

Invisible,  
Not Viable, Venal, and Now Vibrant:

Metadata Research throughout InterPARES

---

# Metadata Research throughout InterPARES

---

# Metadata Research throughout InterPARES

---

*For the translators:*

Record = Documento archivístico. *Véase:* documento de archive

*For Twitter:*

@josephttennis

# Metadata Research throughout InterPARES

---

# Metadata Research throughout InterPARES

InterPARES 1 – 1999-2002

InterPARES 2 – 2002-2007

InterPARES 3 – 2007-2012

InterPARES Trust – 2013-2018

## Metadata Research throughout InterPARES

---

InterPARES 1 – 1999-2002 – Invisible

InterPARES 2 – 2002-2007 – Not Viable

InterPARES 3 – 2007-2012 – Venal

InterPARES Trust – 2013-2018 – Vibrant

# Invisible



## Invisible

---

From the work of the Authenticity Task Force, starting in 1999, metadata has played various roles in InterPARES. The publication of the 2002 Report from the Authenticity Task Force [1], and specifically the work done on the Benchmark and Baseline Requirements for maintaining authentic records in electronic systems, provided us with the seeds of powerful metadata specifications.



# Invisible – Benchmark and Baseline Requirements

## Benchmark Requirements (2002)

“The benchmark requirements are the conditions that serve as a basis for the preserver’s assessment of the authenticity of the creator’s electronic records. Satisfaction of these benchmark requirements will enable the preserver to infer a record’s authenticity on the basis of the manner in which the records have been created, handled, and maintained by the creator,” [1 p 4]

# Invisible – Benchmark and Baseline Requirements

## Benchmark Requirements (2002)

Req. A1 – Expression of Record Attributes and Linkage to Record

Req. A2 – Access Privileges

Req. A3 – Protective Procedures: Loss and Corruption of Records

Req. A4 – Protective Procedures: Media and Technology

Req. A5 – Establishment of Documentary Forms

Req. A6 – Authentication of Records

Req. A7 – Identification of Authoritative Record

Req. A8 – Removal and Transfer of Relevant Documentation

# Invisible – Benchmark and Baseline Requirements

## Benchmark Requirements (2002)

Req. A1 – Expression of Record Attributes and Linkage to Record

Req. A2 – Access Privileges

Req. A3 – Protective Procedures: Loss and Corruption of Records

Req. A4 – Protective Procedures: Media and Technology

Req. A5 – Establishment of Documentary Forms

Req. A6 – Authentication of Records

Req. A7 – Identification of Authoritative Record

Req. A8 – Removal and Transfer of Relevant Documentation

# Invisible – Benchmark and Baseline Requirements

## Benchmark Requirements (2002)

### Req. A1 – Expression of Record Attributes and Linkage to Record

#### *Identity of the Record*

Names of Persons (author, writer, originator, addressee, recipient\*)

Action or Matter

Dates (chronological date, received date, archival date, transmission date)

Expression of Archival Bond (classification code, file identifier)

Indication of Attachments

# Invisible – Benchmark and Baseline Requirements

## Benchmark Requirements (2002)

### Req. A1 – Expression of Record Attributes and Linkage to Record

#### *Integrity of the Record*

Name of handling office

Name of office of primary responsibility

Indication of types of annotations added to the record

Indication of technical modifications

# Invisible – Benchmark and Baseline Requirements

## Benchmark Requirements (2002)

### Req. A1 – Expression of Record Attributes and Linkage to Record

“The case studies undertaken as part of the work of the task force revealed very little consistency in the way the attributes that specifically establish the identity of a record are captured and expressed from one electronic system to another. In certain systems, some attributes were explicitly mentioned on the face of the record; in others they could be found in a wide range of **metadata** linked to the record ... The task force’s concern is that, in the absence of a precise and explicit statement of the **basic facts** concerning a record’s identity and integrity, it will be necessary for the preserver to acquire enormous, and otherwise unnecessary, quantities of data and documentation simply to establish those facts,” [1 p 8].

# Invisible – Benchmark and Baseline Requirements

## Benchmark Requirements (2002)

In this context **metadata** is something generated by a system ignorant of archival purpose. It is, at this time, not an archival concept, but rather a concept imported into the field, not visible in the field of archival diplomatics on which the Benchmark and Baseline Requirements are built.

This will change as digital records systems evolve.

# Invisible – Benchmark and Baseline Requirements

---

## Baseline Requirements



# Invisible – Benchmark and Baseline Requirements

## Baseline Requirements (2002)

“The baseline requirements outline the minimum conditions necessary to enable the preserver to attest to the authenticity of copies of inactive electronic records,” [1 p. 7].

# Invisible – Benchmark and Baseline Requirements

## Baseline Requirements (2002)

Req. B1 – Controls over Records Transfer, Maintenance, and  
Reproduction

Req. B2 – Documentation of Reproduction Processes and Effects

Req. B3 – Archival Description

# Invisible – Benchmark and Baseline Requirements

## Baseline Requirements (2002)

“Traditionally it has been a function of **archival description** to authenticate the records and perpetuate their administrative and documentary relationships. With electronic records, this function becomes critical. Once the records no longer exist except as authentic copies, the archival description is the primary source of information about the history of the record, that is, its various reproductions and the changes to the record that have resulted from them,” [1 p. 11].

## Invisible – Benchmark and Baseline Requirements

---

It is this orthodox archival activity of **archival description** that serves as a forcing function in the next conversations about metadata – specifically how it is not viable as an alternative.

One final note. The term *metadata* does not appear in the InterPARES 1 glossary. This underlines the fact that it had not yet been seen as a term of art for the purposes of InterPARES research [3].

# Not Viable



# Not Viable: 1 Metadata vs. Archival Description

---

## Not Viable: 1 Metadata vs. Archival Description

In 1995 Heather MacNeil separated out archival description from *metadata strategies* [2].

Her paper outlines the rationale of of archival description and the limitations of contemporary metadata systems. This view, that metadata is not a viable alternative to archival description is received by most InterPARES researchers as correct during InterPARES 2 (2002-2007).

## Not Viable: 1 Metadata vs. Archival Description

---

A notable defector from this orthodox view was Gilliland who is listed as lead on the Description Cross Domain report, [4].

This report documents the successes and confusion over the role of metadata in preserving authentic records. After 5 years of work this Report closes with a set of questions that connects the second wave of InterPARES to MacNeil:



## Not Viable: 1 Metadata vs. Archival Description

---

“Can metadata associated with the creation and active use of records ever contribute to archival description, particularly in the capture and elucidation of certain kinds of context and fundamental identification and arrangement information relating to the records,” [4]?

So as of 2006, metadata was not yet viable as a replacement of or even as a source for archival description.

## Not Viable: 2 Models and Metadata

---

## Not Viable: 2 Models and Metadata

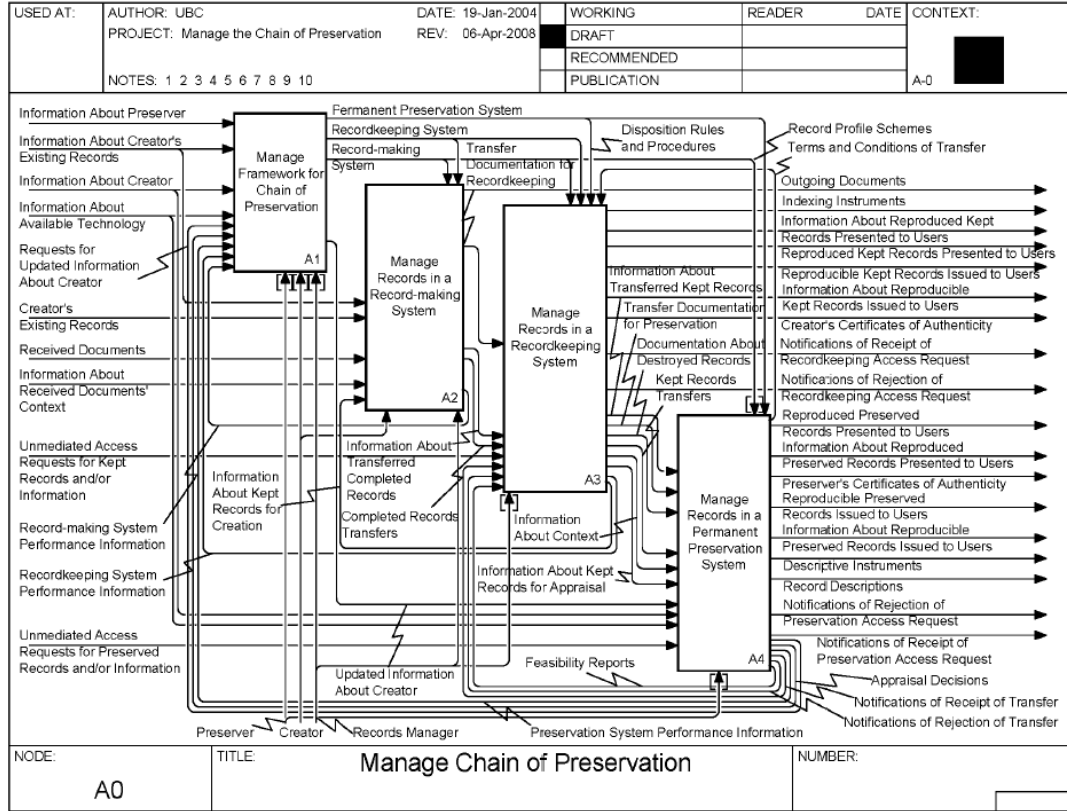
---

A second problem with the viability of metadata in service of preserving authentic digital records is the question of **how much metadata do we need?**

With the creation of the **Chain of Preservation Model** there grew an interest in identifying where metadata would be created from records creation, maintenance, past the threshold of preservation. This resulted in the **IP2 Metadata Specification Model**.

## Not Viable: 2 Models and Metadata

- Box A2 = Manage Records in Record-making System
- Box A3 = Manage Records in a Recordkeeping system
- Box A4 = Manage Records in a Permanent Preservation System

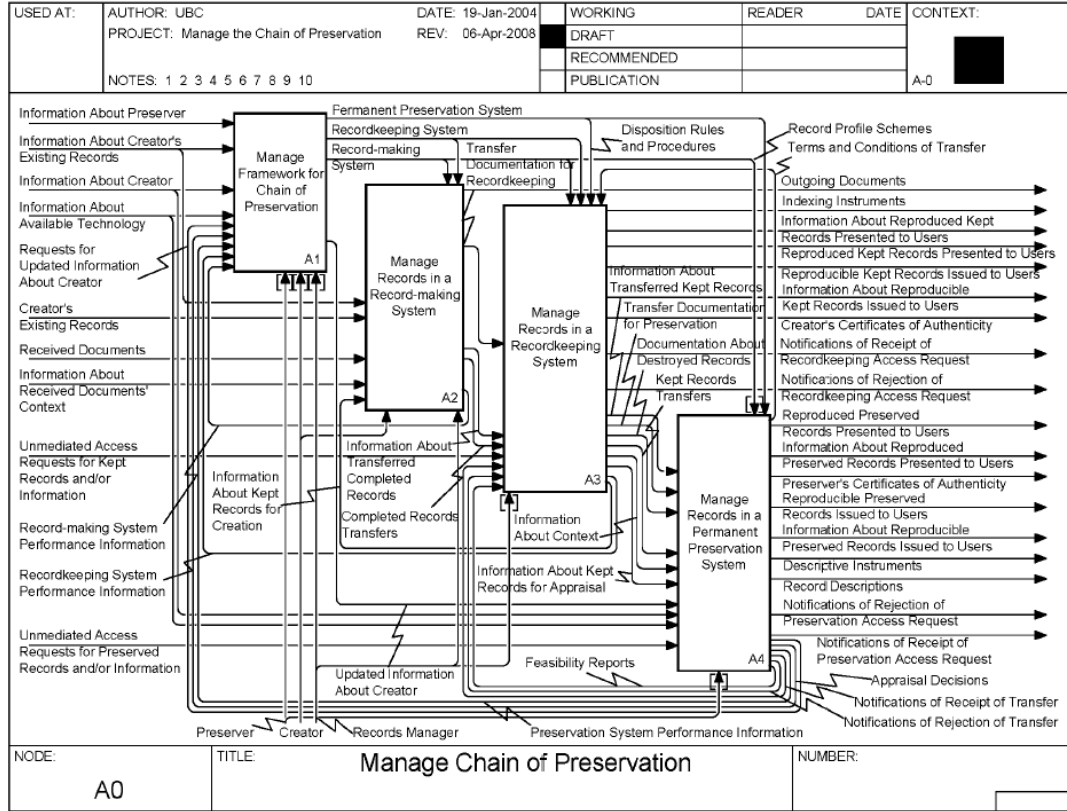


# Not Viable: 2 Models and Metadata

From this we derived an initial set of 137 unique metadata assertions.

The viability of this was questioned immediately. *This is too much metadata!*

This is where **venality** creeps in.



## Not Viable:

---

Again, in the InterPARES 2 Glossary none of the definitions from archival science are used. They are documented from MoReq, SAA, and the ISO standard, but are not sanctioned for use in InterPARES documentation this point. Still not viable.

# Venal

---



Information School  
UNIVERSITY of WASHINGTON



# Venal: Persistent Uselessness

---



## Venal: Persistent Uselessness

---

At the beginning of InterPARES 3 we set about an attempt to formalize the **Metadata Specification Model (MSM)** into Dublin Core Metadata Initiative standard application profile [4, 5].

The rationale for this was, if the metadata were both human and machine readable we could see how practitioners would adopt the **MSM**.

However, the presence of metadata discussions were not always welcome. A subset of the InterPARES 3 researchers saw metadata work as busy-work, not amounting to solutions of on-the-ground archivists and records managers.

## Venal: Persistent Uselessness

---

The exercise seemed purely academic to some, and those of us championing the **InterPARES Authenticity Metadata Application Profile (IPAM AP)** were clearly trying to persuade the others, perhaps not with bribery, but whatever means necessary. We felt that we needed to bridge the methodological and technological divide that was now more than ten years old in the project. Metadata should be an archival term as well.

So perhaps *venality* is not exactly right here, but the value of metadata did not circulate through InterPARES 3 without advocates paying with social and intellectual capital to win their case.

## Venal: Persistent Uselessness

---

We also, for the first time, combined the purpose of archival descriptive work and diplomatics with the design of metadata. This was fully supported by the work from the **Metadata Specification Model**, drawn from the **Chain of Preservation Model**.

Our metadata requirements for this phase of InterPARES were “should be necessary and sufficient to support the presumption of authenticity of records, interoperate between systems and across time, be adequate for archival description, and be useful for both retrieval and meaningful display of records,” [5 p4].

## Venal: Persistent Uselessness

---

We also, for the first time, combined the purpose of archival descriptive work and diplomatics with the design of metadata. This was fully supported by the work from the **Metadata Specification Model**, drawn from the **Chain of Preservation Model**.

Our metadata requirements for this phase of InterPARES were “should be necessary and sufficient to support the presumption of authenticity of records, interoperate between systems and across time, **be adequate for archival description**, and be useful for both retrieval and meaningful display of records,” [5 p4].

## Venal: Persistent Uselessness

---

At the end of InterPARES 3 we felt we had grown to understand the contours of concerns metadata had presented from 1999-2012.

1. Metadata was an imported concept and tool
2. It would take work, but we had the ability to use or shape metadata for the purposes of archival description
3. It is possible, based on our modeling of archival processes to wed metadata to the work of archivists and records managers

## Venal: Persistent Uselessness

---

But our understanding as of 2012 was there was too much metadata to manage.

What was left, was to identify the minimal set drawn from fieldwork based on our ideal understanding of archival processes and the metadata that can be drawn from them.

# Vibrant



## Vibrant

---

That was our task in InterPARES Trust.

We examined the requirements of working archivists and records managers to see what metadata they were using to presume they were managing or preserving authentic records.





## Vibrant

---

It comes as no shock to us, that because of current nature of the work of archivists and records managers, that while identity is important, integrity overrides that concern for presuming authenticity of digital records.

If the system is intact and legible, then we presume, if we are the preservers, that the creators handed us all the metadata they care about. And that this is the minimal set.

## Vibrant

---

The concerns about identity metadata (persons, actions and matters, etc.) are vibrantly present in the body of records moved from active systems to permanent preservation systems. However, they are not called out in our current practices and placed discretely in fields in databases.

We can, however, **audit** them at any time against our idealized set pulled from the Metadata Specification Model and the InterPARES Authenticity Metadata Application Profile (IPAM AP).

## Vibrant

---

Further, without the sewing together of archival concerns and investigations into metadata, we would not have the full robust analytical tools of digital diplomatics [6] and digital forensics [7].

And this brings me back to the Benchmark and Baseline Requirements [1].

# Vibrant – Benchmark and Baseline Requirements

## Benchmark Requirements (2002)

Req. A1 – Expression of Record Attributes and Linkage to Record

Req. A2 – Access Privileges

Req. A3 – Protective Procedures: Loss and Corruption of Records

Req. A4 – Protective Procedures: Media and Technology

Req. A5 – Establishment of Documentary Forms – Metadata and Form

Req. A6 – Authentication of Records

Req. A7 – Identification of Authoritative Record

Req. A8 – Removal and Transfer of Relevant Documentation

# Vibrant – Benchmark and Baseline Requirements

Benchmark Requirements (2002)

*Req. A5 – Establishment of Documentary Forms – Metadata and Form*

It is in this Requirement A5 that we return to metadata for the purposes of digital diplomacy and digital forensics.

# Metadata Research throughout InterPARES

---

## Metadata Research throughout InterPARES

---

We have learned a lot from incorporating metadata into the theory and conceptualizations of InterPARES.

I have tried to point out different perspectives on metadata research over four generations of InterPARES research from 1999 to the present. But in so doing I hope to have pointed out how we have learned that we are now in a *vibrant* stage of metadata inquiry.

# Metadata Research throughout InterPARES

---

And, perhaps surprisingly, I mean vibrant as in *shaking to and fro*.

*Latin: Vibrare.*

We have now a practice that under-specifies identity metadata. But needs metadata to understand documentary form in the context of records creation.

And we expect to be able to audit the integrity of our systems.



## Metadata Research throughout InterPARES

---

These interlocking parts of our investigation into requirements for preserving authentic records moves us back and forth across various units of analysis, and demands that we broaden our view of metadata.

Though “data about data” is a broad definition of metadata, we have a more full understanding, thanks to InterPARES research as to what these “data” are.

## Metadata Research throughout InterPARES

---

We now, in 2020, understand we need to know that the creator's system has not been tampered with. We need the full range of integrity metadata which includes systems upgrades, file conversions (if any), and server migration or cloud storage information.

We need to know the intrinsic and extrinsic forms of digital documents.

And we can link contemporary identification practices to diplomatic theory.

## Metadata Research throughout InterPARES

---

Our data about data are no longer simply generated by accident from poorly designed recordkeeping systems, but a range of information about the handling, structure, and perhaps the identity of records.

## References

---

1. MacNeil et al., (2002). “Appendix 2: Requirements for Assessing and Maintaining the Authenticity of Electronic Records.” In Mac Neil et al., *Authenticity Task for Report*. Available: [http://www.interpares.org/display\\_file.cfm?doc=interpares\\_book\\_k\\_app02.pdf](http://www.interpares.org/display_file.cfm?doc=interpares_book_k_app02.pdf)
2. MacNeil, H. (1995). “Metadata strategies and archival description: Comparing apples and oranges.”
3. InterPARES (2001). The InterPARES Glossary. Available: [http://www.interpares.org/display\\_file.cfm?doc=ip1\\_glossary.pdf](http://www.interpares.org/display_file.cfm?doc=ip1_glossary.pdf)
4. Gilliland, A. et al., (2006). Investigating The Roles and Requirements, Manifestations and Management Of Metadata in The Creation of Reliable And Preservation of Authentic Electronic Entities: Report on The Work and Findings of The InterPARES 2 Description Cross Domain Group.
5. Rogers, C. and Tennis, J. T. (2012). General Study 15 – Application Profile for Authenticity Metadata. Available:
6. Rogers, C. (2015). “Diplomatics of born digital documents – considering documentary form in a digital environment.” In *Records Management Journal* 25(1): 6-20.
7. Duranti, L. (2009). “From Digital Diplomatics to Digital Records Forensics.” In *Archivaria* 68 (Fall): 39-66.

# Invisible, Not Viable, Venal, and Now Vibrant:

---

Thank you.

Joseph T. Tennis

Google scholar: jt tennis