

# Information Continuity: A Temporal Approach to Assessing Metadata and Organizational Quality in an Institutional Repository

Erik Radio[1]

[1]The University of Kansas, Watson Library, Lawrence, KS

**Abstract:** Repositories provide a vital infrastructure for an institution to aggregate and disseminate creative output, yet this task is only as successful as the effective organization of its content. The University of Kansas is currently undergoing a systematic review to analyze metadata and content organization in its own repository. This paper argues that for a full assessment to be achieved it is necessary to not view the repository as a fixed item, but as an entity with its own continuity. This temporal approach has a significant impact on establishing resource provenance for metadata policy adjustments, disciplinary migration, and resource extensibility. For any repository it is essential for ensuring long-term viability.

**Keywords:** Institutional Repositories, Information Hierarchy, Author Generated Metadata, Temporality

## 1 Introduction

The organization of an institutional repository (IR) has a significant impact on its ability to enable discovery and disseminate resources. Processes to analyze both metadata and information hierarchy are critical to ensure consistency of records and clarity of structure. Repositories can be understood as having both horizontal and vertical dimensions, the former concerning records' conformance to metadata standards across the corpus and the latter to ensure consistency between organizing classes. While examining both dimensions is necessary to any analysis as a step towards identifying needed remediation, they are not entirely sufficient in providing the clearest picture of an IR's structure.

Information continuity is the idea that a resource, be it item, collection, or repository, has an ongoing narrative. This temporal aspect to information systems plays a significant role in providing vital provenance information about resources and their place within an IR. This information also provides an important way to organize content that is consistent with the history of the institution, disciplinary migration, and IR policies. This paper will define this principle and its impact on these critical aspects, and provide examples of how they may be understood by using the University of Kansas' (KU) repository, ScholarWorks as a model. [1] Finally, it will provide possible use cases for implementing information continuity as well as the next steps to be taken in the context of KU ScholarWorks (KUSW).

## 2 Dimensional Aspects of Institutional Repositories

An IR can be described most fully as having three dimensions: horizontal, vertical, and temporal. Briefly, the horizontal aspect can be thought of as a way of assessing resources at the item level with regards to their metadata. Consistency regarding adherence to policies about required elements, content standards, and other similar qualities can be analyzed from this perspective to ensure maximal discoverability of resources across the repository. Ensuring that appropriate semantic units are being employed for the corresponding element is required before any scalable enhancement can be considered.

The vertical dimension of an IR refers to the organizational structures present within the IR in terms of an information hierarchy. For example, the Dspace framework allows for the creation of both communities and collections. In KUSW, this has actually been a source of inconsistency with communities like 'Office of the Provost' being listed alongside 'Mathematics'. While the aim of this paper is not to assess the repository on this level, this is merely provided to show an instance of when classes may not have equivalency and in turn obfuscate the hierarchy of the IR; an unnecessary challenge for users to navigate.

The horizontal and vertical approaches to repository quality should be familiar as they reflect a common, persistent way of viewing information resources. In this case metadata and the organizing structures that contain it are understood as relatively unchanging objects. However, these dimensions provide a misleading notion of the repository and its resources as fixed items. This view is perhaps owing to the time of the physical card catalog in which the metadata record, as a mostly static object, stood as a surrogate for the information resource. When the most atomic metadata in a system is seen as unchanging, so also will its organizing containers. This persistent view fails to take into account not only the capabilities of current technology to track changes to information over time, but the fact that information is not a static entity. Rather, information's ability to change is a critical component of any system. It is important that records may change over time, and that a repository's structure may adapt to accomplish new organizing objectives.

Concern for temporality is required to gain the fullest context for information resources. Significant work has already been achieved in this area by Isabelle Boydens, who begins from the notion that an unequivocal interpretation of 'observable reality' is impossible. [2] As closed world systems, databases, the ubiquitous system for storing what we will for simplicity's sake call 'facts', are generally in opposition to this fundamental idea. Her thesis, neatly distilled by Bade, is that

"...The world changes for us because we ask and expect different things of it at different times; that being so, both our data and our metadata need to change to reflect that new state of the world." [3]

Information stored in a database represents a particular, fixed view of reality. That picture should change with time and by the functional needs of those using it to remain consistent with the changing world around it. An IR is no exception. Being databases with a culturally significant bent, they are acutely tied to temporal changes. Furthermore, there are many ways in which these temporal elements affect IR administrators' ability to manage and enhance metadata. By incorporating elements that capture continuity, administrators will be able to present a clearer, hopefully more useful, bibliographic system to users.

### **3 Temporal Perspectives**

#### **3.1 Policy Adjustments**

Policies for metadata creation are necessary to ensure consistency across records. Before any standardization can occur (e.g. naming conventions), it is necessary to know what guidelines are already in place. However, these policies change by the adoption of new schema, identifiers, content standards, and other similar considerations. Fortunately, the time a record is created can be captured by repository technology, making it a straightforward process to determine which records were created after a specific policy had been put in place. It is then possible to more accurately adjust records to reflect a newer policy.

While this capability is certainly essential to effective record management, it is unfortunately insufficient, especially for an IR. KUSW metadata consists of author generated (AG) records provided at the time of a resource's submission into the repository. While this process has saved cataloging staff from a substantial workload, it has understandably led to creative interpretations of schema elements, causing a

wide level of inconsistency across records. This obviously raises concerns for those interested in providing an effective retrieval system. The remediation process will require substantial work, particularly as it pertains to subject terms.

Some examples of author-generated subject terms include 'parent-child talk', 'agency', 'reading with distraction', and 'future selves' to name only a few. These types of terms, in some cases at once both specific and vague, make up the majority of subjects populating the corpus. There is a clear disparity between these terms and what one would expect to see had they been populated from an authority source such as Library of Congress Subject Headings or another more discipline appropriate vocabulary. From a retrieval perspective they are generally not very useful given their frequency of use; most are employed only once. Indeed, collocation is essentially impossible for those browsing by subjects. From a linked data perspective they are useless as there is no URI available for them to reference.

It was initially suggested that these terms be mapped onto terms from an authority source, an idea that will likely be implemented in the future. But should AG terms then be discarded? As Lubas notes, it is recommended to retain terms from AG metadata, however this does come at the cost of increased 'noise' within a repository, an obvious hindrance to any bibliographic system. [4] While the terms may not be useful in the context of a retrieval system, they remain valuable as artifacts. AG metadata provides a unique prism to view how an author understood their work, with subject terms adding a particularly acute distillation of what terms they thought might also be used by other practitioners.

Preserving AG records raises questions of access and audience. Tracking provenance as it conforms to the Dublin Core element set has been built into Dspace. However, there currently is no mechanism to store meta-metadata or what changes were made and why. Even if such a system existed, such changes would probably only be viewable from an administrator perspective, which while still valuable, does not lend itself to openness. Indeed part of the motivation for saving AG records is so that users would be able to observe these changes, in turn informing their own assessment of the resource and its relevance. A straight normalization of subject terms, or their mapping onto a pre-existing controlled vocabulary eliminates this unique perspective of the author regarding their work. Yet if we continue to think of databases as a snapshot of a particular instance in time, an edited record, while positively adding consistency to the corpus, will come at the cost of provenance or even more precisely, a cultural artifact unique to the institution. Capturing the evolution of metadata allows for the ability to analyze how it changes and understand why it changes, becoming itself an object of critical inquiry.

Transparency, usually referring to openness about the internal workflows and policies a trusted repository follows, should be extended to providing the clearest accounting of a resource's lifespan, including adjustments to its metadata. By retaining an original record and making it accessible to users as an optional view on the same splash page with the new record is an important way of establishing trust through providing additional context for a resource rather than covering it up. For KUSW, since there was little guidance regarding what metadata should look like, this will play an important part for an administrator's ability to understand how and why records changed over time and more effectively monitor the utility of records or lack thereof.

Continuity is important on many different timescales. While the previous example focused on metadata through its entire lifecycle, metadata may also be improved by understanding its current usage. Working from a micro-temporal level, a dynamic prototype interface designed by Boydens and Hooland allows for users to choose which elements to search by in the interface through a drag and drop mechanism, customizing their experience to their specific information seeking behavior. [5] Besides its flexibility for users, it also allows collection managers to prioritize work based on the most frequently used elements and monitor how this changes over time. This demonstrates promising possibility for determining appropriate metadata scope that is informed by temporal considerations. Furthermore it addresses capturing the narrative of the record, in this case, why certain elements are more important at a given time than others. As a basic way of

understanding user behavior, dynamic interaction with metadata will provide insight into the effectiveness of interfaces and inform user experiences decisions in a more rigorous way.

### 3.2 Disciplinary Migration

If policy adjustments over time may be understood as referring to the horizontal dimension of an IR, disciplinary migration represents the vertical. Disciplinary migration is, as its name suggests, when a particular discipline undergoes some sort of rebranding or refocusing to reflect more current appellations consistent with the larger academic community. At KU, what was once the *Human Development and Family Life* department broadened in 2004 to be *Applied Behavioral Science*. Such changes are ubiquitous at academic institutions and their changes can be no small feat to navigate.

These migrations present a difficult challenge for IR with static worldviews. Recently deposited resources stored under the most current appellation of a discipline present no problem, but relevance collisions emerge when one considers aggregating resources under the same collection from when the discipline went by other names. Are the resources consistent enough in subject matter to form a cohesive collection? For more recent changes this is likely true, but for subjects that can trace back several decades, even into other centuries, the similarities grow increasingly thin. Likewise, it may be counterproductive to aggregate resources for disciplines that have undergone disruptive foundational changes, particularly in the physical sciences.

There are a couple of ways in which repositories have sought to mediate this issue. In the University of Illinois at Urbana-Champaign repository, IDEALS, separate top-level collections were created for each iteration of a discipline or department, with dates listed to delineate when the department existed or went by a different name. [6] Within each collection a description was provided indicating the origin of the discipline/department and when it transitioned into something else. From a statically based perspective, this is an acceptable if somewhat less than ideal solution. The tracking of migrations is documented and it avoids the problems of disparity caused by aggregating resources across a century of development under one roof. However, for users navigating the system it creates an even grander domain to traverse. Similarly, relations between iterations of a discipline, while noted, are not built into the infrastructure. There is no attempt to meaningfully connect these relationships on a deeper level.

Another method has been to ignore it as a problem. In KUSW a collection of pre-1923 Theses and Dissertations have been aggregated under the broader collection of Theses and Dissertations that includes those resources from 1923 to the present. While the more recent ETDs are also housed within their respective disciplines, the pre-1923 works can only be found in this one collection. One work, 'Aspects of the Gothic Romance', has a discipline marked unambiguously as 'English' and if desired could easily be mapped into the existing discipline collection of the same name, especially since the content is less likely to be dated.

By contrast, another work, 'Additional Peculiarities in the Spermatogenesis of Phrynotettix Magnus' has for its discipline 'Zoology'. While valuable as representing an important historical perspective on an evolving subject, it has not been joined to its current related discipline 'Biological Sciences'. It is possible that its content has been superseded by new research, but if that is the case, who will go looking for it if it is not stored under what would be its current discipline? For similar examples, separating these resources from their current disciplinary descendent may be a bit of detraction; they appear now as simply relics of a certain time period, perhaps not worthy of consideration alongside more recent research. The solution of aggregating resources in an appropriate collection, even though it may not represent the best, most specific location for it categorically speaking, is fraught with inconsistencies and difficultly resolved content arrangements.

The limits of assessing a repository without a temporal perspective are significant. It may be argued that the methods mentioned above to address disciplinary migration are really more compromises than permanent solutions as some element of the information hierarchy and continuity is lost after implementation. In the context of KUSW, Dspace does not have any built in method of applying these nuanced details within its architecture. Indeed, this is beyond the scope of what a Postgres database can offer on its own. However, linked data provides a possible solution to the challenges of representing information continuity.

As stated above the specter of the physical card catalog continues to influence bibliographic systems. Memory institutions are still strongly tied to the record format, which while useful still suggests a view of the record as an immutable object. Of course current technologies allow for a degree of interoperability that in the light of our discussion can go some way to addressing these temporal concerns, transitioning a record more to data along the spectrum outlined by Glushko and McGrath. [7] In a promising study, Peponakis describes a system for building semantic relationships into an IR, primarily by a combination of RDF informed by FRBR and FRAD principles to create a network of authority files for individuals and departments that can be tied to works. [8] Additionally, a time span was attached to each triple to allow an even more powerful method of coordinated searching, though as Peponakis notes, the implementation of time with RDF has not reached a community consensus despite research into this area. [8, 9] Nevertheless, it remains a unique way of addressing these temporal constraints in an elegant and progressive way. By serializing the history of the departments and disciplines in RDF, one is able to view the relationships between different academic units and their evolution over time through a semantically rich and meaningful way.

As a method of assessing a repository, the continuity of its vertical organization is a critical element with significant implications. To remain with a static view of disciplines, an image of the world at a given time, is to compromise on the meaningful structure of the information hierarchy in addition to creating a system that is at best flawed for navigation and at worst inconsistent and confusing. Initial steps for assessing an IR from this perspective must first to determine what disciplines are represented and if there has been an attempt at documenting the continuity of migrations, then some steps towards integration into a semantic framework may be already in place. If, as in the case of KUSW, the view is of a static present then much of this historical context will need to be discerned before it can be integrated into an authority file and incorporated into a linked environment.

### **3.3 Resource Extensibility**

#### *Format Types*

Digital resource types and the software that create them are in a constant state of development. It goes without saying that ensuring resources remain in a currently accessible format is a natural concern for any repository. Fortunately, the push for open formats is a positive step towards ensuring the longevity of repository resources. [10] However, long term utility of resources is not quite so simple a matter of choosing an initial format. Rather, it is an increasingly nuanced universe, especially as media types continue to change and grow in complexity.

The majority of resources in KUSW are of a text-based nature, consisting of articles, books, project reports, and other similar items. However, there is a slowly growing body of code, data, and videos; formats for which are much more in flux regarding adoption than their text-based counterparts. For example, the resources listed as 'videos' in KUSW often point to embedded videos in web pages. Ignoring the misleading type designation, the resource is subject to a similar problem regarding extensibility. Ensuring that the HTML remains consistent with modern browsers ability to read the page is one concern, but the video in the page,

which currently runs on the Adobe Flash, may be replaced by functionalities built into HTML 5. It is inevitable that over time the current format of the video will become obsolete.

The predominance of text resources may diminish over time, requiring the already needed attention given to preservation of more complex resources and consideration for those that are yet to emerge. Resources in KUSW are generally available in the format they were ingested in. Little effort has been made to upgrade resources to their most current, stable format which fortunately has not been a significant problem as most of the resources are in formats that remain widely used (e.g. PDF, CSV). However, if Python code were to be ingested today, how useful will that resource be to the user who finds it a decade from now and by which time a new version of the language will have been released? Similarly, XML metadata for GIS datasets may need conversion to JSON to support future use.

Evolving resource types are an excellent way of demonstrating how the fixed-world view of databases is itself a contradiction. As a repository ages it will grow to contain several iterations of a particular resource type, becoming an inconsistent disseminator of anachronistic resources. Understandably this is in direct conflict with the definition of a trusted repository put forth by RLG-OCLC, particularly that a trusted repository “design its system(s) in accordance with commonly accepted conventions and standards to ensure the ongoing management, access, and security of materials deposited within it.” [11] Enabling access is a continuous process, not an event, and ensuring that resources can be not only accessed but also used is a part of that process.

It is evident that resource extensibility is a critical part of any repository’s assessment, and particularly that format types, dependencies, and other information needed to enable its use be documented in the preservation metadata. Provided this is accomplished, we may return to the horizontal dimension of the repository and be able to observe that after a certain date, for example, a policy regarding accepted file types was implemented and that previous types were migrated to a more current format. The question of retaining those older versions and how metadata records should reflect their replacements goes beyond the scope of this paper, but it is mentioned now as a possible source of future discussion.

### *Preservation Narratives*

The use of PREMIS for documenting the technical components of a resource at a very granular level is widespread. Given the limitations of other schema to record information of this quality in any sort of all encompassing way, its existence is a needed tool for managers of repositories. However, the claim that PREMIS data dictionary contains all the information required to ensure preservation has been rightly questioned by Wilson, who notes that the dictionary does not capture anything related to the preservation process, only the resources, events, and related agents. [12] Additional provenance information, he continues, is required to gain the clearest picture of a resource’s lifespan beginning from before its ingest and into the future.

Wilson hints at the temporal element by his mention of preservation as a process. Defined by the PREMIS data dictionary, a *representation* is ‘a single digital instance of an intellectual entity held in a preservation repository’. [13] This concept seems to exist somewhere between a FRBR manifestation and item. It uses as an example a TIFF image of a statue of a horse as one representation, while a derived JPEG image represents yet another, both being representations of the statue image. The ability to track events and relationships between representations and agents goes some way towards introducing a temporal element, but it is not sufficient. We only know that preservation events happened, but not necessarily why they happened. While the <eventOutcomeDetailNote> element can capture some of this information, it reflects only a small part of the ongoing preservation process.

The contextual element is essential to the notion of information continuity. Without the proper documentation for why certain processes were implemented, we are left with guesswork. But if this information is not to be fully captured in PREMIS, where can it be? While METS does allow for some of this

more administrative information to be recorded, it is neither exactly what is needed nor on a scale appropriate to the myriad iterations of digital resources that may populate a system; the structure of a METS record would quickly be overwhelmed. Of course some degree of rational can usually be inferred from a PREMIS event, but leaving these deductions to future stewards when more should be captured at the time of the event is hardly a favor.

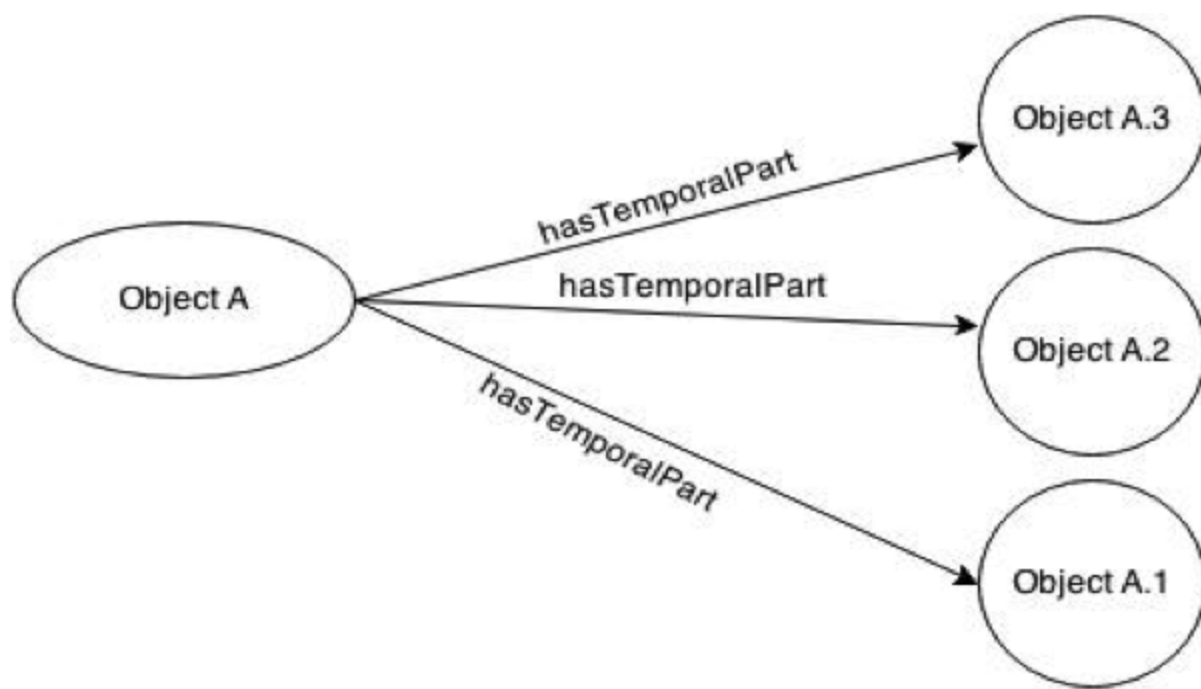
It is recognized that to some extent we are reaching the limits of what can be expressed by metadata, at least currently. Again, while events recorded by PREMIS can be given additional details, the ability to form a narrative outlining these changes is really what is required to fully reflect information continuity. How these discrete metadata units can be rebuilt to form a larger picture remains to be seen, but its achievement is required for future administrators to have the most complete picture of the preservation continuum and how it informs their own work. For KUSW, which is only in its initial stages of implementing a preservation system, care will have to be given to these concerns to create an effective and understandable paper trail from the outset to ensure it is achieved.

## **4 Temporal Modeling and Implementation**

### **4.1 Temporal Parts**

Any consideration for modeling how resources exist in time will inevitably require a decision between two trains of thought. The perdurantist view posits that entities have both a spatial and temporal parts, in essence that an institution does not just have its present version, but has in equally important measure its version from yesterday and tomorrow; in essence, a sum of parts. The opposing view, endurantism, says that an entity is wholly present whenever it exists, that parts do not extend beyond the current moment. [14] While not delving into the nuances behind each school of thought, each has significant implications for temporal representation of an IR. Information continuity is most closely aligned with perdurantism as it has been argued that the narrative of a metadata element, document, and institution is important.

Before any implementation of a system can occur it is necessary to model what exactly comprises a perdurant object. Despite the fundamental differences between a resource and an entity in an IR there are similarities in the way they are viewed from temporal perspectives. Figure 1 represents how an admittedly oversimplified perdurant view might represent an object (e.g. article, department):



**Fig. 1.** An object with its perdurant components

The prime object itself is not identical with any one of its iterations, but rather is the sum of the three. While Object A.3 may represent the one currently accessible in a repository, it does not represent the totality of the object. Rather, Object A, an intangible resource, might be considered the prime object, and is not unlike a *Work* in the FRBR sense. This correlation is only meant to illustrate and not to incorporate another model in any systematic way.

It may be judged that temporal parts are just a matter of versioning, and while the similarities are obvious, the primary difference is that continuity is more interested in the matter of the objects as aggregates; one does not supersede another. Indeed, that is the critical element of continuity. That one may succeed another does not change the fact that all are necessary to form the narrative of the object.

## 4.2. Implementation

It is clear that semantic solutions are required to facilitate the meaningful relationships between resources built around temporal considerations. Otherwise one runs into the same compromises that we have observed in the examples of organizing an IR by discipline. Yet, as noted above, there is no common consensus on how time is to be accounted for within the context of RDF despite research into the area. [9] More recent approaches include the development of a temporal Web Ontology Language (tOWL) though its widespread adoption remains to be seen and it seems likely that its proposed serialization in RDF/XML will pose the same challenges that have encumbered that format so far. [15] Concerns for interoperability are of course critical but beyond the scope of this paper.

This initial exploration seeks to draw on existing models to avoid the proliferation of alternatives among an already competitive ecosystem. RDF is the natural tool for creating these relationships and using JSON-LD will likely provide a more lightweight and flexible method for preliminary investigations into the feasibility of these aggregations. Regarding an appropriate aggregating standard, the most closely related model currently available is that of OAI-ORE [16]; Fig. 1. Is derived from a part of this model. Though the goal of the aggregations are different with OAI-ORE designed to collect resources from various content



providers, the specification provides a broadly analogous model for how various perdurant iterations could likewise be collected. [16] To build in a temporal element one may draw on the W3's OWL-Time, which allows for time to be incorporated into an object's metadata as 'instants' and intervals'. [17] There are fundamental differences regarding representations in OAI-ORE that will need to be reconciled with the non-tangible prime object of the perdurant model. This brief discussion has only been meant as a cursory outline of what potential frameworks may be employed towards a serialization of IR resources and structures. It is implied that a more detailed model and analysis of its implementation will be the immediate focus of future research.

## 5 Conclusions

This paper has outlined that the idea of information continuity is essential to an effective assessment and organization of any information system. By outlining three areas most immediately impacted by temporal concerns, a framework for approaching the needs of a repository can be used and a more thorough analysis achieved. Future work will demonstrate how effectively perceived temporal needs can be implemented in the context of KUSW, with results serving as a model for future development.

While this paper has pointed to potential solutions for each of three areas described, they by no means represent the best or only approaches to consider. As regards preservation metadata there is significant room for research on how to best incorporate the already present sequential aspects into a more meaningful relational context. For archival metadata additional effort is needed to address the final role and perceived utility of author generated records. Finally, the disposition of certain schemas towards capturing important temporal information will provide an important outlet for a wider assessment of schema utility in general.

Finally, the perdurant model outlined above is only an initial step in using existing tools and standards to build a flexible relational model. Since there is yet no standard, allowing for flexibility in proposed models is necessary to ensure responsive adaptability. Further research will involve using the OAI-ORE framework within the context of an IR collection to assess its appropriateness and needed extensions.

Metadata needs to be more flexible. As only one part of the information environment that is a repository, it in some ways can be considered a microcosm of its host structure. If the metadata is considered a static resource, then the organizing structures containing it will follow suit, as there is no temporal element to be considered and built upon. If metadata begins with a concern for capturing temporality and in a way that meaningfully impacts larger structures, a greater symbiosis between resource and repository can be achieved and information continuity more effectively represented.

Information, like the world, is affected by temporality. Monitoring how metadata is used over time and in certain contexts naturally allows for a greater understanding of its effectiveness and how it must change to reflect new information seeking behaviors, social responses, and networked environments. That metadata is an indicator for resources and not just a pointer is a vital way of assessing the narrative of resources themselves. As Boyden notes, this provides a new direction for a type of metadata that is more interactive and responsive to the world around it. [2]

## References

1. The University of Kansas, ScholarWorks, <http://kuscholarworks.ku.edu>
2. Boydens, I. *Informatiques, normes et temps*. Bruxelles: Bruylant. (1999)
3. Bade, D. It's about Time!: Temporal Aspects of Metadata Management in the Work of Isabelle Boydens. *Cataloging & Classification Quarterly*, 49(4), 328–338 (2011) doi:10.1080/01639374.2011.571096
4. Lubas, R. L. Defining Best Practices in Electronic Thesis and Dissertation Metadata. *Journal of Library Metadata*, 9(3-4), 252–263 (2009) doi:10.1080/19386380903405165
5. Boydens, I., & Hooland, S. Van. Hermeneutics applied to the quality of empirical databases. *Journal of Documentation*, 67(2), 279–289 (2011) doi:10.1108/00220411111109476
6. The University of Illinois at Urbana-Champaign, IDEALS, <https://ideals.illinois.edu/>
7. Glushko, R. & McGrath, T. *Document Engineering*. Cambridge, MA: MIT Press. (2005).
8. Peponakis, M. Libraries' Metadata as Data in the Era of the Semantic Web: Modeling a Repository of Master Theses and PhD Dissertations for the Web of Data. *Journal of Library Metadata*, 13(4), 330–348 (2013) doi:10.1080/19386389.2013.846618
9. Gutierrez, C., Hurtado, C. & A. Vaisman, S. Introducing time into RDF. *IEEE Transactions on Knowledge and Data Engineering*, 19(2), 207–218 (2007)
10. Rimkus, K., Padilla, T., Popp, T., & Martin, G. Digital Preservation File Format Policies of ARL Member Libraries: An Analysis. *D-Lib Magazine*, 20(3-4) (2014) doi: 10.1045/march2014-rimkus
11. RLG-OCLC. *Trusted Digital Repositories : Attributes and Responsibilities*. Mountain View CA. (2002)
12. Wilson, A. How Much Is Enough: Metadata for Preserving Digital Data. *Journal of Library Metadata*, 10(2-3), 205–217 (2010) doi:10.1080/19386389.2010.506395
13. PREMIS Editorial Committee. *Premis Data Dictionary for Preservation Metadata, version 2.2*. Washington, D.C. (2012)
14. Hawley, K. Temporal Parts. *The Stanford Encyclopedia of Philosophy*. Edward N. Zalta (ed.), (Winter 2010 Edition), <http://plato.stanford.edu/archives/win2010/entries/temporal-parts/>
15. Milea, V., Frasincar, F., & Kaymak, U. tOWL: A Temporal Web Ontology Language. *IEEE Transactions on Systems, Man, and Cybernetics*. 42(1), 268-281. (2012)
16. Open Archives Initiative-Object Reuse and Exchange, <http://www.openarchives.org/ore/>
17. OWL-Time, <http://www.w3.org/2006/time>