Information in support of responsible fisheries and aquaculture

ISSN: 1728-4392

Guidelines on digital publishing

A practical approach for small organizations with limited resources

FAO/FishCode
Review No. 20

Global Partnerships for Responsible Fisheries

fishCODE

FAO FIAT PANIS

Cover graphic by Emanuela D'Antoni
Layout by Nicholas Rubery

Printed on ecological paper

Information in support of responsible fisheries and aquaculture

Guidelines on digital publishing: a practical approach for small organizations with limited resources

by

Simon Wilkinson and Jean Collins

Global Partnerships for Responsible Fisheries (FishCode)

Food and Agriculture Organization of the United Nations
Rome, 2007

The designations employed and the presentation of the material in this information product do not imply the expression of any opinion whatsoever on the part of the Food and Agriculture Organization of the United Nations concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries.

All rights reserved. Reproduction and dissemination of material in this information product for educational or other non-commercial purposes are authorized without any prior written permission from the copyright holders provided the source is fully acknowledged. Reproduction of material in this information product for resale or other commercial purposes is prohibited without written permission of the copyright holders. Applications for such permission should be addressed to the Chief, Electronic Publishing Policy and Support Branch, Communication Division, FAO, Viale delle Terme di Caracalla, 00153 Rome, Italy or by e-mail to copyright@fao.org

© FAO 2007

Foreword

These Guidelines on digital publishing were prepared by Simon Wilkinson (Network of Aquaculture Centres in Asia-Pacific) and Jean Collins (Librarian, Fisheries and Aquaculture Branch Library, FAO Fisheries and Aquaculture Department) with the support of the FAO FishCode Programme, through the FishCode Trust (MTF/GLO/125/MUL) and through the FishCode SIDS Project (GCP/INT/823/JPN – "Responsible Fisheries for Small Islands Developing States"). The Guidelines are targeted primarily at small organizations with limited resources in developing countries, in order to facilitate decision-making on how to publish and disseminate their information.

The *FishCode Review* series publishes results of studies, missions, consultations, workshops, meetings and other project activities undertaken through the Programme, in furtherance of the objective of facilitating implementation of the 1995 FAO Code of Conduct for Responsible Fisheries and related international fisheries instruments and plans of action. Individual issues in the series are distributed to appropriate governments, regional bodies, meeting participants and Programme partners. For further information on Programme background, publications and activities, please consult the Web site at **http://www.fao.org/fi/fishcode.htm.**

J. Eric Reynolds
Programme Coordinator, FishCode
FAO Fisheries and Aquaculture Department
Rome, Italy
fishcode@fao.org

FAO/FishCode Review, No. 20
Distribution:

NACA
IAMSLIC
Fisheries and Aquaculture Department Divisions and Services
Fisheries and Aquaculture Department Branch Library
Other interested agencies

Wilkinson, S.; Collins, J.
Information in support of responsible fisheries and aquaculture. Guidelines on digital publishing: a practical approach for small organizations with limited resources.
FAO/FishCode Review. No. 20. Rome, FAO. 2007. 68p.

ABSTRACT

The importance of research in fisheries and aquaculture is referred to throughout the Code of Conduct, as is the need to disseminate and share the results of research. Stakeholders in developing countries generally are still waiting for reliable, high speed and cost-effective Internet access that is widely available in the industrialized world. Once connected, users must grapple with, and make decisions about, myriad technological solutions that exist. These Guidelines on digital publishing are targeted primarily at small organizations with limited resources in developing countries, in order to facilitate decision-making on how to publish and disseminate their information. The Guidelines are based on the years of experience of the Network of Aquaculture Centres in Asia-Pacific (NACA) and its partners. The approach is practical in orientation, covering topics including: (a) planning, building and maintaining a sustainable digital publishing system, focusing on a common scenario of setting up a Web site as a digital publishing platform; (b) producing user-friendly digital publications and making them accessible; (c) some recent international developments in digital publishing; and (d) recommended software tools and technical resources for further reading.

Keywords: Code of Conduct for Responsible Fisheries; digital publishing; fisheries information; aquaculture information; fisheries research; aquaculture research; information commons; aquatic commons

CONTENTS

Abbreviations and acronyms **5**

1. Background **6**

1.1 The need to publish **7**

1.1.1 Validating the results of research and development *7*

1.1.2 Sharing the results of research *8*

1.1.3 Lessons learned in the development process *8*

1.1.4 Learning from local knowledge *8*

1.2 Where to publish **8**

1.3 Control of the publishing process **9**

1.4 Emphasis on the users and target audience **9**

1.5 Aquaculture development **10**

2. Guidelines on digital publishing. A practical approach for small organizations with limited resources **11**

2.1 Introduction **11**

2.1.1 What is digital publishing, really? *12*

2.1.2 Advantages and limitations of digital publishing *12*

2.1.3 Getting started *14*

2.1.4 Where to get help *14*

2.2 Planning a digital publishing system **14**

2.2.1 Getting approval for the project *15*

2.2.2 Determining requirements *17*

2.2.3 Designing a digital publishing system *18*

2.2.4 Staff considerations: recruitment and training *23*

2.2.5 Building an information system *24*

2.2.6 Choosing a content management system *28*

2.2.7 Suggested content management systems *29*

2.2.8 Surveying products *30*

2.2.9 Short-listing products *30*

2.2.10 Testing products 31

2.2.11 Getting a domain name 32

2.2.12 Hosting your Web site 32

2.2.13 Optimising your site for search engines 34

2.2.14 Security, backups and disaster recovery 37

2.3 NACA Web site and publications 40

2.4 Creating accessible digital publications 44

2.4.1 Increasing accessibility through a multimedia approach 44

2.4.2 Increasing awareness of publications 45

2.4.3 Metadata for the Web 45

2.4.4 Improving physical accessibility: a crash course in compression 47

2.4.5 Adopting an open access policy 51

2.5 Successful digital publishing 52

2.5.1 Provide high quality targeted content 52

2.5.2 Explore syndication and partnerships 53

2.5.3 Establishing an online community 53

2.5.4 Performance monitoring and improvement 56

3. Collaboration and the next steps 57

3.1 Where NACA is today 57

3.2 Integration and long-term security of publications 58

3.3 Joining the information commons 59

3.4 The aquatic commons 60

References 61

Appendix 1: Glossary of technical terms 63

Appendix 2: Open source digital library and digital repository software 65

Abbreviations and acronyms

ABAFR	Aquatic Biology Aquaculture and Fisheries Resources
AGRIS	International Information System for Agricultural Sciences and Technology
AJOL	African Journals Online
ASFA	Aquatic Sciences and Fisheries Abstracts
CC	Creative Commons
CD-ROM	Compact disc read-only memory
CMS	Content management systems
CSS	Cascading style sheets
DLC	Digital Library of the Commons
DOALOS	Division for Ocean Affairs and the Law of the Sea, Office of Legal Affairs, United Nations
FAO	Food and Agriculture Organization of the United Nations
FOSS	Free and open source software
GIF	Graphics interchange format
GNU	"GNU's not Unix"
GIMP	GNU Image Manipulation Program
HTML	Hypertext markup language
IAMSLIC	International Association of Aquatic and Marine Science Libraries and Information Centers
IAS	Information access survey
ICES	International Council for the Exploration of the Sea
ICT	Information and communication technology
IMARK	Information Management Resource Kit
IOC	Intergovernmental Oceanographic Commission
IP	Internet Protocol
IR	Institutional repository
IT	Information technology
JPEG	Joint Photographic Expert Group
LAMP	Linux, Apache, MySQL and PHP
MARC	Machine readable catalogue
NACA	Network of Aquaculture Centres in Asia-Pacific
OAI	Open Archives Initiative
OSS	Open source software
PDF	Portable document format
PHP	HyperText Preprocessor
PNG	Portable network graphics
RSS	Really simple syndication
SPC	Secretariat of the Pacific Community
STREAM	Support to Regional Aquatic Resources Management
UNESCO	United Nations Educational, Scientific and Cultural Organization
URL	Uniform resource locator
XHTML	Extensible hypertext markup language

1. Background

The 1995 FAO Code of Conduct for Responsible Fisheries (Code of Conduct) (FAO, 1995) provides a policy framework for sustainable fisheries and aquaculture management. Many FAO Members have reported that the lack of reliable and timely information constrains the Code's implementation. As recognized in articles 7 to 12, two broad categories of information need to be available. These are general information about the Code, its goals, coverage, etc., and specialized and technical information of a research nature required to permit officials and stakeholders to make informed decisions about options and approaches to support the Code's implementation.

To gain a better understanding of these issues, a FAO study was carried out in 2004 to assess the nature of the information used and produced by selected specialists working in fisheries management (Webster and Collins, 2004). The surveys, case studies, citation analysis and literature review highlighted the breadth of subjects required, the historic depth of relevant information, the scale of information from local to global, and the diversity of information sources. Given this complexity, it is not surprising that a major effort is required to obtain the best information upon which to base decisions and policy. The lack of global information resources in developing countries presents significant, but not insurmountable, challenges and the study proposes strategies to help meet them.

Work so far undertaken by FAO to analyze information needs in support of the Code's implementation mainly reflects the situation of people affiliated with institutions and organizations with an appropriate infrastructure, usually a library and information and communication technology (ICT) support and information services. Hosting the 31st Annual Conference of the International Association of Aquatic and Marine Science Libraries and Information Centers (IAMSLIC, 2006) at FAO in October 2005 provided an opportunity to discuss information resource sharing and networking as the most cost-effective means to meet information needs. Emerging from this discussion was the fact that few organizations have a mandate that permits them to extend their library and information services beyond their own defined community. There appears to be no formalized mechanisms for a trickle down, or "request up", information flow between the various types and levels of library and information centre i.e. large universities, small research institutions, non-government organizations, community information centres, etc. Channels for improving the information flow should be investigated, especially given the trends towards the decentralization of fisheries management or at least some form of community participation in management. The opportunities for stakeholders in small, resource-poor organizations or at local level to obtain and disseminate information are limited and their needs should also be understood and met. Using the expertise and experience of larger information entities for the benefit of smaller organizations is one mechanism that these guidelines seek to address.

It has also long been recognized that the results of research and the development lessons learned are often lost because of inadequate opportunities to publish findings in developing countries. Capturing what is published has never been totally effective and much needs to be done to improve dissemination and the sharing of information, as well as its preservation for future generations. The digitization of information and its availability via the World Wide Web offer enormous potential and opportunities for improved access and dissemination. Better integration of the information generated in developing countries into the global pool of fisheries and aquaculture publications will increase the use and validation of research and development results and avoid costly and wasteful duplication of effort.

Stakeholders in developing countries generally are still waiting for reliable, high speed and cost-effective Internet access that is widely available in the industrialized world. Once connected, users must grapple with, and make decisions about, myriad technological

solutions that exist. These Guidelines on digital publishing are targeted primarily at small institutions and organizations in developing countries, in order to facilitate decision-making on how to publish and disseminate their information. The Guidelines are based on the years of experience of the Network of Aquaculture Centres in Asia-Pacific (NACA) and its partners. They focus on a practical approach to the production and dissemination of digital information for small institutions and organizations with limited resources.

1.1 The need to publish

The importance of research in fisheries and aquaculture is referred to throughout the Code of Conduct, as is the need to disseminate and share the results of research. Many of the highlighted quotes below from the text of the Code are extracted from its Article 12, which covers research in general. It is universally agreed that there is little point in doing research if results cannot be used, particularly in developing countries, where applied research is an essential component of economic and social development and has a high opportunity cost. Researchers in developing countries face huge difficulties in being able to publish the results of their work. Information technology (IT), and the general lack of access to scholarly literature, has been referred to as the Achilles Heel of African fisheries science and technology. Low scientific publication rates and a high rejection rate by publishers have a demoralizing effect on researchers, resulting in their emigration and loss to the economy. This situation leads to slow technological progress, poor management, the re-invention of wheels and a waste of scarce resources (Hecht, 2003).

> *12.3 States should ensure that data generated by research are analyzed, that the results of such analyses are published, respecting confidentiality where appropriate, and distributed in a timely and readily understood fashion, in order that the best scientific evidence is made available as a contribution to fisheries conservation, management and development. In the absence of adequate scientific information, appropriate research should be initiated as soon as possible*.

The "best scientific evidence" referred to in Article 12.3 also should be available and understandable to non-scientists i.e. managers, policy-makers and stakeholders at community level. This implies synthesis as well as analysis and communication in appropriate formats and languages. Information is not available and of little value if it is not comprehensible to the intended audience, whose needs should also determine the way information is packaged and distributed.

1.1.1 Validating the results of research and development

The information on the fisheries in a specific locality or a specific country is normally the most important information for the sector as a whole in that country. Organizing, managing and disseminating this information is an essential part of the research and development process. Ensuring that it is accessible by all stakeholders, including policy makers, research and development workers and resource users is a major challenge. Validating locally generated information and the results of research is impossible unless it reaches the appropriate audience and is ***used***.

1.1.2 Sharing the results of research

The Code emphasizes the need to create and share information across political boundaries. Sharing information at regional and international level introduces additional considerations, including possible language barriers and the need for standards and compatibility.

> *12.16 States should, where appropriate, support the establishment of mechanisms including, inter alia, the adoption of uniform guidelines, to facilitate research at the sub-regional or regional level and should encourage the sharing of the results of such research with other regions.*

Information ownership is an important issue, particularly in developing countries where publishing can present insurmountable difficulties. In these situations researchers often feel the need to protect their research results rather than share them with colleagues. The whole concept of information sharing is based on recognition of the originator. Such acknowledgement is difficult to achieve when there are limited opportunities to publish.

1.1.3 Lessons learned in the development process

Sharing the lessons learned in one community or situation can avoid repeating mistakes and can try to build and improve upon the successes. What may have been seen as a success in the past can suddenly become a disaster in the present: e.g. current high levels of overfishing and environmental degradation are probably the result of historical “success stories”. Only by documenting, preserving and making available this information over the longer term will lessons be learned. Time and resources can be saved by following the lead or example of others. Why start from “scratch” when somebody else has already done half of the work?

1.1.4 Learning from local knowledge

> *12.12 States should investigate and document traditional fisheries knowledge and technologies...in order to assess their application to sustainable fisheries conservation, management and development.*

Even though the management of local fisheries information can be problematic, its utility in fisheries management is important. As has been pointed out: "Local knowledge can be used to corroborate science data and to fill in gaps in the scientifically generated data. While local knowledge typically is not subject to the same peer review as scientific knowledge, triangulation with other data sources and comparative techniques can help validate it" (Scholz *et al*., 2004).

1.2 Where to publish

The debate on whether to publish in the peer-reviewed literature rather than at the other end of the spectrum as “grey literature” continues, particularly in the western academic world. Grey literature covers those publications outside of the readily available commercial publishing realm. Scientists in developing countries often do not have the same opportunities to publish in commercial journals. However, the emergence and improved dissemination of scholarly journals published in developing countries is helping to fill this gap. These journals often provide a better mechanism for publishing on applied research and development. Many secondary information services are being developed to give them a much wider

dissemination and higher visibility. A good example is African Journals Online (AJOL) <http://www.ajol.info>.

Science often relies on the peer-review process to validate information. However, fisheries and aquaculture management draws upon information from a much wider variety of sources, including those outside the peer-reviewed, and the quantity and value of grey literature should not be under-estimated. The emergence and increasing availability of the Internet together with the influence of the Open Access Movement are radically changing the debate on where and how to publish. Digital technologies and international networks are providing the opportunity to reduce publishing costs, allowing self-publishing and collaborative or syndicated publishing. In subject areas like fisheries and aquaculture with such a large proportion of grey literature, there are increasing opportunities to improve information access and dissemination.

1.3 Control of the publishing process

Small organizations and institutions with limited resources are faced by many constraints in implementing their own digital publishing system. These guidelines, based on the experience of NACA and its partners, are intended to assist them in deciding how best to manage and control their own digital publishing and dissemination. An important consideration from the outset is whether to go it alone, to work in partnership with related organizations or to investigate other opportunities, even at international level.

NACA, an intergovernmental organization created in the early 1980's, initially as an FAO project, promotes rural development in 17 Asia-Pacific countries through sustainable aquaculture <http://www.enaca.org>. NACA conducts development assistance projects throughout the region in various partnership arrangements. It supports institutional strengthening, technical exchange and the development of policies for sustainable aquaculture and aquatic resource management.

One of NACA's core activities is the development of communication and information networks amongst the member countries. Its Information Centre, staffed by a manager and one computer engineer, is charged with producing publications, developing a Web-based platform for storing and distributing digital documents, and providing technical assistance in communications to the members. The Information and Communication Program focuses on information in digital format as the best means of sharing it widely, quickly and affordably. Consequently, NACA produces all publications in digital form, delivered primarily by download from the Web site. CD-ROMs are also created for distribution to offline centres. Most publications are also produced in printed form as part of an ongoing commitment to accessibility. NACA creates its own content as well as collecting other appropriate digital content and communities of users are supported through its Web site.

The expertise gained by NACA and the lessons learned along the way of developing and implementing a digital publishing system could benefit other organizations that have not yet started, or are beginning to plan, their own system.

1.4 Emphasis on the users and target audience

NACA works in partnership with its Members, national centres that serve as focal points and regional centres selected by the Members to serve as lead centres to share in the regional activities and responsibilities of NACA. As such, its information products and services must respond to the needs and requirements of a broad spectrum of users. Many of the partners are small research organizations with limited resources and they need training and capacity

development in order to participate effectively in the information and publishing process. The feedback which they provide to NACA becomes an integral part of the evolving system. So, the users include those who publish as well as the target audiences that the authors seek to reach. Added to these user groups is the worldwide audience of Internet users seeking aquaculture information and who are not specifically targeted.

The partnership arrangement is a main feature of NACA's expertise in the planning and management of a digital publishing system that focuses on user requirements. Of prime importance is an understanding of the human interactions needed to build and sustain the system – i.e., technology is not the solution but a tool. It is important to choose the best and most appropriate system, but the biggest and most important investment will always be in human resources and training. This is why not all institutions have the resources to build and manage their own information and publishing systems. In many cases a better solution is to pool available resources in partnership with others. The sharing of resources includes the expertise of staff as well as the resources of information.

1.5 Aquaculture development

Article 9 of the Code of Conduct on Aquaculture Development includes:

> *9.2.4 States should establish appropriate mechanisms such as databases and information networks to collect, share and disseminate data related to their aquaculture activities to facilitate cooperation on planning for aquaculture development at the global level.*

Though aquaculture has been practiced for centuries in some countries, management of the sector is a fairly new concern. In fact, aquaculture was recognized only recently (March 2001) as an independent economic activity by the United Nations Statistical Commission. Accordingly, the collection of statistical data and other information on aquaculture separately from fisheries data is a recent endeavour at the national, regional and global levels and lags well behind systems for agriculture and capture fisheries.

However, the growing interest in aquaculture and the implications of its expansion, together with strategic concerns for sustainable development and trade, and for social and economic development, have created a need for a better array of numerical data of reliable quality and for other information that measures and describes trends of the sector. In many countries, the sector is developing rapidly, or is expected to do so, adding to the need for close and regular monitoring.

In recent years the demand for reliable data and information and for reporting on aquaculture has greatly increased, driven not only by the need to formulate and monitor sound policies and development plans, but also by new information and reporting requirements of international agreements and initiatives, and by the increasing public demand for transparency and accountability (FAO, 2005). The information needs of the aquaculture sector are diverse and include aspects of technology, education, training, legislation, and business management. The specific format for this diversity of information will depend on the user, and specific types of information may need to be available in several different formats in order the make the information widely accessible and useable.

NACA's philosophy and policy on publishing and disseminating information provide a good model for other organizations (i.e. the use of free and open source software; collaborative development; free and open access, fully digital information; additional output formats, such

as print and CD-ROM, in response to demand). These policies are part of the process of cooperation required for aquaculture development at the global level.

2. Guidelines on digital publishing. A practical approach for small organizations with limited resources

An old tradition and a new technology have converged to make possible an unprecedented public good. The old tradition is the willingness of scientists and scholars to publish the fruits of their research in scholarly journals without payment, for the sake of inquiry and knowledge. The new technology is the Internet. The public good they make possible is the world-wide electronic distribution of the peer-reviewed journal literature and completely free and unrestricted access to it by all scientists, scholars, teachers, students, and other curious minds. Removing access barriers to this literature will accelerate research, enrich education, share the learning of the rich with the poor and the poor with the rich, make this literature as useful as it can be, and lay the foundation for uniting humanity in a common intellectual conversation and quest for knowledge

- Budapest Open Access Initiative <http://www.soros.org/openaccess/read.shtml>

"When we speak of free software, we are referring to freedom, not price. Our General Public Licenses are designed to make sure that you have the freedom to distribute copies of free software (and charge for this service if you wish), that you receive source code or can get it if you want it, that you can change the software or use pieces of it in new free programs; and that you know you can do these things."

- Preamble to the GNU General Public License <http://www.gnu.org>

2.1 Introduction

The Internet is the most powerful network for exchanging information that has ever existed in human society. Its scope of coverage, accessibility and influence grow every day. Since you have taken the trouble to look at this document you probably recognize the potential of "digital publishing" – circulating your publications via the Internet – to increase the distribution and impact of your organization's work. However, if this is your first experience in working with Web sites and "computer issues" then you may be somewhat apprehensive about the technical difficulties of getting involved.

The good news is that even a small organization with a shoe-string budget and limited staff resources can establish an effective digital publishing system with a global reach. It just takes a bit of careful planning to ensure that the system you develop meets the needs of your stakeholders, is accepted by your staff, and importantly, that it is sustainable within the resource constraints of your organization.

Our objective is to provide **practical** guidelines that will lead you through the process of planning and implementing a sustainable digital publishing system. The guidelines cover:

- planning, building and maintaining a sustainable digital publishing system, focusing on a common scenario of setting up a Web site as a digital publishing platform for your organization;
- producing user-friendly digital publications and making them accessible;
- some recent international developments in digital publishing; and
- recommended software tools and technical resources for further reading.

The guidelines are specifically intended to serve organizations in developing countries with limited financial resources and few technical staff. We do not assume that you have access to a large team of people, a professional librarian or an IT expert. Hence, these guidelines concern development of a *simple and effective system that can be sustained with limited financial and technical support.*

The guidelines were prepared for the FAO FishCode Programme <http://www.fao.org/fi/fishcode.htm> as a contribution to assist the dissemination, accessibility and preservation of digital publications in support of implementation of the 1995 FAO Code of Conduct for Responsible Fisheries.

2.1.1 What is digital publishing, really?

When people think of digital *publications* they generally start by thinking about their written publications, but there are many other possibilities. You can publish anything that can be stored as a file on a computer including digital photographs, audio recordings, video footage, computer programs and databases. When we use the term “digital publication” throughout these guidelines we are really talking about *any kind of information that is packaged in a digital format*.

Digital *publishing* is the process of making that information publicly available. The most common way to do this is by putting your publications “online” as a Web page that can be viewed on screen or as a document file that can be downloaded. However, it can also include “offline” publishing through CD’s and other portable media. When we refer to digital publishing, we are talking about *the dissemination of information through any digital medium,* although most of our discussion is about applications that make use of the Web.

2.1.2 Advantages and limitations of digital publishing

Should you go digital? Digital publishing has many potential applications and benefits but for many organizations it is not a complete communication solution in itself. It is important to evaluate the potential benefits and limitations with regard to your stakeholders before you decide to get involved.

With careful planning, digital publishing offers the following advantages.

- ***Massively improved accessibility and circulation of publications.*** The sheer scale and worldwide nature of the Internet means that even the simplest of Web pages can be a highly effective communication tool.
- ***Low publishing costs.*** Most of the tools required to build Web sites and make electronic publications can be obtained for free. A small Web site can be developed and hosted for around $20 per year (yes, it’s true). Large offline collections can be circulated on CD for only a few dollars.

- ***Fast publishing.*** It is possible to publish a new document and inform people of its availability in only a few minutes. This makes “real time” reporting possible, as well as the provision of time-sensitive services such as market information.
- ***Control of publishing.*** The ability to self-publish frees organizations from constraints such as the strict language and subject matter controls imposed by journals. This allows publication of “grey literature” and other information that may not previously have been available in any form.
- ***Community participation.*** Many Web-based digital publishing tools are designed to be interactive, allowing groups of people to communicate and collaborate in the process of creating and publishing documents online. This allows the publishing process to be decentralized.

There are, of course, many limitations.

- Digital publishing does not improve accessibility for everyone. In most cases it is useful only to the sub-set of people who have access to computers and/or the Internet. However, as the Internet is a global medium the number of people that you can reach may be very large – perhaps a lot larger than your intended audience.
- Some degree of computer literacy is required to make effective use of a digital publishing system, and a somewhat higher level to plan, install and administer such a system. These skills are usually limited or absent in organizations that are just starting to get involved.
- Effective use of a digital publishing system often requires changes in an organization's internal workflow. Staff resistance to change is often very high, particularly in projects that introduce new information technology.
- Security and data loss are real and serious risks on the Internet. Many small organizations do not have a good understanding of how to protect digital information.
- While digital publishing is cheap, it is not free. Implementing digital publishing is unlikely to “save money”. It provides another, highly efficient way of circulating information – but at an additional cost.

The value of digital publishing to your organization will depend to a large extent on the nature of your stakeholders. In most situations it is best seen as a suite of *additional* tools that should be used in combination with other media for communicating with people, preferably through an integrated communications strategy tailored to meet the needs of your stakeholders.

Digital and Print

Four years ago when we began making electronic versions of NACA publications available for download from the Web, we attracted a lot of criticism. There is no substitute for printed publications, we were told. There's not enough Internet access in rural areas. You're wasting your time. These were, and are, valid points. We accept that only a small percentage of our stakeholders have Internet access. However, in terms of absolute numbers this group is very large because it is a global audience, and they can be reached very cheaply. Circulation of publications via download overtook printed copies in the second year of operation. We are now reaching more people than ever before, and with hardly any additional cost, but we still produce the printed publications. The people that access our publications by print and by Internet are two separate audiences.

2.1.3 Getting started

Development of a digital publishing system can be divided into three phases:

(a) ***Planning:*** Making a case for establishing a system, determining what it must do and designing a system that will meet them.
(b) ***Implementation:*** Construction and deployment of the system according to your plan, both in terms of setting up the technology and establishing staff workflow and operational procedures within your organization.
(c) ***Maintenance:*** Ongoing publishing activities and administration of the system including backups and security patches. This is where the "real work" starts, requiring an ongoing commitment of staff and resources to keep it moving.

We will deal with these issues in the succeeding sections but first, a note of caution: ***People are more important than technology.*** Developing a digital publishing system is not just about selecting and implementing effective technology, it is just as much about people because they are the most important part of the publishing system (Boiko, 1995). A lot of your time will actually be spent helping people to see the benefits of the system, to change the way they work and to make effective use of it. When we speak of digital publishing "systems", we are referring to both the technology and the people who use it. You need to engage your colleagues and your stakeholders in every step of the development process. Setting up the technology is actually the easy part!

2.1.4 Where to get help

If this is your first experience with digital publishing it may well be worth seeking feedback and advice during the planning and development of your project from people with prior experience. If you don't know anyone that has done this before, you may find it useful to consult people via an online discussion forum or e-mail list that is active in this area. Most digital publishing software projects maintain their own fora where you can ask questions of other users. They will usually be happy to provide advice on general project development and management issues as well as technical advice about the software. Another useful resource is the **Information Management Resource Kit (IMARK)**, a partnership-based e-learning initiative to train individuals and support institutions and networks world-wide in the effective management of agricultural information. IMARK consists of a suite of distance learning resources, tools and communities on information management. Four IMARK modules are available on CD ROM and via the Internet: *Management of Electronic Documents; Building Electronic Communities and Networks; Investing in Information for Development; Digitization and Digital Libraries* <http://www.imarkgroup.org/index_en.asp>.

2.2 Planning a digital publishing system

The planning phase is by far the most important in setting up a sustainable digital publishing system. It is also the most difficult to get right. The decisions you make will determine how effectively the system meets the needs of your stakeholders, how easy it is to set up and to maintain, and ultimately how successful it is. Your main tasks in the planning phase are to:

- Set a clear objective for the system, with some measurable indicators of success.
- Gather the support of your colleagues and management for the initiative. This is crucial – if you want ongoing funding and staffing support then you need to have the people in your organization on your side.
- Determine the requirements of the system.

- Design a system that will meet the requirements, both in terms of technology and staff workflow within your organization. Note that this comes *after* determining requirements!

2.2.1 Getting approval for the project

Projects involving information technology have a reputation for being technically difficult and expensive, which tends to make them unpopular with management. The benefits can also be hard to explain to staff without an IT background. How can you build support within your organization to turn an idea into reality? How can you persuade your manager to provide the resources needed to build and sustain the system? There are some general steps you can take to build the initial consensus needed to get the project moving.

Set a clear objective

The first questions you are likely to be asked will be: Why do you want to do this? What is the benefit? A clear objective is crucial to gaining support for your project and provides the reference around which it will be built. Issues to consider in setting an objective include (FAO and UNESCO, 2005):

- What is the need?
- Who has expressed this need and how do they relate to your organization?
- What is the purpose of the digital publishing system?
- Who is the target user group?
- What are the characteristics of this group, their capabilities and limitations?

The objective for the project should:

- Meet a real need of your stakeholders.
- Be consistent with the goals and priorities of your organization.
- Be measurable. You should also establish indicators of success.

Once you have established these points in your own mind you should discuss them with the key stakeholders and intended beneficiaries of the system, perhaps through meetings or informal discussions. This will help you to develop your ideas more fully and validate any assumptions you have made. It will also give you an indication of the existing level of support for the initiative and begin to engage people in its development.

An example of a *good* objective could be “to provide daily market price information to help fish farmers make informed decisions about stocking and harvesting their ponds”. Indicators of success could be the number of fish farmers subscribing to an e-mail news alert service and number of page views on a market price Web site, or an offline survey of shrimp farmer opinion conducted through target associations. An example of a *poor* objective (which we hear all too often) is “to build a Web site for our organization”. Developing a digital publishing system is not an end in itself! What do you plan to *do* with it?

> ***A clear objective***
>
> *It is quite common for organizations to begin developing a digital publishing system by focusing on the technology, without clearly defined goals as to what the system is actually meant to achieve or how to measure its success. When it becomes apparent that the system is not performing as desired, it often has to be modified, rebuilt or even abandoned due to cost constraints. If you find yourself in this situation, examine the project's goals carefully. Be wary of proposing additional technology and complexity as a solution. You may just need a better-defined objective.*

Identify potential project "champions"

To function effectively the system must be integrated into your organization's daily activities as closely as possible. To integrate successfully, the system needs to (Boiko, 2005):

- Capture information from a variety of internal and external sources, usually requiring a common editorial approach.
- Organize the publishing activities of different parts of your organization into a single overall system.
- Integrate within your organization's existing IT and information management infrastructure.

Different areas within your organization will play different roles in this process.

- "Core staff" will create most of the information and / or publications.
- Editors will help pull publications together and organize them into coherent collections.
- IT staff (if you have them) will build and maintain the Web site and other computer infrastructure that will make the digital publishing system possible.

Your project will be much more likely to succeed if you can identify the key people that can provide access to these skills and obtain their support. These are the potential "champions" of the project, people whose support you need to make it a success (Boiko, 2005). They may be managers within the organization, someone who can influence key decision makers or a donor that can provide essential resources. Ideally, they will be people who share a common interest in addressing the needs of the target group that the system will serve. It may be easier to approach them if you spend some time studying their interests, so that you can discuss how the project will be of benefit in terms that they value.

You should also give some thought to potential detractors. Does the project impinge on the activities traditionally "owned" by some other part of your organization? Who would be affected and why? You may be able to defuse potential problems before they arise by bringing potential detractors into the process early, giving them ownership of the project and the opportunity to address their concerns.

Publishing is not an IT issue

A common mistake is for an organization to view a digital publishing system as a "computer issue" that should be handled by the "computer people". IT groups generally do not have strong editorial skills and lack technical knowledge about the subject matter needed to organize collections of publications effectively. If your business units are dumping unstructured information directly onto your IT staff without a separate editorial process, they may not know what to do with it, and they may not have enough people to perform this role.

Develop a concept proposal

Equipped with a clear objective and hopefully some champions to support the initiative, you are now ready to prepare a concept proposal for submission to management. This is often the first formal step in gaining approval to undertake the project. The objective is to bring the key stakeholders together so that they can build consensus on the goals of the project and to develop a written proposal that can be used to "sell" the concept.

You could begin by discussing the concept with your champions. Draft an outline as the basis for discussions and, with their support, convene meetings with the key stakeholders to develop it more fully. The outcome of the meetings should be a formal concept proposal endorsed by the major stakeholders. The concept proposal should contain:

- A project statement that clearly summarizes the need for the project and its objectives as they relate to the goals of the organization.
- A tentative indication of the resources that will be required (funding, staff, equipment) and how long it will take to implement.
- Some key indicators that can be used to measure the success of the project.

As digital publishing is usually an ongoing activity, it is vital that the proposal makes it clear that the initiative will require ongoing support. If management approves the concept proposal, you will be ready to begin detailed planning of the project.

2.2.2 Determining requirements

The first step in planning a digital publishing system is to determine the operational requirements. These should relate to the *needs* of the people it will serve, both within your organization and externally. Forget about technology for the moment. What kinds of information do you need to provide to meet your objective, and what are the best ways to provide it to stakeholders?

Analyzing stakeholder requirements

Conducting an information access survey (IAS) (STREAM, 2005) is a good way to get accurate information about your stakeholders' requirements, and how you can communicate with them. The purpose of an IAS is to:

- Identify key issues about people and what information needs they have.
- Identify what media sources are available, what strategies people use to get their information and how cost-effective these are.
- Suggest methods of communication that are useful for different groups of people.

An IAS should:

- Take into consideration the needs of the target group.
- Involve as many people as possible.
- Be socially and culturally acceptable.
- Be flexible, so that it can be modified to suit different circumstances.
- Provide recommendations that are easy to put into practice.

For each stakeholder group, issues to consider include:

- What is the geographical area of the survey (if relevant)?
- What are the existing communications networks available to your stakeholders (digital, non-digital and personal)?
- What are their needs?
- What kinds of information would be useful to them? Where can you get it?
- How would it help them?
- How do people prefer to get this kind of information?
- What works well, and why?

An IAS will give you an indication of how effective different media are to reach your stakeholders. You may, for example, find a clear preference towards e-mail. You may also find that digital media are not a very effective way to reach your target group, or that they are only useful to part of it. It is quite likely that you will discover that a broader publishing strategy using several media channels, both digital and non-digital, may be more effective.

Obviously, the extent to which you can carry out an IAS and the way in which you do it will depend upon your resources and the level of access you have to your stakeholder groups. However, it is quite likely that your organization already has a substantial amount of information about its stakeholders, if you can find it. This may be contained within previous studies and surveys, contact and distribution lists. Field officers, publications people and librarians, if you have them, may be able to give you some good information about your stakeholders requirements based on their experience.

You will also have stakeholder groups within your organization. Most commonly, these are people producing information that they want to publish for a particular client group. They will inevitably see a "need" to publish all of it, but in reality only some kinds will actually be useful to the organization's clients. Look for areas where the objectives of the organization overlap the needs of its clients. The strong points of your digital publishing system will likely be in these areas of common interest.

Technical requirements

Your organization may have policies that restrict or dictate what technology you can use. Issues to consider include (Boiko, 2005):

- Does the digital publishing system have to be integrated into a larger, existing Web site framework or with other IT infrastructure? You may well have to use the same technology to ensure compatibility.
- Does your organization have an existing investment in software licenses? They might expect you to make use of what is already available.
- Does your existing staff expertise lie in a particular software environment or platform? If so, you would need a very strong reason to use anything else. Retraining staff is slow and expensive.

If your organization's resources are limited, try *not* to spread it across multiple technologies. This is one of the worst things you can possibly do. If you have an IT department, do not be surprised if they *tell* you what you are going to use.

Documenting the requirements

After you have finished determining the requirements of the project you should produce a document outlining:

- Which groups of people the system will serve.
- What their information needs are.
- Where the information will come from, in-house or external sources.
- The most appropriate way(s) to deliver it.
- Technical requirements and constraints of the project.

You can think of this document as the "deliverables" of the system. You should seek to have it endorsed by key players within your organization, and by management, as it will form the basis for justifying requests for funding and staff resources.

2.2.3 Designing a digital publishing system

Good planning will save you an incredible amount of time, and possibly a lot of money as well. Now that you have identified the requirements of the system, the remaining planning tasks may be broken down into deciding:

(a) What the digital publishing system will actually do, i.e. which functions does it need to perform to meet the requirements.
(b) Who will do what, i.e. developing a staff workflow to operate the system.
(c) How the system will actually work, i.e. the hardware and software infrastructure needed to build it.

Steps (a) and (c) may be respectively considered as the *logical design* and *physical design* of the system. In logical design you decide at a conceptual level what the system should achieve. In physical design, you decide how to actually build a system using hardware and software that will meet these goals (Boiko 2005).

> ***Don't put the cart before the horse***
>
> *A common mistake is to buy computer hardware and software before developing a detailed plan on what the system should do. You may later discover that your purchase is unsuitable or imposes undesirable constraints on what you wish to achieve. Plan first, and then look for the most suitable technology to implement your goals. The functionality required of the system should determine what technology you use, not the other way round!*

Logical design: What will it do?

Your goal in logical design is to outline the main components of the publishing system. This is where you take the needs expressed by your stakeholders and try to match them with the real-world capacities and resources of your organization. It should be hardware and software neutral – that is, it should not make any assumptions about the technology that will be used to create the system (Boiko, 2005). You will look for the right technology to implement your design later on. The basic considerations are:

- ***Stakeholders:*** Who are the main stakeholders the system will serve?
- ***Information needs***: What kinds of information do you have that your stakeholders value?
- ***Preferred communication methods:*** What formats would be the most useful and convenient for your stakeholders?
- ***Authors:*** Which groups of people will create publications for the system?
- ***Information sources:*** Where do you expect to get your information? Is it produced by internal sources? Can it be obtained from external partners?

One of the benefits of conducting an information access survey is that you will already have most of this information to hand, so you should be able to prepare a logical design quite easily. For example, if you worked in a small aquaculture research organization, funded primarily by donors, whose mission is to promote usage of better management practices in fish farming you might produce a simple logical design that could be summarized as in Table 1.

Table 1

External stakeholder group	Fish farmers	Aquaculture researchers	Donors
Information needs	▪ Aquaculture news ▪ Market prices ▪ Better farming techniques ▪ Health management	▪ Research findings ▪ Trends in aquaculture production	▪ Alleviation of rural poverty
Requirements to meet needs	▪ Fish farmers are provided with practical information on better farm management practices to help reduce the incidence of disease and stock loss, and timely market information that will help them plan their next harvest.	▪ Researchers are informed of the latest developments in aquaculture research, to help them support fish farmers	▪ Donors are kept informed of our organization's achievements in relation to improving the livelihoods of fish farmers.
Information sources	▪ Commercial news providers ▪ Industry newsletters ▪ Internal sources	▪ Journals ▪ Personal contacts with partner institutes	▪ Internal sources
Preferred communication methods	▪ Monthly meetings of farmer cooperatives (personal contact) ▪ Printed newsletter (no Internet access)	▪ Web site. They spend most of their time working on the computer. They like to download long publications as PDF files.	▪ E-mail newsletter. They only react to information we send them personally. Printed publications rarely elicit feedback.

Based on this very simple example, you can begin to make informed decisions about what your digital publishing system should include. For instance:

- A Web site focused primarily on aquaculture research news, targeted at aquaculture scientists. This group prefers that short "news" stories be presented as Web pages that they can browse, and they like to download longer documents as PDF files for easy print out and to e-mail to one another.
- An e-mail newsletter carrying "good news" press releases about the organization's achievements in improving the livelihoods of farmers and rural communities, targeted primarily at donors.

It is also apparent that the digital components of the publishing system will not reach the farmer stakeholder group. This illustrates the need to use digital publishing in combination with non-digital media: a printed newsletter will also be needed, carrying information on market prices and practical information on better management practices that farmers can apply in their daily work. It may be useful to send staff with this information to participate in meetings of fish farmer associations (or at least, to ensure the printed newsletter is available at the meetings).

As with most other steps in the planning process, you should summarize your conclusions in a brief report and circulate it for discussion, refinement and approval. Keep your colleagues involved!

Workflow: Who will do it?

> ***Big planning mistakes***
>
> *Digital publishing projects often fail. There are three main causes that seem to occur over and over again. We have listed them here so that you can address them throughout your planning process and avoid becoming another depressing statistic! The top three "killers" of digital publishing projects are:*
>
> - ***The staff are ignored.*** *Project managers often see digital publishing as a technical challenge. They focus on the technology issues and neglect to develop the human side of the digital publishing system. You need to develop a network of people that will supply and update information or the system will be empty and useless! You also need to help your staff come to terms with the new system - don't assume that they will use it just because you built it.*
>
> - ***No ongoing support.*** *We have observed that Web sites and other information systems are often developed on a "project" basis, with no funding or staff to maintain them once the initial construction is complete. They quickly fall into disrepair and become obsolete. There is no point in developing a digital publishing system unless your organization is willing to make an ongoing commitment to maintain it (cash, information and staff). You must ensure that your organization fully understands this point and is supportive of the initiative before you start.*
>
> - ***The system is too complex.*** *Organizations often build a system that is simply beyond their technical capacity to maintain. It is best to start with a simple system that requires a minimum of financial and technical input, particularly if you don't have much money or expert IT staff. You can expand it later on as you gain experience.*

Digital publishing systems do not run themselves. Writers are needed to prepare new publications, editors to ensure consistency and quality, and IT staff to maintain the infrastructure and perhaps to update content. Ideally, this is a job best performed by a team. However, it is quite likely that your colleagues will not understand this initially. They may think that it should all be done by the "computer person". You can help people to understand the processes that are required by developing a workflow for the system. A workflow is simply an agreed set of operational procedures that need to be carried out to maintain it. The benefits of developing a workflow include (FAO and UNESCO, 2005):

- It creates awareness of the different tasks that must be undertaken and helps people to understand their role in the overall process.
- It helps to ensure that management allocate adequate resources to perform each task.
- It creates a trail of visibility and accountability for ensuring that jobs are completed and on time.
- It promotes efficiency by setting out a standard procedure, roles and responsibilities.

A workflow document should:

- Identify the major tasks that must be performed in the publishing process, including creation and editing of publications and maintenance of the Web site.
- Identify who is responsible for each task and the role they will perform.
- Set a time frame for completing each task.

- Identify key decision makers and authorities throughout the process.

Examples of issues to consider are:

- What publications will be created?
- Who are the authors?
- Who are the editors?
- Who will package them into electronic formats for distribution?
- Who will approve release of publications?
- Who will update content on the Web site?
- Who is responsible for technical maintenance of the Web site?
- Who has managerial responsibility for different aspects of the publishing system?

Editorial responsibilities are one of the areas most often overlooked in establishing any kind of digital publishing system. Without an editorial process, it can be difficult to ensure quality control and to standardize approaches. The need for a strong editorial process increases with the number of authors and with the degree of decentralization.

Workflows must be agreed by the staff who will perform them to be of any practical value. An agreed workflow may emerge naturally as part of your discussions on planning the site. If not, hold consultations with key players to discuss roles and responsibilities. You should document the outcomes of discussions and seek the endorsement of your management. The Workflow Document will provide people with the authority they need to perform the required tasks and also ensure that they receive appropriate recognition for the work that they do.

A simple workflow for the hypothetical aquaculture research organization's publishing activities could be summarized as shown in the following table (your own workflow would also include deadlines for each step):

Table 2

Publication / step	Who	Timeframe
Web site		**Updated weekly**
Authors	Research officers	1 week
Editor	Press officer	1 day
Format / upload	IT officer	1 day
Approve release	Director	1 day
E-mail newsletter		**Updated weekly**
Authors	Journalist (consultant)	1 week
Editor	Press officer	2 days
Format	IT officer	1 day
Approve release	Director	2 days
Publish	IT officer	N/A - instant
Printed newsletter		**Produced quarterly**
Authors	Extension officers	3 months
Editor	Press officer	1 week
Layout	Outsource to printing company	1 week
Approve release	Director	1 week
Print	Printing company	1 week
Distribute	Administrative staff	2 weeks

In this example, the organization's own research staff will write the technical content, while the press officer will write general news items. A professional journalist will be hired to prepare sensitive publications aimed at donors. The organization's press officer has agreed to take on editorial responsibility for all publications, but internal policy dictates that the

Director must give final approval before release. The e-mail newsletter will carry information about new resources on the Web site, so it will be synchronized with Web site updates.

Good governance

Information and communication technology (ICT) use needs to be integrated with very basic content generation and organizational structures so that ICT contributes effectively and efficiently to useful information flows and services in an organization. Content management also includes appropriate governance, which includes policies, strategies, monitoring and evaluation and quality assurance. The publication of digital documents is only the tip of an institutional "iceberg" in terms of the organization needed to make effective and efficient use of ICT (APAARI 2004).

2.2.4 Staff considerations: recruitment and training

If you can afford it, it is a good idea to hire a professional Webmaster to plan and maintain your Web site, preferably one with some managerial and communication skills in addition to general IT expertise. Having appropriate technical expertise in-house will help get the system up and running quickly, but a Webmaster with additional managerial and communication skills will be able to guide you in planning an effective system, in helping to plan workflows and in building the capacity of your regular staff to contribute to digital publishing activities. Note that IT is a very broad field with many specializations. If you hire IT staff, do not assume they can do everything. Make sure they have the specific skills you require, i.e. Web site *management*.

Unfortunately, most small organizations cannot afford to hire additional specialized staff. In practice, the responsibility for Web publishing is generally handed to existing staff, who frequently have limited prior experience or no IT background at all. This is not an insurmountable problem, particularly if you use content management systems (CMS) that minimize the need for technical knowledge, but you should invest in staff training as far as your budget permits. Basic training courses in Web publishing and related computer use are often surprisingly cheap. If you cannot afford to send people for training, you may still be able to provide them with books instead – along with time to read them and to learn. Passing on whatever is learnt to other staff through regular in-house tutorial sessions or discussions can also be a valuable exercise to build capacity. The important thing is to provide some sort of regular mechanism for staff learning and sharing, and to ensure that this activity is clearly seen as sanctioned by management. You should be able to get a broad indication of training needs (both IT and non-IT) from your workflow document, and through discussions with the staff concerned.

The type of training that is needed will depend on the role of each staff member and also on the type of technology you use. However, training courses in Web publishing generally fall into three categories:

- ***Basic Web site design***. Most commonly, these focus on how to hand-code Web pages in HTML, usually with the assistance a visual design program such as Front Page or Dreamweaver that provides a graphic user interface to make the process a bit easier. The emphasis is usually on learning HTML code, creating simple page layouts, and basic skills such as uploading files to a remote server. These courses are suitable for anyone who is reasonably computer literate.

- ***Advanced Web site design***. These are essentially about development of custom database-driven Web sites using scripting languages such as PHP. They essentially concern programming, and are not suitable for beginners.
- ***Building Web sites with content management systems*** (usually, a specific portal system will be named such as XOOPS or Joomla!). These kinds of courses are offered much less frequently, but are probably the most appropriate for small organizations as they concern use of programs that automate Web publishing. Such courses are generally suitable for anyone that is reasonably computer literate.

HTML is certainly a useful thing for anyone involved in Web publishing to know, but it is not essential to have an in-depth understanding. As discussed in the next section, modern technologies are increasingly making the hand coding of Web pages redundant. In most situations we feel that it is better to focus on tools that automate the technical aspects of Web page construction. Why build pages by hand when a machine can do it for you faster and better? In our view, training people to use a content management system is a much more sustainable proposition in a resource-limited organization than teaching them how to code pages in HTML.

2.2.5 Building an information system

Once you have decided the services the Web site should provide, you can begin making decisions about what it should look like, and what software you will need to build it. This is the physical design stage.

Introducing content management systems

The "traditional" way to build a Web site is to code "static" Web pages by hand, perhaps with the assistance of Dreamweaver or a similar visual editor. In a static Web site each page is a fixed object that must be created and saved in advance before it can be viewed, just like a book. The problem with hand coding pages is that it is laborious and requires specific technical knowledge that is in short supply in small organizations. As static sites grow they become increasingly difficult to maintain, as all changes must be made manually. Eventually, IT staff cannot keep up with demand and become a "technical bottleneck" to updating the site.

Small organizations need tools to make the process of Web publishing easier. These tools need to be simple to use, because IT expertise may be low; they need to be flexible, so that they can be customized to suit different applications and purposes; and they need to be cheap because of resource constraints. These are the needs that CMS are designed to address (Pedersen, 2004).

CMS are computer programs for building and managing Web sites. They automate most of the technically demanding tasks such as HTML coding of page layout and formatting of text through the use of templates. They also provide tools to organize your information and to administer your Web site. To update a CMS Web site you simply submit information via a form in your browser.

> ***Updating a CMS Web site is as simple as pasting information into a submission form and pressing a button. The alternative – manually preparing HTML code – is much slower and requires specialized technical knowledge.***

CMS Web sites are usually "dynamic" or "database driven" Web sites. The concept of a "page" is quite different in a dynamic site. Pages are not fixed objects prepared in advance;

rather they are instantaneously constructed on demand and the contents can change to suit different circumstances! When a page is requested the server draws the necessary information (say, a news story) from a supporting database and assembles the HTML code for the page layout "on the fly" using preset templates to create the requested page elements, much in the way that a set of Lego blocks can be stuck together in different ways to create a range of different objects. As pages are automatically generated from templates rather than individually hand-coded, dynamic Web sites offer tremendous efficiency advantages.

One of the main advantages of using a CMS is that it removes the need for technical knowledge to update a Web site (knowledge of HTML code is not necessary for general use), allowing non-technical staff to publish their own information directly or even manage sections of a Web site themselves. This increases the pool of staff available to contribute to the site and creates incentive for them to get involved. Since information can be submitted and the site managed via a browser, both the content submission and management processes can be decentralized. Staff can even post updates from the field or while traveling, and contributions can also be submitted by the organizations' partners or even by the general public if you choose to allow it.

There are many different kinds of CMS available, but they may loosely be divided into two categories:

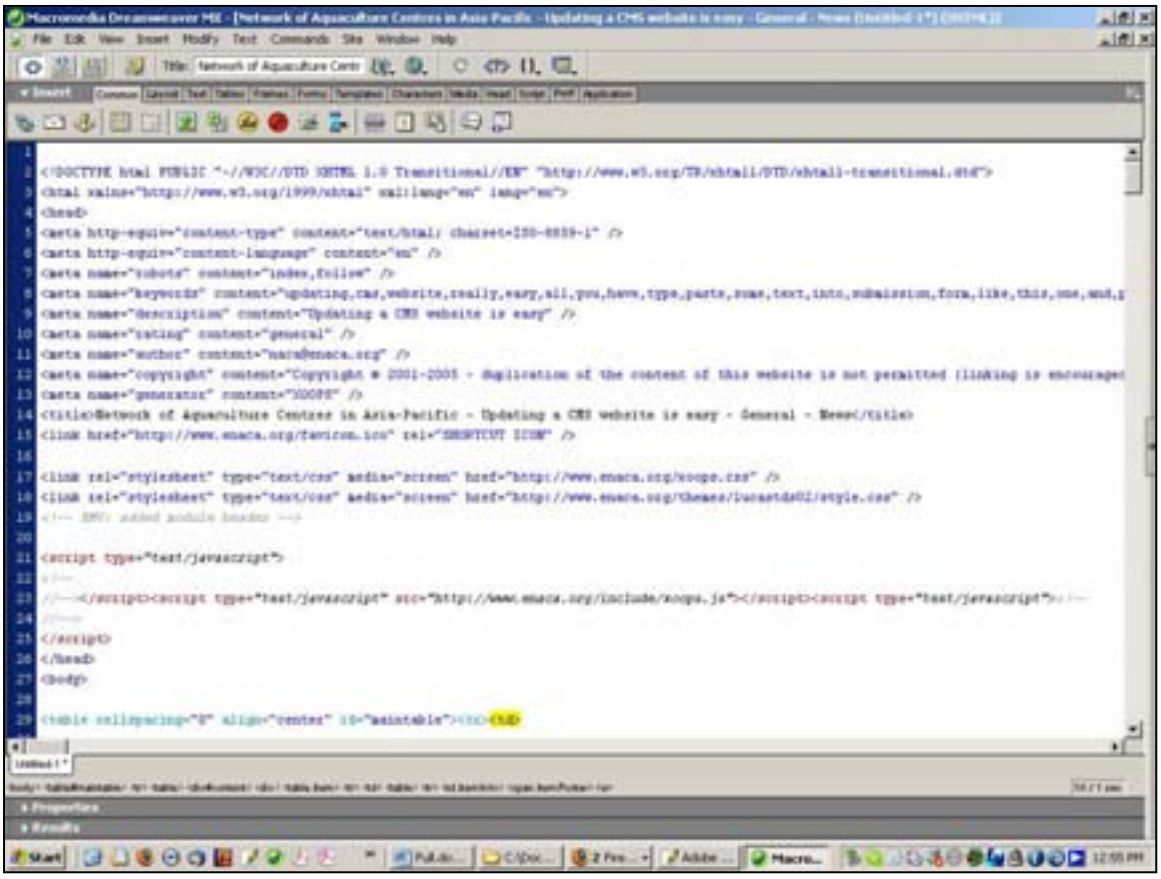
- ***Portal systems***, which are designed as flexible platforms for managing many different kinds of content. The functionality of a portal can be extended by adding "modules" or "plug ins" that are designed to perform specific tasks such as managing news stories, books or photographs. They usually come with a wide range of functionality and are easily customized, making them excellent tools for building Web sites to suit most common purposes. The strength of portals is that they provide a common framework for linking different applications together.
- ***Purpose-built CMS systems***, which are designed to manage a specific kind of content. Digital repositories are a good example, but photo galleries and most kinds of "information system" can also be considered purpose-built CMS. They tend to be much better at their designated task but less flexible and very difficult to integrate into other systems as they lack the common application framework of portals.

CMS make digital publishing on the Web fast, efficient and *easy*. They remove many of the technical barriers and allow better use of staff resources, making them excellent tools for resource-limited organizations, and a superior alternative to hand-coding HTML pages.

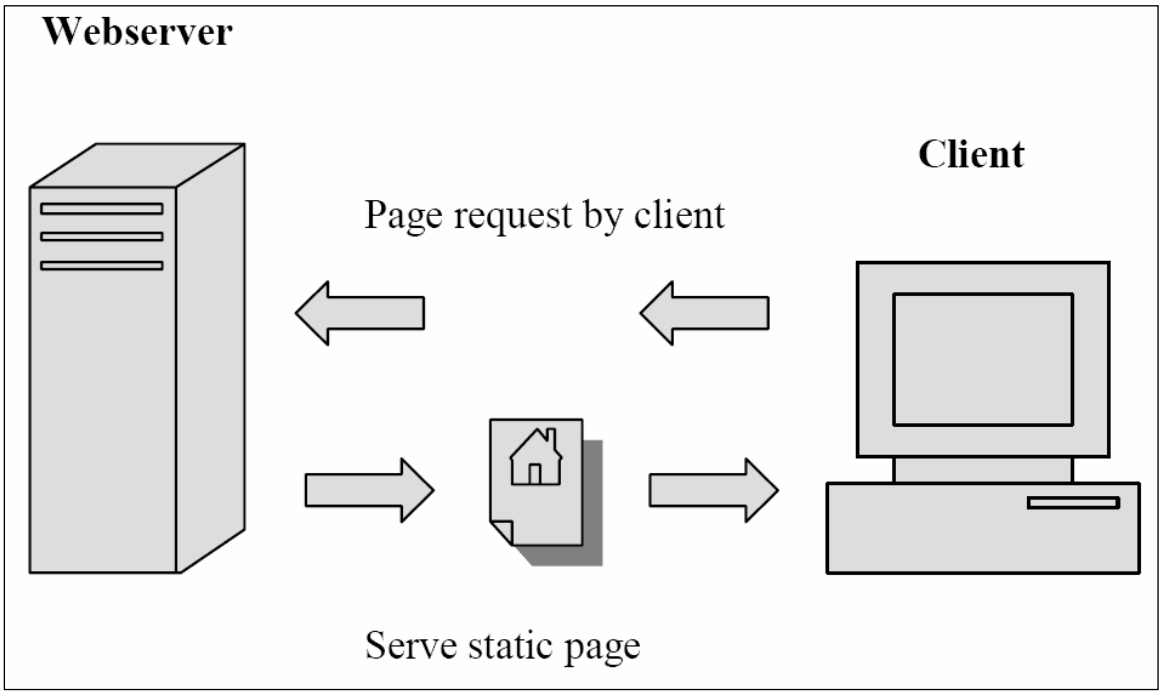
The kind of CMS you choose will depend upon your specific goals and circumstances. In the next section we will look at CMS options for building a general-purpose portal Web site suitable for use as your organization's digital publishing platform.

CMS Submission Form

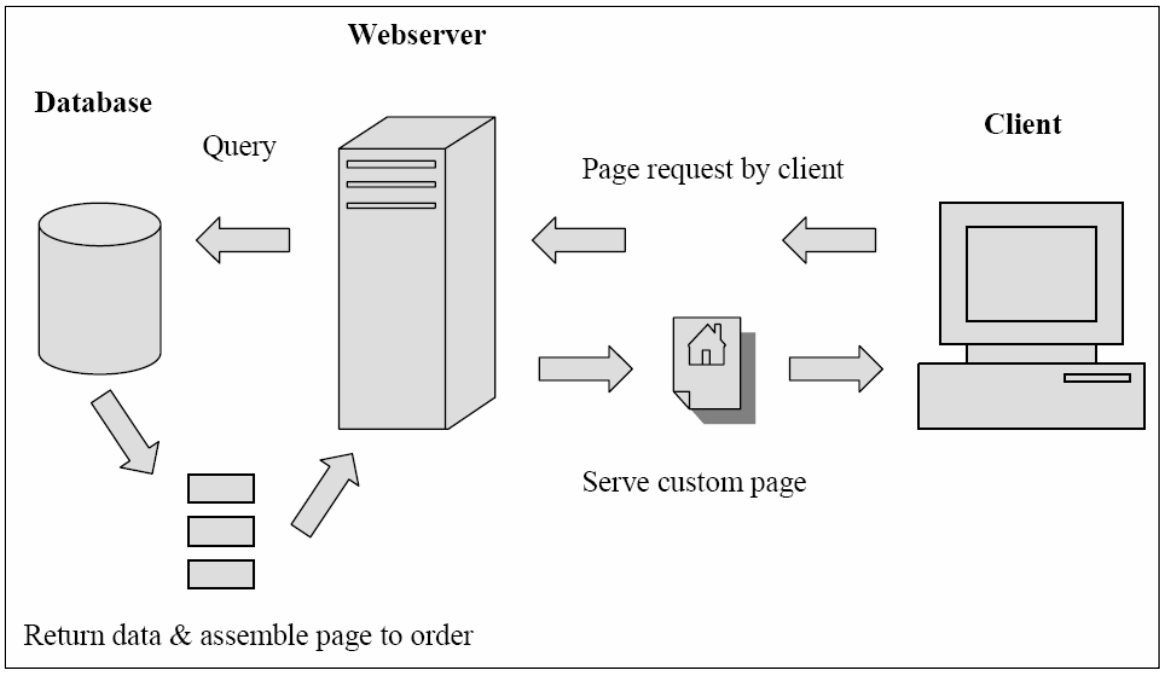
HTML coding



Static Web pages are prepared in advance



Dynamic Web pages are generated "on the fly"



2.2.6 Choosing a content management system

The primary considerations in choosing a content management system – or any software really – are that it performs the functions that you *need*, and that it is *sustainable* within the resource constraints of your organization – financial, technical and human.

Open source software (OSS)

Open source software is different from proprietary software. With OSS, the programming code used to create software solutions is available for inspection, modification, re-use and distribution by others. It is often assumed that open source software is free of charge but in some cases it can be purchased for a fee as well. The concept of free, in this context, emphasizes what can be done with the source code rather than its cost. OSS incorporates the concepts of community, public good, non-commercialism and issues of intellectual property, copyrights and patents. Underlying much of the discussion is that "information" in general, and "software´ as a means of delivery is unlike other goods and services. Central to the discussion are the issues of when and if information should be owned versus shared, what is the value of software, and when is it considered a commodity (Dravis, 2003).

If your resources are limited then it makes sense to consider using an "open source" CMS. One of the main attractions of open source CMS is that the total cost of ownership is generally very low, compared to commercial "closed source" systems. This is because:

- The software itself is usually provided for free, or for a relatively small fee.
- They make use of other open source software to provide essential supporting services such as databases, which can be very expensive to buy in their own right.

- Support is typically obtained for free through online communities of users.
- Ongoing development of the product and software upgrades are usually provided free by the user community.
- The source code is provided so that you can customize it if you wish, and if you have the knowledge to do so (Pedersen, 2004).

Further information on Free and Open Source Software (FOSS) and the Free Software Directory can be found at UNESCO's FOSS Portal <http://www.unesco.org/cgi-bin/Webworld/portal_freesoftware/cgi/page.cgi?d=1>. Also useful is the Web site of the Free Software Foundation <http://www.fsf.org/>.

2.2.7 Suggested content management systems

There is an intimidating array of CMS available. Finding the one that is best for you can be difficult and may require a lot of research and experimentation. To spare you from this experience we have suggested a few good open source CMS below that will meet most "typical" requirements. All are **free**, and make use of other mainstream open source software to provide supporting technologies, specifically the "LAMP" combination of Linux, the Apache Webserver, PHP scripting language and MySQL database. A convenient way to install a LAMP environment on a local computer for testing purposes is to download Appserv, which bundles them together with a single, easy to use installer. It is available from <http://www.appservnetwork.com/>. A glossary of technical terms in Appendix 1 explains some of the IT terminology.

XOOPS
XOOPS is the portal system selected and used by NACA. We like it because it is simple to use, offers a wide range of functionality and is very stable, secure and easily customized. The product is under active development and excellent support is available through a large online community. Translations of the XOOPS interface are available in many languages, including Chinese, Spanish, French, Arabic, Japanese, Russian and local support sites are available in more than 20 countries. Really simple syndication (RSS) support is available for many modules, making it easy to share your information with other Web sites via syndication feeds. You can download XOOPS from the main support site <http://www.xoops.org>. Good documentation is available including an installation guide and an operations manual, along with documentation for some of the more popular modules. If you would like an overview of what XOOPS is capable of then download the document "A visual introduction to XOOPS".

Wordpress
Wordpress is a Web publishing system designed for personal use, commonly used to create Weblogs or small discussion sites. It excels in terms of simplicity and ease of use, and offers a limited but sensible range of features with tools to manage multiple-author publishing, user interaction and exchange of data/interaction with other sites through RSS feeds, pings and trackbacks. The functionality of the system can be extended through many third-party modules. Translations are available for many languages and it is relatively easy to develop new ones. A disadvantage of the system is that unlike the more advanced portal CMS, it does not have an integrated system for controlling page layout and customizations must be made manually, although minor alterations are relatively simple to achieve. WordPress is well suited to an organization that wants to run a simple "news" site, with or without reader feedback/comments. Wordpress is available from <http://www.wordpress.org>. Excellent documentation is available.

Joomla!
Joomla! is a very popular portal system that has recently arisen as an offshoot of Mambo. Installation is very simple. An attractive, intuitive user interface makes it highly effective at

presenting information and easy to use and administer. Joomla! is a good choice for simple Web sites but is somewhat more difficult to upgrade and customize than XOOPS. Download Joomla! from <http://www.joomla.org>.

Xaraya
Xaraya is an advanced content management system. It is an extremely flexible system, well suited to sites requiring a high degree of customization. While installation is easy, the flexibility of the system also makes it considerably more complex and difficult to configure. It is a powerful tool but is best used after some experience has been gained with simpler systems. Download Xaraya from <http://www.xaraya.com>.

2.2.8 Surveying products

If you have specialized requirements or would like to investigate alternatives to those suggested above, a good way to start is to draw up a list of features required by your logical design, and then survey the range of available products to find those that provide the features you need and fall within your budget range. There are several Web sites about content management systems that provide information and tools that you can use to screen potential candidates:

- CMS Matrix: Provides very detailed tools that you can use to compare the feature sets of more than 370 CMS, both commercial and open source (but be aware in evaluating the information that CMS Matrix is the vendor of one product). <http://www.cmsmatrix.org/>.
- Open Source CMS: Provides demonstration installations of more than 30 open source portal systems that you can experiment with online, without having to install anything yourself. The demonstration installations are reset periodically. <http://www.opensourcecms.com/>.
- CMS Watch: Provides analysis and reports on content management, records management and search systems. There are some excellent articles available on various aspects of software selection, project management and standards. <http://www.cmswatch.com/>.

2.2.9 Short-listing products

If your requirements are fairly “normal” then the list of possibilities will still probably be quite long. However some of the “non-feature” qualities of a CMS are also important, and you can use these to filter out the better systems. Things to look for in a good CMS are:

- ***A large user base.*** This is an indicator of how robust software projects are. Larger projects also tend to offer more features and evolve at a faster rate as there are more people developing the system.
- ***Good support options.*** Can you get help if you have a problem? Large open source software projects tend to provide good (and free) support via online user communities, but small projects may not offer any support at all. Take a look at the project’s discussion forums – is there any sign of recent activity? Sometimes paid support packages are available from the developers, or from third party providers.
- ***A clear development path.*** All software has to be upgraded eventually. There is nothing worse than getting “stuck” in an abandoned product! Check that it is still under active development, and that it has a coherent development process. As community efforts, open source projects sometimes fragment into different projects or collapse when key developers leave. Large, well-organized projects will publish a “development roadmap” that outlines the direction and future enhancements of the software.

- ***Compliance with Web standards*** such as XHTML and CSS to ensure consistent representation of pages between browsers. Support for other standards such as Dublin Core metadata, and data import / export via XML may also be useful.
- ***Use of "mainstream" supporting technologies*** (such as "LAMP" – the Linux operating system, Apache Webserver, MySQL database and PHP scripting language) will be easier to integrate with other "mainstream" IT infrastructure.
- ***Compatibility with existing IT staff skills and infrastructure.*** If your staff is experienced with a particular set of technologies – such as LAMP – then you should seriously consider using a CMS that makes use of their expertise.
- ***International.*** Check that the system is available in the language(s) that you need. Can it support multiple languages simultaneously? If you will be working with non-Latin character sets then support for Unicode (UTF-8) is highly desirable. Unicode is a character encoding scheme that allows consistent display of virtually all character sets.
- ***Good documentation.*** A user-friendly manual can save you an *enormous* amount of time and frustration, and cut down on training requirements. Many software projects are badly documented, or have no documentation at all, particularly in the open source world.

It is important to note that most of these points also have financial implications. In calculating the total cost of owning a system you should consider the staff salaries and support services that will be required to learn how to use it and to maintain it in addition to any licensing fees. Even though the software may be free, the total cost of implementing a CMS can be quite high if you are not careful!

2.2.10 Testing products

The best way to select a final product is to test all of the systems on your shortlist. One option may look better than another on paper, but the only way you can *really* find out if a CMS meets your needs is to try it. A good place to start is the live demonstration installations at **opensourcecms.com**, mentioned above.

When evaluating a CMS it is a good idea to seek a cross-section of opinion from the people that will be using it. IT staff tend to place too much focus on the technical merits of a system. Non-IT staff are more interested in its actual "usability" and can provide feedback about whether a system is convenient and intuitive to use…or not!

There is considerable variation between CMS in the way they implement "standard" features. Important considerations are:

- ***User rights management system.*** How much control does the system offer over what parts of the site people can access, and what they are allowed to do?
- ***Workflow management.*** Does the system offer tools for managing submissions from users and enforcing editorial policies?
- ***Page layout controls.*** How easy is it to set up or change the page layout? Can you apply changes to individual pages or do they affect the whole site?
- ***Module installation.*** Is it easy to customize the system by installing new "modules" (the term varies between systems) that add extra functionality?
- ***Site configuration.*** How easy is it to set preferences for your site? What metadata does the system allow you to assign to pages and other resources?

2.2.11 Getting a domain name

Your Web site needs a domain name. You can obtain a vacant domain name by "registering" it with a company that provides domain registration services, usually known as a "domain registrar", in return for an annual fee. Registrar companies will normally provide a search tool on their Web site where you can type in domain names to see if they are still available. If the name you want is taken, some of them will also suggest similar names or alternative extensions (like .net or .org).

Registration is for a fixed period, normally for a minimum of one year but some companies will allow you to register it for up to 100 years if you really feel it is necessary. Most people usually register their domain for between 1-5 years. Normally, the company will e-mail a reminder notice to you when it is time to renew your domain registration. It is important that you keep your contact details up to date because once your domain registration expires the name goes back into the "vacant pool", where it can be registered by someone else.

When registering a domain be aware that many companies offering this service (including some of the best known companies on the Internet) are actually resellers. It is better to register your domain name through a primary registrar company as resellers may not be able to offer you full control over your domain name directly. For example if you later wish to transfer ownership of the domain later on you may find that a reseller will refer you back to the underlying registrar company, who may not be able to act until the reseller has agreed to release the domain and you have proven your identity to both of them, leading to substantial delays and frustration. NACA has had positive experiences with Network Solutions <http://www.networksolutions.com>, although they are relatively expensive. In any event, it is best to register your domain with a reputable primary registrar company. Avoid companies that offer "free" domain registration, unless it comes as part of a hosting package that you are paying for.

2.2.12 Hosting your Web site

Once you have chosen a CMS you need somewhere to put it. Making your Web site available to the public requires that it be placed ("hosted") on a Web server with good access to the Internet. There are two main ways to do this:

- ***Rent disk space on a server maintained by a commercial Web-hosting company.*** This is by far the best option if your resources are limited. At time of writing (mid-2006), small hosting packages with reliable companies start from just **$1 per month** (really!). A major advantage is that the company will look after server administration and equipment upgrades, reducing your need for IT staff and time wasted solving technical problems. Minor disadvantages are that hosts impose limitations on disk space and the amount of data transfer (but you can buy more, at a modest rate) and they may not offer all the software tools you could want. Finding a reliable hosting company with good customer service can be difficult, particularly at the "cheap" end of the market.
- ***Buy (or rent) a server and maintain it yourself.*** This normally requires that you also physically place the server in the data centre of an Internet service provider, which will provide high-quality facilities to connect it to the Internet. The major advantage is that you can run any software you need and have practically unlimited disk space. However, it is very expensive and **requires** IT staff with experience in server administration. This option is best for very large or complex sites that require a customized software environment and will not be considered further here.

Do you "really" need to buy your own equipment?

We have noticed that government agencies and institutions tend to buy their own servers – even when they do not have IT personnel to maintain them! However, the idea that you need to buy expensive hardware to run a Web site and that it should be physically located in your own office is a ***myth****. It is actually far more difficult and expensive. We suggest that wherever possible you rent disk space from a reputable commercial hosting company with quality facilities specifically designed for this purpose. You can use the money you save to train your staff in Web publishing.*

Dealing with the "Internet dark"

If Internet connections or electricity supplies are unreliable in your area then consider renting disk space from a hosting company overseas. Even when your office is "offline", your Web site can still be presenting a high-quality presence to the rest of the world.

Choosing a host for your Web site

We cannot stress how important it is to have a good host for your Web site. Unfortunately, Web-hosting is not a very profitable business. Many companies are unreliable or provide poor customer support, or may disappear overnight taking your Web site with them (so make sure you keep your own up-to-date backup!).

Issues to consider in selecting a host include:

- ***Functionality.*** Does the host server meet all of the system requirements that you need to run your Web site? Do you require a particular operating system, server side language or database? How much disk space do you need/get? Is there a limit on the amount of data transfer and is it acceptable?
- ***Connectivity.*** What is the bandwidth of the server's connection to the Internet? Is the available bandwidth used lightly or heavily? The larger companies will guarantee that overall usage/sharing of the advertised bandwidth will not exceed a certain limit, as they will undertake to increase their capacity in response to growing traffic loads.
- ***Redundancy.*** Where are the company's servers located? Are they in a data center with backup links to the Internet, backup power supplies and good physical security, or is it in a shed in someone's backyard?
- ***Service.*** Does the company have a responsive helpdesk that you can contact in the event of problems? It is worth testing the helpdesk by sending them a couple of questions, for example you could ask them about their connectivity and redundancy arrangements. If you get a slow response, or no response at all, find another company.

Again, it is best to choose a reputable company, rather than the cheapest you can find, as many of the bottom range hosts are "hobby businesses", often without dedicated customer support staff. At the budget end of hosting providers, NACA has had good experiences with a company based in the United States of America <http://www.3ix.org>. At the time of writing (mid 2006) their hosting plans started from US$1 per month and came with excellent help desk service. However, you should conduct your own research into the current reputation of any hosting company before you sign up. There are many Web hosting discussion forums on the Internet where you can consult others about their experiences with particular companies. If you encounter difficulties with a hosting company it is often best just to find a new one. The

cost of switching to another host is often much less than the cost of trying to deal with ongoing problems.

2.2.13 Optimising your site for search engines

There's more to publishing a Web site than just turning it on. You need to make sure that the world knows it exists! The key to success is to get your site indexed with a high ranking in the major search engines, especially by Google, which at time of writing was the dominant search engine, at least in the English-speaking world (other search engines may be dominant for other languages or particular countries). Getting your site onto the first page of Google's search results for your topic will make it much more "visible" to potential visitors... but competition for the top ranking is fierce.

Every search engine is different, and exactly how they work is a closely guarded secret (Google claims to evaluate more than 100 different factors), but most offer some tips on their Web sites that you can use to optimize the ranking of your site. The following discussion is specific to Google, which as the dominant player should be your primary concern.

How Google works
Mechanisms that Google uses to rank the relevance of a page with regards to a particular search query include:

- ***Content***. Google analyzes the content of Web pages to determine the search queries for which they are most relevant. The "density" of search terms on a page is an important indicator. Search terms found in the title and heading of a page and in links are weighted more heavily than those found in the body.
- ***Structure***. The position of an item on a Web page can give an indication of how relevant it is to the search engine. Items found at the top of a page (where headings usually go) and in the top left column of a page (where the most important navigational links are usually located) are weighted more heavily.
- ***Popularity***. Google considers how many other sites link to the site in question. sites with many referrers are presumed to be more relevant. However, Google also considers the relevance of the *referring* sites. Incoming links from highly ranked Web sites are worth more.

You can use this information to make your site "Google friendly". Consider the main subject(s) of each page on your Web site:

- Give each page a title relevant to its content (title is a field found in the page header).
- Give each page a heading relevant to its content.
- Give links meaningful names that are relevant to their subject (do not give links names like "click here").
- Put navigational links relevant to the page content in the upper, left column areas.
- Try to get *highly popular* and relevant Web sites to publish a link to yours, preferably on their home page.

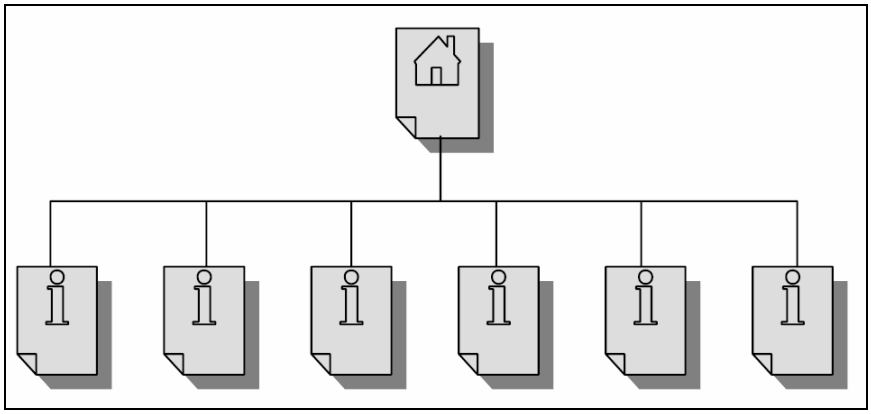
Note that at time of writing Google *ignores* the keyword metadata field on Web sites – presumably because this is commonly abused by Web masters trying to boost their ranking for particular search terms. Google also tends to ignore referring links from "link farm" sites. You can find more information on optimizing your site from Google's Web site, at: <http://www.google.com/support/Webmasters>.

None of this should distract you from the fact that the best way to get a high ranking in Google is simply to **provide high quality information**. If you do that, people will naturally be drawn and link to your site. Beware of companies that offer "search engine optimization" services that guarantee a high ranking: placements cannot be bought.

Keep your Web site structure "flat"

One of the problems with database-driven sites is that search engines do not always index them fully. Database driven sites can be extremely complex, so search engines tend to crawl the top few "layers" of links, but they will not follow chains of links indefinitely down into the depths of your database. This is another good reason to keep your site "flat" with all resources available within 2-3 clicks from the home page; sites built with a "hub" architecture tend to be indexed better than those with a deep, hierarchical structure.

Flat Web sites are easier to navigate and index



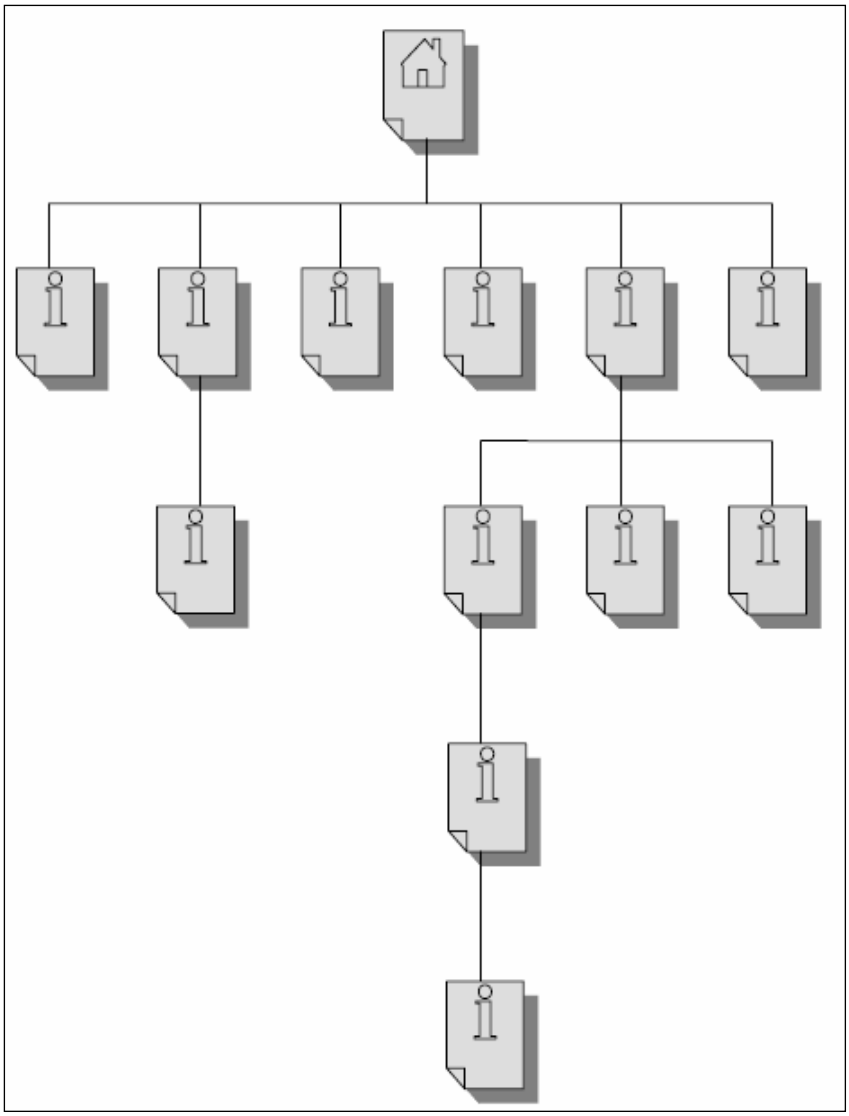
Submitting your site to search engines

A good way to draw attention to a new site is to submit it to the major search engines for indexing. Instructions on how to do this are usually available from a link on the search engine's home page (look for "Webmaster resources" or equivalent). The submission process usually just consists of entering the URL for your site in an online form, and the search engine will crawl it sometime thereafter. The Google submission form can be found at: <http://www.google.com/addurl/?continue=/addurl>.

Advertise widely

Do not forget to advertise your site through conventional means – through newsletters, magazines, stakeholder meetings etc. Make sure that as many stakeholders are aware of your site as possible. If you are offering a useful service, word will spread quickly.

Information on the bottom layers of a strongly hierarchical Web site is hard for users to find and may be less visible to search engines



2.2.14 Security, backups and disaster recovery

Like most things we have discussed, security ultimately depends on people. Unfortunately, the reality is that most people do not take security seriously, or see it as an inconvenience to be circumvented. It is also a fact that many malicious attacks against computer systems – probably most – are "inside jobs" by an organization's own employees.

Security issues can be broken down into:

- ***Physical security threats.*** The computer your system exists on can be stolen, or a fire or some other disaster can destroy it.
- ***Information security threats***, including both loss and theft of information. These can happen in many ways: a hard disk may crash; a malicious "hacker" may break into the system; or disgruntled employees may abuse otherwise legitimate access rights.

An in-depth review of security is beyond these guidelines, but a simple approach is to:

- Restrict access to both the computer hardware that hosts your Web site and the internal administration of the Web site as much as is practical. Provide access on a "needs basis" only, and be strict about it!
- Ensure that you have a good backup procedure and disaster recovery plan. That way the complete destruction of your system will just be an inconvenience, rather than a disaster.
- Assign someone independent to review the plan several times a year to ensure that it is adequate and that it is actually being carried out.

Controlling access rights

Collaborative digital publishing tools usually provide ways to restrict "access rights" to the system, most often by allocating people to different "user groups" with different access privileges. To use a "group rights" effectively:

- Identify the system's main groups of users, in terms of their role.
- Consider the tasks each group will perform. What areas of the site do they need access to?

Commonly used groups and typical roles are outlined in Table 3. The level of access required for each of these groups is lowest for guests and progressively increases to Webmasters, who usually have access to everything. You may well need to create additional groups for specific functions on your own site. Editorial and Webmaster groups are positions of trust. Do *not* give people these rights lightly.

Development of a security policy that documents roles, access rights and other issues can help to raise staff awareness of security issues and give the system administrator the authority needed to enforce it. If the document allocates responsibilities and accountabilities it can help create some semblance of compliance among internal users, but do not depend on it. Fortunately most systems provide built-in mechanisms for enforcing security policies.

If your system permits, it is a good policy to set a minimum length of passwords to make them more difficult to "crack". It is also a good, though unpopular, policy to require that people rotate passwords periodically. This prevents people from using the same password for everything, and ensures that if a password is discovered, limits its usage. You should also ensure that employee's access to the system is promptly suspended when they leave the organization.

Table 3

User group	Who they are / role	Typical privileges			
		View site	**Submit content**	**Approve content**	**Admin rights**
Anonymous users	Visiting members of the public (unregistered).	Yes (some or all)	No	No	No
Registered users	Members of the public (or staff) that have "registered" an account on your site. This provides a mechanism for identifying their contributions and limited ability to apply disciplinary measures, if needed. They are often allowed to contribute information in community sites.	Yes (all)	Yes	No	No
Editors	Members of staff, or people closely affiliated with your organization. Editors review submissions by members (or staff).	Yes	Yes	Yes	No
Webmasters	Trusted members of staff, usually with IT background. They administer the structure and configuration of the site. They are often called on to provide technical assistance to members or editors.	Yes	Yes	Yes	Yes

Technical issues

The technical aspects of computer security are extremely complex, but from a practical standpoint most problems originate from either:

- Improperly configured or unpatched servers (operating systems and Webserver software); or
- Security "holes" in the CMS software itself.

Securing a server effectively requires the attention of IT staff specialized in server administration. If you do not have access to such staff then it is advisable to host your Web site remotely on the server of a reputable commercial hosting company, who will take on this responsibility for you.

Security holes in your CMS software are something you *can and should* do something about. No matter what software you use, from time to time new security problems will be discovered. If the software is well maintained, the developers will usually release a new version or a "patch" that you can apply to your system to fix the security hole. It is important that you stay aware of security fixes and keep your software up to date. Announcements are typically made on a software project's Web site and in e-mail support lists.

One of the risks of patches and upgrades is that they often introduce new bugs (or new security holes). A pragmatic approach is to wait a couple of weeks after a patch is released

to see if any major problems are reported with it. If no serious issues are announced, go ahead and patch your Web site. **Always** backup your site before attempting to upgrade the software because you may not be able to uninstall it!

We have barely touched on security here, and encourage you to spend some time reading about it. It is quite an interesting area, and even if you do not have a technical background you will still find the general principles quite useful.

Backup your system, or you will lose it

Despite your best intentions we can **absolutely guarantee** that sooner or later, something will go horribly wrong. When it does go wrong, you had better have a disaster recovery plan in place. Imagine trying to explain that your organization's collective output for the past five years has…vanished. Your job may well vanish, too.

If you have good backup procedures then you need not fear disaster. Backing up is quite easy, the main points are to ensure that you:

- ***Backup everything*** – it is good practice to maintain a list of important information resources that must be included in your backup procedure.
- ***Do it regularly***.
- ***Know how to restore*** the system from a backup.
- ***Keep a copy of your backup "off site"*** in a physically separate location from the machine that has your information on it. People like to steal computers, and fires do happen.

Keep your old backups in a collection as a "time series". When you discover a problem you may need to be able to restore the system from some point in the past when you *know* it was working properly - or perhaps to retrieve a single record that was accidentally deleted. It is good practice to also **test your backups** occasionally to make sure that they (a) actually work and (b) are complete.

The procedures you need to follow to back up your system will depend on its nature and on how much information you have.

- If you have a simple HTML Web site, you may only need to backup the directories containing all the files.
- If you have a dynamic database driven Web site, you need to backup both the directories containing the files and the contents of the database itself.
- Be aware that if you have a decentralized system with your information assets residing in several locations, you will need to back up all of them! For this reason you need to know exactly where all of your information assets are stored. Do not assume they are all sitting on the same server, on the Web resources can be anywhere.

For a small Web site or digital library collection, backing up to CD or DVD may be the cheapest and most convenient solution. For larger systems (generally beyond the scope of these guidelines), you may need a tape backup or dedicated storage infrastructure. Commercial Web site hosting companies often have some form of limited backup service, but there is no substitute for making your own. Web hosting companies frequently go out of business!

Planning for disaster

A disaster recovery plan is a document that describes the organization's policies and procedures for backup and recovery of the system. The plan should document:

- All important information assets that need to be backed up.
- How often they should be backed up.

- Provide instructions on how to access them. You could include logins and passwords (in which case this document should be kept in a very safe place) or identify people that have the required access.
- If your assets are hosted on a remote server, provide contact details for the owner of the machine(s) where the assets physically reside. If you become unable to access the information via the Internet for whatever reason (such as a problem with your hosting service), you need to be able to contact them.
- Provide step by step instructions on how to make a copy.
- Document the procedure for storing backups – where they are and who has access.
- Provide step-by-step procedures for restoring the system from a backup.
- Identify the officer responsible for carrying out backup.
- A procedure and time frame for reviewing the disaster recovery plan. As with all contingency plans, they should be reviewed periodically to ensure that they still meet current needs.

If your organization handles confidential information, you should keep your backups appropriately secured. It may also be a good policy for backups to be entrusted to someone that has nothing to do with the computer system: one of the worst forms of disaster is the disgruntled IT officer who has access to everything.

It is useful to lay a contingency plan for relocating your Web site. This will enable you to restore the system quickly in the event that your hosting company goes bankrupt, or your computer room burns down.

2.3 NACA Web site and publications

Until 2000 NACA conducted networking activities exclusively through traditional methods such as workshops, training courses, study tours, conferences and printed publications. As a highly decentralized organization with a relatively small budget the cost of international travel, communications and postage is the main constraint to networking and sharing of information.

In 2000 NACA established its first static Web site. Public utilization of the site increased substantially with the addition of NACA's first online information services – an e-mail newsletter and later free access to full-text publications. In March 2004 the Web site was rebuilt using a content management system, with a view to establishing an “online community” of network participants. The site is now interactive: visitors can post their own news stories, events and publications on the Web site and share them with others. A discussion forum has also been added, where visitors can ask questions and interact with each other directly online. This has enormous potential as a networking tool, particularly in a highly decentralized organization such as NACA where individual participants are located in different countries. As of July 2006, more than 2 500 people had registered as members of the Web site, which has become an integral part of the network’s communications framework.

The NACA Web site has been constructed using the XOOPS CMS, a free, open source product available from <http://www.xoops.org>. XOOPS relies on other open source supporting technologies, has good documentation and free support via an online user community so expenditure has been minimal (it was built without a budget). The major cost related to staff time while we learned how to use XOOPS and about Web publishing in general. The NACA Web site is physically located on a US-based server maintained by a commercial Webhosting company, 3ix <http://www.3ix.org> at a cost of US$ 3 per month.

Since 2002, all of NACA's publications have been made available on the Web site in electronic form (in addition to paper and CD) for free download, with new titles actively publicized through a variety of channels. These are organized as a simple electronic library with more than 700 titles as of June 2006, ranging from news clipping services to periodicals such as Aquaculture Asia Magazine, project reports and monographs. Most publications are provided in PDF format to facilitate local reproduction and secondary distribution. Bandwidth is an issue in many member countries so PDF files are carefully optimized to minimize their file size to facilitate easy download.

All of NACA's electronic publications are made available for **free** (hard copies are distributed for nominal cost). Key NACA publications such as Aquaculture Asia Magazine are now released under the Creative Commons License, which allows for non-commercial reproduction and distribution by third parties.

The Web site houses a number of e-mail newsletters and a notification system that can be used actively to inform people of new publications and information resources both on the NACA Web site and elsewhere. More recently, discussion fora have also been added where members of the public can communicate and exchange information directly.

The simple steps of making NACA's information **available, accessible and free** have greatly increased awareness of and access to the results of the NACA's activities and to the sharing of network information. The Web site has become the core of NACA's information system, averaging around 35 000 visits and 179 000 page views per month in 2005, with approximately 142 000 publications distributed via download over the full year.

The NACA Web site <http://www.enaca.org>

Digital publications on the NACA Web site are arranged in simple categories

Individual publications can also be accessed by the site's built in search engine

2.4 Creating accessible digital publications

Successful digital publishing is all about making your information accessible to people. For a publication to be truly accessible to your stakeholders:

- it must be available in a format(s) that they can make use of;
- they must know it exists, or be able to find it when searching the Internet;
- it has to be physically obtainable, which in the context of the Internet generally means easily downloadable; and
- it should be available for free.

2.4.1 Increasing accessibility through a multimedia approach

A common and valid criticism of digital publishing is that most people do not have access to the Internet or even to a computer. The best ways to distribute your publications will depend on the capabilities and preferences of your key stakeholder groups. You may need to publish your work in a variety of formats to meet the requirements of different stakeholders, potentially making use of a range of electronic and traditional media (such as printed publications, radio and television). The relative importance of different media should be a consideration in your allocation of resources, and based on the profile of your stakeholders. The most important thing is that, whatever their capabilities, your stakeholders should be able to access your publications easily and in a form that they can use.

In choosing a suitable format, the basic considerations are whether the publication needs to be viewed online or offline, and whether it needs to be portable (i.e. easily copied and exchanged). Covering these options will help people to access and share your work, maximizing its distribution and impact.

Web sites are the standard way for presenting information for online dissemination. Information embedded within a Web page (i.e. in HTML), such as a news story composed of text and photographs, can be directly viewed through any standard browser and indexed by search engines, so it is easily accessible to the casual reader. The disadvantage of presenting text publications as Web pages is that they can be difficult for people to copy, exchange and print. This is not so much of an issue with one-page articles, but long HTML documents such as books and reference manuals are typically split into multiple pages for easy viewing. Each page may itself consist of many separate files, so handling of these documents is not very convenient.

You can address this issue by additionally offering your publications in a “portable” format that is easy for people to copy and distribute, for example as a Microsoft Word Document or PDF file that they can download from your Web site via a link.

If your stakeholders do not have good Internet access, you might also consider offering your publications on CD-ROM, which provide a relatively cheap and easy way to distribute large amounts of digital content. This can be as simple as putting your publication files on a CD, but if you have a basic knowledge of HTML you can easily create a simple Webpage interface for it. In fact, if you have a static Web site, it is a relatively simple to put your entire Web site onto CD for distribution to offline stakeholders, publications and all!

2.4.2 Increasing awareness of publications

To make publications available on the Web is not enough. You need to find ways to let people know they exist, or to help them find them when they search. Adding a new publication to your Web site is a bit like putting a new book on a shelf in a large library. People may stumble over it eventually while browsing the shelves, but they will find it a lot sooner if you advertise its presence and index it properly.

Publicity

There are many ways that you can advertise new publications. Two simple and effective measures are to (a) make conspicuous announcements on your Web site homepage (we find that including a picture of the cover increases the number of downloads by three to four times) and (b) provide an e-mail newsletter service, which contains direct links to files. Announcements in other media, such as your regular print newsletter and magazines, also create incentive for people to visit your site. If your software provides RSS or ATOM syndication feeds you may be able to automatically share your latest listings with other Web sites (refer to the section on “Exploring syndication and partnerships”).

Indexing publications: Metadata

When you index, say, a book for a library catalogue you create a record about it that describes important attributes such as the title, who wrote it, when it was published and where it can be found. This record is *metadata*: information that describes a data resource. In most cases when you search a library catalogue you are searching metadata (records about publications), not the publications themselves. Accurate metadata improves the precision of searches, making information easier to find. Metadata is also commonly used to help interpret and administer data.

When preparing metadata (i.e. indexing publications) you should always try to follow an existing standard, if one exists, rather than develop your own customized scheme. Using an accepted metadata standard will help keep your data **portable**, making it easy for you to exchange records with other systems and to migrate your records to another software platform, a key consideration in their long-term preservation.

What kinds of metadata should you collect? There are many different metadata “schemes” available for describing different kinds of information, but few have achieved the widespread acceptance and formalization required to be considered a “standard”. Metadata schemes for describing Internet resources are still at an early stage of development, although standards are beginning to emerge around the Dublin Core Metadata Initiative. However, libraries have been around for much longer and have developed fairly well accepted metadata schemes for cataloguing conventional publications.

2.4.3 Metadata for the Web

Web site metadata

There are several commonly used metadata fields (“meta tags”) that you can include in your Web pages to help people find them. Meta tags are embedded in the header (code) of each page, so they do not appear on the public face of the page itself. These are read by search engines, which may use them to help index your Web pages, and some tags are read by browsers, which use them to display pages correctly. Table 4 covers some of the more commonly used Web site metadata fields and their use.

Table 4

Meta tag	Application
Content-type	Indicates the nature of the resource e.g. “text/HTML”, and the encoding scheme for the character set e.g. “charset=iso-8859-1”. This information is used by browsers to display text correctly.
Title	The name of the page. Strictly speaking this is not defined as a meta tag, but functionally it is used as one.
Content-language	Specifies the language of the Web site. Each language has its own code, for example “en” for English.
Robots	Provides instructions to the indexing agents of search engines (sometimes called “robots” or “spiders”) as to what content they should index. You can request that certain pages be included or excluded, and whether links to other pages should be followed or not. However, not all will honor these tags – a robot’s behaviour is dictated by the policies of its master! For example, spammers often send robots to harvest e-mail addresses from Web sites to add to their databases.
Keywords	Keywords are used to indicate the subject matter of the Web site.
Description	A general description of what is contained in your Web page.
Author	The author meta tag defines the name of the author of the document being read. Supported data formats include the name, e-mail address of the Webmaster, company name or URL.
Copyright	The copyright meta tag defines any copyright statements you wish to disclose about your Web page documents.

The title is probably the most important element, followed by the “keywords” and “description” meta tags, as these convey information about the subject matter of a Webpage, and may affect the way your page is indexed by search engines. The Content-type tag is strongly advised (and automatically generated by most Web-authoring tools) for reasons of accessibility.

HTML meta tags are optional and users are free to define their own. As a consequence there is little real standardization. If you are hand-coding your own Web pages, it is good practice to include the Dublin Core metadata elements (see below) to describe each page. However, at this point in time very few Web sites do this and few open-source content management systems offer native support for Dublin Core metadata.

An additional complication is a trend for search engines to *ignore* subject-related meta tags, as they are commonly abused by Web masters trying to artificially “boost” the rankings of their Web site. Search engines do not trust metadata supplied by authors (with good reason) and are moving towards reliance on their own increasingly sophisticated automated analysis of Webpage content for indexing purposes. In our view such policies make the preparation of detailed metadata for *Web pages* somewhat redundant as far as the major search engines are concerned (refer to previous discussion on how Google works). However, metadata can be important for the indexing of resources within your site’s internal databases, or that you wish to share across a federated network of sites.

The Dublin Core: metadata for networked resources

The Dublin Core is a metadata standard consisting of 15 elements designed for description of networked resources (i.e. digital objects available on the Internet). It is designed as a generic, interoperable standard that can be used to describe any kind of resource. As a broad standard, the precise usage or encoding values of Dublin Core metadata elements is not mandated. However, the specification does make recommendations on best practice, generally advocating the use of other standards, controlled vocabularies and thesauri to represent dates, coverage, subject etc. The reference description is available from <http://dublincore.org/documents/dces/>.

The main attraction of Dublin Core, particularly for resource-limited organizations, is that it is relatively simple and therefore makes it relatively fast and easy to create records. This is not the case in more complex systems such as MARC 21, which is used by many library catalogues, although the simplicity of Dublin Core may be seen as a disadvantage for some applications. The Dublin Core is emerging as a generic standard for interoperable metadata, and is the basis of the Open Archives Initiative protocol for metadata harvesting. It is attracting support both from major libraries and the Internet community in general, and is supported natively in some advanced open source digital repository software such as Greenstone.

The standard is maintained by the Dublin Core Metadata Initiative. For more information including the list of terms and detailed explanation of their usage, visit <http://dublincore.org/>. Guidelines for Encoding Bibliographic Citation Information in Dublin Core Metadata are available from <http://dublincore.org/documents/dc-citation-guidelines/>.

2.4.4 Improving physical accessibility: a crash course in compression

The best way to improve the physical accessibility of digital publications on the Internet is to reduce their total file size, thereby making them easier to download. This is particularly important if your stakeholders are likely to have low-speed Internet access, for example through dial-up modems. There are two main ways you can do this:

- choose an efficient file format;
- judiciously apply various forms of compression.

Below we provide some general guidance on how to optimize the file size of your publications for use on the Web beginning with images, which are the source of most problems.

Optimizing images for download

Images take up a tremendous amount of space, and often account for most of the file size in a text document or Webpage. For this reason you should minimize the use of images wherever possible: only use them when you have a real need. Fortunately, with a bit of care and experimentation it is possible to drastically reduce image file size:

- JPEG (file extension: JPG) is the most common format for saving colour photographs and images with many colours. Most software will ask you what “quality” to save at. Lower quality values give smaller file size but the trade off is loss of clarity. How low to go is a matter of personal taste, but a value around 45-60% (in Adobe Photoshop) will usually give a reasonably clear picture with a substantially reduced file size. You should only compress a JPEG once: opening and resaving will compress it again, causing further degradation of image quality, so keep a copy of the original as a master in case you need to use it again somewhere else or in a printed publication (where much higher quality is required). Note that different software packages may give different results with the same settings, so experiment and examine the results critically until you find settings that you are satisfied with.

- GIF and 8-bit PNG support a maximum of 256 colours and are the most efficient formats for simple artwork that contain few colours such as logos. They are **not** suitable formats for photographs. If your software allows you to set preferences when saving GIF or PNG, you can experiment with reducing the number of colours in the images until you start to see noticeable degradation in quality. Reducing the number of colours lowers the file size.
- Converting colour images to grayscale will reduce their file size by around two thirds. This is a very effective way to reduce the size of publications containing many photographs.

The best available tool for optimizing images is Adobe Photoshop (www.adobe.com). This provides a "save for Web" function that allows you to preview the effect of different compression settings on the appearance and file size of images, side by side. Unfortunately it is very expensive and there are no open source alternatives of the same caliber, but the GNU Image Manipulation Program (GIMP) (http://www.gimp.org) provides reasonable image editing functionality, and is free. The Adobe Photoshop Bible (McClelland 2001) provides an excellent overview of image optimization and image editing in general for Photoshop users. Documentation for GIMP is available from the GIMP Web site.

Optimizing Web pages for download

Since Web pages are often made up of multiple files, you must optimize each page element individually. The most important things to do are:

- Minimize the use of images, and optimize them as per the instructions above.
- Resize images down to the actual size, in pixels, that you want them to appear on the screen. Do not use HTML functions to "scale" images to appear smaller than they really are. It is inefficient, since the full image must still be downloaded even if you do not display it at the full size.
- Minimize the number of page elements by combining images and consolidating multiple javascript (.js) and style sheet files (.css) into a single file where possible. Each object on a page must be separately requested from the Webserver, which takes time. Pages that have many objects on them, such as a lot of photographs or "sliced" images, can take considerably longer to load. Combining objects reduces the number of requests and associated latency.
- Some content management systems offer an option to compress Webpages (GZIP), which reduces download times for the end user but increases processing load on the server.
- Eliminating comments and white space from your HTML code can reduce file size, though it may also make maintenance of the site more difficult.

Optimizing text documents for download

Text documents come in many different formats but the dominant formats in commercial use today are Microsoft Word, Adobe Portable Document Format (PDF) and plain text. Word is the defacto standard where people need to collaborate in writing a document, but it is not necessarily the best format for distributing publications over the Internet. The relative merits of these document formats are summarized in Table 5.

Table 5.

Format	Advantages	Disadvantages	Application
Microsoft Word	• Defacto standard • Easy to modify and good for collaborative writing	• Proprietary format • No built in compression tools available; documents that include photographs may be very large • Documents may not be readable if the character set is not installed on the host computer • Only supports simple document layouts • Longevity of format is not known	• Collaborative writing and exchange
Open Office	• Easy to modify and good for collaborative writing • Based on XML • Supports internal compression (ZIP).	• Nowhere near as common as Microsoft Word.	• Collaborative writing and exchange
Adobe PDF	• Supports complex page layouts, including from commercially produced publications • Consistent reproduction of document across platforms • Embedding of fonts allows consistent display of documents in different languages / character sets • Internal compression tools available including for images • Widely supported • Reader software is free • Can be generated from most common office software	• Very difficult to modify – intended as a format for viewing, not for editing • Requires dedicated software to view files (but available for free) • Proprietary format (but specification has been published) • Longevity of format is not known although a variant of PDF is in development as an archival format, which may offer a very good mechanism for long-term preservation of digital publications.	• Distribution and exchange of documents that do not require further modification (view only) • Distribution and exchange of documents with complex / commercially produced page layouts or character sets
Plain text	• Non-proprietary format • Can be read and modified with the most basic software • Compact • Likely to be supported for a very long time to come	• Does not support images or complex formatting • Not favored for documents where presentation is important	• Favoured for text-only documents where accessibility is paramount. • Basis for many common document types (HTML, XML, software source code and “help” files.

In terms of minimizing file size, Word is usually more efficient than PDF for producing **small** documents (several pages in length) that do not contain photographs. However, PDF is usually more efficient for longer documents. PDF also offers internal compression of images which, when properly configured, make it massively superior to equivalent documents in Word. Added advantages of the PDF format are that it can be generated from professionally laid out publications, and that it supports embedding of fonts, enabling documents to be consistently viewed on any computer regardless of character set. PDF is our preferred format for distribution purposes.

The following tips can help you to optimize the file size (and accessibility) of PDF files:

- Some PDF creation tools offer quality "presets" or "job options" that are optimized for different purposes. If you do not have much technical knowledge, the easiest way is to make use of an existing preset designed for Web use. The most visible difference to the end user is in the quality of the images, so experiment with the various presets until you find one that produces an acceptable trade off between file size and image quality. We prefer the "standard" preset in Adobe Acrobat Distiller 7 (called "eBook" in previous versions). The lowest quality presets tend to produce unusable images, so something a bit higher generally gives a good compromise.
- If creating your own custom quality settings, the biggest saving in file size comes from compressing images, for which resolution is the key issue. We find a resolution of around 150dpi (dots per inch) is a good trade off between file size and on-screen image quality. However, you should be aware that images printed at this resolution will not be entirely clear on paper. Print-quality images need to be of at least 267dpi resolution, but this produces very large files that are difficult to download. If you need print quality images it may be worth creating separate Web- and print-optimized versions of the file.
- If offered a choice in the compression scheme for images, choose JPG over ZIP as it is more efficient and use a medium quality setting.
- Consider producing a grayscale PDF file. If your publication contains lots of colour images, this will cut file size substantially.
- Do not embed thumbnails in the PDF file. Each thumbnail may take up several KB of space, which can add substantially to file size in long documents. Thumbnails will be automatically generated when the document is opened anyway, so there is no real need for this option; it just makes the file open slightly faster.
- Embed all fonts. This actually *increases* file size but it assures accessibility. People who do not have the right fonts installed on their computer will still be able to view the document. This is particularly important for organizations working with multiple languages and character sets.
- Use as few fonts as possible in the document. Embedding fonts takes up space, so using fewer fonts will reduce the size of the document.

PDF documents can be created directly from most standard word processing and page layout applications including Microsoft Word, Open Office, Pagemaker, Quark Express and InDesign. The level of control varies from program to program. If your software does not permit manipulation of PDF output internally, there are a number of free programs available that allow you to create PDF files by printing the output of any program to a "virtual printer", including:

- PrimoPDF: <http://www.primopdf.com/>
- PDF995: <http://www.pdf995.com/>
- PDFCreator: <http://sourceforge.net/projects/pdfcreator/>

Adobe distribute free PDF viewer software i.e. Adobe Acrobat Reader, which can be downloaded from <http://www.adobe.com>. They also sell advanced PDF creation software, Adobe Acrobat, which is excellent although expensive.

It is also possible to compress plain text and Microsoft Word files using third-party compression utilities or "zip" tools, which may be particularly useful for distributing documents that are long or which contain images. One advantage of these tools is that they can typically bind multiple files into a single compressed archive, facilitating exchange of collections of documents. This is a very useful way to distribute multi-page HTML documents. A disadvantage is that the recipient will need their own compression software to extract the contents of the archive. Many open source compression tools are available, including:

- 7-Zip: <http://www.7-zip.org/>
- QuickZip: <http://www.quickzip.org/>
- ExtractNow: <http://www.extractnow.com/>

2.4.5 Adopting an open access policy

Print publications are expensive to produce and to post. It is quite common for organizations to attempt to recover their costs by charging for their publications. Unfortunately, this thinking is often carried over into the online world, even though distribution costs are much lower. Sometimes fees are charged on electronic publications to help pay for the cost of the printed version!

If your organization is serious about making its publications accessible then consider making them available for **free** to avoid creating financial barriers to access. The fact is that very few people are willing to pay for information services over the Internet. Many people, particularly in developing countries, either do not have the capacity to pay, or do not have a mechanism to pay, especially where banks, credit cards and international transactions are involved. In practical terms, charging even US$1 for your publications will massively reduce their circulation and also their impact. Why produce information if you are going to lock it up?

There are significant intangible benefits to providing free access to your publications. They will achieve a much greater circulation. More people will receive your information and more people will use it, contributing more effectively towards your organization's goals. More people will also become aware of your organization's activities, enhancing its reputation and standing with your stakeholders and peers. This in turn can make it easier for your organization to obtain funding for projects and to establish partnerships with other organizations. If you need to recover funding, then you could consider advertising and sponsorship arrangements as alternative models.

In addition to making your publications available for free, you may wish to consider releasing them under a copyright license that permits the public to copy and distribute them amongst their own networks. This is bound to happen to some extent anyway, but formalizing permission to do so will help recruit government agencies and other "official" bodies in the secondary circulation of your publications, ensuring their widest possible dissemination. Creative Commons (CC) licenses, for example, offer a flexible scheme that retains your copyright but allows people to copy and distribute your work, so long as they give you credit. Several variations of CC licenses are available, allowing you to specify whether your work can be modified or used for commercial purposes or not. For more information about Creative Commons licenses (including logos and code you can use to mark Web sites) visit <http://creativecommons.org>.

2.5 Successful digital publishing

Having designed and built a technically excellent and needs-driven digital publishing system, success will now depend on how effectively you *use* it. The most important thing is to concentrate on providing high quality information. However, you can increase the effectiveness of the system by taking steps to ensure that people *know* when you publish something, and by adopting policies that promote open access to your work. Every aspect of publishing needs to be considered primarily from the point of view of the end user (Gutman, Ayers and Booth 2003).

2.5.1 Provide high quality targeted content

What kinds of information does your target user community *want?* Organizations quite often fall into the trap of providing information that reflects their *own* interests, much of which may not be useful to their target user group. Too much irrelevant content will reduce the value of your site. It is better to adopt a demand driven focus and provide a smaller amount of information that is well targeted to meet the needs of your stakeholders.

Consider the profile of your stakeholders and their primary interests. Take stock of the information your organization produces:

- Which information is likely to be useful to your stakeholders? How will it help *them?* These are your most valuable information assets and the highest priority for publication. Publishing this kind of material will encourage return visits to your Web site and enhance the reputation of your organization.
- What information is *not* useful to your stakeholders? Most organizations tend to produce a lot of information that, if we may be honest, is only read internally or by their peers (or if we may be even more honest, by nobody at all). You will inevitably be asked to publish some of this, but be aware of its real value to your stakeholders. Do not let low-value information dominate your publishing activities. Make your colleagues aware of the real interest of readers, for example from the Web site log file analysis. An understanding of which information is of most interest should help when deciding what to publish next.
- Are there any gaps where your organization is not meeting the information needs of your stakeholders? Is there an opportunity to provide them with additional high-value services?

Once your site is publicly available it is fairly easy to get good, quantifiable information from Web site log files about the main interests of users by examining the kinds of information they prefer to access (refer to the section on “performance monitoring and improvement” for further information).

Publish frequently

On the Web it is much better to publish small amounts frequently than a lot of material occasionally. Publishing new material on a regular basis creates an expectation that the Web site will be updated and provides incentive for people to return. Circulation of your publications will be greatly increased if you can make your site useful enough that it becomes part of people's daily routine.

Actively distribute publications
You can dramatically increase the distribution of your publications simply by letting people know that they are available. When you release a new publication on your Web site it is a good idea to look for ways to advertise the fact. E-mail newsletters are a simple and highly effective way to inform users about new resources (most CMS will allow you to e-mail registered members of your Web site). If your stakeholders are not major computer users, it may also be beneficial to submit announcements via other channels - perhaps in newspapers or magazines that they are likely to read.

Exploit cross-media linkages
You can often leverage different media off of one another to raise awareness of new resources. For example, if you produce a printed magazine it could carry a column announcing new downloads and other developments on your organization's Web site to stimulate interest. Your Web site can be used to promote your printed publications. This mutual value-adding effect of linking different media can make an important contribution to your communications strategy.

2.5.2 Explore syndication and partnerships

Every organization wants to have its own Web site, but it is a fact that few are willing to pay for the staff and computer facilities needed to do it well. Most make a token investment and try to get by with their existing resources. The Internet seethes with small, fragmented and poorly maintained Web sites.

If you do not have access to sufficient resources to run your Web site, you are not alone. It makes a lot of sense to approach organizations with common interests about pooling your resources to build a shared Web site. Between you, you may be able to find enough cash and people to develop something that is a lot better than you could afford alone. Not only will you have better infrastructure, but by publishing together you will create a larger and more frequently updated Web site, which will be more useful to your stakeholders and more frequently visited. An additional benefit of sharing a Web site is that it is a very practical way of improving ties and sharing of information between organizations. The people in participating organizations will be more likely to read each others publications if they are working off a common Web site.

If you do have your own Web site it is still worth considering sharing or publicizing your content with other organizations. A simple way to do this is to discuss exchanging announcements with other Web sites. A good option for sharing content between Web sites is RSS Feeds, if your Web site supports this. RSS stands for “Really Simple Syndication”. An RSS Feed is just a file on your Web site that contains a summary of the latest information that has been added to it. Other Web sites can link to this file to display your latest information on their own site. As the feed is updated (usually automatically) all sites linking to it will also be updated, so no additional maintenance is necessary. You can also use RSS in reverse to automatically bring in the headlines of other sites that are willing to share. This is a great way for partner Web sites to exchange information.

2.5.3 Establishing an online community

We have already discussed how Web publishing tools can allow members of your organization to collaborate online. Community Web sites take this concept one step further by allowing members of the *public* to participate as well. Instead of merely presenting information to people, community Web sites allow their members to communicate and exchange information among themselves (Boiko 2005). The most common form of

community Web site is a “discussion forum”, but the community concept can be applied to nearly any form of Web site.

Online communities are a unique tool in that they allow an individual to access the collective knowledge of a large group of people that may be scattered all over the world. They provide a “venue” where people with similar interests can “meet” each other, share experiences and solve common problems. One of the most powerful applications of online communities are “self-help” groups. If you have limited resources, empowering your stakeholders to help each other through a community Web site may be more practical than trying to help them yourself.

Establishing an online community presents some challenges. How can you (a) draw enough people to your Web site, and keep them coming back, to create a “community” and (b) how do you manage submissions, potentially from thousands of people - and prevent abuse of the system?

Bringing people together

The key to establishing a viable online community lies in attracting a “critical mass” of active members posting questions and information and responding to the posts of others. Once activity reaches a certain level the feedback and mutual interaction among members will become largely self-sustaining. In a community site the most valuable content is usually that contributed by its members. They become, in a sense, the primary authors.

Achieving the critical mass of members needed to initiate an ongoing “conversation” can be difficult but as general principles, the following apply.

- The most important aspect of establishing an online community is to identify an area of common interest to your target stakeholders, something that will bind them together as a social group. Ideally, it should be a subject that they would have difficulty getting information on elsewhere (Boiko 2005).
- To attract a lot of members, the site must provide high quality information that your target stakeholders will *value*.
- To encourage people to return, the site must be regularly updated with new information. You can encourage participation by offering competitions (for example, best book review), opinion polls and other opportunities for people to interact.
- E-mail newsletters carrying links to new posts can be a very useful tool for encouraging debates. These can be used to alert members of new contributions and help to bring a lot of people into the site at around the same time, increasing the chance of interaction.
- Advertising the site widely through mail outs, partner Web sites, magazines and other media can help bring members in.

During the early stages of community formation, while membership is low, more questions will be asked than are answered. People rapidly lose interest in community sites if they cannot get feedback, so it is useful to arrange for staff to provide support until the community matures.

Managing community input

While the rewards of a successful online community can be very high, allowing members of the public to contribute to your Web site creates some risk. Quality control and copyright can be issues, and sooner or later somebody will abuse the system by posting an advertisement or unwelcome message. Fortunately, most community-oriented content management systems provide editorial tools that you can use to manage member submissions. Most allow you to control:

- Which areas of the site visitors are allowed to see.
- Which areas visitors of the site visitors can contribute to.

- Whether visitor submissions are published immediately, or withheld pending review and approval by an editor.

The basis for this is usually some sort of password protected member registration system that allows people to be assigned to "user groups" with varying levels of privileges, in terms of what they are allowed to do (refer to the section on security for more detail on user groups). In setting editorial policies for your community Web site you should carefully consider the sensitivity of the site and the potential for abuse. In a highly sensitive site, say for a political party, allowing the public to post comments without moderation would be disastrous. However, if your Web site is about growing flowers you might be less concerned.

Different levels of "openness" may be appropriate in different areas of your site. Imposing an editorial process in a discussion forum will impede the flow of conversations (and create a lot of work) so general practice is to allow people to post messages freely once they have registered as member. However, in a photo gallery delays are less important and the potential for abuse is high, so you might want editors to review submissions before they are published.

Moderators and community "culture"

Moderators provide another useful mechanism for managing community input, particularly in discussion forums. The role of moderators includes:

(a) Leadership: Moderators set the "tone" or "spirit" of the forum community through their behaviour: visitors follow their example. The importance of moderators in establishing a friendly and cooperative community cannot be understated.
(b) Managing the forum: Encouraging discussion, removing irrelevant topics, referring participants to sources of expertise.
(c) Keeping the peace: Stepping in to stop arguments, reminding people of the rules and (where necessary) disciplining troublemakers.

Online communities develop their own culture, just like any other social group. The "spirit" of the group is something that evolves over time, but it is strongly influenced by the behavior of the moderators and other site officials, who serve as role models. To establish a friendly, helpful and constructive community it is important that the moderators demonstrate these qualities, and encourage them in others.

Dealing with troublemakers

If someone seems troublesome, this is often because they are not aware of accepted forms of online behavior – sometimes called "netiquette". In some cases it is just a language problem: non-native speakers cannot be expected to be masters of diplomatic language! Most "difficult" people will respond well to some politely worded "attitude counseling" and can go on to become constructive and useful members of the community. However, there are a number of ways that you can deal with inappropriate behavior:

(a) Do nothing. Let the community discipline them. Peer pressure is a powerful social force. Often other members will shame offenders into good behavior. Many people respond well to this approach, if it is done politely.
(b) Issue a polite but firm reminder about the "rules" of the forum. This is a good way to deal with minor or unintentional offences, or to stop arguments.
(c) Lock the topic. Most discussion forums enable you to close a particular discussion topic so that no further posts can be made.
(d) Issue a warning to the offender(s) that their behavior is unacceptable and that they will be banned if they continue, for serious offences. You can do this privately an e-mail, or publicly.

(e) Ban serious or repeat offenders. This is best used as a last resort; people that are excluded from an online community may try to cause you problems later on.

2.5.4 Performance monitoring and improvement

Your Web site is up and running. It is loaded with huge amounts of useful information. Everything seems to be going smoothly… but are people *using* it? What parts of the site do they use most? What kinds of information do they prefer to access? What do they *really* want?

Most Web servers keep a detailed record of visitor activity in a "log file". This provides a lot of information about what pages people looked at, what files they downloaded, how long they stayed and many other things. Most commercial Web hosting companies also provide free software tools that automatically generate graphs and statistical analyses from your Weblog file. The kinds of information recorded vary but some of the most commonly used statistics are shown in Table 6.

Table 6.

Web log statistics	**Meaning / application**
Number of page views	More page views means greater usage of the site. The average number of page views per visit is also a useful indicator.
Number of visits	More visits means greater usage of your site.
Number of returning visitors	A high percentage of returning visitors suggests that people find your site useful.
Number of unique visitors	An indicator of how many people actually visited the site. It is usually overestimated as it is based on Internet Protocol (IP) numbers – anyone with a dynamic IP number (e.g. dial up modem users) will register as a unique visitor every time the enter the site.
Median visit length	People tend to spend longer on sites or pages that they find useful. Visit length is strongly influenced by type of information and the way it is presented, so you can only really make direct comparisons between similar resources.
Top pages	Shows which pages are the most popular on your site. A good indication of visitor preferences.
Most downloaded files	Shows which downloadable files are the most popular on your site. Also a good indicator of visitor preferences.
Top referring sites	A useful indicator of which sites are linking to yours, and their relative importance. If you choose to syndicate content with other sites, this can help you decide which sites may be best to approach.
Top search phrases / keywords	Identifies the keywords that lead the most visitors to the site, and their relative importance. Browsing this list can give you a feel for the main interests of visitors to your site – but bear in mind that it this reflects, to an extent, the nature of your *existing* content.
Number of hits	A popular but generally misleading measure of Web site activity. One "hit" simply means one action by the Web server. If a visitor opens one page that has ten images on it, this will count as eleven hits! Therefore "hits" are not very useful for general monitoring purposes and you should generally avoid using this statistic. Other Webmasters will not be impressed with how many hits you have!

Targeting content more effectively
The real value of Web log statistics lies in monitoring trends over time. Careful examination of log files can provide a lot of useful information about visitors and their interests. Observing the kinds of information that people prefer to access can help you to validate assumptions about what subjects they are interested in and what kinds of information they most value, and also to detect interests that you did not anticipate. You can use this to identify and further develop the "high value" aspects of your Web site, increasing its utilization. For example, on the NACA Web site we found a high level of visitor interest in job advertisements, a service that was not of much interest to the organization's core objectives. However, we decided to continue to provide this service as it was attracting visitors to the site who were exposed to our core information services while they were there.

Improving structure and presentation
Log file analysis can also help you to improve the design of your Web site by providing a way to "see" it through the eyes of your visitors. Through observation of visitor reaction to changes in the way information is presented you can gauge the effectiveness of changes – what is successful, and what is not. As an example, we found that we could increase the number of downloads of a publication by about four times simply by including a small picture of the cover in the description. Over time you may be able to make many changes that substantially improve the usability of your site. It is useful to maintain a record of changes that you make to your site, and the date that they occur, to help you correlate these with changes observed in log file statistics.

While Web log analysis can be a highly effective tool, it does have some shortcomings. It is often not possible to know, except in very general terms, where your visitors come from. The statistics will represent the behavior of a mixture of both your intended audience and other visitors to the site, so you should be careful about the conclusions that you draw.

3. Collaboration and the next steps

3.1 Where NACA is today

NACA currently has 700+ publications on its Web site, incorporated in the XOOPS CMS. These are available for browsing via a Publications tab and they are full text searchable via the inbuilt search engine. The experience of NACA and user feedback indicates that this is a satisfactory means of organizing and disseminating this volume of publications. The whole Web site, including publications, is searchable full text and this integration of different types of information is what most users have said they prefer. If they wish to limit the results, an advanced search allows them to search by specified information categories, such as publications, news items, background documents, etc.

A more advanced option for NACA would be the implementation of a Digital Library or Digital Repository in order to store its publications in a more structured system or archive. This would allow the production of standard metadata for each publication, which can then be retrieved using a more structured search strategy than is possible with free text e.g. by author, title, subject. Appendix 2 gives a brief overview of some open source digital library and digital repository software that NACA investigated with this option in mind. Note that the list of software solutions is continuously changing and it is not our intention to review any system. The features that NACA looked at are provided for the benefit of institutions which may be looking for a similar solution. The most important thing to consider is: first decide what you need and then carefully check the features and functionality that the various options offer. It is also a good idea to talk to experienced users of a system before making a final decision.

NACA has decided that their needs are met for the present and foreseeable future by a CMS and dissemination of publications via the eNACA Web site. Foreseeable in the context of rapidly developing technology probably need not be longer than five or six years. Many of the NACA partners produce fewer publications, have more limited resources and less IT support and will probably also continue with Web site dissemination of information.

Increasingly, digital libraries are becoming integrated as components within larger Web site frameworks that offer all kinds of information and communication services. These include tools to create and publish new documents and to organize and manage the collection. If a digital library or repository solution is considered by NACA and partner institutions in the future, it would make sense to take advantage of developments in the management of decentralized digital collections. Collaboration with regional or international developments in digital publishing and dissemination offers better opportunities and less of a strain on resources for smaller institutions.

3.2 Integration and long-term security of publications

Having decided that an in-house digital repository is too resource-demanding for NACA's needs, there are still several issues to be resolved. Maintaining publications on an in-house server and disseminating only via the Web site has several disadvantages, unless additional action is taken. Some obvious disadvantages are related to computer security and long-term availability and preservation. But NACA must also ensure that its publications reach as wide an audience as possible as quickly as possible if they are to contribute to the "*cooperation on planning for aquaculture development at the global level*" (Code of Conduct Article 9.2.4). This is best achieved by using a variety of dissemination channels in addition to the Web site. Google is very good but it is not good for specialized subject searching and there are many alternatives for the retrieval of subject based information.

NACA publications should be integrated into the mainstream of fisheries and aquaculture literature if they are to reach a worldwide audience – i.e. they can be retrieved by specialized systems designed to search aquatic sciences publications worldwide. There are several bibliographic databases that cover the global literature, usually agricultural or aquatic sciences-specific, which also cover aquaculture. Making sure that the publications of NACA are indexed in these would help to reach a wider audience, possibly with a bias towards the academic community but not exclusively. A major drawback of most of these databases is that they are commercially published. Hence they are relatively expensive and in general they do not link very well to full text documents (other than those of other commercial publishers).

Examples of relevant bibliographic databases are:

- Aquatic Sciences and Fisheries Abstracts (ASFA)
  <http://www.fao.org/fi/asfa/asfa.asp>

  Aquatic Biology, Aquaculture and Fisheries Resources (ABAFR)
  <http://www.nisc.com/factsheets/qabf.asp>

- CAB Abstracts
  <http://www.cabi-publishing.org/AbstractDatabases.asp?SubjectArea=&PID=125>

- AGRIS
  <http://www.fao.org/agris/>

Secure publications
The hazards of a computer crashing (especially when somebody forgot to make a recent backup), of file corruption and the increasing risk of virus attacks are well known. Publications which reside only on a Web site or only in one digital copy are seriously at risk of being lost. Added to that is the uncertainty caused by rapid developments in computer hardware and software and how sustainable the system is in the long-term if larger investments are needed.

Preservation of publications
Libraries were often the permanent repositories for print publications, ensuring that several copies were available in different locations long after the publication was out of print. There are also inter-library mechanisms that allow borrowing of those copies in many other locations, as well as a national depository library which keeps a permanent copy of its own country's publications. Similar mechanisms have not been put in place for digital publications and the onus of distribution and preservation falls upon the publishing organization. If we believe that we have an obligation to preserve our publications so that they are available for future generations we need a more long-term solution than a Web site publishing system

3.3 Joining the information commons

This section addresses the need for NACA publications to enter the mainstream of aquaculture publications to increase their visibility and to ensure their long-term viability. The Open Access Movement and the Information Commons have generated much strong debate and volumes of literature in recent years. We will simply point to some examples and further reading because it is on the basis of these two initiatives that NACA and other fisheries and aquaculture organizations may find a solution for the long-term availability of their publications as an integral part of the most appropriate "Commons".

The Open Access research literature is composed of free, online copies of peer-reviewed journal articles and conference papers as well as technical reports, theses and working papers. In most cases there are no licensing restrictions on their use by readers. They can therefore be used freely for research, teaching and other purposes. There are various misunderstandings about Open Access. It is not self-publishing, nor a way to bypass peer-review and publication, nor is it a kind of second-class, cut-price publishing route. It is simply a means to make research results freely available online to the whole research community (Joint Information Systems Committee (JISC), 2005).

NACA is already an Open Access publisher so the remaining question is to see if an appropriate Commons exists or whether NACA can participate in any being developed. There are thousands of examples of "Information Commons" already implemented, being planned, at project level etc. Some are discipline based, while others are national or regional in scope. An interesting example for a whole continent is "Commons-sense: towards an African Digital Information Commons" <http://www.commons-sense.org/>. The Web site presents some of the principles and the philosophy of the Information Commons as well as practical guidelines for participation. The "Commons-sense" project develops tools for African organizations and individuals to make copyright work in the service of economic development, public heritage and a shared history for all. An example of a discipline based repository is the Digital Library of the Commons (DLC), which is a gateway to the international literature on the natural resources commons. It allows author submission of documents and includes research on the commons, common-pool resources, and common property. Further information at: <http://dlc.dlib.indiana.edu/>.

An international, discipline-based Commons for the literature of the aquatic sciences, fisheries and aquaculture so far does not exist.

3.4 The aquatic commons

Building a Commons for the literature of the aquatic sciences, including fisheries and aquaculture, is being planned by the International Association of Aquatic and Marine Science Libraries and Information Centers (IAMSLIC) <http://www.iamslic.org>. IAMSLIC is a professional association with over 300 libraries worldwide, including large university libraries in the US and Europe, International Organizations such as FAO, the Intergovernmental Oceanographic Commission (IOC) and the Division for Ocean Affairs and the Law of the Sea (DOALOS), intergovernmental and regional organizations such as the Secretariat of the Pacific Community (SPC), the Network of Aquaculture Centers in Asia-Pacific (NACA) and the International Council for the Exploration of the Sea (ICES) and research institutions from many developing as well as developed countries. For the past three years, IAMSLIC has acknowledged the value of institutional repositories (IRs) and has encouraged members to implement them within their own institutions. Repositories are now spreading because they provide stewardship and global access to the digital assets of institutions and organizations. Impetus for the development of IRs has also come from the Open Access and Open Archives movements.

A key benefit of being an IAMSLIC member, in addition to sharing information and expertise, are the various resource sharing activities which facilitate improved access to global information. Small libraries in developing countries are encouraged to join and to participate in these activities. As was mentioned earlier, much of the literature of fisheries and aquaculture is categorized as “grey” and as such is often difficult to locate and sometimes impossible to obtain. Much of the grey literature from the developed world is being produced in digital form and made available digitally, on CD ROM, via the World Wide Web and from repositories. Many institutions with limited resources, a common feature of small fisheries institutions in developing countries (though also not uncommon in developed ones) are producing publications in digital form but are not able to make them readily available.

The Aquatic Commons is intended to provide a central portal to the literature in the aquatic sciences. It proposes to harvest metadata from existing repositories, mostly at institutional level, and to provide a facility for those who do not have the resources to build their own. These will generally be smaller institutions and organizations which need a secure archive where their documents will be integrated with the global literature of the aquatic sciences. A model has been designed to integrate the efforts of the total community by harvesting metadata where available and by creating repository and harvesting opportunities where needed. The benefit for all IAMSLIC members will be better access to literature which they otherwise do not easily locate.

Hardware and software components of the Aquatic Commons will include, for example, an Aquatics repository, existing OAI-capable repositories, a metadata harvester, an OAI data provider, a search and retrieval interface, a database, and a zebra Z39.50 server. The latter would interface with the existing IAMSLIC Z39.50 Distributed Library, which has been operational for several years (<http://library.csumb.edu/iamslic/ill/search.php>).

Establishing the Aquatic Commons infrastructure will require funding for hardware and software, although cost estimates are relatively low in the order of under US$10 000 start up and US$ 2 000 annual maintenance. Functionality, quality and content will rely upon the expertise of IAMSLIC members and the hosting institution. For organizations like NACA and FAO, a significant longer term investment would need to focus on capacity building to ensure the full participation of developing countries.

References

Asia-Pacific Association of Agricultural Research Institutions. 2004. Information and communication technologies in agricultural research development in the Asia-Pacific region: a status report. Bangkok (Thailand), APAARI. 130 pp.

Boiko, B. 2005. Content management bible. Indianapolis (USA), Wiley Publishing Inc. 1122 pp.

Dravis, P. 2004. Open source software - perspectives for development. Report prepared for the World Bank infoDev Symposium. Available at: <http://www.infodev.org/content/library/detail/837>.

Dspace at MIT. Available at: <http://libraries.mit.edu/dspace-mit/mit/services.html>.

Dspace Federation. Available at: <http://www.dspace.org/>.

FAO and UNESCO. 2005. Digitization and digital libraries module. Information Management Resource Kit (IMARK). CD-ROM. Rome, FAO. Available at: <http://www.imarkgroup.org>.

FAO. 1995. Code of conduct for responsible fisheries. Rome, FAO 1995, 46 pp. Also available at: <ftp://ftp.fao.org/docrep/fao/005/v9878e/v9878e00.pdf>.

FAO. 2005. Towards improving global information on aquaculture. *FAO Fisheries Technical Paper.* No. 480. Rome, FAO. 172 pp. Also available at <ftp://ftp.fao.org/docrep/fao/008/a0066e/a0066e00.pdf>.

GNU EPrints. 2005. The eprints.org software Web site. Available at: <http://www.eprints.org/software.php>.

Gutman, L., Ayers, P. and Booth, D. 2003. Inside Dreamweaver MX. Indianapolis (USA), New Riders Publishing, 1301 pp.

Hecht, T. (2004). Access to information - the Achilles Heel of African fisheries science and technology. *In Report of and papers presented at the Regional Workshop on Networking for Improved Access to Fisheries and Aquaculture Information in Africa. Grahamstown, South Africa, 3-7 November 2003.* pp. 57-60. *FAO Fisheries Report.* No. 740. Rome, FAO. 2004. 232 pp. (Also available at: <ftp://ftp.fao.org/docrep/fao/007/y5519b/y5519b00.pdf>)

IAMSLIC. 2006. Information for responsible fisheries: libraries as mediators. Proceedings of the 31st Annual IAMSLIC Conference, Rome, Oct.10-14, 2005. Fort Pierce (USA), IAMSLIC, 2006.

Joint Information Systems Committee (JISC). 2005. Open Access. Available at: <http://www.jisc.ac.uk/uploaded_documents/JISC-BP-OpenAccess-v1-final.pdf>.

McClelland, D. 2001. Adobe Photoshop bible. New York, Hungry Minds Inc. 1268 pp.

Open Society Institute. 2004. A guide to institutional repository software. Available at: <http://www.soros.org/openaccess/software/>.

Pedersen, J.K. 2004. A study of open source content management. Available at: <http://www.Web-udvikling.dk/uploads/comparison.pdf>.

Scholz, A., Bonzon, K., Fujita, R., Benjamin, N., Woodling, N. Black, P. and Steinback, C. 2004. Participatory socioeconomic analysis: drawing on fishermen's knowledge for marine protected area planning in California. *Marine Policy,* 28(4): 335-49.

STREAM. 2005. Information access surveys (IAS) V1.0. Better Practice Guidelines No. 3. 4 pp. Available at: <http://www.streaminitiative.org>.

Webster, J.G. and Collins, J. 2004. Fisheries information in developing countries. Support to the implementation of the 1995 FAO Code of Conduct for Responsible Fisheries. *FAO Fisheries Circular* No. 1006. Rome, FAO. 2004. 127 pp. Also available at ftp://ftp.fao.org/docrep/fao/007/y5847e/ y5847e00.pdf.

Appendix 1

Glossary of technical terms

Apache: The dominant Web server on the Internet, its function is to serve Web pages and other objects requested by client browsers.

ATOM: A file format that is used to feed information such as Web site headlines and summaries to other Web sites or to aggregators (programs that aggregate feeds from many sites for convenient viewing). ATOM files are typically generated and updated automatically when new information is posted on a Web site. Other sites and aggregators can link to the file and periodically retrieve it to extract the new information for local display. ATOM is often used to share headlines between Web sites.

Bandwidth: In the context of the Internet bandwidth is usually used to refer to the data carrying capacity or maximum transfer rate of an Internet connection.

CMS: In the context of the Web, content management systems are programs that facilitate many aspects of Web publishing, including the collaborative authoring of content by communities of users, the automatic generation of Web pages, organization of collections of documents and search and retrieval services. Most dynamic Web sites can be considered CMS.

Compression: Encoding data in a more efficient manner so as to use less bits (reduce file size). Compression may be “lossy”, in which case some data may be discarded with some loss of data quality (for example in JPEG photographs) or lossless.

CSS: Cascading Style Sheets is a language for presenting the elements of documents written in a markup language such as HTML or XML. CSS is most commonly used to format and position the elements of Web pages.

Digital publications: Any kind of document or information resource published in digital form, for example text documents, databases and audio recordings.

Digital publishing: The process of making information publicly available in digital form, whether it is on the Web, CD or some other electronic media.

GIF: Graphics Interchange Format. A popular graphics format commonly used on the Web.

Host: A Web host is a computer or a company that offers disk space for hosting a Web site.

HTML: HyperText Markup Language is a mark up language designed for creating Web pages. HTML code is interpreted by a user’s browser to display a Web page.

IAS: An Information Access Survey is a tool to identify the key issues about people and what information needs they have, what media sources are available to them and what strategies they use to get information, and how cost effective these are.

IP number: A unique number that devices use in order to identify and communicate with each other on a computer network (such as the Web) utilizing the Internet Protocol standard.

JPEG: A common file format for digital photographs. It employs highly efficient but lossy compression that makes it well suited to use on the Web. At high compression rates image quality is reduced and visual compression artifacts begin to appear.

LAMP: A commonly used combination of open source software consisting of the Linux operating system, the Apache Web sever, MySQL database and PHP scripting language.

Lossy compression: Refers to data compression techniques in which some amount of data is lost. Lossy compression technologies attempt to eliminate redundant or unnecessary information.

MySQL: A popular open source SQL database management system that is commonly used in Web applications and dynamic Web sites.

Open source: Computer software for which the source code is made available for others to view, modify and adapt for new purposes.

PDF: Portable Document Format, a popular format for distributing electronic documents. A major advantage of PDF is that it preserves the original document layout.

PHP: HyperText Prepropressor is a popular scripting language commonly used to write the code to run dynamic Web sites.

PNG: Portable Network Graphics is a popular image format commonly used on the Web, often as an alternative to GIF.

RSS: Really simple syndication. A Web-feed format. It is similar in function to ATOM (see above) but more popular.

Search engine: A program for searching a computer system, database or network for specific kinds of information. The most famous example is Google, the dominant search engine on the Web.

Server: A computer (hardware) that provides services to others, in the context of the Web typically running a software Web server program that will serve pages to clients.

Trackbacks: A mechanism for communication between Web sites (most often between personal Weblogs). If a new item is published on a Web site referring to an entry on another, and if both sites support the TrackBack protocol, then the referring site can notify the other of a comment with an automatically generated “TrackBack ping”. This referred site will often show summaries or links to comments about its story that have been made on other sites below the original entry.

Unicode: An industry standard designed to allow text and symbols from all of the writing systems of the world to be consistently represented and manipulated by computers. It is useful to ensure consistent representation of Web pages and text documents. Web pages may not display correctly if a user does not have the correct character set installed on their computer, but Unicode helps overcome this issue.

URL: A Uniform Resource Locator, commonly called a Web site address, is a string of characters conforming to a standardized format, which refers to a resource on the Internet (such as a document or an image) by its location.

Web server: A computer program that is responsible for accepting HTTP requests from clients, usually through a browser, and serving them Web pages which are usually HTML documents and linked objects.

XHTML: Extensible HyperText Markup Language, it is a reformulation of HTML with stricter syntax in order to comply with XML.

Appendix 2

Open source digital library and digital repository software

A digital library or repository?

There are many competing definitions of the term "digital library". Many of these are complex attempts to translate roles traditionally undertaken by large public or institutional "paper" libraries into the digital world, notably with regards to indexing, archiving and preservation of publications. The reality is that digital publishing is in its infancy. Much of the technology and standards required to perform these functions for digital publications is still experimental, not yet widely accepted or simply too complex to be of practical application in a resource-limited organization.

Just as traditional libraries vary widely in their scale and complexity, so do digital ones. We favour a simple working definition that does not make assumptions about the nature of the information, the scale of the collection, the way it is organized or the capacity of the organization that maintains it: *A digital library is a collection of publications in organized electronic form*. We regard a facility that performs the additional function of long-term preservation and archiving as a *digital repository*.

Early (pre-Internet) digital libraries were usually designed as standalone applications designed to perform a specific task in isolation or over a limited network. However, the Internet has expanded their scope and role by providing a framework for computers to share information. Increasingly, digital libraries are becoming integrated as components within larger Web site frameworks that offer all kinds of information and communication services, including tools to create and publish new documents and to organize and manage the collection, and even to combine and create new decentralized collections out of resources that are physically scattered across the world.

What kind of system do you need?

A wide range of open source digital library CMS are available. However, there is a clear segmentation between those designed for use by individuals or small organizations, and those designed to serve large, diverse institutions such as universities.

Simple digital library systems provide basic catalogue services, i.e. record entry and retrieval. Most have evolved from library automation or personal reference management tools and are designed to manage access to a single collection of text publications. Search functionality is restricted to the metadata (record), and full-text publications are retrieved via a simple hyperlink to an external file, which may be located anywhere on the Web, i.e. not necessarily in a central repository. These systems typically make use of standard open source technologies such as LAMP (Linux, Apache, MySQL and PHP), making them relatively easy to install and maintain, and opening the possibility to integrate them within larger Web site frameworks. They are well suited to managing small collections in small, resource-constrained organizations.

Advanced digital library systems offer additional tools to help create and manage collections, not just of text but all kinds of "digital objects". They offer sophisticated user management and workflow tools that allow communities of people to collaborate in the creation and validation of new records. Some provide mechanisms to manage multiple collections and "repository" functionality to preserve and archive publications. Search functionality may extend beyond metadata to the actual content of publications. While these systems are very powerful and technically far superior, their complexity renders them considerably more difficult to install, configure and maintain. They are more difficult to integrate within a larger Web site framework. They are best suited to management of large collections in large organizations that have good internal technical support.

Most organizations which NACA works with are relatively small, with a basic need to organize their publications and make them available to the public. For this, a simple system will suffice, with the added advantage that it can be installed quickly with "off the shelf" packages and with a minimum of training. However, organizations that publish in high volume, have specialized needs or that are seeking to preserve publications for the long term, i.e. to take on an archival function as well, may require a more advanced system.

Just like portal CMS, digital library CMS vary widely in the features that they offer and it is important to take a close look at the way features have been implemented when comparing products. In addition to the general considerations previously suggested for portals, (such as use of standards, internationalization, documentation and support), features to look for in digital library CMS include:

- ***Extensive search functionality*.** Does it support full-text search of documents, or metadata only? What formats?
- ***Metadata management*.** What metadata set does it use internally? Can it be customized? What formats can it import and export? The system must be able to export data in widely interchangeable formats in order to exchange data with other systems.
- ***Collection management*.** Can you create and manage multiple collections from one data set?
- ***Tools for batch adding records and files*.** Important if you need to import large volumes of information.
- ***Archival support*.** Does the system offer tools for archiving publications? What formats?
- ***Scalability.*** Will the system be able to cope with the projected growth of your collection?
- ***Interoperability support for sharing collections.*** For example, the Open Archives Initiative Protocol for Metadata Harvesting provides a way to make your metadata (i.e. publication records) available through the Web, facilitating exchange between systems. There are several different versions of the protocol.
- ***RSS support*.** Syndication feeds provide an easy (if limited) way to share new publications with other Web sites. This can play an important role in raising awareness of new resources.
- ***Platform and system requirements*.** Can you find a commercial host that provides all necessary services or will you need to maintain your own server?

A full review of digital library systems is beyond the scope of these guidelines, but details of some systems we regard as useful are given below. For a summary covering most advanced digital library systems download *A Guide to Institutional Repository Software*, available from the Budapest Open Access Initiative Web site <http://www.soros.org/openaccess/software/>. A tutorial on the use of the Greenstone system (which is not covered by the above mentioned guide) to create a collection is available on the *Digitization and Digital Libraries* Module of the Information Management Resource Kit (IMARK), distributed on CD-ROM and via Internet <http://www.imarkgroup.org> by FAO and UNESCO, which also includes the Greenstone software.

Simple digital library / Web-capable bibliographic systems based on LAMP technologies

RefBase

RefBase is a Web-based, platform independent application for managing scientific literature and citations. Disadvantages of Refbase are that it is not currently well internationalized and there is no documentation, beyond some basic installation instructions. However, the authors have advised that they intend to implement a more comprehensive language system in future versions.

Refbase may be downloaded from <http://refbase.sourceforge.net/>. A community support forum and listserv e-mail group is available, although at time of writing activity was limited. Refbase is a member of the Bibliophile project, which aims to promote collaboration between developers and end users of bibliographic databases. Bibliophile aims to establish a technology that will enable cross-site searches for references.

WIKINDX

WIKINDX is a Web-based bibliographic and quotations / notes management system that is designed for both personal use and as a multi-user, collaborative tool. The customizable home page and integrated news page give WIKINDX potential to be used as a simple Web site in addition to providing a digital library service. A disadvantage of the system is that most functionality, including the search, is contained in drop down menus that are not immediately visible. Like Refbase, WIKINDX is a member of the Bibliophile Web literature database initiative. WIKINDX is available from <http://wikindx.sourceforge.net/>.

Basilic

Basilic is a bibliography server designed for research laboratories, intended to automate the distribution of publications via the Internet. Each publication has an associated summary page that provides the download link, abstract, citation and associated images. Basilic is exceptionally well presented with a very clean layout and navigation. Basilic's weakness is its limited search functionality, which only extends to title, author and year, limiting its utility to managing small collections of documents. Basilic is available from <http://artis.imag.fr/Software/Basilic/>.

Advanced digital repository systems

If you are likely to develop a large collection (probably in the thousands of documents) and if you *do* have access to good IT and support resources, you might want to consider one of the "industrial strength" systems. With the possible exception of Greenstone we consider them over-powered for small collections.

Greenstone

Greenstone is an excellent suite of suite of software for building and distributing digital library collections. It is produced by the New Zealand Digital Library Project at the University of Waikato, and developed and distributed in cooperation with UNESCO and the Human Info NGO. Greenstone is available for download from <http://www.greenstone.org/cgi-bin/library>.

EPrints

Eprints was designed to allow research institutions to create online archives that comply with the Open Access Initiative protocol for metadata harvesting (OAI – see chapter 4). It was designed primarily to be used by researchers and institutions to maximize the access to and impact of their peer-reviewed research output, but it is also suitable for use to create OAI-compliant archives for open-access journals. It was developed at the University of Southampton and is biased towards text documents.

We feel that the documentation of Eprints is somewhat lacking for software of this complexity. A detailed user guide does exist, but it is quite technical in nature and not suitable for people without an IT background. This will be a substantial barrier to implementation of Eprints by organizations with limited in-house IT support. Eprints is available from <http://www.eprints.org>.

DSpace

DSpace is a digital repository system designed by MIT to capture, store, index, preserve, and redistribute research data from large, multidisciplinary research organizations. The user interface of DSpace is not yet extensively translated (we have references to partially completed or work in progress for German, French, Portuguese, Hungarian and Bahasa). Localization does not appear to be a simple matter. DSpace is available from <http://www.dspace.org/>.

FAO FishCode Reviews

1 Pintz, W.S. Tuna and bottom fishery licence management: Tonga. *FAO/FishCode Review*. No. 1. Rome, FAO. 2003. 35p.

Fish are now the largest single export from the Kingdom of Tonga. However, expansion of the industry faces severe infrastructure constraints, and granting substantial numbers of new longline licences without resolving the constraints could seriously affect all Tongan commercial fisheries.

2 Gillett, R. Aspects of fisheries management in the Maldives. *FAO/FishCode Review.* No. 2. Rome, FAO. 2003. 61p. (*Restricted distribution)*

The inshore marine resources of the Maldives, an atoll environment, are being increasingly exploited for baitfishing, food for local residents, consumption by tourists, exports and non-extractive uses such as dive tourism. This situation must be reconciled with the limited nature of the resources.

3 Die, D.L.; Alió, J.; Ferreira, L.; Marcano, L.; Soomai, S. Assessment of demersal stocks shared by Trinidad and Tobago and Venezuela. *FAO/FishCode Review*. No. 3. Rome, FAO. 2004. 32p.

The FAO/WECAFC Workshop on assessment of demersal stocks shared by Trinidad and Tobago and Venezuela (2002) initiated an assessment of the shrimp stocks shared by the two countries. The main conclusion of the assessment is that some shrimp stocks are being severely overfished and are suffering as a result.

4 Gillett, R. The marine fisheries of Cambodia. *FAO/FishCode Review.* No. 4. Rome, FAO. 2004. 57p.

Excess fishing effort and associated declines in abundance of target species are the most serious problems facing Cambodia's marine fisheries: resource sustainability will require restrictions on resource access.

5EN FAO/FishCode. Seminar on responsible fisheries management in large rivers and reservoirs of Latin America. *FAO/FishCode Review*. No. 5. Rome, FAO. 2004. 72p. [En]

This report of the Seminar on Responsible Fisheries Management in Large Rivers and Reservoirs in Latin America (2003), attended by experts from member countries of the Commission, observers from other regional bodies and representatives from local fishing communities in El Salvador, presents the principles of responsible fishery management in Latin America as well as a selection of national reports.

5SP FAO/FishCode. Seminario sobre ordenación pesquera responsable en grandes ríos y embalses de América Latina. *FAO/FishCode Revista.* No. 5. Roma, FAO. 2004. 78 p. [Sp]

El Seminario sobre Ordenación Pesquera Responsable en Grandes Ríos y Embalses de América Latina (2003) se efectuó en San Salvador en asociación con la novena reunión de la Comisión de Pesca Continental para América Latina (COPESCAL). Participaron expertos de países miembros de la Comisión; observadores de otros organismos regionales y representantes de comunidades pesqueras locales de El Salvador. Se presentaron dos documentos sobre los principios de la ordenación pesquera responsable en grandes ríos y embalses en América Latina y una selección de informes nacionales.

6 Swan, J. National Plans to combat illegal, unreported and unregulated fishing: models for coastal and small island developing states. *FAO/FishCode Review.* No. 6. Rome, FAO. 2003. 76p.

These case studies for use in FAO regional and subregional workshops were prepared in accordance with the FAO International Plan of Action to Prevent, Deter and Eliminate IUU Fishing. The "Republic of Galactia" and the "Alpha Islands" are fictitious, but the fisheries profiles presented draw on typical existing circumstances.

7 Kuemlangan, B. Creating legal space for community-based fisheries and customary marine tenure in the Pacific: issues and opportunities. *FAO/FishCode Review.* No. 7. Rome, FAO. 2004. 65p.

The laws of Pacific Island countries generally support traditional fisheries management with only modest efforts to encourage the use of customary marine tenure-based community fisheries management. Government commitment for the role of customary marine tenure in community-based fisheries management, with support from interested stakeholders, will complement efforts for promoting sustainable utilization of fisheries resources and improved livelihoods in the Pacific region.

8 FAO/FishCode. Report of the Workshop on Development of a Management Plan for Tomini Bay Fisheries, Indonesia. *FAO/FishCode Review.* No. 8. Rome, FAO. 2004. 31p.

Tomini Bay fishery resources are still considered to be underexploited, but annual catches have increased dramatically over the past ten years. In the absence of a fisheries management body, The FAO/Government of Indonesia Workshop on the Development of a Management Plan for Tomini Bay Fisheries (2003) provided a starting point for addressing responsible fisheries issues and laying the groundwork for a fisheries management plan

9 FAO/FishCode. Report of the National Conference on Responsible Fisheries in Viet Nam, Hanoi, Viet Nam, 29–30 September 2003. *FAO/FishCode Review.* No. 9. Rome, FAO. 2004. 94p.

This national conference was organized in the context of increasing problems faced by Vietnamese fishers in maintaining and improving their livelihoods through coastal and offshore fisheries; some coastal fish resources in particular are being heavily over-exploited.

10 Stanley, J. Institutional review of the National Fishing Corporation and the Fisheries Department of Tuvalu. *FAO/FishCode Review.* No. 10. Rome, FAO. 2004. 47p. (*Restricted distribution)*

The economic growth and development of Tuvalu depend on its marine resources and especially its relatively rich tuna resources. Although the primary concern of the government is the sustainable economic development and management of tuna, there is also potential for the development of other marine products, particularly deep bottom fish.

11 García Mesinas, A. Lineamientos para un Código de Ética de Pesca y Acuicultura para El Salvador. *FAO/FishCode Revista.* No. 11. Roma, FAO. 2004. 59p. [Sp] (*Restricted distribution)*

Este documento presenta los resultados de un proyecto llevado a cabo a través del Programa FishCode de la FAO a petición del Gobierno de El Salvador para desarrollar los lineamientos a nivel nacional del Código de Ética de la Pesca y Acuicultura. El trabajo se realizó coordinado a través de la Oficina Regional de América Latina (RLC) y la Representación de FAO de El Salvador.

12 FAO/FishCode. Report of the National Workshop on the Code of Conduct for Responsible Fisheries and its practical application to coastal aquaculture development in Viet Nam. *FAO/FishCode Review.* No. 12. Rome, FAO. 2004. 47p.

The National Workshop on the Code of Conduct for Responsible Fisheries and its Practical Application to Coastal Aquaculture Development in Viet Nam took place in Hué from 3 to 4 October 2003. The Workshop aimed to build awareness among national and provincial stakeholders about the need to develop and implement an Aquaculture Code of Conduct for Viet Nam. Coastal aquaculture in Viet Nam, particularly shrimp culture, has developed rapidly in recent years. Although shrimp farming has brought many benefits to coastal communities, it is associated with high social and environmental risks.

13 FAO/FishCode. Report of the National Seminar on the reduction and management of commercial fishing capacity in Thailand. *FAO/FishCode Review.* No. 13. Rome, FAO. 2005. 59p.

The marine capture fisheries sector is more capital intensive than is appropriate for Thailand's resource endowment, and there is an urgent need for fishing capacity reduction for improved fisheries management and protection and conservation of fish habitats and other threatened coastal resources. Failure to achieve this will have serious consequences for the most vulnerable people in coastal communities, fish consumers and society at large.

14 FAO/FishCode. Reports of the regional vessel monitoring systems workshops: Southwest Indian Ocean, Central America, the Caribbean and Southeast Asia *FAO/FishCode Review.* No. 14. Rome, FAO. 2005. 91p.

Four regional workshops on vessel monitoring systems (VMS), respectively covering the South West Indian Ocean, Central America, the Caribbean and Southeast Asia, were organized and implemented in succession from September 2003 to October 2004. The workshops were intended to promote the use of VMS as an additional instrument for the management of fisheries, both at a national level and in cooperation with regional fisheries bodies. They comprise one aspect of FAO's larger set of activities to implement the International Plan of Action (IPOA) to Prevent Deter or Eliminate Illegal, Unreported and Unregulated (IUU) Fishing. The document includes a CD-ROM.

15 FAO/FishCode. Fishery policy in the Marshall Islands. *FAO/FishCode Review*. No. 15. Rome, FAO. 2005. 33p.

Fisheries play a key role in the economy of the Republic of the Marshall Islands (RMI) and in the lives of its people. Substantial tuna resources are exploited from the country's vast exclusive economic zone, largely by foreign fishing vessels operating under license. Coastal fisheries are important for subsistence purposes, and also generate income for atoll communities. RMI's well-recognized remote and pristine outer atoll lagoons are considered suitable for targeted commercial mariculture development. The Marshall Islands Marine Resources Authority is investing heavily in formulating its outer island work programmes, involving both coastal fisheries and mariculture research and development. A cautious and transparent approach is needed, with attention to partnerships between communities and private business concerns and the use of incentives involving seed funding, technical assistance, transport facilitation, and other support activities.

16 FAO/FishCode. Report of the Conference on the National Strategy for Marine Fisheries Management and Development in Viet Nam. *FAO/FishCode Review.* No. 16. Rome. FAO. 2005. 64p.

The Conference on the Strategy for Marine Fisheries Management and Development in Viet Nam, (Hanoi, 26 – 27 April 2005) was organized by the Ministry of Fisheries of Viet Nam (MOFI) in close collaboration with the Research Institute Marine Fisheries, the DANIDA Fisheries Sector Programme Support (FSPS) and the FAO FishCode Programme. It represented the culmination of a process that started in 2003 with the Conference on Responsible Fisheries in Viet Nam and that included a number of local level consultations as well as a senior expert meeting in 2004. The 2005 Strategy Conference was attended by a wide range of sectoral stakeholders, representing local and commercial fisheries interests, national and provincial government bodies, bilateral development assistance agencies and international organizations. Observations and recommendations received from the Conference have provided a basis for MOFI to finalize the Strategy for official Government approval.

17 Macfadyen, G.; Cacaud, P.; Kuemlangan, B. Policy and legislative frameworks for co-management. Paper prepared for the APFIC Regional Workshop on Mainstreaming Fisheries Co-management in Asia Pacific. Siem Reap, Cambodia, 9–12 August 2005. *FAO/FishCode Review.* No. 17. Rome, FAO. 2005. 51p.

This paper was prepared for the Asia-Pacific Fisheries Commission workshop on mainstreaming fisheries co-management, held in Cambodia in August 2005. It examines the policy and legislative frameworks for co-management in thirteen countries in Asia and the Pacific, and the extent to which these frameworks hinder or support co-management practices. The nature of policy and legislative frameworks is varied, as is commitment by governments to co-management – in some cases support is more rhetoric than reality, with insufficient real transfer of powers and financial resources to local levels. Through an analysis of the different case studies, "lessons learned" are presented and a number of conclusions drawn about the key characteristics of a supportive policy and legislative framework based on some ideas about "best practice." The adoption of these characteristics by governments would demonstrate their commitment to co-management and increase the likelihood of co-management success.

18 FAO/FishCode. Report of the Global Fisheries Enforcement Training Workshop. Kuala Lumpur, Malaysia, 18–22 July 2005. *FAO/FishCode Review*. No. 18. Rome, FAO. 2007. 66p.

The Global Fisheries Enforcement Training Workshop (Kuala Lumpur, Malaysia, 18–22 July 2005) brought together operational-level monitoring, control and surveillance (MCS) professionals for the global community who are dedicated to resolving illegal, unreported and unregulated (IUU) fishing issues. Hosted by the Government of Malaysia in cooperation with the MCS Network, the FAO FishCode Programme and the European Union, the Workshop provided participants with training on a wide range of MCS topics and gave them the opportunity to share information and experiences, latest developments and new ways to improve fisheries enforcement. Among other subjects, the Workshop reviewed enforcement techniques and MCS operations through individual presentations, case studies and panel discussions. Participants discussed a wide range of tools available to assist countries in dealing more efficiently with IUU fishing, as well as methods of applying these tools through legal systems.

19 Gillett, R. and Moy, W. Spearfishing in the Pacific Islands. Current status and management issues. *FAO/FishCode Review*. No. 19. Rome, FAO. 2006. 76p.

Spearfishing is growing in importance in the Pacific Islands. While its management has featured as a topic in some regional-level meetings, detailed information on spearfishing is surprisingly scarce. In early 1994, the Secretariat of the Pacific Community (SPC) proposed to consolidate information on spearfishing in the Pacific Islands. The original intent was to undertake a review of the available literature through a desk study. With the realization that many issues related to spearfishing are undocumented, the strategy was changed to include some field work. These activities were supported by the FAO FishCode Programme. This report reviews spearfishing in selected Pacific Island countries and identifies the important species caught by and the major problems associated with the method. It further considers possible interventions to mitigate these problems and the assistance that is likely to be required by Pacific Island countries in the management of their spearfisheries. For several reasons, a complete ban of scuba spearfishing coupled with effective enforcement is the single most important spearfishing management measure.

20 Wilkinson, S.; Collins, J. Information in support of responsible fisheries and aquaculture. Guidelines on digital publishing: a practical approach for small organizations with limited resources. *FAO/FishCode Review.* No. 20. Rome, FAO. 2007. 68p.

These Guidelines on digital publishing are targeted primarily at small organizations with limited resources in developing countries, in order to facilitate decision-making on how to publish and disseminate their information. The Guidelines are based on the years of experience of the Network of Aquaculture Centres in Asia-Pacific (NACA) and its partners. The approach is practical in orientation, covering topics including: (a) planning, building and maintaining a sustainable digital publishing system; focusing on a common scenario of setting up a Web site as a digital publishing platform; (b) producing user-friendly digital publications and making them accessible; (c) some recent international developments in digital publishing; and (d) recommended software tools and technical resources for further reading.

For further information, or to obtain copies,
please contact:
FishCode Programme/Fisheries Department
Food and Agriculture Organization of the United Nations
Viale delle Terme di Caracalla
00153 Rome, Italy
Tel.: (+39) 06 5705 5396/6807 **Fax:** (+39) 06 5705 6500
E-mail: FishCode@fao.org
http://**www.fao.org/fi/fishcode.htm**

fishCODE

The FAO Programme of Global Partnerships for Implementation of the Code of Conduct for Responsible Fisheries

These Guidelines on digital publishing are targeted primarily at small organizations with limited resources in developing countries, in order to facilitate decision-making on how to publish and disseminate their information. The Guidelines are based on the years of experience of the Network of Aquaculture Centres in Asia-Pacific (NACA) and its partners. The approach is practical in orientation, covering topics including: (a) planning, building and maintaining a sustainable digital publishing system; focusing on a common scenario of setting up a Web site as a digital publishing platform; (b) producing user-friendly digital publications and making them accessible; (c) some recent international developments in digital publishing; and (d) recommended software tools and technical resources for further reading.

NACA

For further information:

FishCode Programme
Fisheries and Aquaculture Department
Food and Agriculture Organization of the United Nations
Viale delle Terme di Caracalla
00153 Rome, Italy
Tel: (+39) 06 57055396/6807 Fax: (+39) 06 57056500
Email: FishCode@fao.org
www.fao.org/fi/fishcode.htm

FAO FIAT PANIS