Edition 2.2

Author date: 2025-03-17

Archive date: 2025-03-17

Why Publish Baseprint Document Successions

E. Castedo Ellerman (castedo@castedo.com)

Citation:
E. Castedo Ellerman (2025) "Why Publish Baseprint Document Successions" https://perm.pub/wk1LzCaCSKkIvLAYObAvaoLNGPc/2.2

Abstract

Baseprint document successions offer benefits beyond preprints. They can be used for other document types, such as supplemental material and technical blog posts. Additionally, researchers can access these citable documents through multiple compatible websites and software applications. Despite being distributed across independent platforms, authors still retain exclusive control to amend their Baseprint successions. Baseprint successions also provide key benefits of preprints, including long-term preservation through archiving, preservation of multiple document versions, and the ability for authors to add new document versions. This allows the research community to cite specific fixed versions while also having the option to discover newer versions.

Summary

Baseprint document successions are preserved in Git-compatible[1] repositories and archives. Three such examples are the Software Heritage Archive, GitHub, and GitLab. See the Relationship to Git section for more details. The data archived is the machine-readable content of a document, separated from how it is visually presented. The benefits of this separation are discussed in the Diversity of Reading Venues section. A key component to making this separation possible is discussed in the Document Succession Identifiers section. These identifiers can be used as permanent references similar to a DOI[2]. This extends one of the great features of traditional academic publishing, namely, the ability of researchers to reference a static archived document long into the future.

Diversity of Reading Venues

Multiple independent websites and applications can access a Baseprint document succession to present the document in various ways. Baseprint successions are not tied to a single website, allowing readers the freedom to choose between alternative reading venues.

As of 2024, JATS XML[3] is used to encode the contents of Baseprint snapshots that make up a Baseprint succession. A Baseprint snapshot gives readers the choice between multiple web page formats, similar to the choice between PubMed Central (US), Europe PubMed Central, and journal websites.

Researchers with accessibility needs, such as those with vision impairments, benefit from a variety of websites and software applications. By encoding new research documents in a format like JATS XML, rather than PDF, certain websites and software applications can present a Baseprint document in ways that enhance accessibility.

Document Succession Identifiers

A Document Succession Identifier (DSI)[4] is an intrinsic persistent identifier[5] of a Baseprint document succession. It is a textual identifier similar to a DOI[2] or a web address URL. This document itself is archived as a Baseprint document succession with DSI:

`dsi:wk1LzCaCSKkIvLAYObAvaoLNGPc`

This is a *base DSI*, which identifies all snapshots within a Baseprint document succession, both current and future. Similar to textbooks and preprints, Baseprint document successions contain multiple editions (or versions) of a Baseprint document, each encoded as a Baseprint document snapshot[6]. A base DSI identifies all the editions of a document added to a Baseprint succession. Usually, readers are interested in the latest edition in the Baseprint succession.

An edition number can follow a *base DSI* to identify a specific static edition. For instance:

`dsi:wk1LzCaCSKkIvLAYObAvaoLNGPc/1.1`

identifies the earlier edition 1.1 of this document. Edition 1.1 will never change.

Author-Owned Identifiers

A Baseprint succession is a work by an author, as declared by the author over time. Unlike a traditional journal article, a Baseprint succession is not a single final published result. Unlike preprints on a preprint server, a Baseprint succession is not a sequence of deposits at a specific preprint server. A DSI identifies an author's work independently of where it is stored or viewed.

An author determines the editions in a Baseprint succession by digitally signing the Baseprint succession with an SSH signing key.

Multilevel Edition Numbering

Multilevel numbering is found in the numbering of chapters, sections, and subsections (e.g., Chapter 2, Section 2.4, Subsection 2.4.3) as well as software release versions (e.g., software release 2.19.2). Authors of Baseprint successions can use multilevel numbering to identify editions or stick to simple edition numbers of just positive integers, like textbook editions and preprint version numbers.

Multilevel numbering is particularly useful when amending editions for a binary change between Baseprint document snapshots but not in the intellectual content. The DSI specification does not specify the meaning of different levels in edition numbers, except that larger integers come after smaller integers, and higher-level edition numbers identify subordinate sequences of lower-level edition numbers (e.g., the entire sequence of edition numbers 2.1, 2.2, 2.3, … can be identified by edition number 2).

Relationship to Git

While Git is normally used to store source code revisions, it is not used for that exact purpose with Baseprint successions. Git is used as a pragmatic implementation layer because it can handle distributed data replication, digital signing, and cryptographic hashes. The only history stored in a Baseprint succession is primarily the history of additions of new editions to the succession.

Conclusion

Baseprint successions have been implemented and used by the author for years. Authors interested in publishing Baseprint successions can visit try.perm.pub to get started.

For technical details on how DSIs are implemented, see the Document Succession Identifier Specification[4], Document Succession Git Layout[7], or the software library at gitlab.com/perm.pub/hidos.

Acknowledgments

This document was copyedited using CopyAid.it, which uses OpenAI GPT-4.

Changes

From Edition 1.1 to 2.1

- The “digital succession” terminology has been updated to “Baseprint document succession” terminology.

From Edition 2.1 to 2.2

- XML restyling
- Minor changes in wording
- Discuss accessibility in reading venue section

References

1. Git — Wikipedia, the free encyclopedia. 2023. Available: https://en.wikipedia.org/w/index.php?title=Git&oldid=1177307938
2. Digital object identifier — Wikipedia, the free encyclopedia. 2023. Available: https://en.wikipedia.org/w/index.php?title=Digital_object_identifier&oldid=1177404866
3. Journal Article Tag Suite — Wikipedia, the free encyclopedia. 2023. Available: https://en.wikipedia.org/w/index.php?title=Journal_Article_Tag_Suite&oldid=1172109468
4. Ellerman EC. Document Succession Identifiers. 2024. Available: https://perm.pub/1wFGhvmv8XZfPx0O5Hya2e9AyXo/2
5. Heritage S. Intrinsic and extrinsic identifiers. 2020. Available: https://web.archive.org/web/20221019201056/https://www.softwareheritage.org/2020/07/09/intrinsic-vs-extrinsic-identifiers/
6. Ellerman EC. What is a baseprint? 2023. Available: https://perm.pub/HKSI5NPzMFmgRlb4Vboi71OTKYo/1
7. Ellerman EC. Document Succession Git Layout. 2024. Available: https://perm.pub/VGajCjaNP1Ugz58Khn1JWOEdMZ8/1