An Information Website for Tree Crops

Producing Oils, Pharmaceuticals, Fodder, Timbers, and Beverages

A report for the Rural Industries Research and Development Corporation

by David Noël
Tree Crops Centre

November 1999
RIRDC Publication No. 99/144
RIRDC Project No. TCC-1A
Foreword

This research project has resulted in a public access website on the Internet which is freely available for public use at any time.

The aim was to present extensive and regularly-updated information on a major group of rural industries for a major world region. The industries concerned were those based on useful perennial plants ('tree crops' for short), and the region was Australasia. This information was to include not only data on the plants involved, which changes relatively slowly, but also information and contact details for all the players involved, from industry, research, government, and commerce — an area in constant flux.

In contrast to most research and development projects, the results of this one can be appraised immediately by users of the website, in that the website is the results. However, this report on paper can still be of value to other classes of user, including those who are not concerned specifically with the actual data within the website, but more with the information and experience gained in the development process. This information and experience may well be applicable and of value to those considering projects to build public-access information sources on the World Wide Web where the information concerned is quite different to the present case, but the processes involved are parallel.

This report, a new addition to RIRDC’s diverse range of over 400 research publications, forms part of our Human Capital, Communications and Information Systems R & D Program which aims to enhance human capital and to facilitate innovation in rural industries and communities.

Most of our publications are available for viewing, downloading or purchasing online through our website:

- downloads at www.rirdc.gov.au/reports/Index.htm

Peter Core
Managing Director
Rural Industries Research and Development Corporation
Note on terminology and glossary

Terms which are underlined in the body of this report are listed with a brief explanation in the Glossary at the end of the report.
Foreword

Contents v

Executive Summary vi

1. Origins of Atcros on Internet (AOI) 1
2. Building the website 3
3. Website design for speed, appearance, and maintainability 6
4. Database to HTML conversion methodology 10
5. Original plan and modifications made 13
6. Conclusions 14

Appendices

1. Brief Glossary 15
2. Examples of calculated text fields and their usage 16
3. Original project proposal 18
4. Interim report (May 1998) 23
Executive Summary

The ATCROS on Internet project (AOI) had the aim of gathering and converting a very large body of data on all aspects of commercial and research exploitation of useful perennial plants and making this data available for general public access on the Internet.

This information was to include not only data on the plants involved, which changes relatively slowly, but also information and contact details for all the players involved, ranging from industry, research, government, and commerce, an area in constant flux.

The plants involved and their products are called, for convenience, Tree Crops. 'Tree Crops' is used in this context to relate to any useful perennial plants. This includes not only conventional horticulture, but also edible perennial-plant oils, timber production, vine crops, plants for land stabilization and improvement, and plant sources of animal fodder, pharmaceuticals such as quinine, beverages such as coffee, essential oils such as eucalyptus, and a big range of industrial products such as jojoba oil, rubber, and cork.

Similarly, the area of focus is called, for convenience, Australasia. While the emphasis is on Australia and New Zealand, the area as defined extends outward to island nations north and east of these, comprising over 300 million people and with a spread of climates and social and agronomic characteristics to match those of any other world region.

The actual scope of AOI is further broadened by two other factors. First, the range of climates available within the region means that a useful plant from virtually any world source can also feasibly be grown here, somewhere. The data pages in AOI concerning useful plants have therefore necessarily had to include ones originating all over the world.

Second, the players, that is organizations and groupings of every sort with an involvement in tree crops, are undergoing the same sort of globalization and communications accessibility as any other fields of human endeavour. So while Atcros directories of plant suppliers are largely restricted to the region itself, tables of seed suppliers, key research institutes, grower groups, government contacts, and highly specialist suppliers and services which may have only a few instances over the whole world, have world-wide coverage attempted. In an increasing number of cases these contacts are accessible via the Internet or the Web, which makes them distance-neutral for those in Australasia with Internet access.

In the three years since this Project was first proposed, general familiarity with and usage of the Internet has increased enormously. For those who are not fully familiar with the jargon involved, a brief glossary and explanation is included as an appendix to this Report.

The initial source of the data used in AOI was two printed editions of 'ATCROS: Australasian Tree Crops Sourcebook' which were compiled by the Tree Crops Centre and published by the West Australian Nut & Tree Crop Association Inc. and by ACOTANC, the Australasian Council on Tree and Nut Crops. In the on-line manifestation of Atcros produced by this Project, this data was supplemented from later sources, and in particular the Atcros Directory information on contacts was comprehensively updated and expanded by reference back to the organizations involved and a continuing scan of trade and published sources.
The design criteria set for implementation of the Atcros On Internet website required: (1) fast downloading/display of the web pages; (2) an attractive, user-friendly presentation of pages; and (3) an underlying process system which could cope with frequent and regular updating of the pages. Innovation in design and display systems was desirable.

How these criteria were met is the main subject of this written Report. Their importance is reflected in the point that a high proportion of websites, both in the past and today, are annoyingly slow to access and move about in, often because of their emphasis on graphics and flashy gimmicks, and increasingly because of their selling of advertising space on their pages. Success in implementing the first two design criteria mentioned above can be judged from the website itself. Achievement of the third criterion is explained later in this report.

The Atcros website is at <www.AOI.com.au/ateros>. As well as the data which forms the main reference material for the site, included also are descriptive and explanatory pages covering such things as a news release when the site was launched, and an explanation of the area covered.

The site is currently receiving in excess of 4000 accesses ('hits') per year, which must be regarded as highly successful for a specialist site of this nature.

The foresight and confidence of RIRDC in funding the development of the AOI website is gratefully acknowledged. However the website's viability and future as a public-access facility depends on identifying a continuing, non-development source of funding.
1. Origins of Atrcros on Internet (AOI)

Atrcros on Internet as implemented consists of a Website, a set of interlinked World Wide Web pages sitting on the Internet and able to be referenced by anyone anywhere who has a connection to the Internet and a Web Browser such as Netscape.

Most of the data pages on AOI are Web equivalents of tables which appeared in printed form in the two editions of the book ATCROS: Australasian Tree Crops Sourcebook.

Both editions were edited by David Noël and published by Cornucopia Press. The first was published in 1991 for the West Australian Nut & Tree Crop Association Inc., the second in 1994 for ACOTANC Inc., the Australasian Council on Tree and Nut Crops.

The printed editions of Atrcros were themselves an extension and update of tables which appeared from 1980 onwards in WANATCA Yearbook, the annual reference publication of the West Australian Nut & Tree Crop Association. These tables aimed to list contact addresses of sister organizations of WANATCA elsewhere in Australia and overseas, and local suppliers of nut and fruit trees and seeds of perennial plant species.

The 1991 printed ATCROS was a courageous and innovative venture by WANATCA, who funded the first edition and distributed copies to all its members as a membership benefit. The four pages of tables in the 1988 WANATCA Yearbook (the last to contain the Useful Contact Addresses section) were expanded to 80 pages in the first ATCROS. To the original WANATCA tables were added a host of new tables, including some which originated from the California Rare Fruit Growers and the Rare Fruit Council of Australia, plus some local material not published previously.

The new material was in two sections. The first section covered mostly the characteristics, names, and recommendations for culture and use of the plants figuring in the Tree Crops area. The second section was a greatly expanded and updated version of the WANATCA Useful Contacts tables, now divided into 12 new classifications under the heading Atrcros Directory. The majority of the Directory information was drawn from a database compiled at the Tree Crops Centre, originally on hand-written catalogue cards, later transferred to a computer database.

The appearance of the first ATCROS in 1991 was warmly welcomed around the world, as was the second edition three years later, in 1994. However by 1995 difficulties were becoming apparent, especially in the use of the Directory section material. We live in a rapidly changing world, any printed contact list begins to lose its currentness as soon as it is issued, and after three years it may be downright unreliable. When a new edition of ATCROS was considered in 1995, it was realized that a new approach which would allow more frequent and rapid updating of contact information was needed.

1995 was probably the first year in which the World Wide Web started to make a significant appearance, and in which its potential began to be realized. In October 1995, a Preliminary Proposal (handwritten) was submitted to the RIRDC for the conversion and transfer of the ATCROS data into a Website on the Internet.
This was followed in 1996 by a Full Proposal which was eventually accepted, and a contract signed in October 1966. This was a significant event — not just new R & D, but a new sort of R & D, in a new and untried area. The foresight, courage, and perspicacity of the RIRDC in accepting such a proposal should not be underrated.

At this stage, funding was in place, and the serious work of implementation had to begin. The project implementor knew what was possible in general terms, but like almost everyone else at that time had very little experience on how to do it. This is not a bad basis on which to undertake research and development — if everything is known in advance about how to implement a project, it can scarcely be classed as research.

In what follows, the outcome of this development process is described, together with some insights into the design constraints and criteria involved, and the processes and techniques evolved to meet the overall goals.

2. Building the Website
It is commonly said that it makes no sense to try and reinvent the wheel. The implication is that it is always best to find out how others have gone about a research project first and use their experience in tackling one's own similar problem. This is accepted as almost self-evident.

Even so, there are virtues in starting a research project from scratch, looking rather at the basics and theoretical bounds of the task and ignoring how others have gone about it. In that way, a new approach or result may be obtained — ending up with the ball-bearing, or the caterpillar tread, perhaps, rather than the wheel.

In the present project, starting from basics led to a rather different result in the appearance of web pages, but one which was efficient and answered the design criteria (Figure 1).

**Frames**

In this design, the screen image presented by a Web Browser such as Netscape or Microsoft Internet Explorer is divided up into a number of sectors called Frames. Each frame can be manipulated independently by the Browser.

The top of a typical screen holds the menu and button items belonging to the browser itself, and these remain unchanged as the browser goes to different Web sites. Below this are the frames belonging to Atcros. Each frame will normally remain the same size as the data within it changes according to the user's commands.

A standard Atcros display screen contains 9 frames. The five help boxes across the top are each a separate frame, as is the banner right across the top of the main frame, the largest part of the screen. Below the main frame is a small counter and update box, on the left, and on the right is the Scratchpad box.

**Main Frame**

The text information being looked at by the user appears in the main frame, and the user moves this data up or down by moving the scroll bar on the right of it. A specific name or piece of text within the screen can be found by using the browser's FIND button — when a word such as 'pecan' is entered at this button, the browser highlights its first occurrence, and a simple keyboard instruction moves the FIND on to the next occurrence.

**Scratchpad**

The Scratchpad, which is an Atcros innovation, can be used in a number of different ways. Text can be captured from the main frame by dragging the mouse across the required text, COPYING it with a single keystroke, and PASTING it in the Scratchpad with another keystroke. This process can be repeated to build up a list of relevant entries or parts such as phone numbers in the Scratchpad. When the list is complete, the whole of the scratchpad can be copied with a single command and pasted to a Notepad or Word Processor application running independently on the user's computer.

**Help Boxes**

In a second usage, the Scratchpad is used to display the contents of any of the Help Boxes which appear above the main display. If the user clicks on the 'Unfold Box' link of the '2. Usage Hints Box', the stored contents of this box will be displayed in the Scratchpad area. These box
contents, which may vary from one table to another according to what they contain, give help to the user and may contain links to other pages within the site or to completely different sites. For example, the '1. About This Table' box will unfold to give links which can lead to the RIRDC's own website.

**Web Boxes**
The third use of the Scratchpad is to hold another Atcros innovation, the Web Box. Some of the Atcros entries have clickable links on their Reference number at the end of their entry (e.g., at the end of the 'Hamel Nursery' entry. If such a link is clicked, a small file of information about the organization featured in the entry opens up in the Scratchpad area.

This Web Box operation is similar in nature to the usage of a Classified (e.g., Yellow Pages) telephone directory, where an alphabetical entry for a company may have a note such as *See Advertisement page 2298*. With the Atcros Web box, clicking the link causes the relevant organization information to appear at the Scratchpad box, and this information may contain further links, for example to an organization's full web pages.

The special value of using a separate frame for such uses is that the information displayed in the Web Box does not affect that in the Main Frame. Most web page help facilities rewrite the whole screen, and to go back to the previous screen it is usually necessary to re-load the whole underlying file, which is time-consuming for a long file.

**Counter/Update Box**
The small box on the bottom left of the screen contains a Page Counter and an Update Notice as well as the pointer to the Scratchpad. The Update Notice is updated manually by Atcros when a new version of the data displayed in the Main Frame is installed.

The Page Counter is quite different in nature to the rest of the Atcros data. This counter is actually an information slot which sits on the computer of the Internet Service Provider or ISP. The intention is that each time the relevant page is accessed by a Web user, the ISP increases the count in the information slot for this page by 1. This is in contrast to the rest of the page data, which is uploaded by the website controller to the ISP whenever a file or page is updated.

In practice, these page counters give an indication of usage, but no more than that. The problem is that a minor breakdown at the ISP may cause all the counters to be reset to zero. In addition, the ISP counters are sensitive to how the page counters are described in the site's HTML code and to where they are referenced.

In addition, page counters only increment when the page they relate to is accessed, as would be expected. Regular users of a particular Atcros table page will 'bookmark' that page so that it can be accessed directly in a later use. The counter on the Atcros Home Page will not count such accesses because they bypass the Home Page.

Figure 1. Sample layout for Atcros Web pages.
3. Website Design for Speed, Appearance, and Maintainability

The previous section described the general appearance and features of Atcros web pages, without specifying the reasons for the design. A description of how and why these particular aspects were arrived at may be of value to others designing similar websites.

The frame structure used gave a number of benefits.

**Updating**
Each frame is the result of an underlying HTML file. A complete Atcros screen in fact draws on 15 underlying HTML files, only 10 of which are active at any given instant. This is apart from any additional files which may be pulled into a Web Box.

By breaking the screen up into frames, only the underlying files for data which has changed need to be modified and uploaded for an update. With a routine Atcros update, only the files for the Main Frame and the Counter/Update Box need to be replaced at the ISP. The other 13 files are only replaced when the conditions for the services they provide change.

**Appearance**
Browsers allow the background colour of a frame to be specified very simply. Using a number of frames, each with a characteristic background colour, gives a distinctive and pleasing appearance to the screen.

Background colours may also have some function. For example, the Banner Frame has the same background colour as the Main Frame which it refers to, and which sits immediately below it. It acts similarly to the column headings on a table, but remains in position as the data in the frame below it is scrolled up and down. If a Help Box is opened, its contents appear in the Scratchpad area with the same colour as the Help Box chosen, tying the two together.

When visible, the Scratchpad itself is white, suggesting blank paper, and the Counter/Update box has a black background, suggesting it is fixed and complete.

**Screen Size**
It is a feature of Web pages that when interpreted by a Browser, the latter interprets the HTML files to fill the required screen or display size. With a Netscape browser page, at the bottom right-hand corner of the page is a small box called the Size Box. If this Size Box is picked up with a mouse click and dragged to another position on the computer screen, to give a new position for the the bottom right-hand corner of the page displayed, the browser re-loads all the underlying HTML files to fit the new proportions of the frames.

This is a required feature of HTML files underlying Web pages, so that the browser is usable on a computer screen of any size or resolution, and with character sets and fonts which suit the user. A browser will normally have a default set of character fonts and styles interpreted in a standard way, for
example it will use HTML markings to make some parts of the text bold, italic, coloured, a sans-serif
script, or a large size. But the user, perhaps someone with limited vision, can set the browser to display
much larger characters than normal if they wish. The underlying HTML code has to take account of
these possibilities.

Tables and Running Entries
For the reasons just outlined, the satisfactory display of tables on Web pages can be quite tricky. A
typical table made up of a number of columns can be tackled in various ways. One is to write it in the
HTML as a fixed item, with words appearing at fixed places on the screen. This is unsatisfactory when
the user has a small screen (and part of the table will be always offscreen) or is using unusually large
characters. The user may still scroll left and right to bring parts of the table into view, but this is very
annoying if the table is wide.

Another way is to use HTML table code to define a table, each column of which has a fixed percentage
of the screen width. This can give a satisfactory result with a small table with a limited number of
columns, but was unacceptable for Atcros, which typically would have needed 10 columns. For many
entries, some columns would be almost empty, while others might have 10 or 20 words or part words
strung one above the other, and there would be a struggle to fit 10 readable columns of fixed widths
across a typical screen.

The solution adopted with Atcros, possibly another innovation in this area, was to present tables as
running entries — a single stream of text broken up into 10 parts by leading flags (such as ‘A:’ before the
address) and by presenting different parts in different styles, such as bold and italic. Such a format
adjusts quickly and effortlessly to varying screen sizes. The leading flags are explained in the Banner
immediately above the entries.

This solution also had two other major benefits as outlined below.

Downloading rates
In its structure, an HTML file consists essentially of a stream of characters. Some of these characters
are ones which will be displayed on the final screen, as text. The others are instruction characters which
tell the browser how to lay out the page, assign styles to the text, and so on.

If the web page is simple text, as in a letter or a novel, the percentage of the character stream which is
instruction may be quite low, say less than 10%. With a defined table, the position is quite different. A
single character sitting in a column, or even no character, may require 20 or more instruction characters
to define its position and style and the width and position of its table element. A standard table
definition of some text may therefore be twenty times the size of the text it contains.

The running-entry approach used in Atcros is very economical in its use of instruction characters, again
probably less than 10% in most cases.

The rate at which a browser builds up a page on the screen mostly depends on two things: the size of
the HTML files being downloaded, and the speed and congestion of the modem or communications link
used to connect the user's computer to the ISP holding the pages referenced.
Typical modems used for Web browsing have theoretical downloading rates of between 1400 and 5600 characters per second (1.4-5.6 kb/s). Congestion and switching limitations in the path travelled — which may involve dozens of links and steps of tens of thousands of kilometres — can cut the effective transfer rate down considerably, to below 100 characters a second in a bad case.

It is therefore vital to keep the character size of HTML files down to the smallest possible where they can still function efficiently. If at all possible, HTML files should be under 100 kb (100 thousand characters) for ease of use.

The way that the data is divided into files may also have an effect on perceived speed of downloading and screen presentation. Most of the Atcros frame files are less than a couple of hundred characters long. The browser takes only seconds to download these and quickly build up the screen picture, and the Main Frame data fills in the first entries and presents a complete screen. At this stage, as much of the main data as has been downloaded, is available for use (eg a FIND can be done on it), while the browser quietly completes the download.

In a conventional HTML table definition, it has been mentioned that the instruction characters for the table can swell the size of a table by perhaps 20 times, and this multiplies the download time by a similar factor. The running entry approach used in Atcros avoids this problem.

There is another approach to HTML table definition, in which only the bare bones of the instruction character tags are included, and the browser is left to make its own best assessment of how to lay out the table column widths and other parameters. However the outcome of this for the user can be annoying, as even though the download time may be reduced for the total file, the browser cannot display anything at all of the table until it has everything in for its assessment, which itself takes appreciable time. Such a table is therefore perceived to load very slowly, and in the case of Atcros this technique has been used only on one or two very small tables.

**Automatic update procedures**

Earlier in this this report it has been explained how the Atcros user views the website pages, and the reasons for their overt design and appearance. These features are immediately apparent to the novice user, while an experienced user who knows about HTML and wants to find out how different features are achieved can do so by looking the HTML source code — briefly, this is the raw HTML file, including instruction characters, which is delivered to the user's browser for interpretation as a web page.

The methodology used to generate the HTML source code is normally unknown to the web page user. In the case of Atcros, this methodology is explained in a later main section. Briefly, the original data underlying tables such as those in the Atcros Directory pages is maintained and updated on a conventional computer database. To satisfy the requirements for easy regular updates of pages, the Tree Crops Centre developed what is believed to be an innovative methodology to take the information involved through from its database form to the HTML form which can be used by a browser.

The 'running entry' approach used in most Atcros tables is itself well suited to this 'Intelligent Autoconversion' methodology. Methods of presenting tables on websites have been discussed above. Of these methods, the fixed-item table is totally unsuited to an automatic conversion approach.
Conventional HTML tables can be generated automatically, but still end up clumsy and excessively large.

**Graphics**
Graphics have their place on Web pages where the website's purpose is to entertain or impress, but are annoying to the web user whose main purpose is to find specific information. Many organizations will place a large graphic of a prominent building on their Home Page, without any appreciation of how long this takes to download, while others will place pictures of objects to represent links from their pages — a picture of a book to represent a link to their publications list, for example.

Graphics take a long time to download because they are very large compared to text. A fully-functional text-based Home Page may be only 200 characters long, while a single large graphic may be a thousand times bigger.

In order to deliver speed of access, Aicros pages generally dispense with graphics altogether. The result is efficient, although it may be criticized as somewhat plain or severe.
4. Database to HTML Conversion Methodology

As already described, the underlying data used to generate the Atcros web pages exists as a series of computer database files.

These files are mounted on Macintosh computers at the Tree Crops Centre. They are maintained, updated, and manipulated using the database facilities of ClarisWorks Office, a simple and effective piece of software which is normally supplied free with the purchase of a Macintosh computer.

Database description
As with normal database packages, ClarisWorks Office maintains the information for a given record or entry in a series of Fields, for example the 'Name' field in an Atcros Directory database record will normally hold the name of the organization described.

ClarisWorks Office also allows the data within records to be presented in various formats, called Layouts. For example, a Phone Directory Layout might include only two items from the complete record (having perhaps 20 fields), presented in two columns, 'Name' and 'Phone'. Records from the complete database can be selected and sorted according to quite complex criteria and the required fields displayed according to the parameters of the particular layout. Importantly, the data actually displayed in a particular layout can be copied and pasted into another document, such as a word processor document.

In effect, the layout gives a window capable of selecting and looking at particular parts of particular entries in a particular way. The records showing up in the window, say name and phone, may be sorted alphabetically by name, or even by a field which is not used in the layout, say postcode. The data as it appears in the layout can be copied and used elsewhere. If the underlying full record is amended, say when an organization changes its phone number, then a new look at the database with the same layout will show the new number.

Calculated text fields
A vital feature of the database facilities of ClarisWorks Office in the present context is its ability to generate calculated text fields. A calculated field in a record is one which is derived from other fields in the same record and/or from constants by applying an algorithm, that is a calculation or logical operation. These operations can involve text fields as well as other types of field, say numbers or dates or times. So it would be possible to have a calculated field 'Last Report' which took data from fields containing first name and last name, and the current time and date from the computer operating system, and added in spaces, punctuation, and text to give the result in a calculated field of:

:John Smith last reported in at 11.33 on February 10.

In the Atcros Intelligent Autoconversion methodology, calculated fields are used extensively to generate HTML codes (the instruction characters added to displayed text) directly from the contents and according to the contents of fields present.
For example, in Figure 1, each organization listed in the Plant Supplier table, Table Z, has HTML codes inserted so that its name appears in bold type and the description of its activities appears in italics. However, of the organizations shown, one, Hamel Nursery, has a Web Box on the Atcros website which can be displayed by clicking the link on its reference number. The Intelligent Autoconversion routines detect the fact that Hamel Nursery owns a Web Box, and insert HTML tags to make its name appear in red and activate a link from its reference number.

**Database Shells**

Another feature used in the database manipulation is that of database shells. The fields in a database record may be any of a range of types (text, number, date, logical, calculated etc), according to the capabilities of the particular database package.

The main working database for Atcros is a full one, in which the majority of the fields are text fields, with a few calculated, number, and date fields. It is this version which is amended as changes are discovered or notified for individual entries. This version contains a number of 'administrative' fields, in which dates of amendment are noted, the outcome of queries to the subjects of the entries, and the like.

In the database shell technique, a copy is made of the main database, and its characteristics are modified. A number of fields might be cut out, a calculated text field might be re-assigned as an ordinary text field (with contents as they stood and no longer dependent on other fields), and new calculated fields might be created, with values dependent on the new field contents and types of the database records.

If the modified secondary database is emptied of active records, it may be saved as a database shell. If then a copy of the original database or a selected portion of it is copied and dropped into this shell, the operation automatically culls, reformats, and creates fields in the same way as was done manually in creating the shell.

In the Atcros 'Database to HTML conversion methodology', this database shell technique is used extensively. Because calculated fields are generally calculated using the values of non-calculated fields, a number of successive shell conversions may be needed to achieve a desired result. In general, the HTML code will be added to the applicable text in the last active database shell, immediately before data is stripped off in word-processor form.

Figure 2 shows a notional flow diagram of the conversion stages needed to generate individual Atcros frame tables to the stage where they can be uploaded to the ISP. The flow operations are cycled, ie one cycle through the flowchart is needed to generate each table, with variations to generate additional tables such as indexes.

Appendix 2 gives some sample calculated-field expressions used in generating Atcros HTML tables, together with some notes and cautions on their use. These expressions are those used by ClarisWorks Office version 4, although the expressions are believed valid in some earlier and later versions of this software.
Figure 2. Flowchart for generation of Atcros tables

5. Original Plan and Modifications Made
Appendix 3 gives the original plan for setting up the Atcros website, as approved by RIRDC, with the expected table structure and page hierarchy. This Appendix offers further background material on the conception, justification, and expected implementation of the project.

The final structure as it stands at the date of this report follows the original plan quite closely. The main differences have been in two areas.

First, extra indexes have been added. In the first few months after public release of the website, it became obvious that the Directory tables were the ones getting most use, and that an extra index should be provided listing names and contact persons of organizations listed and the tables in which they appeared. Clickable links to the individual tables were incorporated, not in the main frame index itself, but in the Banner Box above this frame.

A convenience of web page use is that a phrase typed into a FIND box can be used repeatedly in zeroing in on the required information. For example, 'olive' can be used in the Directory index to locate organizations with this word in their name, the relevant table can be opened, and the same FIND can be used without re-typing to locate, say, suppliers or processors of olives.

Second, the Atcros Home Page was revised and greatly simplified, with non-essential material pushed out to side pages accessible by links from the Home Page. The original Table of Tables, listing the contents of the 24 Atcros tables, was split into two pages, one listing the 12 Directory pages, the other the 12 Knowledge Base tables. The Home Page was streamlined and divided into an Experienced User panel and a New User panel, with obvious direct links to the most used facilities.

Appendix 4 gives an interim progress report on the project. The most noteworthy item in this report was that Atcros reference numbers were being quoted in Quandong, the magazine of the West Australian Nut & Tree Crop Association Inc. For example, news items on research work being done at a particular institution would have its name underlined, with its Atcros Reference number, eg &lt;A1234&gt;, listed at the end of the article. An advertisement for a pecan processor might also quote his Atcros number.

The value of this practice is that a reader seeing a quoted Atcros Reference number would know that the Atcros website could be referred to at any time to get current contact details for the organization referred to. Printed details in a magazine are only current as at the time of printing, the use of Atcros Reference numbers allows details to be checked as the time of reading, which might be some years later.

6. Conclusions
The Atcros on Internet (AOI) project as described in this report represents a significant RIRDC-supported advance in a very fluid and rapidly-advancing area of information technology.

The result, the Atcros Website, is believed to still be unique in its coverage of a major area of rural production and research, which is that relating to useful perennial plants ('Tree Crops'), for a major world region ('Australasia'). It is believed that this project has created a valuable information-access tool which can continue to be of value to workers in this field, into the future.

In addition, the existence of this facility may serve as a useful model or prototype for others looking to set up a similar tool for other subject areas and geographical regions. The detail given in the present written report is intended to assist in this process.

The site is currently receiving in excess of 4000 accesses ('hits') per year, which must be regarded as highly successful for a specialist site of this nature.

However, the future outlook for the site is uncertain. The financial support from the RIRDC was limited to that needed to create the facility, as being in the nature of a research project, and has now ended. The continuance of Atcros as a public-benefit facility in the future is dependent on identification of funds for its maintenance and upgrading to meet future standards.

Appendix 1. Brief Glossary
(see also Section B7 of Appendix 3 for some more detail and background)

**Browser**: a computer program or application sitting on a user's computer and designed to allow access to the **Internet**. Common examples are Netscape and Microsoft Internet Explorer.

**Character Set**: the set of displayable characters in a text font — letters, numbers, punctuation signs etc.

**Download**: the act of transferring digital data from a host to the user's computer.

**Field**: a specific section of a database record.

**Font**: the graphic style name of a character set, eg Times or Helvetica.

**Frame**: a section of the screen displayed by a browser, the contents of which can be manipulated independently of the rest of the screen.

**Home Page**: the first point of access for a website.

**Host**: a computer assembly linked directly and permanently into the Internet and typically holding Web pages intended for access by remote users.

**HTML**: a system of marking Web page data to give specified screen appearances and page access and linkage when interpreted by a Web browser.

**Internet**: a world-straddling grouping of publicly-accessible computers (hosts), linked by very high speed communications links designed for rapid exchange of digital data. Facilities include the Web, e-mail, and other types of usage.

**Internet Service Provider**: A commercial host offering Internet access on a fee-paying basis.

**ISP**: abbreviation for Internet Service Provider.

**Upload**: the act of transferring digital data from the user's computer to a host.


**Website**: a set of interlinked World Wide Web pages sitting on the Internet and able to be referenced by anyone anywhere who has a connection to the Internet and a Web browser such as Netscape.

**World Wide Web**: The totality of all HTML-marked data files held on all hosts of the Internet.

---

**Appendix 2. Examples of Calculated Text Fields and their Usage**
Below are given most of the calculated text fields used in the last stage of generating HTML entries for Atcros tables. These expressions are for Claris Works v. 4.0 and other versions of this software.

Field names are delimited by single quotes and literals (text) by double quotes. Two consecutive double quotes imply a field or literal which is empty. A blank is represented by a blank between double quotes.

Expression List.

Field: Line5. Algorithm: CONCAT('Addr3',' ','Postcode')
Notes: This simple expression joins the third address line with a blank and the postcode or country.

Field: FullRef. Algorithm: CONCAT('File','Ref#')
Notes: This converts the 4-digit entry reference number (eg 1234) into text and appends this to the file code 'A' (to give eg "A1234").

Field: FullFax. Algorithm: IF('CountryCode:+'<>"",CONCAT('+','CountryCode:+','-','Fax'),CONCAT('0','Fax'))
Notes: If the country code field is not empty, the field contents and a hyphen are placed before the fax number, otherwise the Australian national code of '0' is placed before the fax number.

Field: OpString1. Algorithm: CONCAT(" L: ','Area',",". N:<b>")
Notes: In this first part of the HTML expression, the entry's area, name flag, and start of bold are written.

Field: OpString1a. Algorithm: IF('Box'="Y",CONCAT("<font color="#ff0000","Q","#ff0000","Q",">","Name","<font>"),"Name")
Notes: In HTML string 2, the entry name is written in red if a boxholder, otherwise as is.

Notes: HTML string 3: bolding closed off, first line of address.

Notes: HTML string 4: parts of the address and postcode or country.

PROJECT DESCRIPTION (Original Proposal)

Project Title: ATCROS ON INTERNET (AOI)
RIRDC Program: 4.3 Education, Extension & Information Systems
Objective.

1. To mount the data files contained in "ATCROS", the Australasian Tree Crops Sourcebook (as published) as a World Wide Web site on a suitable Internet host, for general public access.
2. To maintain, regularly update, and progressively enhance this information

B4. Background to Proposal.

1. The Internet, and in particular the World Wide Web, represents the fastest-growing and most rapidly developing new technology in human history. Usage is doubling or tripling each year.

2. ATCROS is believed to be unique in the world in its coverage of information on useful perennial crop plants and on public, research, government, and commercial organizations involved in current and future production and research.

3. This is a unique and timely opportunity to construct a Web site based on the ATCROS data, to serve as a continuing and vital tool to all involved with perennial plant crops, particularly those which are the subject of ‘sunrise’ industries which must access less readily-available information and contacts if they are to succeed.

4. The ATCROS data exists as a series of computer files held at the Tree Crops Centre. The last published edition was in 1994. Use of Web technology permits file data to be updated regularly and rapidly, avoiding the drawback of printed information which tends rapidly to become out-of-date. This is particularly true of contact details (especially phone and fax numbers) of organizations such as suppliers of plants, seeds, and services, but also applies to research and education contacts.

B5. Relevance and Benefits.

1. ‘Tree Crops’ is used in our context to relate to any useful perennial plants. This includes not only conventional horticulture, but also timber production, vine crops, plants for land stabilization and improvement, and plant sources of animal fodder, pharmaceuticals such as quinine, beverages such as coffee, essential oils such as eucalyptus, and a big range of industrial products such as jojoba oil, rubber, and cork. All such industries should benefit from the availability of ATCROS on Internet.

2. Benefits should be particularly evident for new and sunrise industries in this area. It is a central operational point that however much effort is put into informing people and publicizing the virtues of a proposed new crop, this effort may be largely wasted if those interested are unable to obtain plants or materials essential to the raising of that crop. It is a basic aim with ATCROS to put people interested in a new crop in touch with suppliers and information relevant to that crop.

3. Tree crops differ from conventional broadacre crops and stock raising in that they are biologically extremely diverse, involving thousands of species, with the scope for hundreds of specialized and ‘niche’ crops. This in itself leads to a diverse and stable economy with built-in buffers against climatic and trade variations. The whole AOI project can also serve as a useful template for other promising areas of primary production which have similar biological diversity, such as fish farming, small meat-animal production, and production of Australian-species plants for cut flowers.
4. AOI has the potential to be a heuristic entity, continually evolving and improving through actual use. For example, if a user attempts to look up propagation details of a particular species and finds information in AOI deficient or in conflict with what the user has, they can leave a message for the AOI editor to that effect at the Web site, immediately and at no cost. This is effectively equivalent to being able to annotate an encyclopedia with a note on an obscure paragraph, and have that annotation transmitted immediately to the encyclopedia editor, with the likelihood that the next user will find the obscure paragraph amended or corrected.

5. The original ATCROS, in its two printed editions, has benefitted from feedback from industry and research users in the field. However, a printed version requires a particular effort and cost to enter such feedback, which in any case could not have effect till the next printed version. With the AOI proposal, such feedback is easy and without cost, and can have a rapid effect.

6. AOI also has the capability for expansion to offer other associated tools for users. For example, ATCROS has two main divisions, the first a Databank with information on particular species or applications, the second a Directory of organizations and other contacts. On the Databank side, a user looking up cultivation details for Mabolo, a new fruit in the persimmon family, could be offered a link not only to literature references, but also to an encyclopedia-style account of Mabolo linked to the main AOI database. This is a possible future development and is not included in the present proposal for funding.

7. On the Directory side, it would be possible to offer organizations included, such as colleges or seed suppliers, two additional facilities, a Web Box and a Web Page. The Web Box would be similar to a box advertisement in a classified directory, or a box office at a multi-screen cinema, in that it would include amplified information on the organization concerned, plus links or gateways to organization-maintained facilities such as its e-mail message receiver or its own web pages, all accessible by the press of a button. A Web page, with parallels to the actual film seen after accessing the box office, might include a list of courses in the case of a horticultural college, or a list of seeds available in the case of a seed supplier. Web Pages can hang off the AOI facility, or an organization's own facility, or both, with no obvious difference to the user.

8. The suggestion contained in (7) is a possible future development, is here for information, and is not included in the present proposal for funding. It is believed that such a development could be self-supporting financially, and if proceeded with, would not engender any call for RIRDC or other outside funds. Obviously it does require the basic proposal to be implemented first. However, as far as the RIRDC is concerned, it would be offered free use of a Web Box and a Web Page, if desired. The latter might include a file of research reports available, or even on-line research project applications and conditions.

9. As far as impacts on ecological sustainability and biodiversity are concerned, this project would be highly positive in every way.

10. Dissemination and extension are inbuilt in the project, no further action is required. In addition, every visitor to the AOI Web site would be alerted with a tag 'A project supported by the RIRDC, the Australian Government's Rural Industries Research & Development Corporation' or similar, plus a button to access an RIRDC Web Box.
B6. Review and Interaction.

1. The database to be mounted on the Internet in this proposal originated with a 'List of Useful Contacts' published in the WANATCA Yearbooks of 1987 and 1988. This information was greatly expanded and updated in the first separately-published ATCROS of 1990, and further updated and expanded in the second edition of 1994. A very large variety of literature and other sources was accessed in this work.

B7. Methodology.

1. Because of the unsurpassed rate of development of the Internet and the World Wide Web, some explanation of the terms involved in this still widely unfamiliar field is warranted.

2. Files of information intended for use on the Web are written in a language called HyperText Markup Language (HTML). HTML files are ordinary text files, with two additions. The first class of addition is formatting marks such as those used by word processors, for example `<br>` for a line break.

3. The second class of addition is the more important one. These additions consists of links or 'anchors' which point to information elsewhere, perhaps within the same file, to a file elsewhere in the same database, or to a location within a different database, possibly half a world away.

4. It is these links which transform the various items of data into 'hypertext', a mass of linked text items whose components can be distributed anywhere, or 'hypermedia', when graphics, sound items, or video clips are included.

5. The Internet is a set of linked computer sites containing information and interacting according to an agreed set of standards ('protocol') which allows the various types of data involved to be transmitted, interpreted, and operated on.

6. Public access to the Internet is via an Internet Service Provider (ISP), who leases lines, from a communications provider such as Telstra, which are permanently connected to the world network of ISPs. A large university may have its own ISP node, an individual user is more likely to dial-in to a local commercial ISP using a modem-equipped computer.

7. The World Wide Web is one of the available modes of using the Internet. It includes two main parts: a Web Browser, which is a computer program set up specifically to access and manipulate Web data, and the data itself, which consists of any and all files written in HTML and accessible through the Internet.

8. Simple Web Browsers may sit on the computer provided by the ISP, in which case the user is essentially using a terminal onto the ISP system. However, more recent and advanced Web Browsers such as Netscape sit on the user's own computer, usually after having been downloaded (sent down a communications line) from the ISP.

9. Web Pages, the files of data intended for Web Browser access, usually sit on the ISP computer, where they are accessible at all times via high-speed links, rather than on the user's computer (where they would only be accessible at low speed and only when the user was connected).
10. With a reasonably complex database such as AOI, revisions are not normally made directly on the version of the database accessed by remote Web Browsers via the local ISP. Instead, revisions, extensions, upgrades, and new linkings are made on a 'Front-End' computer maintained by the user, and the latest 'tuned' version of the data is downloaded or otherwise transferred to the ISP at intervals. In so doing the previous version is overwritten.

11. In the case of AOI, new Browser versions would be expected to be in place on our Internet Service Provider at least once per month, and possibly fortnightly.

12. This arrangement not only allows 24-hour, high-speed access from anywhere in the world, it also protects the original data from corruption by accidental computer failure at the ISP or elsewhere on the Internet, and by hacking.

13. The basic ATCROS data already exists as computer files. Once a Front-End computer is available, it is expected to take only 3 months to satisfactorily mount and trial this data on a local ISP site. At the end of this time, AOI should be 'live' — accessible via a Web Browser, albeit with a proviso ("Web site under construction, data to be refined").

14. At this stage the site could be publicly demonstrated and promoted.

15. The next stages could proceed concurrently. Of the tasks involved, the biggest and most laborious is Directory Vetting.

16. Directory Vetting involves contacting existing organizations listed in the AOI Directory and verifying details, also discovering and adding new candidates for entry on the directory. The latter is done partly by scanning the several hundred journals from allied organizations received at the Tree Crops Centre, partly from other directories and databases or through other contacts.

17. Directory Vetting is continuous, on-going, and irregular. A 'mature' organization which already contains an AOI contact with e-mail access, which has already been vetted once, may respond unprompted with new details whenever these change. An isolated third-world research centre without fax or e-mail facilities and limited English or secretarial help may take over a year to respond (a response was received in January 1996 from an ATCROS organization which was queried late in 1993). In the GANTT chart, this process is listed for one year, but its erratic nature must be appreciated.

18. Databank Vetting involves making similar reviews for the databank information in AOI. However, such vetting is more similar to that of revising a book for a new edition, making use of newly-available data sources and feedback received from users. There is also an aspect of databank enhancement, both by adding new files and by improving links between existing data. As an example, the basic AOI files include a file of fruit common names showing their botanical names, and a separate file of cultural information under botanical name. With improved linking, a user searching the common name list could be provided with a 'button' to press to link directly to the botanical name file at the correct point.

19. Databank Vetting is included in the GANTT Chart for 21 months in the first instance. However, the actual process would yield precedence to the Directory Vetting and also depend on progress elsewhere;
rather than duplicate work already done elsewhere, giving data available on the Internet, it would be more cost-effective simply to provide an in-built link to that work.

20. As the rate of change on the Internet is so fast, future developments must be considered mostly in the light of future circumstances and demands. However, one anticipated problem is user questions. The Tree Crops Centre already receives many phone and written requests for information about a particular new crop. There is no provision in the above scheme or funding outline for answering a hundred such e-mail requests each week from all over the world.

21. As regards milestones and benchmarks, physical events are easy to define, intellectual perceptions more difficult. There are no major technical problems in mounting the AOI Web Site as described. Most accesses of the Site can be recorded and analysed to give an indication of use. However, the ultimate degree of success for this project will depend on the level of user opinion, and that itself depends on the design skills, foresight, and knowledge of the field possessed by the implementor.


The last 12 month's work has seen the Atcros project through all major development phases and on to public release.

The project places two groups of information relevant to tree and other perennial plant crops on the World Wide Web, in 24 tables or pages. Half of these tables relate to relatively stable data on crop tree properties, culture, names etc.

The other 12 tables comprise a Directory of contact details and interest areas of all relevant Australasian organizations -- grower groups, plant and seed suppliers, research and education bodies, government entities, traders, etc. This information is constantly changing and needs regular updating.

The whole website was publicized and released for public use, 1997 August 1. Initial take-up of the site was pleasing, in terms of number of accesses, and has increased regularly ever since, currently achieving a very satisfactory 150 accesses each month.
Current work is mainly on verifying, updating, and adding to Directory entries, now about 1200 in number, up from the original 1000. Some entries have varied as much as 5 times since August. Table updates have been regularly achieved about every 24 days during 1998. All data will be verified at least once by October 1998.

A proposal has been submitted for a major enhancement, the Consortium Web Project, which would allow non-profit Atcros organizations to place public-value information from their fields of expertise onto Web pages at negligible cost.

A recent innovation is the quoting of Atcros 4-digit codes with articles and notices in the WA Nut & Tree Crop Association magazine 'Quandong'. This allows 'Quandong' readers to quickly locate full details of organizations mentioned through the Atcros site.