FEATURE

It’s About Time

Use of The Extended Date/Time Format in the Digital Public Library of America

Annamarie C. Klose, Scott Goldstein, and Morris S. Levy

Date metadata can be a powerful tool for browsing, search, and visualization. Although metadata quality and consistency are key, a broad array of practices exists in the Libraries, Archives, and Museums (LAMs) community’s metadata, which is stored in repositories, library catalogs, and other platforms. Cultural heritage materials with “squishy” dates can be difficult to express and display. While institutions may find local approaches, challenges arise when the metadata is shared elsewhere. The Extended Date/Time Format (EDTF) is a solution to encode dates that conform to a broad variety of scenarios, including ambiguous and circa dates. This article assesses the impact of EDTF in the Digital Public Library of America (DPLA), an aggregator of LAMs metadata. Compared to a prior study in 2015, date values that correspond with EDTF have increased. The overall number of records with any date value, however, has declined. The authors note nuances regarding the results and provide recommendations for the community.

Introduction

Within the Library, Archives, and Museums (LAMs) community, date metadata is a challenge. Digital collections can exist in a variety of platforms, including library catalogs, repositories, digital asset management systems, and digital exhibits—all with their own unique architecture and limitations. A work may have one or more dates associated with its life cycle. It can be difficult to express various uncertain dates and date ranges. Date searching is also important to users but frustrating due to various issues.[1] Meanwhile, one survey demonstrates date as the third most evaluated field by metadata creators.[2]

Machine-readable values support discovery in terms of search limits, results grouping, and sort options.[3] Consequently, encoding standards bridge the gap between humans and computers. As per Shaw and Maloney, “an encoding scheme is a specialized writing system or syntax for particular types of values.”[4] For instance, the date September 17, 1873, can be expressed as 1873-09-17 as defined by the encoding standard International Organization for Standardization (ISO) 8601. Unfortunately, encoded values often do not appear as human-readable for display and require query parsing for search indexing.[5] Additionally, encoding standards for more challenging date scenarios in relation to cultural heritage materials have only been provided fairly recently.

Annamarie C. Klose (klose.16@osu.edu) is the Metadata Initiatives Librarian at The Ohio State University, https://orcid.org/0000-0002-8926-8892. **Scott Goldstein** (scott.goldstein@mcgill.ca) is Coordinator, Web Services & Library Technology at McGill University. **Morris S. Levy** (mslevy69@gmail.com), https://orcid.org/0000-0002-7801-132X.

https://doi.org/10.5860/lrts.69n4.8549

The creation of the Extended Date/Time Format standard (EDTF), now a profile of ISO 8601, provides solutions for encoding many complex date scenarios. However, it is unclear how often this standard is used. This article discusses the current state of date metadata encoding with EDTF in the Digital Public Library of America (DPLA). With millions of records harvested from thousands of institutions across the United States, DPLA serves as a microcosm of metadata entry trends in the community. The authors use a 2015 study from Zavalina, Phillips, Alemneh, Tarver, and Kizhakkethill as a point of comparison and find that there has been progress with EDTF. However, there is still a need to standardize date metadata entry in the community with machine-readable dates in order to support search and retrieval.

Literature Review

Cultural heritage metadata—how it is created, stored, and displayed—is diverse. Across the literature and related guidance, however, some themes emerge, particularly for date metadata. Standards and guidelines have evolved. Although this section is arranged thematically, the concepts are interconnected.

Best Practices for Metadata and Sharing It

There are best practices to improve outcomes for metadata locally and for sharing via aggregators. Shreeves, Riley, and Milewicz define the six C's of shareable metadata: content, consistency, coherence, context, communication, and conformance to standards.[6] Bruce and Hillman's quality measures for metadata are completeness, accuracy, provenance, conformance to expectations, logical consistency and coherence, timeliness, and accessibility.[7] When sharing metadata, interoperability is key. Shreeves, Riley, and Milewicz point out that shareable metadata should support search interoperability.[8] Caplan states that search interoperability is "the ability to perform a search over diverse sets of metadata records and obtain meaningful results."[9] Foulonneau and Riley defined interoperability as "the ability of different systems to talk to each other."[10] They designated interoperability factors as "facets": technical interoperability, content-related interoperability, and organization interoperability.[11] Recently, the FAIR Principles for data and metadata provided guidance in terms of Findability, Accessibility, Interoperability, and Reuse.[12] Within FAIR, the second findability (F2) and first interoperability (I1) principles, respectively, speak to the importance of rich metadata and interoperability through language and syntax, i.e., machine readability. Although FAIR was intended for scientific data and metadata, it has gained international attention for cultural heritage.[13] FAIR has been evaluated in terms of describing digital cultural heritage and MAchine-Readable Cataloging (MARC) bibliographic records.[14]

The literature highlights various challenges. A lack of consistency can inhibit crosswalking metadata easily.[15] "Localisms"—data used to facilitate searching or display in a local system—lose context when the records are shared.[16] Also, one-to-one relationships between metadata standards are easier than converting between records or standards with one-to-many and many-to-many relationships.[17] Normalization is frequently mentioned as a solution to remediate metadata values. Programmatic approaches can be addressed at an aggregator level; however, metadata creators themselves best know their materials.[18] In addition, the concept of technical debt has been extended to library metadata quality.[19] Intentional metadata code debt includes cataloging issues meant to be fixed later in response to current constraints. Unintentional metadata code debt includes errors unwittingly made during cataloging, such as typos. Other debt issues with metadata are applied to design and architecture, the environment, documentation, and requirements.

Date Encoding Standards

Machine readability matters. Globally, there are different colloquial representations of dates. This can lead to confusion for both humans and computers. For example, the date May 19, 2013, could also be expressed as 19th of May 2013, 19 de mayo de 2013, 2013-05-19, 201300519, 19.05.2013, 05/19/2013, and 19/05/2013. Despite not being comprehensive, this nonetheless demonstrates why date encoding can be useful if applied consistently. For encoding date and time values, ISO 8601 is widely referenced and used in a variety of industries. It established YYYYMMDD in its basic format, or YYYY-MM-DD in its extended format, as the de facto date format. It has rules for various scenarios for expressing date and time.[20] The World Wide Web Consortium's Date and Time Formats (W3CDTF) may be the best-known profile of ISO 8601's extended format. Formalized in 1998, it provides a simple way to encode and display date and time with six levels of precision.[21]

Gaps remained for cultural heritage institutions in terms of common date scenarios, including Before Common Era (BCE) dates, questionable dates, approximate dates, and open-ended date ranges.[22] In 2012, the Library of Congress (LOC) drafted the Extended Date/Time Format with the intention to build on ISO 8601-2004 while addressing more date scenarios with "various semantic qualifiers and concepts," including ambiguous and approximate date ranges in applications.[23] These had been supported "using ad hoc conventions; EDTF provides a standard syntax for their representation."[24] The standard was subsequently revised in 2019 and adopted as an extension of ISO 8601-2019.[25]

EDTF has three levels: 0, 1, and 2. Level 0 is the most basic. In terms of encoding dates and time, part of it matches W3CDTF. Consequently, it is the most recognizable part of the standard. For time intervals, i.e., date ranges, it employs a "/" as a separator. This may seem foreign to metadata creators who are used to using a hyphen between date ranges. Other separators have clear usage as well. For instance, a hyphen is used between calendar components, e.g., YYYY-MM, and a colon is used for clock components. In Level 1, the concepts that can be encoded get more complicated. Some examples include seasons, unspecified digits, open intervals, and negative calendar years (i.e., Before Common Era dates). Level 2 introduces more encoding options for increasingly complex date scenarios, including sub-year groupings and set representation. The levels build upon each other like rules. For instance, intervals can have negative, circa, and questionable dates in Level 1. EDTF gets complicated as one goes through the levels and its logic. In addition, LOC provides the EDTF datatypes scheme.[26]

Adoption of EDTF has grown. University of North Texas (UNT) Libraries, the Portal to Texas History, and the Gateway to Oklahoma History were early adopters when EDTF was a draft standard.[27] UNT Libraries also found the EDTF seasons' usefulness for serials' dates.[28] In 2015, Zavalina, Phillips, Alemneh, Tarver, and Kizhakkethill analyzed the use of EDTF in the Digital Public Library of America

(DPLA). They presented a picture that EDTF was slowly being adopted by LAMs. This paper was used significantly as a point of comparison in this study. In a survey of twenty-four Metadata Application Profiles (MAPs), Green reported that EDTF and W3CDTF were both indicated in five, while ISO-8601 was indicated in nine.[29] EDTF has been used in MARC records at institutions including Yale.[30] But recent literature has noted challenges. McGrath opined about the complexities of using EDTF in the MARC 046 field.[31] Klose, Goldstein, and Levy reflected on problems with using EDTF for BCE dates in Metadata Object Description Schema (MODS) records when minting digital object identifiers (DOIs) through DataCite.[32] Meanwhile, Manifold et al. noted the successful application of EDTF dates with Dublin Core in Samvera Hyrax repositories.[33] In this case, scripting was used to convert encoded EDTF dates to human-readable date values in the display. Meanwhile, the EDTF values were utilized to support search by date with a timeline slider.

Date-Specific Metadata Issues

Date metadata is well researched. Issues include the values, the metadata fields used, and how metadata is transformed, e.g., for harvesting. The authors provide this summary of metadata date issues highlighted in the literature:

- Location of values: Dates may be recorded in multiple locations in a single record.[34] For instance, the date of creation may be found in the free text in a description field.[35]
- Data entry errors: Misspellings and various errors occur.[36] For example, "(ca.)" without any date information.[37]
- Syntax differences: Using both numeric and text values.[38] Different encoding standards, e.g., YYYY-MM-DD and MM-DD-YYYY.[39]
- Linguistic differences: Dates in different languages in records.[40]
- Ambiguity: There are unclear strings and characters, such as "c" could represent circa or copyright.[41]
- Vague values: Values such as "unknown" and "no date" as colloquially referred to as "noise"[42] or "garbage values."[43] Those are not useful for search and retrieval.
- Lack of values: According to one survey, date was only required at 45 percent of the participants' institutions.[44]

Date value encoding variations are perceived as inhibiting searching.[45] The use of a single encoding standard, e.g., W3CDTF, is frequently recommended.[46] Indeed, user-focused studies point to issues related to searching and browsing for records by it.[47] Normalization, a process of transforming data to make it consistent, is also recommended.

Dublin Core-Specific Date Metadata Issues

Although various metadata standards are used by LAMs, Dublin Core metadata is referenced and used frequently. Unqualified Dublin Core, also referred to as Simple Dublin Core, provides one date field, while Qualified Dublin Core has more granular options, e.g., creation and copyright date. In addition, Unqualified Dublin Core is a minimum requirement for harvesting records through the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH).[48] A summary of issues follows:

- Semantic interpretation: Institutions apply fields differently. For instance, the date and coverage elements are often used interchangeably.[49]
- One-to-one principle: Dublin Core has a one-to-one principle that states an original work and its digital representation should be described separately.[50] In practice, this principle is often dismissed. Consequently, a record's date values may include when the work was created, when it was digitized, or both. Yet the date of creation may be considered the most useful for searching regardless.[51]
- Harvesting issues: When converting from a robust metadata standard, such as MARC, to a simpler standard, such as Dublin Core, there is often data loss.[52]
- Pseudoqualifiers: Sometimes metadata creators add pseudo-qualifiers, e.g., "Digitized:" at the start of the value.[53]
- Custom date fields: Local date fields often get muddled when shared outside their local environments.
- Loss of context: In some cases, different date fields, e.g., created date and digitized date, are mapped together to the unqualified date field. Lumped together, they lack context about their values.
- Mismapping: Non-date values, e.g., publishing information, sometimes get mismapped to date fields. In other cases, date metadata sometimes does not get mapped at all.

Consequently, users searching for content by a specific date range may not find the content at all.[54]

Dates in MARC

MARC metadata can be transformed into other standards for digital objects and harvesting. The predominant areas for recording encoded date metadata in the *MARC 21 Format for Bibliographic Data* are the 008 field bytes 06-14 (Type of date/Publication status; Date 1; Date 2) and the 046 field (Special Coded Dates).[55] The code in 008/06 provides context for the date(s) in 008/07-14. Field 046 can be used to record BCE dates, incorrect dates, creation dates, modification dates for electronic resources, as well as starting and ending dates for aggregated content or the validity period of a resource.[56] It can also be used to specify the type of date when the information in the 008/06-14 is ambiguous. By the nature of only having four bytes available for Date 1 and Date 2, respectively, it is only possible to record years (known and unknown) in the 008 field. One can also use a more detailed date with code e. Across those eight bytes, one can then express an individual date as a month and year (YYYY and MM) or month, day, and year (YYYY and MMDD). The default format for recording dates in the 046 field is ISO 8601, which requires 8 numeric characters in the pattern YYYYMMDD. It is possible to use the EDTF field in the 046 field if it is appropriately coded as such in the subfield $2. However, EDTF was a later development.

Research Methods

Launched in 2013, DPLA aggregates metadata from cultural heritage institutions, including libraries, archives, and museums. Records are received through organizations referred to as hubs. Large organizations can serve as content hubs. Smaller organizations often band together as service hubs, typically via state and regional collaborations. Metadata comes from a variety of formats, including MARC, MODS, Dublin Core, and local customized standards. A common scenario is sharing Dublin Core metadata through OAI-PMH. But standards and harvesting differ by hub. DPLA has a date field labeled in the interface as Created Date.[57] According to DPLA's metadata application profile, it maps to dc:date, and the usage is "Date values as supplied by Data Provider."[58] It is a recommended field but not required. It can also store multiple date values or strings. DPLA recommends the use of the Extended Date/Time Format in this field.[59] Harvested date values may then be repurposed for the Date search feature in DPLA. The temporal enrichment process converts date values, as applicable, to begin and end dates.

The complete data set is available as a Bulk Download from DPLA.[60] The February 2021 data set was downloaded as JSON files from their public Amazon Simple Storage Service (S3) bucket. The data was processed using R and Python with the help of an advanced research computing cluster.[61] In the process of analyzing data, it was discovered that duplicated records were present in the data set, comprising about 0.1 percent of the total number of records. They were deduplicated before continuing with further analysis. Zavalina et al. did not mention any duplicate records in their data set. Consequently, it is unclear how often this issue occurs in the data sets supplied by DPLA in buckets or its Application Programming Interface (API).

Within the Date field, multiple values or strings can be in a record. As with the prior study, the values in the records were separated for analysis. In the following results, common syntax for values, e.g., YYYY, is referred to as patterns. To identify common patterns, all dates were run through a function that performed the following transformations: (1) each digit was converted to a 9, (2) each full month name (e.g., April) was converted to "Month," and (3) each abbreviated month name (e.g., Aug) was converted to "Mon." The rest of the content of the date field was left untouched. As examples, "November 20, 1913" would be transformed into "Month 99, 9999" whereas "Afternoon of 10 Feb 2011" would be transformed into "Afternoon of 99 Mon 9999." By transforming the dates in this manner, similarly formatted dates could be collapsed together to calculate cross-tabulations.

This study utilized the University of North Texas's EDTF validation Python library to identify EDTF date values.[62] The date patterns were reviewed according to the EDTF rules. Consequently, six patterns (>600 values) were eliminated as invalid according to EDTF rules. In addition, further analysis was needed to present this study's results. In the case of 9999-99 date patterns, it was necessary to separate which values represented calendar dates with reduced precision for month and year (YYYY-MM), seasons, sub-year groupings, or invalid patterns.

Findings

From 8,012,390 records in Zavalina et al.'s study, the DPLA had grown significantly to 41,740,821 records.

Frequency of Date Values

The total number of date values was 26,358,343. While trending toward fewer values, the number of date values in a record varied from zero to as many as 881. There were also 16,598,377 records without dates. Compared to 83 percent of records having at least one date value in the prior study, there was a drop to 60 percent of records having at least one date.[63] The newer percentage is in line with prior studies.[64] Content hubs had results on both extremes (figure 1). David Rumsey's records all had at least one date value. Meanwhile, the National Archives and Records Administration had the most records with no date values—more than 80 percent of their records in DPLA. Service hubs were more consistent as a set and fared better (figure 2). On the low end, less than 1 percent of Digital Maryland records lacked any date value. On the other extreme, 45 percent of Digital Library of Tennessee records lacked dates. Overall, content hubs and service hubs had dates, respectively, for 48.9 percent and 84.5 percent of records. These results stand in contrast to Zavalina et al., who found little difference between content and service hubs, respectively, with 83.4 percent and 80.9 percent of records including dates.

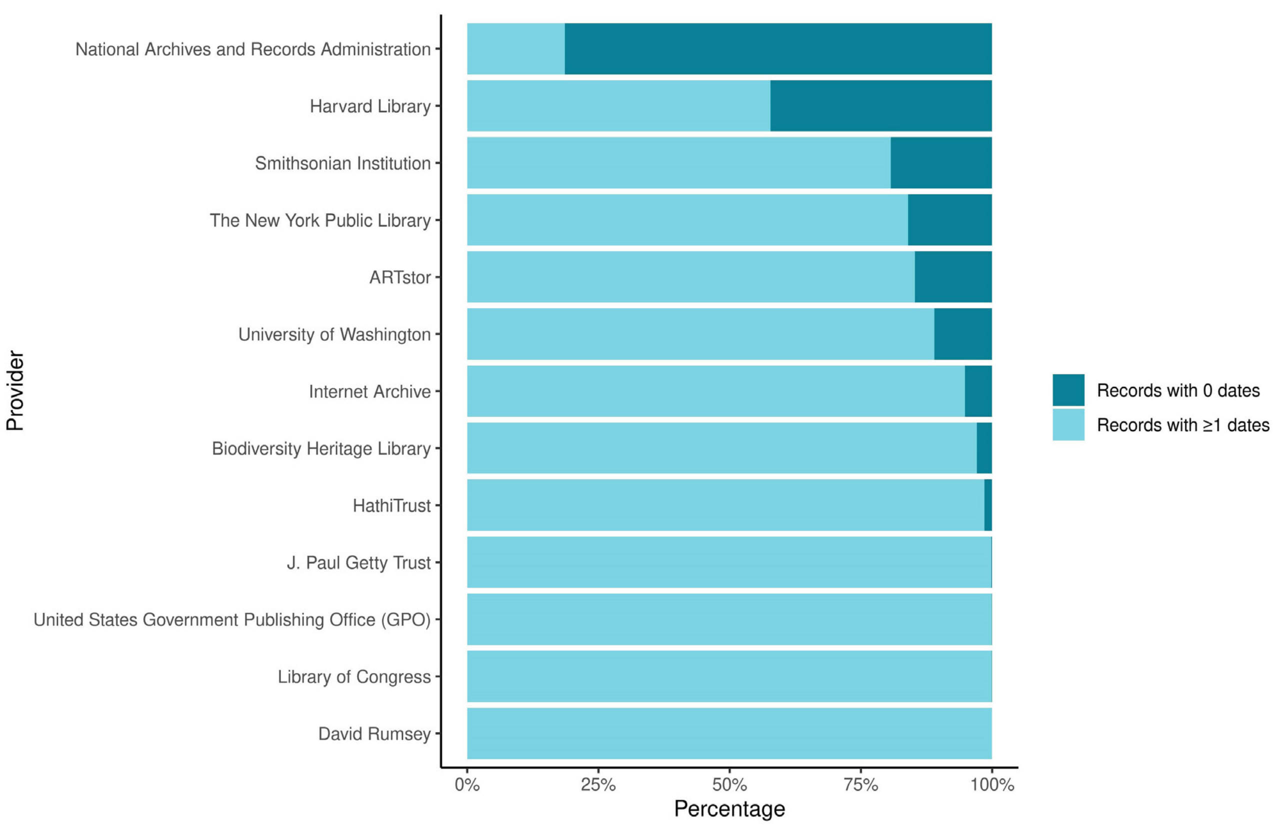

Figure 1. Bar chart for DPLA content hub records with and without date values.

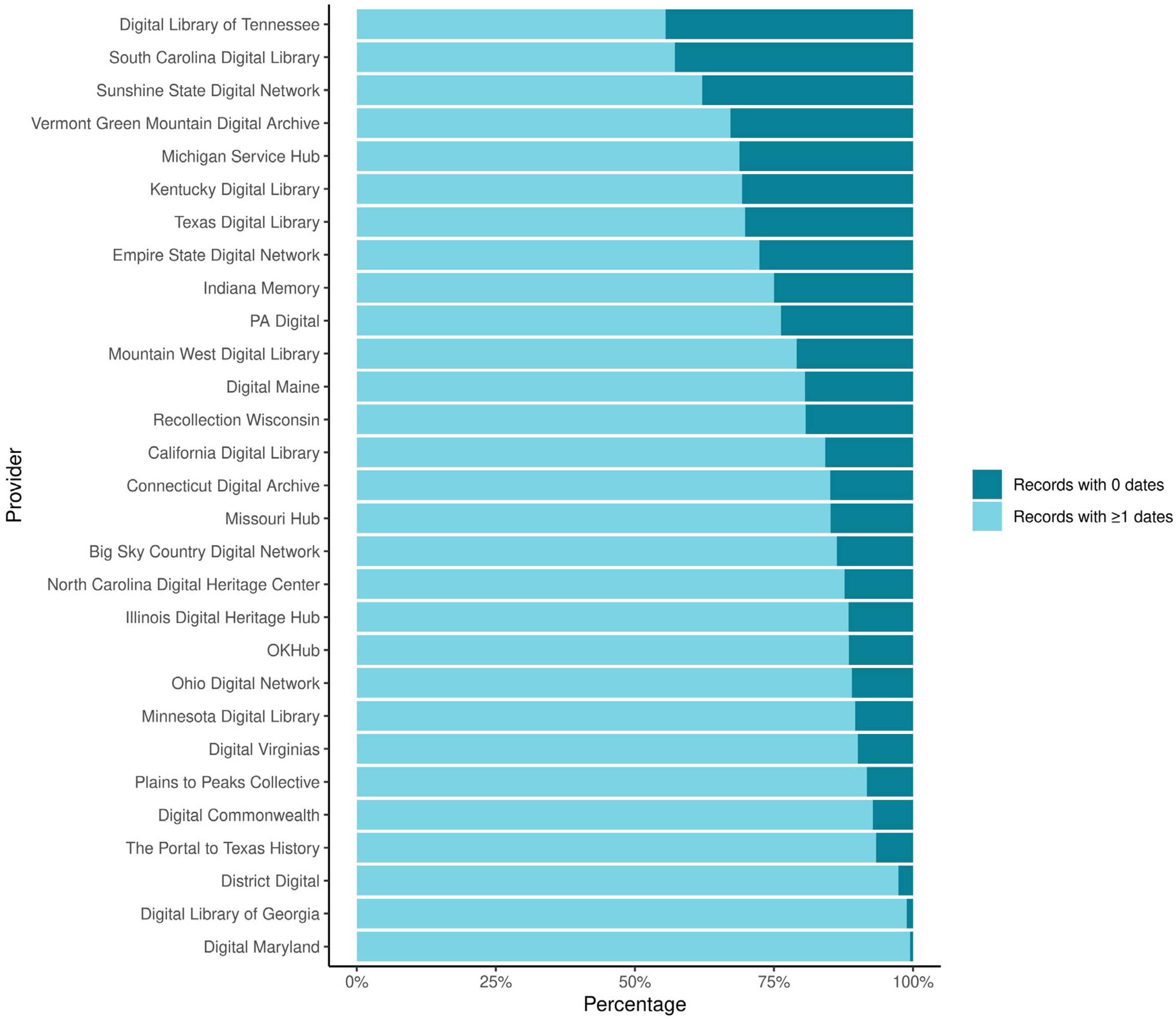

Figure 2. Bar chart for DPLA service hub records with and without date values.

EDTF Compliance

Of 26,358,343 date patterns, 68 percent were initially deemed valid according to EDTF rules and represented 176 different types of valid date patterns. This is an increase compared to the 51 percent EDTF validity for date values in 2015, according to Zavalina et al.[65] While 67 percent of the values conformed to Level 0, a mere 0.5 percent and 0.3 percent of values followed Levels 1 and 2, respectively (figure 3). The following sections detail this study's findings in detail for each EDTF level.

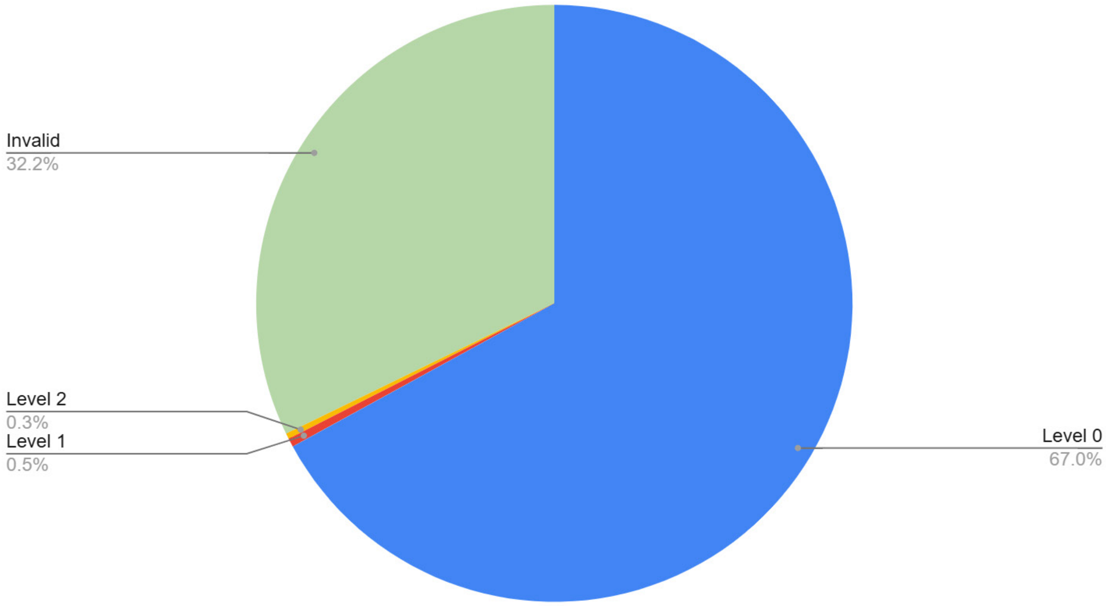

Figure 3. Pie chart comparing date patterns by EDTF validity.

Level 0

Level 0 (L0) date patterns represented 98.9 percent of EDTF-compliant date strings. This is a less than 1 percent drop from 2015, as per Zavalina et al. The widely available W3CDTF date patterns represent the bulk of the results. Every hub had values that matched L0 and overlapped with WC3DTF—58.6 percent of all date patterns in DPLA. As per tables 1 and 2, the most common patterns, in order of results, were year (YYYY), exact date (YYYY-MM-DD), and time intervals with exact date (9999-99-99/9999-99-99). The YYYY and YYYY-MM-DD patterns were also the most used of all date patterns in the DPLA, regardless of compliance with EDTF.

Table 1. EDTF Level 0 patterns that conform to W3CDTF and ISO 8601

Pattern	Number of patterns	Percentage of all patterns
9999	8,292,952	31.5
YYYY-MM	641,777	2.4
9999-99-99	5,741,424	21.8
9999-99-99T99:99:99	126,941	0.5
9999-99-99T99:99:99Z	655,142	2.5

There are fewer results for intervals, i.e., date ranges with a solidus (/) between them. An example with calendar year precision is “1993/1994,” which indicates that the date begins sometime in 1993 and ends sometime in 1994. Some variations provide more or less precision in terms of whether a month and day are known. While every hub had date patterns that matched at least one W3CDTF date pattern, fifteen hubs did not use time intervals at all. The National Archives and Records Administration had the most significant number of L0 time intervals: 67 percent of values. However, most hubs with results in this category had far fewer. The 9999-99-99/9999-99-99 and 9999/9999 patterns are most common in this set, with the latter representing less than 1 percent of all date values.

Table 2. EDTF Level 0 time intervals

Pattern	Number of patterns	Percentage of all patterns
9999-99-99/9999	636	<0.1
9999-99-99/9999-99	114	<0.1
9999-99-99/9999-99-99	2,002,535	7.6
9999-99/9999	46	<0.1
9999-99/9999-99	6,160	<0.1
9999-99/9999-99-99	318	<0.1
9999/9999	201,918	0.8
9999/9999-99	48	<0.1
9999/9999-99-99	845	<0.1

Level 1

For Level 1 (L1), more rules are added to the basics established for Level 0 dates. However, the results were far fewer than Level 0 at 0.5 percent. A total of 121,919 values represent fifty-five date pattern types. Theoretically, the rules allow for more patterns, but there were as few as one result for some patterns identified. Nine patterns have 1,000 or more values (table 3). Notably, there were no values that conformed to a letter-prefixed calendar year, i.e., dates that exceeded four digits—either before -9999 or after 9999—with a Y at the start of the string. Theoretically, these could be used for items such as prehistoric artifacts.

Some L1 date patterns in this set may express the same meaning in EDTF but overlap with another content standard. The most common pattern in this set is "YYYY?" with 39,395 instances. This format was valid under the *Anglo-American Cataloguing Rules*, Second Edition (AACR2) and had been used in MARC records for many years before the introduction of EDTF. Another interesting case is seasons: patterns of YYYY-21, YYYY-22, YYYY-23, and YYYY-24 where 21, 22, 23, and 24 represent spring, summer, autumn, and winter, respectively. The authors sampled values for review. Except for the Portal to Texas History, most season values are likely non-EDTF values. This is discussed more in the Analysis section. Nine hubs do not utilize L1 patterns at all: Biodiversity Heritage Library, David Rumsey, Digital Commonwealth, Digital Maine, Internet Archive, Kentucky Digital Library, Library of Congress, National Archives and Records Administration, and Texas Digital Library.[66]

Table 3. Level 1 patterns with 1,000 or more results

L0 Rules	L1 Rules	Pattern(s)	Number of results
Date	Seasons	YYYY-21 YYYY-22 YYYY-23 YYYY-24	6,370[a]
Date	Qualification of a date (complete)	9999-99~	3,566
Date	Qualification of a date (complete)	9999?	39,395
Date	Qualification of date (complete)	9999~	27,107
Date	Unspecified digit(s) from the right	999X	10,144
Date	Negative calendar year	-9999	3,489
Time Interval	Negative calendar year	1	4,004
Time Interval	Negative calendar year	-1	1,068
Time Interval	Qualification of date (complete)	9999/9999?	19,256
Time Interval	Qualification of date (complete)	9999~/9999~	4,666

[a] Due to sampling, many of these values may be considered invalid. While they match EDTF season encoding, they appear to be improperly formatted date spans, i.e., YYYY-YY. The Portal of Texas History's 3,166 values may be the only valid season patterns.

Level 2

Level 2's (L2) options build upon the rules of L0 and L1 dates. However, they were little utilized at 0.3 percent of all date values. With an infinite number of potential date patterns, 109 L1 patterns were observed in a total of 80,460 date values. There were no Level 2 patterns that aligned to the rules for exponential years, significant digits, or intervals. While L2 rules are robust, they can be confusing. As with L1's validation results, some matches may be considered false hits. As with the L1 seasons, sub-year groupings were sampled. No valid examples were found.

Table 4. Level 2 patterns with 100 or more results

L0 Rules	L1 Rules	L2 Rules	Pattern	Number of results
Date	Seasons	Sub-year groupings	YYYY-21 to YYYY-31	6,441[a]
Date	N/A	Set representation - All members	{9999-99-99,9999-99-99}	105
Date	N/A	Set representation - One of a set	[9999-99-99,9999-99-99]	650
Date	N/A	Set representation - One of a set	[9999-99-99..9999-99-99]	49,994
Date	N/A	Set representation - One of a set	[9999-99..9999-99]	204
Date	N/A	Set representation - One of a set	[9999,9999]	3,831
Date	N/A	Set representation - One of a set	[9999..]	2,825
Date	N/A	Set representation - One of a set	[9999..9999]	15,624

[a] Due to sampling, most or all of these values may be considered invalid. While they match EDTF season encoding, they appear to be improperly formatted date spans, i.e., YYYY-YY.

Compliance at the Hub Level

Content and service hubs represented both the highest and lowest percentages of EDTF valid records, with service hubs faring only slightly better than content hubs. Only three hubs had all their values in complete compliance with EDTF: Biodiversity Heritage Library, Digital Maine, and Kentucky Digital Library. However, they utilized fewer date pattern types and only complied with Level 0. Digital Maine solely used YYYY-MM-DD, while the Biodiversity Heritage Library and Kentucky Digital Library restrict values to the YYYY pattern. As the YYYY and YYYY-MM-DD patterns are well established in ISO 8601 and W3CDTF, the compliance with EDTF may be correlational only. Some hubs only employed a single date pattern within Level 0. Though each had results that did not match EDTF, David Rumsey typically used the YYYY pattern, while the Internet Archive results often matched the 9999-99-99 and 9999-99-99T99:99:99Z patterns. Eleven hubs had 90 percent to nearly 100 percent of EDTF valid date patterns. In this set, the Portal to Texas History was a standout, which employed more than 100 unique EDTF date pattern types with nearly 100 percent EDTF validity across a million and a half date patterns. When complex EDTF date examples from this hub were sampled, the examination proved satisfactory regarding their accuracy. This hub's results demonstrate both the value of EDTF encoding to express complicated dates and provide human-readable values in its own repository through scripting.

Non-EDTF Values

As with Zavalina et al., the results included a wide variety of date patterns that encompassed encoded dates, text dates, and values with both text and encoded dates. Including EDTF date patterns, at least 40,000 unique date patterns were recognized. Assorted languages and syntax were observed. Some values include text as pseudo-qualifiers, e.g., "Digitized:" and "Transcribed:" at the beginning. There were historical periods and dynasties. However, some values were unclear in terms of their semantics. For example, "c." before a string may correspond to copyright or circa. There were also garbage values, e.g., "undated" and "unknown." Some values were mismapped to the date field, including publisher name and location.

Analysis

The findings show mixed results. In relation to Zavalina et al. in 2015, EDTF compliance rose 17 percent by 2021. Some hubs only utilized one date pattern that corresponds to a Level 0 pattern. In addition, some hubs complied with Level 0 but did not support Levels 1 and 2. Biodiversity Heritage Library, Digital Commonwealth, Kentucky Digital Library, Library of Congress, National Archives and Records Administration, and Texas Digital Library fell in that category. Conversely, the number of records with any date values decreased by 23 percent. EDTF has the potential to be a useful encoding standard, and its adoption is slowly rising. But the results suggest that the more complicated the date scenarios and rules, the less likely they are to be employed. In addition, date metadata in an unruly form may be preferable to a growing number of records without any date metadata. The robust options in Levels 1 and 2 appear to be less frequently used by LAMs. It is not clear if it is because they are confused or if the standard is not mature enough. The Portal to Texas History shows that EDTF can be used extensively. However, time will tell if this vanguard institution will have more join it in this endeavor.

Levels of Precision

Level 0 had the most use in DPLA. However, EDTF intervals, e.g., YYYY/YYYY, remain an oddity. The non-EDTF pattern YYYY-YYYY was more common by comparison. Another consideration is that DPLA encourages the use of EDTF but uses YYYY-YYYY for its temporal enrichment process to support searching. It is not clear if that impacts the use of this non-standard pattern that is commonly used. There is flexibility when applying EDTF rules in terms of precision. These considerations may be subjective or practical. For instance, the National Archives and Records Administration (NARA) has a record with a date of 1941-01-01/1945-12-31.[67] One can infer from the value that the month and day related to this work may be unknown. With a large date range of 1941 to 1945, it could be equally appropriate to reduce precision to 1941/1945. The source record has collection-level text that suggests a date range of 1942 to 1945.[68] In addition, the Internet Archive uses zeroes as placeholders for minutes and seconds in date values, e.g., “1986-01-01T00:00:00Z.”[69] This level of precision, while not erroneous, could be deemed excessive.

Platform-Specific EDTF Inaccuracies

However, rigid compliance with EDTF rules did not necessarily ensure metadata accuracy. Through sampling, it was revealed that Digital Maine has a large number of records dated January 1 of a specific year, i.e., YYYY-01-01. But most of these do not correspond to January 1. The hub’s metadata documentation indicates that there are four date patterns that are permitted with a text box to enter the year.[70] It also notes issues with bulk entering dates that its Bepress platform changes to the first date of the year. For instance, a record has the date value of “1940-01-01” but states that it is “dated June 14, 1940” in the description field.[71]

False Matches to EDTF

Conversely, there are cases where the EDTF validation matches a pattern but is likely a false match. A case is L1 dates with hyphens for negative dates, i.e., BCE dates. As per Klose, Goldstein, and Levy (2021), date values that fall into this category are sometimes modern dates that represent open date ranges.[72] Another set of false matches was supposed EDTF Seasons (Level 1) and sub-year groupings (Level 2). Most appeared to be improperly created date intervals, i.e., YYYY-YY, where the second year in the date span was only indicated with the last two digits. Of the 6,370 values that matched EDTF seasons encoding, only sampled values from the Portal of Texas History’s 3,166 records were true EDTF encoded seasons. In these cases, the Portal to Texas History has converted the values into human-readable dates in its platform for display. For instance, a harvested DPLA record has “2017-21” as Created Date.[73] In the institution’s portal, this record displays as “Spring 2017” as the Creation Date.[74]

Non-EDTF Values

The non-EDTF values mimicked what was observed in Zavalina et al. and prior studies. While a percentage of these values could be converted easily to EDTF, the diverse variations are largely outside of the capacity of an aggregator, including the DPLA, to normalize. Some of the date patterns represented commonly understood syntax that could easily be converted to EDTF. However, some values require interpretation, e.g., historical periods and dynasties, to convert to a Gregorian calendar for EDTF. In addition, historical periods are more appropriate for the coverage field and its temporal qualifier. Some values would require the source institution to review the records due to ambiguous syntax, localizations, and mismappings. The UNT EDTF Python script identified six date patterns, which were eliminated in this study as invalid. These represented a mere 619 date values. Owing to the complexity of EDTF and the reality of metadata creation, programmatic approaches to EDTF validation may be challenging.

Date Issues with Representing Works

Various oddities and inconsistencies were observed and explored through sampling. A monograph, *The Sneads of Fluvanna* by Oranie Virginia Snead, had radically different date values depending on the source. One institution had 880 valid YYYY dates from 1000 to 1910. However, the broad date range is related to aboutness rather than creation. The publication's title page lists a date of 1910. Two other records for this book avoid this error. As with prior studies, there were mismappings and conflicts with the one-to-one principle. For instance, the Smithsonian Institution hub has records for amphoras—tall, handled jars or vases—entered in different ways. Depending on the museum, the date values were mapped from the date of the artifact, the accession date, or there was no date at all.

Limitations

The results still leave knowledge gaps. First, metadata records are in a state of flux. As per sampling, some metadata records had their values updated later. In addition, records remain in place for lapsed hubs. Second, metadata mappings vary across institutions and hubs. They can also change over time. There is likely relevant date metadata in records that did not get mapped to the DPLA's date field. The DPLA field label "Created Date," for instance, is problematic semantically. Lastly, date values that correlate to EDTF rules do not necessarily mean they were intended to do so. While the source metadata from the institution or hub may or may not indicate an encoding standard, the resulting DPLA record does not include that information.[75] Hubs may also convert data before sending it to DPLA. Another consideration with these findings is the high number of date patterns that are valid in another descriptive cataloging standard but do not conform to EDTF. Essentially, one would need to analyze the source records to get a more complete picture of EDTF compliance.

Recommendations

While EDTF use may be slowly growing in LAMs, there are ways to increase its implementation and improve outcomes for searching by date. Some recommendations are to create shareable metadata from the outset, be consistent and normalize values, think programmatically, document your decisions, check your mappings, educate and update guidelines, and advocate for date encoding support. These recommendations are not ranked. Their application depends on each institution's capability to do so.

Create shareable Metadata from the Outset

If metadata is not consistent in a collection or platform, it will have less usefulness for users, whether or not it is shared. Consequently, it is best to create metadata that is shareable from the outset, preferably with encoded date values. Even if your platform(s) do not support all of EDTF, normalizing applicable date values to W3CDTF or a subset of EDTF is usually an improvement. There is a potential workaround for platforms that do not support EDTF. Martin notes using a hidden encoded date field and a public text date field in CONTENTdm.[76] However, that requires careful mapping for harvesting. Even if there are no current plans to share metadata, consider the potential impact for researchers. Data sets for cultural heritage materials are beneficial in aggregate, e.g., collections as data.

Be Consistent and Normalize Values

Normalize values in legacy records. They may represent a wide variety of metadata creators, content standards, and local practices. Periodically review your institution's metadata to ensure consistency through remediation. Look for gaps. Avoid garbage values like "undated" as they do not help users. Sometimes, a broad date range, e.g., a century, based on a collection-level date range or educated assessment, may be appropriate. Details about the creation date found in the description could be used to augment records lacking suitable values.[77]

Think Programmatically

When possible, organizations should look for programmatic solutions for remediating date metadata. For example, Phillips and Tarver show strategies for normalizing EDTF dates with OpenRefine.[78] In addition, there is a UNT tool to aid with identifying or validating EDTF date strings, which is highly beneficial to metadata creators with varying experience levels.[79] Even with simpler date patterns, e.g., YYYY-MM-DD, audit tools that automate error checking can be beneficial to ensure consistency.[80] Automation can also be employed to remove characters that may impact the interoperability of date metadata for harvesting. For instance, square brackets for dates in AACR2 and *Resource Description & Access* (RDA) have a different meaning than in EDTF.

Remediation

There will always be legacy records that do not reflect the current best practices. Remediation of records is often time-consuming and expensive. However, it is a worthwhile process for providing access to LAMs' content, including digital collections. While cataloging and metadata creation have often been seen as a one-time activity, they should be seen as part of an iterative process in terms of providing impact.[81] A useful framework is Tarver and Phillips' EPIC model.[82]

Document Decisions

Document internal metadata decisions for platforms(s) and specific projects so that there is an understanding of why and how values are entered. As staff turnover and memories fade, internal platform and project documentation (including application profiles) is critical. If there are barriers to moving forward with encoded values, understanding the *why, how*, and *where* can ensure that future staff can pick up where current efforts cannot. An example is local workarounds for platform limitations, e.g., having separate fields for text and searchable date values.

Check Your Mappings

Institutions that decide to share their metadata should check their metadata mappings. Often, large-scale crosswalks and metadata transformations available are a one-size-fits-all approach. As there are often localizations, reviewing metadata mapped for sharing, e.g., OAI feed, can lead to useful discoveries. In addition, it is helpful to crosswalk your local metadata to ensure a best fit for sharing metadata. By understanding the strengths and weaknesses in your institutional and/or project metadata, you can better find ways to deal with them. This also goes hand-in-hand with normalizing metadata and documentation. As encoded date values are preferable to text date values, knowing which ones are appropriate matters. This task can also help with discovering what you do not want to share, e.g., date digitized and publisher name, and what you do want to share, e.g., date created and date published. Even sampling records can be advantageous.

When mapping from catalog records, there are also better and worse options. MARC's encoded dates are not displayable to users. Mapping from fixed fields in MARC, e.g., the 008 field, rather than display fields, could be beneficial. In display dates, square brackets have a clear meaning. Those characters in MARC could be used to convert existing values into EDTF as well. This concept could be extended to text dates that apply AACR2 and RDA with non-MARC metadata. However, this means spending time converting dates to be consistent. Also, there are limitations to using the 046 field. Not every system—if any—knows what to do with a 19uu. As EDTF is a newer innovation, it is not clear how many MARC catalogers use EDTF in the 046 field in bibliographic records. In the past, catalogers were discouraged from using this field. It is only meant to be used as an expansion of the 008 field. The user-friendly way of creating dates is in the 264 field. Encoding is used in case libraries convert to a system that could use this information, much like preparation for the conversion to the Bibliographic Framework Initiative (BIBFRAME). Time periods can be addressed in the 388 field.

Educate and Update Guidelines

EDTF can be confusing, and education is necessary. If catalogers and metadata creators do not understand it, they will not use it. Within the American Library Association, the Association for Library Collections & Technical Services (ALCTS) approved the Core Competencies for Cataloging and Metadata Professional Librarians in January 2017.[83] It pointed to knowledge competencies in relation to data standardization, structure standards, data encoding, format, exchange standards, and value standards.[84] The December 2023 revision of the Core Competencies, which incorporated survey feedback from the library and metadata community, added EDTF and W3C among data value encoding schemes in its appendix.[85] Therefore, it is hopeful that resources such as this will serve as a model for LAMs in terms of best practices and that an awareness of EDTF can permeate the field.

Advocate

Platforms that host metadata for cultural objects need to support date encoding, including EDTF. Standards are only beneficial if leveraged to support search and browse while also presenting dates with a human-readable output. In addition to the list of EDTF implementations on its website,[86] there are additional GitHub repositories that support it in various programming languages. The future of EDTF support depends on coders and implementers. One such example is a Ruby-based GitHub repository that supports a subset of EDTF.[87] The Marva BIBFRAME Editor supports EDTF.[88] In terms of repositories, some Samvera Hyrax adopters have utilized a subset of EDTF. But date encoding support is not a given on many platforms. Therefore, it behooves institutions to demand encoded date support from vendors. EDTF may seem odd initially. However, every technological innovation seemed new when it was first introduced.

Conclusion

Date metadata needs to be standardized to be useful. Traditionally, dates and date ranges have been expressed in diverse ways through both text and encoding. The values themselves can be confusing to many users. While there have been gaps in how to encode more complex date scenarios, including "squishy" dates, EDTF supports those.[89] Consequently, there is a need for LAMs to use and advocate for this standard to provide a baseline for consistent, quality date metadata. Not only is EDTF useful for providing discovery across diverse records, but it could also allow the records to be leveraged in new and exciting ways, including scholarship and visualization. This requires the aid of coders and an investment by vendors.

Sadly, date metadata can be aggravating for metadata creators and end users alike. It is particularly relevant to users searching digital collections for cultural heritage. Unfortunately, date metadata inconsistency inhibits search and retrieval. A delay in providing machine-readable dates also hampers the progress of big data and collections as data projects. The plethora of date patterns observed in the DPLA suggests there is still work to be done. The variety of date strings suggests that much of it is beyond the scope of an aggregator, including the DPLA, to normalize through automation. This work of standardizing date metadata is best approached by the metadata creators themselves. In addition, there is related work that can be addressed by LAMs.

Increasing the number of records that both have dates and conform to EDTF would have a significant impact on search and retrieval, as well as users' comprehension of records. Lastly, the most troubling aspect of this study is the rise of records that have no date values at all. Sampling also found structural and semantic issues that have plagued the community for decades. More research is needed to understand the choices and challenges that metadata creators face with date metadata.

Acknowledgments

The authors thank Hannah Tarver and Amanda Rinehart for reviewing an earlier draft of this article and giving feedback on it.

Notes

1. Elina Late, Hille Ruotsalainen, and Sanna Kumpulainen, "Image Searching in an Open Photograph Archive: Search Tactics and Faced Barriers in Historical Research," *International Journal on Digital Libraries* 25, no. 4 (2024): 715–728, https://doi.org/10.1007/s00799-023-00390-1.
2. Steven Gentry et al., *Survey of Benchmarks in Metadata: Initial Findings* (Digital Library Federation Assessment Interest Group Metadata Assessment Working Group, May 2020): 15–17, https://dlfmetadataassessment.github.io/assets/pdf/2020-dlf-mawg-mqb-white-paper.pdf.
3. Muriel Foulonneau and Jenn Riley, *Metadata for Digital Resources: Implementation, Systems Design and Interoperability* (Chandos Publishing, 2008), 92–93.
4. Ryan Shaw and Murray Maloney, "The Forms of Resource Description," in *The Discipline of Organizing*, ed. Robert J. Glushko (MIT Press, 2013), 338.
5. Foulonneau and Riley, *Metadata for Digital Resources*, 92–93.
6. Sarah L. Shreeves, Jenn Riley, and Liz Milewicz, "Moving towards Shareable Metadata," *First Monday* 11, no. 8 (2006), https://doi.org/10.5210/fm.v11i8.1386.
7. Thomas R. Bruce and Diane I. Hillmann, "The Continuum of Metadata Quality: Defining, Expressing, Exploiting," in *Metadata in Practice*, ed. Diane I. Hillmann and Elaine L. Westbrooks (American Library Association, 2004), 243.
8. Shreeves, Riley, and Milewicz, "Moving towards Shareable Metadata."
9. Priscilla Caplan, *Metadata Fundamentals for All Librarians* (ALA Editions, 2003), 33.
10. Foulonneau and Riley, *Metadata for Digital Resources,* 119.
11. Foulonneau and Riley, *Metadata for Digital Resources,* 119–122.
12. "FAIR Principles," GO FAIR, accessed May 23, 2025, www.go-fair.org/fair-principles/.
13. *International Journal on Digital Libraries* devoted volume 22, issue 3, to FAIR data and cultural heritage.

14. Nicola Barbuti, "Thinking Digital Libraries for Preservation as Digital Cultural Heritage: By R to R4 Facet of FAIR Principles," *International Journal on Digital Libraries* 22, no. 3 (2020), https://doi.org/10.1007/s00799-020-00291-7; Brian Dobreski, Heather Moulaison-Sandy, and Bradley Wade Bishop, "How FAIR is MARC?: FAIR Data Principles and Bibliographic Data," *International Conference on Dublin Core and Metadata Applications, 2022* (2023), https://doi.org/10.23106/dcmi.953165517.

15. Kyle Banerjee and Terry Reese, Jr., *Building Digital Libraries: A How-To-Do-It Manual* (ALA Neal-Schuman, 2019), 167.

16. Banjeree and Reese, *Building Digital Libraries,* 169–170.

17. Banjeree and Reese, *Building Digital Libraries,* 168.

18. Foulonneau and Riley, *Metadata for Digital Resources,* 180.

19. Kevin Clair, "Technical Debt as an Indicator of Library Metadata Quality," *D-Lib Magazine* 22, no. 11/12 (2016): section 3, https://doi.org/10.1045/november2016-clair.

20. "Date and time—Representations for information interchange—Part 1: Basic rules," ISO 8601-1, First edition (Geneva, Switzerland: International Organization for Standards, February 2019), 9.

21. Misha Wolf and Charles Wicksteed, "Date and Time Formats," W3C, August 27, 1998, www.w3.org/TR/NOTE-datetime.

22. Arwen Hutt and Jenn Riley, "Semantics and Syntax of Dublin Core Usage in Open Archives Initiative Data Providers of Cultural Heritage Materials," *Proceedings of the 5th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL '05)* (2005), 265, https://doi.org//10.1145/1065385.1065447.

23. "EDTF -- Background," Library of Congress, January 28, 2022, www.loc.gov/standards/datetime/background.html.

24. "EDTF -- Background."

25. "Extended Date/Time Format (EDTF) Specification" (Library of Congress, February 4, 2019), www.loc.gov/standards/datetime/.; "EDTF -- Background." Notably, ISO 8601-2:2019(E) cites a number of Library of Congress EDTF concepts while expanding upon them. In addition, it provides EDTF as a profile in Annex A. "Date and time — Representations for information interchange — Part 2: Extensions," ISO 8601-2 (Geneva, Switzerland: International Organization for Standards, February 2019), 12–13, 28–29, 35, 39–40, 57–62.

26. "Extended Date/Time Format, Datatypes Scheme," Linked Data Service, Library of Congress, modified February 7, 2023, http://id.loc.gov/datatypes/EDTFScheme.

27. Hannah Tarver and Mark Phillips, "Lessons Learned in Implementing the Extended Date/Time Format in a Large Digital Library," *International Conference on Dublin Core and Metadata Applications, 2013* (2013): 228–233, https://doi.org/10.23106/dcmi.952136100.

28. Hannah Tarver et al., "Managing Serials in a Large Digital Library: Case Study of the UNT Libraries Digital Collections," *The Serials Librarian* 68, no. 1–4 (2015): 357, https://doi.org/10.1080/0361526X.2015.1017715.

29. Ashlea M. Green, "Metadata Application Profiles in U. S. Academic Libraries: A Document Analysis," *Journal of Library Metadata* 21, no. 3–4 (2022): 123–124, https://doi.org/10.1080/19386389.2022.2030172.

30. "Special Coded Dates (MARC 046)," Irving S. Gilmore Library, Yale University Libraries, accessed June 19, 2025, https://web.library.yale.edu/cataloging/music/MARC046.

31. Kelley McGrath, "Musings on Faceted Search, Metadata, and Library Discovery Interfaces," *Cataloging & Classification Quarterly* 61, no. 5–6 (2023): 455, https://doi.org/10.1080/01639374.2023.2222120.

32. Annamarie C. Klose, Scott Goldstein, and Morris S. Levy, "Numismatics & Bibliographic Description: How Rutgers University Libraries Described Coins with MODS," *Journal of Library Metadata* 22, no. 1–2 (2022): 88–90, https://doi.org/10.1080/19386389.2022.2051979.

33. Katherine Manifold et al., "Fragments in Cyberspace: The Ohio State University's Digital Collection of Medieval and Renaissance Manuscript Fragments," *Journal of Library Metadata* 25, no. 1 (2025): 18, https://doi.org/10.1080/19386389.2024.2403927.

34. Sarah L. Shreeves, Ellen M. Knutson, Besiki Stvilia, Carole L. Palmer, Michael B. Twidale, and Timothy W. Cole, "Is 'Quality' Metadata 'Shareable' Metadata? The Implications of Local Metadata Practices for Federated Collections," *ACRL Twelfth National Conference* (2005): 228–229, http://hdl.handle.net/2142/145.

35. Mary Burke and Oksana L. Zavalina, "Descriptive Richness of Free-Text Metadata: A Comparative Analysis of Three Language Archives," *Proceedings of the Association for Information Science and Technology* 57, e429 (2020): 3, https://doi.org/10.1002/pra2.429.

36. Jung-Ran Park, "Metadata Quality in Digital Repositories: A Survey of the Current State of the Art," *Cataloging & Classification Quarterly* 47, no. 3–4 (2009): 220, https://doi.org/10.1080/01639370902737240.

37. Hutt and Riley, "Semantics and Syntax of Dublin Core Usage," 264.

38. Hutt and Riley, "Semantics and Syntax of Dublin Core Usage," 263.

39. Foulonneau and Riley, *Metadata for Digital Resources,* 120–121.

40. Foulonneau and Riley, *Metadata for Digital Resources,* 121.

41. Greta de Groat, *Future Directions in Metadata Remediation for Metadata Aggregators* (Digital Library Federation, 2009), 18, https://old.diglib.org/aquifer/dlf110.pdf.

42. Diane I. Hillmann, Naomi Dushay, and Jon Phipps, "Improving Metadata Quality: Augmentation and Recombination," *International Conference on Dublin Core and Metadata Applications, 2004* (2004): 2, https://doi.org/10.23106/dcmi.952107705.

43. Steven Jack Miller, *Metadata for Digital Collections*, 2nd ed. (ALA Neal-Schuman, 2022), 98.

44. Gentry et al., "Survey of Benchmarks," 12–13.

45. "Bitter Harvest," Roy Tennant, last modified May 14, 2004, https://web.archive.org/web/20211128083729/http://roytennant.com/bitter_harvest.html.

46. Shreeves, Riley, and Milewicz, "Moving towards Shareable Metadata."; Tennant, "Bitter Harvest."

47. Late, Ruotsalainen, and Kumpulainen, "Image searching in an open photograph archive," 718, 721–723.

48. "Open Archives Protocol for Metadata Harvesting: Protocol Version 2.0 of 2002-06-14," Open Archives Initiative, modified January 8, 2015, section 2.5 Record, www.openarchives.org/OAI/openarchivesprotocol.html#Record.

49. DCMI defines the coverage field as "The spatial or temporal topic of the resource, spatial applicability of the resource, or jurisdiction under which the resource is relevant." It also notes that "Temporal topic may be a named period, date, or date range." However, its guidance includes the recommendation that "Because coverage is so broadly defined, it is preferable to use the more specific subproperties Temporal Coverage and Spatial Coverage." "DCMI Metadata Terms," Dublin Core, January 20, 2020, www.dublincore.org/specifications/dublin-core/dcmi-terms/.; Timothy W. Cole and Sarah L. Shreeves, "Lessons Learned from the Illinois OAI Metadata Harvesting Project," in *Metadata in Practice*, ed. Diane Hillman and Elaine L. Westbrooks (Chicago: American Library Association, 2004), 187.; Shreeves, Riley, and Milewicz, "Moving towards Shareable Metadata."

50. "One-to-One Principle," Dublin Core, accessed May 11, 2025, www.dublincore.org/resources/glossary/one-to-one_principle/.

51. Miller, *Metadata for Digital Collections*, 100.

52. Banjeree and Reese, *Building Digital Libraries,* 169.

53. Hutt and Riley, "Semantics and Syntax of Dublin Core Usage," 263–264.

54. Amalia Beisler and Glee Willis, "Beyond Theory: Preparing Dublin Core Metadata for OAI-PMH Harvesting," *Journal of Library Metadata*, 9, no. 1–2 (2009), 78, https://doi.org/10.1080/19386380903095099.

55. Library of Congress, Network Development and MARC Standards Office, "MARC 21 Format for Bibliographic Data: 008 - Fixed-Length Data Elements-General Information (NR)" (Library of Congress, February 2010), revised March 5, 2010, www.loc.gov/marc/bibliographic/bd008.html.

56. Library of Congress, Network Development and MARC Standards Office, "MARC 21 Format for Bibliographic Data: 046 - Special Coded Dates (R)" (Library of Congress, June 2021), revised June 3, 2021, www.loc.gov/marc/bibliographic/bd046.html.

57. Temporal Coverage (dcterms:temporal) exists in harvested DPLA metadata. However, it is not displayed in the DPLA portal. It is only available to users via the API or the bulk download.

58. Digital Public Library of America and DPLA MAP Working Group, "Metadata Application Profile, version 5.0" (December 7, 2017), DPLA Pro, https://pro.dp.la/hubs/metadata-application-profile.

59. Digital Public Library of America, "DPLA Metadata Quality Guidelines," GoogleDocs, modified June 2022, 5, http://bit.ly/dpla-metadata-qual.

60. Digital Public Library of America, "Bulk Download," DPLA Pro, accessed February 15, 2021, https://pro.dp.la/developers/bulk-download.

61. The computing environment was originally called Compute Canada (https://computecanada.ca/); its operations have since become the Digital Research Alliance of Canada (https://alliancecan.ca/).

62. University of North Texas Libraries, "edtf-validate," GitHub, modified 2021, https://github.com/unt-libraries/edtf-validate.

63. Oksana L. Zavalina et al., "Extended date/time format (EDTF) in the Digital Public Library of America's metadata: exploratory analysis," *Proceedings of the Association for Information Science and Technology* 52, no. 1 (2015): 3, https://doi.org/10.1002/pra2.2015.145052010066.

64. Cole and Shreeves studied simple Dublin Core metadata for harvesting with OAI-PMH. Their results indicate the percentage of records with dc:date as follows: academic libraries (52%), digital library projects (63%), and museums and cultural/historical societies (64%). Cole and Shreeves, "Lessons Learned," 183–184.
65. Zavalina et al., "Extended Date/Time Format (EDTF)," 3.
66. Note: The DigitalMaine Repository is listed as Digital Maine on the DPLA website.
67. Office for Emergency Management. Office of War Information. Domestic Operations Branch. Bureau of Special Services. 3/9/1943-9/15/1945, "Cover Coughs, Cover Sneezes!" DPLA, accessed June 1, 2025, https://dp.la/item/cc1c8afb28c8f7e5abdbb1127fee6dc5.
68. Office for Emergency Management. Office of War Information. Domestic Operations Branch. Bureau of Special Services. 3/9/1943-9/15/1945, "Cover Coughs, Cover Sneezes!" National Archives Catalog, accessed June 1, 2025, https://catalog.archives.gov/id/514081.
69. "Annual report of the Town of Stewartstown, New Hampshire, 1985," DPLA, accessed June 1, 2025, https://dp.la/item/0004d056787cbf85e5f2cf5c564e43c5.
70. Maine State Library and Kate F. Herbert, "DigitalMaine Repository Master Metadata Guide," DigitalMaine Repository, Library Documents, 2017, https://digitalmaine.com/msl_docs/118.
71. "Alien Registration- Crossman, Benjamin H. (Hampden, Penobscot County)," DPLA, accessed June 22, 2025, https://dp.la/item/b00f92d0172b686804d01b7b8352eb2e.
72. Annamarie C. Klose, Scott Goldstein, and Morris Levy, "Out of the Past: The Challenges of B.C.E. Dates in Metadata," *Proceedings for the Association for Information Science and Technology* 58, no. 1 (2021): 757, https://doi.org/10.1002/pra2.552.
73. Junebug Clark, "[Undergraduate Students Walking into the Coliseum for Their Ceremony]," DPLA, accessed June 1, 2025, https://dp.la/item/347134b94a657c3f178006807ef7b096.
74. Junebug Clark, "[Undergraduate Students Walking into the Coliseum for Their Ceremony]," The Portal to Texas History, modified August 12, 2019, https://texashistory.unt.edu/ark:/67531/metadc1039463/.
75. In the DPLA record in the JSON format, there is the harvested record and the history of the source metadata, including its original encodings. However, DPLA's own date metadata does not support encoding within it.
76. Kristin E. Martin, "Marrying Local Metadata Needs with Accepted Standards: The Creation of a Data Dictionary at the University of Illinois at Chicago," *Journal of Library Metadata* 11, no. 1 (2011): 44, https://doi.org/10.1080/19386389.2011.545006.
77. Mary Burke and Oksana L. Zavalina, "Descriptive Richness of Free-Text Metadata: A Comparative Analysis of Three Language Archives," *Proceedings of the Association for Information Science and Technology* 57, e429 (2020): 5–6, https://doi.org/10.1002/pra2.429.
78. Mark Edward Phillips and Hannah Tarver, "Experiments in Operationalizing Metadata Quality Interfaces: A Case Study at the University of North Texas Libraries," *Proceedings of the International Conference on Dublin Core and Metadata Applications, 2018* (2018): 20–21, https://doi.org/10.23106/dcmi.952139004.
79. With the exception of the few invalid patterns, UNT's EDTF validation service (https://digital2.library.unt.edu/edtf/) and Python code (https://github.com/unt-libraries/edtf-validate) are highly suggested tools for EDTF verification and validation.

80. Corinne Chatnik and Jame Gaskell, "Quality Control Automation for Student Driven Digitization Workflows," *Code4Lib* 60 (2025), https://journal.code4lib.org/articles/18340.

81. Dana Reijerkerk, Rebecca Fried, and Mandy Mastrovita, "Five Tips for Rapid Metadata Assessment," *Digital Library Federation* (blog), May 15, 2023, www.diglib.org/five-tips-for-rapid-metadata-assessment/.

82. Hannah Tarver and Mark Edward Phillips, "EPIC: A Proposed Model for Approaching Metadata Improvement," in *Metadata and Semantic Research: 14th International Conference, MTSR 2020, Madrid, Spain, December 2–4, 2020, Revised Selected Papers*, ed. Emmanouel Garoufallou and María-Antonia Ovalle-Perandones (Springer Cham, 2021), 228–233, https://doi.org/10.1007/978-3-030-71903-6_22.

83. The Association for Library Collections & Technical Services (ALCTS) is now Core Metadata & Collections section (see https://core.ala.org/sections/).

84. Cataloging Competencies Task Force, *Core Competencies for Cataloging and Metadata Professional Librarians* (American Library Association, January 2017), http://hdl.handle.net/11213/7853.

85. Karen Snow et al., *Core Competencies for Cataloging and Metadata Professional Librarians* (American Library Association, December 2023), http://hdl.handle.net/11213/20799.

86. "Extended Date Time Format: Implementations," Library of Congress, modified October 31, 2023, www.loc.gov/standards/datetime/implementations.html.

87. Sylvester Keil and Ilja Srna, "EDTF-Ruby," GitHub, modified 2023, https://github.com/inukshuk/edtf-ruby.

88. Greta Heng, Patricia Lampron, and Myung-Ja Han, "Linked Data Editors: Where Are We Now, Challenges and Considerations," *Library Resources & Technical Services* 68, no. 3 (2024): 4, https://doi.org/10.5860/lrts.68n3.8226.

89. The term "squishy" for date metadata can relate to meaning and representation. Maggie Dickson, "It's Date Night Here at Digital Projects and Production Services," *Bitstreams: Notes from the Digital Projects Team* (blog), Duke University Libraries, January 29, 2016, https://blogs.library.duke.edu/bitstreams/2016/01/29/3439/.