

Word to Web publishing for Agricultural Research

Roger C. Johnson

The Regional Institute Ltd

email rjohnson@regional.org.au web site <http://www.regional.org.au/>

Abstract

This paper describes a project to make available, via the Internet, the complete proceedings of The Australian Society of Agronomy. As part of a larger collection of agricultural research at regional, national and international scales the papers will provide a valuable educational resource for the wider agricultural community. The project addresses the political and technical issues associated with making agricultural research information that is, or ought to be, in the public domain available via the Internet for free.

Introduction

The first known treatise on agriculture (De Re Rustica) was written by a Roman soldier and farmer, Lucius Junius Moderatus Columella, in the first century A.D. The original Latin text is on the Internet (2) .

The first agronomy conference was held in 1980, the same year that IBM introduced the personal computer. In the past 20 years, advances in agricultural research have been rapid. Yet apart from the existence of various decision support systems, abstract databases and electronic journals, the computerisation of agricultural research information is poorly developed. The standardisation and computerisation of Australian law, by contrast, began in 1977 with the full text now freely available via the Internet (7) .

The primary output from agricultural research is improved production. The information derived from the research is a by-product for which there is no collective responsibility. As a result, the output from a generation of agricultural researchers has been dissipated through numerous industry or discipline based extension pathways, causing extreme fragmentation of the knowledge base.

Two years ago this conference addressed the environmental imperative for agriculture and the need to equip people as agents of change to understand and manage complex systems. Pratley (12) called on agronomists as a profession to accept these challenges.

These challenges give rise to the information imperative to improve the way agricultural research information is gathered and disseminated. To achieve this aim, a commitment is needed to overcome the social barriers and organisational politics and policy restrictions that inhibit the free flow of public information. Copyright is a key issue. In Australia, copyright protection is free and automatic. The copyright policy of many research organisations is more concerned with legal liability and protecting commercial interests than maximising public access to research findings. A study of journal pricing recommended that:

Academics including researchers must be aware of the implications of their publishing patterns and the cost to institutions and their readers. These include the possibility of not supporting journals ... which have excessive costs, journals which take copyright unto themselves for profit, and journals which restrict access to information by high use or subscription costs (3)

The fundamental idea of free access to a substantial part of the agricultural research knowledge base is behind this ambitious project to make all ten proceedings available via the internet. It is inspired and supported by the Australian Society of Agronomy and Charles Sturt University who first published the 1998 proceedings on line. In overcoming the policy issues these organizations provide leadership in supporting research into the computerisation of agricultural research information.

The Regional Institute – more than just a web site

The Regional Institute was established in 1999 as a not-for-profit, educational and research forum for information and ideas of importance to regional Australia. Its web site, modelled on AustLII, a full text legal web site (www.austlii.edu), publishes information that is, or ought to be, in the public domain via the Internet for free. It presents a contributory model that shows, by example that: research information can be made easily accessible by the public; the massive task of building such a resource can be distributed amongst the authors; and, by enabling most of the publishing effort to be automated, access to the information can be provided for free with moderate public funding.

Building the knowledge base.

We quickly made two observations. Firstly, most authors use Microsoft Word. Word is a proprietary, but ubiquitous standard for electronic document creation. Secondly, most user classes, such as farmers and agronomists have a telephone, which means they have access to the Internet (cost and bandwidth in rural areas are controversial but separate issues). The internet is based on open standards (www.w3c.org). What happens in the middle is very messy. We began by trying to understand how Word works.

Styles and templates in Word

Two of the most important but least understood features of Microsoft Word are styles and templates. Styles and templates are the essence of Word, the inner logic of the whole program (11). A style is a combination of formats (font size, type face, tab stops etc.) that are stored as a function. When a style is selected the formats are applied to the text. Styles are structural elements and define the logical organisation of the document. Styles can be stored in a template (8) and applied to other documents.

As a minimum requirement, authors should use the default styles in the normal template (Figure 1a) and always check the logical organization of the document using the outline view (Figure 1b).



Figure 1a: Style elements

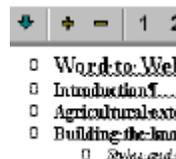


Figure 1b: Outline view

Understanding styles, and following basic guidelines for the formatting of text, tables and images is the key to distributing the effort of building a critical mass of agricultural research information. Microsoft Word provides us with a common unit of exchange for information that meets technical and usability standards.

Case study – Agronomy Conference Proceedings 1996 to 2001

The transition from paper-only to fully automated on-line publishing is illustrated by the Agronomy conference proceedings from 1996 to 2001. The papers from the 1996 conference were the earliest available in electronic format. The initial (and most time consuming) task was to convert the source documents to a standard Word format. From this exercise, a methodology for automating the task of publishing future proceedings would emerge.

The authors' original versions of the 1996 papers on some 400 floppy disks had been meticulously catalogued. Multiple revisions of many papers were also kept in a variety of formats including early

(Windows 3.X) versions of WordPerfect and Word with a variety of embedded graphics. Sorting and loading the disks alone took five days. A standard word template with macros to automate page layout and formatting was applied to each document in turn. The documents were marked up by applying styles to the key structural elements including Title, Heading 1, Heading 2, author, keywords and references.

The 1998 papers had previously been published on the Web. The authors' original electronic documents no longer existed, having been discarded after conversion to Ventura, a proprietary electronic publishing format. The web pages were saved as Word files and marked up as for the 1996 papers. The 2001 papers were provided on CD ROM. As for the 1996 papers, a standard template was applied to each paper in turn, prior to marking up. The collection of 850 source documents in standardised Word format was now ready for publishing to the web.

The electronic publishing system – a fruit packing process

A bottom-up approach to building the web site is made possible by using an electronic publishing system to convert the source documents from their native Word format (Figure 3a) to a complete web site. The HTML documents were published to a standard directory structure (Figure 3b,c and d) by applying templates for a consistent look and feel and to incorporate navigation features. Graphic design is done last using basic colours and existing logos to establish the identity of each part of the web site.

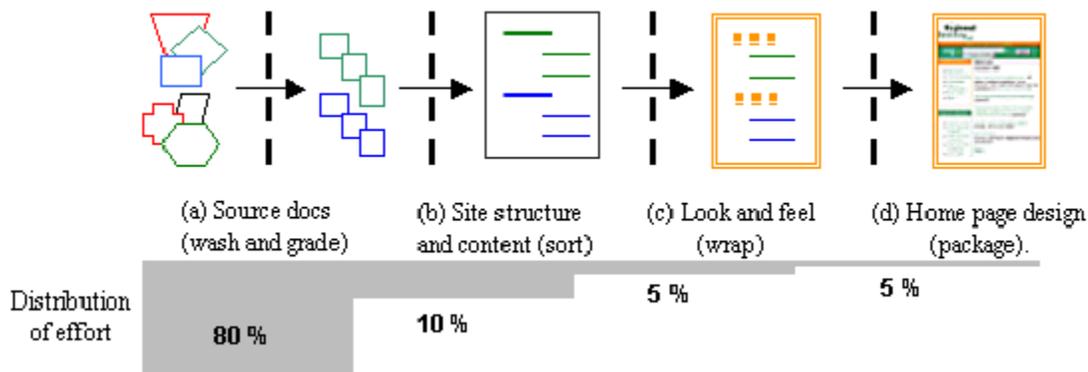


Figure 3 Electronic publishing as a fruit packaging process

The electronic publishing approach means that 80% of the effort of building a complex, content rich web site can be shared by the authors who benefit from more efficient internal document management techniques. It avoids the need to use document database systems which have poor usability and overcomes the structural problems associated with sites built with web page authoring tools like Microsoft Front page.

Web site design: accessibility and usability for agronomists

Since the purpose of the web site is to make information accessible, it is designed for simplicity of form and function (1) ,(10) . The electronic publication of a large collection of documents requires organization. The electronic publishing system provides a user centred approach (4) by allowing the site content and structure to be determined by the nature of the source material supplied by the authors who are also users. The simple (Unix) web site directory structure provides stability (the URL is constant) and scalability (new collections can be added without changing the established URL concordance).

Heuristic usability testing is applied to the site based on criteria of simplicity, consistency and familiarity. The electronic publishing system enables each collection to have a similar structure with navigation buttons automatically inserted in each page to aid browsing. Internal tables of contents can be automatically generated for large complex documents and global links inserted as a header/footer. Frames are not used because they do not provide transparent links. Proprietary delivery formats like

Word and PDF are avoided because on the web, their accessibility, searchability and hypertext functionality are poor.

The resulting web site is an integrated electronic library of publications, each having the usability features of a book, including table of contents, author and references index and an “electronic finger” or search function. Meta data (<http://purl.oclc.org/dc/>) and named anchors imbedded in the HTML are generated automatically. Standardisation of the source material in Microsoft Word 2000 format and an XML compliant publishing system (<http://www.w3.org/>) mean that any part of the web site can be republished from the source with little effort. Those involved in education or extension are able to add value to the material by linking from their own web sites to specific paragraphs and creating stored searches of the individual publications.

Conclusion

Agronomists play a pivotal role in information exchange but need to adopt best practice principles and view this role in terms of processes rather than outcomes (12) .

Organisations play an important role in extension policy. A review of the communication efforts of rural industry research and development corporations (9) called for an approach which allows the continuous flow of information across all levels of rural industry management; and communication strategies that not only facilitate the rapid exchange of information, but also create dynamic social learning processes.

The electronic publishing approach used in building the public knowledge base overcomes technical and political constraints. The author of the information becomes the publisher and retains copyright of the source material. The organization retains copyright where applicable of the collection and the agricultural community and the public have free access to the information. No collaboration is required and anyone is free to contribute.

Once the knowledge base is established, we can make it more intelligent by incorporating decision support tools and improved indexing and text retrieval mechanisms tailored to our known user classes. Finally we must connect it with people. Synchronous text, audio and video communication are ubiquitous on the web and developments in Internet based groupware will provide new opportunities for agricultural research and extension in cyberspace. Dreher (5) describes ‘swimming’ in a vast sea of potentially useful information. We must explore new ways in which hypertext can be used to support and empower human cognitive activity.

Embracing these challenges requires a significant shift in thinking. Best practice compels us to look outside agriculture for models to suit our needs. The criteria we set are firstly, independence and secondly, to use the power of information technology to bring information and people together in a new and innovative way.

References

- (1) Ausinfo (1999) Guidelines for Commonwealth information published in electronic formats http://www.dofa.gov.au/ausinfo/guidelines/guidelines_for_commonwealth_in.html
- (2) Camden, D. (2000) Forum Romanum. <http://www.geocities.com/Athens/Forum/6946/literature/>.
- (3) Cornell University (1998) Journal Price Study - Core Agricultural and Biological Journals. Cornell University <http://jan.mannlib.cornell.edu/jps/>.
- (4) De Troyer, O. M. F. and Leune, C. J. (1998). WSDM: a user centred design method for WEB sites. Proceedings of WWW7 Conference, April 14-18, Brisbane, Australia. <http://www7.scu.edu.au/programme/fullpapers/1853/com1853.htm>

- (5) Dreher, H.V. (1997) Empowering human cognitive activity through hypertext technology. PhD Thesis. Curtin University of Technology, Perth Western Australia
<http://john.curtin.edu.au/theses/public/adt-WCU2000046.121219/>
- (6) Green, D (1999) ITC130: On-Line Publishing. <http://life.csu.edu.au/itc130/>
- (7) Greeleaf, G., Mowbray, A. and King, G. (1996) AustLII – Technology and politics of law on the net. 4th International Conference on Substantive Technology in the Law School and in Law Practice. Univ of Montreal, Quebec, 3-6 July 1996.
- (8) Johnson R.C. (2000) Author guidelines for Word to Web Publishing. The Regional Institute Ltd.
<http://www.regional.org.au>
- (9) Keen, M. and Stocklmayer, S. (1999) Communicating Research: An Overview of Communication Efforts of Rural Industry Research Funding Bodies. (RIRDC Publication No 99/54. MS Word)
<http://www.rirdc.gov.au/reports/HCC/ANU-26A.doc>
- (10) Lynch, P. and Horton, S. (1997) Web Style Guide: Basic Design Principles for Creating Web Sites
<http://info.med.yale.edu/caim/manual/>.
- (11) Maardt, M. (1998) Get Going With Word. e-Ways Pty Ltd www.knowwareglobal.com
- (12) Michalk, D.L. and Pratley, J.E. (1998) Foreward *Proceedings 9th Australian Agronomy Conference*.
- (13) Williams, R.T. and Walcott, J.J. (1998) *Proceedings 9th Australian Agronomy Conference*.