BIBFRAME, Europeana and DPLA: The Future of Open Cultural Heritage?

Laura Brown
Pratt Institute

Ellie Horowitz
Pratt Institute

Emory Johnson
Pratt Institute

Meredith Powers
CUNY York College

Sarah Quick
Pratt Institute

How does access to this work benefit you? Let us know!

Follow this and additional works at: https://academicworks.cuny.edu/yc_pubs

Part of the Cataloging and Metadata Commons

Recommended Citation
Brown, Laura; Horowitz, Ellie; Johnson, Emory; Powers, Meredith; and Quick, Sarah, "BIBFRAME, Europeana and DPLA: The Future of Open Cultural Heritage?" (2014). CUNY Academic Works.
https://academicworks.cuny.edu/yc_pubs/200

This Working Paper is brought to you for free and open access by the York College at CUNY Academic Works. It has been accepted for inclusion in Publications and Research by an authorized administrator of CUNY Academic Works. For more information, please contact AcademicWorks@cuny.edu.
BIBFRAME, Europeana and DPLA:
The Future of Open Cultural Heritage?

Cultural Heritage: Description and Access

LIS 670-01 Spring 2014

Laura Brown, Ellie Horowitz, Emory Johnson, Meredith Powers, Sarah Quick
Table of Contents

1. Introduction
2. Background and Objectives
   2.1. Europeana
      2.1.1. Audience
      2.1.2. Interface and Functionality
      2.1.3. Partners
      2.1.4. Knowledge Structure and Organization
         2.1.4.1. Europeana Data Model (EDM)
         2.1.4.2. Linked Open Data
   2.2. The Digital Public Library of America (DPLA)
      2.2.1. Audience
      2.2.2. Interface and Functionality
      2.2.3. Partners
      2.2.4. Knowledge Organization and Structure
         2.2.4.1. DPLA Metadata Application Profile
         2.2.4.2. Data Harvesting
         2.2.4.3. Linked Open Data and API
   2.3. BIBFRAME
      2.3.1. Metadata Standards
         2.3.1.1. MARC (Machine Readable Cataloging Standards)
         2.3.1.2. RDA (Resource Description and Access)
         2.3.1.3. FRBR (Functional Requirements for Bibliographic Records)
         2.3.1.4. RDF (Resource Description Framework)
      2.3.2. Goals and Model
      2.3.3. Entities
      2.3.4. Conversion and Tools
      2.3.5. Case Studies
         2.3.5.1. DNB (Deutsche Nationalbibliothek)
         2.3.5.2. British Library
         2.3.5.3. George Washington University
3. Challenges
   3.1. Access
   3.2. Outreach, Funding, Sustainability
   3.3. Openness
4. Opportunities
5. Conclusion
6. References
The digital environment has changed the way that museums, libraries and archives provide access to their collections. Traditionally these types of ‘memory institutions’ have each operated in their own discrete intellectual spaces, employing customized metadata standards, organizational practices and methods of engaging their users (Trant, 2009). Over the last couple of decades, millions of items from these institutions have been digitized and made available on the Web, enabling users to globally access cultural objects in a virtual space, where once they would only have been able to interact with them in person. While the metadata describing digital objects remains heterogeneous, a commonality has emerged in digital objects’ shared technologies and delivery platforms (Trant, 2009), setting the groundwork for a convergence between the holdings of diverse memory institutions that was once thought impossible.

Simultaneously, the information behaviour and needs of users have evolved as they enter the digital research environment. Digital tools are expected to operate seamlessly and an institution is expected to go beyond simply providing one-dimensional access to their collections (Trant, 2009). Increasingly instead of simply consuming information, users expect to interact with, re-use and produce it, too (Valtysson, 2011). The creative audience of Web 2.0 demands that the online culture transforms from read-only to read-write, and that institutions increase the trust they place in their audiences (Valtysson, 2011).

Examples of cultural heritage projects that have emerged in the last ten years, include two mass-digitization projects, Google Books and HathiTrust. These resources host and provide
online access to millions of items. Google Books, launched in 2004 with the intention of scanning all known existing books, was believed to have (as at November 2013) 30 million books in its database (Eichenlaub, 2013). Google’s objectives and incentives in this project are not completely understood and there are concerns about such a vast trove of material being held by a corporate entity which could at any time seek to monetize its pursuit or become commercially unviable and disappear (Eichenlaub, 2013). HathiTrust began in 2008 as a response to these fears with a focus on preservation and providing access to digitized books and journal content provided by partner collections, including Google Books and the Internet Archive. As of November 2013, HathiTrust was the host of more than 10.7 million total volumes and more than 5.6 million book titles (Eichenlaub, 2013).

While the access these services provide is unprecedented, three more recent projects demonstrate a move towards greater institutional convergence, data openness and read-write culture. The first, Europeana, launched in 2008 as Europe’s response to Google Books, provides access to 30 million records from 2,300 European cultural institutions. The second, the Digital Public Library of America (DPLA), launched in 2013, provides access to over 7,000,000 items and counting from American cultural institutions. Both of these organizations aggregate and make accessible the metadata records of digital objects through their portals, while the digital object themselves remain housed in the online repositories of partner institutions. The collections include objects in various formats from various types of memory institution, such as manuscripts, sound recordings, photographs, and more. Europeana and DPLA both face the challenges of making these records interoperable on a syntactic and semantic level (Hyvönen, 2012). Employing semantic web models and developing their own, they have sought to
harmonize these metadata records and simultaneously enrich the existing data to provide links between people, places, concepts, etc. Both organizations have also made their massive datasets available for re-use through APIs, enabling users to write their own content with it.

Finally, a third project, BIBFRAME (or or the Bibliographic Framework Initiative) was released by the Library of Congress in 2012. Drawing on influences, such as MARC, FRBR and RDA, BIBFRAME provides “a foundation for the future of bibliographic description, both on the web, and in the broader networked world” (Library of Congress, n.d.). While not yet broadly adopted, if properly instituted BIBFRAME will greatly increase the access to information that is normally captured within libraries’ individual catalogs and connect it to information contained across other memory institutions.

In this paper, we discuss these three recent initiatives in detail. We will make clear their contextual backgrounds, objectives and functionality, and compare the common challenges they face as well as opportunities for the future.

2. **Background and Objectives**

2.1. **Europeana**

Europeana is “an internet portal that acts as an interface to millions of books, paintings, films, museum objects and archival records that have been digitised throughout Europe” (Wikimedia Foundation, 2014). Launched by the European Commission in 2008, its mission is to “create new ways for people to engage with their cultural history, whether it's for work, learning or pleasure”. Its vision encompasses the cultural goals of “making cultural heritage
Jon Purday of the Royal Library, The Hague, (now listed on Europeana website as ‘Head of Communications’) described the path of Europeana from “concept to construction” in his detailed 2009 report for The Electronic Library. He explains that the concept of a European Digital Library was conceived by several heads of state, sparked by Google’s announcement of the launch of its Google Books project in 2005. There was a concern that Google’s project would transfer “a significant amount of public domain intellectual resource into the private sector” and a motivation to complement Google’s Anglophonic content with European content in European languages (Purday, 2009). This was coupled with the recognition that Web 2.0 had advanced users’ expectations, requiring cultural heritage institutions to find new ways to remain relevant and to meet user needs when searching and retrieving materials. The concept was “to create a space in which all manifestations of Europe’s cultural and scientific heritage could be connected and integrated within a single portal, in a multilingual environment” (Purday, 2009).

As at November 2013, Europeana contains over 30 million records from 2,300 institutions from 36 countries (Europeana, n.d.). In an interview, Jill Cousins, executive director of Europeana, explained that “the whole Europeana concept is not about creating a destination site in Europeana.eu but about distributing the aggregated data into other systems, mobile applications... so that content can be used in many different ways and sustain different ways of looking at the material...” (Cranfield, 2012). The four key principles that “make up Europeana thinking”, as set out by Jill Cousins and on Europeana.eu are set out in Figure 1:
Jill Cousins further explains there has been a movement from aggregating a “critical mass of content” at the launch of Europeana, to “distributing the resources to others” in more recent years (Cranfield, 2012). One of the most notable movements was in 2012 when Europeana opened up its dataset of over 20 million items for free re-use.

2.1.1. Audience

The audience of the Europeana portal was actually determined as a result of extensive research undertaken during the planning stages of the portal development. Purday explained that the efforts were led by Daniel Terrugi of the Institut national de l’audiovisuel, and that the following user groups were identified:

1. General users
2. School students
3. Academic users
4. Expert researchers
5. Professional users

These users have varying levels of domain knowledge, search methods and skills, expectations for information presentation and re-use, etc. This research also provided four types of objectives:

1. The user wants to be entertained.
2. The user wants to know about a cultural or historic subject or person.
3. The user wants to know the current whereabouts of cultural heritage materials.
4. The user wants to be part of a community of interest.

With this understanding, end user research was undertaken with methods including focus groups and surveys to gain further insight into user behaviour in the fields of searching, refining, browsings, saving, personalising, tagging, sharing and community building (Purday, 2009).

2.1.2. Interface and Functionality

The Europeana portal was designed to be “engaging, reflect a diversity of content in a contemporary framework, and centered on the search box” (Purday, 2009). Hosted at europeana.eu, the portal is fronted with a homepage that is simple in design and does indeed have a prominent search box as well as an engaging banner inviting users to browse the website of the current online exhibition: Europeana 1914-1918. Until Other items of note are articles from the blog, featured items, featured partners and links to social media platforms, such as Pinterest, Twitter and Facebook. The user is given the option to sign up to receive the Europeana newsletter and to create an account in My Europeana - the benefits of which include the ability to save searches, items, tags and API Keys.
Few (if any) studies have examined Europeana’s interface and search and retrieval functions in detail. Inevitably, looking at these elements brings up questions regarding metadata, language, etc. which will be discussed in more detail later. One simple examination of search and retrieval undertaken by Rachel E. Scott in her article for Music Reference Services Quarterly provided some insights, which a user would experience when conducting individual simple searches. With regard to search and retrieval, there is no ‘advanced search’ but searches can be specifically focused on options chosen in the search box dropdown menu: Titles, Creators, Subjects, Dates/Periods and Places. There is no option to browse items - a search begins with typing the name of a person, place, subject, etc. Typing ‘picasso’ brings up several retrieval options as displayed in Figure 3. This highlights the metadata discrepancies between
institutions, such as name variations. Each of these variations only brings up records with that exact presentation (Scott, 2013). Scott, when searching for the composer Erik Satie, additionally discovered discrepancies in name ordering. This is clearly problematic and impacts on usability of the portal and users’ ability to conduct complete searches.

![Europeana search box](image)

Figure 3. searching for picasso using the Europeana search box

Search results are displayed in a grid formation with a thumbnail image - the order of results is not explained (it seems to be by provider) and often the records are missing images, in which case the thumbnail is a stock placeholder. Results can be refined by keyword, media type (image, sound, text or video), language of description, year, providing country, copyright status, provider, data provider. Scott found the ‘sound’ limiter unhelpful - for music scholars, it is important to differentiate between sound recordings and scores. Upon selecting an item of interest, a user is brought to a brief metadata record. In order to view the complete record, media, higher resolution image etc. the user must click through to the provider’s website, which will be in the language of the provider institution.
The majority of Europeana’s exhibitions are hosted on a separate website, exhibitions.europeana.eu and Scott found these to be “beautifully curated, but simplistic” (Scott, 2013). The homepage displays 31 exhibitions that include Darwinism: Reception in Spain and Catalonia, European Sport Heritage, and Leaving Europe: A new life In America (a collaboration with DPLA). The exhibitions are image-heavy, visually appealing and within an exhibition, the user is able to browse individual items, navigate by theme and (for some) to view the themes on a map. Scott found the map feature to be “somewhat cumbersome” and “misleading” as it shows the locations of materials contributed by providers without necessarily being relevant to the exhibition itself (Scott, 2013). Users can contribute comments on some exhibitions through a somewhat uninviting comments form.

![Figure 4. Roma in Festa exhibition Themes Map](image)

In a development aimed to more successfully engage users, Europeana launched its *Europeana 1914-1918* exhibition, a collaboration between Europeana and DPLA, which brings together stories from the public, national collections and film archives (Europeana, n.d.).
Cousins, in a press release states that “Europeana brings a new approach to cultural history, linking people's own stories to the official histories and showing the many-sided views of the same slice of history” (Europeana, n.d.). As mentioned above, the previous exhibitions had been “read only” with an uninviting comments form for users to provide feedback and a somewhat static level of interaction. Hosted on a separate website: http://europeana1914-1918.eu/en, Europeana 1914-1918 provides an interactive format and an increased level of community participation. Users are invited to ‘Add your story’ and there are posts about recent news and upcoming events. The exhibition can be searched or browsed (by type, subject of content or particular war fronts) and the content can be refined by whether it was provided by a user or a library, with further filters available upon selection of one of those options.

2.1.3. Partners

During the development stage of the Europeana prototype portal, research was conducted to gauge the type of data that Europeana would be collecting from its partners, such as subject matter, quantity, file formats and metadata standards (Purday, 2009). It was decided during this time that on the Europeana portal, users would see and search the surrogates of digital objects with the full object being housed at the provider’s website - one of the advantages of this is that Europeana avoids the need for massive servers to house duplicate content.

In preparation for the launch of the prototype, Europeana worked closely with content providers to manually harvest their datasets - this process enabled Europeana to develop mapping standards. and produce the Specification for the European Semantic Elements (ESE) (discussed below). Managing individual content from each of the 2,300 institutions/content
providers would be an overwhelming duty for Europeana’s relatively small team. Therefore, Europeana works with numerous aggregators to “collect the data on a national, thematic, or project-based level, harmonise it, and deliver to Europeana” (Europeana, n.d.), reducing the number of channels through which data is delivered. Each type of aggregator can represent a geographic (local, national or European) or domain level (cross, single, thematic) (Europeana, n.d.). A full list of Europeana’s providers can be found here:


Data is submitted through a multistage procedure, during which institutions are encouraged to work closely with the Europeana Aggregation team. A guideline of the process is here: http://pro.europeana.eu/procedure, which includes legal and technical requirements. A legal requirement is that institutions must sign the Data Exchange Agreement which sets out the relationship between Europeana and data providers, particularly with regard to the CC0 licence that applies to third party re-use of descriptive metadata, and, Europeana’s use of previews provided to them, which are not included in the CC0 licence and must, therefore, include a copyright statement (European, n.d.). The technical requirements are that the metadata will comply with the Europeana Data Model (EDM) or ESE (both discussed in more detail below), that the metadata records for each object must include a URI to the digital object on a provider’s website, and that metadata can be harvested through Open Archives Initiative Protocol for Harvesting Metadata (OAI-PMH) requiring that a provider has implemented this in their system (Europeana, n.d.).

Metadata is indexed, ingested and displayed in the language in which it was submitted by a data provider. Therefore, if a metadata record is provided in Greek, it will be indexed and will
be retrievable in Greek. A user is able to select one of 37 languages when using Europeana. Providers are encouraged to provide metadata in more than one language, using the XML attribute ‘xml:lang’ in all metadata elements in which it is able to provide multilingual metadata. Doing so ensures increases the visibility of a provider’s objects across languages.

2.1.4. Knowledge Structure and Organization

2.1.4.1. Europeana Data Model (EDM)

Europeana’s first data structure, ESE, was based on unqualified Dublin Core with a few Europeana specific elements. ESE records were structured in a flat list, which made linking to other resources within Europeana extremely difficult. ESE also caused confusion by not allowing institutions to distinguish between their collection’s physical objects and their digital surrogates. As a solution to these shortcomings, Europeana developed EDM using Resource Description Framework (RDF) with the help of leaders in the library, archive and museum fields.

As an alternative to creating a whole new set of elements, EDM’s developers chose to use existing vocabularies, such as Dublin Core, Open Archives Initiative Object Reuse and Exchange (OAI-ORE), Simple Knowledge Organizatin System (SKOS) and CIDOC-CRM http://labs.europeana.eu/api/linked-open-data/data-structure/, in order to ensure its interoperability and to reduce the cost of development. The structure of EDM is compliant with OAI-ORE, which allows institutions to provide Europeana with different metadata sets for a physical item and it’s digital surrogate, allowing the information to be “distinct, yet together.” Like ESE, this metadata is represented in Dublin Core with additional Europeana specific elements. The difference is that EDM allows contextual information (places, people) and
conceptual relationships (using the SKOS vocabulary) to be represented in the data and linked to additional, related cultural objects. Because of the emphasis on linking, the EDM metadata rules are more focused on proper RDF structure than providing specific direction on the content itself, which has resulted in certain issues with the quality of the metadata provided by partners.

When submitting a record in EDM format, the institution will provide a “bundle” that includes three core “classes” to represent a cultural heritage object (CHO), each with both mandatory and optional elements. edm:ProvidedCHO represents the physical object, such as a painting or sculpture, while edm:WebResource represents the digital image of that object. Each will have its own metadata, but will be linked by ore:Aggregation.

![Diagram of three core EDM classes](image)

Figure 5. Visualization of the three core EDM classes for data provider

In addition to the core classes, there are also four contextual classes including edm:Agent, edm:Place, edm:TimeSpan and skos:Concept. These allow the institution to not only provide data separate from the object, but provide data linked to an authority record, allowing the user to discover other, similar objects held in institutions all over Europe. In figure 6, we see an example of an EDM record for Leonardo’s *Mona Lisa*. Information such as preferred spelling and time frame have been included using the contextual classes. Linked open data will be discussed in more detail below.
By the time Europeana introduced EDM, over 20 million ESE records had already been provided. These records were not discarded but rather converted using a simple mapping system which can be found here: http://pro.europeana.eu/moving-to-edm. Institutions were given the choice to allow Europeana to handle the conversion of their data, create a new map using the source data (not recommended), use their originally provided elements to map to EDM, or take advantage of the new elements to create a more detailed record. Although there was no obligation, this was a good opportunity for institutions to improve the richness of their records, making them and their collection more discoverable. Europeana is still accepting records in ESE, but will now convert them before making them available online.

2.1.4.2. Linked Open Data

In September 2012, Europeana announced that it had opened up its dataset of over 20 million cultural objects for free re-use under a Creative Commons CC0 Public Domain
Dedication, “meaning that anyone can use the data for any purpose - creative, educational, commercial - with no restrictions” (Europeana, 2012). The goals of this action were in keeping with Europeana’s mission in term of economics: to provide “a new boost to the digital economy”, and culture: demonstrating a “move away from the world of closed and controlled data” as an example for memory organizations. Details of the dataset and the opportunities for using and re-using it are housed here: http://labs.europeana.eu/ with the instruction: “This is your code, this is your heritage, these are your labs”.

EDM’s core classes, as mentioned above, can be enriched by certain properties to allow for the development of a network of EDM resources. For example, edm:providedCHOS URIs are the main entry points of the dataset. With the addition of 'owl:sameAs' statements, items are linked with other resources about the same object. Currently no metadata is attached to the URIs themselves. Instead, metadata is attached to the proxies that represent a view of the object. Properties for Europeana providers’ proxies include ore:proxyFor, which connects a proxy to the item it represents a facet of, and ore:proxyIn, which attaches a proxy to the aggregation that contextualises it. Europeana’s proxies, ore:proxy, can be connected to items they represent using the same properties as above for Europeana providers, and can also be enriched by normalized dates, generic EDM fields taken from ESE and statements connecting the proxy to external sources, such as GeoNames (for places), DBPedia (for names of people), GEMET (for concepts) and Semium Time vocabulary (for time periods) - this is primarily done through Annocultor (Europeana, 2014). The class ore:Aggregation, mentioned above, provides data regarding a Europeana provider’s aggregation of resources for an item. The property edm:object, for instance, relates an item to a digital resource representing it and the property edm:isShownBy
relates the item to a webpage displaying the object in context. Europeana’s aggregation, edm:EuropeanaAggregation, groups together the result of all data creation and aggregation efforts for a given item, including a provider’s aggregation, which would be connected to a provider’s proxy (Europeana labs, n.d.).

The dataset is available in several ways. It can be downloaded in RDF form: http://data.europeana.eu/download/2.0/; queried through a SPARQL endpoint: http://europeana.ontotext.com and through two APIs: REST-API which is “suited for dynamic search and retrieval of our data”; and a second API which is “more experimental and supports download of complete datasets and advanced semantic search and retrieval of our data via the SPARQL query language” (Europeana labs, n.d.).

Europeana provides case studies on Linked Open Data Uses: the Linked Logainm dataset and Location LODer, which used Europeana’s linked open data to feed its Irish bilingual place names database; and on API Implementations, of which there are many, including ATHENA, a project “developing metadata exchange standards for the museum domain and aggregating museum content for Europeana”, the Digital Humanities Observatory, which uses Europeana’s API implementation to enrich its search results of e-resources in the arts and humanities (europeana, n.d.)

2.2. The Digital Public Library of America (DPLA)

With a mission to freely connect the country’s digital cultural heritage artifacts and records together via the Web, the Digital Public Library of America (DPLA) is moving into its second year of existence with great support and a sense of momentum. The DPLA, which “strives to contain the full breadth of human expression, from the written word, to works of art
and culture, to records of America’s heritage, to the efforts and data of science,” (DPLA, n.d.) is currently implemented at the web address dp.la, where users can search for over 7,000,000 digital items from libraries, archives, and museums across the U.S.A. via its online portal. Metadata records are also freely available via an API platform. Within the first year of implementation, the organization has seen increasing participation on the state and institutional level, as well as an increase in funding (as of March 2014, it was most recently awarded a $594,000 grant from the Andrew W. Mellon Foundation to research and pursue a sustainability model). Today it is clear that the DPLA has arrived, and is growing swiftly and strongly.

While the DPLA—the Digital Public Library of America—may sound like a digital national library offering its own holdings (as if it were a branch of the Library of Congress, perhaps), it should be made clear that the non-profit relies on the participation of public state and local libraries, institutions of higher education, archives, museums, and other cultural heritage institutions across the country to help populate its portal and data platform through metadata aggregation. Much like Europeana, when a user searches for an item in DPLA’s portal, she will find a record that links back to its holding institution, where the digital item is actually located. State or regional digital libraries, or “service hubs” “aggregate the metadata records from institutions in each state or region and then donate those records to the DPLA” (Scardilli, 2014). “Content hubs” “are large institutions that provide more than 250,000 metadata records directly to DPLA’s platform, and they maintain and edit the records themselves” (Scardilli, 2014). DPLA’s partnerships with both service and content hubs are its lifeblood; without them, the organization could not meet its mission in containing “the full breadth of human expression,”
and more pragmatically, users wouldn’t turn up any search results while using its portal or platform. These metadata records are DPLA’s holdings.

A project of this scale—one that requires participation from so many institutions across the country—is unprecedented in the United States. While a global union catalog like WorldCat has existed for some time, the DPLA differs in its vision to provide not only the locality of an item, but easy access to the digitized item itself. This portal model was first established by the Europeana Foundation, whose mission is to provide similar access to the millions of digital artifacts scattered across Europe’s cultural heritage institutions. In abiding by the Europeana Data Model (EDM), the DPLA aims to make their metadata records interoperable with Europeana, which will essentially create a portal to a huge portion of the Western world’s digital cultural heritage artifacts and records—quite an impressive feat.

In this section, we will strive to further describe its relationship with user communities and institutions in this country; uncover and clearly elucidate DPLA’s metadata model and standards, and how its metadata lives outwardly in the world; and finally investigate any challenges DPLA has faced already, and could potentially face as it continues to grow. Although DPLA appears to be developing quickly and robustly, it is still an extremely young organization—and this makes it difficult to be too critical of its practices. Through this report, we will instead offer a coherent overview of its seemingly most vital operations, and how those operations materialize in practice.

2.2.1. Audience
In name alone, the Digital Public Library of America plainly states the user group to whom it intends to serve: the American Public. Its mission somewhat clarifies the notion of the “American Public” as “students, teachers, scholars” (DPLA, n.d.), stressing its potential use in an educational environment—whether it’s a K-12 classroom or scholarly research lab—while still equally laying emphasis on the general public as an intended user group. Further, the DPLA states that it will be “an advocate for a strong public option in the twenty-first century” (DPLA, n.d.). In a digital world ruled by Google, the DPLA intends to be the mostly-publicly-funded option for digitized cultural heritage information retrieval in the United States (this perhaps in contrast to something like the Google Cultural Institute).

2.2.2. Interface and Functionality

So what is the DPLA like in practice? From a user’s perspective, the DPLA offers a somewhat simple interface to search for materials. Its homepage highlights its search portal and three browsing or “exploring” functions (by exhibition, by place, and by date). When searching for a material, the user can refine their results by format (text, image, sound, etc.), contributing institution, partner, date, location, and subject. For example, when searching for “Fitzgerald, Ella,” the user can refine their results to just sound recordings. Upon clicking on the first result, the user is brought to DPLA’s record page, which includes metadata like Creator, Created Date, Partner, Contributing Institution, Description, and so on (Figure 7). From here, the user has the option of accessing the object directly on its contributing institution’s page (in this case, from the National Museum of American History of the Smithsonian Institution). Additionally, because the user has the option of creating a profile, they also have the option of “saving” searches and
materials for later, and creating lists of those saved searches and materials. This function will certainly be of use to scholars, and we imagine younger students also engaging with this tool for school projects. There are no bells and whistles hanging off of the DPLA search portal, and this contributes to a comfortable user experience and unambiguous access to materials.

The DPLA offers a few different ways to browse its aggregated materials. First, it has created a digital exhibitions page that compound items into discernable, browseable collections. As of April 2014, some of these exhibitions include *Activism in the U.S.*, *This Land is Your Land: Parks and Public Spaces*, *America’s Great Depression and Roosevelt’s New Deal*, and *Leave Europe: A New Life in America* (a collaborative exhibition with Europeana). Each exhibition includes general contextual information as well as background information on each item. Exhibitions might be further broken down into smaller, thematic collections (for example, *This Land is Your Land* consists of six distinct collections).
Secondly, users have the option of browsing by location. Using an OpenStreetMap application, users can browse items on the state level, and upon zooming in, on the county, town, and institutional level (Figures 8, 9).

![Figure 8. All items in the U.S.](image)

![Figure 9. Users can browse items on the county, town, and institutional level.](image)

When browsing using a specific search term—again, we will use “Fitzgerald, Ella”—results appear in accordance to where their contributing institutions are located in the United States (Figure 10). Results will only appear if location data is available. For scholars in particular, this browsing function may shed new light onto where exactly many items and collections are held, and open the door to new research opportunities.

![Figure 10. Four items associated with “Fitzgerald, Ella” are represented on the map.](image)
Lastly, users can browse by time. Using a timeline, the user can view a timeline that stretches from 1000 A.D. until today, and zoom in on one decade at a time. Items created in that decade will appear to the user. When searching for “Fitzgerald, Ella,” the timeline visualizes the years in which items relating to “Fitzgerald, Ella” were created, and clicking on those years will exhibit short records of those items (Figure 11). Again, this visualized manner of browsing and searching may illuminate as to when exactly many items were created at once, and may lead to further questions as to why they were created when they were.

![Figure 11. Items associated with Ella Fitzgerald according to their creation date.](image)

However, despite these leaps in browsing functionality, the user must remember that the DPLA is and will continue to be in the process of aggregating materials. Browsing through the map and timeline may give an idea as to where and when items are located and created, respectively, but the user should keep in mind that these results are not wholly representative; there will always be materials and institutions that are under-represented or unrepresented completely. Additionally, the map and timeline only illustrates results that contain location and date creation metadata. Because visualizations make data easy to swallow, it is recommended
that one use these browsing tools with a grain of salt. At this point in time, a user might be better off using DPLA’s traditional portal for searching and browsing.

2.2.3. Partners

As previously stated, the DPLA garners its metadata records from contributing institutions from across the United States, and these institutions either deliver their records to a Service Hub or Content Hub. “The DPLA Digital Hubs Program is designed to establish a national network out of the over forty state/regional digital libraries and myriad large digital libraries in the US, bringing together digitized content from across the country into a single access point for end users, and an open platform for developers” (DPLA, n.d.). The DPLA Digital Hubs Program is, then, the heart of DPLA itself.

DPLA Service Hubs are state or regional libraries that aggregate resource descriptions from across their state or region, harmonizes them with the DPLA-MAP (to be discussed), and then contributes these records to DPLA. (DPLA maintains these records.) Further, Service Hubs offer “its state or regional partners a full menu of standardized digital services, including digitization, metadata, data aggregation and storage services, as well as locally hosted community outreach programs, bringing users in contact with digital content of local relevance” (DPLA, n.d.). Some Service Hubs include the Connecticut Digital Archive, the Digital Commonwealth of Massachusetts, the Montana Memory Project, and the Portal to Texas History.
Unlike Service Hubs, Content Hubs are usually large digital libraries, museums, archives, or repositories that are committed to maintaining the metadata records they contribute to DPLA. As a rule of thumb, Content Hubs generally provide and maintain at least 200,000 unique metadata records to DPLA. Some Content Hubs include ARTstor, the California Digital Library, the U.S. Government Printing Office, the Internet Archive, The Smithsonian Institution, and the University of Virginia.

2.2.4. Knowledge Organization and Structure

2.2.4.1. DPLA Metadata Application Profile

The DPLA Metadata Application Profile (MAP) builds from EDM and reuses definitions from RDF and the RDF Schema (RDFS), OAI-ORE, Dublin Core namespaces for elements and DC terms, and the Basic Geo (WGS84 lat/long) vocabulary (DPLA, 2013). Additionally, the MAP incorporates feedback from the DPLA community and digital hub pilot participants to create a balanced framework that allows DPLA to accommodate existing and emerging data models for library, archive, and museum resources (DPLA, 2013). DPLA’s metadata schema reflects multiple levels of representation. Like EDM, the DPLA-MAP is built with classes and properties (Figure 12) in order to structure data hierarchy and manage data values. For example, descriptive metadata (i.e., the kind used to search the DPLA web portal) is contained within the SourceResource class. Many of the properties in the SourceResource class (title, date, format, etc.) are based on Dublin Core Metadata Initiative vocabularies, which allows partners to crosswalk their elements to DPLA-MAP. The SourceResource class also links to other classes in DPLA MAP: WebResource, which stores information about digital versions; Place and TimePeriod, which allow for further field enrichments; Collection, which gathers information.
about locally defined sets or collections; and Aggregation, which packages all the information from the other classes together and also stores important information about direct collaborators (Hubs) and data providers (Contributing Institutions).

Figure 12. DPLA Domain Model v.3 with core classes highlighted (DPLA, 2013).

The structure of DPLA-MAP allows it to interface with most other metadata standards, and so far Dublin Core (simple and qualified), MODS, METS-wrapped MODS, MARC XML, and a few others have been crosswalked to the DPLA-MAP. Furthermore, any metadata standard which has been mapped to EDM can easily be mapped to DPLA-MAP. Other standards and models, such as VRA Core, CDWA, and CIDOC, have not yet been crosswalked but are compatible with the MAP (Rudersdorf, 2014).

The DPLA MAP documentation, published only in the initial stages of formal specification in RDF of the DPLA classes and properties, does not account for all of properties used by current or future DPLA content providers (DPLA, 2013). One of the proposed changes for the upcoming version of DPLA MAP is replacing the dpla:Place class with a new edm:Place
class; other classes, such as EDM TimeSpace, will be expressed in greater detail in a future version of the application profile as DPLA transitions their "linking" values from literals to URIs (DPLA, 2013). Furthermore, many of the DC-based DPLA properties will be replaced with properties from GeoNames and the Basic Geo vocabularies in order to increase interoperability as well as assist in the transition towards using URIs instead of literal values (DPLA, 2013). Despite the continued revision of the MAP, the documentation is available in order to gather additional community feedback and provide developers with a baseline of properties (DPLA, 2013).

2.2.4.2. Data Harvesting

DPLA is able to receive data from its Hubs in many forms. Currently, OAI-PMH is the most common type of data feed used by DPLA's Hubs (usually in the Dublin Core standard or MODS), but many institutions also use locally defined APIs to provide data in multiple formats. Rarely, DPLA will even download static data in the form of text files, XML files, or something else; though this is not ideal, the DPLA MAP is still able to harvest this kind of data. More frequently, data from smaller or under-funded institutions are aggregated by a Service Hub (for example, the North Carolina Digital Heritage Center Hub, which is responsible for aggregating the data from 150 institutions) and then shared with DPLA in a single feed (Rudersdorf, 2014).

2.2.4.3. Linked Open Data and API

DPLA is committed to facilitating interoperability and operating as part of the global linked data environment. As such, the DPLA uses existing open source code wherever possible, supports open standards, and makes all DPLA-funded code available in free, open source, and
reusable form; furthermore, all tools created or funded by DPLA will be made freely available for reuse and extension (DPLA, n.d.).

DPLA data is available as bulk downloadable files using JavaScript Object Notation for Linked Data (JSON-LD) serialization (http://dp.la/info/developers) and the DPLA MAP code and structure is available on Github (DPLA, n.d.). DPLA data is also maintained in an open API in order to allow the development of applications, tools, and resources (DPLA, 2013). The DPLA API is created with the same principles of openness that underpins its broader mission, and consequently the API is a model of accessibility featuring a simple HTTP request/response model allowing anyone to begin querying the API, free access to API keys, open access to "meta-metadata" (e.g. information from the Aggregation class), and no rate limiting (DPLA, n.d.). The open API serves the purpose of engaging community input as well as ultimately extending the functionality of the DPLA public portal by allowing developers free access to build new tools or services.

2.3. **BIBFRAME**

The invention and proliferation of the web has greatly influenced the manner in which we share and produce information. This shift has been particularly important in regard to traditional information institutions such as Libraries, Archives and Museums (LAMs). Much information and resources are now available at the tips of our fingers, with very little searching necessary. The web is linked to similar and related information via the ‘back end’ of the internet. This version of the internet is known as Semantic Web, a term that was popularized by Tim Berners Lee in 2001. The ideal of the semantic web is an interlinking of information in which one topic connects to multiple access points across a broad spectrum of information. With this new
interpretation of information sharing comes the realization that many of the great traditional knowledge structures are holding their information in unattainable silos, quite separate from the constantly exchanging web model.

BIBFRAME, or the Bibliographic Framework Initiative is an initiative undertaken by the Library of Congress and released in 2012. It attempts to bring the library catalog of tomorrow into the information age of today. If properly instituted BIBFRAME will greatly increase the access to information that is normally captured within each libraries individual catalogs. Drawing on many influences (everything from the current MARC records to FRBR to RDA) BIBFRAME offers a structure for connecting and creating an interlinking web of information resources. Some of the goals of BIBFRAME include a concentration on the difference between conceptual concepts and physical manifestations, a way to ensure the effectiveness of authority files and the use of the semantic web to link library catalogs to the outside information community. The original library card catalog and the following early cataloging modules all focus on the library item as a physical object. This poses an issue in this day and age as documents and items take increasingly diverse shapes, the majority of them not physical. BIBFRAME also endeavors to improve the functionality of authority files. Authority files allow for a disambiguous identification of entities and would be possible in the BIBFRAME model through a series of linking entities to their authority files. This linking would allow no room for uncertainty as to which entity is being referred to. The final goal is to link library catalogs more effectively to the outside web search community. It is important that the information available in the library catalog is also available through the normal searching of the web. The following sections will set out the history and basic foundations of the BIBFRAME initiative. Each section
will support and explain the main goals and objectives of BIBFRAME.

The transition away from MARC has been underway for at least 10 years. In 2002 Roy Tennant (who currently works for OCLC) published a paper entitled “MARC Must Die” in which he questioned the relevance of MARC and suggested (as the title implies) that it was time for a new standard. While Tennant admits that the whole system is warped he points out that the MARC syntax, MARC data elements and AACR (Anglo American Cataloging Rules) are so closely intertwined that it’s almost impossible to locate and remove the ‘culprits’ in the current model. Tennant also points out that MARC is too inflexible to be useful in the current information environment, mainly because it essentially denigrates all books to being solely physical objects. Tennant suggests the XML (Extensible Markup Language) as a solution to this inflexibility. (Kroeger, 874).

The BIBFRAME Transition Initiative Update Forum had its first meeting on January 22, 2012. This meeting was run by Beacher Wiggings, the director for Acquisitions and Bibliographic Access at the Library of Congress (LOC). 110 people attended this meeting; including representatives from the Deutsche Nationalbibliothek (DNB, German National Library), the Biblioteca Nacional de Espana (National Library of Spain) and Library and Archives of Canada. Among the speakers at the meeting included Deanna Marcum, the then associate librarian of Congress for Library Services (she now works in the nonprofit hemisphere). In her speech she cited 2 events that convinced her it was time for a new bibliographic framework. The first was the January 2008 meeting of the Working Group on the Future of Bibliographic Control. During this meeting 108 recommendation were made the Library of Congress and the library community in general. These recommendations include that
the LOC suspend any further work on RDA and instead start work on a more flexible and extensible metadata carrier. The second event was in regards to the RDA Test Coordinating Committee. They similarly reported that it is necessary to work on a replacement of MARC before RDA can be properly implemented. (Bothmann 350). As Marcum has pointed out, continuing to attempt to force MARC to handle a web environment is an expensive and most likely futile goal. The plan to supersede MARC with a new framework more appropriate for the WWW, with machine readable language. On November 21, 2012 LOC officially announced BIBFRAME, a new bibliographic framework meant to replace MARC with a more modern metadata encoding format. BIBFRAME will be based on linked data entity relationships that are taken from the current RDF (Resource Description Framework). (Kroeger 873). The project is a collaboration between LOC and Zepheira. Zepheira is an information architecture company lead by Eric Miller that provides technical and software support to the LOC. (Adamich, 9). BIBFRAME is still a work in progress and as such the framework itself is still open to interpretation. The more libraries that adopt the model, the stronger and more easily adoptable it will become.

2.3.1 Metadata Standards

2.3.1.1 MARC (Machine Readable Cataloging Standards)

MARC (Machine Readable Cataloging Standard) was originally implemented in 1968 and has been in use ever since. The original MARC was based on the primary content of printed catalog cards. It was developed to provide a manner for data exchange and control using machines. MARC allows bibliographic information to be encoded, exchanged and interpreted
using computers. This machine processable format allows for more effective searching and storing methods. MARC records typically have two levels; bibliographic records and holdings records. MARC records contain data about the intellectual essence of the work, data about the actual instance of the work and the record metadata such as control numbers, record handling codes and other annotations. (Library 7). Variations on MARC including USMARC (1980s) and MARC21 (late 1990s) came to be over the years, each providing rules and standards to match the current information community. (Adamich 8). An example includes the addition of the 856 field in 2008, this field allows for the inclusion of a URL, usually linking to an external table of contents or other related material. Lately it has come to light that MARC is no longer suitable to the current age of the WWW. Most notably MARC considers the format of the majority of the objects to be text, or at least physical objects. With the increasing existence and prevalence of digital items it is necessary to have a library catalog that can represent both these digital items and the more traditional physical items. Not only can MARC not sufficiently describe the various formats of items it also lacks the ability to connect the bibliographic information in the catalog to related resources. Linked Data (LD) has been suggested as a solution to this problem.

LD connects similar entities together via a system of interlinking bibliographic information and descriptive annotations. Tim Berners-Lee referred to the final result of this interlinking as the “Giant Global Graph”. Linked Data has 4 main characteristics or rules. The first is the URIs, (Uniform Resource Identifiers) are used to identify all resources. The second is the HTTP (Hyper Text Transfer Protocol) URIs are used so that people can access and look up these resources. These are in the form of URL strings and unambiguously identify bibliographic data while connecting it with other data. This process is known as dereferencing, or following a
resource back to an area containing more information about it. This is the third characteristic of LD, that useful information be available once an item is dereferenced. The final rule of linked data is that URIs provide connections to other similar URIs, this provides a means for resource discovery. (Library, 24). For example, using the book Harry Potter, a link could be created between a book and its author in the form of: “Harry of Potter” “hasAuthor” “J.K. Rowling” and in turn “J.K. Rowling” “hasProfessionOf” “Author”. This system creates a web of connected data. In this scenario the URI for “J.K. Rowling” would then direct you to all instances of “J.K. Rowling” within the web of data. This system has the potential to greatly increase access within the library and information community.

Unfortunately MARCs format is not flexible enough to support a LOD model. MARC makes it difficult to represent information using URIs, hence the creation of BIBFRAME. But referring back to Tennant’s article, MARC is not dead yet. It has been in existence for many years and has evolved quite well until recently, clearly it contains much vital information it its format and structure. It is the intention of the creators of BIBFRAME (mainly the LOC) to continue using MARC for as long as it still makes sense. BIBFRAME is not necessarily meant as an all-out replacement of MARC than as a sort of drastic face lift to move the system into the current information community.

2.3.1.2 RDA (Resource Description and Access)

RDA (Resource Description and Access) is a cataloging model was meant to succeed AACR2. It is an application of the FRBR and the FRAD models which will be covered in the next section. (Library 35). It is important to note that at this time the implementation of RDA is both separate from BIBFRAME and is on somewhat of a hiatus. As mentioned earlier the LOC
decided that it is necessary to complete the implementation of the LOD BIBFRAME model before RDA can be successfully using in library communities. RDA was released fairly recently (2010) and was intended to link a work to its related forms. RDA was a precursor to BIBFRAME in a way in that it attempted to confront the limitations of MARC. RDA is based on the FRBR model which will be discussed in detail in the next section. (Adamich 8).

2.3.1.3. FRBR (Functional Requirements for Bibliographic Records)

FRBR (Functional Requirements of Bibliographic Records) is a conceptual model that describes the various relationships between bibliographic entities or data points. FRBR was first released in the late 1990s and like many before and after it looked to improve the MARC system by making it easier to express relationships between bibliographic information. (Lindquist 874). FRBR was created based on 4 main user tasks; find, identify, select and obtain. Find refers the the ability to find entities based on the search criteria supplied by the user. Next the user must be able to correctly identify the entity. The following task is to select the entity most appropriate to the users needs. Finally the user must acquire or obtain access to the entity. Entities in FRBR are divided into 3 main groups. Group 1 represents the products of intellectual and artistic endeavors. Group 2 contains those entities responsible for the intellectual and artistic content, these can be either people or corporate bodies. The 3rd group includes concepts, objects events, places and the entities from groups 1 and 2. (Library, 37) Group 1 in turn consists of 4 elements, abbreviated to WEMI. They are Work, Expression, Manifestation and Item and they represent various forms or states of a given work. Work represents the creator’s creative idea. At this stage everything is technically intangible although it is the intellectual property of the creator. Consequently an Expression is the tangible form of the creator’s idea. This is also the work in its
first and original, physical form. The Manifestation is the work in a variety of format, such as translations, books on tape, etc. This entity allows for a great deal of flexibility regarding the description of a work in non-traditional and non-print formats. Finally an Item is the actual number of copies or units of a work. Item level records include specifics such as page number, call numbers, ISBNs, etc. (Adamich 8). FRBR was also released with two standards for authority and subject authority data, known as FRAD (Functional Requirements for Authority Data) and FRSAD “Functional Requirements for Subject Authority Data”. FRBR has been of great influence within the information community because it allows for proper description of the more abstract works of this age. In turn the FRBR model has influenced many schemas and models.

2.3.1.4. RDF (Resource Description Framework)

RDF is a metadata language that represents bibliographic data in a linked data format. RDF creates descriptive “sentences” about data in the form of “triples”. Triples highly resembles the annotated relationships used in LOD that was covered earlier. These triples are constructed using the subject, predicate and object format of a simple sentence. Although the values of each are tweaked to work within the web community. Using the Harry Potter example again; the subject, represented by a URI, “Harry Potter” “hasAuthor” “J.K. Rowling”, also represented by a URI. There are various published RDF languages that can be used in the predicate including vocabularies for DC (Dublin Core), MODS (Metadata Object Description Schema), EAD (Encoded Archival Description), TEI (Text Encoding Initiative), VRA Core (Visual Resources Association), ISBD (International Standard Bibliographic Description), FRBR (Functional Requirements for Bibliographic Records), RDA (Resource Description and Access) and MARC21. There are also many already created URIs for subjects and objects through sources
such as DBPedia and VIAF (Virtual International Authority File). To assist in the interlinking of data it makes sense to use URIs that already exist as opposed to making your own.

The main strength of BIBFRAME is its ability to use RDF to pull from all of these preexisting models, schemas and vocabularies to describe bibliographic records in a machine translatable manner. RDF is, so far, the most machine readable metadata schema in existence, mostly due to the language standards it adopted from the World Wide Web Consortium (W3C). (Kroeger 876). Of course the machines we are talking about here are computers. RDF represents these sentence statements using XML (Extensible Markup Language) in the form of subject URI, a URL that represents the predicate accompanied by which vocabulary set the predicate is from and an object URI. This method is easily read by the internet and allows the information once only available in library catalogs to now be visible via the World Wide Web. The next sections covers the manners in which BIBFRAME will facilitate this process and how it can be applied in libraries and other information communities.

2.3.2 Goals and Model

In this day and age the vast majority of research head to resource aggregators on the web such as Google and Wikipedia as their first step in information gathering, very few people use the library catalog as a vital step in the research process. BIBFRAME is the most recent step in the organization of bibliographic information and allows for resources to be available through the wider web. It supports the conversion of text strings into data and makes the library data accessible to the larger data community. This method of bibliographic information storage can be machine read and allows for computers to make connections between resources. (NISO, 17).
MARC and MARC21 have always provided solid and high quality data that is formatted for a specialized library market. This data could be helpful to the web as a whole and can be made more accessible when transferred into a linked data format. (Library, 7). BIBFRAME can also assist in the interchange of information from institutions that don’t use library data information exchange formats such as library vendors. BIBFRAME could also facilitate administrative functions such as these. (NISO, 15).

One of the most important and unique parts of BIBFRAME is its deference to outside standards and rules. In other words BIBFRAME is agnostic to cataloging rules, it is flexible enough to contain a variety of content rules and data models. In the evolving bibliographic information community there are many specific models such as those for archives and the hard sciences. From the beginning BIBFRAME had to work to ensure that all these models could be expressed using BIBFRAME. For example, BIBFRAME supports RDA, DACS (Describing Archives: A Content Standard), and CCO (Cataloging Cultural Objects) among many others. (Bothmann 350). This agnostic cataloging model could partly be influenced by the inextricability of MARC and AACR2. This new model needs to be general yet independent enough not to become entangled in any one model or set of rules. (Kroeger 881).

The BIBFRAME Initiative has 3 main objectives to increase discovery and access between libraries and the entire information community. The first it to “differentiate between conceptual concepts and physical manifestations (e.g. works and instances)”. The 4 main entities of BIBFRAME will be covered more explicitly in the next section. This objective is clearly influenced by the FRBR model in that it recognizes the difference between the tangible and intangible. The next objective is to “unambiguously identify information entities
(authorities)”. This objective refers to BIBFRAME’s use of URIs. URIs allow for the unambiguous identification of bibliographic information because each URI is not only unique but also connected to a “record” that includes absolutely identifying information about the person, place or that is being described. Authorities allow for transparency in situations such as pen names and alternative spellings, linking all similar entities together. The third and last objective is to “leverage and expose relations between entities”. This is in reference to the use of LOD in the BIBFRAME model. The model uses RDF to express the LOD and hence, links it to related information. This process links library cataloging information not just to similar entries within the catalog but also to related resources within the entire WWW. This is the leveraging of relations. By expressing the relations in XML they became a visible part of the entire web of data; or as mentioned in the opening paragraph; an integral part of Tim Berners-Lees Semantic Web.

Another Feature of BIBFRAME is its “atomization” of bibliographic information. In a traditional bibliographic record information is organized within a record which represents a work. This structure is rigid and the information is not nebulous. BIBFRAME essentially deconstructs each record into its individual parts and then reconstructs it in multiple, single subject sentences. This process makes BIBFRAME very flexible and allows for more connections between resources (Mitchell 27). This deconstruction and reconstruction actually adds a more detailed level of description that was unattainable before. This is particularly helpful in the case of cultural heritage information. Using BIBFRAME allows CH works to become aligned in a landscape rich with related resources that provide context for the object. This can be anything from research paper on the work to other works from the same creator.
2.3.3. Entities

The BIBFRAME model contains 4 entities; Creative Works, Instances, Authorities, and Annotations. Entity 1, Creative Works reflects the conceptual version of the resource; similar to the FRBR work, at this stage it is still intangible. (BIBFRAME). A Creative Work has 27 core properties; including title, classification, language and content scope. It also has 13 predefined types including audio, cartography and data sets (Mitchell 30).

Entity 2 or Instances are actual physical or material embodiments of creative works. Instances can take the form of both digital and physical manifestations adding a layer of granularity to the description. Instances have 79 properties although many of them are the same as the properties for creative work. Some of the properties include title, alternate title, ISSN, publication date and UPC. Instances also have special properties to show relationships such as “hasAnnotation” and “instanceOf” (Mitchell 30).

Entity 3, Authorities is very similar to the traditional concept of authorities. Authorities can be people, places, things, topics, organization and events that are in some way related to the work. Authorities are reflected using external vocabularies or ontologies such as those mentioned in the section on RDF. As authorities these entries are unambiguous and point towards other related items (Mitchell 30).

Entity 4 or Annotations are a fairly new concept within the information community. Annotations provide applicable related knowledge about the work at hand. (BIBFRAME). There are only 3 properties within the annotation including “annotates”, “annotationAssertedBy” and “annotationBody”. The types include “coverArt”, “holdings”, and “reviews” among other things.
and provides a context for the work (Mitchell 30). Annotation can be either in-band or out of band. In-band annotation are created by the holdings library and provide a more personalized record. Out of band annotations are imported data from a 3rd party service. (Library, 11).

Figure 13. BIBFRAME Linked Data Model Source: [http://www.loc.gov/BIBFRAME/docs/model.html](http://www.loc.gov/BIBFRAME/docs/model.html)

2.3.4. Conversion and Tools

As with any new framework conversion and implementation is easier said than done. MARC takes several different forms as it consists of layers of several different standards, not to mention the older versions of MARC. This makes it difficult to create one translation tool for all of MARC. Some solutions include creating multiple translation tools, converting MARC to MARC21 before converting and creating variants of BIBFRAME (Kroeger 875). At the moment BIBFRAME has been experimenting with using RDF/XML language to convey the bibliographic data and relationships. BIBFRAME is still a work in progress and other language have been attempted including RDF/Turtle and N-triple. (NISO, 19).

Obviously conversion in general is an expensive process and the additional cost of noncompliant metadata can be too much for small libraries. In addition to the price there is still a general lack of BIBFRAME software. The software vendors don’t have an incentive to create affordable versions of the software since the framework has not yet been generally adapted by the community. Most libraries do not even have the ability to adapt BIBFRAME because their software systems won’t support it. (Kroeger 884). These problems should be solved as BIBFRAME becomes more popular but there are clearly many obstacles. On the upside there are several tools available to assist in the conversion of MARC metadata to BIBFRAME format. An example is Viewshare, which was created in conjunction with the BIBFRAME Initiative (LOC and Zepheira) and converts MARC records into LOD format. This tool, still in its first version translates MARCXML into RDA using the BIBFRAME data model. (Lindquist 919). The caveat to this translation tool is that the bibliographic data has to already be in MARCXML format, if the library at hand does not use this format already it will be necessary to translate all MARC
records into XML prior to conversion, a lengthy process. This tool also creates LOD visualization to allow for a view of the relationships between metadata (Lindquist 924).

In May 2014 Sally McCallum at the Library of Congress announced that a tool called the BIBRAME Editor (BFE) was now available on the BIBFRAME website. This tool is able to create BIBFRAME records using a front end input tool. It also, similar to a traditional cataloging module; is able to be paired with a back end storage component to save and reopen records later for viewing and editing. (McCallum, 1). Further experimentation with the current tools will allow for experimentation and innovation in the field.

![Figure 16. MARC21 Record and BIBFRAME Resource. Source: http://bibframe.org/resources/sample-gwu/exhibit.html](http://bibframe.org/resources/sample-gwu/exhibit.html)
2.3.5. Case Studies

2.3.5.1. DNB (Deutsche Nationalbibliothek)

Deutsche Nationalbibliothek (DNB) or the German National Library has been involved in the creation of MARC since the beginning. They are part of the early experimenters group, a group meant to shed light on the actual implementation of MARC. This group will execute BIBFRAME within their libraries and report back to the LOC with any important information that has been discovered (Bothmann 354). Lars Svennson from the DNB points out that the card catalog was not built or “pivot points” or authorities files. It was created in representation of a more sequential display. Svennson and his team are very invested in implemented a way to make this library data not just available but an integral part of the web. (NISO, 5).
Prior to April 2012 DNB has stored its authorities in four different authority files. Since then they have created one integrated authority files that includes all persons, topics, geographic areas, corporate bodies and work titles. This has insured the when the authorities are translated to linked data and URIs there will be no redundancy. The same URI is used for the entity regardless of the amount of occurrences of said entity. For example, the same author URI is used in the instances of different books they created. (NISO, 11). DNB began to publish its authority data as linked data in 2010. The DNB uses RDF/XML and RDF/Turtle to serialize its records. Currently they are not of the same complexity as the MARC records but work is being done to correct this. DNB only uses existing ontologies such as DBPedia and FOAF to ensure the easy conversion of data. (Library, 32).

2.3.5.2. British Library

The British Library is currently developing a Linked Data version of the British National Bibliography which contains 3.3 million titles. At the moment the project only includes published books although future plans are to include serials, multi-part work and kits among other more complex records. The British Library is modeling things of interest related to items, such as people, places and events. The following is an example of the British Library data model. (Library, 29).
2.3.5.3 **George Washington University**

George Washington University (GW), is also part of the BIBFRAME initiative as an early experimenter. GW has a fairly small library compared to some of the other early experimenters but the data is contains has been created and collected over a long period of time. It has also been migrated from a variety of platforms and formats. Currently GW is only using monographic records, nor is it including multiple version records or those with complex holdings locations. During the first phase of the project authority files were also not included. The GW team had catalogers and programmers working together on this project. Catalogers can provide insight on how to describe an item and the programmer understand how to describe the resulting data effectively and connect it to similar data. (NISO, 20).
3. **Challenges**

These three initiatives, each ground-breaking in their own respect, are forging a path towards increased openness and collaboration in the cultural heritage sphere. OpenGLAM (Galleries, Libraries, Archives and Museums), founded in 2011, is “an initiative coordinated by the Open Knowledge Foundation that is committed to building a global cultural commons for everyone to use, access and enjoy.” (Open Knowledge Foundation, n.d.). Encouraging cultural institutions around the world to open their content and metadata to the public, OpenGLAM praises Europeana and DPLA for their “exemplary open metadata licensing policies”. The success of Europeana, DPLA and BIBFRAME is, in part, due to their flexibility, responsiveness, and forward-thinking strategies. They do, however, face challenges. Adoption, community engagement, and support are essential to their sustainability.

3.1. **Access**

Europeana, the most mature of the three, has set the groundwork on which the more youthful and agile DPLA and BIBFRAME are able to take new steps. As the largest (and most cross-cultural) initiative, Europeana has faced some unique challenges, particularly with regard to access in terms of language, metadata and enrichment.

Europeana follows the European Library’s language policy, issuing the statement: "As far as possible, the aim is to provide the public with the information they are looking for in their own language. However, due to constraints on resources, this isn't always possible." The goals stated are that top-level pages (i.e. the navigation, search, retrieval, and display interfaces) in all full partner languages; prioritise speed of publication over availability of translation; and, allow
limited language mixing (i.e. more than one language on a page), where a warning sign needs to be displayed. As mentioned above, Europeana suggests that data providers provide metadata in multiple languages (europeana, n.d.). In practice, Europeana’s multilingual functionality is less than perfect resulting in inconsistent top-level page translations, etc. A user is able to limit search results to a specific language using the ‘By language of description’ facet. Using our previous ‘picasso pablo’ example, the most populated language options in the results are Espanol (77), Italiano (57) and Multilingue (26). This is not entirely representative of the norm, with the descriptive text of the majority of items being in German and French (Valtysson, 2012) in line with those countries also supplying the largest number of items. While Europeana does offer language support (for instance, the language-specific metadata element headings), once a user locates an item of interest, she must go to the item repository’s website where she might find that the complete metadata record in another language without any translating tools

With regard to metadata, the main issue is the extent of how heterogeneous Europeana’s data is. EDM, in providing a framework without specific content direction, enables Europeana to harvest this diverse data but, however, also allows incorrect, inconsistent or minimal metadata through the gates. Europeana sees enrichment of metadata as a way of ensuring “that highly curated content from providers gets represented correctly across different languages” (Isaac, A., Petras, V., Stiller, J., 2014).

Europeana states the semantic and multilingual enrichment of metadata in Europeana is a core concern as it “improves access to the material, defines relations among objects and enables cross-lingual retrieval of documents” (Isaac, A., Petras, V., Stiller, J., 2014). In a EuropeanaTech Task Force report on Multilingual and Semantic Enrichment Strategy, released in March, 2014,
six sample multilingual datasets were examined. It was concluded that metadata quality was the main causes of poor enrichments and that enrichment of poor metadata worsened the problem by distributing it across multiple languages. Suggestions were made on how Europeana might tackle the following key issues with regard to metadata quality:

- **Quality of the original metadata:** Examples include vague or incorrect metadata, inconsistent use of separators (such as semicolons) to divide subjects, etc. Additionally, several unresolvable links were found. Suggestions for Europeana include user feedback forms to allow flagging of incorrect metadata or enrichments, increased collaboration with providers, establishing rules for field formatting, an automatic link/checking process, and more encouragement for providers to use persistent URIs as metadata values. For providers, a tool like OpenRefine might help clean up data.
- **Mapping to EDM:** It was found that the majority of enrichment flaws originated in the mapping process. Therefore, the task force suggested Europeana provides: clearer answers on specific areas of metadata entry that seem to cause the most confusion; increased documentation on enrichment so that providers are aware of which fields are enriched; supporting tools for providers to enable them, for instance, to test their mappings; and, increased training through clinics.
- **Checking metadata at ingestion:** Generating a metadata-quality score so that low quality data can be identified at ingestion, not enriched and flagged for the provider.

In terms of vocabularies used and enrichment, suggestions include fitting the vocabulary with the context and language of the metadata and collection and that enrichment rules and documentation be established for all fields. The report concluded by stressing the importance that users could play in enriching metadata.

### 3.2 Outreach, Funding, Sustainability

Europeana defines sustainability as “the ability of a project to maintain its operations, services and benefits during its life-time and possibly beyond” (Europeana Photography, 2013). This can be interpreted in two ways: the need to maintain the digital infrastructure through the inevitable changes in technology and the need for the site to maintain its relevance with the
population it serves. In this section, we will address the latter issue. The challenge is to not only to continue Europeana’s expansion, but to attract and engage users in a way that will make them advocates for Europeana. Without support from the European population, the site risks losing funding from the European Union and additional funds from individual countries that its come to rely on. In the words of Communication Director Jon Purday, “If you don’t spread, you’re dead.”

Europeana is attempting to attract new users through advertising and a presence on social media, but how can it keep users coming back after that initial visit? Pulling from what we already know about Europeana’s user groups and their objectives, we can assume that the more scholarly visitors will return because Europeana provides them a service, but what about the users that want to be entertained or to feel like part of a community?

Europeana has addressed these concerns by creating a “Europeana Awareness” project, which will serve the site in four ways (Europeana, n.d.):

- Increase publicity to users, policy makers, politicians and cultural heritage institutions to raise awareness and encourage contribution.
- Encourage use of the site for hobbies, research, learning, genealogy and tourism. Engage users with user-generated content, creation of digital stories and social networking.
- Develop a partnership with underutilized institutions such as public libraries, local archival groups; broadcast organizations and open culture re-users.
- Encourage cultural institutions to continue to provide content by raising awareness of the opportunities provided by the new Europeana Licensing framework; developing mechanisms for collective rights management; and increasing the amount of content in Europeana that can be freely re-used.

Perhaps the most interesting feature of the project was the creation of an activity planner. This is a document available on the Europeana Professional website that charts events to promote Europeana, and accepts suggestions for events via email. The log is arranged in a spreadsheet and includes information such a description of the activity, completion status, country, category and the name of the person who submitted the suggestion. Examples of events include “WWI Collection Day and PR Event” in Italy, “Fall of the Iron Curtain Day” in Hungary and “50th
Anniversary of the Beatles First Album” in England. Real world activities like this will allow users to get out and actually see some of the objects in their original institutions. Perhaps more important, there will be a chance to meet fellow users in their own towns. Although it may be too soon to tell how effective this approach is, by reaching out beyond the computer screen, creating a sense of community, Europeana is positioning itself to be a leader in local cultural heritage.

Although Europeana is taking great steps to form a sense of community outside of the digital world, their presence on social media sites such as Facebook, Pinterest and Twitter also helps to reach people in their daily lives. Europeana has active accounts on all of these platforms, with constant updates including interesting finds from their collection or announcements about upcoming projects, posted by Europeana or one of their contributing institutions. Users are encouraged to participate by commenting, re-pinning or sharing posts. Every share has the potential to drive a new user to the site, and every new user is a potential advocate.

Europeana has seen the value in these social media uses, and takes steps to provide them with interesting, high quality posts. Europeana maintains an editorial calendar that charts upcoming posts for all of it’s social media platforms and invites all partners to contribute. Europeana also provides user and traffic statistics, so partners can have an idea of the kind of exposure their posts will receive on each platform. This helps partners to post where their targeted users are most likely to be, increasing the possibility of site traffic and providing the best content to the users.

In 2012 Europeana reported that only 1.7% of the site’s traffic was directed from social media. However, the report found that those users spent longer on the site, visited more pages and were likely to “share” while they were there (Bates, 2012). Although this is quite low
compare to visitors who arrive via search engine, Europeana values these users highly, taking a
quality over quantity approach.

The site and its additional projects take an estimated 30 million euros per year to run, and
starting in 2015 Europeana will be funded by the European Commission’s Connecting Europe
Facility (CEF). CEF is the funding source for many digital and broadband projects in Europe,
which means that Europeana must now compete for enough funding to keep it running properly.
Unfortunately, a proposal was recently introduced to slash CEF’s budget from 9 to 1 billion
euros through 2020, leaving Europeana to compete for a smaller part of an already strained
budget.

To raise public awareness about their potential funding cut, Europeana launched a social
media campaign aimed at rallying their users and getting the word out to politicians. Using the
hashtag #AllezCulture, a dialog was encouraged on Facebook and Twitter, and petition was
started that gathered 6,800 signatures. Europeana was very vocal about the service they provide
and what the public stands to lose they cannot operate at full capacity. Although funding was
ultimately slashed, Europeana is hoping to sustain itself by raising additional funds with the help
of individual countries and for profit projects.

Since it’s very beginning Europeana has received additional funding from the individual
governments of the European Union; twenty three countries have been able to contribute a total
of €3,600,200 from 2008-2013 (europa.eu, 2013). Europeana estimates it’s funding gap for the
upcoming year and requests a certain amount through meetings and conferences attended by
political and cultural leaders of various countries. Although there has historically been an
increase in the funds requested each year, Europeana requested only €446,000 for 2014, the
lowest request since 2009 (europa.eu, 2013). This was possible because of a decrease in costs for the Europeana home office.

Thanks to a contribution from twelve countries including The Netherlands, Germany, Estonia and Ireland, Europeana is on its way to meeting financial goals for 2014/2015, however a gap of over €200,000 still exists for each year. Securing the additional funds could be challenging, as funding can be inconsistent. Europeana could attempt to fill this gap by lobbying the seven countries that have yet to contribute any money to the site. Countries such as Bulgaria, Malta and Portugal have never provided funding from their annual budgets (europa.eu, 2013).

Learning from Europeana, DPLA has approached outreach with organized, on-the-ground strategies that tend to be more focused than those executed by Europeana. Where Europeana creates expansive campaigns targeting their entire audience, DPLA has focused on particular target communities. Because the DPLA intends to both serve the public while drawing on cultural heritage materials of the public, throughout its early development it has focused much of its energy on finding ways to engage and expand its user communities. At the 2013 DPLAFest, workshops and discussions took place on how to share its materials with K-12 educators and higher education faculty, concluding that the DPLA must prove its worth as both a pedagogical and digital literacy tool that can scaffold curriculum throughout the grade levels (Steffel, 2013). Another workshop challenged the fact that the DPLA’s Board “does not reflect the U.S.” (Blackmer, et al., 2013), and questioned how the DPLA can further represent historically underrepresented communities. They noted that archives of the underrepresented are usually underfunded and understaffed, and the DPLA and its service hubs should remain aware of this fact as they aggregate metadata from across the country. In order to both engage and
represent diverse user communities of the American public, the workshops concluded strong outreach programs are necessary for the future functioning of the DPLA.

In 2013, the DPLA introduced its Community Reps program, which recruited about 100 representatives from thirty-six states and two international countries (Abbott, 2014) to help disseminate information about DPLA. DPLA Community Reps are required, in addition to attending an introductory webinar, to plan at least one DPLA outreach event during a single calendar year. Since January 2014, about forty-five outreach events have already taken place, including such events as classroom activities in higher education and K-12 classrooms; presentations for library patrons and professional organizations; Wikipedia edit-a-thons that use DPLA as a reference source; and conference poster sessions (Abbott, 2014). Community Reps come from a variety of professional backgrounds including state and public libraries, colleges and universities, and other LAM organizations. Geographic and “professional diversity is very important for DPLA’s reach since Reps are tasked with engaging their communities and providing feedback [to DPLA staff]” (Abbott, 2014). As DPLA so heavily relies on participation from its users and contributing institutions, it has vigorously pursued their support, opinions, and recommendations through a unique and enthusiastic outreach strategy, as demonstrated in the Community Reps program.

3.3. Openness

According to John Palfrey, once-president of the DPLA Board of Directors, copyright is the “the specter hanging in the background of conversations about libraries and about the DPLA” (Laskow, 2013). Unlike the Google Books Project, which scanned and made searchable millions of books before dealing with legal ramifications, both DPLA and Europeana are establishing a
more cautious approach to copyrighted material. Because they are not-for-profit institutions, the legal issues pertaining to copyright are different than those of the Google Books Project. DPLA is working from a public-interested perspective (Palfrey, 2011) and Europeana states that “Europe’s digital library, museum and archive, belongs to the public and must represent the public interest” (Europeana, 2010).

In his article *Europeana: the digital construction of Europe’s collective memory*, written in 2010 and published in 2011, Bjarki Valtýsson discusses the paradox of Europeana’s goals of reaching a Web 2.0 audience while taking a seemingly traditional approach to copyright. He states that Europeana’s collecting policy is not transparent and that the Commission has adhered to copyright laws and thus focused on content that is in the public domain. This “top-down” approach has meant that Europeana is “locked in a cultural landscape based on permission” - the resulting content is predominantly in French and German and is more classical in theme. There is far more content on Mozart, for instance, than contemporary artists. Valtýsson goes on to say that Europeana’s focus has not been on user-generated content, which removed it from a “bottom-up approach”. Europeana has continued a “read only” culture in a time where users expect to be able to “read and write”. Since the publication of his article, Europeana has addressed several issues raised by Valtýsson. As mentioned above, in September 2012, Europeana released its metadata for a dataset of 20 million objects in “an important new international precedent, a decisive move away from the world of closed and controlled data” (europeana, n.d.). This is intended to promote re-use and user-generated linked open data projects. However, the Creative Commons licence does not apply to the thumbnail preview, which must contain a copyright statement.
DPLA launched with access to public domain works only (Laskow, 2013), and one of the biggest challenges DPLA faces is how to facilitate access to commercially available titles legally and without negating their economic value (Palfrey, 2011). Members of the board of directors have differing ideas for tackling copyright issues. Palfrey is interested in interpreting existing law in favor of lending out materials, and advocating and creating new legislation for copyright concerns; other viewpoints include hammering out licensing systems to allow copyrighted works to be digitally loaned, and allowing authors to earn fees based on the circulation of their books (Laskow, 2013). In this case, the DPLA may prove to be revenue enhancing, rather than revenue threatening, for content creators (Palfrey, 2011).

One of the specific challenges DPLA faces, should it opt for a collective licensing approach, is the balancing act of setting up digital rights management (DRM) and incidentally limiting or restricting rights already in existence. First, fair and private use can be affected incidentally by DRM measures. While collective licensing does not inherently violate individual interests in fair and private use, there is concern that the implementation of DRM measures combined with other aspects of U.S. law which restrict circumventions of DRM (the DMCA legislation, for example) may end up restricting uses of works that would otherwise be considered fair use (Palfrey, 2011). The second concern is privacy: to take advantage of the offerings of a digital library, end users may be forced to give out more personal information than if they simply browsed the stacks or checked out books at a physical public library (Palfrey, 2011). DRM measures often require authentication and tracking of users in ways that lead to a loss of personal informational privacy (Palfrey, 2011). Thus, the implementation of measures to
reach a collective licensing agreement may have unintended consequences relating to the erosion of other kinds of rights.

Still, it is unlikely that the issues will be worked out via licensing agreements alone. The issue of orphan works, in particular, may only be resolved with legislation -- and the orphan works approach may even augur in an era of copyright legislative reforms (Palfrey, 2011) if DPLA is also able to defend its interpretation of the doctrine of fair use in court (Darnton, 2012). Ultimately, the end result DPLA may look similar to the Google Books Project, though they will have arrived by different measures. After all, in order to serve as a digital library, a level of content lending or sharing is expected -- even for a library which doesn't hold copies of the works it directs users to find (Laskow, 2011). If the goal of DPLA is to democratically enhance access to digital materials, legal avenues may need to be carved anew. Strengthening fair use, limiting DRM, and developing better pricing models is becoming a large part of the advocacy work of librarians who are determined to preserve and provide access to digital materials over the long term (Manoff, 2013).

4. **Opportunities**

One area in which both DPLA and Europeana can progress lies within the user interface of each site. Currently, DPLA offers several ways to browse materials: by curated digital exhibitions, location, and time period. Similarly, Europeana offers searches narrowed by place, time, creator and language. The homepage also offers users the choice to browse a featured item, partner institution, or current blog post. But as both initiatives move towards a linked open data model, they can begin to take advantage of more powerful browsing capabilities.
Both Europeana and DPLA offer the ability to download complete datasets, in order to support development of new tools, portals, or services. Additionally, DPLA actively encourages the engagement of third-party developers: in November 2012, DPLA hosted an "appfest hackathon" to support developers interested in using DPLA content to create web and mobile apps. Participants were encouraged to build, or even simply sketch an idea for, apps to recommend cultural heritage content: maps or photographs based on certain criteria, for example; or a way to help students find content related to paper topics; or even just a way to visualize metadata in interesting ways (DPLA, n.d.). Ideas generated during the Appfest included the use of DPLA content to create visualizations—timelines, galleries, maps, etc.—for users of the public portal. Jeff Goldenson and Jessica Yurkofsy created an app called Catalog the Whole Earth which allows anyone to catalog their own environment with a simple process. The idea is that the user can email a picture of any object or place to the webapp, and then log metadata and submit online; further developments, such as a more stream-lined metadata entering process or ways of displaying collections, are still being explored (DPLA, n.d.). Another potentially useful app was developed by Ed Summers. Called simply DPLA Map, the app uses the DPLA API and Google Maps to take advantage of the geo-location feature embedded in modern web browsers in order to show the user things that relate to the user's location (Figure 19).
Summers was interested in it as "mainly a good sniff test for the geo data that the DPLA API makes available" (DPLA, n.d.) and the results offer an alternative means of browsing, distinct from the way DPLA and Europeana are currently deploying geo data. Future efforts by Europeana and DPLA to encourage developers to continue making use of their respective APIs may yield more ways to enhance their public portals.

Europeana and DPLA may also be well-served to examine the practices of other cultural heritage aggregation sites. CultureSampo (http://www.kulttuurisampo.fi/about.shtml), or Finnish Culture on the Semantic Web, is a national communal publishing conduit for both institutional
memory organizations as well as private citizens. It offers a powerful semantic web portal, harnessing cross-domain semantic models, semantic searching and browsing methods, and semantic visualizations for users (Hyvönen et al, 2009). The methods used by CultureSampo may serve as a useful model off of which both DPLA and Europeana can build. In 2009, Hyvönen et al. describe CultureSampo’s nine different thematic browsing perspectives, some of which might easily be reused by DPLA and Europeana to augment the current browsing interfaces.

The maps search and browse view allows for the display of collection objects geolocated using Google Maps and shows the semantic relations of objects to places using twelve different spatial relations, including place of acquirement, place of subject, and place of manufacture. The map perspective also uses historical map rasterization to display historical areas as well as current geographical boundaries. The relational search perspective uses association identification to find relation chains between objects (Hyvönen et al, 2009). The user can type in two names (of people or organizations) and CultureSampo displays the social network (based on social roles such as parent-of, teacher-of, etc.) to show how individuals or organizations are related to one another (Figure 20).
The skills and cultural narratives perspective (Figure 21) documents traditional skills on video; the processes are then described semantically, generating a table of contents which allows the user to jump to any part of the video by selecting items from the list (Hyvönen et al, 2009).
Both DPLA and Europeana already allow users to explore their collections by place, time period, and through curated exhibitions. Building on CultureSampo's Semantic Web platform might add new dimension and depth to the end-user experience of browsing cultural content in the broader contexts of DPLA and Europeana.

Realistically, the CultureSampo portal perspectives are just the tip of what could lie ahead for the user interfaces of both Europeana and DPLA. Perhaps the most significant factor leading to new innovations is the open invitation for developers to write new applications using DPLA and Europeana resources. No single organization or individual can think of all the application possibilities for the data maintained by the two organization, but by allowing unfettered access, DPLA and Europeana are positioned to harness the creative energies of a large community of developers. Rather than becoming a repository for data, DPLA and Europeana’s utility can only increase as more people use and engage with its materials.

5. **Conclusion**

In spite of the challenges and shortcomings Europeana, DPLA, and BIBFRAME respectively face, their enthusiastic ideals and innovative practices seem, at this point, to outweigh most concerns for their future prospects. With strong mission statements highlighting the importance of public access to cultural heritage objects and environments, Europeana and DPLA have posed themselves as beacons of engagement in the cultural heritage community, shedding light on the location of millions of digital objects in the Western world and how to access them. While Europeana admittedly has had a slow start reaching out to its user base, both Europeana and DPLA are now pro-actively involved with their broad and distinct user
communities, vigorously publicizing their own worth and potential use in both public and educational contexts.

However, their greatest achievements so far, perhaps, have been somewhat invisible to the greater public: their data models. Today, the EDM remains an exemplary model for cultural heritage institutions encountering and creating linked data for the first time; its success is evident in the fact that the DPLA chose to use and build off of the same model when creating the DPLA MAP. Together, their data stores not only provide strong navigational infrastructures for front-end searching and browsing interfaces, but they also offer a new way to query and inference data through the availability of APIs.

BIBFRAME does not as exclusively and enthusiastically devote itself to disseminating cultural heritage objects throughout the world. However, because it provides a new, rich vocabulary for LOD resource description, its data will absolutely have the capacity to encounter and link to other vocabularies, thereby building new relationships (and thus new knowledge), particularly in the bibliographic domain. BIBFRAME is in a unique and exciting position as it is poised to replace MARC and become the new standard for bibliographic description. The fact that the Library of Congress has offered a LOD framework for a MARC replacement signifies a great intellectual shift not only in regards to resource description and records, but also to the greater, World Wide Web: that the effort to convert the “Web of documents” into the “Web of data” has positively begun, as also evidenced by major organizations like Europeana and DPLA.
6. References

Europeana


http://go.galegroup.com/ps/i.do?id=GALE%7CA201442292&v=2.1&u=nysl_me_pml&it=r&p=AONE&sw=w&asid=35a2e4b642c13039920465267d712e72


http://www.europeana.eu/portal/europeana-providers.html


http://labs.europeana.eu/api/linked-open-data/data-structure/


http://www.pro.europeana.eu/web/guest/content


Europeana. (2014, April 7). EuropeanaTech Task Force on a Multilingual and Semantic
Enrichment Strategy: final report. Retrieved April 15, 2014, from Europeana Professional:

http://pro.europeana.eu/documents/468623/8b75b054-712e-432b-a0f7-761898e6f60e


http://pro.europeana.eu/press-release?p_p_id=itemsindexportlet_WAR_europeanaportlet_INSTANCE_AtV7&lifecycle=1&p_p_state=normal&p_p_mode=view&p_p_col_id=column-2&p_p_col_pos=1&p_p_col_count=2&itemsindexportlet_WAR_europeanaportlet_INSTANCE_AtV7_itemId=1284453&itemsindexportlet_WAR_europeanaportlet_INSTANCE_AtV7_javax.portlet.action=setItemId


http://pro.europeana.eu/web/guest/about


Information, Communication & Society, 15(2), 151-170. doi:  
http://dx.doi.org/10.1080/1369118X.2011.586433


DPLA

Abbott, F. (2014, April 28). Community reps: A handy guide to who they are and what they are doing. Retrieved from  

http://www.americanlibrariesmagazine.org/article/digital-library-everyone

Blackmer, K., Castillo-Speed, L., & Darnton, R. (n.d.). [DPLAfest notes][afternoon] Focus on community: Enhancing community contributions to, and use of, the DPLA. Retrieved from  
https://docs.google.com/document/d/1k2tbriyeho_sYiVm5OeMW66tfGEppST7VSHE2oZLPP8/edit


DPLA. (2013 February 8). Metadata application profile, version 3. Retrieved from  
http://dp.la/about/map


http://dp.la/info/developers/codex/policies/philosophy/


The Johns Hopkins University Press. Retrieved from

http://muse.jhu.edu/journals/portal_libraries_and_the_academy/v013/13.3.manoff.html

for Privacy, Private Use and Fair Use. *Columbia Journal of Law & the Arts*, 34


https://docs.google.com/document/d/1ak7v1vzaA-yADlwQeewm8jbJd_isGfTeZDDI_kgH9o/edit

**BIBFRAME**


http://go.galegroup.com.ezproxy.pratt.edu:2048/ps/i.do?id=GALE%7CA342676507&v=2.1&u=nysl_me_pml&it=r&p=AONE&sw=w&asid=d4e127158a6d6ce38a06cece7464eef3.


General

