

MAGDIR

Managing Australian Government

Digital Information Resources

Presentation by Pam Gatenby, National Library of Australia

third annual AGLIN conference, July 2006¹

¹ The contribution of Colin Webb, Director Preservation Services, National Library of Australia, to preparation of the text is acknowledged.

I would like to thank the organisers of the AGLIN conference for inviting me to talk about the Working Group on Management of Australian Government Digital Information Resources – also known as MAGDIR, a more memorable and manageable name! This is still a fairly new undertaking for the National Library and we welcome opportunities to tell people about it.

Before I explain what MAGDIR is I'd like to provide some contextual background to explain how it fits into the Library's strategic priorities. The National Library has been active for over a decade now in responding to the huge impact that digital technology has had on the dissemination of information and on how people want to access and use information.

The Library is mandated by its act to play a leadership role in ensuring Australian library materials remain accessible. However, in the digital environment, many information resources that would once have been made available through libraries are now served from other organisations. The Library has a strong interest in working with others to define and manage resources deemed to have ongoing value.

Our strategic planning document, *Directions for 2006-2008*, states that in order to ensure a significant record of Australia and Australians is collected and safeguarded, we will build a national collection of resources in many forms, including digital form.

As most of you would be aware, we have been collecting and managing digital resources routinely for at least the last 10 years. Significant Australian web sites and online publications have been selected for access through the PANDORA archive since 1996 and last year we carried out our first harvest of the whole Australian web domain. We will conduct another domain harvest next month and aim to collect around 500 million URLs!

The web resources we collect are all freely accessible to the public. There is a huge amount of other digital information out there that we do not attempt to collect – for instance, statistical data sets, geospatial mapping data; raw research data; commercial content and databases.

As with other agencies, both within and outside what you might call the “cultural or heritage sector”, digital information is not just an add-on, but intrinsic to our business.

Collaboration

From the outset, our strategy for collecting Australian digital resources has assumed collaboration. We were well aware that the Library alone could only ever accept responsibility for a small part of the public information being created and disseminated in digital form in Australia. Our first formal collaborative undertaking was with the state and territory libraries to establish and develop the PANDORA archive. Other collecting institutions are now partners in PANDORA as well.

We are now working with the higher education sector on research projects funded by the Department of Education, Science and Technology (DEST) aimed at strengthening the national research infrastructure. The project we are most active in now is the Australian Partnership for Sustainable Repositories (APSR) which is identifying requirements for repositories to be able to manage digital content over time and provide meaningful access to it.

Our envisaged model for national collaboration entails a distributed network of publicly accessible, “trusted” digital archives, involving a range of organisations who accept responsibility for the digital resources they create or collect, and agree to common standards and practices.

MAGDIR

Why we established the MAGDIR working group

It is within this context that the Library decided to explore the possibility of a collaborative arrangement with Commonwealth government agencies that have some responsibility for digital resources.

A particular stimulus for this decision was the international web archiving conference the Library held in late 2004. The conference highlighted the fact that a wide range of digital information resources are being produced in Australia with enduring research, and cultural value. We considered that there was an urgent need to respond to the challenges of managing these resources with a more coherent and cohesive national approach. To test the feasibility of such an approach, we decided to approach some other national agencies to start exploring possibilities. This led to the establishment of the MAGDIR (Managing Access to Government Digital Information Resources) Working Group.

MAGDIR is a small, informally constituted group which currently represents five Australian government agencies - AGIMO, the Australian Bureau of Statistics, GeoScience Australia, the National Archives of Australia, and the National Library of Australia.

There is nothing exclusive about this list of members. It is simply a small group that the Library invited to be involved either because of the range of information resources they manage, the roles they have, and/or the connections that already existed with us.

The group first met in June 2005 and there was a consensus of opinion that we should start small, putting aside the question of a wider membership until we had worked out what we wanted or needed to do. It is still fairly early days for this group, and we are still feeling our way a little, developing an agenda for action, working out what kind of group we are, what we should aim to achieve, who we should invite to work with us, and what kind of status and profile we want to have.

Having said that, we are clear on our primary focus: to keep government digital information resources both safe and useable for as long as they are needed, and to do this as efficiently and effectively as possible.

However, it is important to stress that MAGDIR is not based on any sense that we have to start from scratch, or re-invent the wheel. In fact, the primary purpose of the group is to ensure that we can avoid re-inventing wheels, but rather take advantage of and share developments happening nationally, internationally and within our own agencies.

While many issues remain to be defined and solved, significant progress is being made internationally. There has been a surge of developments in methods and tools for managing digital materials and many players here and overseas have been doing excellent work including developing tools, formats, and standards; raising stakeholder awareness and commitment; and developing policy and business model frameworks.

Despite that progress however, a range of technical, organisational, policy and funding issues threaten the future usefulness of digital information resources managed by Commonwealth government agencies and must be addressed soon if we do not want to lose access these important resources, most of which represent a significant investment of government and public funds.

We were also aware that some sectors were struggling to deal adequately with digital issues, especially issues of long-term management, but that even the lead agencies who seemed to be doing well – including the NLA – were really dealing with the tip of an information iceberg.

We needed, and still need, to make significant investments in policy, research, infrastructure, and perhaps even organisational change, if we are going to manage the huge and diverse kinds of data of potential long-term value as information that the community and government may well be expecting us to manage and keep available for use.

What we envisaged was a coalition of agencies with a big stake in managing digital information, working towards some national, or whole of government, or at least multi-lateral approaches that might give us all better tools, better planning environments, maybe even better community awareness, to help us do this important job.

With this in mind, a key motive for setting up a consortia of government agencies was to use it as a vehicle to raise awareness with government and funding bodies of the urgency of addressing the issues. Unlike in some other countries, to date there has been very little government assistance provided to cultural agencies and other institutions with responsibilities for digital resources. Without this, the future of large amounts of digital content is very uncertain.

To have any impact, we understood that it was necessary to take a whole of government approach and to work within the framework of the AGIMO agenda which emphasises shared accountability, shared learning and expertise; generic solutions and approaches; and minimising unnecessary duplication.

We were also inspired by the achievements of consortia with similar objectives that have been set up in other countries. Two noteworthy examples are the cross-sectoral UK Digital Preservation Coalition (<http://www.dpconline.org>) and the US National Digital Information Infrastructure and Preservation Program (NDIIPP) (<http://www.digitalpreservation.gov/about/index.html>).

Objective and goals of MAGDIR

As already mentioned, the MAGDIR working group was set up initially to explore some fundamentals that might serve as a platform for the involvement of a larger group of agencies. The group has met four times since June 2005 and has agreed on its name – MAGDIR – and on its broad objective and goals, and implemented a work plan for its first phase.

The working group adopted the following statement of objective:

Australian Government information resources in digital form are controlled, managed and preserved so as to permit their ongoing and reliable use by government and the community for as long as needed.

Specific goals of the group are to:

- understand the responsibilities of partner organisations
- document responsibilities in statements of agreement
- identify expertise within partner organisations
- identify opportunities for working together based on an assessment of current performance; and to
- promote the objective and goals of the group.

What MAGDIR has been doing?

The working group has met 4 times since June 2005. Our action agenda has been focused on three areas:

- Surveying our activities
- Mapping responsibilities
- Benchmarking our practices.

Survey of agency issues

First up, we surveyed ourselves. We looked at the following areas:

- The types of digital resources for which each agency has responsibility
- The strategies each agency currently uses to address its responsibility and the strategies required
- Our strategies for providing access.
- The specialist expertise available in our agencies, and notable expertise gaps
- Issues or challenges we thought would warrant collaborative action.

While the survey gave us interesting information, it also served as the basis for the next steps the group has taken.

Mapping responsibilities

From the survey exercise, the working group prepared a paper mapping our respective responsibilities for managing digital information resources. We did this for 2 reasons:

- because we believe that accepting responsibility is a critical first step in ensuring ongoing survival and accessibility of digital information, and
- because we all had a business interest in seeing if we could rationalise any overlaps in responsibility, to save ourselves some unnecessary work.

The overlaps we identified relate to mapping and statistical data. We are now working on draft agreements to clarify and rationalise the overlaps where we can. For instance, The National Library has traditionally collected Australian maps but digital mapping is a very different situation from print maps. GeoScience is much better placed to accept responsibility for maps in digital form as they have the expertise and technical infrastructure required so it is unnecessary for us to have a role as well. The same goes for statistical data. We have archived some online statistical content issued by the Bureau of Statistics in PANDORA but this has not been easy because of technical complexities with how the data is presented. Now that ABS are distributing most of their statistical products online, it makes more sense for them to accept responsibility for their management and ongoing access than for us to do this.

Benchmarking

The second step we have taken, based on our survey, was to start looking at our practices and strategies etc, with the view to establishing some best practices.

We approached this with two thoughts in mind:

- we wanted to be able to find a credible way of benchmarking to assess whether we are doing what is required to meet long-term responsibilities or whether we might need to change our practices to have confidence that things are under control

- we also would like to be able to help other agencies by articulating practical principles and benchmarks they could use.

We see this step as necessary to raise awareness of the state of control of valuable government information assets in digital form.

The path we have chosen to explore benchmarking is a draft standard for Trusted Digital Repositories, which is the result of international effort led by a working group established by the Research Library Group and the US National Archives and Records Administration (NARA). The National Library has contributed to this effort.

The working group has issued an Audit Checklist for Certification of Trusted Digital Repositories and MAGDIR is interested in applying it to our own repositories. While it may not fully meet our needs, we all thought it was a good place to start, looking for areas of compliance in our agency approaches, and areas where compliance would be difficult to achieve.

An initial review of the Checklist has identified a number of areas of broad compliance in this list, and a number of areas of concern, both in terms of things we know we need to do, and things that just don't seem to fit our needs very well. We intend to do more work on this.

One avenue for taking this forward is an international forum on TDR methodologies that we are organising in conjunction with a university group called the Australian Partnership for Sustainable Repositories, of which NLA is a member. The forum is on September 1 at the National Library. We envisage a lively discussion between authors of the Audit Checklist and members of MAGDIR (and others). You are all welcome to attend.

What we have learnt so far

It probably comes as no surprise that we have found some diversity in the group – our agencies are dealing with some quite different issues, different types of data, for different kinds of clients and stakeholders, and trying to meet some quite different needs.

We are also found there is quite a high level of commonality. We are all worried about the ongoing availability of financial and technical resources, and we all wish there were more tools available to help us; many of us struggle to define the scope of future community expectations.

But perhaps one of the most encouraging things is the potential for synergy. The agencies involved in MAGDIR all play a wider leadership role in their different sectors and all have well established networks through which they are trying to improve the management of digital information resources nationally within their sectors. You might be aware of the National Data Network led by ABS, the Australian Spatial Data Infrastructure (ASDI), in which GeoScience Australia plays such an important role; the Australasian Digital Recordkeeping Initiative (ADRI) forged by the National Archives, and I have mentioned the strong links between the National Library and all the State and Territory Libraries and the University Libraries. MAGDIR is keen to build on these networks, not compete with or try to replace them. We see MAGDIR as a vehicle for drawing on the expertise and developments in each of our sectors, and for sharing it across sectoral boundaries to the benefit of all concerned. If we can take advantage of it, there is an opportunity to build on the synergy of these existing networks.

What next?

The next steps in our action plan are to:

- Finalise agreements on responsibilities
- Look more closely at each other's approaches and assess what might be usefully transferred. For example, at the Library we have done some work with the Archives' Xena tool, but we are keen to make a more thorough assessment of whether and how we could use Xena and the format normalisation approach.
- Look at the barriers and constraints to meeting our long term responsibilities, and to articulate what would be needed to overcome them, with reference to the Trusted Digital Repositories Checklist.
- Develop an expertise network that works.

Something else we will do before the end of this year is to make some decisions about the future scope, directions, membership and modus operandi of the group. If there is support for the group continuing to work together we will invite other agencies to join in, establish it on a more formal basis, and launch the collaboration as a whole of government initiative. The aim would be to ensure MAGDIR complements other e-government strategies and is recognised within the official AGIMO agenda.