

CURL Task Force on Digital Content Creation and Curation

Digital Resource Management Sub-Group

Digital Preservation Software – Questionnaire

Section 1: General Supplier and System Information	CONTENTdm	ENCompass	EPrints	CDSWare	FEDORA	MyCoRe	Greenstone
<i>Contact information</i>	<p>Name: DiMeMa Inc. Address: 100 W. Harrison, Suite N480, Seattle, WA 98119 Web address: http://www.contentdm.com Email address: info@contentdm.com Tel: 206-281-1632</p>	<p>Name: Endeavor Information Systems Address: 1350 E. Touhy Ave Suite 200 East Web address: http://www.endinfosys.com Contact name: Sara Randall, Director of Strategic Products Email address: sara.randall@endinfosys.com</p>	<p>Name: GNU EPrints Address: School of ECS University of Southampton SO171BJ Web address: http://www.ecs.soton.ac.uk Contact name: Christopher Gutteridge Email address: cjg@ecs.soton.ac.uk Tel: 02380 594833</p>	<p>Name: CERN Document Server Software Address: CERN 1211 Geneva 23 - Switzerland Web address: http://cdsware.cern.ch Contact name: Jean-Yves Le Meur Email address: Cds.support@cern.ch Tel: +41 22 767 47 45</p>	<p>Name: Fedora Address: University of Virginia, Charlottesville, VA 22904 Web address: http://www.fedora.info Contact name: Ronda A. Grizzle Email address: Rgrizzle@virginia.edu Tel: 434-924-3965</p>	<p>Name: MyCoRe Address: University of Duisburg-Essen, University Library, Universitätsstrasse 9-11, 45141 Essen, Germany Web address: http://www.mycore.de/engl/index.html and http://miless.uni-essen.de/ Contact name: Frank Lützenkirchen Email address: luetzenkirchen@bibl.uni-essen.de Tel: 0049 201 183 2124</p>	<p>Name: New Zealand Digital Library Project Address: Computer Science Department University of Waikato Private Bag 3105 Hamilton New Zealand Web address: http://www.greenstone.org Contact name: Ian Witten Email address: ihw@cs.waikato.ac.nz Tel: +64 7 838 4246</p>
<i>Summary of digital preservation software</i>	CONTENTdm is software designed and developed for complete management of digital collections of all formats. The product was based on	ENCompass for Digital Collections provides institutions with a system for locally creating, managing, and storing	GNU EPrints is software designed to allow an institution, or similar group, to create a web-based archive of their	CDSware allows you to run your own electronic preprint server, your own online library catalogue or a document system on the web. It complies	Fedora is a general purpose repository service developed jointly by The University of Virginia and Cornell University. The Fedora project is devoted to the goal of	MyCoRe is an Open Source project for the development of Digital Library and archive solutions (or, put more generally, "Content Repositories" >> CoRe). In the	Greenstone is a suite of software which has the ability to serve digital library collections and build new collections. It provides a new way of organizing information and making it available over the Internet or on

CURL Task Force on Digital Content Creation and Curation

Digital Resource Management Sub-Group

Digital Preservation Software – Questionnaire

	<p>collaborative input from cataloguers, archivists, and other information professionals.</p> <p>CONTENTdm supports each step of digital image collection development process: capture, index, store, administer, query, display, and share. It is designed for large archives of digital assets.</p> <p>CONTENTdm searches millions of objects and identifies target content in less than one second. The CONTENTdm software enables growth from tens to millions of items.</p> <p>This software tool enables cataloguers and archivists to use the expertise of their professional staff in building digital collections of their primary</p>	<p>collections of digital objects. It is an extremely flexible system, enabling customers to build multiple repositories of digital content and make them available to their end users. Repositories in ENCompass can store multiple file formats, and different repositories can conform to different metadata standards – Dublin Core, EAD, TEI-Lite, and so on.</p>	<p>research papers. It can be adapted purposes easily. It is currently used by more than 100 institutions.</p>	<p>with the Open Archives Initiative metadata harvesting protocol (OAI-PMH) and uses MARC 21 as its underlying bibliographic standard.</p> <p>The CDSware is free software, licensed under GNU General Public Licence (GPL).</p> <p>At CERN, the CDSware manages over 450 collections of data, consisting of over 700,000 bibliographic records, including 320,000 fulltext documents. (Covering preprints, articles, books, journals, photographs, videos and more).</p> <ul style="list-style-type: none"> • Configurable portal-like interfaces for 	<p>providing open-source repository software that can serve as the foundation for many types of information management systems. The software demonstrates how distributed digital information management can be deployed using web-based technologies, including XML and web services.</p>	<p>MyCoRe project a group of universities is working on the development of a shared software core for such applications. This core is adjustable to local requirements and easy to modify. This is expressed by the "My" in MyCoRe, which represents the local adaptability. On the basis of this core which is available under the open source GNU General Public License, specific local applications will emerge at the participating institutes. The technical base of the system is formed of Java class libraries, XML technology and implementations for various database backends, both commercial like IBM Content Manager and IBM DB2 or Open Source solutions like MySQL and XML:DB databases. The Core Functionalities of MyCoRe include the</p>	<p>CD-ROM. It is open-source software, available under the terms of the Gnu public license.</p> <p>A digital library is made up of a set of collections. Each collection of information comprises several (typically several thousand, or even several million) documents, which share a uniform searching and browsing interface. Collections can be organized in many different ways while retaining a strong family resemblance.</p> <p>Greenstone supports a variety of media formats. For textual documents it provides full text indexing of Word, PDF, HTML, email, PowerPoint amongst others, as well as automatic metadata extraction such as Title, Authors, Language etc. Metadata can also be assigned manually, which is the principle technique used in Greenstone for digesting multi-media</p>
--	---	--	--	---	--	--	--

CURL Task Force on Digital Content Creation and Curation

Digital Resource Management Sub-Group

Digital Preservation Software – Questionnaire

	<p>resource materials. CONTENTdm has proven results in building hundreds of valuable digital collections. CONTENTdm also enables libraries to provide leadership, share best practices and collaborate. Currently more than two hundred organizations use CONTENTdm. There are thousands of diverse collections and millions of items of rich content on-line with continuous monthly growth</p>			<p>hosting various kind of collections.</p> <ul style="list-style-type: none"> • Powerful search engine with Google-like syntax. • User personalization, including document baskets and email notification alerts. • Electronic submission and upload of various types of documents. • Running an OAI data and service provider enabling the metadata exchange between heterogeneous repositories. <p>We aim at user friendliness and speed</p>		<p>following:</p> <ul style="list-style-type: none"> • Flexible and configurable metadata model • Internal, hierarchical filesystem • Hierarchical Classification System • User and Rights Management • User and Author Interface with Online Editor • Distributed Search Function and an Interface for OAI 2.0 	<p>documents. Greenstone comes equipped for handling images such as TIFF, JPG, GIF, and audio files such as MP3, and Ogg</p>
--	---	--	--	---	--	---	---

CURL Task Force on Digital Content Creation and Curation

Digital Resource Management Sub-Group

Digital Preservation Software – Questionnaire

<p><i>Mailing lists</i></p>	<p>A free subscription to the CONTENTdm listserv is available to CONTENTdm customers.</p>	<p>Endigit-I is a mailing list that has been set up by ENCompass customers for ENCompass customers.</p>	<p>http://software.eprints.org/maillist.php</p>	<p>project-cdsware-announce@cern.ch [Archives] Low-volume read-only mailing list to announce new CDSware releases and other major news concerning the project. Every CDSware user is strongly encouraged to subscribe to this list. Moderated: yes</p> <p>project-cdsware-users@cern.ch [Archives] Possibly high-volume mailing list, intended for discussion among users and developers of CDSware. Moderated: no</p>	<p>Users of Fedora may join the fedora-users mailing list by going to: https://comm.nsd.org/mailman/listinfo/fedora-users</p>	<p>Send email to majordomo@uni-essen.de with the line: subscribe miles-community your@email.address in the mail message body</p>	<p>We have two mailing lists, one for Greenstone users, and one for Greenstone developers. Our main support page is http://www.greenstone.org/help, and both mailing lists can be joined by following links from this page. Archives of the mailing lists are also provided.</p>
<p><i>Information available on websites or professional journals</i></p>	<p>See www.contentdm.com/news/</p>	<p>No</p>	<p>http://www.aria.dne.ac.uk/issue37/nixon/</p>	<p>Yes, many, some are publicly available, others are more confidential. Contact:</p>	<p>All publications about Fedora are linked from the Fedora project home page. http://www.fedora.info/pubs.shtml</p>	<p>See "A Guide to Institutional Repository Software v 2.0" from "Budapest Open Access Initiative", at</p>	<p>We have published numerous research papers about or using the Greenstone software, for example:</p>

CURL Task Force on Digital Content Creation and Curation

Digital Resource Management Sub-Group

Digital Preservation Software – Questionnaire

				cds.support@cern.ch		http://www.soros.org/openaccess/software/	<p>Witten, I.H., Bainbridge, D., and Boddie, S.J. (2001) Greenstone: Open-source DL software. Communications of the ACM, vol 5, pp 47.</p> <p>Witten, I.H., Bainbridge, D., and Boddie, S.J. (2001) Greenstone: Open-source Digital Library Software with end-user collection building. Online Information Review, vol 25, no 5, pp 288-298.</p> <p>Apperley, M., Cunningham, S.J., Keegan, T., and Witten, I.H. (2001) Niupepa: A historical newspaper collection. Communications of the ACM, vol 44, no 5, pp 86-87.</p> <p>We have a publications page at http://www.cs.waikato.ac.nz/~nzdl/publications/. More Greenstone publications can be found at http://www.cs.waikato.ac.nz/pubs/</p>
--	--	--	--	--	--	---	--

CURL Task Force on Digital Content Creation and Curation

Digital Resource Management Sub-Group

Digital Preservation Software – Questionnaire

Section 2: Technical Information	CONTENTdm	ENCompass	EPrints	CDS Ware	FEDORA	MyCoRe	Greenstone
<p><i>What operating system(s) does the software run under?</i></p>	<p>Operating System 1: Windows Server NT 4.0, 2000, or 2003, IIS web server, XML database.</p> <p>Operating System 2: Linux, Apache web server, XML database.</p> <p>Operating System 3: Sun Solaris 8 or higher, Apache web server, XML database</p>	<p>Operating System 1: Sun Solaris, Apache web server, Oracle database.</p> <p>Operating System 2: Red Hat Linux, Apache web server, Oracle database.</p>	<p>Operating System 1: Linux, Apache web server 1.3 or 2, MY SQL 3 or 4.</p> <p>Operating System 2: Solaris, Apache web server 1.3 or 2, MySQL 3 or 4.</p>	<p>Operating System 1: UNIX-like (LINUX, SOLARIS, etc.)</p> <p>Apache web server, MySQL</p>	<p>Operating System 1: Unix, Apache Tomcat 4.1 web server, MySQL/McKoi/Oracle/others.</p> <p>Operating System 2: MacOSX, Apache Tomcat 4.1 web server, MySQL/McKoi/Oracle/others.</p> <p>Operating System 3: Windows, Apache Tomcat 4.1 web server, MySQL/McKoi/Oracle/others.</p>	<p>Operating System 1: Windows, Apache, Tomcat web server, MySQL/Lucene</p> <p>Operating System 2: Linux, Apache Tomcat webserver, MySQL or DB2 or Lucene.</p> <p>Operating System 3: IBM AIX, Apache, Tomcat or IBM WebSphere, IBM DB2, IBM DB2 Content Manager.</p> <p>Operating System 4: Sun Solaris, Apache, Tomcat, MySQL or DB2, Lucene.</p>	<p>Operating System 1: Windows 3.1/9x/2000/NT/XP</p> <p>Collections can be served on all windows OS, but collection building requires Windows 95 or above.</p> <p>Built-in web server, or can be used in conjunction with standard web servers, such as Apache or IIS.</p> <p>Operating System 2: Unix/Linux</p> <p>Can be used in conjunction with standard web servers such as Apache</p> <p>Operating System 3: Mac OS X</p> <p>We use two databases, one for indexing and text retrieval (MG/MGPP –</p>

CURL Task Force on Digital Content Creation and Curation

Digital Resource Management Sub-Group

Digital Preservation Software – Questionnaire

							Managing Gigabytes) and one for metadata storage (GDBM – GNU Database Manager). Both are open source projects in their own right. Because the software is open source, it can be compiled for any additional platform that provides Perl and C/C++.
<i>Please supply minimum hardware requirements</i>	For hardware requirements please see http://www.contentdm.com/products/system-req.html	Based on each customer's local needs, and which components of ENCompass they will be using, Endeavor works with each customer to determine hardware requirements that meet their local storage and usage needs.	Current entry level PC. 256Mb+ RAM. 1GHz+ Processor. Disk space as required for OS + archive (10Gb+). Provision for backups should always be made!	Dual Pentium III 1.4GHz, with 1 GB of memory.	No hardware requirements have been defined. Planned for late 2004.	Any hardware supporting current Java 2 Platform, Enterprise Edition (J2EE) 1.4	Full installation has been successfully used on Pentium 133 MHz, 100M RAM. The software requires about 80MB free disk space excluding any collection content.
<i>Please indicate skills required to run the digital preservation system.</i>	Skills for OS1: Windows System Administration for installation. Skills for OS2: Linux System Administration for installation. Skills for OS3: Solaris System	ENCompass does not require extensive programming knowledge. Library staff with an interest in web resources and knowledge of common web technologies as they relate to the academic library environment can	Basic UNIX admin (installing packages etc). Basic understanding of XML. Advanced configuration requires knowledge of	Some Unix administration skills required for installation. (Python, PHP) See http://cdsware.cern.ch/download/INSTALL	Unix system administrator for setup, Java.	Java, XML, XSL	Initially it requires expertise to set up the web server. Once Greenstone is installed, all interaction can be via the graphical interfaces. No programming is required to run the

CURL Task Force on Digital Content Creation and Curation

Digital Resource Management Sub-Group

Digital Preservation Software – Questionnaire

	<p>Administration for installation.</p> <p>HTML and JavaScript for Web site customization, HTML and JavaScript for Web site customization, HTML and JavaScript for Web site customization</p>	<p>easily undertake the administration of ENCompass.</p> <p>The following skill set is recommended for staff:</p> <ul style="list-style-type: none"> • Basic Microsoft Windows functionality knowledge • Knowledge of basic XML record formats • Knowledge of basic MARC record format • Web skills such as HTML, image manipulation and file naming practises • Knowledge of digital file formats such as pdf, doc, jpeg, gif, etc. <p>A library that wishes to take full advantage of the public interface customization features will want a staff member with more extensive XSL experience.</p> <p>The following skill set is recommended for system administrators:</p> <ul style="list-style-type: none"> • Basic Windows applications • XML and XML- 	<p>PERL.</p>				<p>standard system. Customizations to interface require skill in HTML syntax. Customized enhancements to functionality are possible by modifying Java, Perl, C/C++ code. In practice, such modifications are rare.</p>
--	---	---	--------------	--	--	--	--

CURL Task Force on Digital Content Creation and Curation

Digital Resource Management Sub-Group

Digital Preservation Software – Questionnaire

		<p>formatted data, including well-formed and valid DTDs</p> <ul style="list-style-type: none"> • EAD-structured XML • TEI Lite-structured XML • UNIX operating system commands and file system • Oracle database system • Apache web servers • LDAP servers, Voyager servers, or the type of server where your patron information is stored • A Unix-based text editor, such as edi or vi • A telnet program or SSH secure shell 					
<i>Does the system support OAI-PMH?</i>	Yes	ENCompass supports version 2 of the OAI-PMH as a data provider; using the ENCompass XML Loader and an OAI Harvester, customers can also harvest metadata records using OAI-PMH and load them into ENCompass	1.1 and 2.0 (1.1 deprecated)	V2.0	OAI-PMH 2.0	V2.0	We support ingesting OAI databases, and the next release will support collection serving using OAI.
<i>Does the system support Z39.50?</i>	Yes	ENCompass supports Z39.50 as a Z39.50 client. Endeavor is also planning on providing SRW server	No	Not yet. On going.	No	Planned but not yet supported.	Z39.50 client support

CURL Task Force on Digital Content Creation and Curation

Digital Resource Management Sub-Group

Digital Preservation Software – Questionnaire

		support in a future release of ENCompass.					
<i>Does the system conform to OAIS?</i>	No	No	No	No	Yes	Planned but currently not supported. For information about planned activities on OAIS with MyCoRe, you may contact joerg.lorenz@uni-muenster.de	No
<i>Is an open source licence available?</i>	No	No	GPL	GPL	Mozilla Public License	GPL	GPL

CURL Task Force on Digital Content Creation and Curation

Digital Resource Management Sub-Group

Digital Preservation Software – Questionnaire

Section 3: Preservation system information	CONTENTdm	ENCompass	EPrints	CDSWare	FEDORA	MyCoRe	Greenstone
<p><i>Please describe how digital objects are “ingested or imported” into the digital preservation system and discuss whether remote registered users can make submissions to the archive.</i></p>	<p>CONTENTdm simplifies the process of creating and organizing collections. Each user has personal project space within the CONTENTdm Acquisition Station on their desktop. When image acquisition and metadata entry are complete within a user's project, the items are uploaded to a centralized pending queue for final review or editing by the collection administrator before being added to the collection. CONTENTdm supports distributed collection building. The Acquisition Station can be installed on up to 50 different workstations supporting remote users in diverse locations.</p>	<p>ENCompass uses its XML Loader to enable staff to batch load digital objects into an ENCompass repository. Customers can also add digital objects manually (that is, one at a time) through the Staff Client, or through the Web Forms feature. Using the Web Forms feature, remote authorized users can submit digital objects to the system, which in turn can be reviewed in the ENCompass Staff Client before being published. In each case outlined above, the digital object itself – that is, the file itself, as opposed to the metadata describing it – can be stored in ENCompass, or outside ENCompass and referenced by the digital object metadata stored in ENCompass. When a digital object is accessed in the ENCompass Web</p>	<p>Files and metadata are normally entered via the web interface by registered users. There is an optional approval stage before these items appear in the main archive. A scripting interface is also available for bulk imports (or other special features).</p>	<p>Ingestion:</p> <ul style="list-style-type: none"> • From OAI repositories with module BibHarvest • From non OAI sources with module BibConvert • From registered user submission: • Direct submission to the archive • Submission with monitoring (someone is alerted by email for checking) • Submission with approval (doc available only once it is approved) • Submission to protected 	<p>There is a batch creation process that is used to create large batches of objects. Single submission interfaces have been developed by some implementers. Further development for batch and single submission is planned for the next phase of the project.</p>	<p>Command line interface for batch import of digital objects, used by administrators. Metadata is encoded in XML files, content can be read from directories and files</p> <p>Web-based GUI (HTML forms and Java Applet) for remote registered users, users can upload single files with HTML form or multiple files and directories using Java Applet.</p>	<p>Operationally, documents are added into the import directory of a collection, then converted into a standard XML form (Greenstone proprietary or METS) usable by the indexing software and other external applications if desired. The import process is not automatic by default, however it could be set up as a scheduled task.</p> <p>The most common way users ingest documents is to use the graphical collection management tool, drag and drop new documents into the collection, and press rebuild.</p> <p>This instigates the operational behaviour</p>

CURL Task Force on Digital Content Creation and Curation

Digital Resource Management Sub-Group

Digital Preservation Software – Questionnaire

	<p>Inputting Data and Media: Items can be added to CONTENTdm singly or in batches. Items that are not yet digitized can be scanned directly into CONTENTdm from any TWAIN device. CONTENTdm's batch processes and tools streamline your work, making it easy to add and manage thousands of digital items at once. Create metadata templates to speed and standardize the entry of descriptive and administrative metadata. Apply image settings instantly to items as they are imported into your projects or make global changes to metadata within live collections. CONTENTdm also provides a range of options for working with high resolution JPEG or TIFF files, from automatically creating lower resolution display copies that enable</p>	<p>client, ENCompass retrieves the file from the location specified in the metadata record – whether that's ENCompass or an outside repository of digital objects.</p>		<p>collection plus approval (doc appears in a confidential collection and will appear in public collection only once it is approved).</p>			<p>described above. The graphical tool has been implemented as an applet, allowing remote users to create collections on a central server, though not as yet to work with a central archive.</p>
--	--	--	--	---	--	--	--

CURL Task Force on Digital Content Creation and Curation

Digital Resource Management Sub-Group

Digital Preservation Software – Questionnaire

	<p>modifications to image quality and legibility, to using custom display copies you create. Tab-delimited text import enables easy import of metadata into CONTENTdm from existing tools such as Access, Excel, and FileMaker Pro. Users can also batch import lengthy text files from OCR or transcripts into metadata fields to enable full-text searching of documents. Other batch processes allow users to make global changes to collections or the database.</p> <p>Complex Media Types: CONTENTdm handles audio, video, PDF, image files and more. CONTENTdm also enables users to define items that consist of multiple views so search results will retrieve entire entities rather than individual elements. Add documents and</p>						
--	---	--	--	--	--	--	--

CURL Task Force on Digital Content Creation and Curation

Digital Resource Management Sub-Group

Digital Preservation Software – Questionnaire

	<p>monographs, or multi-sided objects such as post cards or sculptures.</p> <p>For more information on CONTENTdm see www.contentdm.com/products/.</p>						
<p><i>Does the system automatically recognise file types? i.e. support an approved file format function</i></p>	<p>Yes. CONTENTdm can handle any file type that is supported by your browser natively or with a plug-in. There are specific tools for working with JPEG2000 (does not require browser plug-in), TIFF, JPEG, GIF, audio, video, and documents.</p>	<p>Yes. As digital objects are loaded, ENCompass looks at file extensions and compares them to standard MIME types.</p>	<p>Yes. The archive admin may specify a list of accepted types, although the software does not currently check the formats are as specified (this is currently considered to be a human not computer task as it could reject things in annoying ways). An auto format checking feature could be added very easily.</p>	<p>Yes. This is optional. Accepted files format can be defined when configuring the submission process. For pictures, a function allows a preview before the end of the submission.</p>	<p>Yes.</p>	<p>Yes, list of supported file types in configurable and includes PDF, PS, HTML, XML, Word, MPEG, MP3, Real, AVI, ...</p>	<p>Greenstone recognises and processes many file types (based on filename extension) using a series of document recognizer "plugins". Additional plugins can be written (in Perl) to accommodate currently unsupported file types.</p>
<p><i>Can multiple file types be ingested as part of a single deposit?</i></p>	<p>Yes. The CONTENTdm Acquisition Station supports batch loading of multiple files.</p>	<p>Yes. ENCompass doesn't care if files with different file types are ingested at the same time, into the same repository.</p>	<p>Yes. Each deposit has metadata plus "documents" each document has a format (eg. HTML) and a set of files (eg.</p>	<p>Yes. Users can load as many files as they want to attach to a single deposit. Files can also be automatically converted to PDF</p>	<p>Yes.</p>	<p>Yes, a complete directory on local filesystem can be imported, the structure of files and</p>	<p>Yes Each file is treated independently and will be processed by the appropriate plugin.</p>

CURL Task Force on Digital Content Creation and Curation

Digital Resource Management Sub-Group

Digital Preservation Software – Questionnaire

			index.html, fig1.jpg).	using the CDS Conversion server plug-in.		subdirectories is reflected in the repository	
<i>Please list the metadata formats your software supports (e.g. METS, Dublin Core etc.)</i>	Metadata formats can be defined and customized by the collection administrator. Default metadata structures are Dublin Core and VRA. Metadata can be mapped to MARC through OCLC WorldCat.	Out of the box, ENCompass has support for Dublin Core, Qualified Dublin Core, EAD, and TEI-Lite built-in. With some additional set up work, ENCompass can support any DTD-defined metadata format.	Default metadata format may be mapped into dublin core (function provided). Metadata may be locally customised. New mapping functions may be added.	Internal representation: MARC21. OAI Exposure: XML MARC – XML DUBLIN CORE.	FedoraMETS/FOXML, Dublin Core	Qualified Dublin Core as example metadata format, other formats are configurable	METS, Dublin Core, MARC, Refer. Plugins and XML Schemas can be used to extend the metadata sets/formats handled by Greenstone as needed.
<i>Does the software support preservation metadata? If so, please provide further information on this feature.</i>	Yes. CONTENTdm supports JPEG2000 conversion which includes preservation data. Also, preservation data from TIFF and other full-resolution files can be mapped to metadata fields in CONTENTdm.	ENCompass enables customers to use any DTD-defined metadata format, so any DTD-defined metadata format that covers preservation metadata can be used in ENCompass.	Not out of the box. This could be added if desired.	Preservation: <ul style="list-style-type: none"> MARC21 = Library of Congress Standard = aims at long term keeping of metadata. 	Yes. Via FedoraMets/FOXML.	Currently not following any standards (OAIS planned), but MD5 checksum, file size, file format is kept for each file. Versioning is planned.	Not explicitly.
<i>Does the software support persistent naming / identifies. If so, please provide further information on this feature</i>	Yes. CONTENTdm creates a unique URL for each item in the database. In addition, collections can be registered with OCLC WorldCat for global persistence and uniqueness.	Currently, ENCompass does not support persistent identifiers. However, Endeavor plans to support persistent identifiers in a future release of the product.	Yes. Every deposit is assigned an ID number and a URL. If a revised version is submitted then the page for the original item will include a link to	CDSware Record Ids never die. If record is deleted, it becomes invisible but administrators can reactivate it later (very important for URLs lifetime: for	Yes. Each object in a Fedora repository has a persistent identifier that remains with the object throughout its "life."	Each object has a unique ID that can be used to directly reference the object in a request URL	Yes. Greenstone uses internal document identifiers. Providing the document doesn't change, these will remain the same across multiple collection builds as

CURL Task Force on Digital Content Creation and Curation

Digital Resource Management Sub-Group

Digital Preservation Software – Questionnaire

			the newest version.	e.g., if links where created to the record, it will not be dead but inform about the deletion of the record).			they are based on a hash of the document content.
--	--	--	---------------------	---	--	--	---

CURL Task Force on Digital Content Creation and Curation

Digital Resource Management Sub-Group

Digital Preservation Software – Questionnaire

Section 4: User interface and search functionality	CONTENTdm	ENCompass	EPrints	CDSWare	FEDORA	MyCoRe	Greenstone
<i>Please describe to what extent the user interface can be customised, whether multiple languages are supported etc</i>	<p>The end-user interface is fully customizable. All search, advanced search, results, and viewing (image, document, newspapers, etc.) pages are Web templates that can be customized. These templates can be localized for multiple languages. Localized versions for French and Spanish are available.</p> <p>To see examples of customized collections, visit: www.contentdm.com/customers/.</p>	<p>The user interface in ENCompass is extremely flexible and can be customized. Using XSL and XSLT, the interface can be modified to meet local needs, including the ability to display interface labels, as well as confirmation and error messaging, in languages other than English.</p>	<p>Multiple languages are supported, translations exist into: French, German, Portugese, Spanish, Swedish, Italian, Russian and more.</p> <p>All text which appears to users is stored as XML for full unicode support.</p> <p>All "phrases" in the interface can be customised, the new search pages may be added and new 'views' on the data added. Workflow may be customised, but not fully.</p>	<p>All Search and Submit interfaces can be customized (at CERN, more than 450 different collections are supported: http://cdsweb.cern.ch)</p> <p>At this day, languages supported include German, English, French, Italian, Portuguese, Swedish etc.</p>	<p>The user interface is completely customisable and supports application of custom headers/footers and page design and multiple language interfaces.</p>	<p>Multiple languages are supported. Interface can be fully customised using XSL stylesheets and CSS.</p>	<p>Web based user interface, can be customised by editing "macro" files (combines HTML with style parameters). Collection specific customisations are supported via configuration files. Multiple languages are supported (over 30) and new ones can be added by translating the set of text strings for the interface.</p>
<i>Please describe how resources</i>	CONTENTdm supports browse and search	ENCompass uses FAST Data Search	Full text searching is	Search	Selected metadata fields	Search in metadata is	Full text searching, metadata

CURL Task Force on Digital Content Creation and Curation

Digital Resource Management Sub-Group

Digital Preservation Software – Questionnaire

<p><i>ingested into the archive/repository can be searched. For example, is full-text searching supported, or is searching restricted to the metadata? Can the archive be browsed, and if so, how? Can search results be sorted?</i></p>	<p>capability for all collections. Full text searching and cross-collection searching are also supported. CONTENTdm also supports the creation of custom queries. Search results can be sorted by metadata field. For more information on search and browse features, see www.contentdm.com/products/features.html. Visit www.contentdm.com/customers/ to see customer collections.</p>	<p>for indexing and searching metadata stored in ENCompass. Currently, ENCompass supports searching across the metadata stored in ENCompass digital repositories. In addition to searching through digital objects stored in ENCompass repositories, users can browse through the digital objects available in individual repositories. ENCompass supports multiple sorting options, including sorting by creator, title, and date.</p>	<p>supported, with sorting. Basic and Advanced search configured by default, but other searches may be added. Ordering properties of fields can be customised by administrator (with a little PERL), eg. making titles ignore 'The ' at the start of a title when sorting by title.</p>	<p>Options:</p> <ul style="list-style-type: none"> • Index browsing • All metadata altogether • Any metadata field • Full text • References (can be automatically extracted with a plug-in) • Any combination of the above (and, or, and not) • With Sorting according to one or many fields • With User Guidance in case no results are found • With multiple possible output 	<p>in Fedora objects are indexed for search. The archive can be browsed by author, title, issue date, subject term, and collection. Additionally, Fedora objects may be independently indexed to allow users richer search capabilities.</p>	<p>supported, different search masks can be freely configured</p> <p>Full text search in content of ASCII and PDF files supported, other file types supported if IBM Content Manager is used as backend database</p> <p>Browsing in hierarchical classification systems, e.g. PACS, MSC, DDC is supported.</p> <p>Search results can be sorted by any metadata field</p>	<p>searching, and field based searching are all supported. Search results can be unsorted or sorted by rank. Browsing via metadata is also supported, for example via alphabetical lists, or subject hierarchies. Phrase browsing via phrases extracted from the documents is also supported. The features that a particular collection provides are determined by the collection designer and are highly customisable.</p>
--	---	---	---	--	--	--	---

CURL Task Force on Digital Content Creation and Curation

Digital Resource Management Sub-Group

Digital Preservation Software – Questionnaire

				formats (HTML brief, HTML detailed, HTML customized, XML DC, XML Marc, etc)			
<i>Final additional comments</i>	<p>CONTENTdm Digital Collection Management Software is the result of years of development and collaboration with archivists. A flexible, multifunction software package, CONTENTdm provides tools for all aspects of digital collection management. CONTENTdm is a single software solution that consists of three main elements:</p> <ul style="list-style-type: none"> • The Acquisition Station where data and images are imported or added • The CONTENTdm Server where data and images are stored • The customizable Web-based search client <p>CONTENTdm enables collection owners to organize and manage their</p>		The software is very adaptable. Every institution has their own requirements and rather than try and enforce what we see as the Right Answer, we have created very flexible and configurable software with a sensible initial configuration.	Try it!	Fedora is a foundation for many systems rather than a complete application for any one domain. Though Fedora provides some basic search and workflow services, it is assumed that domain specific solutions will be developed on top of the Fedora system. In it's next phase, beginning in October of 2004, the project will be developing more elaborate infrastructure for the implementation	MyCoRe is currently available as a version 0.9 that is already very useful for many application scenarios, but some functions are still under development. MyCoRe is especially interesting if you need a configurable metadata model, support for multi-media file types and integrated audio-/video-streaming server support (Helix universal server, real	Greenstone is widely used by various UN agencies and other non-government organisations.

CURL Task Force on Digital Content Creation and Curation

Digital Resource Management Sub-Group

Digital Preservation Software – Questionnaire

	<p>digital assets and provides easy Web access to collections. Optimized for speedy delivery of exceptionally large datasets, CONTENTdm is scalable for any size collection, whether you're managing 500 or 5,000,000 items. CONTENTdm also provides full search support for Latin-1 foreign character sets so that collection owners can enter, store, display, and search international characters in metadata fields. The software is Z39.50 compatible, and provides unparalleled visibility for digital collections by offering support for OAI harvesting and the ability for libraries to automatically add collections to WorldCat, the world's largest bibliographic database. CONTENTdm provides tools for each stage of the collection building process and when your collection is ready for viewing, CONTENTdm offers a flexible and customizable Web-based search interface.</p> <p>How CONTENTdm Works: Collections reside on a CONTENTdm Server. Media items can be added and</p>				<p>of such services.</p>	<p>server, IBM VideoCharger server) to stream MP3, Real, AVI, MPEG, WAV SMIL etc.</p> <p>The predecessor system is MILESS developed at University of Duisburg-Essen, see http://miless.uni-essen.de/</p> <p>This system is currently used by other german universities too, for example see http://miami.uni-muenster.de/ and http://www.db-thueringen.de/</p>	
--	--	--	--	--	--------------------------	---	--

CURL Task Force on Digital Content Creation and Curation

Digital Resource Management Sub-Group

Digital Preservation Software – Questionnaire

	<p>managed from anywhere using our Acquisition Station software, which can be installed on as many as 50 different workstations for collaborative work or collections can be managed remotely over the Web. The database is searched over the Web using standard Web browsers by any number of users unless you choose to restrict access. Visit www.contentdm.com to learn more, or sign up for a <u>free 60-day evaluation</u> of CONTENTdm.</p>						
--	---	--	--	--	--	--	--