U.S. Government Printing Office
Federal Digital System
System Design Document

Volume XLI:

Data Management Definition (DMD)

Education Reports from ERIC

R1C2 Edition

Prepared by: FDsys Program

**Office of the Chief Information Officer
U.S. Government Printing Office**

April 16, 2010

Revision History

Revision	Date	Description
0.1	May 10, 2009	Initial Version
0.2	June 1, 2009	Initial Draft
0.3	June 15, 2009	Search Technologies Technical Review
0.4	November 5, 2009	Updates based on PDF to Text conversion
0.5	12/21/2009	Search Technologies Architect Review
0.6	3/31/2010	PMO Review
0.7	4/5/2010	Implement PMO comments, add PCS/isFallbackTitle
0.8	4/6/2010	Implement Peer Review Comments
0.9	4/15/10	Changed value for PM/Publisher constant
0.10	4/16/10	Incorporated final Peer Review Comments

Responsibilities

Description	Responsible Party
Co-Owners:	Program Management: Lisa LaPlant Technical: Ronald Matamoros

Documentation Conventions

1. Strings with embedded values are indicated with curly-braces, for example:

   Compilation of Presidential Documents Volume {PCS/volume},
   Issue {PCS/issue}, {PM/dateIssued}

2. References to XML entities and attributes are referenced using the XPath standard. However, to save room, common prefixes may be abbreviated (see next section).

Abbreviations

This document uses the following abbreviations for specifying elements from fdsys.xml:

PH   Package header
       fdsysPackage/packageHdr/

PM   Package metadata
       fdsysPackage/packageHdr/descPkgMd/

PCS   Package collection specific metadata
       fdsysPackage/packageHdr/descPkgMd/collectionSpecific/

GM   Group metadata (generic granule metadata)
       fdsysPackage/mdSect/descMdGroups/descMdGroup

GCS Group collection specific metadata (collection specific granule metadata)
       fdsysPackage/mdSect/descMdGroups/descMdGroup/collectionSpecific/

Table of Contents

1. Introduction

1.1. General Description

The ERIC Collection contains records to education related materials and journals such as:

- Books
- Research syntheses
- Conference papers
- Technical reports
- Policy papers

The data is used primarily by people interested in education policy, instructors and students in teaching programs. The data comes from a myriad of different sources and prefaced by an abstract and metadata on the first page of the PDF file. The GPO collection is static with data going from 1995 to 2004.

GPO Access contains reports on federally funded education research topics from the U.S. Department of Education's Educational Resources Information Center. Reports on *GPO Access* begin with those received in October 2002. Prior reports are available from select Federal depository libraries nationwide in microfiche. A larger selection of ERIC reports is available from the ERIC program. Files are available in Adobe Portable Document Format (PDF) only.

PM/collectionCode	Display Name
ERIC	Education Reports from ERIC

1.2. Document Types

There is only a single type of file for ERIC, the PDF file that contains the educational resource.

PCS/docClass	User-readable (full)	User-readable (abbreviated)	Description
ERIC	Education Resources Information Center	ERIC	ERIC indexes education journals, the majority of which are peer-reviewed. Most of these journals are indexed comprehensively - that is, a record for every article in each issue is included in ERIC. Some journals are indexed selectively - that is, only those articles that are education-related are included.

2. fdsys.xml Schema Elements

2.1. Package-level metadata

Notes:

- See above for XML entity abbreviations (PH, PM, PCS, GM, etc.)
- Items which are <blank> do not need to be included in the fdsys.xml file.
- All of the data values below are assumed to be strings unless specified otherwise.

XML Entity	Description or "constant value"	Source	Arity
Generic Metadata Fields			
PM/collectionCode	"ERIC"	constant	1
PM/quality	Filled out by parser (see section 4.3.1)	parser	0-1
PM/scope	"fdlp"	constant	1
PM/governmentAuthor1	"Department of Education"	constant	1
PM/governmentAuthor2	"Education Resources Information Center"	constant	1
PM/governmentAuthor3	<blank>	n/a	0
PM/starprintNumber	<blank>	n/a	0
PM/category	"Executive Agency Publications "	constant	1
PM/title	The title of the report as parsed from the document	parser	1
PM/title/@info	"from-parsing"	parser	1
PM/sourceContentType	"deposited"	constant	1
PM/ packageDigitalOrigin	"born digital"	constant	1
PM/personalAuthor	<blank> *Note: Personal authors are stored in PCS/personalAuthor so that no changes ar required to the FDsys generic schema or mods-common.xsl.*	n/a	0
PM/branch	"executive"	constant	1
PM/typeOfResource	"text"	constant	1
PM/genre	<genre authority="marcgt">government publication<genre>	constant	1
PM/geographicLocation	<blank>	n/a	0
PM/dateIssued	The date of the document as parsed from its summary metadata.	parser	1
PM/dateCreated	<blank>	n/a	0
PM/dateCopyrighted	<blank>	n/a	0
PM/dateValid	<blank>	n/a	0

PM/dateModified	<blank>	n/a	0
PM/dateIngested	Date ingested into FDsys	submission	1
PM/edition	<blank>	n/a	0
PM/issuance	"monographic"	constant	1
PM/language	"eng"	constant	1
PM/abstract	The abstract parsed from the document metadata.	parser	0-1
PM/tableOfContents	<blank>	n/a	0
PM/topic	<blank>	n/a	0
PM/geographicSubject	<blank>	n/a	0
PM/temporalSubject	<blank>	n/a	0
PM/classification	With authority="sudocs": "ED 1.615:"	constant	1
PM/otherIdentifier [@idStandard= 'ils-system-id']	000573142	n/a	1
PM/otherIdentifier [@idStandard= 'migrated-doc-id']	<blank>	n/a	0
PM/otherIdentifier [@idStandard= 'stock-number']	<blank>	n/a	0
PM/otherIdentifier/ [@idStandard= 'sudoc-item-number']	<blank>	n/a	0
PM/otherIdentifier [@idStandard='isbn']	From document metadata.	parser	0-1
PM/otherIdentifier [@idStandard='issn']	<blank>	n/a	0
PM/part	<blank>	n/a	0
PM/waisDatabaseName	<blank>	n/a	0
PM/notes	<blank>	n/a	0
PM/recordSource	"DGPO"	constant	1
PM/recordCreationDate	Date that the mods.xml is created	processing	1
PM/recordChangeDate	Date that the mods.xml is modified	processing	1
PM/recordOrigin	"machine generated"	constant	1
PM/subTitle	<blank>	n/a	0
PM/creator	<blank>	n/a	0
PM/publisher	U.S. Department of Education	constant	0

PM/pageCount	Computed based on the total page count of the pages in the PDF rendition.	processing	1
PM/frequency	<blank>	n/a	0
Collection-Specific Metadata Fields			
PCS/docClass	"ERIC"	constant	1
PCS/accessId	The access identifier for this package, used to uniquely identify this granule to the public. Example "ERIC-ED463724"	parser	1
PCS/type	The type of education resource. Example, "Creative Works"	parser	1-n
PCS/institution	The "INSTITUTION" is the originating institution for of the resource. Example "Arizona Univ., Tucson. Coll. of Education."	parser	0-1
PCS/sponsorAgency	The sponsoring agency of the education resource. Example "Special Education Programs (ED/OSERS), Washington, DC."	parser	0-n
PCS/subject	The "DESCRIPTORS" attribute found in the metadata section of the PDF file	parser	0-n
PCS/identifiers	The "IDENTIFIERS" attribute found in the metadata section of the PDF file	parser	0-n
PCS/ericNumber	[KEY]The ERIC # for the report. This the file name without the extension. Example "ED464427"	parser	1
PCS/personalAuthor	The author(s) of the ERIC resource as parsed from the document metadata. For example, "Kaput, James J."	parser	0-n
PCS/isFallbackTitle	Will be "true" if a fallback title was used for this package, or "false" if it is a normal, descriptive title.	parser	1

2.2. Standardized References

The ERIC DMD uses no standard references.

3. Renditions and Input Files

3.1. Renditions

Format Name	Mime Type	Classification	From	isPublic	granules	Description
pdf-submitted	application/ pdf	production	Plant	no	no	The original PDF files
pdf	application/ pdf	web-optimized	pdf-submitted	yes	no	The PDF rendition, for public consumption. This is separated from "pdf-submitted" since the files will be renamed and to allow for future digital signing.
text	text/plain	production, derived	pdf-submitted	no	no	Produced from the submitted PDF by processing. NOTE: This is an ACP derived rendition. It is not stored in the AIP.

Notes:
1. The format name will be used for the rendition folder name, which will in turn be used as a component of the URL for access.
2. If isPublic, then the rendition will also have classification="public-access"

3.2. Plant Processing

This is a static collection. There will be no further GPO Plant processing required.

3.3. Migration Input & Packaging

Detailed packaging rules will be computed as each directory of files is migrated from GPO Access to FDsys. The following is a rough guide for how to determine, for each file name, how to determine the package ID to which it will be assigned.

Rendition	Sample File Names	Instructions for determining package ID
pdf-submitted	ed464338.pdf	Remove extension from file name.

text	ed464338.txt	This rendition is generated from the pdf-submitted rendition.
		It should be generated following the output format of the FAST pdftotext tool
		The name of the file is the same as the PDF rendition and replacing the extension with "txt".
		For example, "ed464338.pdf" yields "ed464338.txt"
		The file is set on the "text" rendition folder of the package.

3.4. Day Forward Input

This is a static collection. There will be no day-forward processing required.

4. Parsing

4.1. Renditions

This collection will be parsed using the text rendition produced from the submitted PDF.

The DMD assumes the text output resembles the Adobe "Saves as Text" format.

Parsing notes: (applicable to all renditions)

1. Dates should be converted to YYYY-MM-DD format.
   a. Only dates should be stored. There are no date-time formats.

4.2. Parsing Text

All the metadata for the document is found on the first page.

4.2.1.   Preprocessing

Reduce the parsing scope to the first page by excluding all other pages.

The document will have a top header separating the pages. The second header instance in the document marks the end of the first page.

Below is an example of the header pattern to match:

```
7 organizations that offer help to families, caregivers, and teachers and
an

annotated list of 11 print resources.) (CR)

Reproductions supplied by EDRS are the best that can be made
from the original document.


This is the background image for an Adobe Acrobat Capture OCR page with
image plus hidden text.
Perspectiva General sobre la Sordo-Ceguera. DB-LINK.
Diciembre de 1995
Por
Barbara Miles
```

4.2.2.   Granules

There are no granules associated with this collection.

4.2.3.   Parsing metadata

All the metadata about the content is found on the first page of the document an will be parsed in a similar fashion. The few exceptions will be noted in subsequent sections.

Below is a sample of the metadata from the text rendition.

```
                              DOCUMENT RESUME



AUTHOR          Kleiner, Anne; Lewis, Laurie
TITLE           Internet Access in U.S. Public Schools and Classrooms: 1994-
                2002. E.D. Tabs.
INSTITUTION     National Center for Education Statistics (ED), Washington,
                DC.; Westat, Inc., Rockville, MD.
REPORT NO       NCES-2004-011
PUB DATE        2004-00-00
NOTE            85p.; Project Officer, Bernard Greene. For the 1994-200'1
                edition, see ED 472 678.
AVAILABLE FROM  For full text: http://nces.ed.gov/pubs2004/2004011.pdf.
PUB TYPE        Numerical/Quantitative Data (110) -- Reports - Research (143)
                -- Tests/Questionnaires (160)                    . .
EDRS PRICE      EDRS Price MFOl/PC04 Plus Postage.
DESCRIPTORS     *Classroom Environment; Educational Equipment; Information
                Dissemination; Information Technology; *Internet; Public
                Education; *Public Schools


ABSTRACT
                This report presents data on Internet access in U.S. public
schools from 1994 to 2002 by school characteristics. It provides trend
analysis on the progress of public schools 'and classrooms in connecting to
the Internet and on the ratio of students to instructional computers with
Internet access. For the year 2002, this report also presents data on the
types of Internet connections used; student access to the Internet outside of
regular school hours; laptop computer loans; hand-held computers for students
and teachers; and school Web sites. It also contains information on computer
hardware, software, and Internet support and Web site support at the school;
teacher professional development on how to integrate the use of the Intern'et
into the curriculum; and technologies and procedures to prevent student
access to inappropriate material on the Internet. Appended are the
Methodology and Technical Notes; and Questionnaire. (Contains 43 tables and 4
figures.) (Author)




Reproductions supplied by EDRS are the best that can be made
                        from the original document.
```

All documents start with the text "DOCUMENT RESUME". But the first metadata does not appear until the line starting with "AUTHOR". The author metadata element is typically the first element to be found, but "TITLE" as also shown up where there is no author element.   All metadata element names are all in capital letters. The other important factor to notice is that values for the metadata elements start always from a particular character spacing in document. It the previous example the metadata values start at character 19. Elements with multiple values have the value delineated by a ";" (semi-colon).

Common actions:

- Remove any carriage returns or end of line codes.
- Do not include the last period as part of the value.
- Set the PM/title metadata @info attribute to "from-parsing".
- Remove any * from the values
- Remove any ; from the values
- Remove the extra "ISBN" from the ISBN values only.

Table of mappings from element name in the text to specific metadata values:

Element from text rendition	Metadata Name
TITLE	PM/title
AUTHOR	PCS/personalAuthor
PUB DATE	PM/dateIssued
ISBN	PM/otherIdentifier[@idStandard='isbn']
PUB TYPE	PCS/type
ABSTRACT	PM/abstract
INSTITUTION	PCS/institution
SPONS AGENCY	PCS/sponsorAgency
DESCRIPTORS	PCS/subject
IDENTIFIERS	PCS/identifiers

4.2.4.   PM/title and PCS/isFallbackTitle

If no title can be parsed from the document, set PM/title to the following fallback title:

"Education Report {PCS/ericNumber}"

Where {PCS/ericNumber} is formatted as follows:

Format ERIC number as {alphaprefix}-{ddd}-{ddd}. Where "ddd" segments are from numeric suffix.

For example, ED464761 is formatted as "ED 464 761".

If the fallback title is used, be sure to set PCS/isFallbackTitle to "true". Otherwise, set it to "false"

4.2.5.   PM/dateIssued

The PM/dateIssued for a package is found on the line for the "PUB DATE"

For example:

```
SPONS AGENCY       Special Education Programs (ED/OSERS), Washington, DC.
PUB DATE           1995-12-00
NOTE               1op.; For the English version, see ED 436 056.
```

Here the PM/dateIssued is "1995-12-01".

Note:

- The format is interpreted as YYYY-MM-DD
- The "PUB DATE" can set the month and date value to "00".
- When the month or date is set as "00", default the value to "01".

4.2.6.   PM/abstract

The PM/abstract data can be found at the end of the metadata section. It is preceded by the uppercase "ABSTRACT" markup. It is terminated by a pattern "Reproductions supplied by EDRS are the best that can be made" It is important to note that the value for PM/abstract does not start on the same line as the "ABSTRACT" element. It starts on the next line and frequently does not use the same character starting position.

For example:

```
DESCRIPTORS        *Federal Aid; Higher Education


ABSTRACT

                This guide explains student financial aid programs the U.S.
Department of Education's Federal Student Aid (FSA) office administers. The
first three pages are a quick reference; the rest of the publication provides
more of what you need to know about the financial aid programs offered. (AMT)




                Reproductions supplied by EDRS are the best that can be made
```

Please note the abstract should be truncated to the "." prior to the terminating pattern. As with other metadata elements remove any carriage returns in the metadata, but not the trailing period.

4.2.7. PCS/accessId

The formula for deriving the access identifier:

> {PM/collectionCode}-{PCS/ericNumber}

Example:

> ERIC-ED463948

4.2.8. PCS/type

The PCS/type is found after the "PUB TYPE" tag on the metadata page in the document. There can more than one identifier per document and they should all be indexed. Unlike other metadata values, the values for this element are separated by "--".

For Example:

```
AVAILABLE FROM    For full text: http://nces.ed.gov/pubs2004/2004011.pdf.

PUB TYPE          Numerical/Quantitative Data (110) -- Reports - Research (143)

                  -- Tests/Questionnaires (160)                    . .

EDRS PRICE        EDRS Price MFOl/PC04 Plus Postage.
```

The PCS/type values for this example are "Numerical/Quantitative Data", "Reports – Research", "Tests/Questionnaires"

- As with other metadata elements remove any carriage returns in the metadata.

- Additionally remove any text in parentheses encountered and the "--" characters.

- Frequently values have dash in the data such as " Reports -Evaluative (142)". The dash should have a single space on either side, so the value would appear as "Reports - Evaluative".

- Set each value into a separate PCS/type metadata field.

- Note: If unable to parse out a PCS/type set it to "Other"

4.2.9.   PCS/ericNumber

The PCS/ericNumber is the file name of the document, dropping the extension.

Uppercase the filename before setting the value to PCS/ericNumber.

For Example:

```
Directory of \data\2000\pdf-submitted


05/05/2009  01:40 PM    <DIR>          .
05/05/2009  01:40 PM    <DIR>          ..
04/02/2009  02:55 PM         4,604,128 ed463510.pdf
04/02/2009  02:55 PM         5,670,425 ed463572.pdf
04/02/2009  02:55 PM         4,020,631 ed463617.pdf
04/02/2009  02:55 PM         1,501,243 ed463618.pdf
.
.
.
```

The files highlighted would have the PCS/ericNumber of ED463510, ED463572, ED463617, ED463618 respectively.

4.3. Validation Heuristics

The following validation heuristics should be checked by the parser and added to the element in the fdsys.xml value as a "quality=" attribute or the <quality> element.

4.3.1.   Validating Granules and Packages

There are two quality elements for flagging quality issues encountered when parsing packages and granules.

    PM/quality   (and PM/quality/@quality)

    GM/quality   (and GM/quality/@quality)

In general, quality/@quality should specify either "error", "low", or "medium". The text inside the <quality> element should provide some descriptive text as to why the quality was tagged as such.

Situation	metadata	@quality value	Descriptive Text
Unrecognized file name format	PM/quality	error	"Unrecognized file name format"
Incorrect format	PM/quality	error	"Text file format appears to contain locator data"

5. FDsys Processing

This section describes in detail all of the steps required to process the collection files inside of FDsys. This includes creating renditions, creating the table of contents, mapping metadata, expected manual edits, etc.

5.1. Special Manual Interventions Required

There are no special manual interventions required for this collection.

5.2. Text Creation

The PDF files will need to be converted to text using the FAST "pdftotext" tool.

Text files will:

- Be stored in the "text" rendition
- Have the same file name as the PDF file., except...
- Will have a "txt" extension instead a "pdf"

For example "ed453610.pdf" will be run through the "pdftotext" tool and will result in a "ed453610.txt" which will be stored in the text rendition.

5.3. PDF Processing

5.3.1.   Renaming PDF files

The submitted PDF file for the entire package will need to be copied to the "pdf" rendition and renamed.

The original file name is: ed{digits}.pdf

It should be renamed to: {PCS/accessId}.pdf

For example, the file: ed483019.pdf

Should be copied to the "pdf" rendition and renamed: ERIC-ED483019.pdf

5.4. HTML Processing

There will be no HTML rendition for this collection. Only a PDF rendition will be made public. The text version of the files will not be made public in any way. They are used only for parsing.

6. Content Publishing and Indexing

6.1. Indexing Granularity

There will be an entry in the public web site search engine indexes for:

1. The package as a whole
    o   This entry will be used for the "Browse By Date" feature on FDsys as well as for simple and advanced searches.
    o   The PDF content file for the package will be indexed with this entry, as well as the metadata

6.2. Index-profile field mapping

Notes:

1. {pdfobj} is computed as follows:

    /fdsysPackage/contentSect/rendition[label/mime='application/pdf' and isPublic='true']/digitalObject/techMdGroup

index-profile field	from fdsys.xml entity(ies) and *Special Instructions*	Purpose
Standard ESP Fields		
title*	PM/title	results, sorting
getpath	file://{package-directory-location}/{pdfobj}/filePath *Note: "getpath" is a special FAST field for loading document data by file name URL. It will be indexed into the "body" and "content" fields.*	full content search
teaser	PM/abstract	results
contenttype*	"text/html"	indexing control
language*	PM/language	indexing control
charset*	"utf-8"	indexing control
url	http://www.gpo.gov/fdsys/pkg/{PCS/accessId}/ {$pdfobj/filePath}	For admin and testing through the FAST SFE
Standard Document Types and Identifiers		
accode*	PM/collectionCode	navigator, results
packageid*	PCS/accessId	results
granuleid	<blank>	n/a
docclass*	PCS/docClass	navigator, results
granuleclass	<blank>	n/a
processingcode*	PM/collectionCode	results

Granule and Collection Hierarchy Fields		
nodeclass*	"simple;package;browse;"	search control
treesort*	*(See next section below)*	sorting
ancestors	<blank>	n/a
thisnode	<blank>	n/a
Publishing Specific Fields		
publishdate*	PM/dateIssued	sorting
publishdatehier*	PM/dateIssued *Formatted as "YYYY; YYYY/MM; YYYY/MM/DD"*	hierarchical navigator
publishyear*	PM/dateIssued *Formatted as "YYYY"*	navigator
publishmonth*	PM/dateIssued *Formatted as "MM"*	navigator
publishweek*	PM/dateIssued *Formatted as "MM-DD/W" where the MM-DD/W is the month and day of the Friday on or after the GM/eventDate OR the last of the month, whichever is earlier.*	navigator
publishday*	PM/dateIssued *Formatted as "MM-DD/W" where "W" is the numeric day of the week where 1=Sunday and 7=Saturday*	navigator
publishmonthyear*	PM/dateIssued *Formatted as "YYYY-MM"*	navigator for browse
firstpage	<blank>	results, citation search
lastpage	<blank>	results, citation search
pageprefix	<blank>	
governmentauthor*	{PM/governmentAuthor1}; {PM/governmentAuthor2};	navigator
xml*	*(a copy of the mods.xml)*	advanced search
Fields for Relevancy Ranking Control		
grank1	{PCS/ericNumber}; {ericnum-formatted} *Note:* • *{ericnum-formatted} is PCS/ericNumber formatted as "{alphprefix} {ddd} {ddd}. Where "d" are the digits from the numeric suffix.* *For example:* *ED463412; ED 463 412*	relevancy ranking
grank2	{PM/title}; "Education Reports from ERIC"	relevancy ranking

grank3	{PCS/institution}; {PCS/sponsorAgency}; {PCS/personalAuthor}... *Note:* *1.* *There may be multiple {PCS/personalAuthor} fields, include them all separated by semi-colons.*	relevancy ranking

grank4	{PM/abstract}; {PCS/subject};... {PCS/identifier};...
	
	 Note:
	
	    1. *There may be multiple {PCS/subject} values, include them all separated by semi-colons.*
	
	    2. *There may be multiple {PCS/identifier} values, include them all separated by semi-colons.*
		relevancy ranking

grank5	<blank>	relevancy ranking
grank6	<blank>	relevancy ranking
File Access Fields
pdffile	{pdfobj}/filePath	content delivery
pdfsize	{pdfobj}/fileSize	content delivery
htmlfile	<blank>	n/a
htmlsize	<blank>	n/a
other1file	<blank>	n/a
other1size	<blank>	n/a
other1mime	<blank>	n/a
other2file	<blank>	n/a
other2size	<blank>	n/a
other2mime	<blank>	n/a
Common FDsys Metadata
branch*	PM/branch	navigator
chamber*	<blank>	navigator
category*	PM/category	navigator
Standard Navigator Add-Ins
fdsys_orgs	{PCS/institution}; {PCS/sponsorAgency}	navigator
fdsys_people	{PCS/personalAuthor};… Note • *There may be multiple {PCS/personalAuthor} values, include them all separated by semi-colons.*	navigator
fdsys_locations	<blank>	navigator
fdsys_concepts	{PCS/type}:…{PCS/subject};... {PCS/identifier};... *Note* • *There may be multiple values for the three metadata fields, include them all separated by semi-colons.*	navigator

Collection-Specific Fields		
generic1val	{PCS/type};… *Note:* 1. *There can be multiple subjects and identifiers. Include all of them, separated by semicolons.*	navigator
agencies	{PCS/sponsorAgency}	navigator
categoryhier	Subject/{PCS/subject};… Identifier/{PCS/identifier};… Type/{PCS/type};…. *Note:* 1. *There can be multiple subjects, identifiers and types. Include all of them, separated by semicolons.* 2. *For each entry set the appropriate prefix constant.* *For example:* *Subject/Children;* *Subject/Communication Skills;* *Subject/Communication (Thought Transfer);* *Subject/Deaf Blind;* *Type/Guides;* *Type/Non-Classroom*	hierarchical navigator
r.ericnumber	{PCS/ericNumber}	results
r.fericnumber	{PCS/ericNumber} *Note:* • *Format ERIC number as "{alphaprefix} {ddd} {ddd}". Where d are digits from the numeric suffix.* • *For example, "ED 463 412"*	results
r:isfallbacktitle	{PCS/isFallbackTitle}	results

Note:

1. Fields prefixed with "r:" are stored in the "resultsbundle" index-profile field, as a nested list of name/value pairs. These pairs are unbundled for display as needed by the search API layer.

6.3. Computing "treesort"

"treesort" will be used to sort the documents when being displayed for collection browsing.

Tree sort will be made up of the following components, each separated by a forward slash:

order	fdsys.xml field	Formatting and *Special Instructions*
1	PM/dateIssued	Format to YYYY
2	PCS/ericNumber	

For example:

1996/ED438032

1996/ED439 239

7. Search and Browse

7.1. Search Results Presentation

Line	Pattern (using index-profile fields) and *Special Instructions*
1	{r.fericnumber} – {title} [PDF {pdfsize}] *Note:* *1. The title should link to the PDF file.* *2. If r:isfallbacktitle is "true", use "Education Report from ERIC" instead of {title}*
2	Education Reports from ERIC. {agencies}.{publishdate}. *Note:* *1. {publishdate} is converted to "Month day, Year" format, for example "November 2, 2008"*
3	{teaser} More information. *Note: More information links to the package's content-detail page.*

Example:

---

i)

ED 466 380 – On the Development of Human Representational Competence from an Evolutionary Point of View: From Episodic to Virtual [PDF 1238 KB]
Education Reports from ERIC. Office of Educational Research and Improvement (ED), Washington, DC. October 1, 1999.
… Suggested that the evolutionary perspective needs to complement mathematics educators' other ways of understanding the learning More information.

ii)

ED 464 338 – Learning Together. Parents and Children Together Series [PDF 2984 KB]
Education Reports from ERIC. Office of Educational Research and Improvement (ED), Washington, DC. April 1, 2002.
… Before you read a story, talk about the title or things that might happen in it. Then, after you have finished reading, talk about what happened in the story. …More information.

iii)

ED 464 399 – Education Report from ERIC [PDF 2984 KB]
Education Reports from ERIC. Office of Educational Research and Improvement (ED), Washington, DC. April 1, 2002.
… Before you read a story, talk about the title or things that might happen in it. Then, after you have finished reading, talk about what happened in the story. …More information.

Notes:
1. Truncate {title} to 100 characters, if necessary.
   a. All truncated items should contain a "..." suffix if they are truncated.
   b. All truncated items should be truncated before a whitespace character.

2. {pdfsize} should be converted to KB or MB, as appropriate.
3. Convert {publishdate} to standard GPO "Month day, Year" format, for example "November 2, 2008".

7.2. Navigators

Collection specific navigators are listed here.

Display Name	Navigator	index-profile field	Type	Purpose
Document Category	categoryhiernav	categoryhier	string, hierarchical	results

7.2.1.　Navigator Examples

7.2.1.1.　Document Category

Note: The quality of this navigator will be evaluated in development.

{categoryhier}

Navigator values are already formatted at the index level; no need for post-processing.

+　　　Subject
- Dialog Journals
- Elementary Education
- Learning Activities
- Parent Child Relationship
- …

+　　　Identifier
- Family Activities
- Read Along
- Team Learning
- …

+　　　Type
- Creative Works
- Guides - Non-Classroom
- ERIC Publications
- …

7.3. Search Fields

Notes:

1. All of the standard fields, even though they are not listed below, must be available for this collection.

2. The advanced search form will be made up of all of the fields with the "Srch Form" column equal to "Y".

3. Developers are encouraged to use a pull-down box or query completion for elements with pre-defined valued.

4. When two or more collections are selected by a user on the search form, the metadata in the pull-down should be a join of the common elements.

Display Name	Search field: (all one word)	Fast FQL	Srch Form	Type	Allowed Values	Help Text
ERIC Number	ericnumber	xml:extension:ericNumberFormatted: (${s})	Y	string		Search for an ERIC resource based on its number. Example ERIC Numbers are "ED466380", or "ED 483 022", (either format is allowed).
Subject	ericsubject	xml:extension:subject:(${s})	Y	string		Search for an ERIC resource based on its subject. Examples subjects are "Cognitive Psychology", "Elementary Education" and "Story Reading"
Identifiers	ericidentifier	xml:extension:identifier:(${s})	Y	string		The topical area covered by the ERIC resourse. Example Identifiers are "Family Activities; *Read Along; Team Learning".
Sponsoring Agency	sponsoragency	xml:extension:sponsorAgency:(${s})	Y	string		Search for an ERIC resource based on the Agency that funded the creation of the resource. An example agency is "Office of Educational Research and Improvement (ED), Washington, DC."

Source Institution	institution	xml:extension:institution:(${s})	Y	string		The institution that created the ERIC resource, for example "Family Learning Association, Bloomington, IN."

7.3.1.   Search Form Query Completion

The following fields will be present in the query completion server.

Advanced Search Field	QC Value Prefix	Data comes from fdsys.xml fields
Government Author*	GAU	PM/governmentAuthor1, PM/governmentAuthor2, PM/governmentAuthor3
SuDoc Class Number*	GSU	PM/classification[@authority='sudocs']
Type Education Resource	ERR	PCS/type
Institution	ERI	PCS/institution
Sponsor Agency	ERA	PCS/sponsorAgency
Subject	ERS	PCS/subject
Identifiers	ERD	PCS/identifiers

7.4. Collection Browsing

There will be one method for collection browsing for R1C2: by date and resource type.

7.4.1.   Front Page

The following is the description of the collection to be presented to public users:

> Education Reports from ERIC
>
> Find reports on federally funded education research topics from the U.S. Department of Education's Educational Resources Information Center. Reports on FDsys begin with those received in October 2002. Prior reports are available from select Federal depository libraries nationwide in microfiche. A larger selection of ERIC reports is available from the ERIC program. Files are available in Adobe Portable Document Format (PDF) only.
>
> [About Education Reports from ERIC](#)

[About the Education Reports from ERIC](#)

> links to: Education Reports from ERIC in RoboHelp.
>
> (http://www.gpo.gov/help/index.html#about_education_reports_from_eric.htm)

7.4.2. Browse

Level	Display Name	Navigator	Purpose / Special Instructions
1	Date Issued Year	pubyearnav	Not Applicable

The hierarchy method will be: Ancestor siblings.

7.4.2.1. Presentation of articles

The articles at the leaf of the browse-by-date tree should be presented the same as specified in section 7.1, with the following exceptions:

1. Remove the PDF size reference from line 1.
2. Remove the collection name and date from line 2.
3. Remove line 3.

For example:

ED 464 338 – Learning Together. Parents and Children Together Series.
Office of Educational Research and Improvement (ED), Washington, DC.   PDF | More

Notes:

Articles are to be sorted by "treesort" (see section 6.3).

7.4.2.2. Complete Collection Browsing Example

+ 2004

+ 2003

- 2002

       ED 463 411 - Effective Advisory Committees. In Brief: Fast Facts for Policy and Practice
       Office of Vocational and Adult Education (ED), Washington, DC   PDF | More

       ED 463 445 - High Schools That Work: Best Practices for
       CTE. Practice Application Brief No. 19
       Office of Vocational and Adult Education (ED), Washington, DC   PDF | More

       ED 464 053 - Educating Preservice Teachers: The State of Affairs
       Office of Vocational and Adult Education (ED), Washington, DC   PDF | More

       ED 464 268 - Building Stronger School Counseling Programs: Bringing
       Futuristic Approaches into the Present
       Office of Vocational and Adult Education (ED), Washington, DC   PDF | More

       ...

+ 2001

+ 2000

+ 1999

…

7.4.2.3.　Browse Latest Resources

Purchase Education Publications from the U.S. Government Online Bookstore.

Link to: http://bookstore.gpo.gov/education.jsp

Locate Education Resources Information Center Reports in a local **Federal depository library**.

Link to: http://catalog.gpo.gov/fdlpdir/public.jsp

8. Content Delivery

This section covers any collection-specific needs for accessing the documents. This includes the content detail page, the collection browsing, and providing standard metadata formats.

8.1. Available Downloads

The following are the URLs which will be provided to the public.

- For the package as a whole:
    - PDF file
    - mods.xml
    - premis.xml
    - ZIP file of the entire package contents

All URLs are in the FDsys standard format. See the SDD Volume XIX – Common Metadata and Standard References for more details.

8.2. Content Detail Page

8.2.1.   Header

The header at the top of the content-detail page will read the same as line 1 from section 7.1.

8.2.2.   Fields to display

The following fields are to be displayed on the content-detail pages. Note that the fields are to be displayed in the same order as they are listed below.

Display Name	fdsys.xml Entities and *Special Instructions*	Example
Category	{PM/category}	Executive Agency Publications
Collection	"Education Reports from ERIC"	Education Reports from ERIC
SuDoc Class Number	{PM/classification[@authority='sudocs']}	ED 1.615:
Date Issued	{PM/dateIssued}	December 1, 2002
Author	{PCS/personalAuthor}… Set all entries separated by semi-colons.	Ahearn, Eileen M.; Lange, Cheryl M.; Rhim, Lauren Morando; McLaughlin, Margaret J.
Source Institution	{PCS/institution}	National Association of State Directors of Special Education, Alexandria, VA.
Sponsoring Agency	{PCS/sponsorAgency}	Special Education Programs (ED/OSERS), Washington, DC.

Publication Type	{PCS/type}… *Set all available entries in a comma separated list.*	Reports – Evaluative. Numerical/Quantitative Data
Subject	{PCS/subject}… *Set all available entries in a comma separated list.*	Accountability, Case Studies, Charter Schools, Compliance (Legal), Delivery Systems
Identifiers	{PCS/identifiers}… *Set all available entries in a comma separated list.*	Family Activities, Read Along, Team Learning
Abstract	{PCS/abstract}	The message of this series of books, "Parents and Children Together," is that parents should get together with their children, talk about stories, and learn together. This book, "Learning Together," contains…

Notes:

1. All dates must be formatted as: {month} {day}, {year} - as follows:

    -   Example: October 31, 2008

8.2.3.   Actions

This collection contains only standard actions.

-   Browse Education Reports from ERIC.

-   More Information about Education Reports from ERIC.

-   View in Catalog of U.S. Government Publications

-   Find at a local Federal depository library

-   Purchase Educational Publications from the GPO Bookstore

-   Email a link to this page

Note:

-   "Purchase educational publicatons from the GPO Bookstore" should link to http://bookstore.gpo.gov/education.jsp ...

8.2.4.   Related Publications

This collection does not have related collection entries.

9. mods.xml Mapping

Since this collection has no granules, there will only be a single mods.xml produced for all situations. This mods.xml structure is mapped out in the following section.

9.1. mods.xml structures

This section defines the two different structures for the mods.xml.

mods.xml for the entire issue:

```
<?xml version="1.0" encoding="UTF-8"?>
<mods version="3.3" ID="{PH/@id}"
      xsi:schemaLocation="http://www.loc.gov/mods/v3
http://www.loc.gov/standards/mods/v3/mods-3-3.xsd"
      xlink:href="http://www.gpo.gov/fdsys/pkg/{PCS/accessId}/mods.xml"
      xmlns:xlink="http://www.w3.org/1999/xlink"
      xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
      xmlns="http://www.loc.gov/mods/v3">

  {Component 1:  Publication Metadata independent of package or granule}
  {Component 2:  Package Metadata for the package as a whole}

</mods>
```

9.2. mods.xml Components

This section specifies the mapping of mods.xml collection-specific metadata entities from fdsys.xml entities.

Notes:

1. All elements below are children of the top-level <mods> tag which is specified above.

2. Items indicated with an asterisk (*) below are the same for all collections.

MODS Schema Entry	MODS entity attributes	fdsys.xml entity(ies) & *special instructions*
Component 1: Metadata elements independent of package or granule		
Include all "common publication" metadata specified for FDsys, see SDD Volume XIX – Common Metadata and Standard References for more details.		
Component 2: Metadata elements only for the package as a whole		
Include all "common package" metadata specified for FDsys, see SDD Volume XIX – Common Metadata and Standard References for more details.		
titleInfo/title		PM/title
location/url	displayLabel="PDF rendition" access="raw object"	http://www.gpo.gov/fdsys/pkg/{PCS/accessId}/pdf/{PCS/accessId}.pdf
location/url	displayLabel="Content Detail" access="object in context"	http://www.gpo.gov/fdsys/pkg/{PCS/accessId }/content-detail.html
identifier	type="preferred citation"	{PCS/ericNumber} *Format ERIC number as "ED {ddd} {ddd}". Where "ddd" segments are from numeric suffix.* *For example, ED464761 is formatted as "ED 464 761".*
extension/searchTitle		{PCS/ericNumber}; {PM/title}; {eric-number-formatted} *{eric-number-formatted} = format ERIC number as "ED {ddd} {ddd}". Where "ddd" segments are from numeric suffix.* *For example, ED464761 is formatted as "ED 464 761".*

extension/ericNumberFormatted		{PCS/ericNumber}; {eric-number-formatted}
		{eric-number-formatted} = format ERIC number as "ED {ddd} {ddd}". Where "ddd" segments are from numeric suffix.
		For example, ED464761 is formatted as "ED 464 761".
extension		PCS/*

Note:

- No standard references will be set for this collection.