NISO-RP-8-2008

Journal Article Versions (JAV): Recommendations of the NISO/ALPSP JAV Technical Working Group

April 2008

A Recommended Practice of the
National Information Standards Organization
in partnership with the
Association of Learned and Professional Society Publishers

Prepared by the
NISO/ALPSP Journal Article Versions (JAV) Technical Working Group

About NISO Recommended Practices

A NISO Recommended Practice is a recommended "best practice" or "guideline" for methods, materials, or practices in order to give guidance to the user. Such documents usually represent a leading edge, exceptional model, or proven industry practice. All elements of Recommended Practices are discretionary and may be used as stated or modified by the user to meet specific needs.

This recommended practice may be revised or withdrawn at any time. For current information on the status of this publication contact the NISO office or visit the NISO website (www.niso.org).

Published by
National Information Standards Organization (NISO)
One North Charles Street, Suite 1905
Baltimore, MD 21201
www.niso.org

Printed in the United States of America
ISBN (10): 1-880124-79-3
ISBN (13): 978-1-880124-79-6

Table of Contents

Participants

NISO/ALPSP JAV Technical Working Group Members

The following individuals served on the NISO/ALPSP Journal Article Versions (JAV) Technical Working Group, which developed and approved this Recommended Practice. No endorsement by organizational affiliations is implied.

Beverley Acreman
Taylor and Francis

Claire Bird
Oxford University Press

Catherine Jones
STFC (formerly CCLRC)

Peter McCracken
Serials Solutions

Cliff Morgan (Chair)
John Wiley & Sons

John Ober
California Digital Library (CDL)

Evan Owens
Portico

T. Scott Plutchak
University of Alabama at Birmingham

Bernie Rous (Interim Chair)
ACM

Andrew Wray
IOP Publishing

NISO/ALPSP JAV Review Group Members

In addition, the NISO/ALPSP JAV Technical Working Group would like to thank the following members of the Journal Article Versions Review Group for their helpful feedback:

Helen Atkins
HighWire Press

Lindi Belfield
Science Direct (Elsevier)

Rachel Bruce (Alternate)
Joint Information Systems Committee

Emily Dill
Indiana University

Richard Fidczuk
Sage

Fred Friend
University College London

David Goodman
Long Island University

Toby Green
OECD Publishing

Janet Halsall
CABI Publishing

Chuck Koscher
CrossRef

Ted Koppel
Auto-Graphics (formerly Ex Libris USA)

Barbara Meredith
Association of American Publishers

Cliff Morgan (Chair)
John Wiley & Sons

Norman Paskin
International DOI Foundation

Erik Oltmans
Koninklijke Bibliotheek

Jan Peterson
Infotrieve

Heather Reid
Copyright Clearance Center

Nathan Robertson
University of Maryland, Thurgood Marshall Law Library

Bruce Rosenblum
Inera

Sarah Rosenblum
London School of Economics

continued on next page

NISO/ALPSP JAV Review Group (continued)

Ian Russell
Association of Learned and Professional Society Publishers (ALPSP)

Rebecca Simon
University of California Press

Gavin Swanson
Cambridge University Press

Peter Suber
Earlham College

Anthony Watkinson
Consultant

Candy Zemon
Polaris Library Systems

NISO Topic Committee Members

The Content and Collection Management (CCM) Topic Committee, which provided oversight to the NISO/ALPSP Journal Article Versions Technical Working Group, had the following members at the time it formally approved this Recommended Practice:

Julia Blixrud
Association of Research Libraries (ARL)

Kevin Cohn
Atypon Systems, Inc.

Juha Hakala
The National Library of Finland

Diane Hillmann
Cornell University Library

Ted Koppel (Chair)
Auto-Graphics, Inc.

Katherine Kott
Digital Library Federation (DLF)

Bonnie Lawlor
National Federation of Advanced Information Services (NFAIS)

Denise Troll Covey
Carnegie Mellon University Libraries

Rollo Turner
Association of Subscription Agents (ASA)

Introduction

These NISO/ALPSP Journal Article Versions (JAV) Technical Working Group recommendations provide a simple, practical way of describing the versions of scholarly journal articles that typically appear online before, during, and after formal journal publication.

Researchers, their institutions, and journal publishers are rapidly moving on from using static, single copies of research papers that are essentially "images" of a printed document. Changes in the way we create, edit, circulate, validate, publish, find, use, and update articles are producing multiple versions whose status and provenance is often unclear. Online searching now allows multiple versions to be found but rarely makes clear the relationships between them.

This is not simply a problem for the journals themselves. Repositories want to provide authoritative records of their faculty's work; libraries want to offer "appropriate copies" to different users; readers need to know what has been peer reviewed; and authors may wish to update their work and ensure that the latest version is used.

In September 2005, NISO (the National Information Standards Organization) launched a partnership with ALPSP (the Association of Learned and Professional Society Publishers) to bring together experts from the publishing, library, library systems, and user communities to examine the problems associated with the proliferation of different article versions. A Technical Working Group produced the recommendations in this report, with feedback and advice provided by a wider Review Group (members are listed above).

The recommendations have been agreed by all members of the Technical Working Group. The Review Group provided many points of feedback, advice, and criticism. This process was very valuable and led to improvements throughout the document. It was not possible to obtain a consensus from all members on all issues. However, all the proposals, comments, and responses are available to read on the NISO website and the comments received during the formal review period are incorporated here in Appendix 3 together with the responses to them.

The Technical Working Group created use cases to explore the lifecycle of journal articles, starting from a base case that describes a typical interaction between author, institutional repository, and publisher. Analyzing these led us to identify common lifecycle stages, the dimensions that describe the evolution of articles, and possible attributes of each instance of an article version.

Considering attributes helped us to agree upon the most important variables for describing versions. Ownership, bibliographic context, identifiers (e.g., DOI), relationships, fixity, and peer review are explicitly stated in the terms and definitions that we recommend. They can be described by the article version names and some are already covered by standard metadata elements (e.g., bibliographic reference, date, DOI). Other variables describe digital copies or variants, but are less closely related to the academic content of the article: visibility (limited, general), version status (known/unknown), source (different websites), scope (text, fully-featured, resolution of images, etc.), delivery context, and format (PDF, HTML, etc.). These are important factors and could be described in article metadata; however, in most cases it is

possible to distinguish properties inherent to the article version and properties derived from the delivery system. Our focus has been to describe which version of the intellectual content of the article a reader has encountered.

These variables led us to identify important dimensions for an article version:

- Time: from first draft to latest version
- Added value: from rough draft to polished publication
- Manifestation/Rendition: different document formats and layouts
- Siblings: multiple mappings between technical reports, conference papers, lectures, journal articles, review articles, etc.
- Stakeholders: author, editor, referee, publisher, librarian, reader, funder

Manifestation was largely excluded from the nomenclature as being too dependent upon technology changes and the creator or user's computer systems (though it is relevant for describing our Enhanced Version of Record). Siblings were discussed at length and eventually excluded, as the relationships are complex and conventions vary considerably between disciplines. Our recommendations could be used to describe, for example, a technical report validated by a funding agency and published in an institutional repository, but we stuck strictly to our brief of considering only journal articles as a first goal.

The variables, dimensions, and use cases reveal the difficulty of describing all possible lifecycle stages in clear terminology. However, the vast majority of journal articles do pass through the same milestones and lie within a limited range of use cases. This remains true even though different users will value different versions for different purposes. From these common milestones we have produced six terms to describe journal article versions. The variety of activity illustrated in the 26 use cases (Appendix 2) shows just how much can be described by these high-level semantic terms. We have focused on changes to content, the formal social process of validation, and the ways in which journal articles are used. We have also noted the tradition that journal articles record the "minutes of science" and are intended as a fixed record of a body of work at a moment in time chosen by the scholar. This leads us to the Version of Record as a useful definition for formalizing academic achievement, distributing authoritative information, and building upon the established knowledge in a field.

In line with the JAV Technical Working Group's charge, the group submits the following:

- Background and Rationale: a narrative that explains the background to our project and the rationale for our recommended terms and definitions
- Recommended Terms and Definitions for Journal Article Versions
- Appendix 1 – Graphical Representation of Journal Article Versions and Relationships with Formal and Gray Literature; Assumptions, Primary Challenges, and Best Practices
- Appendix 2 – Use Cases: a set of use cases showing the application of these recommended terms

- Appendix 3 – JAV Review Group Comments: comments received from the JAV Review Group to an earlier Technical Working Group document submission, and the Technical Working Group's responses

We propose that the terms as defined be promulgated by NISO/ALPSP to the full journal article stakeholder community (authors, readers, libraries, publishers, aggregators, archives, repositories, research institutions, funding agencies, and service providers such as search engines and link resolvers). The JAVTWG recognizes that adopting a standard terminology will not be enough; to avoid version confusion, terminology needs to be implemented in such a way that readers (whether human or machine) encountering any version can immediately ascertain which it is and know whether it is trustworthy.

Background and Rationale

The NISO/ALPSP Journal Article Versions (JAV) Technical Working Group work plan was as follows:

- Creation of use cases to identify the most common journal article lifecycles.
- Analysis of use cases to determine common lifecycle stages.
- Selection of preferred vocabulary for the most common lifecycle stages.
- Development of appropriate metadata to identify each variant version and its relationship to other versions, in particular the definitive, fully functional published version.
- Establishment of practical systems for ensuring that the metadata is applied by authors or repository managers and publishers.

In addition to the above, the Technical Working Group spent some time considering abstract data models and the attributes that could apply to various versions of a journal article. The working group website (http://www.niso.org/workrooms/jav) contains a full set of minutes and documents.

As a result of our analysis, the JAV Technical Working Group decided to focus on the following key points:

1. Our brief was limited in scope to *journal articles*—even so, we have recognized the possible and important, if not frequent, relationships between journal articles and other scholarly document types (such as working papers, conference papers, book chapters, wikis, blogs, etc.). Rather than creating a full set of semantics and proposed metadata disambiguating these document types, we focused on the minimum necessary to show the *relationship* between an instance of these document types and one or more journal articles. Of course, some of these other document types will be similar enough to journal articles to be able to use the same (or similar) semantics; others will not.
2. In most cases we believe the relationship needs to be codified through the retrospective act of including an unambiguous reference or link within the metadata of a "previous" version to the version of record. Although this act creates a high and potentially onerous standard of performance for some, enabling it through standard metadata and semantics and its promulgation as a best practice is crucial for establishing the relationships that the use cases suggest are necessary.
3. We decided to concentrate on a reasonably high-level set of semantics—let's say the "phylum" rather than the "species". We believe that these high-level terms give sufficient distinction for 80% of article versions—and distinction where it most matters to the reader and secondarily to the author or the publisher.
4. Each term identifies a significant value-added "state change" in the progress of a journal article from origination to publication. Five of the versions (Author's Original; Submitted Manuscript Under Review; Proof; Corrected Version of Record; Enhanced Version of Record) may have a number of iterative stages. We have not attempted to identify

these stages, although date stamps, version numbers, and metadata records may be used to differentiate them. Two of the versions (Accepted Manuscript; Version of Record) represent fixed stages. A Submitted Manuscript Under Review that is accepted for publication becomes an Accepted Manuscript at the point of acceptance. A Proof that is corrected and published becomes a Version of Record.

5. In our first set of recommendations, which were reviewed by the JAV Review Group, we had proposed only one stage after the Version of Record—the "Updated Version of Record". However, we received strong representation that this was an over-simplification, and that it was important for users of versions to know whether an update was a correction or an enhancement.

See Appendix 1: Graphical Representation of Journal Article Versions and Relationships with Formal and Gray Literature; Assumptions, Primary Challenges, and Best Practices for a graphical representation of formal and gray literature and the related assumptions, primary challenges, and conventions and best practices that were associated with this figure by the JAV Technical Working Group.

Recommended Terms and Definitions for Journal Article Versions

We propose that metadata be associated with each document object designating its status as one of the following:

AO = Author's Original
SMUR = Submitted Manuscript Under Review
AM = Accepted Manuscript
P = Proof
VoR = Version of Record
CVoR = Corrected Version of Record
EVoR = Enhanced Version of Record

Full descriptions, including definitions and notes, follow.

Author's Original

Definition: Any version of a journal article that is considered by the author to be of sufficient quality to be submitted for formal peer review by a second party. The author accepts full responsibility for the article. May have a version number or date stamp. Content and layout as set out by the author.

Notes:

1. In all definitions, the singular "Author" includes the plural "Authors". For multi-authored works, one author (the "corresponding author") takes responsibility for submitting the article for review and dealing with later stages, such as proofs.
2. We propose "Original" rather than "Draft" because "Draft" implies incompleteness, whereas an Author's Original (in our terminology) defines the point at which an article is deemed good enough by the author to be released for review.
3. This stage is *sometimes* described as a "personal version", a "draft", or a "preprint", but these terms are not synonymous.

Submitted Manuscript Under Review

Definition: Any version of a journal article that is under formal review managed by a socially recognized publishing entity. The entity recognizes its responsibility to provide objective expert review and feedback to the author, and, ultimately, to pass judgment on the fitness of the article for publication with an "accept" or "reject" decision. May have a version number or date stamp. Content and layout follow publisher's submission requirements.

Notes:

1. This stage highlights the contribution made by the peer review process to the formal record of scholarship.
2. This stage is typically characterized by intellectual critique that often leads to one or more revisions of the Author's Original, though it is possible that the critiques lead to the next stage, Accepted Manuscript, without any content changes whatsoever. Nonetheless, we propose to recognize this as a significant stage in the evolution of a journal article *whether the content changes or not.*
3. We propose "Submitted Manuscript Under Review" rather than simply "Submitted Manuscript'" to emphasize the *process* taking place and the possibility of iterations. And we propose it rather than "Manuscript Under Review" to emphasize the formal act of submission undertaken by the author to transfer responsibility to a publishing entity.

Accepted Manuscript

Definition: The version of a journal article that has been accepted for publication in a journal. A second party (the "publisher"—see "Version of Record" below for definition) takes permanent responsibility for the article. Content and layout follow publisher's submission requirements.

Notes:

1. Acceptance must follow some review process, even if limited to a single decision point about whether to publish or not. We recommend that there should be a link from the Accepted Manuscript to the journal's website that describes its review process.
2. If the Accepted Manuscript (AM) is processed in such a way that the content and layout is unchanged (e.g., by scanning or converting directly into a PDF), this does not alter its status as an AM. This will also apply to "normalized" files where, for example, an author's Word file is automatically processed into some standardized form by the publisher. The content has not changed so this essentially constitutes a shift of format only, and our terms are format neutral.
3. This stage is also known as "Author's Manuscript" by, for example, the NIH, but we believe that the key point is the acceptance of the manuscript by a second party. Elsevier refers to it as "Author's Accepted Manuscript". SHERPA/RoMEO refer to it as "Postprint", but this term is counterintuitive since it implies that it refers to a version that comes after printing.

Proof

Definition: A version of a journal article that is created as part of the publication process. This includes the copy-edited manuscript, galley proofs (i.e., a typeset version that has not been made up into pages), page proofs, and revised proofs. Some of these versions may remain essentially internal process versions, but others are commonly released from the internal environment (e.g., proofs are sent to authors) and may thus become public, even though they are not authorized to be so. Content has been changed from Accepted Manuscript; layout is the publisher's.

Notes:

1. We recommend "Proof" since this is commonly taken to refer to process stages between accepted manuscript and publication.

Version of Record

Definition: A fixed version of a journal article that has been made available by any organization that acts as a publisher by formally and exclusively declaring the article "published". This includes any "early release" article that *is formally identified as being published* even before the compilation of a volume issue and assignment of associated metadata, as long as it is citable via *some permanent* identifier(s). This does not include any "early release" article that has not yet been "fixed" by processes that are still to be applied, such as copy-editing, proof corrections, layout, and typesetting.

Notes:

1. Version of Record (VoR) is also known as the definitive, authorized, formal, or published version, although these terms may not be synonymous.
2. Many publishers today have adopted the practice of posting articles online prior to printing them and/or prior to compiling them in a particular issue. Some are evolving new ways to cite such articles. These "early release" articles are usually AMs, Proofs, or VoRs. The fact that an "early release" article may be used to establish precedence does not *ipso facto* make it a VoR. The assignment of a DOI does not *ipso facto* make it a VoR. It is a VoR if its content has been fixed by all formal publishing processes save those necessary to create a compiled issue **and** *the publisher declares it to be formally published;* it is a VoR even in the absence of traditional citation data added later when it is assembled within an issue and volume of a particular journal. As long as some permanent citation identifier(s) is provided, it is a *publisher decision* whether to declare the article formally published without issue assignment and pagination, but once so declared, the VoR label applies. Publishers should take extra care to correctly label their "early release" articles. The use of the term "posted" rather than "published" is recommended when the "early release" article is not yet a VoR.

3. The VoR may exist in more than one location (e.g., a publisher's website, an aggregator site, and one or more repositories). That is, there may be more than one *copy* of a VoR but there is only one *version* of a VoR. In Functional Requirements for Bibliographic Records (FRBR) terms, there may be more than one *manifestation* or *instance* of the VoR, but there is only one *expression* of it.
4. The VoR may exist in more than one format.

Corrected Version of Record

Definition: A version of the Version of Record of a journal article in which errors in the VoR have been corrected. The errors may be author errors, publisher errors, or other processing errors.

Notes:

1. See below under Enhanced Version of Record.

Enhanced Version of Record

Definition: A version of the Version of Record of a journal article that has been updated or enhanced by the provision of supplementary material.

Notes:

1. An update is different from a correction. With the latter, the content in the VoR is incorrect at the time of publication and is corrected. This is the equivalent of an erratum or corrigendum. With the former, the VoR is correct at the time of publication but may be amended in the future due to new information or insight.
2. If supplementary material is *linked to* from the VoR rather than being part of the VoR content, changes to the supplementary material or even to the link to the supplementary material would not constitute an Enhanced VoR. But if a link *from* the supplementary material to the VoR itself changes, this would suggest an Enhanced VoR because it suggests an update to the content of the VoR, not an update to material that sits outside the VoR. If the VoR content is amended in light of these new or altered supplemental materials, then this would also suggest an Enhanced VoR.
3. If a party other than the publisher amends a Version of Record, this would not constitute a formal Corrected or Enhanced VoR. , unless approved by the publisher, and therefore should not receive this formal designation. In fact, this version altogether loses its formal status within the typology. It should contain a note to the effect that this version is based on the VoR and was altered by X on date Y, and provide a pointer to the Version of Record. Even in the case of a true Corrected or Enhanced Version of Record, we recommend that the accompanying metadata specify who has made the update (with the default being the publisher), what was changed, and a link to the original Version of Record.

Appendix 1: Graphical Representation of Journal Article Versions and Relationships with Formal and Gray Literature; Assumptions, Primary Challenges, and Best Practices

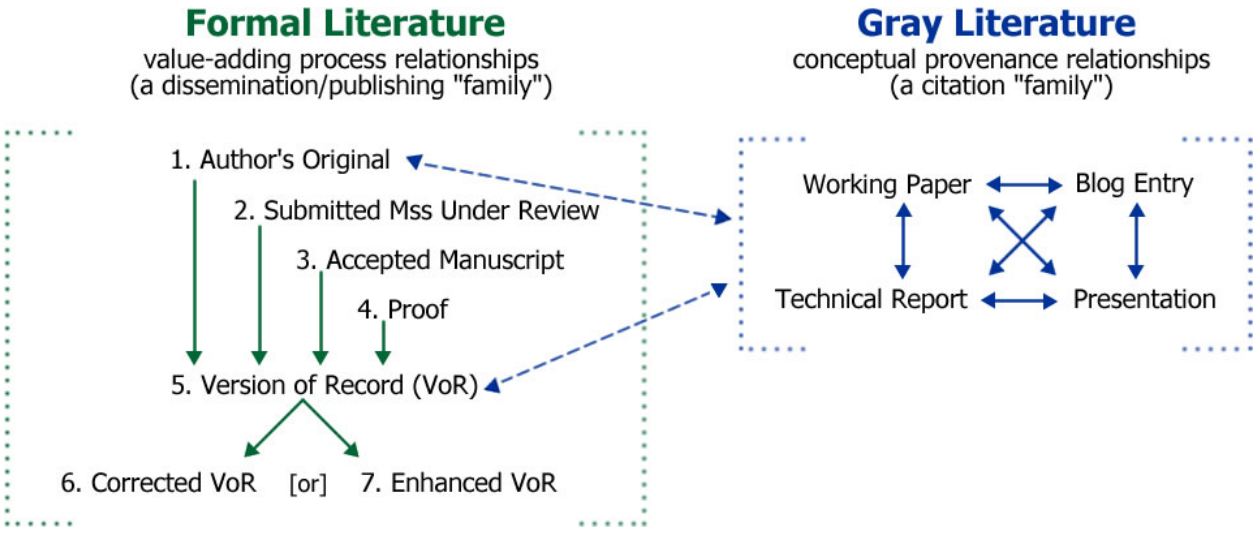

Figure 1: Journal Article Versions and Relationships

Assumptions

In today's world...

1. Any of these content objects can be made public, i.e., can reside on a network node/website.
2. Many of these content objects legally may reside in identical form in more than one place at one time (e.g., a copy on a publisher's website, an aggregator's website, and a preservation archive).
3. All of these content objects "should" have embedded/associated metadata.
4. There are two main relationship types of interest: a) the existence of previous, or more importantly, later value-added versions in the dissemination/publishing process (a sense of "stage" represented with the green solid arrows in Figure 1—the downward pointing arrows under "Formal Literature"); b) the existence of other content objects that convey conceptually similar material (a sense of provenance; represented by the blue solid arrows under "Gray Literature" in Figure 1).
5. A third type of relationship may be of interest: the intersection between a content object that is not originally part of a value-adding dissemination/publishing chain, but that becomes part of one at some time (a sense of movement into a value-adding process; indicated by the dashed arrows in Figure 1).

Primary challenges

Regardless of location...

1. Disambiguate (i.e., disclose) an object's value-added status in its dissemination family (if any), importantly including the entity/agent assuming responsibility for the object's state.
2. Disambiguate (i.e., disclose) the existence of other dissemination family members, especially the Version of Record.
3. Disclose, where possible, an object's conceptually similar family members.

Conventions & best practices

The JAV Working Group believes that standards that codify metadata for the versions above, and use and best practice conventions will need to be promulgated. These include:

- Conceptual provenance (the blue solid arrows under "Gray Literature" in Figure 1 above) should be represented via citations, acknowledgements, cover pages and the like, i.e., conventional citations within the objects rather than required object metadata.
- Objects in the value-adding dissemination process must self-identify what stage they are at and should point forward (and optionally may point backwards—see next note) via relationship metadata.
- Although, as a matter of practicality, not all forward and backward relationship information will be created, the green solid arrows under "Formal Literature" and the dashed arrows in Figure 1 are extremely and deservedly important to publishers, i.e., standards and conventions must make it easy and routine, if not required, to point to the Version of Record when one exists.

Appendix 2: Use Cases

The following use cases were considered by the JAV Technical Working Group. Below, we have indicated how our recommended terms would apply. We have used the following abbreviations for simplicity:

AO = Author's Original
SMUR = Submitted Manuscript Under Review
AM = Accepted Manuscript
P = Proof
VoR = Version of Record
CVoR = Corrected Version of Record
EVoR = Enhanced Version of Record
C/EVoR = Corrected or Enhanced Version of Record, where the corrected or enhanced status is not known from the Use Case

#0: Base Case

Author A writes first draft of article [AO]. This is circulated amongst colleagues (i.e., not formal peer review yet). Author revises article and submits to Journal Z [SMUR]. The Editors of the journal arrange for the article to be peer reviewed. The peer reviewer's comments are sent to the author, who amends the article, and resubmits to Journal Z [SMUR].

The article is accepted for publication [SMUR becomes AM upon acceptance]; the author posts this version [AM] on to her personal website and on to a subject or institutional repository. The publisher processes the article (copy-editing, article conversion, and page composition [all versions P]) to produce page proofs [P], which are sent to the author. The author marks up or annotates the proofs, which are returned to the publisher [P]. The publisher corrects the proofs and publishes the article online, without issue pagination, but with a DOI for identification [VoR]. This constitutes the official publication date.

When the print issue is scheduled to be compiled, the publisher adds volume, issue, and page numbers [VoR—addition of bibliographical details not sufficient enough to change status of VoR]. The article is published in print and electronic form with pagination and added features (e.g., citation tracking), and the author adds a URL link to this version [VoR]. The published version of the article remains available from the publisher's site, or from a third-party aggregator's site, or from an archival site [VoR].

Use Case #1: Author Smith

Author Smith writes an article manuscript [AO]. This manuscript is sent to several individuals to review; they review and provide feedback. Smith submits the article, incorporating the feedback from colleagues, to Journal A for consideration [SMUR]. The journal accepts the article provisionally [SMUR], but contingent upon revisions. Smith reviews the article, accepts the reviewers' comments, and makes a number of the revisions requested by the publisher [AM]. Smith posts the article preprint [AM] to her institution's institutional repository (hereafter

referred to as "IR"), with the approval of her publisher. The article appears in the print version of Journal A [VoR]. The article also appears on the publisher's website, but with the data set that Smith used included (i.e. not just linked), so that other researchers can also see how her work was done [EVoR].

After three months, researcher Jones uses the data set, and discovers valid and unintentional errors in how Smith processed and interpreted the data. Smith reevaluates the data based on Jones' comments, and makes slight modifications to the conclusions in the initial article, citing Jones' comments and clearly identifying the changes from the published version of the article. The corrected version of the article is posted on the publisher's website [CVoR], with a link to the previous version [EVoR], as well. Smith also posts the corrected version [CVoR], as a post-print, to her institution's IR. The preprint [AM] is removed from that site.

Use Case #2: Author Grant

Author Grant composes a manuscript that he does not share with colleagues [AO]. He sends a draft [SMUR] to Journal B, which rejects the manuscript [AO]. Grant makes some modifications to the manuscript [AO] and submits it [SMUR] to Journal C, who accepts it with significant modifications [AM]. Journal C posts the modified version [AM] to its preprint site, where it receives comments from other researchers. Displeased with the modifications and the comments, Grant requests that Journal C remove the manuscript from the preprint site (which it does), and then submits his version of the manuscript—which contains some of the modifications by Journal C and some of his own additional modifications—to his institution's IR [AO]. The institution's IR posts that version (without review), and Grant links to that "final" version [AO] from his personal webpage. Apart from cached versions of the preprint version that appeared for some time on Journal C's webpage [AM], no other version of the article is now available online or in print.

Use Case #3: Author Douglas

Author Douglas writes a manuscript [AO], which she submits to Journal D [SMUR]. Journal D accepts the manuscript without modification [AM], and publishes it in print [VoR]. Journal D is not available electronically, but Douglas receives approval to submit a PDF version of the published article [VoR] to her institution's IR.

In addition (and without prior approval from the publisher), she posts an HTML and/or text version of her document [AM] on her own webpage. This version corrects some minor typographical inconsistencies, but does not change any of the data or conclusions presented in the article. [AM—Journal D accepted the AO without modifications, so it became an AM; author makes minor typographical changes to this version, not the VoR, and the journal would still recognize it as the AM.]

Use Case #4: Author Williams

Author Williams takes the text of an article he published 15 years ago and submits it [VoR] to Journal E, which is only published as a free, online product. Editors at Journal E, not realizing this article has already been published, comment on the lack of current citations, but accept the article [AM—erroneously] nonetheless, and without modification. The article had previously been published in print only, in Journal F, and is now also available in Journal E [misidentified

as VoR since not known to be an updated version]. The only difference in the two articles is a fairly significantly different title; the texts of the two articles are identical.

A reader eventually notices the duplication, and the editors at Journal E decide to remove the article, and post an explanatory note, including a citation to the original article [VoR] in Journal F. On Williams' faculty website, he cites his publications in both Journal E [CVoR] and Journal F [VoR]. After Journal E removes the article, Williams posts the article text on his website [CVoR] and links directly to it, though he still cites the source as Journal E (reasoning that people searching abstracting databases will be able to confirm that it did appear in the journal).

Use Case #5: Conference Papers

Conference papers [not considered since out of scope].

Use Case #6: Author Davis

Journal G accepts a manuscript [AM] from author Davis, and publishes it [VoR]. The publisher also distributes the whole volume of the journal (a combined issue of all four quarterly issues for the year) as a monograph [CVoR because the bibliographic details will be different for the version in the monograph]. Davis does not have the opportunity to have his article appear only in the journal and not in the monograph. The monograph has its own title, but the articles/chapters within are identified as being duplicated from the journal. As a result, Davis' article is a standalone work in the journal [VoR], and a part of the whole in the monograph [CVoR].

Use Case #7: Every Stage Available

Author makes every stage (from initial draft [AO] to published version [VoR]) available on his personal website, and on institutional and subject repositories—every version [AO; SMUR; AM; P; VoR] everywhere.

Use Case #8: Fraudulent Posting

Article [any version] is taken from author's site and posted elsewhere, without the author's knowledge. (Or, worse still, article is stolen from author's site, and reposted with alterations, possibly with malicious or plagiaristic intent.) Such fraudulent behavior will probably lead to malicious misidentifications.

Use Case #9: Author-Updated Version

Author updates version on own (or repository) site [author may identify as CVoR or EVoR but these are not formal VoRs]. They are not formal CVoRs or EVoRs either by our definition: "If a party other than the publisher amends a Version of Record, this would not constitute a formal Corrected or Enhanced VoR. We recommend that the metadata that accompanies a Corrected or Enhanced Version of Record specifies who has made the update (with the default being the publisher), what was changed, and a link to the original Version of Record."], which then diverges over time from the "fixed" version that is the published version [VoR]. However, if the IR contains the AM rather than the VoR, then the author may be creating a new AO once the article changes make it sufficiently different.

Use Case #10: Corrected Version

Corrections to the published version are posted as the equivalent of errata or corrigenda, or the article has these corrections inserted [CVoR], or the errors are so serious (technically or legally) that the article is retracted or removed from publication [VoRs may still exist on various sites but they are no longer formally recognized; the formal publication site identifies the article as having been retracted or removed].

Use Case #11: Discontinued Hosting/Archiving

Publisher goes bankrupt without having made any arrangements for a third party to continue to host or archive the article [may be lost VoR].

Use Case #12: Repository as Journal

Article is not formally published in a journal, but the repository takes on all the formal functions of "publication" [VoR]: accepting the article for inclusion [AM]; processing it for conformance with repository specifications [P]; promoting it to the outside world; maintaining access to it; and dealing with subsidiary rights (i.e., use by other parties). [The JAV Working Group is considering only journal articles. So, if the repository takes on "all the formal functions of publication," then in this use case the repository, or some section of it, must be a journal of some kind. It is irrelevant whether it charges readers or authors and whether it is abstracted or indexed. This case is similar to the operation of some "overlay journals" and it implies that articles accepted for inclusion have undergone some formal review process as they pass through the SMUR stage.]

Use Case #13: Multiple Postings

An article is written by three authors, each of whom works in a different institution. Each institution has its own repository, and each institution requires that the article is deposited in its own repository since it reflects work done by its researchers. The research work is also funded by more than one granting body; each granting body also requires posting in a different repository. Over time, the versions in each of these repositories diverge as they are amended by the authors or the repository managers [various AOs; one version may be the actual AM or SMUR, but others may be author-amended versions of the AM].

Use Case #14: Long-Term Preservation in IR

An article [AO] is submitted to Journal C [SMUR], is accepted after peer-review comment, and the author deposits the version with the corrections [AM] in her IR; Journal C formally publishes it [VoR]. The IR managers convert the AM into a different file format suitable for long-term preservation, which preserves the layout as well as the information content [remains an AM since versions are not format-dependent].

Use Case #15: Discovery of Multiple Versions

User A does a search on Google Scholar for a phrase B. Google Scholar retrieves a list of possible records. User A selects Article C in Journal D and has the following options:

a) full text of the author's accepted manuscript in TeX, accessible via the author's IR [AM];
b) full text of the published version, in PDF [VoR], via a host—access to this version requires that User A's institution subscribes to Journal D;
c) full text of the published version in HTML format [VoR], via an aggregator—access to this version requires that User A's institution subscribes to Aggregation E, which includes Journal D;
d) a link to a Preprint Site F, which offers a Word version of the manuscript, prior to peer review and acceptance in a journal [AO].

Use Case #16: PDF Without Explicit Version

User E starts searching his IR and locates a record that links to the publisher's site for Article F [VoR]. In addition, there is a link to a local (IR-hosted) version of Article F. This version has no explicit version information in the metadata, but it appears to be the published PDF since it has the publisher's copyright line and is paginated within a volume and issue of the journal [VoR].

Use Case #17: Word Document Without Explicit Version

User H starts searching her IR and locates a record that links to the publisher's site for Article I [VoR]. As for Use Case #16, there is also a link to a local (IR-hosted) version of the article, but in this case, the version is in Word, and there is no clear indication of its relationship to the published article [May be AO or SMUR or AM. The fact that it is in Word does not preclude it from being VoR, since format is not a consideration. But the fact that it does not have formal publication indicia such as volume, issue, pagination in running heads and feet, or permanent identifier indicates that it is not VoR].

Use Case #18: TIFF Format Without Explicit Version

User J starts searching his IR and locates a record that links to the publisher's site for Article K [VoR]. As for Use Cases #16 and #17, there is a link to a local (IR-hosted) version of the article, available as a scanned image in TIFF format. Although it is identified as a "pre-publication" version, its relationship to the published version is unclear [may be AO, SMUR, AM, or even P].

Use Case #19: XML from Different Sources

User M belongs to Institution N, which has an OpenURL resolver that targets its IR as well as publisher resources. User M searches in Journal Article Database O to locate an article; the OpenURL resolver offers User M full text in XML from the publisher [VoR] and full text, also in XML, from within its IR. However, although these are both in XML, one has been produced by the publisher as an end product of the typesetting process [VoR], whereas the other has been created by the IR manager in a separate, post-publication process [not the formal VoR but an unauthorized version of it]. Each XML file complies with a different DTD (one the publisher's and the other the IR's), although the content is identical in each. [Our typology is format-independent, so the fact that the article is in XML is irrelevant. Thus, if the IR XML was derived

from an AO, SMUR, or AM, then it remains that version, as the format is irrelevant. Further, our typology also does not distinguish between authorized and unauthorized copies of an article. This, and all questions of infringement, is a matter orthogonal to the typology. Just as someone can illegally reproduce a printed VoR, so someone can post an unauthorized VoR. Also note that this case claims that the two DTDs and conversion processes result in "identical content". If that is indeed true, then the determination of each version depends solely on the source version used to produce the XML. However, it should be noted that in practice, different DTDs and conversion processes applied to the same original article could produce differences in content in the resultant XML files.]

Use Case #20: Links Accompany Article

User P belongs to an institution that subscribes to the Journal Q, and she uses the publisher's website to locate the final published version of Article R [VoR]. The publisher offers both PDF and HTML versions of the article [VoR], and also provides links to comments on the article and to related articles [not an EVoR since the links lie outside the article itself], some of which appear in the subscribed journal, and others in journals to which the user's institution does not subscribe.

Use Case #21: Search Returns Multiple Articles

User S does a phrase search and locates a) Article T in the author's IR with links to the publisher's full text [VoR]), as well as to b) Articles V and W in two different repositories. It is unclear whether V and W are the same or different versions of T, or a different article entirely. (This use case does not provide enough information to identify the version.)

Use Case #22: Multiple Formats Throughout Lifecycle

Author A submits an article to Journal Z in LaTeX, Word, PostScript, or PDF with separate figure files in PS, JPG, TIFF, etc. The editors of the journal arrange for the article to be peer reviewed and prepare a PDF for the reviewers [SMUR].

The peer reviewers' comments are sent to the author, who amends the article, and resubmits to Journal Z. The Editor prepares a PDF of the amended version and sends it back to the referees. The loop of review and revision can be repeated a number of times [SMUR].

The article is accepted for publication [AM]. The accepted version is the last version submitted by the author in his own format. The publisher also creates a PDF version of the article [AM since no content processing].

The publisher processes the article: copy-editing, article conversion, and page composition, creating intermediate "versions" [P]. This process produces PDF page proofs [P], which are sent to the author. The author marks up or annotates the proofs [P], which are returned to the publisher. The publisher corrects the proofs and occasionally there will be a second round of proof checking [P].

The publisher publishes the article online (without issue pagination, but with a DOI for identification) and in one or more file formats: HTML, PS, PDF [VoR]. This constitutes the official publication date. When the print issue is compiled, the publisher adds volume, issue, and

page numbers [remains the VoR]. A high-resolution file is sent to the printer. The online files are changed to include pagination and features are added (e.g., citation links, citation tracking, links to supplementary data) [VoR—if links to these additional features are provided, it remains VoR; it becomes EVoR only if the additional material is incorporated in the work itself, not merely pointed to. See #0: Base Case]. A full-text XML file and figure files exist behind the scenes and could be used online instead of HTML to regenerate new PDFs [C/EVoR—on the assumption that the new PDFs will have some corrections or enhancements compared with the original PDF] or to make other future formats [VoR or C/EVoR—depending on whether the future format is a new format of the original VoR or of the C/EVoR].

Use Case #23: Author Alerts Repository of Version Change

Author (or sanctioned author's agent, such as a department or research unit) deposits her wholly owned working paper [AO] into an institutional repository (first converting, if necessary, to one of a limited number of acceptable formats). Depositor supplies an abstract and descriptors. Repository converts the item to PDF, assigns a unique ID, constructs a metadata record, and makes the paper accessible with a "cover page" supplying full bibliographic information and suggested citation [AO, if not yet AM]. Author is encouraged, but not required, to alert the agent and/or repository, and supply details when/if:

- there is a revised version of the item [AO, SMUR, or AM if accepted for publication in a journal], in which case the repository suggests deposit of the revised version and updates the metadata of both the previous and revised versions and provides forward and backward links to both;
- a version of the item is formally published [VoR], in which case the repository version's metadata is updated with the bibliographic information and URL of the published version; or
- the author wishes to remove the item (e.g., due to "prior publication" concerns of a publisher), in which case the item will be replaced by a metadata record indicating that the item has been removed and, where applicable and supplied, the bibliographic information and URL of the published version.

Use Case #24: Post-print Deposit in Repository

Repository harvests citation data for papers written by institutional authors during Time Period X, filters them for "post-print friendly" publishers (according to SHERPA/RoMEO and locally updated data), and sends a request to the author to deposit an "author's copy" [may be AO, SMUR, or AM] or, when clearly allowed, a publisher's copy [VoR].

Upon deposit the author (or her agent) verifies bibliographic metadata, supplies an abstract and descriptors. Repository double-checks rights status and publisher post-print policy and, if all is in order, converts the item to PDF [AO, SMUR or AM], assigns a unique repository ID, constructs a metadata record, and makes the paper accessible with a cover page, supplying full bibliographic information of published version, suggested citation, and required publisher's statement, if any. The author is encouraged, but not required, to:

- provide bibliographic and location information about "preprint" versions [may be AO, SMUR, or AM] of the item, and/or
- alert the repository if the author wishes to remove the item, in which case the item will be replaced by a metadata record indicating that the item has been removed, along with the bibliographic information and URL of the published version.

Use Case #25: Summary Version

A publisher publishes a short summary version (identified, for example, as an "abridged PDF") of an article in print, and both the short and full version of the article online. These online versions have the same DOI. Only the long version is regarded as the VoR; the short version is regarded as an extended abstract, rather than an alternative version of the article. [NB: If the publisher were to assign distinct DOIs to the long and short versions (thus indicating they are independent publications), then each one is a VoR, one being a derivative work of the other. The metadata and reference list should clarify the relationship between the two.]

Appendix 3: Comments from JAV Review Group on Recommendations

Below are listed comments received from the Journal Article Versions (JAV) Review Group to an earlier JAV Technical Working Group document submission, and the Technical Working Group's responses.

Section 1.01 Subject	***Section 1.02 Comment***	***Section 1.03 Reply***
1) Terminology	Should we use completely new terminology in order to avoid "loaded" terms (the associations that different groups may attach to terms)?	We discussed this possibility early on but decided against it. We tried to avoid loaded terms like pre-print, post-print, authoritative version and chose to use understandable terms with more explicit definitions.
	Top-level vocabulary is excellent. However, within each top-level category there are many subclasses. Next step may be to define a "similarly brief but powerful" vocabulary of sub-types.	We feel that the next step is to operationalize the definitions. Then we would want to keep the typology simple in order to facilitate work implementing standard citing statements and metadata and to build linking mechanisms.
	Alternative view above: proposals are practical and balanced; likes the simplicity of having only five categories, and we should resist temptation to create sub-divisions. Date-stamping may help users navigate through the sub-divisions without the need for new terms	Thank you – and we agree about the date stamping.
	Don't follow definitions with "also known as". This just leads to argument over semantic equivalence of various loaded terms	We feel it is helpful to link our standard terms to other terms in common use by particular communities, although they are not always synonymous, and we make that clear in the narrative.

Section 1.01 Subject	***Section 1.02 Comment***	***Section 1.03 Reply***
a) Author's Original	Author may revise AO as part of peer review process. Calling such a version "Author's Original" doesn't seem right—may need an "Author's Revised" stage?	We agree that it does not seem right to call revised versions arising from the peer review process Author's Original. We have named all such versions Submitted Manuscript Under Review (SMUR). Note that this new stage also designates a public act of formal submission and review whether or not any revisions arise from it.
	Term and definition are fine.	Thank you.
b) Accepted Manuscript	Are iterations possible? E.g., article withdrawn after acceptance and subsequently accepted by another journal?	The example is rare enough not to be included in high-level semantics. We think it is OK to deal with this type of rare situation as exception handling.
	If an IR Manager changes an AM (e.g., format conversion), is this a "versioned AM"?	We did not generally want to consider formats or media as different versions at the top level. In general, we think these cases would be refinements to major version categories. However, there is some technology dependence here. There are such things as "lossy" conversions, e.g., Google's html'izing of PDFs. These are "defective copies" rather than new versions.

Section 1.01 Subject	*Section 1.02 Comment*	*Section 1.03 Reply*
	Does the AO designation remain with the "as submitted" article version (use case #0)?	Yes. This is a very important point because it highlights one of the key criteria we considered in differentiating our types, viz. formal public "acts" by a recognized party. So, the resubmitted version might incorporate all the changes required by the peer review process, and in such a case the content is identical to the AM, but it is still considered an AO until changes are externally verified by a formal act of acceptance. A consequence of this approach is that there must be retroactive version re-naming. If the author deposits the peer-reviewed and corrected AO in an IR at the same time as re-submitting, the "act" that converts the type to AM must apply to all copies.
	Not clear what is meant by the phrase "takes responsibility"—does it mean "asserts publishers' rights"? The description of the review process is unclear—the decision taken by an IR manager, for example, is not a review process	Good point. We did discuss "takes responsibility" initially in a legal sense, but decided instead use the term in the sense of "publicly lends its imprimatur to the scholarship" or "stands behind and supports the work as a valued contribution to scholarship"
	The Author's Original after peer review and just prior to the act of acceptance is essentially the same file in terms of content but may be quite different from the Author's Original prior to submission. Thus your scheme is confusing because it calls two different files, (the one submitted by the author and the one amended by peer review), by the same name, i.e., Author Original, and the same	We agree. Our scheme fails to distinguish an Authors Original Manuscript from all the versions that arise from the onset of a formal value-adding process. This failure does lead to a non-intuitive result: the AO is identical in content to the AM at one moment in time. It takes too much explaining to rationalize this counter-intuitive result. Our scheme incorporates two principles. We look at content changes from a value-adding evolutionary

Section 1.01 Subject	*Section 1.02 Comment*	*Section 1.03 Reply*
	file, the one amended by peer review just prior to act of acceptance and the one after acceptance, by two different names, i.e., Author Original and Accepted Manuscript. In this respect, your taxonomy does not explicitly account for the importance of the peer review process in journal article versioning and therefore diverges markedly from the STM Stages 1-3.	perspective AND we look at points of public authoritative affirmations about articles regardless of content changes. We have subsumed the first principle in the second to such an extent that AO content changes are not recognized as significant in themselves but ONLY by virtue of the declaration "ACCEPTED" by the publicly recognized entity. Both principles of our schema should be operative. And we therefore amend our schema to include a version called Submitted Manuscript Under Review (SMUR) between AO and AM. We note a practical side to this distinction as well in terms of authority and responsibility. The author may do as he/she wishes with an AO. That is not true with a SMUR. Once the author submits, he is obliged to follow the publisher's rules. These invariably include one that states that the author is not free to submit elsewhere until the review process is complete or the manuscript is formally withdrawn. It also may mean that a would-be plagiarizer is going to have to deal with a significant organization rather than an individual author. It is fair to say that submission is a public act with consequences for who is responsible. In fact, some publishers require copyright transfer at the moment of submission; upon rejection, the copyright automatically reverts to the author and the version reverts from SMUR to AO. And while it is true that rejection is usually not a

Section 1.01 Subject	***Section 1.02 Comment***	***Section 1.03 Reply***
		formally recognized PUBLIC stage in the evolution of an article, it is true that "REJECTED" is an expression of the authority of a publicly recognized entity with respect to the document over which it has been exercising its responsibilities.
c) Proof	No one has made any specific comments about this term, apart from one general comment about consideration being given to sub-types (e.g., uncorrected proof, corrected proof, revised proof)	At this point, we would like to keep to the "phylum" level. There may be many further subdivisions, but these are more fine-grained than is necessary for our high-level view.
d) Version of Record	Could we have multiple VoRs, i.e., same VoR but in different locations? (See use case #13.)	Yes. Copies of VoRs proliferate online just as in print. Initially, we did consider "location" as a criterion for distinguishing versions, but decided against that. In other words, our typology ignores the fact that one VoR a user finds might be "an unauthorized copy" or infringing IP from a legal standpoint.
	What difference does it make if a VoR is released in various file formats a) simultaneously or b) asynchronously.	For us, formats would be sub-types of the progressive staged versions we laid out, not different versions at the same hierarchical level.
	Term is "rather inflated" and doesn't take into account the occasions when the published version is inferior to the Author's Original. "Published version" is simpler, more accurate, and more neutral than "Version of Record". [But see other views below.]	We disagree. Version of Record was chosen to avoid inflation of Authoritative Version or Definitive Version. "Published Version" is a highly ambiguous term. It does not distinguish between "published" as in "making publicly available" and "published" as the end result of a formal process that changes content and fixes it in a citable stage. Our entirely typology of versions has explicitly built into it the notion of progression.

Section 1.01 Subject	*Section 1.02 Comment*	*Section 1.03 Reply*
	In the response above: "There will always be cases that don't quite fit. The author's version [i.e., Accepted Manuscript] may be better in some way but there needs to be at some point a defined final record for working purposes."	We agree.
	I endorse the use of the term "Version of Record": "The academic community does want one fixed version that will be the default version for citing."	We agree.
	I endorse the term "Version of Record." All terms are loaded.	We agree.
	There may be multiple VoRs (and UVoRs) floating around. It would be helpful to have some way of indicating which VoR (and UVoR) is the "authentic" one, i.e., the one that is identified as such by the rightsholder. This may also include information about which format (e.g., XML, HTML, PDF, print) constitutes the VoR.	Yes, there will be multiple VoRs (see first VoR comment above). Copies of VoRs abound. There is no need to distinguish them in this typology unless one is not a VoR. There is a parallel with printed copies: everyone has their own copy. We are not concerned with copies but rather with stages. Introducing authenticity as a function of legal rights to expose, display, or serve a copy of a VoR is, we think, outside the scope of our group's work. Formats are sub-types that we are not concerned with in our high-level typology.
	Definition is too loose—it says that "any organization" that makes public can declare a VoR, but this may lead to confusion with IRs.	We agree, and we will clarify in the narrative. Our key is the declaration "fit for publication" and the process, implied by the context of the stage that leads to it. We have a progressive staging model. One cannot skip directly from AO→VoR.

Section 1.01 Subject	***Section 1.02 Comment***	***Section 1.03 Reply***
e) Updated Version of Record	VoR needs to differentiate between changes to content and more "peripheral" changes, such as pagination, formatting, linked errata, comments or datasets (see also note below re AVoR and EVoR).	OK. A number of Review group comments indicate that Updated VoR is too broad a type because it conflates things that are significantly different. We will change the recommendations to distinguish between "Corrected VoR" (which would include all errata and corrigenda) and "Enhanced VoR" (which may contain extra material).
	How does one deal with linked errata and post-publication comments?	Our scope is the journal article itself. Links to supplemental material should not be considered changes/updates to VoR. If the content is amended in light of these supplemental materials, then it is either a Corrected VoR (if the amendments are in the nature of a correction) or an Enhanced VoR (if the extra material is added to the VoR, not just linked from it). Corrected VoRs and Enhanced VoRs should always link back to the original VoR.
	What happens when the destination of an embedded link changes? Does that constitute a change to the VoR?	No—the VoR hasn't changed.
	Is the article in use case #10 a "lost VoR" rather than a UVoR?	We agree. A retracted article usually disappears from a site; a change is often made in the Citation Page that remains visible. This could be formalized in a "metadata" editorial note: "Retracted because…." We do not think this affects the versioning scheme that we have. We think it is a rare case that requires non-standardized exception handling. In theory, the extant VoR copies should have a watermark "Retracted" inserted.
	The print version of a VoR is a VoR, not a UVoR (use case #0).	We agree and will amend the use case text.

Section 1.01 Subject	***Section 1.02 Comment***	***Section 1.03 Reply***
	Should we have "AVoR" for "Alternative VoR" (e.g., same content in different locations; maybe also same content in different format)?	We don't think so. In our scheme, location is not a factor in determining VoR and formats would be more refined sub-types.
	If above accepted, would also lead to requiring a UAVoR.	See above.
	"EVoR" for "Enhanced VoR", e.g., links added but content not changed?	See above re proposal to introduce "EVoR". Added links constitute an Enhanced VoR since the content has changed by the addition of these links.
	"EUVoR" for "Enhanced UVoR"?	We will not distinguish between further iterations of EVoRs—as with other iterative versions, version numbering or date-stamping could be used to disambiguate.
	See above comments re VoR – by the same logic, prefers "Updated Published Version."	We disagree—see comments re VoR versus "Published Version" above.
	Corrections should be treated as a special case and always be linked to the VoR. Retractions are a special form of UVoR.	See comments above re "Corrected Version of Record" and comments re retracted articles.
	The VoR should be the one and only definitive version.	Yes but copies abound; Versions of Record don't.
2) Metadata	"Practical systems for ensuring that the metadata is applied by authors or repository managers and publishers" may be impediments to authors self-archiving.	See below.

Section 1.01 Subject	*Section 1.02 Comment*	*Section 1.03 Reply*
	In response to above: "It is not our concern to establish self-archiving or any other model. We are intending to describe what might be found; in a standard way; we are not intending to specify how it might get there, or why one might choose to use a particular form."	We agree.
	In response to note re self-archiving above: "no standard can compel. The decision of some body to require the standard is what would compel."	We agree.
3) Enforcement/compliance	Who will be responsible for making sure that articles are correctly identified and point properly to variants (both backward and forward)?	We agree that this is a key question, but we think it is outside the remit of our working group, and may be covered by the evolution of best-practice guidelines. CrossRef has been working toward practical methods of version identification. One activity is Publisher-to-Author communication of both standard wordings and supply of particular DOIs; see *CrossRef Guidelines for Standard Citations in Author Postings*, available at: http://crossref.org/08downloads/author_guidelines.pdf. The second activity to implement a system for interlinking of versions are CrossRef-IR processes. These have been discussed with a number of IRs; some are actively retrieving DOIs for VoR and inserting link into IR versions; and plans are underway to build necessary support for automated look-ups and DOI retrievals into some IR software.

Section 1.01 Subject	*Section 1.02 Comment*	*Section 1.03 Reply*
	How would "fraudulent misidentification" (use case #8) be identified?	Well, it certainly won't be embedded in the metadata or full text by the perpetrator! Once discovered, there could be various resolutions, but we think this situation is rare enough to be dealt with by special case exception handling.
	Will there be suggestions from the group on how libraries/repositories, etc. should adopt and use the standards?	See CrossRef references above. In addition, the JISC-funded VERSIONS project (http://www.lse.ac.uk/library/versions/about.html) may address this issue.
4) Use cases	Specific comments and wordsmithing.	We will address these and amend the use cases text accordingly.
	Various specific comments.	As above.
	Specific comments re Case 10 (as a variant of case 0) and further use cases.	As above.
	#1: Should avoid even mentioning terms "preprint" and "postprint". Despite the SHERPA/RoMEO definition, many people (intuitively?) use "postprint" to mean the publisher's PDF. #7: not very realistic. Authors don't want this – but could be an edge case. #9: a bit speculative, but possible. Might be interesting to develop #12: IRs should not be able to give an article "VoR" status. But maybe we need to clarify whether VoR can be assigned to an article that is available only in an IR and not otherwise published? #14: not clear which version the IR	See previous comment on utility of referring our standard terms and definitions to other looser and ambiguous uses. The use cases were written prior to the Terms & Definitions so we will rewrite the Use Cases using only our typology. #12: In our typology, a VoR has to have been formally published. It may exist in more than one location (e.g., on a publisher's site, in an aggregator service, in one or more repositories) but its designation as a VoR indicates that it is the version of the article that has been through a process of verification and certification. Even if an IR contains a VoR, it is not the IR that has taken responsibility for declaring the version to be the VoR—this has been done by the entity acting as the publisher. If an IR does take on this declarative responsibility, it is

Section 1.01 Subject	*Section 1.02 Comment*	*Section 1.03 Reply*
	manager is converting	acting as a *de facto* publisher. Other typologies or vocabulary sets take a different view. For example, the CrossRef Glossary (http://crossref.org/02publishers/glossary.html) is not restricted to journal articles and therefore makes a distinction between "Version of Record" understood as a fixed end-point in the life-cycle of a work and "Definitive Work" understood to be a formally published fixed end-point which incorporates processes of community validation and publisher approval. In this way, IR contents can, and will often be "Versions of Record" in CrossRef terminology, when not followed by formal publication. (Although there has been some crossover in the work of JAV Technical Working Group and CrossRef's Institutional Repository Committee that is evident in the Glossary, the perspectives and scope of the two groups differed, so there is also divergence.)
5) The data model	Relationship metadata requirements: change "must" to "should"	OK – we will amend.
	The diagrams portray a division that doesn't exist from the author's point of view. "Dissemination/publication" and "citation" are not two separate timelines	The conceptual family represents objects that are beyond the scope of our Journal Article Version definitions. Nonetheless, relations exist between these informal, gray literature, works and formal journal articles. Authors recognize the conceptual difference even while wanting to be cited prior to formal journal publication. The left part of diagram might be super-labeled "Formal Literature" and the right side "Gray Literature" and we might add that Conventions and Best Practices have the goal of interlinking the two.

Section 1.01 Subject	*Section 1.02 Comment*	*Section 1.03 Reply*
		Internal, embedded, citations are required but formal pointers external to text are also important. References are not usually called metadata, but are often treated as such in online offerings: they are often extracted from text and exposed external to it with live links. But retroactive forward linking to chronologically later members of the conceptual family (accomplished to a degree by search engine "clustering") should be incorporated in metadata with pointers.
6) General	Make comment in narrative that date of first "making available" potentially important in claims of priority and patent registration	Articles may be made available without being formally published, e.g., preprints in arXiv. The relevance of these postings to claims of priority and patent registration raises issues that are outside the scope of our scheme.
	Clarify which article versions could be iterative. (Narrative implies only AM and VoR are "fixed", but is this true? Graphic in model shows two-way arrows, implying potential iteration.)	We will fix the diagram. The narrative states that AMs and VoRs are not iterative, but all other versions may be.
	Looking for terminology that will help automated update of holdings, automated resolution of citations, ability to retrieve related materials—check against these requirements.	We feel that a standard terminology is at least a start. We do not want to go too far outside our specific scope as a Working Group.
	Comment: "for the less important journals, there is now no functional separation between editing to accommodate peer review, and copyediting."	This is a matter of opinion. Peer-review editing and copyediting may be much the same thing or may differ considerably. Our typology makes no assumptions regarding degree of difference.

Section 1.01 Subject	*Section 1.02 Comment*	*Section 1.03 Reply*
	In the preamble, would prefer that we did not refer to various stages as "added value". More accurate and neutral to say "revisions".	We disagree. Value is not a commercial term; it points to something that *may* be commercialized *or not.* Added value is actually a critical part of our *developmental, progressive stages* model. The VoR is not merely a Fixed Point, but it is a Valued Reference Point that everyone wants identified. All other versions relate to the VoR.
	Endorses comment above.	See response to above. The River Project produced a terrific report.[1] It takes many nuances into account that are purposely omitted by the JAV Technical Working Group. One of their nuances is to understand that different users will value different versions for different purposes. By trying to place version identification into a relationship with the user, the complexities rise exponentially.
	Commends the balanced approach and avoidance of business model issues. Hopes we can keep this up.	Thank you.

[1] *Scoping Study on Repository Version Identification (RIVER) Final Report* (March 2006), available at http://www.jisc.ac.uk/media/documents/themes/infoenvironment/riverfinalreport.pdf