What is single source publishing?

Paul Hoadley, Logic Squad

The short answer is that single source publishing is exactly what you would expect: the publishing of information from a *single source* into a *variety of different output formats*—for example, HTML for the web and PDF for print. The aim of this paper is to briefly explain the process from a high level viewpoint, and to remain as concrete as possible. At this point, the technical details are not as important as the overview. At the same time, though, it is of paramount importance to know exactly what we can achieve, not just what might be possible in an abstract sense.

An overview of the process

This section will be the only abstract section, and we will keep it brief. If you are reading this paper in an attempt to decide whether single source publishing has anything to offer you, there is no point in starting from theory. We will proceed to discuss concrete, practical issues in the very next section.

Single source publishing is an umbrella term for any process that starts with a single source of information and leads eventually to the production of multiple different formats of output. In practice, though, the source of information is a document written in a language called XML. The acronym XML (which stands for eXtensible Markup Language) is very much a computer industry buzzword, but essentially just refers to a format for describing the structure and meaning of an otherwise nebulous lump of text.

The only important concept to retain at this point is that XML imposes no formatting, typesetting or design constraints on the information. This is the key. The way we present a body of information on paper may be significantly different to the way we present it on a screen, for example, so there is no benefit in describing the presentation until we are just about to present the information.

The information stored as XML is then ‘transformed’ into other formats, and during the transformation presentational features are added. For example, a heading might be set in 14pt Times Roman for a web page, and 24pt Helvetica in the PDF for printing. This is all decided once, and in advance. The XML is then just transformed according to a set of rules into some concrete presentational format.

What output formats can be produced?

The general output formats include HTML (for web pages), PDF (for printing and viewing on screen, including slideshows for presentations), PostScript (for printing), RTF (for printing and importing into **Microsoft Word**), HTML Help (for pop-up help applications) and even plain text (for emailing). In theory, any output format for which a suitable set of transformation rules can be written could be supported.

What documents are suitable?

Virtually any kind of document is suitable, though some types of work will benefit from the model more than others. For example, highly structured documents are well suited to this process:

- articles
- essays
- manuals
- documentation for:
    - projects
    - computer software and hardware
    - processes
- operating instructions
- policy documents
- guidelines and recommendations
- rules and regulations.

What documents are unsuitable?

The rule of thumb is that if your work involves a high degree of very visual-oriented design *that is dependent on the content*, there are probably better tools. This is not to say that the output of the single source publishing model involves no elements of design, just that the design involved is strictly separate from the content. What does

all this mean? If you are using **PageMaker** to design a shopping catalogue where lots of different images and small snippets of text are being placed together in arbitrary (though probably visually appealing) ways, then you are probably already using the best tool. If you are using **Microsoft Word** to run off a couple of memos a week that don't even need to be saved, you are probably already using the best tool.

How do I know if I could benefit?

The warning signs are fairly straightforward:

- If you are using **Microsoft Word** to edit long, highly structured and complex documents that need to be published to PDF, paper and the Web, you should consider single source publishing.

- If you are maintaining your group's informational publications using **Microsoft Word** for the printed format, and a HTML editor for the Web, you should consider single source publishing.

- If you are starting a new project and know that you will need to provide information in a range of formats, but can't see how to keep it all coherent, you should consider single source publishing.