Developing an Archiving Strategy: What content to archive?

It seems unlikely that it is either necessary or feasible to archive every aspect of a publication. The first step is to consider how important the various components of the publication are to the understanding of the work by a future reader, and prioritise the most important components. We recommend four prioritised categories.

Category 1. Original content created for and core to the publication that is not being archived elsewhere.
This is most likely to include the core text written by the author as well as content such as images, audio/visual and other content which are embedded or linked to within the work. It may also include specific layouts for the content, such as in poetry or texts with line referencing. Metadata for the publication should also be included in this category.
This is the content that it will be most important to develop archiving solutions for.

Category 2. Non-original core content (such as third party images) that is very important to the understanding of the work, are embedded or linked to within the work, but for which there is no alternative archiving solution available. This, by definition, is not original content created for this publication - but readers will need access to this content to understand the author's work and original contribution, so addressing ways to ensure this content is archived and remains accessible to future readers will clearly be very important.

Category 3. Content important for the understanding of the work, but for which it seems reasonable to assume that archiving solutions are already in place? This might include third party content embedded in the work, important citations or original datasets related to the work but being archived elsewhere. In this case providing permanent links to the archived versions of the content or sufficient information for the archived work to be discovered by a future reader will be the priority, rather than necessarily archiving the content directly.

Category 4. Nice to have, but inessential content. Content that is deemed less essential for future readers to access can be considered as the lowest priorty for archiving. Providing permanent links to archived versions of this content (if available) is clearly still desirable, as is directly archiving that content if it is feasible to do so - but this category of content does not need to drive the archiving strategy.

Revision #1
Created 2026-04-28 15:08:46 UTC by Rupert Gatti

Updated 2026-04-30 16:10:06 UTC by Rupert Gatti