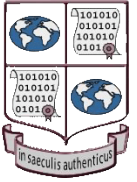




相關背景



- Library and Archives Canada (LAC)
- 2004年由前國家圖書館和國家檔案館合併而成
 - 從各自管理對象的“出版物”和“文件/檔案”到籠統的“文化遺產”
 - 在概念上認同二者的不同
 - 在管理上趨於統一
- 管轄加拿大聯邦政府（GC）部門和部長的文件銷毀和移交



聯邦政府數字文件生成



- 文字處理軟件產生的計算機“文檔”
 - Word, Excel, PPT
- 電子郵件(手機短信，即時短信, etc.)
- 數字圖像
- 數字聲音、錄像
- 數字取證影像
- 數據庫的匯報輸出
- 數據庫本身
- 靜止網頁
- 動態/互動的網頁
- 一組這樣的網頁
- 整個網站...

文件生成方式
並非
文件生成系統



聯邦政府數字文件存放處



- 員工台式PC 硬盤
- 機構內部網絡分享空間
 - SharePoint
 - Intranet
- 機構內部服務器
 - 電子郵件 (MS Outlook)
 - Wiki
- 備份盤/磁帶
- 移動存儲設備
 - ...
 - Blackberry
 - 員工筆記本電腦
- 遠程服務器
 - 簽約IT公司
 - 數據中心
 - “雲端”: Web 2.0
- ...

文件存放處
並非
文件維護系統

電子文件管理系統
EDRMS



DoD5015.2認證
RDIMS

工作流程

線上簽核



LAC數字檔案保管



- 構建一個TDR (Trusted Digital Repository)
 - 管理其所有擁有的並將要擁有的數字資源
 - 提供移交，元數據處理，儲存，保存和利用功能
- GC的數字文件是這些資源中的一部分，亦將經由TDR進入LAC，並被TDR保存及提供利用

文件存放處
且是
文件保管系統



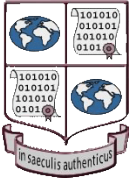
TDR概況



- 類似與NARA的ERA
- 聯邦政府中央撥款
- 團隊由LAC人員和諮詢人員組成
- 基於OAIS功能結構
- 軟件：商業，定制，源代碼開放



TDA構建



- 2005開始構思
- 2008項目開始實施
- 第一階段針對
 - 法定存放的電子出版物
 - 文字類政府文件
- 現階段
 - 提交信息包 (SIP)
 - 虛擬上傳台測試



GC – LAC數字文件移交



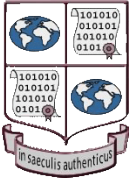
- GC 文件管理系統
 - 文件分類和保管期限系統
- LAC文件處置權系統
 - 跨部門文件處置權
 - 部門專業文件處置權
 - 特殊文件處置權

RDACS

1. 鑑定決定
2. 移交辦法



LAC鑑定決定



- 徵集原則：
 - 重要性 (significance) : ...
 - 充足性 (Sufficiency) : 其質與量是否值得或應該本館收藏
 - 可持續性 (Sustainability) : 本館是否有能力收藏並提供利用
 - 社會性 (Society) : 是否映射加國社會及是否會對加國社會做貢獻
- 應用：
 - 四條原則並用

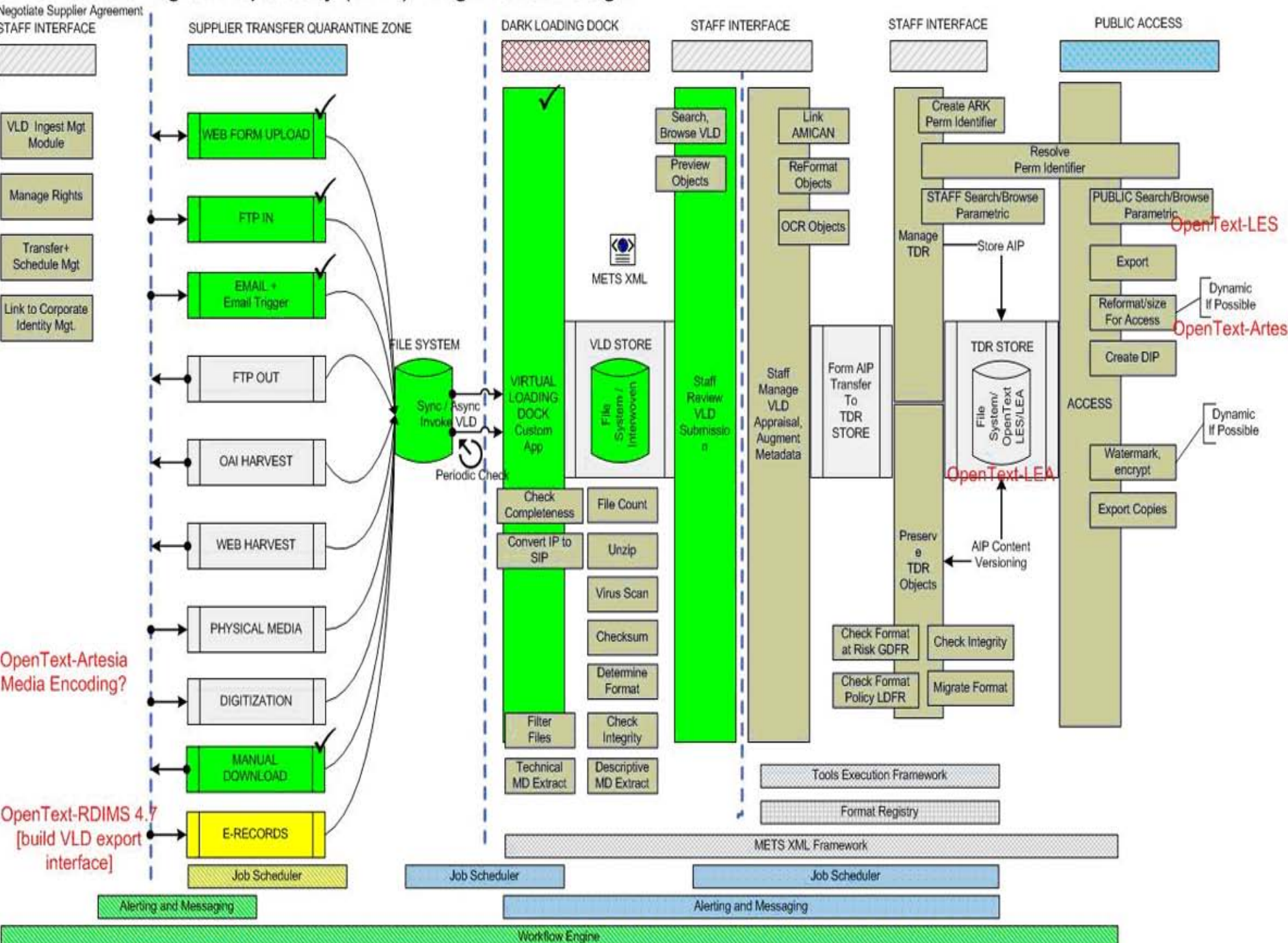


通過TDR的數字文件移交



- 設計專門移交軟件
- 接口聯邦政府的電子文件管理系統RDIMS
- 與LAC的文件處置權系統 (RDACS) 整合
- 識別移交文件，打包，加密，並通過SFTP傳輸

LAC Trusted Digital Repository (TDR) - High Level Design

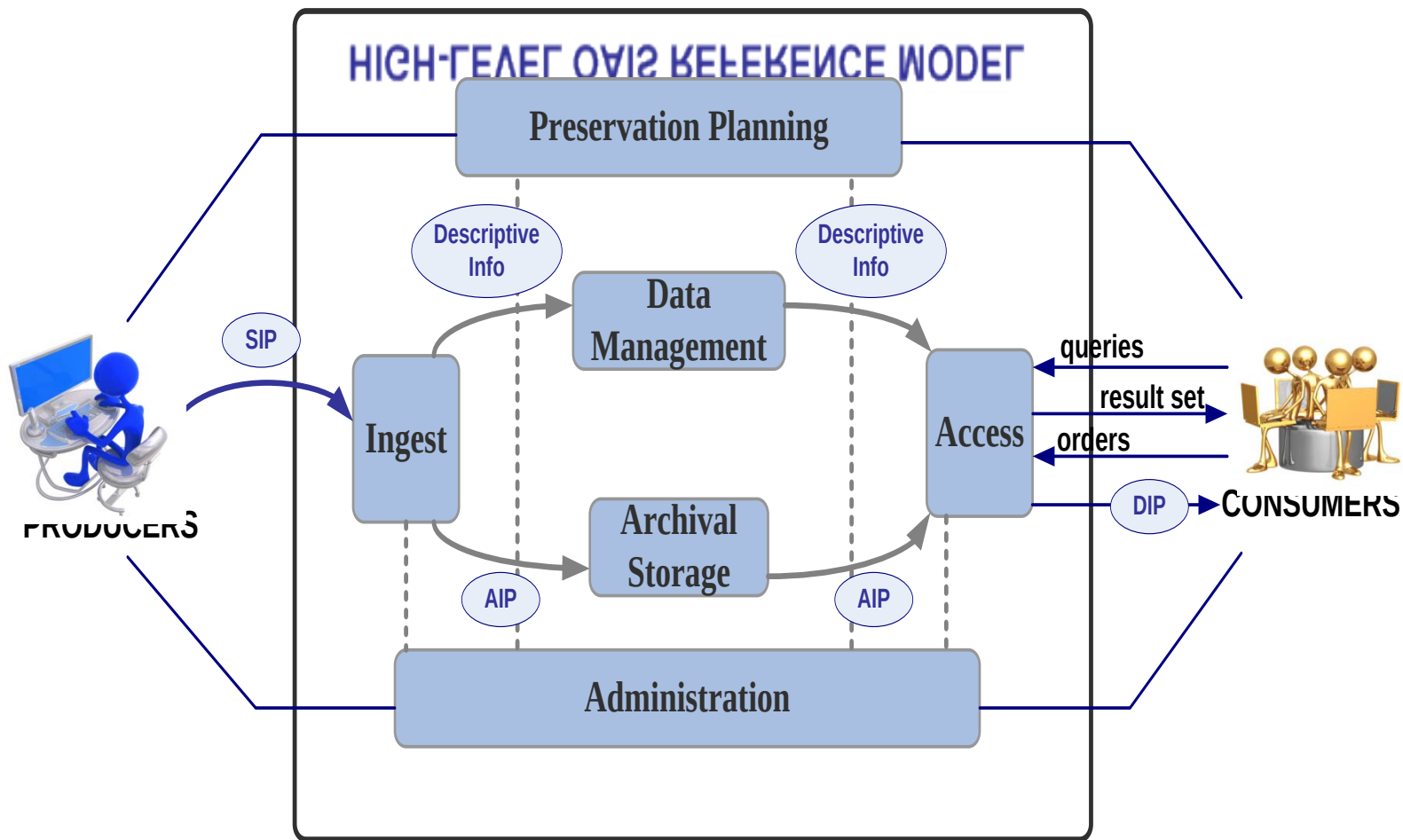




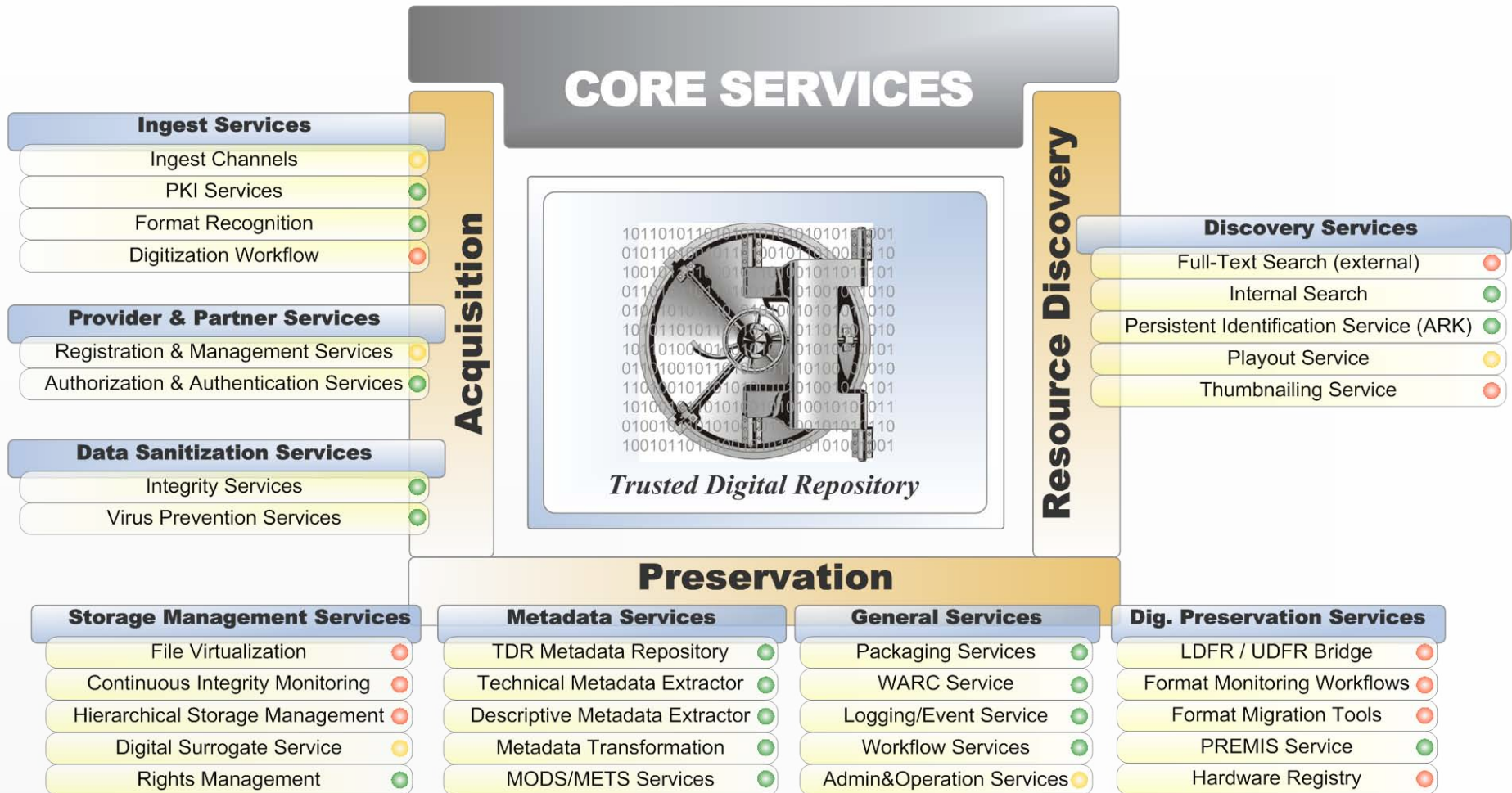
OAIS模型



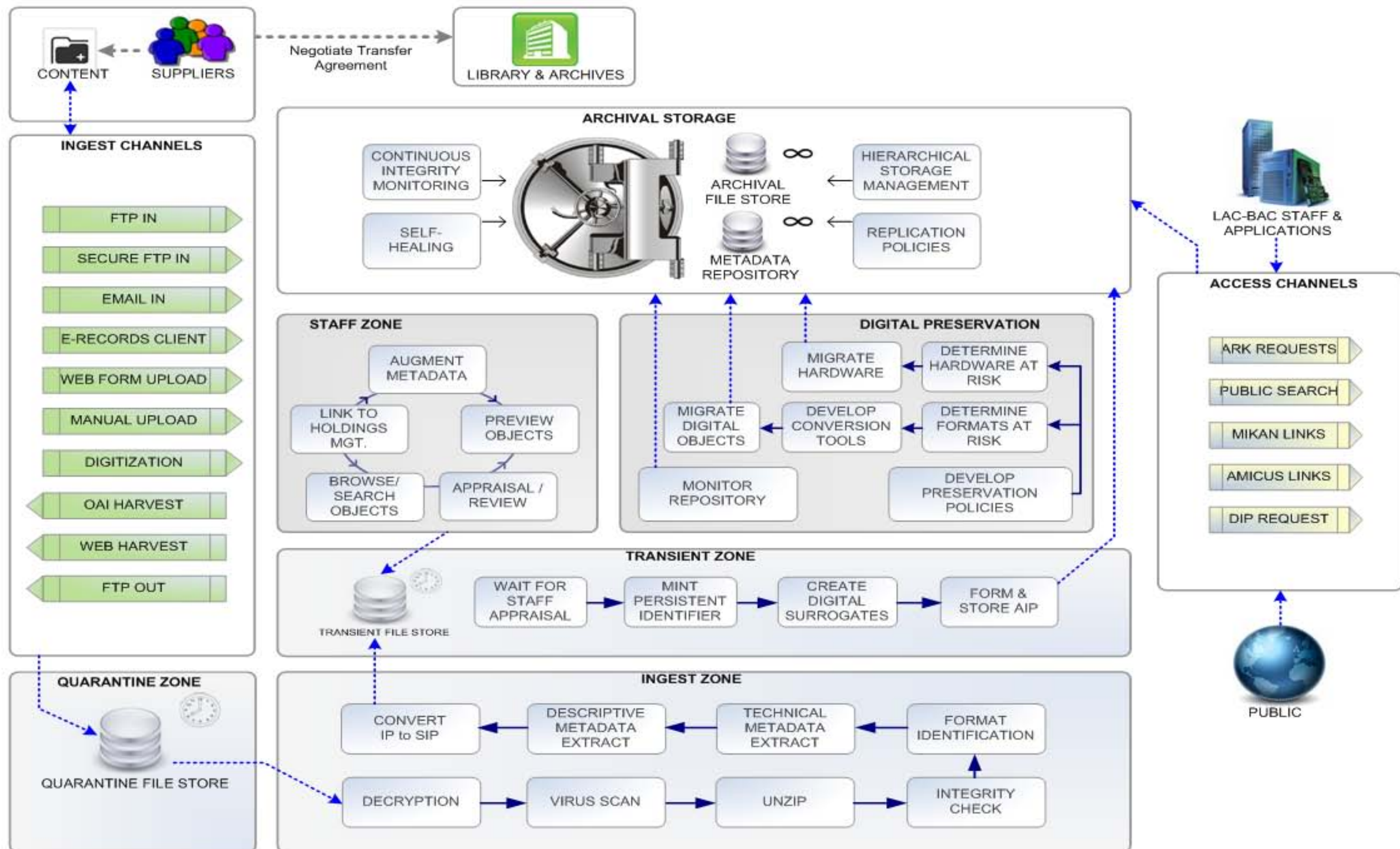
HIGH-LEVEL OAIS REFERENCE MODEL



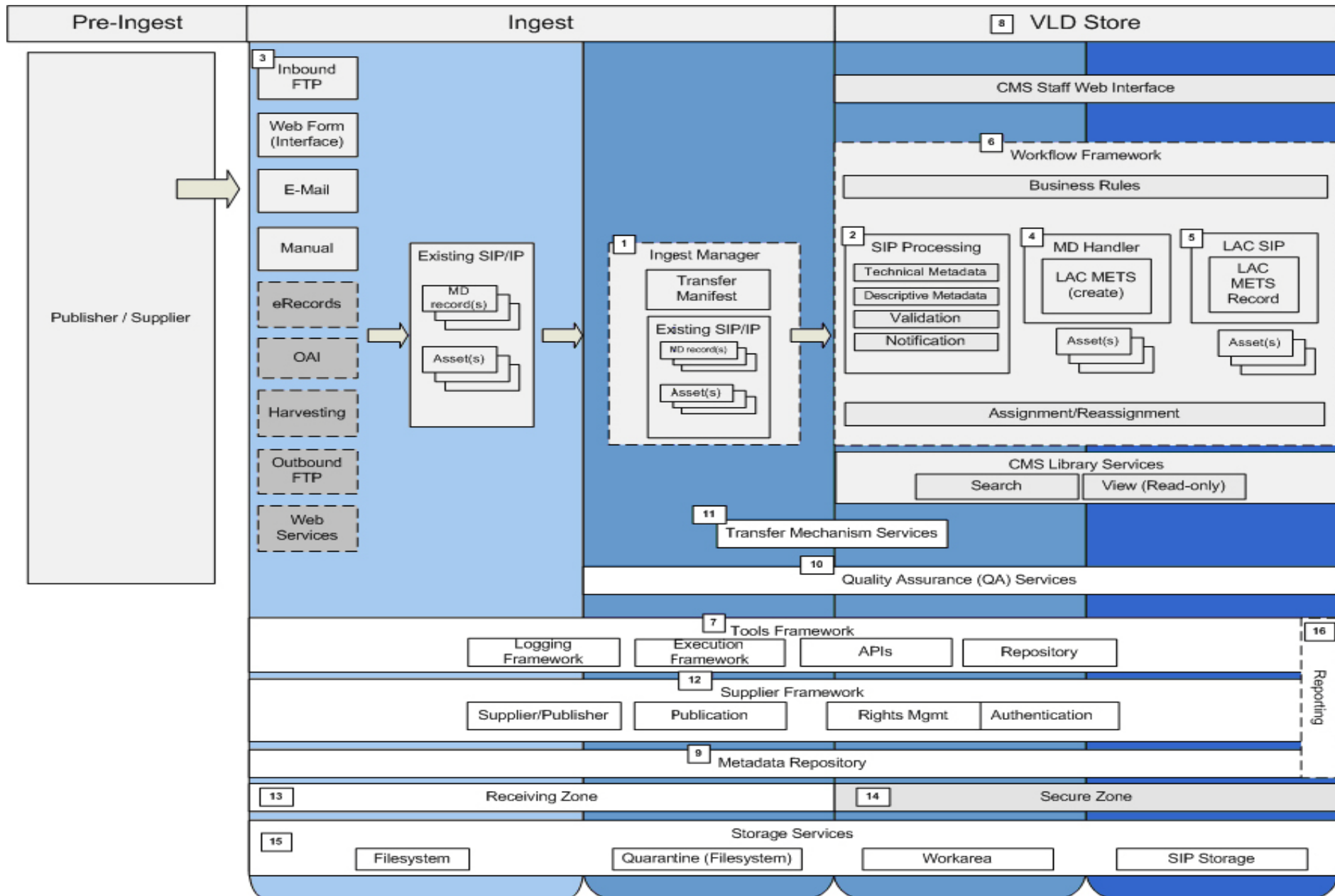
TDR 核心功能



LAC TDR OAIS 實施



虛擬上傳台構建

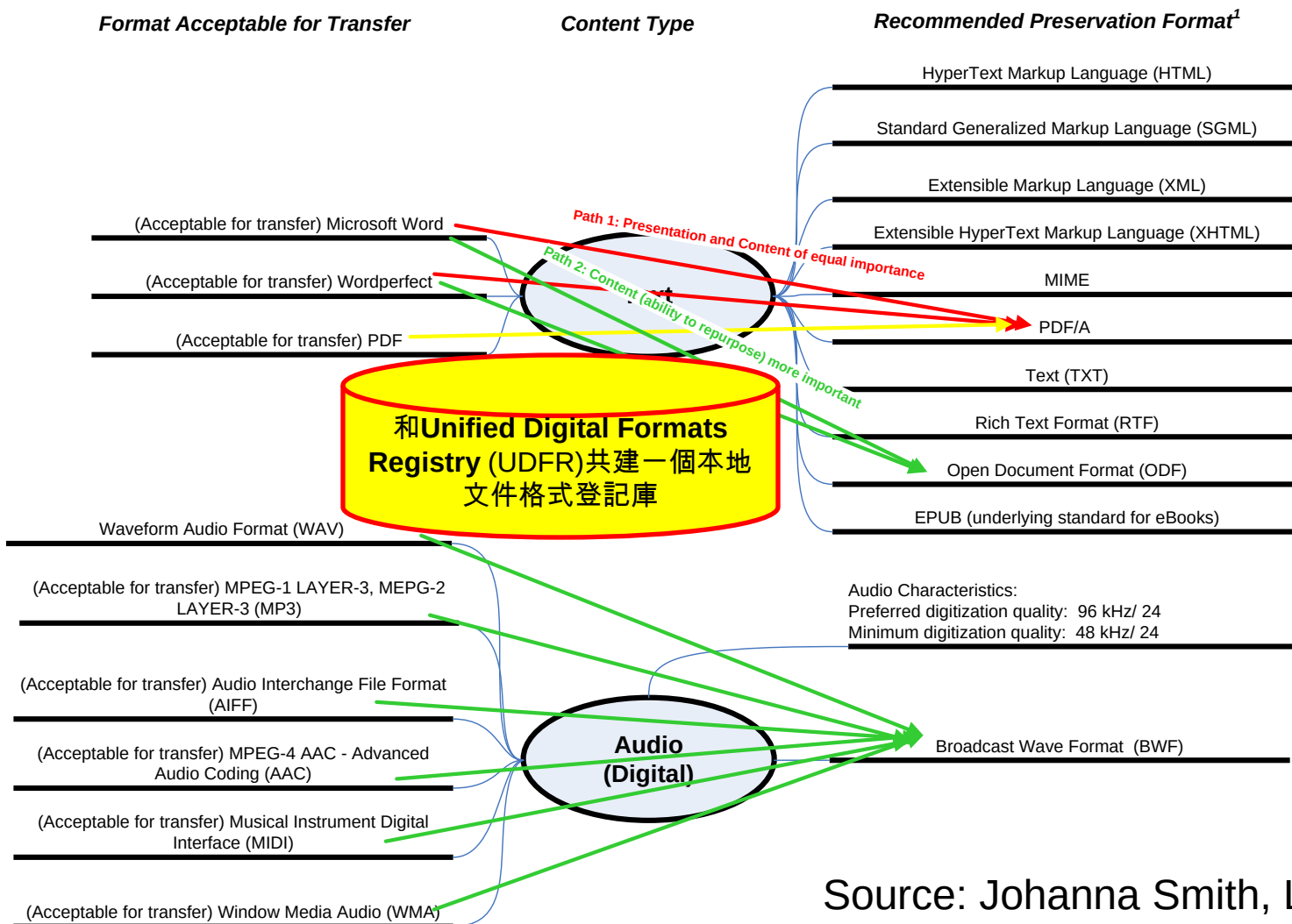




文件格式處理



Library and Archives Canada Format Guidelines – Normalization/Migration Paths

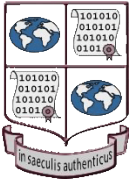


Source: Johanna Smith, LAC

¹ All recommended preservation formats are acceptable for transfer



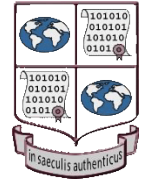
元數據處理



- 提取
 - 在流程步驟產生的元數據
- 整合
 - 描述數字體（數字文件）的元數據
- 生成
 - 符合LAC要求的語義和語法
- 相關元數據標準
 - METS - Metadata Encoding Transmission Schema;
 - MODS - Metadata Object Description Schema;
 - MARC - Machine Readable Cataloguing;
 - PREMIS - Preservation Metadata: Implementation Strategies;
 - METS Simple Rights Schema;
 - Dublin Core; and
 - The Government of Canada Records Management Metadata Std



源代碼開放協議和軟件



- PureFTP - Supports LAC security zone;
- JHOVE - Metadata extract + checksum creation;
- DROID - Digital object recognition;
- Pronom - File format registry;
- HTTP Apache, Tomcat
- LDAP - Authentication and role management
- Clam AV (anti-virus)
- NLNZ Metadata Extractor
- DROID/PRONOM (TNA)
- Jacksum
- md5sum
- XENA (NAA)
- File Information Tool Set (FITS)
- Library of Congress BagIt script



徵集原則之重要性



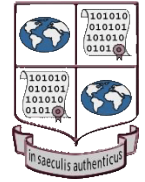
- 重要性 (significance)

- 形成者
- 來源
- 美學價值
- 稀缺性
- 研究價值
- 完整性
- 市場價值





RDIMS可處理的文件格式



- Audio Video Interleave (AVI)
- Bitmap (BMP)
- MS Word (DOC)
- MS Word Office Open XML with Macros Enabled(DOCM)
- Microsoft Open Office XML (DOCX)
- Plain text graph descriptions(DOT)
- Graphics Interchange Format (GIF)
- Hypertext Markup Language(HTM and HTML)
- Adobe InDesign Document (INDD)
- JPEG, Joint Photographic Experts Group (JPG)
- Javascript Source Code (JS)
- MS Access Database (MDB)
- MS Project File (MPP)
- MS Outlook (MSG)
- Portable Document Format (PDF)
- Portable Network Graphics (PNG)
- MS PowerPoint (PPT)
- MS PowerPoint Office Open XML (PPTX)
- Report (RPT)
- Rich Text (RTF)
- Tagged Image File Format (TIFF or TIF)
- Text (TXT)
- Virtual Mailbox (VMBX - what TRIM converts emails to)
- MS Visio Drawing (VSD)
- Windows Media Video (WMV)
- Word Perfect Document (WPD)
- MS Windows Media Redirector (WVX)
- MS Excel (XLS)
- MS Excel Office Open XML (XLSX)
- MS Excel Template (XLT)
- Data Compression file (ZIP)