

The Agricultural Document Library (ADLib) and its applications

J. Tzilivakis ^a and K. A. Lewis ^a

*a Agriculture and Environment Research Unit (AERU), Science and Technology Research Institute (STRI), University of Hertfordshire, AL10 9AB,
J.Tzilivakis@herts.ac.uk, K.A.Lewis@herts.ac.uk*

Abstract

In April 2004 the Agricultural Document Library (ADLib) was launched (www.adlib.ac.uk). It has been developed jointly by the Agriculture and Environment Research Unit (AERU) at the University of Hertfordshire and TLR Everysite Ltd, as a primary electronic resource for the industry. This online system is not just a library of documents, it is a knowledge base that is being used to support numerous applications and web sites. The content of ADLib covers England, Scotland and Wales and works across the industry for farmers, advisors, retailers, educators, policy and planning organisations. Examples of how ADLib is being utilised are described within this paper. These include libraries for Environmental Management for Agriculture (EMA), the Fertiliser Advisers Certification and Training Scheme (FACTS), the Horticultural Development Council (HDC) and the British Potato Council (BPC). Additionally, ADLib is being used by the Department for Environment, Food and Rural Affairs (Defra) to support its Whole Farm Approach (WFA), via the Appraisal, and Assured Produce utilise ADLib to support its online self-assessment audit. The principal objective is to make the vast quantity of agricultural information that is available more easily accessible to farmers and growers and to ensure that information is consistent, up to date, timely, relevant and targeted. In an information age with rapidly changing societal demands, agricultural businesses need the tools and information to be able to respond. ADLib aims to provide the agricultural and horticultural industries with the information they need to push forward the sustainability of the industry, in line with evolving public desires and legal demands.

Key words: ADLib, Knowledge Transfer, Sustainability

1 Background

The origins of ADLib can be traced back to 1994 when the development of the award winning Environmental Management for Agriculture (EMA) software began (Lewis and Bardon, 1998). This software was funded by numerous organisations including the Ministry of Agriculture, Fisheries and Food (MAFF, now Defra), Scottish Executive, Milk Development Council and Horticultural Development Council (HDC) and was developed in collaboration with ADAS, Rothamsted, Cranfield University and Central Science Laboratory. The software is comprised of a suite of auditing and decision support tools to aid farmers and advisers develop environmental management systems and/or improve the environmental performance of a farm. At the outset of the development it was recognised that related documentation and literature was vital to support the auditing and decision support systems. So a process of converting key best practice documents into an electronic format was started, initially with the three codes of good agricultural practice for air, soil and water (MAFF, 1998 a, b and c). This conversation process, although now substantially different in technique, has not stopped since 1994.

The initial documents were created in Windows Help file format (HLP). This involved creating rich text file (RTF) documents, with hidden text used to create hyperlinks, and then the HLP files created using a DOS compiler. In 2001, the library had grown to such an extent that the HLP system was no longer viable. So a programme of work was begun to upgrade and convert the entire library to Windows Compiled HTML (CHM) format. This was completed in 2002 and was launched with the EMA 2002

software. The EMA software has been in circulation amongst the farming industry since 1998. During that time it became apparent that one of the most valued components of the EMA software was its underlying reference library. Having all the documents in one place was valuable in itself, however the cross-referencing via hyperlinks considerably enhanced its value. By the time EMA 2004 was released there had been many demands to make the library available online. Time and technical constraints had prevented this from happening before 2004, but during 2003-2004 the EMA development team had forged links with a local IT company, TLR Everysite Ltd., who had the capability to convert the CHM format of the existing library into an HTML format web based system. This process was completed early in 2004 and ADLib was launched in April 2004 containing 250 documents.

2 The Base ADLib Resource

2.1 The Content

The base ADLib resource now contains over 1500 individual documents (~30000+ pages) (as of March 2007) including Government codes of practice, industry guidelines, information and topic sheets, as well as a vast amount of supporting material, e.g. legislation/regulation summaries, glossary items, photograph galleries and a contacts directory. A few of these documents are not available anywhere else, either due to them being out of print, no electronic version elsewhere or both. All the documents are held on the ADLib server, i.e. ADLib is not a web portal providing links to documents on other web sites. All too often when a document is published on a web site at some point in the future the URL where it is located is changed and thus any external links become broken. As ADLib holds its own versions of the documents, this ensures that when a user clicks on a link for a document, that document exists and is delivered to the user.

The documents are held in a content management system in two formats (where possible), Complete HTML and PDF Versions. 80% are available in HTML format (as well as PDF). The advantage of this version is that documents can be cross-referenced using hyperlinks to individual pages or paragraphs ('deep-linking'), thus taking the user to the most relevant information within a document or related documents. There are also extensive searching and book-marking facilities, including a new keyword search facility that will search down to individual pages. By having all this information in one place and in a common electronic format, it becomes a very powerful and flexible knowledge base that can be utilised in a number of different ways. Firstly it can be simply used to construct various electronic libraries (section 3) designed to suit the needs of different user groups. Secondly it can be used to support other more interactive applications, such as audits and appraisals including those of third parties (sections 4 and 5).

2.2 Management, Maintenance and Quality

Only the base ADLib resource has to be kept up to date. When a base document is amended all libraries and applications that contain that document are automatically updated in real time. This ensures that all users have access to the same current version of the document. Additionally, ADLib is managed with a strict set of protocols that ensure all the information is reviewed and updated on a regular basis. New advice and guidance documents from over 40 organisations including:

- Agricultural Industries Confederation (AIC)
- Assured Food Standards (AFS)
- Crop Protection Association (CPA)
- Defra
- English Heritage
- Environment Agency
- Farming Wildlife Advisory Group (FWAG)
- Food Standards Agency
- Forestry Commission
- Health and Safety Executive (HSE)
- Home Grown Cereals Authority (HGCA)
- Horticultural Development Council (HDC)
- Meat and Livestock Commission
- Milk Development Council (MDC)
- National Assembly of Wales
- Natural England
- Potash Development Association (PDA)
- Royal Society for the Protection of Birds (RSPB)
- Scottish Environmental Protection Agency
- SEERAD
- Scottish Natural Heritage (SNH)
- Veterinary Medicines Directorate (VMD)

Locating documents can be quite difficult. Sometimes published material can be deeply buried within web sites, is available in various formats, old versions of documents are not always removed and occasionally links to a document can be broken, as the document has been moved. As such ADLib has an ongoing process in place where the above organisations, and many others, are searched on a daily basis to locate documents and when a document is obtained it is checked to ensure it is the most current version. Also some organisations notify ADLib directly when they publish a new document to ensure that it becomes immediately available.

ADLib have established relationships with all these organisations and agreements are in place handling the copyright and intellectual property right issues. This is an important issue as it allows full and proper duplication of the documents. In the past copyright restrictions resulted in documents having to be interpreted by different organisations, for publishing purposes, often resulting in different interpretations and thus different messages communicated.

Upon receipt of a new document, the PDF is placed in ADLib immediately and then the whole document is converted to HTML. This is usually completed within 1 day to 3 months depending on the size of the document. Once converted to HTML, the document is then reviewed by a second person (independent of the person who created it) to ensure it is correct and relevant links have been added. Then every document in ADLib is reviewed annually to ensure it is still current and whether any new links can be added to any material that has been added to ADLib in the past 12 months.

All the documents in ADLib are automatically scanned on a weekly basis to extract new keywords for the taxonomy that is used for the search. A taxonomy editor is then used to approve new terms and discard others to a 'stop' list. The stop list is used to ignore terms during the keyword search and next time ADLib is scanned for new words. The 'live' taxonomy consists of over 50000 words and search terms that are directly linked to all the pages in ADLib where they exist. Each search term can also have equivalent terms associated with it, so a user can open a second level of search (by clicking on 'See Also' in the search results) to view related pages to their original search term. For example, a user might type in a common name for a disease, such as 'sudden oak death'. This term will have an equivalent term of 'Phytophthora ramorum', the scientific name for the disease, which will be presented to the user under 'See Also', thus directing them to pages where only the scientific name is used.

A number of management tools are also in place to check that links between documents are maintained and are not broken or link to incorrect or old information. For example, when a code of practice is updated with a new version, the old version is archived. A tool is then used to identify all the links to the old archived version from other documents in ADLib and these are then updated to link to the new version.

When documents are added to ADLib they are also audited to assess their quality in terms of their content, presentation, relevancy to the industry and how up to date they are. This helps us identify those documents that of the highest quality and thus likely to be key documents for providing advice.

3 ADLib Libraries

A number of different libraries have been constructed, each with its own particular focus. The general public can subscribe to four of these. These are the EMA Library; Cross Compliance; Crop Nutrition and Soil Management; and Pest and Disease Management libraries (the subscription currently ranges from £25-50 per annum). These libraries also contain some online and downloadable tools such as the pesticides database and the RB209 fertiliser recommendation calculator, developed from the EMA software. There are also libraries exclusively for members of the Horticultural Development Council (HDC) and British Potato Council (BPC). These libraries contain HDC and BPC publications complimented with some publications from the base ADLib resource. People who are members of FACTS (Fertiliser Advisers Certification and Training Scheme) can access a nutrient management library supplied by ADLib as part of their website (www.factsinfo.org.uk). Finally, there is also a library available to members of Assured Produce which has been developed using documents related to their self-assessment audit (see section 5). The contents index (including document summaries) for most of these libraries can be browsed and searched for free on the ADLib website (www.adlib.ac.uk), but access to the actual document content is restricted to subscribers only.

The documents within each library are organised using a section index structure. Typically this will be a subject-based structure, into which the documents are placed – equivalent to placing different books on different library shelves according to the subject matter. In some instances documents will cover a range of subjects, in which case they are placed in a number of different sections. There may also be other sections within a library, for example 'What's New' or 'Newsletters', to highlight recent additions to the library or a specific type of document. Each library also has book-marking and search facilities. The book-marking facility allows the user to store the location of particular documents or pages, so that next time they login they can browse their bookmarks and go directly to a document or page without having to find them again in the index or by searching. The user can search the library by entering a search term and then look for that term in either the title of the document, its abstract or on a page within a document. The first two of these are fairly simple and only return links to whole documents. The third search utilises a keyword taxonomy that is derived from all the words and phrases on all the pages in the entire library. This taxonomy is edited on a weekly basis (see section 2.2) to ensure it is tailored and focused towards the most relevant terms. When the user undertakes a keyword search the results are a set of links to pages within documents, thus taking the user directly to where the search term exists within the library. If a user has subscription to multiple libraries they can also search one or all the libraries that the user is a member of. This is available at: <http://www.adlib.ac.uk/adlib/search.asp>

4 Defra's Whole Farm Approach

The Whole Farm Approach (WFA) (www.defra.gov.uk/farm/wholefarm/) is designed to lighten the regulatory burden on farmers and growers. It consists of a number of elements including online services, on-farm inspections and the Farm Advisory System. The online services include a web portal and the "Appraisal", an electronic self-assessment questionnaire. This questionnaire utilises ADLib as a means of providing help and advice to farmers. Where necessary each question within the appraisal has some general guidance (level 1) usually consisting of no more than a couple of paragraphs of text. This guidance is accompanied with more detailed information in the form of links (level 2) to appropriate documents within ADLib. These links can be to a whole document or to individual pages or paragraphs within a document. Consequently the combination of level 1 and level 2 guidance provides a highly tailored approach to presenting advice and guidance to farmers. It presents them with the most relevant and up to date information and documents for the issue they are considering at the time. Additionally, the hyperlinks between documents within ADLib means that farmers are not restricted to just the information presented in the level 1 and level 2 guidance, they can explore related information to meet their own specific requirements.

The appraisal was developed and piloted during 2004-2005 and its first official release was in March 2006. The questionnaire has been updated on a regular basis to ensure it covers changes to regulations, such as cross compliance and also to introduce new modules. A second release took place in December 2006 to update cross-compliance and add the December survey module and the next release will be in May 2007 to cover changes to the waste regulations that come into force on 15 May 2007. As ADLib supplies the guidance, WFA also has access to the most up to date versions of documents, providing farmers with a comprehensive and current service. The appraisal is available as both an online questionnaire and on CD and this requires the guidance to be available in two different formats (HTML for online and CHM for the CD). ADLib's has the flexibility to export both the required formats, thus only one set of documents are maintained in the base resource for both versions.

5 Assured Produce

In 2005 a 3-year Defra funded project was undertaken to strengthen the chain for knowledge transfer for members of the Assured Produce (AP) scheme. The aim is to deliver information via ADLib on a 'pull down' basis (i.e. on request of the grower rather than pushed on them by the document producer) by accessing documents relevant to the land-use industries. This saves time and money for growers and increases the uptake of appropriate best practice information. At the same time the AP website (<http://www.assuredproduce.co.uk/>) has been undergoing redevelopment, including the creation of a new online version of the self-assessment audit/checklist. Following a similar model to Defra's WFA (see section 4), this audit is constructed in such a way that when presented with a conformance question, the user is also presented with links to the most up-to-date relevant information in ADLib. For example, a crop protection question might have links to a specific section in the code of practice for using plant

protection products and/or a link to an HDC or BPC fact-sheet on how to control a particular disease. In so doing the user has access to the most appropriate information for the issues they are considering during the self-assessment. The redevelopment of the AP website has also substantially improved the process for developing the crop protocols. The standards are now, collectively, held in a database and tools have been developed for the authors to edit and manage the protocols online, with strict version controls. These protocols can then be displayed online using a protocol viewer or as PDF documents.

As the approach used in Assured Produce audit is similar to that used in Defra's WFA and as they both draw upon ADLib, farmers are presented information in a familiar format and are viewing the same current versions of documents. This consistent approach is important so that a farmer can place trust in the information they are viewing and that they are not presented with different and/or contradictory advice.

6 The Future

It is envisaged that ADLib will continue to grow in all of the areas described above, including the online libraries, its role in Defra's whole farm approach and utilisation within the assurance schemes. There are a number of ongoing and future initiatives in all these areas and some of these outlined below.

The online libraries are continuing to grow in content, in the past year over 180 documents (~3500 pages) have been added to ADLib as a whole. As an expanding resource an important aspect to maintain is for users to have the ability to easily find the information they require. It is all too easy to loose the one page of relevant information to a user amongst the thousands of irrelevant pages. The new keyword search facility that has been implemented in the past year has been a big step towards helping users find the information they require. In the future additional taxonomies could be developed and attached to the documents within ADLib. For example, documents could be 'tagged' with taxonomic terms relating the season in which their content is most relevant and/or the type of agricultural enterprises they apply to and/or the geographical location they are relevant to. If the user then enters information about their farm (a farm profile), search results can be filtered to return those documents that are of most relevance to the user in terms of their location, farm type and the current time of year. Another future development to enhance the online libraries would be the provision of information that is related to the content of a document. There may be instances when users should be aware of other documents they should read when viewing a page of information or, with some older documents, there may be content in a document that has been superseded with more current information (e.g. a change in regulation). We can change hyperlinks to ensure that any links on the page go to the most current documents, but due to copyright restrictions we cannot alter the text of the document without permission. However, what we could do is provide additional notes or links elsewhere on the delivery page, that can provide the user with updates or links to related documents. Such enhancements to the delivery of information could add significant value to the libraries where such facilities are available to the users.

The next step for the WFA, after the May 2007 release, is to begin work on developing a module for Catchment Sensitive Farming (CSF) for release in August 2007. This CSF module will be the first in a programme of 're-basing' the appraisal that will involve a completely new design to the delivery of the appraisal online to make it more user-friendly and concise. There are also plans to improve the user interface of the WFA web portal, part of which will be to deliver some of the benefits of ADLib more directly to farmers and growers. At the moment, users are unable to access any of the guidance until they have started completing the appraisal, which includes a lengthy baseline module. As such the aim will be to create an access point to the guidance closer to entry point of the portal. This may be in the form of an advisory library, similar to those described in section 3, thus also providing valuable book-marking and search facilities for the farmer that currently do not exist within WFA. In the longer term, there are aspirations to integrate WFA with other IT initiatives to present a more integrated experience for users. This includes working with electronic versions of applications for the single payment scheme and entry level stewardship and mapping projects such as the Shared Spatial Information Services (SPIRE) project. It is envisaged that ADLib could be used to underpin such integrated services, so users have consistent access to the most relevant information at any point when using these systems.

The work with Assured Produce is in its final year (starting March 2007) of Defra funding. During this time the self-assessment audit, links to documents and the advisory library are being kept up to date, particularly in relation to any changes in the AP protocols. There are also plans to provide links to ADLib

documents from the online protocol viewer, drawing upon the work done on providing guidance in the audit. An email alert service is also planned that will enable users to register for alerts for different crops. When the standards are changed or updated or when new guidance has been published for that crop, registered growers will be informed directly by email. This would provide another valuable step towards providing growers with the most valuable information that is targeted towards their own particular needs and interests. It is hoped that the work undertaken for Assured Produce will provide a good example of what can be achieved within the context of ensuring food standards and best practice. As a consequence, other schemes that fall under the umbrella of Assured Food Standards (AFS) (Red Tractor) may also wish to develop similar initiatives to provide their members with such a valuable service.

ADLib is a large and valuable resource that needs careful and expert management. Its application so far is relatively minimal in comparison to its potential. The use of ADLib to support applications has in the first instance been quite simple, which in many respects is perhaps the best approach in order to ensure ease of use and greatest value to the agricultural and horticultural industries. There is scope to utilise ADLib in more sophisticated ways and/or to support more complex applications. This is something that is under consideration in the longer-term. In the short-term the priority is to make ADLib and the value it provides to farmers and growers more widely available. It is anticipated that its use will expand as the use of IT grows in the sector and tools such as Defra's WFA become commonplace. With regard to the libraries the ideal would be to make access to them free to all, but this is dependent on some means of central funding. At the moment the basic maintenance costs for ADLib are funded by a combination of research and development funds from government and industry and from revenue from subscriptions. It is hoped that with time and further co-ordination and collaboration between ADLib, government and industry a sustainable means of making ADLib free to all may be found. In so doing it will help the agricultural and horticultural industries evolve in a sustainable direction.

For further information please contact: info@adlib.ac.uk or visit www.adlib.ac.uk

References

- Lewis, K. A. and Bardon, K. S., 1998. A computer based informal environmental management system for agriculture. *Journal of Environmental Modelling and Software* 13, 123-127
- MAFF, 1998a. Code of Good Agricultural Practice for the Protection of Air. Ministry of Agriculture, Fisheries and Food (MAFF) publications, now Department for Environment, Food and Rural Affairs (Defra). PB0618. October 1998
- MAFF, 1998b. Code of Good Agricultural Practice for the Protection of Soil. Ministry of Agriculture, Fisheries and Food (MAFF) publications, now Department for Environment, Food and Rural Affairs (Defra). PB0617. October 1998
- MAFF, 1998c. Code of Good Agricultural Practice for the Protection of Water. Ministry of Agriculture, Fisheries and Food (MAFF) publications, now Department for Environment, Food and Rural Affairs (Defra). PB0587. October 1998