

Free Tools for Your Preservation Toolbelt

Harvard Library
Spencer McEwen, Randy Stern
June 9, 2011



Agenda

1. Harvard's DRS 2
2. DRS 2 Objects
3. Open Source Tools from DRS 2
 1. FITS – on Google Code
 2. OTS-Schemas – on Google Code
 3. OTS – ?
 4. BatchBuilder - ?
4. Questions/Feedback

What is DRS 2?

DRS 2

- ❑ The second generation of the Harvard Library Digital Repository Service
- ❑ A preservation and access repository for digital objects stewarded by Libraries, Archives and Museums at Harvard
- ❑ DRS 2 models “objects” rather than files
- ❑ e.g. – delivery, archival master, and production master images comprise one object
- ❑ e.g. - All images and ocr text for a book comprise one object

What's in the DRS?

What's in the DRS?



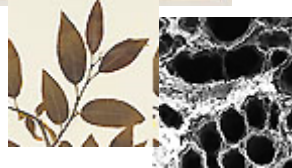


What's in the DRS?



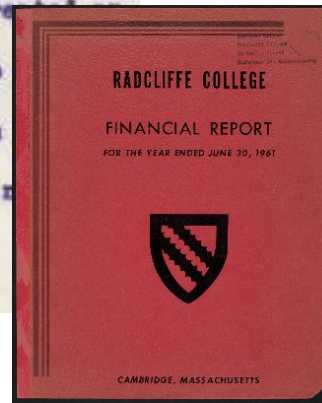


What's in the DRS?



(Is the country now better off or worse off than it would have been if Stalin had not come to power?) Stalin created an artificial famine because the peasants did not want to go to the kolkhozes. Lenin would not have done that. Lenin had taken severe measures in the Civil War, but there was no state treasury then.

Google books



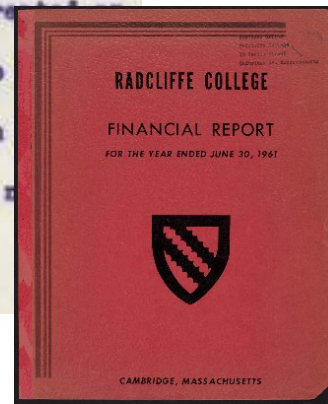


What's in the DRS?



(Is the country now better off or worse off than it would have been if Stalin had not come to power?) Stalin created an artificial famine because the peasants did not want to work on the kolkhozes. Lenin would not have done that. Lenin had taken severe measures in the Civil War, but there was no state treasury then.

Google books

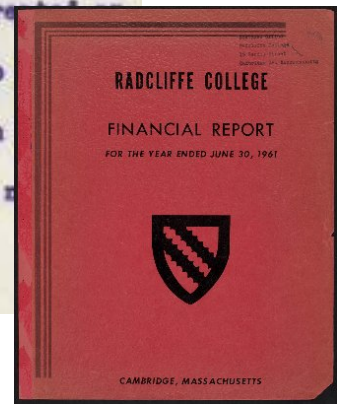


What's in the DRS?



(Is the country now better off or worse off than it would have been if Stalin had not come to power?) Stalin created an artificial famine because the peasants did not want to work the kolkhozes. Lenin would not have done that. Lenin had taken severe measures in the Civil War, but there was no state treasury then.

Google books





What's in the DRS?



(Is the country now better off or worse off than it would have been if Stalin had not come to power?) Stalin created artificial famine because the peasants did not want to the kolkhozes. Lenin would not have done that. Lenin taken severe measures in the Civil War, but there was state treasury then.

Harvard University Library
Web Archive Collection Service
Why was this page archived?
Archived on: November 05, 2007 15:00:56. Archived versions of this web page
Collection: Constitutional Revision in Japan Research Project
Archived by: Edwin O. Reischauer Institute of Japanese Studies, Harvard University

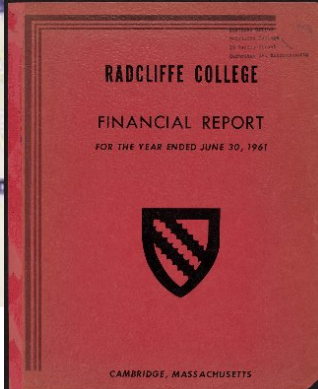
東京自治労連

更新日 10月23日
アクセスカウンタ

東京自治労連組合連合会
東京都豊島区南大塚2-33-10東京労働会館4F
TEL: 03-5910-7951 FAX: 03-5910-7957

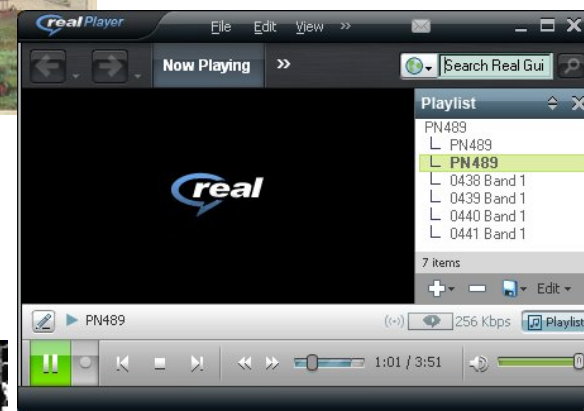


- ### 声明・見解
- 2007年 10月17日 [2007年都区人事委員会動向に対する見解](#) **News!**
 - 2007年 9月19日 [福祉・教育など市民生活の充実と地方自治体運営のための地方財政の確保を求めずして](#)
 - 2007年 9月 5日 [保育施設増設の遅れを断った東京都における新しい到達点](#)
 - 2007年 8月10日 [公立保育園における「派遣・兼員」問題について](#) 派遣減速・兼員削減を一掃し、真に保育の質を確保できる執行体制を確立しよう
 - 2007年 8月 8日 [内閣府による京口東区（住基・戸籍）長期委託化問題市町村アンケートに基づいた新案の提議・法政省長閣僚について](#)
 - 2007年 8月 8日 [2007年人事院勧告について](#)
 - 2007年 7月25日 [東京都「新橋キャンパス再編整備推進委員会」に対する見解-何ひとつとしてネットの無い地方自治体推進人化策-](#)
 - 2007年 7月18日 [文京区分譲後継サービス事業実施取り消し問題に対する見解と対応について](#)
 - 2007年 7月 4日 [2007年参院選選挙で実現をめざす東京自治労連基本要綱](#)



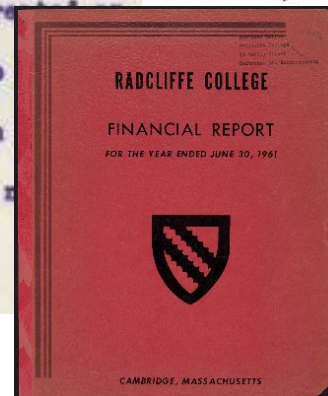


What's in the DRS?



Harvard University Library Web Archive Collection Service. Why was this page archived? Archived on: November 05, 2007 15:00:56. Archived versions of this web page. Collection: Constitutional Revision in Japan Research Project. Archived by: Edwin O. Reischauer Institute of Japanese Studies, Harvard University.

(Is the country now better off or worse off than it would have been if Stalin had not come to power?) Stalin created artificial famine because the peasants did not want to the kolkhozes. Lenin would not have done that. Lenin taken severe measures in the Civil War, but there was state treasury then.



東京自治労連

東京自治労務労働組合連合会
東京都豊島区南大塚2-33-10東京労働会館4F
TEL: 03-5910-7951 FAX: 03-5910-7957

更新日: 10月23日
アクセスカウンタ: 10000

10月23日 東京自治労連 第19回定期大会 説明・見解

- 2007年 10月17日 2007年都区人事委員会勧告に対する見解 **New!**
- 2007年 9月19日 福祉・教育など民生福祉の充実と地方自治体運営のための地方財政の確保をめぐって
- 2007年 9月 5日 保育施設増設の遅れを断った東京都における懸念の到達点
- 2007年 8月10日 公立保育園における「派遣・兼員」問題について 派遣削減・兼員削減を一掃し、真に保育の質を確保できる執行体制を確立しよう
- 2007年 8月 8日 内閣府による窓口業務（住民・戸籍）業務委託化問題に関する町民アンケートに基づいた新たな提議・法政省見解について
- 2007年 8月 8日 2007年人事院勧告について
- 2007年 7月25日 東京都「新橋キャンパス再編整備推進本部」に対する見解-何ひとつとしてメリットの無い地方自治体法人化について
- 2007年 7月18日 文京区分譲地売却サービス事業実施取り消し問題に対する見解と対応について
- 2007年 7月 4日 2007年参院選選挙で実現をめざす東京自治労連基本要綱

Use of WAX is subject to our Terms of Use. Questions and Co...

Object Preservation Metadata

- Our motivation:
 - Digital preservation requires accurate and sufficient technical metadata to support preservation planning and activities
 - Descriptive metadata is also valuable for identification and management by curators
 - Standards-based schemas maximize tool support and ability to exchange data with other repositories

DRS2

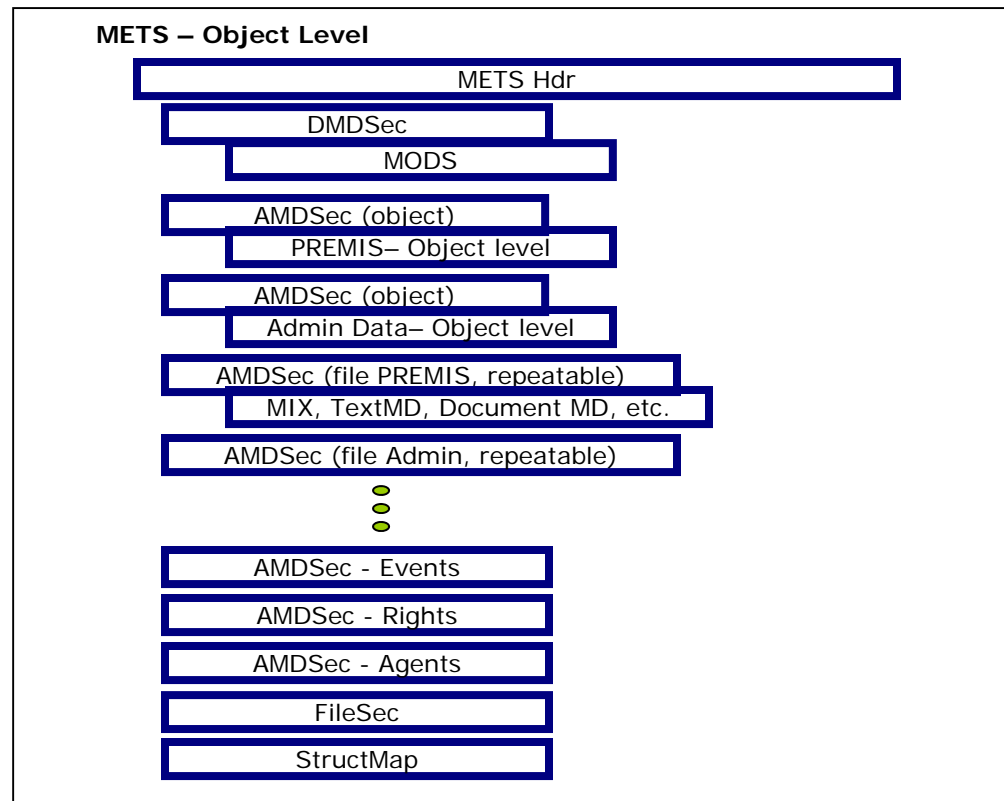
Object Descriptor Files

- A single METS container file per object
 - Descriptive, administrative, preservation, technical, structural, and rights metadata
 - Describes the object, all its files and bitstreams and related significant events
- Self-contained, portable objects - stored in the file management layer alongside the content files

Object Descriptor Files

- Each METS file embeds standard metadata schemas
 - MODS – Descriptive metadata
 - HuIDRSAdmin – administrative metadata
 - METS/Structmap – structural metadata
 - PREMIS – technical metadata, agents, rights, events
 - Format dependent technical metadata
 - MIX for images
 - TextMD for text
 - DocumentMD for PDF and MS Office documents
 - AES Core Audio for audio

METS Object Descriptors

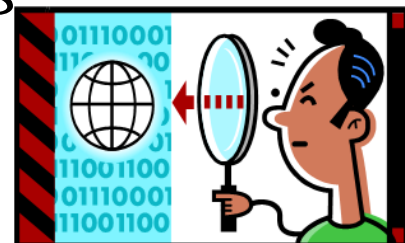


4 Object Descriptor Tools

- ❑ Tool 1 - FITS (File Information Tool Set)
- ❑ Tool 2 - OTS-Schemas (Object Tool Set Schemas)
- ❑ Tool 3 - OTS (Object Tool Set)
- ❑ Tool 4 - BatchBuilder

Tool 1 - FITS

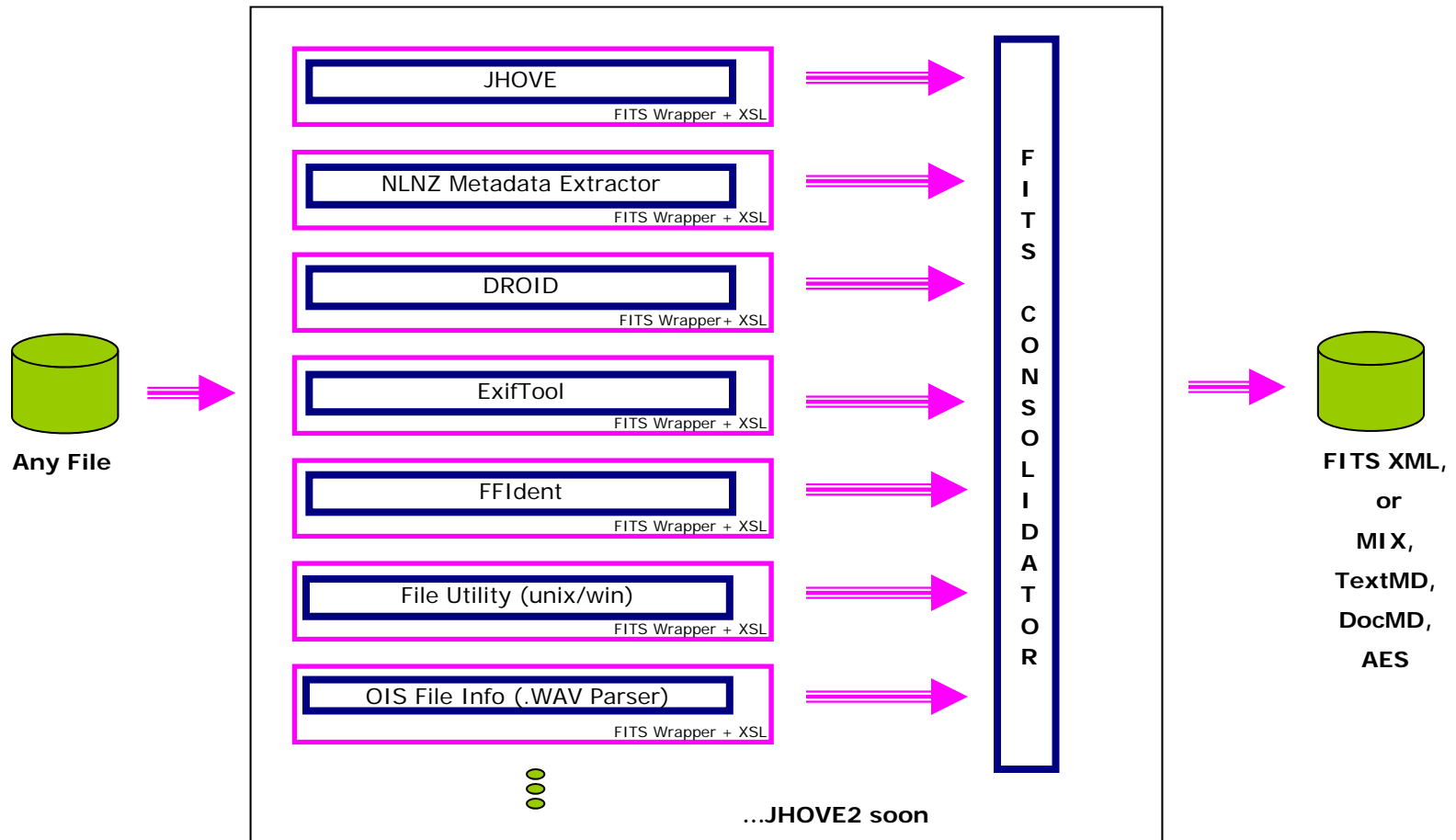
- ❑ Identifies, validates, and extracts technical metadata from files
- ❑ Wraps many tools, resolves tool conflicts and consolidates results
- ❑ Outputs MIX, TextMD, DocMD, AES Core Audio, etc. (using OTS-Schemas)
- ❑ Command line and java API
- ❑ 162 downloads of latest version
- ❑ Used by other projects: Archivematica, KEEP, SCAPE
- ❑ <http://code.google.com/p/fits/>





FITS - File Information Tool Set

File Characterization and Metadata Extraction



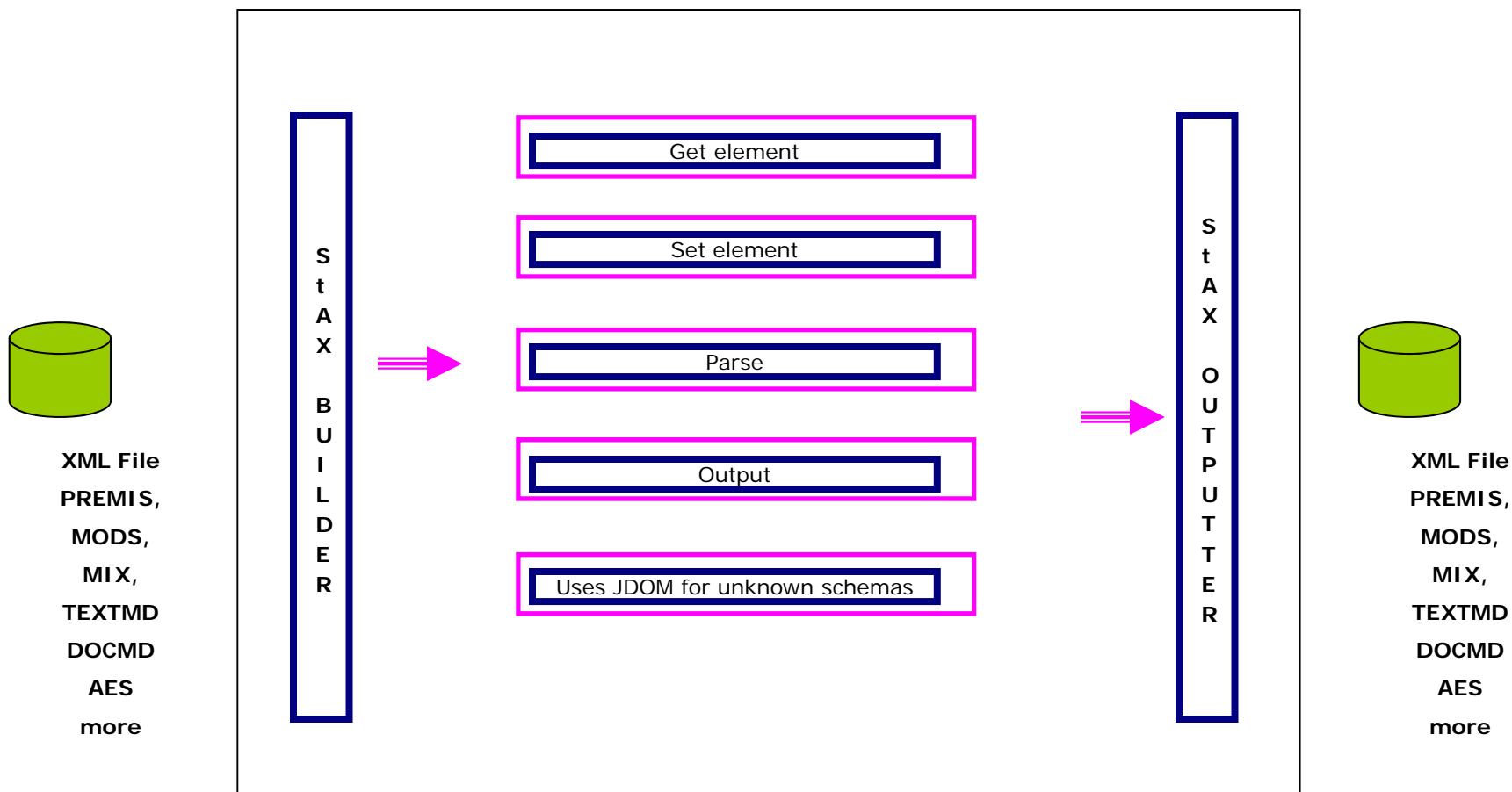
Tool 2 - OTS-Schemas

- ❑ Java library for reading and writing documents in common XML schemas.
- ❑ API exposes metadata elements for get/set as Java Objects
- ❑ Uses efficient StAX parser
- ❑ PREMIS, MODS, TextMD, DocumentMD, MIX, AES Core Audio, METS (partial support), and several custom formats.
- ❑ FITS "-x" switch uses OTS-Schemas
- ❑ <http://code.google.com/p/ots-schemas/>



OTS-Schemas

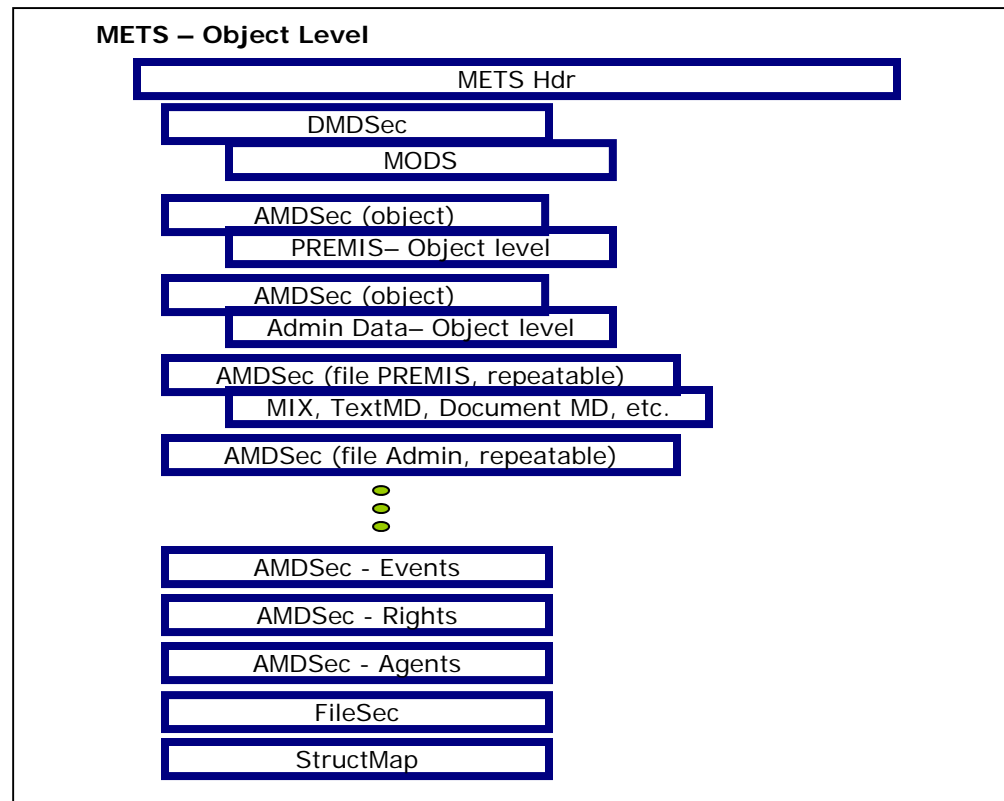
Standard Metadata Schema Read/Write



Tool 3 – OTS (Object Tool Set)

- ❑ Java library for creating, reading, updating, and writing METS Object Descriptors
- ❑ API for interacting with descriptors at embedded schema level
- ❑ Embedded PREMIS, MODS, TextMD, DocumentMD, MIX, AES Core Audio, METS (partial support), via OTS-Schemas
- ❑ Very efficient java objects – for large XML, impractical as jdom

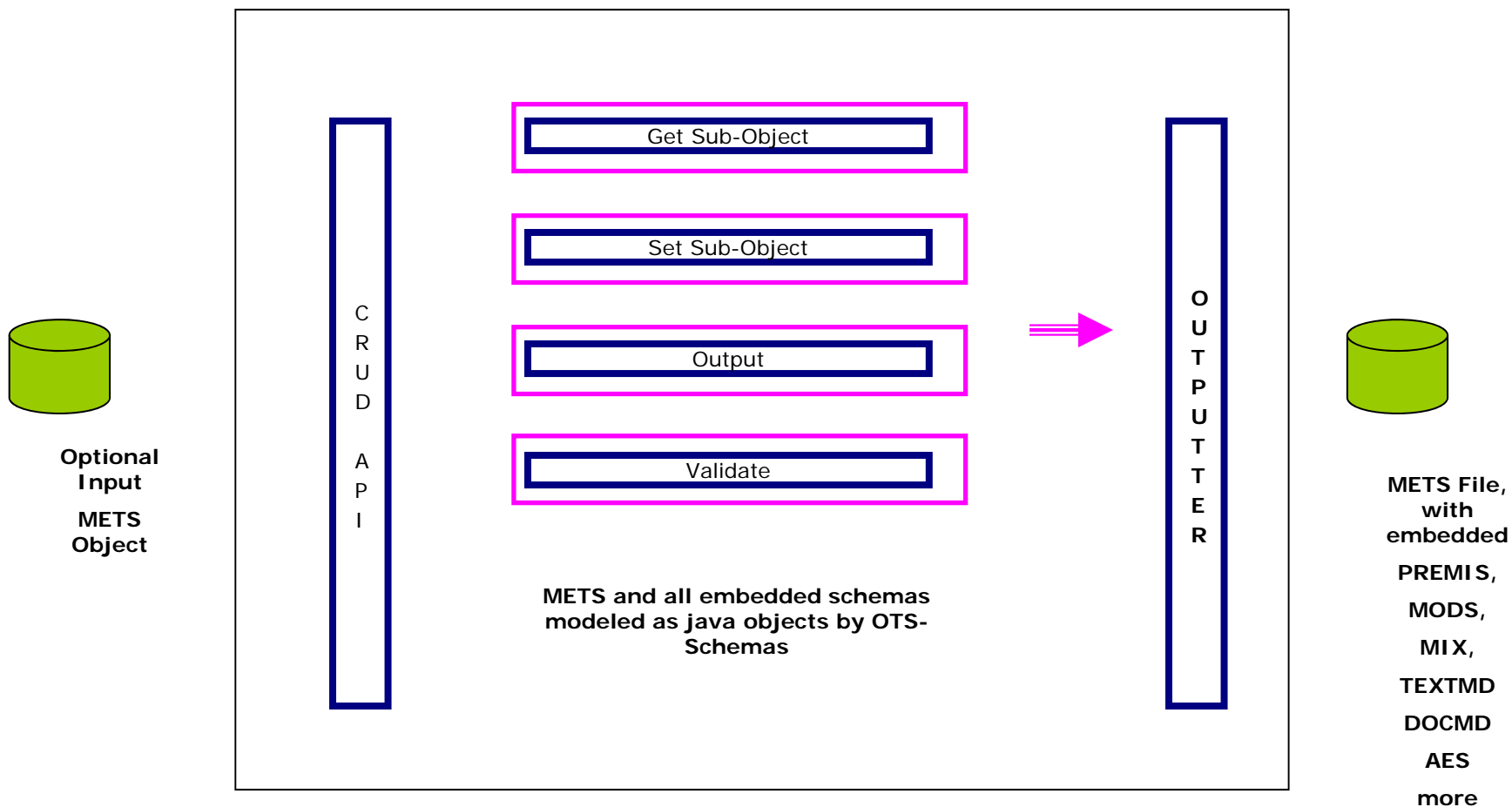
OTS METS Objects





OTS

JAVA API for Create, Read, Update, Write of DRS METS Objects



OTS

- Validates against defined content models
- Content models supported include:
 - Still Image
 - Page Turned Object
 - Audio
 - Document
 - Opaque
 - Web harvest
 - Email
 - Text ... And more...
- <http://code.google.com/p/???????>

Tool 4 - BatchBuilder

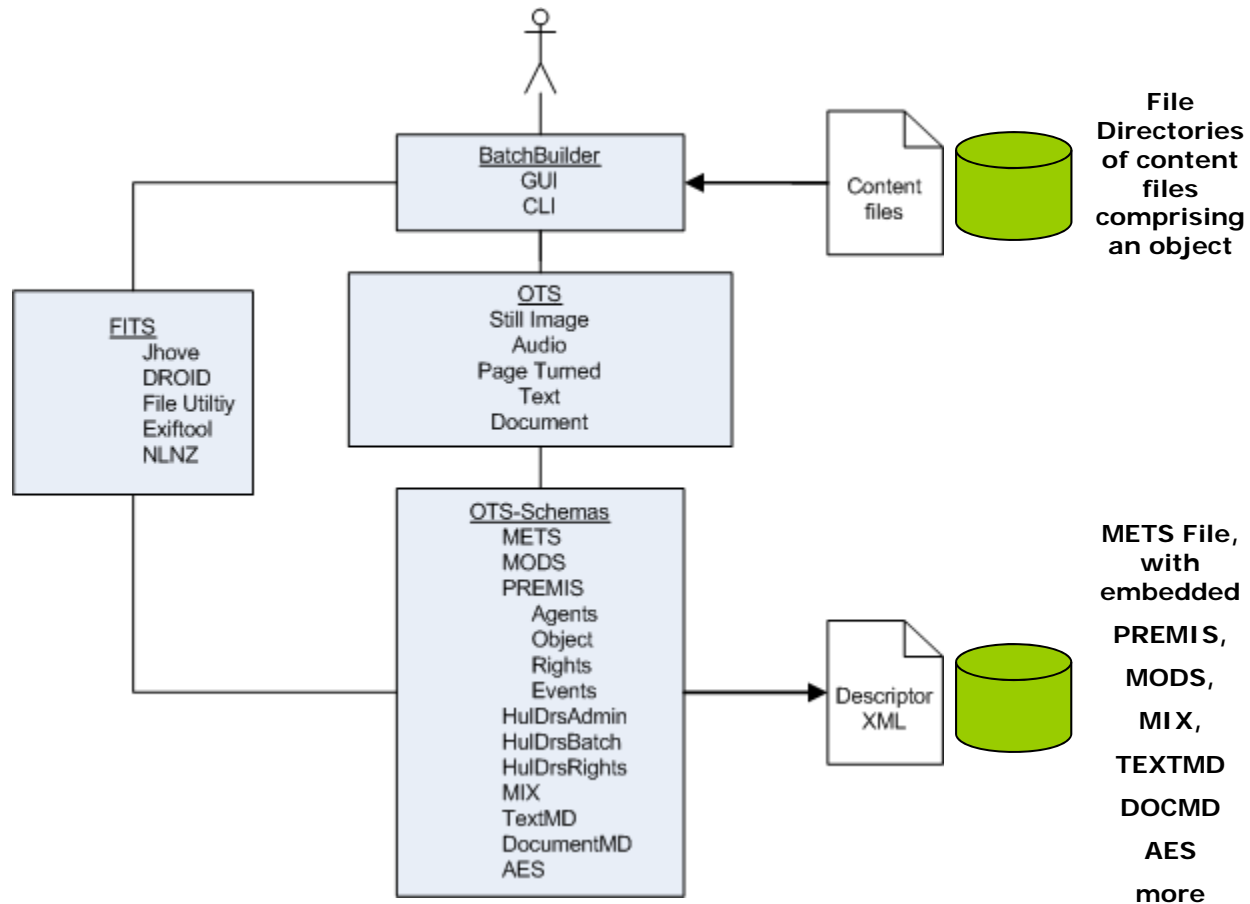
- ❑ Builds OTS METS objects (and SIP) from directory hierarchies of content files
- ❑ Invokes FITS to characterize files and extract technical metadata
- ❑ GUI for setting administrative and descriptive metadata (with optional catalog record import)
- ❑ CLI for scripted workflows
- ❑ <http://code.google.com/p/???????>





BatchBuilder

Object and METS Descriptor Builder



Object Descriptor Tools

- ❑ FITS (File Information Tool Set) *
 - ❑ OTS-Schemas (Object Tool Set Schemas) *
 - ❑ OTS (Object Tool Set) **
 - ❑ BatchBuilder **
-
- ❑ * - Open Source now
 - ❑ ** - Open source soon – if demand

Questions/Feedback?

spencer_mcewen@harvard.edu

randy_stern@harvard.edu

<http://hul.harvard.edu/ois/systems/drs/>

DRS 2 Metadata Analyst: robin_wendler@harvard.edu

Digital Preservation Manager: andrea_goethals@harvard.edu

DRS 2 Services Developer: jonathan_kennedy@harvard.edu

DRS 2 Schema Developer: gary_mcgath@harvard.edu