to describe the act of “making discoveries by accident and sagacity, of things which one is not on quest of” (Foster and Ellis 2014). Though it reads more as whimsy than tech jargon, the term migrated from its fairy tale origins to the domains of sociological study and scientific discovery, eventually making its way into the discourse of information behavior theorists as a way to describe “the discovery of information by chance or accident” (Agarwal 2015).

Among library and information science scholars, serendipity is a nested phenomenon within any number of non-linear models of information behavior that have gained traction in the past three decades, from Marcia Bates “berrypicking” (Bates 1989) to Sandra Erdelez’s “information encountering” (Erdelez 2005). Difficult to observe, due to its instantaneous and accidental nature, it is a continually emergent category of event, resting somewhere between active and passive modes of acquiring information. Its adaptation to data science is a commentary on the unknowns that arise from the well-knowns: improbable outcomes of predictive or algorithmic systems. Taken as such, it is a reverent testament to the impossibility of total control over any domain, especially as areas of knowledge become simultaneously more specialized and more interdisciplinary.

While capturing occurrences of serendipity in research and lab environments has proven as difficult as photographing Sasquatch in the wild, the phenomenon has found a domesticated habitat online, where organizational and algorithmic systems, backed by vast computer memory and processing power, quicken the pace

of encountering information. On the World Wide Web, serendipity comprises two states, both with consequences for collections metadata and interface design. First, it encompasses the personal experience of surprise discovery. Any web searches that begin with, say, a recipe for Beef Wellington and end with ABBA demonstrate the symptoms of such unexpected narratives. This “wending progress” is an important part of what museums have provided the public for decades; in this concern, serendipitous design is their natural talent. The second application is systematic, wherein uncharted geographies are created by the cartography itself. This is where museums fall short, and so an unprecedented interest in extensible data management practices has taken hold, encouraging adaptive descriptive metadata cataloging so that records can be more fluidly combined in broader, global repositories that dissolve the boundaries between disparate collections.

The list of over 1,200 cultural institutions that partner with Google Arts & Culture to “bring the world’s treasures online,” (Google Arts & Culture 2017) is as diverse as the mission is broad. It is hard to imagine that a single resource could bring together missions and collections so variant as the National Archives of the United Arab Emirates, the Georgian Theatre Royal (UK), Delhi Photo Festival (India), the China Intangible Heritage Industry Alliance and the US Navy Memorial. The site as a whole is a smorgasbord of cultural enrichment for a new generation of armchair historians and sofa bed aesthetes; it feels as though everything that one could ever want to know is just clicks away in any direction. Visitors are invited “to explore by time and color” (Google Arts & Culture 2018) choosing adventures ad infinitum, always in the driver’s seat. Additionally, Google harnesses contributions from their vast network of individual partners into curated “Stories,” “Topics,” and “Themes.” The result is a sprawling online exhibition space that comingles the divergent collections and perspectives of their Partners from all over the world in a single global village, a truly digital museum, albeit with some significant limitations.

While Google’s general imperative is to “organize the world’s information and make it universally accessible and useful,” a long history of variant practices and objectives in cultural metadata impedes universal access and use, leaving Google to rely on self-selected “clean” data samples from Partner collections that adhere to copyright permissions and Google’s submission guidelines. This distinguishes the work of Google Arts & Culture from other initiatives in which Google’s methods bump up against pre-digital practices. Efforts to guide the broad transition of libraries and museums to the web bring fundamental differences between these institutions into stark contrast while attempting to reconcile those differences by transposing all collections to data. Google Books, for example, is an ongoing massive ingest of the world’s published books into Optical Character Recognition (OCR) files made searchable by the same web crawlers that index the World Wide Web. Without a human-scaled index or hierarchical organizational structure, the textual content of these books becomes indistinguishable from their metadata. Google Books operates on the premise that, freed from the physical bounds of paper, monographic content can be hoarded, unsorted, until it is searched. This strategy moves information management away from complex descriptive organization of materials towards complex search methodologies that are only possible with the help of computers: unprecedented processing capabilities are the literal engine that makes Google Books “accessible and useful.”

Conversely, the substance of Google Arts & Culture relies entirely on volunteered museum contributions, so while their digital museum platform is more universally “useful” than any single Partner’s collection alone, it falls short of a claim to make Partner collections wholly “accessible.” The obvious reason for this is that museum objects themselves are less “accessible” than print editions of books. Another reason is that, from the vantage point of Google’s servers, networks and code, museum objects are also less “useful” as data. It is one thing to circumnavigate library metadata by turning endless volumes of text into an equally endless index, much like Jorge Luis Borges’ poetically impractical “point for point” cartography. However, building an online museum is impossible without museum metadata. The Google Books approach cannot succeed in museums because the content is more variable than library holdings and requires greater levels of mediation to establish representation online. Making museum collections “accessible and useful” cannot be accomplished by adding data processing power. Instead, Google Arts & Culture intervenes as an aggregating interpretive platform for the Partner organizations who themselves are providing only selective samples of their collections for inclusion. Instead of a global museum, the site demonstrates the complexity and limitations of representing and connecting cultural artifacts online.

Like the libraries that preceded the web, Google’s ingenuity is in manipulating, rather than creating, content. Beyond their “Themes” and “Stories,” “Experiments” are collaborations between programmers and artists that use the massive store of images and data provided by their Partners as the basis of digital tools for interacting with artworks online. These Experiments offer a glimpse into the “datafication” of cultural artifacts, where objects are succumbing to a machine-readable “common representational code” (Manovich 2002). In one, “X Degrees of Separation,” created by Mario Kilngemann and Simon Doury, image recognition is employed to suggest a chain of images that visually link any two artworks in 9 or 10 steps. This “network of connected artworks” (Google Arts & Culture 2018) behaves less like an online collection and more like an addictive game, a slot machine where the user seeds the algorithm with visual divergence and watches cultural history fall into an aesthetic continuum. This experiment, like so many “novel representational schemas and interactive modalities” that rest on the premise that mash-up is metaphor, relies on the emergence of “remix as a form of literacy for the 21st century” (West, Malina, et al. 2015). An element of surprise is anticipated; a new emphasis on discovery, rather than study, requires unpredictability and a level of volatility sufficient to propagate endless newness. With pre-Internet prescience, Walter Ong’s theory of secondary orality predicts this predilection:

where primary orality promotes spontaneity because the analytic reflectiveness implemented by writing is unavailable, secondary orality promotes spontaneity because through analytic reflection we have decided that spontaneity is a good thing. We plan our happenings carefully to be sure that they are thoroughly spontaneous (Hartley 2002).



Figure 1: Example of a result set in X Degrees of Separation, an experiment by the Google Cultural Institute, <https://artsexperiments.withgoogle.com/xdegrees/>

The hypnotic disruption of “X Degrees of Separation” is a symbiosis of the coded technology that eschews curatorial expertise and the audience that craves interaction. Not only are the result sets uncanny, presenting themselves like the frame-by-frame deconstruction of a magic trick, but each one is a contradiction to traditional curatorial precedent by enabling a quick comparison and contrast between objects that might have no other reason to find themselves adjacently displayed if not for the happenstance of the color and shape of like materials or the angle and lighting of their photography. The event of their mediation by the camera is as much a quality of the artwork’s representation as the artworks themselves. The result is a fun and fast-moving interface that feels like a roll of the dice but that rests atop layers of controlled systems, programs, practices and protocols. The spontaneity of “secondary orality,” and the serendipitous affect of the web, which moves with the pace and responsiveness of oral exchange, is “based permanently on the use of writing and print, which are essential for the manufacture and operation of the equipment and for its use” (Hartley 2002).

3. Total Recall and The Serendipity Pattern

Across disciplines, digital technologies excel at conjuring cohesion from the endless re-assemblage of data. While it is as impossible to reverse engineer the World Wide Web into its historical confines as it would be to repackage the big bang neatly into its own prehistory, reviewing the technical models of information organization that predate the web can help museums imbue meaningful narratives within their collections data throughout their transition to a

digital mode. The opposing forces of *precision* on the one hand and *recall* on the other have helped information professionals make meaning of search, browsing and retrieval behaviors among ever-expanding bodies of knowledge. As a cursory definition, precision and recall provide quantitative evaluation of information retrieval methods, balancing the number of retrieved items that are relevant (precision) with the number of relevant items retrieved (recall). These forces exist in tandem; “if you tune for precision, your recall will go down, and vice versa” (Feldman 2012). Consequently, information systems are designed to balance them like humors upon evaluative scales of mediation and memory.

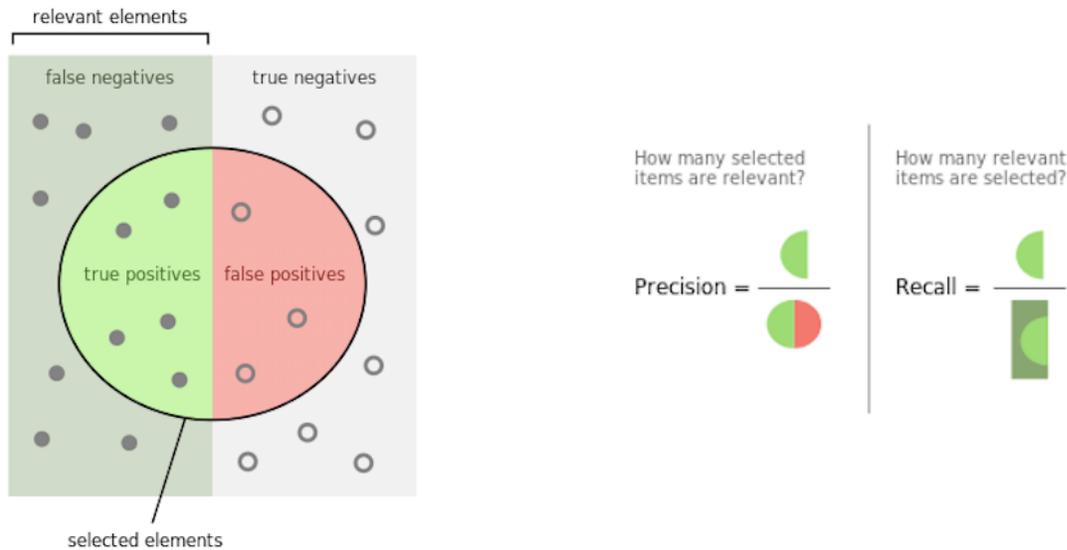


Figure 2: Precision and recall illustrated. By Walber - Own work, CC BY-SA 4.0, <https://commons.wikimedia.org/w/index.php?curid=369262>

Recall is, in part, a measure of the breadth of memory, be it actual human recollection or, on the web, the memory power of computer processing. In *Everything is Miscellaneous*, David Weinberg presents a strategy of “include and postpone” as a means of “going miscellaneous” (Weinberg 2007). This process, ingesting data wholesale and arranging it to order, “on-the-fly,” by algorithmic criteria aims “not to deliver the single right answer but to provide the maximum *potential* knowledge” (Weinberg 2007). It is exactly this strategy that Google Books and some of the Google Arts & Culture Experiments employ, retrieving information based on linguistic – or in the case of “X Degrees of Separation,” visual – patterns. Elevated recall makes it incumbent on users to contextualize search results, translating them into their own informational dialects. At it’s worst, an overabundance of recall reduces serendipity to mere miscellany.

Conversely, precision is arrived at through mediation, by culling and atomizing concepts within a larger body and classifying them in accordance with standardized semantic constructs. Precision is distinctly literate and authoritative, dependent entirely on the closure of the written word and its capacity for hierarchical control. It requires institutional bodies to create standards that dictate exclusions, subordinations, and partialities in order to function. In information science, the tools that increase precision are explicit about these tendencies, even in their names: authority lists, controlled vocabularies, stop words, and preferred terms. Rather than rely on the user’s perspective, the institution provides the context, necessarily concentrating meaning as an organizational tactic. While the benefits of precision are precise, unambiguous result sets, at worst, they disguise prejudice within hegemonic systems of knowledge organization that preclude serendipitous discoveries.

Considering whether information systems are coded for algorithmic recall, prescribed precision, or some combination of the two, can provide a lodestar for an emergent global museum network. When they are balanced, they promote serendipitous experiences, enriching the act of chance discovery with contextual relevance. While an

academic definition of serendipity is elusive and changes in varying contexts, a 1948 paper by sociologist Robert K. Merton identifies three qualities of “The Serendipity Pattern:” a serendipitous finding must be at once unanticipated; “anomalous...either because it seems inconsistent with prevailing theory or other established facts;” and strategic “i.e. that it must permit of implications which bear upon generalized theory” (Merton 1948). The Serendipity Pattern was used to determine the relationship between controlled, empirical research and the advancement of scientific theory, taking a long-overdue look at the complicated, and often suppressed, influence of unexpected events or accidents in the lab on scientific work. Of consequence to serendipitous design, Merton emphasizes that successful chance findings have far more to do with “what the observer brings to the datum than to the datum itself.”²

The last criteria of the Serendipity Pattern, outlining the strategic application of serendipity, yields a critical commentary on all chance encounters that happen by design, whether in “X Degrees of Separation,” or in your local library, where serendipity was first systematized. In libraries, subject indexing and cataloging systems are aimed at not only aiding the findability of specific resources, but on anticipating the usefulness of like materials that are then shelved in proximity to one another. While Merton’s theory observes happy accidents within, and often in spite of, controlled environments, information science investigates the necessity and efficacy of controlled environments to produce happy accidents. These controlled systems that have been honed for decades to aid in information search and retrieval support a strategic scaffold of western knowledge hegemony and historical narrative that “X Degrees of Separation” has no regard for or interest in. In the historical context, Google’s program has no use for Merton’s strategic “implications which bear upon generalized theory” whatsoever. The game is captivating regardless, or perhaps because, it defies context and as such, Google Art & Culture’s approach lacks the structure necessary to actualize the full potential for web-enabled cultural repositories. This is not a design oversight. These limitations arise because Google is working with the metadata museums have submitted to them – the majority of which was created by individuals for whom, and at a time when, “web optimization” was an alien concept. However, shifting strategies, Google is utilizing cultural metadata to strengthen its own influence on cultural engagement and information seeking by tipping the scales towards powerful recall. What it does fantastically well with Partner museums’ records is emphasize the power of the web to create what information behavior theorist Sandra Erdelz calls “super-encounterers” (Erdelz 2005), information seekers who consider chance indispensable to their information seeking behavior and who thrive on content that is at once “unbound, decontextualized and increase(s) their exposure to new knowledge” (Foster and Ellis 2014).

Google has spent nearly two decades making “super-encounterers” of us all. Trends in both scientific and humanities inquiry have embraced the confluence of empirical and theoretical methods provided by the processing power of computers. Historian Daniel Boorstin identifies these practices as an epistemological inversion, by which the analysis of big data upends the scientific method so that discovery supersedes hypothesis. Even before the World Wide Web came into popular use, he joked that the sciences were moving from a state of being “meaning rich, data poor” to being, “data rich, meaning poor” (Boorstin 1994). Similarly, practitioners in the humanities are adjusting their work to incorporate digital tools to an extent that is changing the nature of their professions. In a broad assessment of the state of, and potential for, the fast growing field of digital humanities, Johanna Drucker acknowledges that

humanities content met digital methods and created projects in which the terms of production were, necessarily, set by technological restraints. (The forms of print media and their rhetorics, by contrast, were established by humanist scholars for whom debate, commentary, and interpretive exposition were so essential they drove the development of the book format and the paratextual apparatus.) (Drucker 2012)

She gestures towards the same inversion that Boorstin identified, situating the ascension of discovery within humanities scholarship and, like Briet and Hoskins, acknowledges that meaning is derived from method. In as much

² It is worth noting here that Robert King Merton is actually a stage name carried over from the National Medal of Science winner’s early interest in magic, a field that is also rife with unanticipated, anomalous and strategic encounters.

as the humanities are concerned with making a laboratory of the written word, using text as a material for analysis, Drucker’s parenthetical comparison identifies a massive shift in methodological influence from the literate to digital modes. She isn’t arguing that “debate, commentary and interpretive exposition” are waning, but that the tools and platforms that facilitate these activities are less an extension of the literate mind and more the product of a networked, data-driven society.

While projects in the digital humanities vary, all put computational power to work as an extension of the human capacity for critical thought. As described in their first pamphlet, members of Stanford University’s Literary Lab utilized Docuscope and Most Frequent Words (MFW) software to enact “distant reading,” which contrasts the critical tradition of close reading by looking at vast swaths of textual input. The resulting text analyses identified patterns in the formal construction of writing samples – something like a combination of the Google Books approach with innovative algorithms for pattern recognition like those employed in “X Degrees of Separation” – to classify Shakespeare’s plays, and a sample of novels, by genre. The programs performed reasonably well, though the major conclusions were that they were in need of more human qualities: “...an analytical tool capable to quantify plot is still missing. And as long as that is the case, the generic distribution effected by Docuscope and MFW was too random to support a good literary taxonomy, let alone an exploration of the archive” (Moretti, et al. 2011). Interestingly, the patterns of articles and pronouns upon which the machine classifications were made were less easily apparent to scholars, writers, and humans generally (Moretti, et al. 2011).

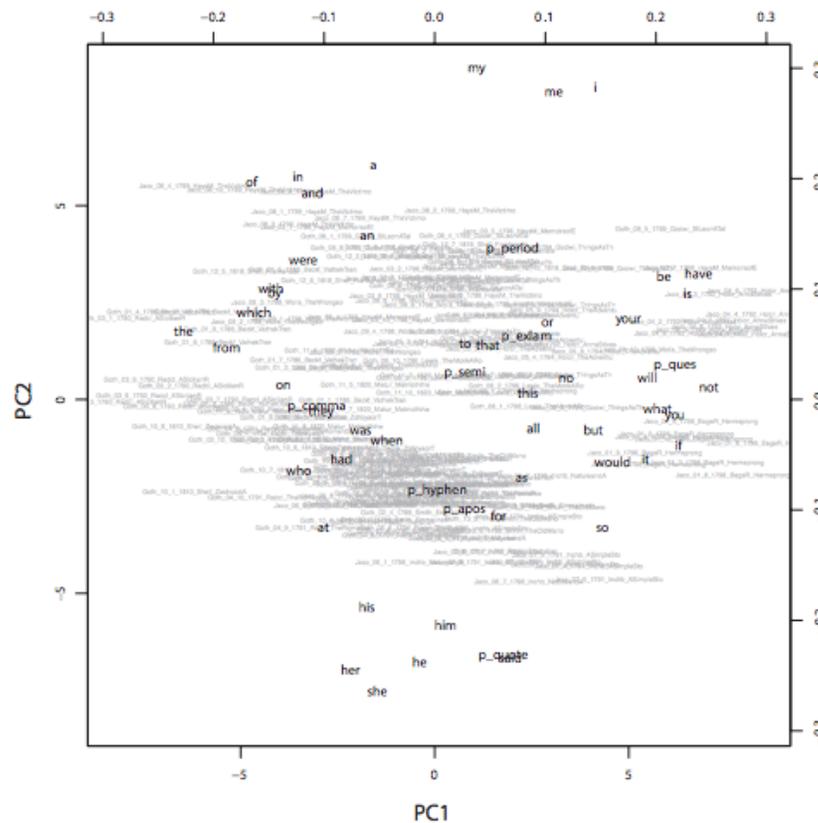


Figure 3: Most Frequent Words scatterplot with titles (light grey) and component loadings (black). Literary Lab Pamphlet 1, “Quantitative Formalism: An Experiment.” (Moretti, et al. 2011)

Docuscope and MFW’s ability to discern genres from textual input has the feel of yet another magic trick like Google’s “X Degrees of Separation” or another Google Experiment, “TAGS,” that uses an algorithmic program to apply over 4,000 tags to images of artworks without human intervention or accompanying metadata to see whether “Machine Learning could help people browse artworks similarly to how they search the web” (Google Arts

& Culture 2017). These simulations of serendipity teach us more about how machines think than how we do. Rapid retrieval and processing has the potential to expose the biases of human classification and make connections both larger and smaller than human scale, no matter if these new machine biases and connections are sometimes without meaning, like the frequency of pronouns as an indicator of Gothic literature (Moretti, et al. 2011) or the color blue as a link between a Givenchy hat and an ancient Egyptian tile (Google Arts & Culture 2018). While the scientific method, close reading, and traditional museum exhibition curation are distinct techniques that have defined their fields and the humanities as a whole, the new perspectives available through data mining prompt humanists and scientists alike to perform exploratory quantitative maneuvers that unite these diverse domains by denominating knowledge as data. By offering an unpredictable perspective, data processing of this sort is extremely useful for critical application within defined contexts, but it would be hard to argue that a Most Frequent Words scatterplot is a good introduction to the enjoyment of literature.

4. Prepared Minds

The first mention that Merton makes of serendipity appears in a footnote of an essay in a 1945 issue of the *American Journal of Sociology*. There, he defines the phenomenon as “the discovery, by chance or sagacity, of valid results which were not sought for” (Merton 1945). Merton’s inclusion of “sagacity” recalls the famous lines of Seneca, “Luck is what happens when preparation meets opportunity” and Louis Pasteur, “Dans les champs de l’observation le hazard ne favorise que les esprits préparés” (*In the fields of observation, chance favors only the prepared minds*) (Foster and Ellis 2014). Preparation, here, can have a number of meanings. In the sciences in which Pasteur was so accomplished, and which Merton studied, it would refer to empirical acumen. Preparation, in information systems, is the classification system itself, a referential sagacity. In both cases, there is a strong relationship between one’s preparation and the strategic factor of the Serendipity Pattern. The algorithmic structures that simulate serendipity in “X Degrees of Separation” accelerate logical processing of cultural metadata to expedite fleeting iterations and connections that are not in service to history, education or institutional agendas; their strategy is to validate digital technologies themselves. In as much as the “study of serendipity spans practitioners, computer systems, psychology, information science and philosophy” (Foster and Ellis 2014), the function that sagacity fulfills in serendipitous discovery is to provide a context or point of view. If Google’s Experiments lack sagacity and strategy from an art historical perspective, they are overrun with those qualities from a computational and data processing perspective. Programmers experimenting with Google datasets are building tools that derive from, and service, the strategies of their field just as much as the scholars in the Stanford Literary Lab are advancing their own agendas. Outlining these strategies reveals a gap between content and context in museum web identities. To assure that interactions with cultural collections online are designed for the advancement of art historical scholarship and for the enduring education of general audiences it is incumbent on cultural heritage professionals to prepare their own digital techniques, or else risk liquidating the sagacity of their domain and relevance of their collections.

The Stanford’s Literary Lab experiments are, intentionally or not, literal enactments of Lev Manovich’s dichotomy between database and narrative. In his 2002 book *The Language of New Media* he astutely observes that “a database can support narrative, but there is nothing in the logic of the medium which would foster its generation” (Manovich 2002). The web, too, enacts that tension between database and narrative in real time with every search, click, and scroll. Who tells the story, and how the databases and narratives find balance, is of utmost importance as the historical narrative becomes more fast-paced and fleeting. In the introduction to his book, *The Modern Invention of Information: Discourse, History and Power*, Ronald E. Day expresses concern that

analysis of information and society and culture have almost totally been given over to so-called information specialists and public policy planners, mainly from computer science, business and business schools, the government, and the quantitative social sciences...The overwhelming trend has been to place responsibility for the creation of an “information society” into ideologically conformist, “professional” hands, which inhibits truly critical analysis and discussions where the

fundamental premises and political stakes of information and communication might be shown and put into question (Day 2001).

Similarly, Alan Liu writes, in *The Laws of Cool: Knowledge Work and the Culture of Information*, “The corporate sector is where the new paradigm of IT-enabled knowledge *as* knowledge (and not just as the bottom-line mentality academics customarily dismiss) has been fashioned for wholesale application to other realms” (Liu 2004). Here, both Liu and Day examine the bureaucratic tendencies that manipulate techniques for criticism and free thought into ideological conformity. Their analyses share the perception that “information culture” is a construct that serves a purpose beyond the metanarratives of accessibility and usefulness. Information is instead becoming a technology in its own right, subsuming prior knowledge practices. Liu acknowledges the potential for interplay between modes:

There is no necessary reason why interfacing through rigorous practices of close “reading” cannot continue in a virtual reality space, though there is also no reason why old standards of rigor invested in particular knowledge interfaces cannot themselves be critiqued by new interfaces in which quick, hypertextual jumps are often more incisive than the most patient sequential or hierarchical probing (Liu 2004).

This cannot be accomplished, however, if the only arbiters for modeling the storage, presentation and access to knowledge are situated outside the broader historical context, conjuring a representative, contemporary canon from a database while they themselves are divorced and divested from the methods that preceded them.

In a 2016 essay, Elizabeth Kolbert, author of *The Sixth Extinction: An Unnatural History*, wrote about job automation and its imminent threat to economic livelihood. She describes this sociotechnical death march thusly, “Each new technology displaced a new brand of workers: first knitters, then farmers, then machinists. The world as we know it today is a product of these successive waves of displacement, and of the social and artistic movements they inspired: Romanticism, socialism, progressivism, Communism” (Kolbert 2016). It is compelling to think of cultural progress as an act of displacement. As our appetite for Big Data requires us to rely on computers to organize the world’s information, it is worthwhile to take a moment and observe the costs alongside the benefits of this convenience and inquire as to the counterintuitive loss of all this access. It is not that the cultural products coming from these network experiments are invalid; it is more that they are intentionally disembodied, tracing a widening orbit around the gravitational pull of the historical record. This is inherent in shifting a mindset from online “search” to a “discovery,” relying on algorithms not only to provide answers to existing questions, but also to present questions that were not previously apparent. Liu asks “how, in other words, is the progress of knowledge constituted from...internally rifted negotiations with historical knowledges, such that every...innovation creates, in its shadow, not just a dark hemisphere of obsolete peoples...consigned to the social margins, but also a repurposing and recirculation of the knowledges of the people of the margin...” (Liu 2004)? There is a term that describes the phenomenon by which data is lost in compression or migration from one format to another, “lossiness.” A certain amount of “lossiness” can be expected, and observed, when migrating cultural life to the web. These risks have been discussed on psychological and culture levels (Jenkins 2006; Turkle 2011; Heffernan 2016) and have a bearing on historical and institutional identities as well. By designing online engagement in cultural heritage for only the unanticipated and anomalous aspects of serendipity without strategy or cohesive context curatorial interpretation is lost and the trajectory of that knowledge is occluded in the margins.

5. Informed Instantiation

Google’s Experiments prove that computational recall can’t fabricate the connections that constitute art historical narratives alone. New techniques are developing to build transverse pathways for online visitors to experience cultural collections on the web in a global capacity, extrapolating relationships between objects by large-scale aggregation through innovative data modeling that elucidates the relationships between objects, events, people, and concepts that are absent from so many current examples. The most effective initiatives to unify, strengthen and deepen collections information for the web fall into the category of Linked Open Data (LOD), a term that

encapsulates a new sociotechnical paradigm, joining a set of technologies and policies that aim to remove barriers between databases for access to assets in a wider, more robust knowledge base than any one institution could hold and thus creating information networks that are hospitable to the autonomy of content. The representation of museum “content” – a sizeable portion of the world’s historical record – is of invaluable importance in determining the future of a World Wide Web built on conceptual relationships. To that end, museums have a unique and indispensable contribution to make in order to fulfill their commitment to the public trust within this most broad public forum. Furthermore, this network is itself an asset for better coping with the new demands of expanding cultural materials and modes.

In her 2015 book *Big Data, Little Data, No Data: Scholarship in the Networked World*, Christine L. Borgman dedicates a chapter to defining data, acknowledging that it

has yet to acquire a consensus definition. It is not a pure concept nor are data natural objects of their own. The most inclusive summary is to say that data are representations of observations, objects, or other entities used as evidence of phenomena for the purposes of research and scholarship. An entity is ‘something that has a real existence,’ per the Oxford English Dictionary, ‘as distinguished from a mere function, attribute, relation, etc.’ (Borgman 2015)

Linked Open Data amends this definition by including functions, attributes, and relations as entities in their own right. Through this elevation of relationships, machines can infer the contextual significance of cultural works, constituents and historical events to one another. This inference is the missing piece that keeps Google Experiments, and museum online collections experientially flat. The only take away from groupings of artifacts by likeness in color or by chronology is the empirical facts that those qualities are so, which is fine for school house recitation but falls far short of the capacities of the Web 2.0 arcade. LOD stands to deepen interoperability and standardization within and across collections. This, in turn, advances institutional mandates of cultural stewardship, interpretation, and engagement by enabling broader interplay between artifacts, widening search results and leading to greater, more meaningful, serendipitous discovery.

However, this is not quick, easy, nor inexpensive work and “cultural institutions have struggled with admitting that their resource-sharing mandates will necessitate use of these technically challenging measures” (Mayer 2015). For a sense of scope, a case study linking data between Oregon State University Library and the University of Oregon Libraries resulted in open access to 300,000 items across 31 digital collections and involved the regular contributions of 21 staff across two campuses (Simic and Seymore 2016). It took more than two years to complete. There are reasonable obstacles to getting institutional buy-in for dedicating resources to these initiatives, but the pay-offs are remarkable. As could be expected, Google is also applying its vast resources to Linked Open Data initiatives, defining a schema for scientific datasets. Osmo Suominen, who built the LOD publishing protocol for the National Library of Finland, says while the involvement of such big influencers on the web confirms the use case for LOD, “if the libraries themselves cannot agree on a model then this will leave the decisions out of their control” (Suominen 2017). This comment confirms that libraries, too, feel the infringement of excessive recall on their professional practices. Google Arts & Culture Projects fall short of what can be made possible in a fully realized LOD environment because they solicit sample sets of data from partnering institutions and centrally govern that data by moving it from the museum of origin and hosting it themselves. In contrast, Linked Open Data simultaneously enables and relies upon the free flow of information, promoting precision alongside recall. This challenges a habituated attitude that the information about and interpretation of collections within museums is somehow proprietary. If museums are to facilitate, rather than resist, their transformation from custodians of objects to repositories of knowledge, it is imperative to bring the history and influence of curatorial practice into the digital realm through the thoughtful presentation of relationships between works online. Precisely put, collection object exhibition histories are the raw data of that knowledge.

Exposing the history of museological activity by linking objects across institutions in an LOD exhibition history repository would enact, in digital space, what art historian and critic Claire Bishop refers to as “dialectical contemporaneity.” The phrase situates museums as the custodians not only of artifacts, but of their “recursive

temporality,” wherein exhibition and display “reboot the future through the unexpected appearance of a relevant past” (Bishop 2014). This concept is laid out in her book, *Radical Museology*, where she presents case studies of historically responsive exhibition practices in three European museums: the Van Abbemuseum in Eindhoven, the Museo Nacional de Reina Sofía in Madrid and the Museum of Contemporary Art Metelkova (MSUM) in Ljubljana. These examples invert the born-from-the-head authority of curation by displaying archival materials and entire past exhibits to expose the critical dominance of contemporaneity. For Bishop, the next step for museums is to be more transparent about their role in creating the present moment through the lens of the past, a lens that they themselves ground and polished. As such, “dialectical contemporaneity” recasts the relationships between viewers and objects, situating the museum in between as

an active, historical agent that speaks in the name not of national pride or hegemony but of creative questioning and dissent. It suggests a spectator no longer focused on the auratic contemplation of individual works, but one who is aware of being presented with arguments and positions to read or contest. Finally, it defetishizes objects by continually juxtaposing works of art with documentary materials, copies and reconstructions. The contemporary becomes less a question of periodization or discourse than a *method* or practice, potentially applicable to all historical periods. (Bishop 2014, emphasis in the original)

The potential for a linked meta-history of exhibitions capable of interoperating with collections of many origins and formats is not only a bounding step towards fashioning the web as a common archive of practices, ideas, arguments and positions, it is the premise for new discoveries within museum practices for a serendipitous approach to history itself. Bishop’s notion of contemporaneity as a method makes the role that museums have to play in this historical moment more urgent, not so that the curatorial viewpoints of history can be preserved as impenetrable, but so that they be used as yardsticks to measure the expanding and contracting temporal distances between cultural moments. Adopting a perspective of cultural reflexivity and evanescence is crucial for museums to retain relevance and marshal meaning into their web experiences. Implicitly, museums need to expand their mandate beyond preserving objects to include preserving their own histories and practices as a living archive for interpretation and research.

Revealing the history of interpretation through curatorial insights is a counterbalance to collective objectivity, by which an equalizing cacophony of voices blunts critical thought. Especially in a volatile historical climate, where consensus is hard to come by, it is important to declare the facts of history. If the interpretations of artifacts are not themselves always facts, their occurrences are. A shared initiative to catalog the careers of artifacts assembles the stories of how objects have moved through the world, congregating with one another in exhibition halls and gaining and ceding influence in the historical canon in order to construct the interminable present. Crucial to such a repository, Linked Open Data structurally supports a data network in which curatorial interpretation of artifacts and the artifacts themselves can occupy a continuum of events through a shared syntax. Without that, the museum mandate migrates to the web with a dilute focus on preserving objects solely in perpetuity, rather than with purpose. The intent is not to rally around exhibition histories as mechanisms by which to centralize interpretive control and resist the democratization of the historical record but to invite continued dialogue among a wider array of cultural stakeholders. Similar to Vivian van Saaze’s assessment of the cohesion of curation, conservation and exhibition as a collaborative act that she calls “doing artworks” (van Saaze 2013). Bishop’s “dialectical contemporaneity” is an invigorating reframing of museums as “doing history,” providing content with context and inference with intention. Viewed from this wide angle, the prospects of Linked Open Data serve an objective of *informed instantiation*. Since it is no longer possible to hold fast to fixed concepts on the World Wide Web, the way forward is to design for serendipitous interactions that directly leverage the strategy of museological narratives.

The web gathers and assembles its contents in a way that Andrew Hoskins argues is changing the very nature of memory, both personal and cultural. “Contemporary memory is thoroughly interpenetrated by a technological unconscious in that a ‘co-evolution’ of memory and technology in its current phase renders everyday memory as readily and dynamically configured ‘on-the-fly’” (Hoskins 2009). What Hoskins calls “everyday memory” is increasingly seen as an iterative process, reconstructed from related fragments under the sway of external influences. Moving away from models of retrieval and representation, memory and, more broadly, history

take on characteristics of the Internet, a thoroughly permutable environment where content and context are correlative. The present moment's sociotechnical growing pains – distrust of social and mainstream media and a broad reckoning with privacy concerns online – are shaped by the dizzying, disjuncting aftereffect of an Internet driven by the dogma that “content is king.” Content is only ever half of the story. Content is a Most Frequent Words scatterplot. The context of critical inference by the academics in the Stanford University Literary Lab mines that terrain for meaning. When considering the future of museums on the web, a disproportionate focus on first-level, “tombstone” object data excludes the curatorial “prepared minds” that provide historical context. The exclusion is two-fold: first it fails to represent historians in the digital museum construct and it further robs them of the potential for discoveries that can only be deduced from the powerful vantage of distant reading. In an “on-the-fly” digital network that runs on countless distributed systems, content alone is merely confounding, not serendipitous. Content alone drives Google's Experiments. The spectrum of discovery made available by a Linked Open Data network of exhibition histories satisfies the needs of rigorous scholarly inquiry while staying attentive to broad audiences that follow the lure of the periphery and have an appetite for the unexpected. The result is a true digital representation of museums. The particular challenges of collections management and audience engagement in the 21st century are a microcosm of the larger issues of control that Alexander R. Galloway identifies in his book *Protocol: How Control Exists After Decentralization*. In it, he identifies “a certain kind of rhizomatic protocol on the web” one that is “in fact both poles of this machinic movement, territorializing structure and anarchical distribution” (Galloway 2004). To best secure access to and preservation of cultural heritage, museum practices need to understand and occupy these poles – structure and distribution, precision and recall, authority and inclusion – and model themselves to accentuate the truly serendipitous places where those forces are in balance.

BIBLIOGRAPHY

- Agarwal, Naresh Kumar. 2015. “Towards a Definition of Serendipity in Information Behavior.” *Information Research* 20, no. 3. Accessed July 23, 2018. <http://www.informationr.net/ir/20-3/paper675.html#.WM7G2xIrLeQ>
- Bates, Marcia J. 1989. “The Design of Browsing and Berrypicking Techniques for the Online Search Interface.” *Online Information Review* 13, no. 5: 407–424.
- Bishop, Claire. 2014. *Radical Museology*. 2nd ed. London: Koenig Books.
- Boorstin, Daniel. 1994. “The Age of Negative Discovery.” In *Cleopatra's Nose: Essays on the Unexpected*, 3–17. New York: Random House.
- Borges, Jorge Luis. 1999. “On Exactitude in Science” in *Collected Fictions*. Translated by Andrew Hurley, 325. New York: Penguin Books.
- Borgman, Christine L. 2015. *Big Data, Little Data, No Data: Scholarship in the Networked World*. Cambridge, MA: MIT Press.
- Brawley-Barker, Tessa. 2016. “Integrating Library, Archives and Museum Collections in an Open Source Information Management System: A Case Study at Glenstone.” In *Art Documentation* 35, no. 1 (Spring): 86–113.
- Day, Ronald E., Laurent Martinet, and Hermina G. B. Anghelescu, eds. 2006. *What Is Documentation?: English Translation of the Classic French Text*. Lanham, MD: Scarecrow Press.
- International Council of Museums (ICOM). 2018. “ICOM Definition of a Museum.” Accessed July 21, 2018. <http://archives.icom.museum/definition.html>
- Day, Ronald E. 2001. *The Modern Invention of Information: Discourse, History and Power*. Carbondale, IL: Southern Illinois University Press.

- Drucker, Johanna. 2012 "Humanistic Theory and Digital Scholarship." In *Debates in the Digital Humanities*, edited by Matthew Gold. *Minneapolis*: University of Minnesota Press. Accessed July 23, 2018. <http://dhdebates.gc.cuny.edu/debates/text/34>.
- Erdelez, Sandra. 2005. "Information Encountering." In *Theories of Information Behavior*, edited by Karen E. Fisher, Sandra Erdelez and Lynne E. F. McKechnie, 179–184. Medford, NJ: Information Today, Inc.
- Feldman, Susan. 2012. *Synthesis Lectures on Information Concepts, Retrieval, and Services: The Answer Machine*. San Rafael, CA: Morgan & Claypool Publishers.
- Foster, Allen Edward, and David Ellis. 2014. "Serendipity and its Study." *Journal of Documentation* 70, 6: 1015–1038.
- Galloway, Alexander R. 2004. *Protocol: How Control Exists After Decentralization*. Cambridge, MA: MIT Press.
- Gilchrest, Alison. 2003. "Factors Affecting Controlled Vocabulary Usage in Art Museum Information Systems." In *Art Documentation* 22, no. 1 (Spring): 13–20.
- Gitelman, Lisa. 2006. *Always Already New: Media, History and the Data of Culture*. Cambridge, MA: MIT Press.
- Google Arts & Culture. 2017. "Google Arts & Culture." Accessed February 9, 2017. <https://www.google.com/culturalinstitute/beta/?hl=en>
- . 2017. "Experiments." Accessed April 1, 2017. <https://artsexperiments.withgoogle.com/>
- . 2018. "X Degrees of Separation." Accessed July 4, 2018. <https://artsexperiments.withgoogle.com/xdegrees/> and https://artsexperiments.withgoogle.com/xdegrees/ZQF_H4f0RL06EQ/BwGARp-j6XoXLw
- . 2018. "Explore." Accessed July 14, 2018. <https://artsandculture.google.com/explore>
- Heffernan, Virginia. 2016. *Magic and Loss*. New York, Simon & Schuster.
- Hoskins, Andrew. 2009. "The Digital Distribution of Memory." *CiteSeerX (Pennsylvania State University College of Information Sciences and Technologies)*. Accessed July 23, 2018. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.473.3811&rep=rep1&type=pdf>
- . 2009. "Digital Network Memory." In *Mediation, Remediation and Cultural Memory*, edited by Astrid Erll and Ann Rigley, 91-108. Berlin: Walter De Gruyter.
- Jenkins, Henry. 2006. *Convergence Culture: Where Old and New Media Collide*. New York: New York University Press.
- Kolbert, Elizabeth. 2016. "Rage Against the Machine." *New Yorker*, December 19 & 26.
- Liu, Alan. 2004. *The Laws of Cool: Knowledge Work and the Culture of Information*. Chicago: The University of Chicago Press.
- Maleuvre, Didier. 2006. "A Plea for Silence: Putting Art Back in to the Art Museum." In *Museum Philosophy for the 21st Century*, edited by Hugh H. Genoways, 165–176. Lanham, MD, AltaMira Press.
- Manovich, Lev. 2002. *The Language of New Media*. Cambridge, MA: MIT Press.
- Mayer, Allana. 2015. "Linked Open Data for Artistic and Cultural Resources." In *Art Documentation* 34, no. 1 (Spring): 2–14.

- Merton, Robert K. 1945. "Sociological Theory." *The American Journal of Sociology* 50, no. 6 (May): 462–473.
- . 1948. "The Bearing of Empirical Research Upon the Development of Sociological Theory." *American Sociological Review*, 13, no. 5 (October): 505-515.
- Moretti, Franco, Sarah Allison, Ryan Heuser, Matthew Jockers, and Michael Witmore. 2011. "Quantitative Formality: An Experiment." Stanford, CA: Stanford Literary Lab. Accessed February 12, 2017. <https://litlab.stanford.edu/LiteraryLabPamphlet1.pdf>
- Hartley, John, ed.. 2002. Ong, Walter J. *Orality and Literacy: The Technologizing of the Word (30th Anniversary Edition)*, edited by John Hartley. London: Routledge.
- Palmer, Carole. 1999. "Structures and Strategies of Interdisciplinary Science." *Journal of the American Society of Information Science*, 50, no. 3: 242-253.
- Pepi, Mike. 2014. "Is a Museum a Database?: Institutional Conditions in Net Utopia." *e-flux* 60 (December). Accessed December 11, 2016. <http://www.e-flux.com/journal/60/61026/is-a-museum-a-database-institutional-conditions-in-net-utopia/>
- Simic, Julia and Sarah Seymore. 2016. "From Silos to Opaquenamespace: Oregon Digital's Migration to Linked Open Data in Hydra." *Art Documentation* 35, no. 2 (Fall): 305–320.
- Souminen, Osmo. 2017. "From MARC Silos to Linked Data Silos?: Data Models for Bibliographic Linked Data." (paper presented online as a webinar hosted by the Association for Information Science and Technology and the Dublin Core Metadata Initiative (ASIS&T/DCMI), February 28).
- Svenonius, Elaine. 2000. *The Intellectual Foundation of Information Organization*. Cambridge, MA: MIT Press.
- Turkle, Sherry. 2011. *Together Alone*. Cambridge, MA: MIT Press.
- van Saaze, Vivian. 2013. *Installation Art and the Museum: Presentation and Conservation of Changing Artworks*. Amsterdam: Amsterdam University Press.
- West, Ruth, Roger Malina et al. 2015. "DataRemix: Ddesigning the Datamade." *Leonardo*. 48, no. 5: 466–467.
- Weinberg David. 2007. *Everything is Miscellaneous: The Power of the New Digital Disorder*. New York: Holt Paperbacks.
- Wood, David, Marsha Zaidman, and Luke Ruth. 2014. *Linked Data: Structured Data on the Web*. Shelter Island, NY: Manning Publications.