February 2005

People Profile: Jake Wallis

Editor
Policy: Where characters in ordinary words can easily be represented by keyboard characters, e.g., the diphthong as can easily be typed as so, then the plain text version is used. However, in place names, personal names and other proper nouns, the original form of the word is retained and represented by using HTML entities. Although this makes the names difficult to search for, this is usually compensated for by including the names in an index.

9. Typefaces and typographic conventions

Issue: Books may use typefaces that are not available to Web users. Many older books use typsetting conventions that do not translate well to online publication, e.g., upper-case headings, full stops after headings, and open quotation marks at the start of every line within long quotations.

Policy: Any use of bold, italic or underlined text is retained in the eBook. However, the font used for printing is not retained. External stylesheets are used to specify fonts currently regarded as suitable for online usage, such as Trebuchet or Verdana. Typographic conventions that were common in older books are regarded as artefacts of publishing that do not need to be preserved in eBooks that use more modern formatting conventions. It is therefore considered acceptable to change the case of headings, punctuation in headings, and the appearance of quotations, provided the text itself is unchanged.

Image Management Issues

10. Image file format

Issue: A compromise is required between image quality, file size and usability.

Policy: JPEG format is considered adequate for eBook publication, but if digital preservation is important as well as online publication, images are archived as lossless TIFF files and converted to JPEG for publication. A resolution of 300 dots per inch is often used for preservation, but lower resolutions of 150dpi or 200dpi are adequate for some types of image, e.g., line drawings.

11. Image captions

Issue: Images in books may or may not have captions.

Policy: Any existing captions are faithfully reproduced, though not necessarily in exactly the same position in relation to the image, and are used as the Alt text for `<img>` tags. If no printed caption exists, one is created for use as Alt tag text, to meet accessibility criteria, but is not displayed on the published page. Lists of captions may be used to create a linked list of illustrations.

12. Image positioning

Issue: Images in books may be printed on a different page or section to the referring text.

Policy: Images can be moved. It makes more sense to include an image in the appropriate place in an eBook than to faithfully reproduce its original position in the book.

Publication Issues

13. Content structuring

Issue: Books may or may not be conveniently divided into chapters and chunks of a suitable size for online access (Wilson & Landoni, 2002).

Policy: Wherever possible, the natural structure of a book (chapters, sections and paragraphs, but not pages) is used to determine the structure of the corresponding eBook. Similarly, structures such as lists, tables, notes and quotations are retained in the eBook. However, in order to produce Web pages of reasonable length, eBooks may have a finer structure than the original, with additional section or paragraph breaks inserted if necessary. If a book has a running heading, applying to one or more pages, this may be adopted to function as a section heading. New section headings may be inserted as a last resort if existing ones are misleading, and should be identified to users as a component of the eBook rather than the original book. In some books an extensive index helps compensate for inadequate section headings.

14. Contents pages

Issue: Printed contents pages may include different chapter or section headings from those used in the body of the book.

Policy: The methodology for eBook creation involves automatic generation of contents pages that link to the relevant headings used within the book. This ensures consistency and functionality, but may mean that discrepancies in the original contents pages are lost.

15. Notes

Issue: Many books have footnotes whose location is dependent on specific pagination.

Policy: It is more important for notes to be associated with the referring text than to appear in exactly the same position, or with the same number, as in the original book. Footnotes are therefore converted to endnotes, appearing at the end of eBook sections. If necessary the symbol used to identify notes is changed to a unique character, to enable automatic linking from note to reference. Notes may also be renumbered to ensure uniqueness.

16. Cross-references

Issue: Books may include internal cross-references, to a page or image in the same book, or external references to other publications.

Policy: Links are a fundamental function of the Web, and eBooks should include them where appropriate so that they function properly, even if this requires making minor changes to wording of the text. References to page numbers may therefore be reworded to refer to sections in order to work effectively as links. Similarly, text references to images 'opposite' or 'on the next page' may be changed to read 'above' or 'below' etc., so that the text makes sense in an online context.

Cataloguing and Indexing Issues

17. Title tags and metadata granularity

Issue: In order for eBook sections to be

continued on page 24

<http://www.against-the-grain.com>