



**ODF:
Our Document Future**

Donna Benjamin
Open Source Industry Australia
Creative Contingencies Pty Ltd

XTECH: Building Web 2.0
Open Data Stream
18 May 2006, Amsterdam



Three Things

- Digital Preservation.
Why does it matter?
- OpenDocument Format.
What's the big deal?
- Australian efforts on the
Digital Preservation front

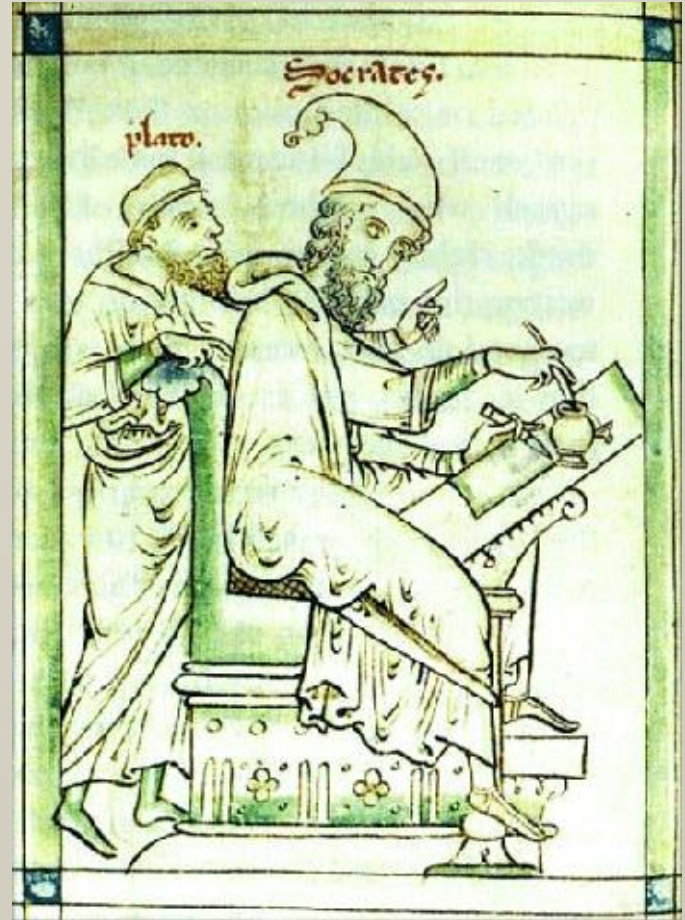


Persistence of Memory

"Humans are the only known species to have devised information storage outside of brains - Libraries (and now the Internet) are a communal memory that permits communication across time."

Steven Dutch <http://www.uwgb.edu/dutchs/CosmosNotes/cosmos11.htm>

Socrates and Plato



Storage Media



Color photo by Jeff Rotherberg

Photo 1: The Rosetta Stone has far outlasted digital media



Four Elements

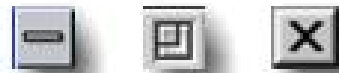
There are four elements of digital technology that threaten our future access to stored data.

- The media on which it is stored.
- The hardware used to design and create it.
- The software used to create, read or edit that data and
- The standards we rely on to record and format data.

Technology Preservation



Emulation



Please Choose an Operating System

Which
OS
do you want to
Emulate
Today?

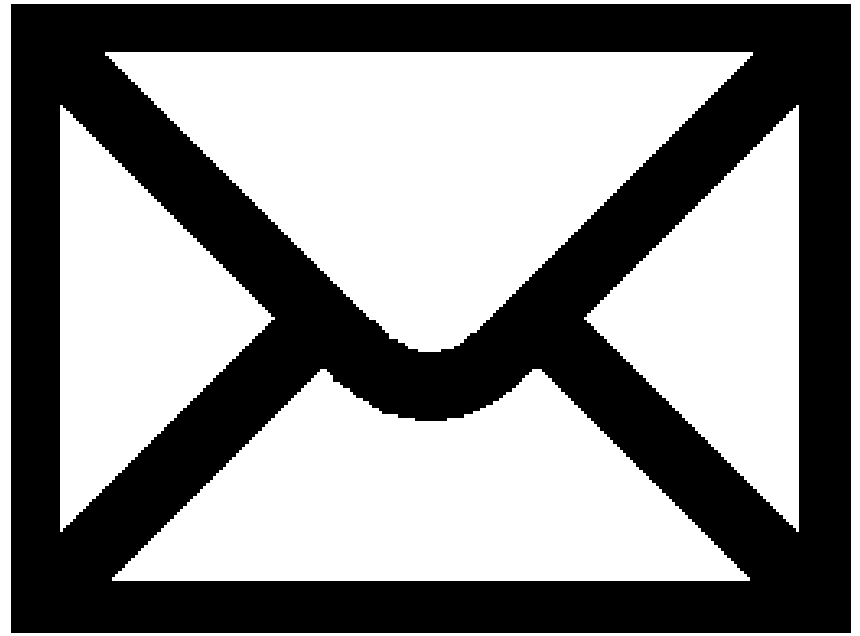
AMIGA



Migration



Encapsulation



XML

```
<Whitepaper>  
  <PaperTitle>XML and Digital Preservation</PaperTitle>  
  <Chapter>  
    <ChapterTitle>XML and its family of standards</ChapterTitle>  
    <Lead>This chapter describes (...)</Lead>  
    <Section>  
      <SectionTitle>Hors-d'oeuvre: form, formatting, structure and content</SectionTitle>  
      <Paragraph> To give a foretaste of the central idea of XML, let us take this  
        document, which you are reading on a computer screen or on paper. In both cases  
        you can see from the <RegulationConcept>formatting</RegulationConcept> of the  
        published version that this sentence is a part of the section (...)</Paragraph>  
    </Section>  
  </Chapter>  
</Whitepaper>
```

OpenDocument Format



The image shows a screenshot of a Wikipedia article page for 'OpenDocument'. The page layout includes a top navigation bar with links for 'article', 'discussion', 'edit this page', and 'history'. A 'Sign in / create account' link is visible in the top right. Below the navigation bar is a banner for 'Your continued donations keep Wikipedia running!'. The main content area features the title 'OpenDocument' and a sub-header 'From Wikipedia, the free encyclopedia (Redirected from Opendocument)'. The article text explains that the OpenDocument format (ODF) is an open document file format for saving and exchanging editable office documents such as text documents, spreadsheets, charts, and presentations. It was developed by the OASIS industry consortium, based upon the XML-based file format originally created by OpenOffice.org, and ODF was approved as an OASIS standard on May 1, 2005. The text also notes that the standard was publicly developed by a variety of organizations, is publicly accessible, and can be implemented by anyone without restriction. It provides an open alternative to proprietary document formats including the popular but undocumented DOC, XLS, and PPT formats used by Microsoft Office, as well as Microsoft Office Open XML format.

[Sign in / create account](#)

[article](#) [discussion](#) [edit this page](#) [history](#)

Your *continued donations* keep Wikipedia running!

OpenDocument

From Wikipedia, the free encyclopedia
(Redirected from [Opendocument](#))

The **OpenDocument** format (**ODF**), short for the **OASIS Open Document Format for Office Applications**, is an [open document file format](#) for saving and exchanging editable office documents such as text documents (including memos, reports, and books), [spreadsheets](#), charts, and presentations. This standard was developed by the [OASIS](#) industry consortium, based upon the [XML](#)-based file format originally created by [OpenOffice.org](#), and ODF was approved as an OASIS standard on [May 1, 2005](#).

The standard was publicly developed by a variety of organizations, is publicly accessible, and can be implemented by anyone without restriction. The OpenDocument format is intended to provide an open alternative to [proprietary](#) document formats including the popular but undocumented [DOC](#), [XLS](#), and [PPT](#) formats used by [Microsoft Office](#), as well as [Microsoft Office Open XML](#) format (this latter format has

WIKIPEDIA
The Free Encyclopedia

navigation

- [Main Page](#)
- [Community Portal](#)
- [Featured articles](#)
- [Current events](#)
- [Recent changes](#)
- [Random article](#)
- [Help](#)
- [Contact Wikipedia](#)
- [Donations](#)

search

ODF vs MSOpenXML










MS XML uses the non-mixed model to represent narrative (word processing). This sort of mismatch leads to markup like this:

```
<w:p>
  <w:r>
    <w:t>This is a </w:t>
  </w:r>
  <w:r>
    <w:rPr>
      <w:b />
    </w:rPr>
    <w:t>very basic</w:t>
  </w:r>
  <w:r>
    <w:t> document </w:t>
  </w:r>
  <w:r>
    <w:rPr>
      <w:i />
    </w:rPr>
    <w:t>with some</w:t>
  </w:r>
  <w:r>
    <w:t> formatting, and a </w:t>
  </w:r>
  <w:hyperlink w:rel="rId4" w:history="1">
    <w:r>
      <w:rPr>
        <w:rStyle w:val="Hyperlink" />
      </w:rPr>
      <w:t>hyperlink</w:t>
    </w:r>
  </w:hyperlink>
</w:p>
```

The equivalent in **OpenDocument** would look like this:

```
<text:p text:style-name="Standard">
  This is a <text:span text:style-name="T1">very basic
</text:span> document <text:span text:style-name="T2">
with some</text:span> formatting, and a <text:a
xlink:type="simple" xlink:href="http://example.com">
hyperlink</text:a>
</text:p>
```

Format Internals

Name	Date Modified	Size	Kind
▶  Configurations2	23/01/2006, 12:05 AM	--	Folder
 content.xml	23/01/2006, 12:05 AM	104 KB	XML file
▶  META-INF	Yesterday, 2:54 PM	--	Folder
 meta.xml	23/01/2006, 12:05 AM	4 KB	XML file
 mimetype	23/01/2006, 12:05 AM	4 KB	Unix Executable File
▶  Pictures	Yesterday, 2:54 PM	--	Folder
 settings.xml	23/01/2006, 12:05 AM	12 KB	XML file
 styles.xml	23/01/2006, 12:05 AM	64 KB	XML file
▶  Thumbnails	Yesterday, 2:54 PM	--	Folder



Australia

- National Library
 - PADI
 - PANDORA
- Public Record Office of Victoria
 - VERS
- National Archives of Australia
 - Xena

PADI is a subject gateway to international digital preservation resources

[Home](#)

[About PADI](#)

[Search](#)

[Browse Topics](#)

[Feedback](#)

[Contributions](#)

[What's New?](#)

[Partners](#)

[Working Groups](#)

[PADI TRAILS](#)

Don't know where to start? Try the new PADI Trails.

[Historical](#)

[New PADI feature!](#)

RESOURCE TYPES

- ▶ [Events](#)
- ▶ [Policies, Strategies & Guidelines](#)
- ▶ [Projects](#)
- ▶ [Organisations & Websites](#)
- ▶ [Bibliographies](#)
- ▶ [Discussion Lists](#)
- ▶ [Glossaries](#)
- ▶ [Journals & Newsletters](#)

DIGITAL PRESERVATION TOPICS

- ▶ [Data Documentation & Standards](#)
- ▶ [Digital Libraries](#)
- ▶ [Digital Records](#)
- ▶ [Digitisation](#)
- ▶ [Formats & Media](#)
- ▶ [General Resources](#)
- ▶ [Issues](#)
- ▶ [Management](#)
- ▶ [National Approaches](#)

PANDORA

PANDORA AUSTRALIA'S WEB ARCHIVE

National Library of Australia and Partners

[Home](#)
[About PANDORA](#)
[News](#)
[Partners](#)
[Notification form](#)
[Services](#)
[Collaboration](#)
[User survey](#)
[Contact us](#)
[Other archives](#)
[Disclaimer](#)

[Search PANDORA](#) [Advanced Search](#) [Search Help](#)

Browse subjects:

Arts & Humanities	Health	News & Media
Business & Economy	History & Geography	Politics & Government
Computers & Internet	Indigenous Peoples	Science & Technology
Education	Juvenile	Society & Culture
Environment	Law & Criminology	Sports & Recreation

View a [complete list of titles](#) available within PANDORA or browse titles alphabetically
1-9 A B C D E F G H I J K L M N O P Q R S T U V W X Z



vers <forever digital>

news faq contact us priv

VERS STORY STANDARD COMPLIANCE PROJECTS DIGITAL ARCHIVE TRAINING TOOLKIT PUBLICATIONS

Home > VERS Story > The Strategy

Search
Enter Keywords
Go

Welcome

Background

The Strategy

The Victorian Electronic Records Strategy specifies a standard format for electronic records. The Strategy:

- is generic but extensible, so that it will work in conjunction with existing recordkeeping and business practices
- ensures that all records are stored in a documented format, to enable viewing of records in the future, regardless of the system that created them
- specifies methods to automate the capture of records from the desktop and business systems
- specifies ways and forms in which to capture information about records and encapsulate this with the records to ensure that records in the future will be understood in context
- details a method for securing records so that any changes are detectable

printer friendly

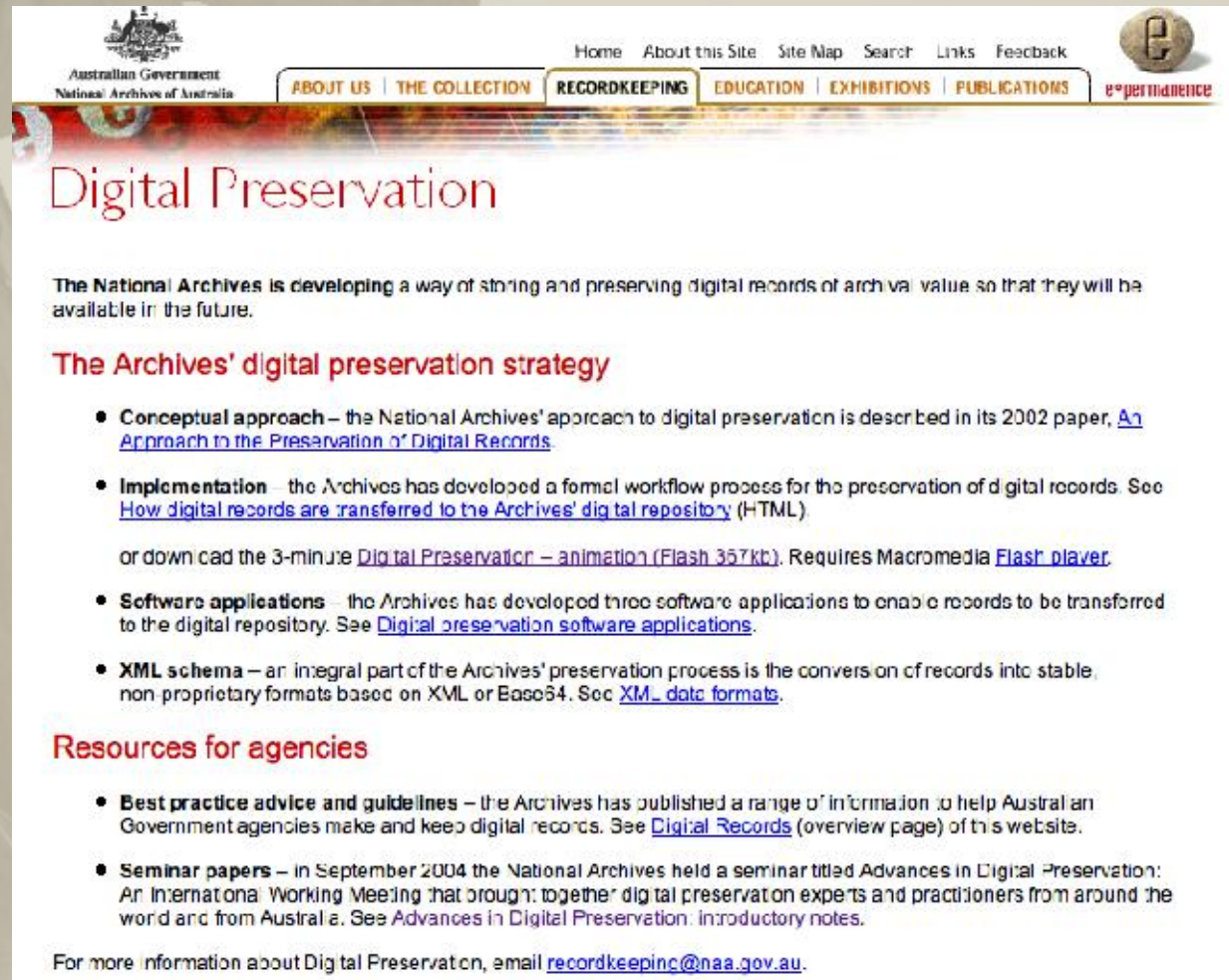
VERS Story | Standard | Compliance | Projects | Digital Archive | Training | Toolkit | Publications

State of Victoria
Victoria Department for Victorian Communities

Site last reviewed: 16 Sep 2004
© State of Victoria 2006 | Disclaimer | Privacy Statement

PUBLIC RECORDS
CREATING HISTORY

National Archives of Oz



The screenshot shows the National Archives of Australia website. At the top left is the Australian Government logo. To the right are navigation links: Home, About this Site, Site Map, Search, Links, Feedback. Below these are menu items: ABOUT US, THE COLLECTION, RECORDKEEPING (highlighted), EDUCATION, EXHIBITIONS, PUBLICATIONS. On the far right is a circular logo with the letter 'E' and the text 'e*permanence'.

Digital Preservation

The National Archives is developing a way of storing and preserving digital records of archival value so that they will be available in the future.

The Archives' digital preservation strategy

- **Conceptual approach** – the National Archives' approach to digital preservation is described in its 2002 paper, [An Approach to the Preservation of Digital Records](#).
- **Implementation** – the Archives has developed a formal workflow process for the preservation of digital records. See [How digital records are transferred to the Archives' digital repository](#) (HTML) or download the 3-minute [Digital Preservation – animation \(Flash 357kb\)](#). Requires Macromedia [Flash player](#).
- **Software applications** – the Archives has developed three software applications to enable records to be transferred to the digital repository. See [Digital preservation software applications](#).
- **XML schema** – an integral part of the Archives' preservation process is the conversion of records into stable, non-proprietary formats based on XML or Base64. See [XML data formats](#).

Resources for agencies

- **Best practice advice and guidelines** – the Archives has published a range of information to help Australian Government agencies make and keep digital records. See [Digital Records](#) (overview page) of this website.
- **Seminar papers** – In September 2004 the National Archives held a seminar titled *Advances in Digital Preservation: An International Working Meeting that brought together digital preservation experts and practitioners from around the world and from Australia*. See [Advances in Digital Preservation: introductory notes](#).

For more information about Digital Preservation, email recordkeeping@naa.gov.au.