A Classification of Digital Emergence: A Critical Approach to the Production of Digital Objects in Special Collections

Robert D. Montoya
University of California, Los Angeles

ABSTRACT
This paper examines the infrastructure of digital libraries and teases out the subtle ways their formation and construction is a digital extension and representation of the social, political, and institutional circumstances by which they are created. Building off lessons learned from UCLA Library Special Collections as a case study site, this paper proposes a classification of digital emergence that provides more transparency about how digital surrogates come to exist in digital libraries and how we can use this information to better contextualize the importance of these surrogates within academic library services. The discussion then situates digital libraries as medial interfacing infrastructures that are fundamentally non-neutral social apparatuses that disappear in the course of daily use. Marcuse’s notion of technological rationality is incorporated to illustrate the extent to which technological infrastructures influence and reformulate the way we understand the research process using special collections and archives and how these infrastructures can function as a mechanism for information control. Finally, Bowker and Star’s text, Sorting Things Out: Classification and Its Consequences, is used to briefly illustrate how librarians can contextualize the emergence of digital objects and how this context, and the concomitant technological biases, can be methodologically brought to light using infrastructural inversion.

Keywords: archives · classification · critical theory · digital libraries · special collections

RÉSUMÉ
Cet article examine l’infrastructure des bibliothèques numériques et s’emploie à cerner en quoi leurs formation et construction constituent une extension numérique et une représentation du contexte social, politique, et institutionnel par lesquels ces bibliothèques sont créées. En utilisant les apprentissages acquis aux collections spéciales de la bibliothèque de UCLA, cet article propose une classification de l’émergence numérique, classification qui permet une plus grande transparence quant au contexte d’émergence de substituts numériques dans les bibliothèques numériques. Il démontre comment il est possible d’utiliser cette information afin de mieux contextualiser l’importance de ces substituts au sein des services des bibliothèques universitaires. La discussion fait ensuite état des bibliothèques numériques comme des infrastructures d’interface à la médiation.
(medial interfacing infrastructures), des appareils sociaux fondamentalement non neutres qui disparaissent dans le cadre de leur utilisation quotidienne. Ensuite est évoquée la notion de rationalité technologique telle qu’énoncée par Marcuse, afin d’illustrer la mesure dans laquelle les infrastructures technologiques influencent et redéfinissent la façon dont nous comprenons le processus de recherche en utilisant les collections spéciales et les archives. Cela permet également de démontrer comment ces infrastructures peuvent être utilisées comme un mécanisme de contrôle de l’information. Enfin, le texte de Bowker et Star, Sorting Things Out: Classification and Its Consequences illustre la façon dont les bibliothèques peuvent contextualiser l’émergence d’objets numériques et comment ce contexte, ainsi que les biais technologiques concomitants, peuvent être exposés de façon méthodologique en utilisant l’inversion infrastructurelle.

**Mots-clés :** archives · bibliothèques numérique · classification · collections spéciales · théorie critique

The presence and use of digital libraries in university settings is becoming more widespread as the availability of digital repositories grows in scope globally. Infrastructures such as the Digital Public Library of America (DPLA) and Canadia.org have become significant players in the dissemination of cultural collections, particularly in library instruction and information literacy sessions (DPLA, 2015; Canadia.org, 2015; Vlastnik, 2014). Repositories such as these have rightly been hailed as “potentially transformative,” providing vital platforms on which academic libraries and other cultural institutions around the world can publish their cultural heritage material in the public realm (Drucker, 2014). Educational institutions at all levels, ranging from K–12 schools to graduate-level courses within the university setting, increasingly rely on these digital resources to incorporate primary-source material into instruction (Gilliland-Swetland, 1998). Without such online access, these resources would be unavailable for classroom use.

Particularly significant efforts have been made to utilize digitized primary sources to enhance research skills, providing students with the literacies necessary to negotiate and interrogate these online repositories as well as the concomitant research skills to incorporate these resources into their scholarly work (Krause, 2010; Mitchell, Seiden, & Taraba, 2012; Bahde, Smedberg, & Taormina, 2014). The Ithaka S+R US Faculty Survey results from 2012 reported a marked increase in the use of primary sources in the classroom, particularly for the humanities and social sciences (Schonfeld, Wulfson, & Housewright, 2013, p. 19). The University of Pennsylvania Libraries initiative, Penn in Hand: Selected Manuscripts online digitization project, is one salient example of how easy it has become for professors to access and incorporate manuscripts dating back to the 1400s in the classroom (Penn Libraries, 2014). Additionally, Calisphere, a University of California–sponsored
program designed to integrate primary-source documents into “sets that support California Content Standard in History–Social Sciences, English–Language Arts, and Visual Arts for use in K–12 classrooms,” is a prime example of how digitized material can essentially transform the way educators contextually frame a student’s relationship to the material artifact of cultural production (University of California, 2014). Digitized content in Calisphere is curricular compatible and pre-organized into modular sets of material for student use. The convenience of digital images is that they can be remotely accessed, downloaded, and manipulated in any number of ways for scholarly use—and these are infinitely valuable qualities in many domains of inquiry. Facilitating this use, library institutions and federated search engines such as Google and Microsoft Academic Search have made the discovery of online sources easier with an array of browsing and searching tools (Schonfeld, Wulfson, & Housewright, 2013, pp. 19–20).

Yet, while a number of faculty do require that students visit the physical archives, as classes grow in size, digital quality and quantity improve, and the pressures upon special collections’ and archives’ capacities increase, digital repositories are quickly becoming the predominant mode of contact with special collections and primary-source material for many beginning researchers. The Ithaka S+R US Faculty Survey notes that the starting point for the research process is overwhelmingly occurring online—via general-purpose search engines, the online library catalogue, and electronic databases—versus at the physical library itself (Schonfeld, Wulfson, & Housewright, 2013, pp. 20–22). Amy Catalano’s review of the information-seeking literature examining the habits of graduate students identifies a similar reliance on electronic resources for research as they become more readily available (2013). Undergraduates, too, demand access to primary sources for required coursework, and such access may become the expected norm, “[satisfying] the vast majority of user needs” (Berger, Meltzer, & Jones, 2012; Conway, 2010, p. 74).

Given the widespread use of digital repositories in library teaching and learning, more work must examine digital library infrastructure within a critical framework, to contextualize the process by which digital objects are chosen for inclusion into digital libraries from the larger corpus of institutional collections. Library information literacy and pedagogical approaches have certainly expressed the importance of “critical thinking about information researching tools,” such as algorithmically retrieved results, but these examinations typically focus on results in federated search engines and how applicable they are for individuated purposes, rather than provide an avenue by which we can interrogate digital libraries and the available information sets from which those results are extracted (Grassian, 2009, pp. 88–91). Bishop, Van House, and Buttenfield’s text, Digital Library Use: Social
Practice in Design and Evaluation, is very helpful here, situating—in a multitude of essay contributions—digital libraries as sociotechnical and economic systems designed for specific types of community uses (2003). The recently published Association of College & Research Libraries’ Framework for Information Literacy for Higher Education indicates that “expert researchers [should] understand that finding relevant results is predicated on knowing where to search and understanding the basic constructs of the system being searched” (ACRL, 2014, p. 1). The situated construction of digital libraries needs to be more closely examined, and to do this one must understand not only the digital library as a stand-alone online unit, but also how it comes to materialize from a larger “information ecology” of digital production on a micro level within library environments (Bishop, Van House, & Buttenfield, 2003, p. 65).

This paper proposes a way to categorize the production of digital objects that arise from within a special collections setting and lays out some critical-theoretical foundations that can frame how the constructed nature of these mediating infrastructures affect our understanding of the information research landscape and the general availability of primary sources. This piece strives to answer two basic questions: (1) What processes within special collections and archives directly contribute to the addition of digital objects in a digital library? (2) How do these institutional digitization workflows help us understand the relationship between digital collections, the physical collections from which they emerge, and the social practices within the repositories in which they are made available? Investigating these questions will help us gain a better understanding of the institutional biases engrained in digital repositories as well as provide academic librarians with tools to expose these hidden layers of infrastructure as part of our research and information literacy services.

Situating Digitization Practices: UCLA Library Special Collections

Before proposing a classification system, it will be helpful to articulate how the work of producing digital images occurs in special collections environments and how internal staff structures support their creation with distinct workflows. To illustrate this, UCLA Library Special Collections (LSC) is briefly described, along with how each of its functional divisions supports digital image production. This description will contextualize the environment and practices in which the classification of digital objects proposed below is based. By virtue of their rare, fragile, and unique nature, and higher percentage of pre-copyright status material, many of the source documents for the digital surrogates that find their way into online repositories are held in university library special collections. Most special collections and archives are engaged in digitization in one form or another (Maron & Pickle, 2013, pp. 9–10). Given these facts, a special collections environment is an appropriate place to examine in
order to better understand how image surrogates find their way into these interfaces. Digital objects and libraries arise from within institutional social spaces, and as such, understanding these structured social mechanisms is essential if we are to fully and critically engage with the end-product infrastructure as a sphere of meaning in and of itself.

UCLA Library Special Collections is a library-wide enterprise, encompassing four major functional divisions: Administrative Services, Public Services, Collection Management, and a team of curatorial subject experts (UCLA Library Special Collections, 2015a). Each division in LSC, with the exception of Administrative Services, produces digital images under different pressures and contexts. Public Services serves as the primary mediator between the internal special collections environment and the exterior researching public (which, broadly speaking, includes staff, faculty, students, and non-campus parties), and as such, is at a crossroads for many forms of digitization. It produces digital images as part of three main operational areas, including instructional initiatives, duplication services, and communication activities such as social media production. Given Public Services’ consistent interaction with students and faculty, it is often the case that larger-scale digitization projects are introduced to LSC via this division. As projects grow in scope and require more coordinated efforts, other divisions participate in or take over digitization coordination and management.

Similar to Public Services, the Curatorial division prompts the digitization of material through its interaction with faculty and students as well as through agreements with potential and existing donors. Curators are charged with developing collections within their subject area, participating in instruction, and coordinating the overall collecting policy for LSC. Given curators’ subject specialty and deep knowledge of collections, they also play a significant role in recommending collections best suited for digitization.

The core digital initiatives for LSC arise from the Collection Management division, where the Center for Primary Research and Training has been an incubator for online interpretive platforms using digitized resources (UCLA Library Special Collections, 2015b). Collection Management also centralizes LSC’s digitization efforts through a committee charged with identifying and prioritizing potential digitization projects. The “Digital LSC” committee solicits project ideas from library staff who are aware of items or collections that need preservation digitization, are appropriate for online exhibit use, or may benefit from increased public accessibility. Large-scale digital projects involving staff and faculty, funded from both internal and external sources, are also coordinated within Collection Management’s robust workflows.
The production of digital objects takes a great deal of collaboration and coordination to implement, so LSC workflows within divisions are not mutually exclusive in practice. LSC also partners with a number of campus constituents to facilitate the creation of digital objects, including other library departments, the Center for Digital Humanities, Department of Information Studies, and Office of Instructional Development. The main online interface for LSC digital objects is the UCLA Digital Library, the central repository for all UCLA Library “digital objects, including text, images, audio, and video in their various digital manifestations and combinations” (UCLA Library, 2014). The UCLA Digital Library has been essential to the success of all digitization projects within LSC. The Digital Library sets long-term preservation standards for all digital assets and maintains the technical infrastructure that supports online access to collections. Given these overlapping domains of library and scholarly activity, who initiates digitization requests and how these projects are funded affect what gets digitized for public consumption. With a broad understanding of LSC department’s function as producers of digital objects, the next section proposes a way in which we can classify the production of digital images, and, in doing so, assess the relative merit of their inclusion into online infrastructure.

The Production of Digital Objects and a Classification of Digital Emergence

As hinted above, university special collections departments are complex bureaucratic organizations, and given this complexity, the reasons that holdings are digitized and deposited into an online repository are varied and far from self-evident in the infrastructure itself. Digitization involves not only the labour time of physically producing digital surrogates but also robust institutional workflows and infrastructural investment, including inter- and intra-departmental project management, quality control mechanisms, the creation and application of metadata, intellectual property assessment, long-term storage, and the maintenance of technical infrastructure to preserve these digital objects in perpetuity. Given these required resources, the management of such digital production processes is a costly endeavour. Somebody, somewhere within the chain of digital creation must absorb the costs of a digital object’s instantiation. In the Survey of Library & Museum Digitization Projects, 2011 edition, there is an extensive breakdown of the budget, staffing, collaboration, outsourcing, technical equipment, and software needed to successfully digitize on an institutional scale (Primary Research Group, 2010). Though this publication is not specific to special collections environments, the workflow considerations are generally the same, given the similar requirements for producing digital objects in any number of domains.

In an effort to render these often opaque institutional digitization processes more transparent, this section proposes a classification of digital emergence, the
goal of which is to expose the biases inherent in the production of these digital objects. Digital emergence is defined here as the institutional (political, economic, and administrative) circumstances that prompted an analog object to be digitized (or a born-digital object to be reproduced) and subsequently included in a publicly accessible digital repository. Fundamentally, the impetus and funding source for an object’s digitization is vital contextual information to make available if we are to engender a critical understanding of an object’s relative importance within an online repository, particularly in relation to the physical collections from which it emerged.

The six most-common activities that prompt the digitization of library material can be classified as researcher-generated requests, instruction and collaborative scholarship, donor driven, foundation supported, digitization for preservation, and collections identified as culturally or institutionally significant. These categories are certainly not exhaustive, nor are they mutually exclusive in practice, but they nonetheless outline core, general mechanisms by which images are systemically generated. Each of these categories represents a different domain and source of funding and administrative control, and thus each classification has vastly different motivations and selection procedures for digitization candidates. In this way, the digital library environment mirrors the institutional, bureaucratic, economic, and power dynamics already present in the physical library setting. These six categories can be roughly separated into three larger groups: on-demand methods, externally funded projects, and internally curated and funded initiatives (Figure 1). These categories are discussed in detail below and draw on examples from UCLA Library Special Collections.

<table>
<thead>
<tr>
<th>Types of Activities Prompting Digitization</th>
<th>Examples of Practices That Produce Surrogates</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. On-demand digitization</td>
<td>a) Research-generated: requests by on-site readers and off-campus researchers</td>
</tr>
<tr>
<td></td>
<td>b) Instruction and collaborative scholarship: digitization arising from instruction and pedagogical use</td>
</tr>
<tr>
<td>2. Externally funded projects</td>
<td>c) Donor-driven: digitization as part of acquisitions or by donor request after processing</td>
</tr>
<tr>
<td></td>
<td>d) Foundation-supported: digitization produced within articulated grant or foundation guidelines</td>
</tr>
<tr>
<td>3. Internally curated and funded initiatives</td>
<td>e) Digitization for preservation</td>
</tr>
<tr>
<td></td>
<td>f) Curator-selected: locally or culturally significant</td>
</tr>
</tbody>
</table>

**Figure 1**  Classifications for the Production of Digital Objects within Special Collections
On-Demand Digitization Methods

On-demand digitization consists of researcher-requested, instructionally prompted, and collaborative scholarship digitization. Digitized materials in this group typically arise with very focused, practical purposes, within intellectual circumstances that are specific to a controlled activity and context-specific pedagogical or research argument. Produced usually from within Public Services, digitization of this material is spurred by requests that arise from public use of the collection and from the normal course of instructional involvement with university classes. Researcher-requested digitization is often but not always smaller in scale and particular to the research project of the requesting entity. Scanning or photography in this case is often fragmentary—of particular pages within a book, or a selection of documents within archival collections—and thus is not typically representative of the larger corpus or volume from which they are selected. It is occasionally the case that institutions decide to digitize more than a patron requests, in order to make the collection of digitized items more useful for ingesting into a digital repository. The application of metadata for these images ranges from extensive to absent. Some institutions choose to apply full-scale metadata to the surrogates created in order to ingest them into their local digital library, while others deliver the image to the user and keep no copy of the file within their repository (Schaffner, Snyder, & Supple, 2011, p. 8). In LSC, fees are charged for mediated reproduction, and as such, researchers who have more funds at their disposal have the freedom to request that more material be digitized and thus potentially place more digital objects related to their particular projects into the digital library. Metadata associated with digitized images prompted by researcher requests should be indicated as such, particularly because digitization of these objects may be incomplete and their inclusion does not necessarily involve curated selection by internal library professionals.

Instructional digitization and collaborative projects are those that arise from engagement with classrooms as part of library teaching and learning initiatives, often from student requests for scanning as part of a given assignment, the creation of scans for classroom presentations by staff, or the digitization of items for exhibits generated as part of collaborative projects. The History of the Book and Literacy Technologies, and Seeing Sunset: Learning Los Angeles are two examples of robust, collaborative digital projects produced at UCLA through special collections partnerships (Drucker, 2015; Reiff, 2015). Within LSC, projects of this nature require a number of library constituents for adequate processing as they grow in scope, including Public Services, Collection Management, and the Digital Library. Digital objects created for this purpose are curated to cater to particular syllabi or student projects. Funding for the production of these images is usually absorbed by the library, since they are produced on an as-needed basis as part of service to the
university—though faculty occasionally have funds to subsidize the production of images for larger digital projects, or projects are foundation supported in some way to cover operating costs. Knowing that images come from instructional projects is significant, because they are produced in line with a university’s teaching mission and tend to be contextualized in specific engagements between faculty, library staff, and students. When possible, surrogates should uniformly point to the university’s published course descriptions, student/faculty project web addresses, or be accompanied by a statement of emergence describing the particular contextualized use of the objects.

**Externally Funded Digitization Projects**

Externally funded projects are supported by funds arising from an outside source and, as such, are generally biased in the direction of the funder’s particular need or area of interest. Donor-driven duplication often arises from requests by donors who require portions of their collection for personal use or who request that entire collections be digitized upon deposit into an institution at their expense. Some donors, though by no means all, have the financial means to request such (often) larger-scale digitization. Donors with more financial resources are better equipped to prompt institutions to scan more of their material. If digitized material finds its way into online repositories by this route, it is important for repository users to understand that these materials are a product of the process of collection development and acquisition but that the inclusion of images was entirely dependent on donor funding and thus not necessarily due to that collection’s relative importance to the collection as a whole.

Foundation-supported digitization might limit duplication of material by subject, format, or other criteria, but institutions in this case may have either applied for a grant acknowledging the importance of digitizing a particular portion of a collection or been given latitude in what collections they can digitize, within broad guidelines. In most instances, however, foundations have clearly articulated guidelines for the types of collections that fall within acceptable grant parameters. Instances can also arise in which a foundation’s interest is directed at very particular collections. The Los Angeles Aqueduct Digital Platform, coordinated by the Collection Management division in LSC, is a good example of a foundation-supported infrastructure designed around primary sources and scholarship related to aqueduct history (UCLA Library Special Collections, 2015d). These digitized resources should be understood within the context of these guidelines and selection procedures. To this end, grant guidelines, grant summaries, articulated selection procedures, or funding sources (if possible) should accompany all surrogates during the course of their life cycle—either described in metadata fields or by pointing to a stable project or foundation URL.
**Internally Curated and Funded**

Internally curated digitization projects are those that require more operational foresight and institutionally coordinated selection. They may be broadly defined as digitization for preservation and the duplication of collections deemed culturally significant. Many institutions, LSC included, have articulated workflows for proposing digital projects—a process that involves staff input and committee approval (UCLA Library Special Collections, 2015c). Curators also play a role in identifying collections that have particular cultural merit or research potential. Approval of projects is determined in light of funding, operational importance (future use in instruction, preservation, etc.), and level of project complexity, including the level of metadata needed, scanning resolution, volume of scans needed, and outsourcing. It should also be noted that duplication outsourced to vendors (for oversized items or large projects, for example) is often considerably more expensive. Metadata associated with internally curated digital objects should include a brief committee statement summarizing the selection criteria by which they were selected as well as any other pertinent information articulating their institutionally specific importance.

Digitization for preservation is performed on the most fragile items within a collection, not necessarily on the most “significant” portions of an institution’s holdings. Large institutions with voluminous holdings and limited financial resources, however, often have more fragile material than they can preserve and must choose the items that are most significant for teaching or research. In this case, the university can weigh research value against preservation digitizing at any given point, scanning only those items that meet particular internal requirements.

Finally, universities also digitize materials because they are deemed culturally significant, important at their local institution, or have high research value (regardless of condition). For example, institutions might digitize a collection based on articulated collecting priorities, high use, or rarity and usefulness to current scholarship. These highly curated modes of reproduction have much more institutional control and are more likely to fall in line with established collection development policies at the local site. As such, users of digital repositories may find it useful to know of the purposive selection in regard to these items. Selection summaries should be made available with all digital surrogates, pointing to or providing curatorial documentation for their selection.

Now that a general framework has been formed regarding how images are produced and can be classified in a special collections environment, the next section of this paper identifies why knowing this institutional context is critical if researchers are to identify and assess the biases these digital collections represent.
Additionally, it lays out a theoretical argument about how digital repositories as medial infrastructures disappear in the course of use, and how this affects the process of conducting research.

**Why Does It Matter How Images Emerge in Digital Libraries? In-Betweenness and Technological Rationality**

If the categories described above for the emergence of digital objects illustrate anything, it is that institutional processes directly correlate with the type of digital content chosen for inclusion into online infrastructures. Knowing why images make their way into digital libraries is essential, because digital libraries are quickly becoming the predominant source of primary research material for university classroom use and undergraduate researchers. As Paul Conway asserts,

> [it] is becoming increasingly clear that if information from analog sources is not readily available in digital form, it simply does not exist from the perspective of the vast majority of potential users.” (Conway, 2014, p. 52)

The digital library illustrates the principle of least research effort and an apparatus of convenience and efficiency, and while convenience is certainly a benefit, there is a danger when physical library spaces are circumvented on the assumption that digital resources are sufficient for a researcher’s particular purposes. Knowing why images are online can help informed users decide upon the best source for their argumentative purpose. But how is it that these infrastructures are taken at face value by students and researchers?

The medial quality of digital libraries, situated between physical archives and users, masks the reproduced quality of these digital documents. Philosopher and information scholar Luciano Floridi, in his book, *The Fourth Revolution: How the Infosphere Is Changing Reality*, speaks to this mediating “in-betweenness” of information communication technologies and how well-designed systems “should become functionally invisible to us” in the course of their use (Floridi, 2014, pp. 33, 38–39). It is this invisibility that we must counteract as librarians and scholars, for these systems—effective at distribution and delivery though they may be—detach the digital surrogate from its institutional and archival context (2014, p. 39). Interfaces interpret, despite how they posit neutrality through their invisibility. As Johanna Drucker states, “The [graphical user interface] is a mediating scrim, a boundary space in which we interact with abstraction of computation, not a window through which information passes like fast food at a drive-through” (2013, p. 216). Digital libraries present the simulacra of abundance and completeness while obfuscating the set of cultural documentation from which these objects were extracted. Floridi also demonstrates how technology is inherently political, either as a space of “detachment
and loss of pristine contact with the natural and authentic” or as an “empowering and enabling form of control” (Floridi, 2014, p. 39). In either case, the fabricated nature of an infrastructure’s in-betweenness is that it coordinates experiences, interprets, fashions visual arguments, and is populated by a coordinated selection of contents and arrangements. Digital libraries are spaces of articulated information control just like any other information system—for better or worse.

Ajit K. Pyati, in his essay, “Critical Theory and Information Studies: A Marcusean Infusion” (2006), posits that Herbert Marcuse’s theoretical notions of technology can be beneficial when interrogating the information structures that are ubiquitous in the field of information studies. “The types of theorizations and frameworks to guide radical democratic action are needed,” Pyati writes, “especially given the growing commercialization of the Internet, and apolitical, ‘neutral’ understandings of ‘information access’ and ‘information retrieval’” (2006, p. 88). Technologies mediate and filter, categorize, and fragment information on a user’s behalf ubiquitously and continuously in our daily lives. The result, as Pyati notes, is that “information, in its modern sense, became dissociated from affective, contextual, and cultural processes, thus making it easier to be commoditized, reified, and abstracted” (2006, p. 85).

In his 1941 essay, “Some Social Implications of Modern Technology,” Marcuse suggests that the rapid assimilation of technological processes into the fabric of social existence produces essential changes to the “rationality . . . and . . . standards of individuality” (1941, p. 139). The technological infrastructures we engage with daily have begun to fundamentally reformulate our interpersonal relationships, external behaviours, and interpretations of the material world. The “in-betweenness” of interfaces, and the increasing availability and use of digital surrogates, is changing the way we understand the availability of special collections and archival resources. We begin to filter our experience of the world through the valences of manufactured technologies and thus cannot disassociate our actions from this “apparatus” (1941, p. 141). Digital repositories become the start and end result of the research process, while our analog, physical (or born-digital) library collections remain outside the realm of easy availability and research possibility. Marcuse calls this general tendency a “technological rationality,” a way of reformulating the self-in-society due to one’s relation to the mediated technological structures we negotiate daily. As technologies integrate into our social experiences, they alter our expectations of how we negotiate our social world. Individuals begin to follow the “general patterns” of technology, further perpetuating notions of how one should act and experience the world (1970, p. 154). Further, given that technologies are inherently political and biased infrastructures, these liminal spaces become potential mechanisms for the exertion of intellectual and social control.
While the digital-library-as-infrastructure certainly seems a far cry from Marcuse’s social critique of the loss of individual autonomy and the “social dominations of technology over the individual,” it nonetheless is informative for helping us understand how these in-between spaces can change the patterns of research behaviour (“Herbert Marcuse,” 2008). The task of the critical user (and teacher) of digital library interfaces is to peel back these constructed layers, in order to understand the fundamentally produced and patterned construction of these mediated forms. This is especially true in a time when critical examination of digital library objects has become secondary to notions of quick availability and convenient access in instructional and research settings. Vicki L. O’Day and Bonnie A. Nardi advocate for an “ecological perspective” of digital libraries, where “systems of people, technologies, practices, and values” all intersect in a network of “interdependencies” that materialize certain infrastructure types (Bishop, Van House, & Buttenfield, 2003, chap. 4). By examining these interdependencies, we can better understand how digital repositories come to exist in the forms that they do, populated as they are.

This critique of mediating infrastructures should not negate the importance of digital libraries in and of themselves, for they are truly vital resources, broadening our access to collective cultural memory and documentation. After all, these online repositories are part of a long historical tradition of making information accessible to the masses, harking “back to the public library movements and the days of Andrew Carnegie and his admirable, successful, and transformative project” (Drucker, 2014). What these theoretical examinations should alert digital library users to, however, are the multifold layers of production that a seemingly neutral web interface embodies as a socially produced product, emergent from institutional selection policies and workflows. These digital libraries are made objects and thus should be understood critically as the subjective formation they are. Furthermore, we need to situate digital libraries as part of a larger institutional corpus of material, rather than as the predominant—and in some cases, only—material used in scholarly research.

**Exposing Institutional Selection Bias: Infrastructural Inversion and Describing Digital Emergence**

The question remains how one is to expose this classification of digital emergence, and, further, how the library field is to systematically execute this critical approach to digital infrastructures in our daily work. On the one hand, it is a simple matter of discourse, discussing these issues as part of our interactions with researchers and students and continually deconstructing these now-ubiquitous infrastructures. Making visible to users what has become invisible in practice has always been a core mission in library science and information studies. To this end, Geoffrey Bowker and Susan Leigh Star’s “infrastructural inversion” approach can be helpful in framing
how we might examine the biases of technology within a practical classroom environment (1999, p. 34). “Infrastructural inversion means recognizing the depths of interdependence of technical networks and standards, on the one hand, and the real work of politics and standards on the other” (1999, p. 34). Seeing technological infrastructure as a manufactured object with numerous inherited systems and decisions, Bowker and Star call for a deconstruction of the systemic layers of infrastructures (in this case, technological) in order to better understand their “moral order” and configurations (1999, pp. 33–34). They offer four methodological themes with which users can approach infrastructural inversion: the ubiquity of classification and standards, the materiality and texture of classification, the indeterminacy of the past and the iterative revision of systems, and practical politics (1999, pp. 37–46). For our purposes, we will discuss only one, practical politics, because the above-proposed classification of digital emergence in special collections is one example of such an interrogation.

Practical politics is an approach to infrastructure that assumes all constructions of technology are facilitated by a series of “institutional negotiations, organizational processes, and conflicts” (Bowker & Star, 1999, p. 44). By invoking infrastructural inversion, the goal is to carefully expose the minute decisions and circumstances that bring a particular technology into being—especially those aspects rendered invisible in a fully functioning, meticulously designed system. In the case of digital libraries, as a start, this process could involve an understanding of why and how digital objects find their way into online systems. The above-proposed classification of digital emergence can be a first step in delineating the institutional biases and economic circumstances that are represented in federated digital libraries. More site-specific examinations of the bureaucratic practices in special collections environments such as LSC can go a long way to contextualize digitization as part of a system of negotiations and operations. As library professionals teach students and scholars how to navigate online repositories in the course of research, they should simultaneously engage in discussions regarding how built systems embody moral principles and political dimensions. As Jonathan Furner states, “we need to appreciate that information science has an ethical dimension,” and to that end, we must ask how a digital library’s formation is or is not providing equitable, diverse, and representative information under current selection practices and workflows in special collections (2012, p. xx). And if not, how then do we change these practices to be more inclusive of divergent voices, communities, and constituents? It is incumbent upon library professionals to continually question our institutional practices with an aim toward equity and diversity.
On the other hand, from a practical point of view digital libraries do not offer uniformly adequate metadata to interrogate the conditions for digital emergence, leaving what Paul Conway calls the “secondary provenance” invisible within the technical system (2014, p. 7, citing Lori Podolsky Nordland). It is quite difficult to assess online repositories if we don’t have the metadata essential for critique. Without metadata attesting to why an object was digitized, a researcher has no way of assessing its relative importance to those entities beside it. We must give researchers the informational tools to interrogate digital infrastructures. This information can reside in multiple locations, including MARC records or finding aids, and as embedded data in the images/sound/video files themselves. At the local level, best practices can be adopted to categorically include information related to digital emergence with all generated objects. Curation development summaries, grant guidelines, and class and instructional descriptions are some statements that can accompany these items. Without such description, we have no basis to assess these infrastructures as library professionals, let alone expect users to assess them on their own terms. Further, by articulating these dimensions in metadata, and by focusing on them in our teaching, metadata itself becomes a more vital and engaged object of attention in the course of student research. As federated digital libraries such as the DPLA concatenate disparate repositories, it becomes all the more important to ensure that proper metadata follows these digital objects as they become associated with millions of other resources on the web from all over the world. The re-use of these surrogates may be out of our control, but we can do our best to describe their emergent qualities.

The classification of intellectual objects within information systems has long been understood as a fundamental exertion of power (Wilson, 1968). Bowker and Star, in their influential work, Sorting Things Out: Classification and Its Consequences, illustrate the extent to which “relatively invisible” classification technologies, “embedded in working infrastructure,” control, situate, and organize the world in which we live (1999, pp. 6, 319). Digital library infrastructure has such embedded practices invisible to the user. Classifying the emergence of digital objects is essential for exposing these biases. At present, institutions are not uniformly providing all of the necessary tools, metadata, and description by which we can deconstruct these technical repositories that have become second nature—the default, and sometimes the only, step in the research process. Though there may be costs involved with such added description, it is vitally important context, as our resources—usefully and productively—fragment into multiple locations and platforms.
Conclusion

In response to the growing use of digital libraries as the predominant source for student research and instruction, this paper has proposed a classification of digital emergence, based on the operations at UCLA Library Special Collections. This narrative has also provided a theoretical foundation upon which we can begin interrogating digital library infrastructures as constructed, political spaces. As digital collections gain in prominence, and as usage of these resources increases, especially in teaching and research environments, it is essential that we understand how these online collections relate to the social practices of selection within the special collections setting. This classification system can provide a starting point for users, giving them the tools to better understand why digital objects are created for public consumption and to gain a sense of their value in relation to their host institution’s physical collection. Finally, examining digital libraries as medial infrastructures, this paper has posited theoretical foundations from which we can understand the constructed and affective nature of digital library systems, including the invocation of Bowker and Star’s method of infrastructural inversion as a starting point to facilitate such examinations.

ABOUT THE AUTHOR

Robert D. Montoya is a doctoral candidate in information studies at UCLA. His dissertation, “Articulating Composite Taxonomies: Epistemology and the Global Unification of Biodiversity Databases,” takes a critical information studies (I/S) approach to examine what I/S can learn from the construction and representation of biodiversity taxonomic databases. His broad interests include classification theory, philosophy of information, documentation studies, print culture, and infrastructure studies. Robert has extensive experience working in special collections, archives, and academic libraries.

NOTES

1. The author would like to thank Michelle Caswell, Mario H. Ramirez, the journal editors, and two anonymous reviewers for their feedback and assistance with this piece.

2. The author acknowledges the distinction between special collections and archives in most academic library environments and the shifting definitions of these two entities. UCLA Library Special Collections is unusual in that both archives and special collections are intermingled in one unit. Therefore, in this paper the term special collections includes archival material, even if it is not explicitly stated.

3. Until September 2015, the author was the Head of Public Services for Library Special Collections and was heavily involved with managing and participating in digitization and digital initiatives.

4. The production of images for researcher use has diminished since the widely adopted practice of allowing personal cameras and smart phones in special collections reading rooms (Miller & Galbraith, 2010). Images produced for researchers are typically high-quality images intended for specific uses in publications.

REFERENCES


