

*Title of the Satellite Meeting: the IFLA Serials and Other Continuing Resources (SOCRS) and National Library Section*

*Date: August 23, 2019*

*Location: Athens, Greece*

## **Collection Strategies for Grey Literature in the Sciences: Implications for Research Libraries in the Digital Era**

**Julia Gelfand**

Collection Strategies Department  
University of California, Irvine Libraries  
Irvine, California, USA  
[jgelfand@uci.edu](mailto:jgelfand@uci.edu)

**Anthony (Tony) Lin**

Reference & Public Services Division  
Irvine Valley College Libraries  
Irvine, California, USA  
[alin@ivc.edu](mailto:alin@ivc.edu)



Copyright © 2019 by Julia Gelfand and Anthony Lin. This work is made available under the terms of the Creative Commons Attribution 4.0 International License:

<http://creativecommons.org/licenses/by/4.0>

---

### **Abstract**

*Grey literature started with content that was difficult to find, acquire, describe, challenging to process, cite and problematic for libraries to manage. Grey literature was considered a low priority commodity that was labor intensive serving a specialized, but limited group of researchers. Today, grey literature has shifted from the shadows into the mainstream. The proliferation of electronic resources and digital services in recent years demonstrate the increasing importance of e-formats in the marketplace transforming how central and ubiquitous they are to libraries. As a result, the value of grey literature skyrocketed to unprecedented heights.*

*Libraries have not only learned to process grey literature more routinely, but also initiated collecting in new formats and scope. These often include data, multimedia, audiovisual, imagery, graphics and a wide range of physical and digital components. Unfortunately, library collection policies have fallen behind the trend especially in consideration of the ephemeral nature of grey literature. Newspapers, maps, games, technical reports, theses and dissertations, data sets, government information are all included as a part of the current research landscape. In time, improved metadata facilitates accurate description and*

*the catalog record enhances discovery. In addition, searching, usability, citations, and references specific to this medium have experienced an evolutionary leap.*

*The value ascribed to grey literature is often determined by its research impact. As new forms of scholarship emerge across disciplines, greyness shifts from dark hues lighter introducing new content and outputs. The research lifecycle benefits from refinements including information citation and repurposing. The investment libraries make in developing collections to respond to user demands for new outputs with social capital contributing to research fulfillment increases. Research using grey literature has transformed from a narrow group to multi-disciplinary environments. Digital scholarship redefined how research in the humanities and social sciences is conducted. Previously traditional published works influenced research outputs. Now a reverse trend with research outputs influencing traditional publishing emerges. Publications that are part of series are no longer an obstacle to handle. Scientific grey literature today is more complex and shares frontiers with museums and other special collections that highlight content using appropriate technologies to exhibit, capture and process it. Physical objects appear more frequently in libraries than before and can be photographed at high resolution, digitized, scanned and described so users can discover and examine such collections with reduced barriers and increased accessibility.*

*Scientific grey literature includes biological pathways, lab notebooks, posters, and videos of scientific procedures, large data collected from observatories, animal testing, healthcare records, models, patents, lab specimens, electronic components and many other sources. The advent of 3D printing, fabrication of new materials and ethnographic processes can tell important stories and contribute to new science in novel ways. This form of innovation and scholarly communication takes into account technology transfer, intellectual property and best practices in sharing resources that make grey literature less grey.*

*This paper addresses how libraries have adopted a more centrist approach to handling scientific grey literature and the successful strategies research libraries practice to acquire grey literature in perpetuity. These new practices highlight how grey literature contributes to cohesive, balanced collections and open science.*

### **Keywords**

*Grey Literature; Collection Development; Collection Management; Science Collections; Publishing Trends; Scientific Information*

---

## **Background**

The definition of Grey Literature has been revised in recent years due to formats and content delivery options, primarily those born digital. Historically, grey literature was characterized as content that was difficult for library selectors to identify, acquire and process because it lacked provenance or series attribution. It often meant material that was referred to as government issued reports, or what we call “Technical Report Literature” today that became known as “intelligence” and were originally related to war generated reports, some requiring security clearances and others that were open. Charles Auger revised and expanded that concept and suggested that one of its advantages was that it had a “growing impact on scientific research.” (Auger, 1989) This alluded to a relationship in scientific literature that was perceived from Auger’s earliest work in the mid-1970s although it was not articulated exactly as that because grey literature was not coined until the second edition of *Information Sources in Grey Literature* was issued in 1989.

At the Third International Conference on Grey Literature in 1997, attendees assembled and approved yet more refinements to include, “that which is produced on all levels of government, academics, business and industry in print and electronic formats, but which is not controlled by commercial publishers.” (Luxembourg Definition, 1997) Nearly a decade later, a clarification was added to include, “. . .not controlled by commercial publishers, where publishing is not the primary activity of the producing body,” (Definition, 2004) and reinforced the economic aspect. “With the changing research environment and new channels of scientific communication, it becomes clear that grey literature needs a new conceptual framework.” (Schopfel, 2011) Schopfel and others began to apply the components of production, dissemination, topology, quality, acquisition and other non-commercial channels that suggested how it could economically be vital when the supply chain was crooked. This meant there had to be circuitous methods to obtain these materials, much less cite them so that they could be found by subsequent readers and rightfully attributed. Metadata as we think of it today, did not have sufficient power to index the content as a particular genre thus the typological approach was inoperative returning to Schopfel’s hypothesis.

These definitions have been challenged by pedagogy and research rigor as it was doubted whether the new forms of grey literature framed in the sciences and described by these definitions, such as models, maps, charts, software, datasets, eMail messages, websites, wikis, blogs, Twitter feeds, social media transcripts are even forms of scholarship or sources of literature. It was clear that these were newer information products and often cited as stand-alone “information” but may have lacked the traditional rigor associated with text, editing and certainly peer review. There is already extensive research among grey literature scholars that explores “new shades of grey.” (Gelfand, 2009) However, will the value or impact of these tools and resources stand the test of time? The British Library and the Library of Congress both point to the emerging value of “new forms of grey literature” and collection development efforts are already underway, leading to new research pathways in social dynamics, informatics and media studies.

Libraries and researchers argued and quibbled over the differences of release by non-commercial publishers as the rise in non-profit organizations was growing. Also, the distinction between peer reviewed and non-peer reviewed caused a major rift as whitepapers and other documents may have been invited, solicited or assigned, but not submitted to blinded peer review. (Gelfand and Lin, Rizal Library Conference, 2012).

### **Contrast to Library Collections**

Today institutional and thematic repositories often considered extensions of libraries, share some of these same quandaries. Should they treat all shades of grey the same or should they blend the hues to be more inclusive? (Gelfand, 2009). Currently, diversity in collections is a noble goal of numerous institutions and is an important consideration when crafting policies. However, diversity also adds a layer of complexity to an already complex situation. For example, should only doctoral dissertations be accepted or should master’s theses, undergraduate papers, and posters that are non-peer reviewed and non-juried also be included? As the ETD evolved this had great implications for format inclusion and other considerations. (Gelfand, 2001) What about preprints or non-subscription viewable versions of publications that force authors to retain their rights in order to enable public browsing of their research? Authors’ rights are critical factors but

should not predicate options for reuse and posting. As we move forward, intellectual property increasingly dictates and reinforces new rules about authentication and access.

Schopfel argues that if grey literature is not the keystone of a library collection or part of the acquisition process, a variation on the standard definition should account for the differences. He advocates for wordsmithing the New York definition to accommodate several important features. This includes distinguishing between archiving, preserving, and a deeper description that allows users to locate the grey literature they seek. Thus, library collections and procedures indicate that value is added through the selection, acquisition and cataloging or processing functions performed by library staff.

In today's more "open" environment, we have mandates for open access, open source publishing, open peer-review processes, open data to share and reuse. We place the onus on the reader or user to determine the quality, value, utility of the information contained therein. Libraries increase their value when their collections contain more unique content. Most library users prefer having their information needs met at the time of need and do not share the academic preferences for uniqueness or specialization or deep collections measured by academic competitiveness and elitism. Our definition diverges from Schopfel as we seek a more democratic stance on these delineations. The shift towards discovery has replaced cataloging and elevated metadata to a higher level in the current research library hierarchy.

### **Library collections work strategies**

The nature of collections work and management has changed as libraries and their collections have reworked their focus and processes. They may still serve their constituencies well, but they face new challenges in this digital age. For the purposes of clarity, we adopt the definition of library collections work as "the operations of managing a library collection, including collection development, acquisition, cataloging, data analysis, providing access and ensuring discovery." (Davis-Kahl, 2019) Library collections are "the total accumulation of books and other materials owned or leased by a library, cataloged and arranged for ease of access, in both print and electronic form. The term may also apply to items we neither own nor lease, but to which we provide ease of access [via OA]." (Davis-Kahl, 2019b)

An abundance of content has driven libraries to expand into a curation and preservation role. This constant state of information overload reflects the new challenges libraries face to select, vet, organize and provide long-term retention and preservation in countless formats. Academic libraries are starting to question the merits of maintaining collection policies to accurately document collections as the newest forms of scholarship become increasingly labor intensive and frequently change and emerging trends sidetrack management efforts.

No academic library today claims to collect exhaustively in any subject area, regardless of the size of the budget for new acquisitions. The emerging relationship between libraries and information providers creates new workflows between the two entities. The library as publisher is an evolving service with university presses and sometimes via institutional repositories serving as frontends as they host journals and scholarly outputs. Thirty percent of US University Presses noted in 2016 that they reported through the institution's library. (Watkinson, 2016) This is a robust administrative function of many libraries worldwide and there are many examples of different models of cooperation between libraries and publishers in achieving their goals of knowledge creation.

## **Changing landscape of library collections**

The Open Access movement functions differently across many spheres. E-prints take on different roles in terms of collections with the focus on distributed institutional collections made available through aggregation. They fit under a larger umbrella of documents including journal articles, conference papers, theses, and form multidisciplinary ePrint archives that are institutionally-based rather than on one focused subject area. Authors and librarians know that complexity in OA is due to the many different flavors that define it – gold to green with different hues emphasizing unique products and accessibility with multiple options for authors. This new normal has fundamentally changed the nature and composition of a library collection.

The shift from acquisition or ownership to licensing or leasing has not only changed what libraries do but also the legal and social implications of licensing including time commitments, consistency, and the number of players or staff involved in activating the content. The range of different models about how content is provided appears to multiply annually. Additional opportunities in library acquisitions and licensing include factoring in practices that allow for data driven and evidence-based examples to dictate what permanent holdings of the collection are and allow libraries to respond to their users at time of need through more services than content provision.

Cooperative collection development opens many new doors and practices. Many of these include institutions sharing specially trained staff, collecting for multiple libraries, coordination and options for faster resource sharing, storage of different formats, or distributing processing responsibilities by language, format, expertise. Libraries join multiple consortia to maximize those collaborations.

## **New Challenges in Changing Landscape**

Today, the adoption of nearly 5,000 operational institutional repositories in the Registry of Open Access (ROAR) suggests the importance of the Open Access movement to the library community. The requirements of complying with the versioning and other publisher requirements conflicts with the ease and convenience of submitting one's work. The coverage of peer-reviewed content remains a challenge and is a huge current dynamic in this area. Grey literature contains mixed peer and non-peer reviewed content, and has found hospitality among repositories for theses and dissertations, data sets, photograph collections, and scores of other products.

Libraries and users can view collection holdings from around the world with services such as OCLC that facilitates the ability to share information among its members. However, digital holdings remain a difficult area. Licensing terms vary widely between publishers. The ability to determine shareable and non-sharable content is not always clear making interlibrary loan/resource sharing even more challenging. Institutional repositories have stepped in to fill some of the gaps as many academic institutions mandate that their faculty and researchers must submit a version of their published outputs into the repository. This is not automatic, and a low percentage of documents find their way into the repository. Specific content, such as grey literature and federally funded sponsored research, contain a structure of guidelines, procedures and mandates in place. Repositories on the ArXiv model follow a less formal approach and enjoy higher contribution rates with thematic repositories in applied physics, physical sciences, biological and life sciences, medicine, psychological and cognitive sciences, information and

computer sciences and other large social science networks.

We are asking the question, what makes collections or content valuable? Collections are informed by different rubrics. For example, the arts inform and influence culture and history, psychology provides context for human relationships and behavior, science explains the power of the environment and potential for the future. New information is generated by knowledge building and creative forms of scholarship often lead to transformational changes, fixing and repairing the world with keener insights and understanding for safer, healthier living.

With new commitments to and promises of equity, diversity and inclusion (EDI) that pledge becomes how to “acknowledge and address historical racial inequities, challenge oppressive systems within academic libraries; value different ways of knowing and identify and work to eliminate barriers to equitable services, spaces, resources and scholarship.” (Munro, 2019) Definitions such as this demonstrate and explore how organizations, published outputs, the publishing process and knowledge generation must make very clear how barriers must be reduced to achieve accessible and open reading environments and content. The open movements aggressively continue to demonstrate this.

### **Ownership to access, format shifts and other models**

Librarians are practicing the “just in case vs just in time” model in defining their collections. Until this century, libraries traditionally collected in an anticipatory or forensic model hoping to have what readers and users wanted at the time of need. Today, there are processes in place to obtain content when requested either via purchase, licensing or borrowing from another library with quick turnaround. The staging of warehousing is a past practice as libraries have computed from usage analysis that often print holdings collect dust or are never accessed at all. Circulation data demonstrates that high percentages of print books have not circulated within a decade of being acquired. Many libraries are weeding and withdrawing their older, less used print volumes and send them to institutional or shared storage. The space previously occupied by print is being reconfigured for other purposes that contribute to student success, collaborations and creative learning.

Since this paper is focused on the sciences, where journal literature is critical, a different but analogous practice is employed. Libraries at least in North America are detecting how 80% percent of users are coming to their collections from Google. Although this lowers barriers to information, the search only takes the reader to an awareness level. Readers are often left with only a taste and not able to access the full content. Actual full text documents are hidden behind user authentication systems. Libraries may not process all open access content. Libraries are investing in digital backfiles, reducing the need for miles of shelving when users prefer to search and read online the shorter articles. Book chapters are often compared to journal articles or conference papers and eBooks are growing as the default format of choice at many libraries.

In years past, libraries traditionally structured their collections by separating print monographs, journals, special collections, and other resources. Librarians have the opportunity to create several teaching moments by expanding on the importance, appropriateness, and uniqueness of each format. Library users could physically handle the monographs, journals, etc. and explore the content of the print format and understand how each resource was a tool to craft the content of their research. It was clear how the differences of each format determined its usefulness in the research process.

Today, more libraries are blending all print resources in one classification system and shelving together. User expectations of library collections overwhelmingly value convenience and speed over tradition and structure. Often, the first exposure to library collections starts with a visit to the library home page. It is highly conceivable that researchers may not even visit the library in-person, especially given the popularity of online and distance education where students are scattered in locations across the globe and may not have the opportunity to travel to the brick and mortar campus to retrieve information needing instead to maximize discovery layers. The virtual presence of the library appears to be more important, than the physical structure. As MOOCs gain stature in continuing education and award credits and degrees, licensing content becomes a greater challenge and is a call for more OA content.

Our previously held notions of instructing library users on the traditional research tools contained within library collections has also evolved into library OPAC systems that favor a Google style search results view without regard to the unique nature of the formats of our traditional research collections. (Rennick, 2019) In the electronic environment, an ebook and its chapters has the same look and feel as a journal and its assemblage of articles. All results are jumbled together on the screen based on “secret sauce” search algorithms that rank all print and electronic library resources based on the user’s search. (Laine, 2019) Although our OPAC vendors have the best of intentions to create products that would minimize the learning curve and increase familiarity between internet search engines and library interfaces, something is lost when users cannot differentiate between books, journals, and web resources. If everything looks the same on the search results page, then users may consider that all library resource formats are all the same and they just need “the answer” regardless of the format. Even in our Internet driven environment, librarians add value by creating learning opportunities and inserting teaching moments to explain the structure of library collections. Filtering certainly helps but users need to select their options.

### **Context of Grey Literature for library collections**

Libraries are constantly under pressure to supply access to high-quality subscription-based resources to their researchers with static or shrinking budgets. It is rare to encounter subscription-based electronic resources that do not incur any annual renewal increases. Even for electronic resources that do not factor in a cost increase will often include an increase for inflation. With a static budget, libraries must make hard management decisions on which resources to keep, discontinue, or weed. (Shenoy, 2018) This dilemma of providing our users with a seemingly limitless supply for electronic resource subscriptions combined with tight budgetary constraints forces libraries to turn to alternatives such as grey literature.

Imagine a library adding or obtaining a world-class electronic resource that has volumes of high quality information available from any computer with an internet connection, free-of-charge. Library users, including community members, could access these resources. Almost every library jumps at the chance to provide their users access to such valuable free resources. The most salient example of high-quality grey literature in the United States is government information. Publicly available information shifts the power from commercial publishing to readers, who are the taxpayers, promising secure access to government funded resources. The costs of data collection and analysis is provided by the federal government through taxpayer dollars. Most unclassified government information in the US is to be available for citizens to

browse, often and increasingly electronically, through resources published by the Government Printing Office (GPO.gov) and managed by the Federal Depository Library Program (FDLP.gov).

Current threats by government agencies to not preserve all federally issued outputs is a serious issue today. Archiving and stewardship comes with a high price that not all current government leaders recognize as a priority. The consequences of this is worrisome because competitive research can't be conducted with gaps in time series.

To a scientist, government information such as population & demographics, health & nutrition, geography & the environment, R&D data, agriculture, forestry, fishing & mining, energy & utilities, manufacturing, materials & trade, transportation, information and communication are all critical elements for comparative data and new contribute to new forms and possibilities of innovation and entrepreneurship. The national census is a prime example of high quality, data rich grey literature freely available for not only citizens, but anyone with an internet connection. (Census.gov) Libraries also provide discovery of Census through online research guides specifically tailored to government information. Science depends on sharing data, adding value and making new conclusions that lead to health, prosperity and advance the future of a democracy.

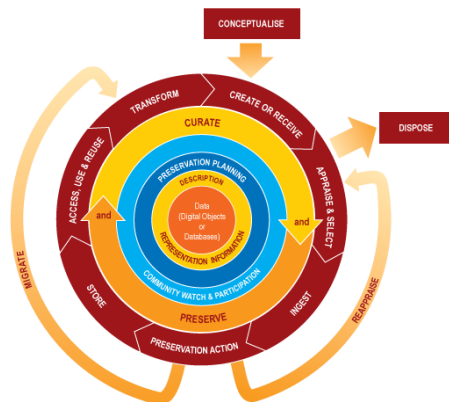
Resources such as publicly funded databases, compendium, union catalogs, indexing services, search engines, suggests how the majority falls outside the realm of grey literature. However, there is content within these resources that are clearly grey.

### **Research Lifecycle – different components**

Within academe, tracking the research process is critical. Documents created during the research process ideally maintain a “research lifecycle.” Several lifecycle models exist. The Digital Curation Centre Curation Lifecycle model is a detailed, yet accessible model that encapsulates the main points of the research lifecycle.

At the center of the model is the Data. We can consider the center as the “birth” phase of the lifecycle. The definition of data is purposefully broad and can include images, datasets, websites, text files, sound files, etc. The next category includes full lifecycle actions, in other words steps to be taken during the active life of the Data. These include the next rings of the circle expanding from the middle such as Description and Representation Information, Preservation Planning, Community Watch and Participation, and Curate and Preserve. The items on the outer rim of the circle are classified as sequential actions, one action must occur before the other action can occur. As we can infer from the model and our earlier work, these are largely management steps to maintain the data. Finally, the outlying items outside the circle are listed as occasional actions where “dispose” signals the end-of-life for the data and the “migrate” is the reincarnation of data in another scholarly space.





Key elements of the DCC Curation Lifecycle Model

Figure 1. DCC Curation Lifecycle Model

The management and curation of the data lifecycle in the sciences is key to innovation, scholarship, and historical value. Most research libraries collaborate internationally and have established campus scholarly data services to serve their research communities to properly curate data. The DCC was created in response to a need to create a national data curation center in the UK to promote best practices for such data curation. Within the organizations who have established data services centers, librarians act as facilitators and advisors for researchers who identify data management needs and match the appropriate tools and resources for the researcher’s data. They also provide guidance about how to make data findable, for citation, repurposing and reuse.

Kaley Leetaru, a popular cultural historian and technologist has opined, “The origins of data in computer science are just one indicator of its value. Future technology needs to blend the social aspects of privacy, civil liberties, and community impact to best understand the relationships between public data management and private data minimization.” (Leetaru) This suggests how data-centric and important the disciplinary mix is to understanding context.

### Climate Change and Global Warming: New Emphases of Grey Science

Fundamental examples of global collaboration that fuels grey literature includes many topics that have political, social and economic implications for the science it covers. The world is experiencing temperature and water rises at unprecedented rates. Climate change could be to blame for spreads of multidrug-resistant fungus that have been deemed by the Centers for Disease Control and Prevention to be a serious global health threat (CDC.gov). There are many examples of grey literature that accompany the study of the effects of climate change. Politicians and their followers have denied some of this information as science worthy. Instead speculations about science fiction describing these serious factors that the next generation must respond to permeate the modern press. To combat this, anticipated reports from the Third International Conference Climate Change held in February 2019 are anxiously awaited.

Previous reports from this conference series have been embraced as well as those from the UN documents. The impacts of climate change are being felt everywhere and are having very real consequences on people's lives. Climate change is disrupting national economies, costing us dearly today and even more tomorrow. However, there is a growing recognition that affordable, scalable solutions are available now that will enable us all to leapfrog to cleaner, more resilient economies. The latest analysis shows that if we act now, we can reduce carbon emissions within 12 years and hold the increase in the global average temperature to well below 2°C and even, as asked by the latest science, to 1.5°C above pre-industrial levels.

Thankfully, we have the Paris Agreement, even though our country has pulled out from being a signatory. Many people think this document is a visionary, viable, forward-looking policy framework that sets out exactly what needs to be done to stop climate disruption and reverse its impact. Regardless, the agreement itself is meaningless without ambitious action.

The Intergovernmental Panel on Climate Change is among the most influential bodies that have addressed climate change. Part of the UN for more than 30+ years, IPCC was "created to provide policy makers with regular scientific assessments on climate change, its implications and potential risks, as well as to put forward adaptation and mitigation options." (IPCC) The sheer volume of reports and data sent to many media outlets is all dated and numbered, widely available but may still be considered as grey by the nature of how it is disseminated. This new permutation of grey shows how the blending of traditional information sources with the grey intensify the values associated with it. The distribution channels and thus the availability is what determines how grey it is.

United Nations Secretary-General António Guterres is calling on all leaders to come to New York on 23 September with concrete, realistic plans to enhance their nationally determined contributions by 2020, in line with reducing greenhouse gas emissions by 45 per cent over the next decade, and to net zero emissions by 2050. We will see then what the next steps will be.

Erosion and earthquakes are other scientific processes that just happen. Predictive data may speculate when they may occur, but the catastrophes and potential loss of life caused by these "acts of God" are analyzed after the fact rather than from a causal perspective. That adds to the greyness and volatility of information value. MacDonald and colleagues concluded a decade ago that "The relevance of grey literature addressing global environmental challenges, such as that produced by intergovernmental organizations in our studies, warrants research engaging the wider and scientific and policy communities. The seriousness of global environmental conditions in the early 21<sup>st</sup> century demands interdisciplinary attention...and is leading to a greater understanding of information life cycles and barriers to the diffusion, use and influence of scientific information." (MacDonald, et al, 2010) This remains central to the unfolding evidence of climate change, global warming and their effects on the global environment.

### **Grey in Science / Science in Grey**

Systematic Reviews are the latest scholarly and increasingly high demand information product that promotes grey literature. A systematic review is a research method that is designed to answer a research question by identifying, coding, appraising and synthesizing a group of studies and is published in traditional literature where it is cited like any other article. It can take

upwards of a year to comb and review the literature and convincingly write up the findings that makes this information a new standard in professional practice that can be adopted becoming a new scientific or healthcare procedure. This method is utilized in many disciplines (medicine, public health, education, social sciences, and more) to answer a variety of questions (effectiveness of an intervention or policy, prevalence, tests/diagnosis, and more). Often systematic reviews must cite and include references to grey studies in order to be inclusive and exhaustive in subject coverage. In the meta-analysis and other methodologies such as rapid reviews, scoping reviews, medicine synthesis, that are performed, all relevant incidences of studies should be noted to provide examples of research that contribute to success as well as failure.

The medical community and increasingly social scientists and engineers are conducting systematic reviews to demonstrate how things are interdependent and these understandings and relationships are based on the outcomes of a thorough literature review. Margaret Foster in her publications and courses that describe best practices for systematic reviews in different disciplines argues for achieving balance in assembling the pieces of a systematic review to include creating the narrative and integrating the quantitative elements. She identifies the activities performed by librarians in authoring a systematic review as:

- Develop and validate search filters or hedges
- Participate or lead in the development standards for review methods
- Design software or other tools for reviews
- Create tutorials and other resources for systematic review authors
- Serve as peer reviewer for searches of review protocols or review reports under consideration for publications
- Develop new techniques for improving the efficiency of citation searching, retrieval, de-duplication or selection
- Track potential opportunities for funding systematic reviews
- Join grant proposal teams and funded research (Foster and Jewell, 2017)

### **Other Shades of Grey**

Language has taken on additional methods of conveyance. One of the most artistic forms of expression has been emoji. A recent posting on the Scholarly Kitchen blog begs, “If Emoji Isn’t a Language, Then What Exactly Does it Do?” (Crotty, 2019). The consensus from linguists is that emoji is not a language but is best described as a “set of gestures.” These forms of expression capture mood and emotion, ambiance and emit a sense of time and being in the moment and usually reflect a personal response to something that was said, shared, read, observed or communicated and received by one person. The analogy to facial expressions or body language and how that is understood makes for a solid comparison because it is nearly impossible without video to capture that moment and emojis have a longer lifespan and can be retained. Today we see books in emoji that are even in the collections of the Library of Congress and other libraries, classics now have emoji editions (Moby Dick, and reference tools like the *Emojipedia* (<https://emojipedia.org/>) document and transcribe categories like:

- Smileys & People
- Animals & Nature
- Food & Drink
- Activity
- Travel & Places
- Objects
- Symbols
- Flags

Incorporated in Twitter and other social media communities of Instagram and Snapchat among others, these visual emojis form a new nomenclature with its own taxonomy giving people a visual and a ‘warm and fuzzy’ or personal sense of what is being shared. The sentiments of joy or frustration can easily be conveyed and as a communication form it has had some staying power since it was conceived nearly a decade ago. In the here today, gone tomorrow mentality, this shows persistence. It is a form of enlightenment and creativity and whether linguists are convinced or dismissive if it is true science, art or another medium, the shades of grey permeate as something to be documented.

Cognitive scientists and linguists have raised the question of why are there different languages. If we accept the premise that “All languages share several basic design features, such as productivity, categorical denotation and compositionality,” (Lupyan and Dale, 2016), then we can explore how the social, ecological and physical niche each plays out and observe how the technological niche gains force. By adding technology’s contribution of software programs and packages such as Instagram or other visual platforms, a new vocabulary illustrates an idea or feeling with new imagery. “Emojis may not simply replace words but they create novel opportunities for self-expression through a type of depiction that is not possible by conventional written language.” (Instagram Engineering, Part 1, 2011) The experience of how Instagram adopted emojis is fascinating and attributes its rise to the emoji keyboard on iOS in 2011 followed by Android platforms in 2013 and the vocabulary has multiplied since then. Natural language processing has drawn parallels to translating emojis using intuition or distributional hypothesis. Due to this work in translation of emojis, “the vocabulary of Instagram is shifting similarly across many different cohorts with a decline in internet slang corresponding to rise in the usage of emoji.” (Instagram Engineering, Part 1, 2011). “Emoji usage is shifting the people’s vocabulary on Instagram and becoming an important means of expression: their use is anti-correlated with internet slang like “lol” and “xoxo.” (Instagram Engineering, Part 1, 2011). By observing words and emoji together we were able to discern representations of both. These representations can help us better understand their semantics and find distinctive characteristics of similar symbols.

If emojis are a form of depiction and depiction is a form of communication as Clark suggests, then these other shades of grey blend description with physicality or spatialization giving a sense of presence or time. Clark suggests that language processing accounts for descriptions and depictions, and thus writing must do the same. (Clark, 2016). Worldwide an increase in emojis used in Instagram texts suggests that a majority could soon surpass figures such as 38% in US postings and 60% from Finland. (Instagram Engineering, Part 1, 2011).

## Conclusion

Change is ongoing and the new normal. As Bryan Sinclair states, “Because digital technologies are assumed in all that we will do in our increasingly visual and connected future, perhaps the most basic definition of digital scholarship is that it is and will be the scholarship of the 21st century.” Sinclair,2014)

Grey literature is the growing branch that reaches out to a new generation of readers and information seekers. Not only does each new generation share different habits of reading and learning, it is curious and seeks to experiment with new enterprising aspects of conducting their professional and personal ways of communicating and applying information with new modes of technology. The born digital population has very different expectations for sharing, belonging, identifying with peers, colleagues and family. The connection between grey literature and communication reminds us of how science is defined by reproducibility, empirical evidence, logical transitions and authenticating previous work that leads to innovation. From the expansion of grey literature we can see that contemporary society is committed to a different set of values that reinforces and promotes unity, honesty, pluralism, empathy, and opportunity and that is what these new forms of grey literature do.

Grey literature, akin to other digital services serves community, leads the charge for change, generates buzz, fosters collaboration, and reduces barriers. For the research library ecosystem, grey becomes the new white as identifying, describing, finding, acquiring, citing and applying the grey has potential for addressing the “fake news” or “false science” that is all too common today. As Open Science continues to evolve science libraries and collections will reflect more grey content and it will be easier to acquire, process, and use. More visualization and augmented or virtual reality will explain and document the data rich and mixed format content from the new and emerging specializations. Museums will partner with libraries more to exhibit forms of the new grey and ways to preserve and archive this work will become increasingly important. Layers of new augmented and virtual reality will add a dimension to the information that is visual, animated, applied and subject to new modes of comprehension. Social media will continue to add forms of grey literature to enhance access and utility of the communication mediums, making them richer and more descriptive. As the interest grows to archive these sites, the new grey literature will naturally expand.

If Sinclair’s sentiments continue to ring true and libraries embrace and develop digital scholarship, new forms of grey will shape the landscape. The structure of library services and collections through this century will be defined by individual and collaborative contributions that illustrate the power of stories and emerging scientific breakthroughs that will have high impact and value for their creative and powerful reaches that document how science evolves within a greying universe of new discoveries with much data. Libraries will be the stewards for different media and findings, will accommodate the traditional, new and misfit elements and collections will blend with new services needed to create, articulate, curate and use the information that is produced. As the open movements expand with more resources and research procedures functioning in an open environment, libraries will be the centers where researchers will continue to learn about organizing and disseminating the digital grey literature their research teams produce. These new processes such as Open Peer Review and Open Educational Resources,

Open Science and hosts of others on the horizon force libraries to promote their roles and products in the mainstream. Science will lead new forms of grey content, in both open and quasi-open environments.

## References

Auger, CP, ed. (1989). *Information Sources in Grey Literature*, 2d ed. London: Bowker-Saur: 11.

Clark, HH. (2016). Depicting as a method of Communication. *Psychological Review* 123 (3): 324-347.

Crotty, D. (2019). "If Emoji Isn't a Language, Then What Exactly Does it Do?" Scholarly Kitchen Blog Post, August 2. <https://scholarlykitchen.sspnet.org/2019/08/02/if-emoji-isnt-a-language-then-what-exactly-does-it-do/?informz=1>

Davis-Kahl, S. (2019). Call for participation in Collections Work Survey. <https://bit.ly/2FCPgDj>

Davis-Kahl, S. (2019b). Adapted with permission from Joan M. Reitz. *Online Dictionary for Library and Information Science*. Santa Barbara, CA: ABC-CLIO, 2004.

Digital Curation Centre Curation Lifecycle. <http://www.dcc.ac.uk/resources/curatio-lifecycle-model>

Federal Depository Library Program. <http://www.fdlp.gov/>

Foster, M. (2019). LibGuide. <https://tamu.libguides.com/c.php?g=574702&p=3963347>

Foster, M. and Jewell, S. (2017). The Intersection of Librarians and Systematic Reviews, in *Assembling the Pieces of a Systematic Review: A Guide for Librarians*. Lanham, MD: Rowman & Littlefield: 204

Gelfand, JM (1993). Academic Libraries and Collection Development Implications of Grey Literature. First International Conference on Grey Literature, Amsterdam, Netherlands: December 14. *Weinberg Report 2000, Conference Proceedings*. Amsterdam, Netherlands, TransAtlantic, 1994, 123-141.

Gelfand, JM. and King, JL. (1998). Grey Market Science: Research Libraries, Grey Literature, and the Legitimization of Scientific Discourse in the Internet Age. Paper presented at the IEEE Socioeconomic Dimensions of Electronic Publishing Workshop, in cooperation with the 1998 IEEE International Conference on Advances in Digital Libraries (ADL'98), Santa Barbara, CA, April 24. In Nielsen, C.S. and J.R. Herkert, eds, *Proceedings, Socioeconomic Dimensions of Electronic Publishing Workshop*, IEEE, 1998, 115-20.  
<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=730719>

Gelfand, JM (1999). Grey Literature Poses New Challenges for Research Libraries. Fiftieth Anniversary of the Center for Research Libraries Conference on Cooperative Collection Development, Atlanta, GA: November 13, 1999. Conference Proceedings co-published simultaneously as *Collection Management*, 24, #1/2 and #3/4, 2000: 137-148 and as *Creating New Strategies for Cooperative Collection Development*, Papers from the Aberdeen Woods Conference, edited by Milton T. Wolf and Marjorie E. Bloss. Binghamton, NY: Harrington Park Press/Haworth Press, 2000.

Gelfand, JM (2001). The ETD: What Color is it - Black, White or Grey?" Third International Conference on the Electronic Thesis/Dissertation (ETD). Pasadena, CA: March 24, 2001.

Gelfand, JM (2005). Knock, Knock, Has Grey Literature Found a Home in Institutional Repositories? Sixth International Grey Literature Conference, New York, NY. December 6, 2004. *GL6 Conference Proceedings: Sixth International Conference on Grey Literature: Work on Grey in Progress 6-7*, December 2004, compiled by D. J. Farace and J. Frantzen. Amsterdam: TextRelease, January 2005: 10-16. Also in *The Grey Journal*, vol. 1 #2 Summer 2005: 61-66.

Gelfand, JM (2006). Grey Literature: Taxonomies and Structures for Collection Development. Eighth International Grey Literature Conference, New Orleans, LA. December 4, 2006. *GL8, Conference Proceedings: Eighth International Conference on Grey Literature: Harnessing the Power of Grey, 4-5 December 2006*, compiled by D. J. Farace and J. Frantzen. Amsterdam: TextRelease, February 2007: CD-ROM. Also published in *The Grey Journal*, vol. 3 #1:7-16 Spring 2007 <http://www.greynet.org/thegreyjournal.html>)

Gelfand, JM (2009). New Shades of Grey: The Emergence of E-Science, Scientific Data and Challenges for Research Libraries. Eleventh International Conference on Grey Literature, Washington, DC. December 12, 2009.

Gelfand, JM (2009b). Changing Collections for eScience: What it Means for Libraries. IFLA Acquisitions and Collection Development Section Mid-Year Meeting, China Academy of Sciences, Beijing, China. February 26

Gelfand, JM and Lin, AB (2012). All Agents of Instruction and Communication: The Intersections between Libraries, Archives and Museums via Social Media. Invited Paper, Fifth Rizal Library International Conference, Manila, Philippines. October 28.

Gelfand, JM and Lin, AB (2013). Research Life Cycle: Exploring Credibility of Metrics and Value in a New Era of eScholarship that Supports Grey Literature. Presented at the International Conference on Grey Literature, Rome, Italy. November 30, 2012. *The Grey Journal* 9 (3) Fall 2013. [http://www.greynet.org/images/GL14-S4P\\_Gelfand\\_and\\_Lin.pdf](http://www.greynet.org/images/GL14-S4P_Gelfand_and_Lin.pdf)

Heidorn, P. B. (2011). The emerging role of libraries in data curation and e-science. *Journal of Library Administration*, 51(7-8), 662-672.

Holzman, A. and Lippincott, S.K. (2019). "Libraries" in Phillips, A. and Bhaskar, M., eds., *Oxford Handbook of Publishing*. Oxford University Press, 379-398.

Instagram Engineering, Part 1, (2011). Machine Learning for Emoji Trends. <https://allthingslinguistic.com/post/124609017512/emojineering-part-1-machine-learning-for-emoji>

Instagram Engineering, Part II (2015). Implementing Hashtag Emoji. <https://instagram-engineering.com/emojineering-part-ii-implementing-hashtag-emoji-7b653b221c82>

Intergovernmental Panel on Climate Change (IPCC). <https://www.ipcc.ch/>

Laine, T. and Laitinen M.A. (2019). The Finna Service: Meeting the New Measurement Challenges in libraries. *Library Management*, 40(1/2), 2-11.

Leetaru, K. (2019). Computer Science Could Learn a lot from Library and Information Science. *Forbes*, August 5. <https://www.forbes.com/sites/kalevleetaru/2019/08/05/computer-science-could-learn-a-lot-from-library-and-information-science/#5537bfa7587d>

Luxembourg Definition (1997). Third International Conference on Grey Literature. Luxembourg.

Lupyan, G. and Dale, R. (2016). Why are There Different Languages? The Role of Adaptation in Linguistic Diversity? *Trends in Cognitive Sciences*, 1594. doi: 10.1016/j.tics.2016.07.005

MacDonald, B.H., Wells, P.G., Cordes, R.E., Hutton, G.E.G., Cossrini, D.M. and Soomai, S.S. (2010). The Use and Influence of Information Produced as Grey Literature by International, Intergovernmental Marine Organizations: Overview of Current Literature. In Farace D.J., Schopfel, J., eds., *Grey Literature in Library and Information Studies*. Berlin: De Gruyter Saur: 167-180.

Munro, K. (2019). Renewing the System: Thinking Broadly About Equity, Diversity and Inclusion in Scholarly Communication. *C&RL News*, July/August: 374.

New York Definition (2004). Sixth International Conference on Grey Literature. New York.

Rennick, B. (2019). Library Services Navigation: Improving the Online User Experience. *Information Technology and Libraries*, 38(1), 14-26.

ROAR:Registry of Open Access Repositories. <http://roarmap.eprints.org/>

Schopfel, J. (2011). Towards a Prague Definition of Grey Literature. Twelfth International Conference on Grey Literature: Transparency in Grey Literature. Grey Tech Approaches to High Tech Issues, 6-7 December 2010, Czech Republic

Shenoy, P., Krossner, L & Koch, T. (2018). Money Doesn't Grow on Trees: Using a Data-Driven Review Process to Add New Resources With No Budget Increases. Charleston Conference Paper. DOI: [10.5703/1288284316664](https://doi.org/10.5703/1288284316664)

Sinclair, B. (2014). The University Library as Incubator for Digital Scholarship. *Educause Review Online*, June 14. <https://er.educause.edu/articles/2014/6/the-university-library-as-incubator-for-digital-scholarship>

Tsang, D. and Gelfand, JM (2016). "The Changing Landscape of Research Library Collections: Ensuring Realistic Sustainability." Presentation. Academic Librarian 4 Conference, Hong Kong University of Science & Technology, Hong Kong. June 2-3, 2016. <http://escholarship.org/uc/item/Osk1f0pz>

United States. Census Bureau. <https://www.census.gov>

United States. Centers for Disease Control and Prevention. <https://cdc.gov>

United States. Government Printing Office. <https://www.gpo.gov>

Watkinson, C. (2016). "Why Marriage Matters: A North American Perspective on Press/Library partnerships" *Learned Publishing*, 29: 342-47. <https://doi.org/10.1002/leap.1044>