

Publications Exchange Working Group

Chairman: Trevor FAULKNER, United Kingdom

Appendix 5: Guidelines for the Creation, Exchange, and Storage of Digital Publications

Issue 2: 9 November 2016

The purposes of this document are:

1. to alert editors, publishers and distributors to issues they should consider in producing speleological papers, journals and books in the 21st century;
2. to set out the standards and procedures endorsed by the International Union of Speleology in relation to publications;
3. to facilitate the exchange of digital publications between members of the Union and with other libraries and institutions; and
4. to attempt to maximise the long-term survival and accessibility of those publications.

It is not the purpose of these guidelines to discuss the respective merits or shortcomings of physical printing of documents on paper versus digital production, nor any copyright issues. It accepts that both methods of “publication” are available to speleological organisations and have their strong points and their advocates.

1. Creation of publications for digital distribution

When creating new publications, editors now need to be conscious of how they may be distributed, viewed and stored. While there is an increasing tendency for digital documents never to be reproduced on paper, in the early 21st century it is desirable that magazines (including scientific journals) and books should be able to be either viewed on screen or printed out as hard copy. This requires the imposition of limitations on the use of any digital techniques which cannot ‘translate’ onto paper, such as the use of HyperText Mark-up Language (HTML – useful for designing web pages and viewing on screens because it works in windows of different sizes, but that flexibility means loss of control of the spacing and arrangement of material on the pages, as well as the actual fonts used), the inclusion of audio or video clips, scrolling windows or pop-ups and the use of hyperlinks to connect to external sources.

Scientific publications (and we include in this caving club newsletters as well as more esoteric speleological journals) are intended for the dissemination of information and the progressive accumulation of knowledge. To fulfil those objectives, they must be widely distributed and accessible. They must also continue to be available for reference in perpetuity and facts reported in them must be able to be cited. The system of referencing developed within the scientific literature is an essential element in the development of knowledge, the cross-checking of information and the prevention of effort-wasting by re-inventing wheels. If the citation system (particularly that referred to as the ‘Harvard system’) is to survive, the reliability of pagination in publications must be maintained. It is essential that every page of a journal (and it would be helpful if each page of a book or monograph) has, within a header or footer, its title and volume and part or issue numbers, as well as page number. (This ensures that if only a page, or a few pages, of a publication is copied or printed out, details of its origin are not lost.)

From mid-2015, the standard format for a digital publication file is Portable Document Format (PDF). Whichever word-processing or layout program is used to put the text and graphics of the publication together, it should be saved as a PDF file. This will preserve its content, pagination and appearance, no matter what computer or other device it is subsequently viewed on. Properly prepared PDF files contain the fonts they need and preserve the exact appearance of the document. Hard copies of the publication can also be produced from the PDF at any time, either in large numbers by a printing house or as single copies at a library or privately. The PDF is now so widely used that one can be certain that when it is replaced by an improved format in the future, provision will be made for PDFs to continue to be read.

It will facilitate all aspects of document exchange, whether as hard copy or digital file, if the following points are also fully considered in publication design:

Pages should be of a standard size that can be economically printed and easily viewed, which means no larger than A4 or US letter size. Smaller sizes can be used but paper will be wasted when the publication is printed on standard paper.

Margins on the binding edge of each page should be adequate, so that important information can still be read in a printed and bound, or at least stapled, copy. Maps and photos should not be run across a centre-fold. If photographs are “bled” (run right to the edge of the paper), nothing vital should be placed right at the edge, because laser printers always leave a small blank margin.

Page numbers and headers or footers that identify the publication should be at least 5 mm or a quarter inch from the edge of the paper. If these rules are not followed, a reduced version of the publication would have to be printed, making type and everything else more difficult to see. (Anyone with the ability to print to only one of the standard paper sizes may have to reduce material designed for the other in any case, as A4 is taller than US letter, and letter is wider than A4.)

Graphics (maps and diagrams) in colour should be designed so that they can still be understood if printed in black and white (or at least greyscale). It would be helpful if points or lines in different colours could also be distinguished by different symbols or dots and dashes. Maps should always have a graphic or bar scale or a caption stating the spacing of grid lines or marks, not a statement like 1:500.

Resolution in graphics should always be set as high as practicable. 600 DPI is ideal for black and white; 300 DPI is adequate for greyscale and colour. Higher resolutions than those will increase the final file size, with little benefit. There is no need to resample graphics to exactly those numbers. Any printing or display software will do its own re-sampling to whatever the actual resolution of the hardware is.

There are many ways to convert a document into a PDF file. Most of them use, directly or indirectly, Adobe Acrobat, but there are other options. Generally there are various choices as to just how the conversion is done, and the default choices are not always the best ones. Make sure that fonts are embedded, that the conversion to PDF does not change the resolution (DPI) of the graphics or at least does not reduce it below 600 DPI for strictly black-and-white material or 300 DPI for greyscale or colour, and that the conversion does not use JPEG compression if there are any greyscale or colour illustrations other than photographs. ZIP compression is acceptable.

A PDF file format that is standard in the printing industry is PDF/A. Files that meet that standard include the fonts, which are important to an editor. They should also contain information to make possible very accurate colour rendition by offset presses, which an editor may not regard as particularly important. Even if a file is not PDF/A, the fonts should at least be embedded. (Very rarely a font cannot be embedded for copyright reasons. In such a case, another font will need to be used and the layout revised. None of the fonts included with any home-computer operating system should have this problem.) Some old printer drivers might have trouble with a PDF file that contains advanced features such as transparent overlays or shadowed text, which should be avoided.

2. Exchange of digital publications

2.1 Posting hard copies

The rising cost of printing and postage (especially international postage) has led many groups to look for more economical ways of carrying out exchanges. A PDF file or collection of files could be mailed to exchange libraries on disks, but CDs and DVDs are rapidly becoming obsolete technology; some home computers already are not equipped to read them. Packaging disks and posting them still costs money, though not as much as printing and mailing paper copies. There is also a question as to the longevity of optical disks.

2.2 Utilising e-mail

Small publications or significantly compressed ones can be sent as digital exchanges by simply attaching them to e-mails. This is fairly convenient to all parties, especially as all the receiving librarian has to do is save the attachment. The quality of the publication may be severely degraded, however, and this technique would not work for large magazines or books.

2.3 Utilising the Internet

A high-quality PDF file of a publication may well be hundreds of megabytes in size. It will almost certainly be too large to e-mail as an attachment. A much smaller PDF might also be made with reduced DPI graphics, more compression, etc., for distribution only for on-screen reading like a web site, but the best possible file should be the one archived and distributed to exchanges. (A receiving librarian may choose to compress the file for storing in an electronic archive, if required.)

There are also various ways to put a PDF file on the Internet where exchange librarians can access it. Many are free unless you exceed specified limits (often ~10GB). Microsoft Cloud, Google Drive, iCloud, and Dropbox are examples. If submitting publications to the Karst Information Portal (see below), it could serve as a cloud source for other librarians, although there might be a delay while it is processed by KIP. Another option is to put the PDF file on your own or another organisation's web site. Note that it can be put there even if public access is denied by giving it a file name that no link points to. This will prevent anyone (including Google) who is not given the exact URL of the file from accessing it.

Exchange librarians then need to be sent e-mail messages that tell them the issue is available and contains the link they need to click to download it. Exchange librarians, who may be dealing with dozens of exchanges in their spare time, *should not be expected to check a number of sites periodically for new material*. Most on-line storage schemes have a way to send a link that does not require the recipient to wrestle with a web site, but just to click instead. The PDF file should be kept available at the stated link for a period of months, because the librarian is probably a volunteer with limited time and no vacation backup.

3. Storage of digital publications

Recipient librarians need to make some decisions regarding storage of exchanged (and, indeed, their own organisation's) digital publications.

3.1 Print and store hard copy

Some librarians may prefer to continue to maintain collections of hard-copy newsletters and journals. They would have the option of printing publications out from the downloaded PDF files (preferably on a colour laser printer using ISO 9706 acid-free, archive-quality paper, as the inks used in ink-jet printers are not regarded as sufficiently stable for archival storage). With increasing use of digital publication, this process could require considerable amounts of time, printing resources, and physical storage space. Additionally, there remains the problem of printing the covers and then binding the document.

3.2 Save digital copies

It is considerably easier, quicker, and cheaper to save copies of the digital files than to print them out.

Removable optical disks (CD-R and DVD±R). These are cheap, inexpensive, and reliable. Properly cared for, the data written on a CD or DVD will likely "outlast the technology." But this may be already happening, as some new computers are not equipped to read them. Another disadvantage of such disks is that they cannot be updated once written to, by, for instance, adding a new issue of a periodical. (Never use rewritable CD-RW or DVD±RW disks; they are not as reliable for long-term storage).

Hard disk drives. These are currently the fastest and easiest way to save digital files. They are, perhaps surprisingly, cheaper per gigabyte than DVDs, with prices continuing to fall. Because they contain moving parts, they are subject to mechanical failure and cannot be counted on to last indefinitely, even if they are not powered most of the time. They are also subject to technological obsolescence, as electronic interfaces or file-system software evolve. Hence, re-copying the information needs to be planned for the future.

USB sticks. These solid state devices can now store 64GB of data and should be more reliable, but are easily lost.

The "cloud". Many commercial services, some of which are mentioned above, are available for storing digital information and accessing it over the Internet. They are not free, but the cost is low and falling rapidly. Store material can be accessed by multiple people from multiple locations, as long as they know the password. These services also have the advantage that they will deal with matters of technological obsolescence in hardware and file systems themselves, and they are reliable because they do their own backups to distributed locations. However, this assumes that subscriptions continue to be received and that the company stays in business.

The Karst Information Portal (www.karstportal.org) could also serve as a cloud facility, but it is oriented toward making everything publicly available on the web, although this policy is evolving somewhat. It requires permissions and other agreements to accept new material. The International Union of Speleology is one of its partners. It is connected with a university library in Florida, USA. It is free, but it is not run by volunteers and long-term funding may not be assured.

3.3 Backups

In any case, a single storage medium is never sufficient, and there should always be at least one backup, with provisions to keep the copies synchronised by updating the backup as changes to the collection occur. Two or more hard disk drives in different locations or hard disk(s) with cloud backup are recommended. As with hard copies, some process of cataloguing accessions is needed, to keep track of publications received, even though searching through digital media is extremely fast and reliable. The catalogue itself, of course, should be backed up along with the data, and it should be made publicly available on the web, so that researchers can discover the receiving organisation as a source for the material.

This guideline document has been drafted by Greg Middleton and Bill Mixon for the International Union of Speleology's Publications Exchange Working Group in April 2016. Comments, suggestions for improvement or clarification, and critical input are invited from interested parties. Please address all correspondence to the authors or to the Chairman.