# Digital forensics and Bodleian Electronic Archives & Manuscripts (BEAM)

Susan Thomas
Digital Archivist, Bodleian Library

# It all started when...

- Barbara Castle's archive accrued digital media in the middle of being catalogued
- While the Paradigm project was running
- Jeremy and the Digital Scriptorium assist with hard disk imaging
- Adopting imaging for digital media
- Enter the futureArch project to improve our capture capacity and extend our forensics work to analysis

# BEAM processing manual

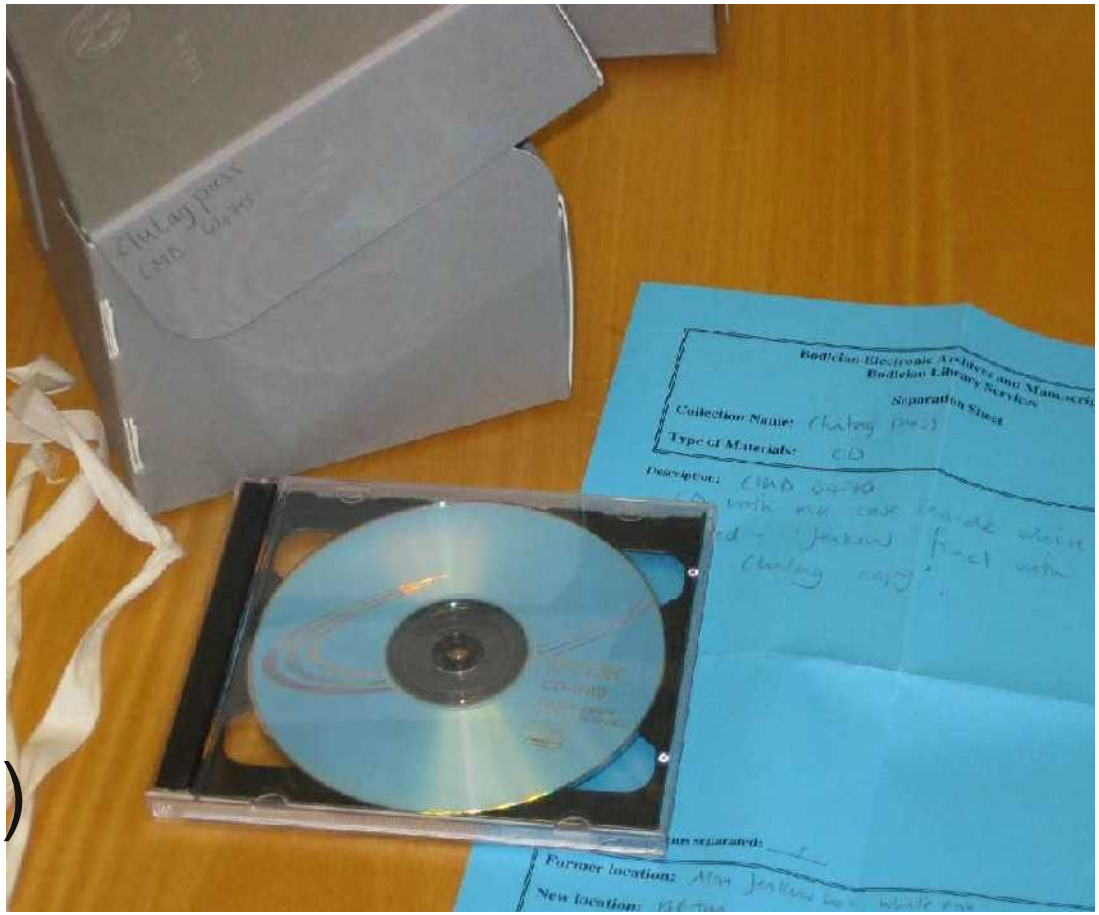
Last modified: 12. May 2010 08:08:08 PM

## Processing Manual for Bodleian Electronic Archives and Manuscripts (BEAM). 2009-

### Table of Contents

# Treating the media

- Record in CMD
- Separate/transfer to BEAM
- Repackage/label
- ID & inventory
- Photograph
- Retain (all? How long?)
- Cataloguing archivist does not have access to original media



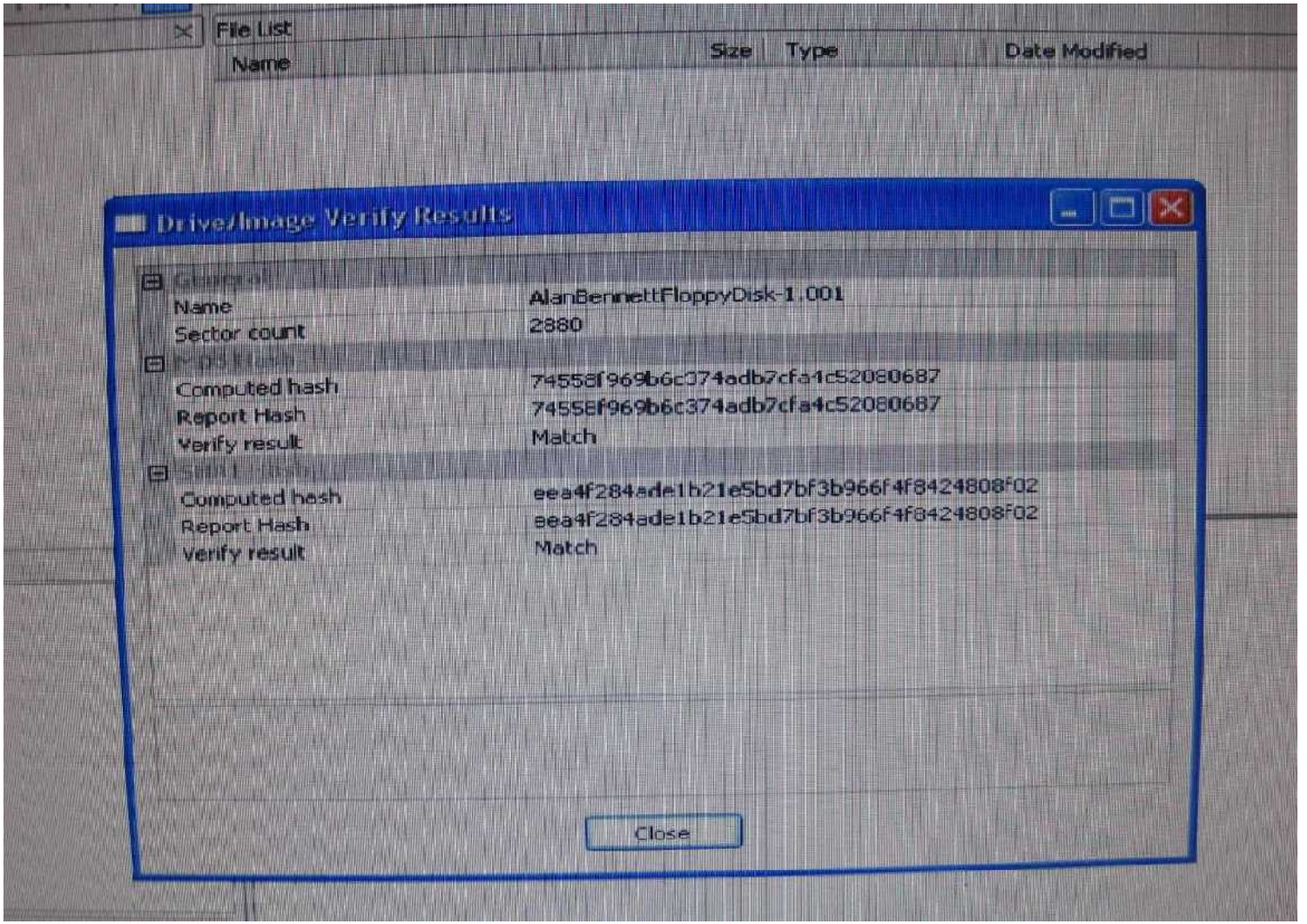
# BEAM processing room

Machines isolated from wider network; venue for capture, appraisal, cataloguing access review; multiple users; multiple platforms

# Imaging with FRED


- Used for most capture
- Supports wide range of media
- Most of our holdings are 3.5” floppies and optical media
- 5 HDDs to-date
- Use FTK Imager

# FTK Imager




Imaging at BEAM (also used on-site with donors

future Arch

# Mass imaging optical media

- Some colls with hundreds of optical media
- Not as straightforward as it seems.
- Used in conjunction with CD/DVD Inspector software

# Imaging issue log

| |
|---|
| Issue ID |
| Nature of issue |
| Affected material |
| Severity of issue |
| Date identified |
| Identified by |
| Proposed resolution |
| Individual/organisation or process passed to for resolution |
| Expected resolution date |
| Actual resolution date |
| Manner of resolution |
| Does issue suggest change of procedure is required? If yes provide details |

| ***Severity scale*** | |
|---|---|
| **1** | Catastrophic. Error causing irrevocable loss of collection materials and/or damage to BEAM hardware or software. |
| **2** | Severe. A road block with no workaround. Performance is so poor that the system is universally regarded as 'pitiful'. |
| **3** | Moderate. No permanent loss of data, but wasted time. There is a workaround to the problem. |
| **4** | Minor but irritating problem. Slows work down slightly. |
| **5** | Minimal error. |

future Arch

# Metadata for media/image

| ***Digital media (physical)*** |
| --- |
| Record ID |
| Accession ID (from CMD) |
| Date record created |
| Date record modified |
| Collection name |
| Disk number |
| Case or sleeve label |
| Physical label |
| Media type |
| Brand |
| Type |
| Capacity |
| Capacity used |
| Virtual disk label (as shown by the operating system once mounted) |
| file system |
| serial/batch number |
| Operating system requirements |
| Comments |
| Associated photos |

| ***Related disk image metadata*** |
| --- |
| Disk image type |
| Imaging software used |
| Imaging hardware configuration used |
| Disk image filename/s |
| Disk image number of parts |
| Disk image MDS hash |
| Date imaged |
| Date ingested |
| Imaging comments |

| ***Disk contents*** |
| --- |
| Covering dates |
| Scope and content |

# Legacy hardware/software



future Arch

# Guides - media/filesystems

**Media Recognition**

It is important to be able to recognise the media you come across in collections. Determining what type of data storage an object is can be easy, since floppy disks look very different from a CD. However, distinguishing between different types of floppy disk or CD can be more challenging. This guide aims to help you identify object characteristics, thereby allowing you to ascertain the specific type of data storage device in your possession.

1. **Floppy Disks:**
   - 8 inch disks
   - 3 inch disks
   - 5.25 inch disks
   - 3.5 inch disks
2. **Optical Media:**
   - Compact Disk
   - Audio and Data CD Comparison
   - CD-ROM
   - CD-R
   - CD-RW
   - DVD
   - HD DVD
   - Blu-Ray
3. **Hard Disk Drives:**
   - AT Attachment Interface
   - Serial ATA Interface
   - SCSI Interface
   - Serial Attached SCSI Interface
   - External Hard Disk Drives
4. **Iomega Zip Drive**
5. **Flash Media:**
   - USB flash drive
   - FireWire flash drive

Blog and document

**File Systems**

1. **Media Types**
2. **File Systems**
3. **File System Extensions**
4. **File Systems in Image Files**
5. **File System Tables**

File systems allocate space on media and organise the data in an easily identifiable and accessible way. The file system, or systems used in different media is determined by the type of media, its content and the platform required. Whilst each file system has different abilities and restrictions some systems have extensions which expand their capabilities in specific ways.

**1. Media Types**

Hard Drives

Hard drives require a file format that can support the constant rewriting and updating of files at incremental times. The File Allocation Table (FAT) was originally the primary file system for several operating systems including MS-DOS and Microsoft Windows. Moreover, FAT is compatible with virtually all operating systems, allowing data to be easily transferred between platforms. However, FAT has to some extent been superseded by the New Technology File System (NTFS), particularly on later Windows platforms from Windows 2000 onwards. This is primarily due to the FAT file size limit of 4GB and NTFS's greater reliability. However, NTFS is not as universal and while compatible with a number of operating systems, this does not include Mac OS.

Optical Media

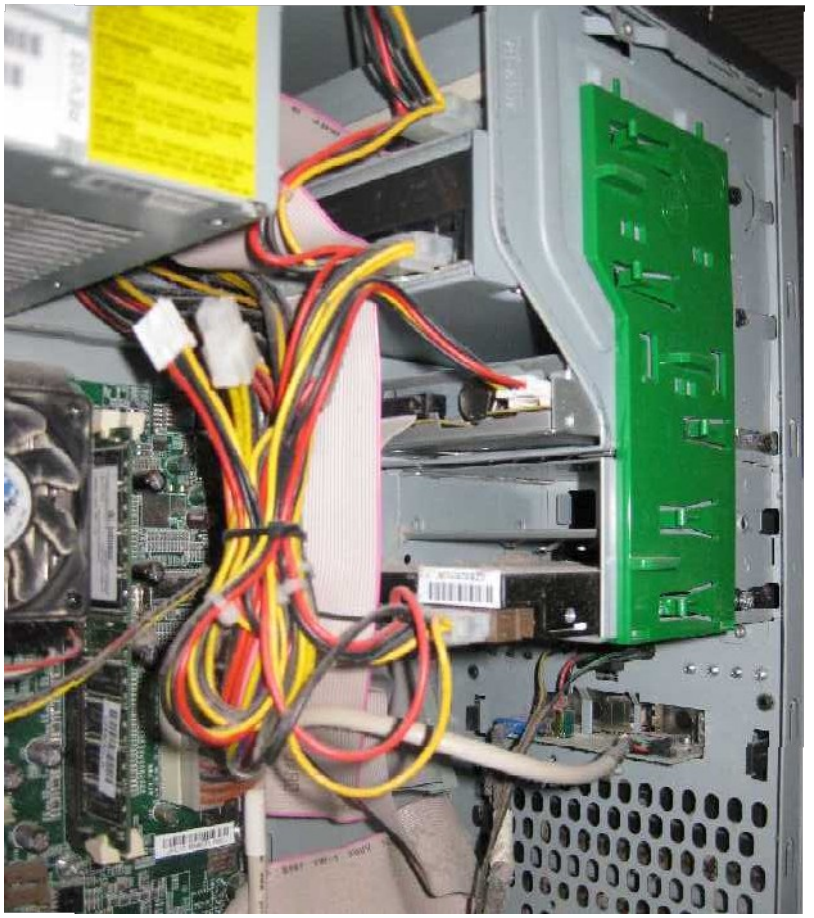
The different types of optical media can be separated into two categories: write-once and re-writeable media. Re-writeable optical media is capable of being supported by FATs and NTFSes, but unlike files on hard drives, optical media files are read-only. Also, there are alternative file systems specifically designed for optical media, therefore FATs and NTFSes are not the standard file systems.

The standard system for DVDs is Universal Disk Format (UDF). This system supports re-writeable media, therefore can also be used for CD-RW. Some file systems have size restrictions, but UDF can support files larger than 4GB, therefore is suited to DVD media, which tends to contain large files. However, write-once DVDs with smaller files can write ISO9660 along with UDF.

The CD-ROM standard is currently ISO9660, although this is intended to change to the latest ISO version, ISO13490.

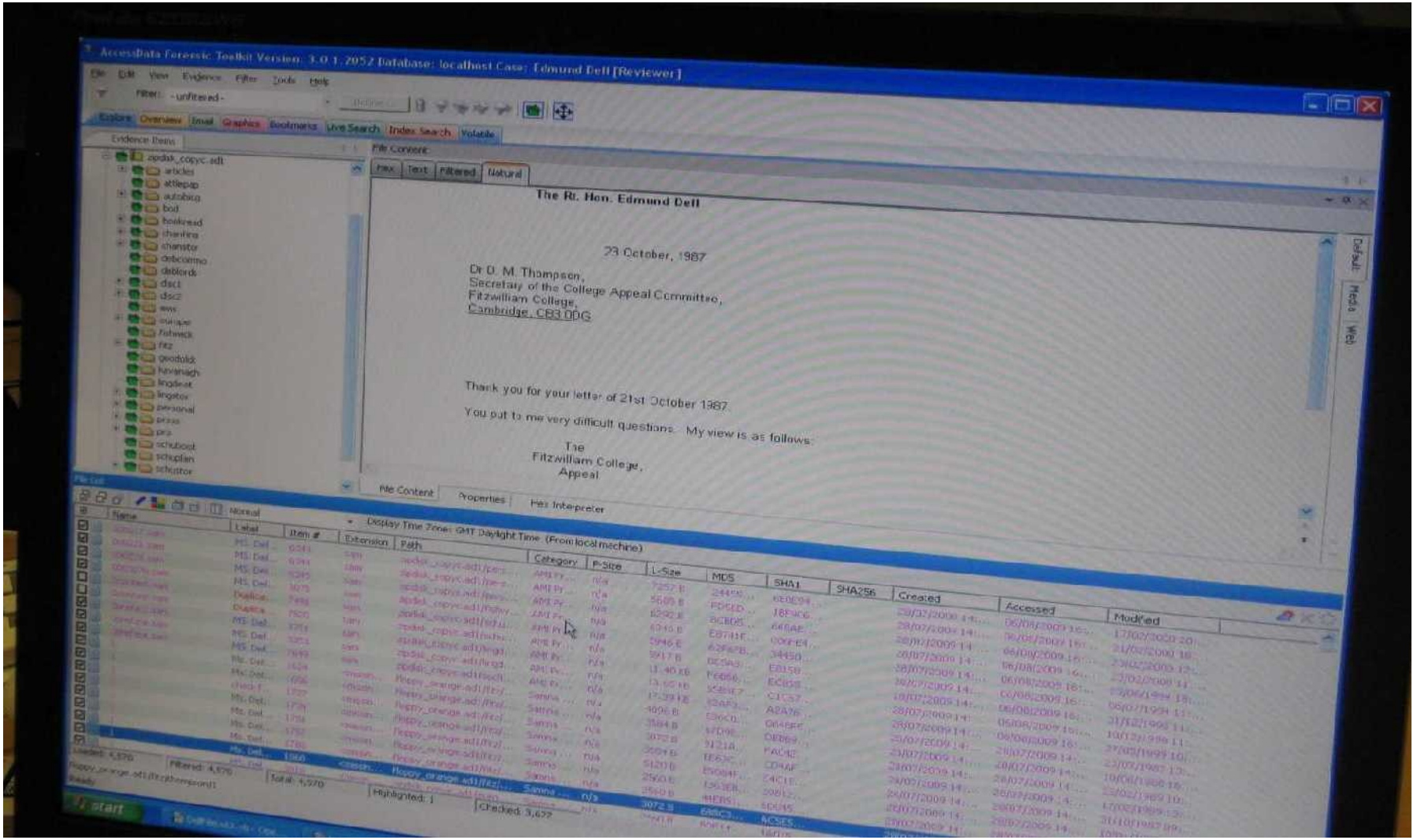
It is only optical media containing data files that has file systems; audio disks do not, they have tracks and track data instead. Originally the technology of the time was

# John Barton



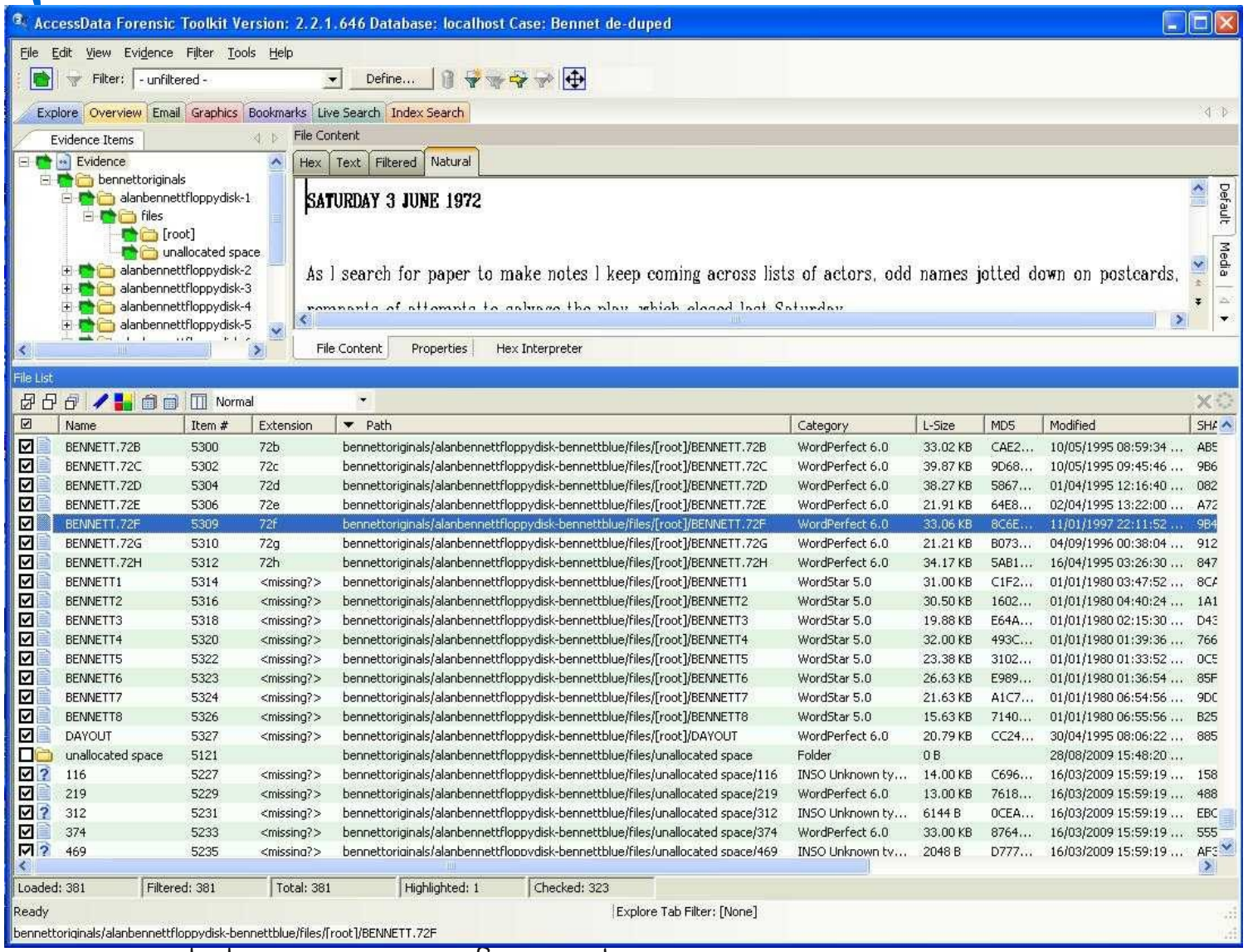
Recent stack discovery – Professor spec. in Roman Law at Oxford

# Appraisal with FTK



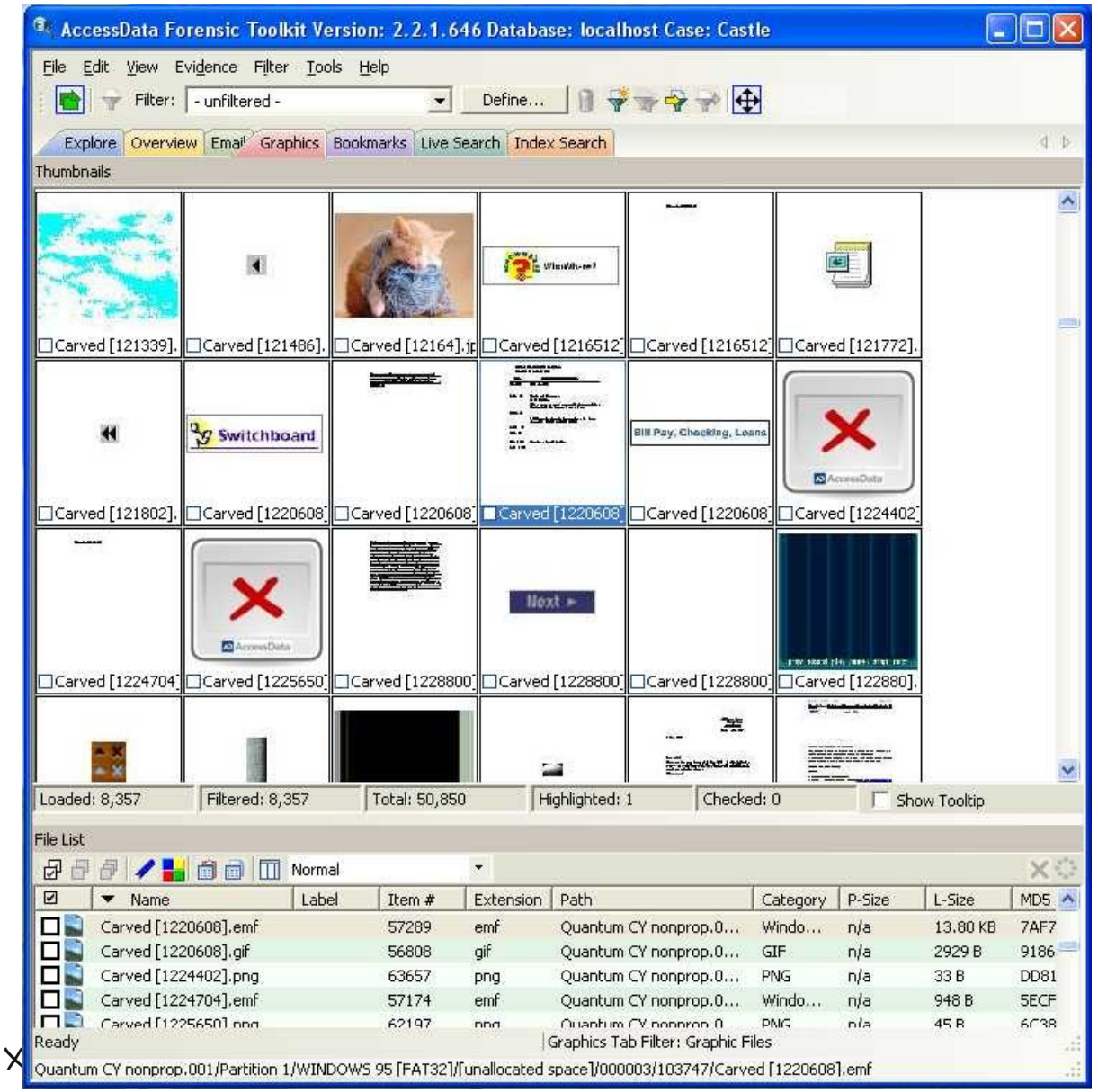

Preview, label, bookmark, id files v sware, dups

# Appraisal - Alan Bennett



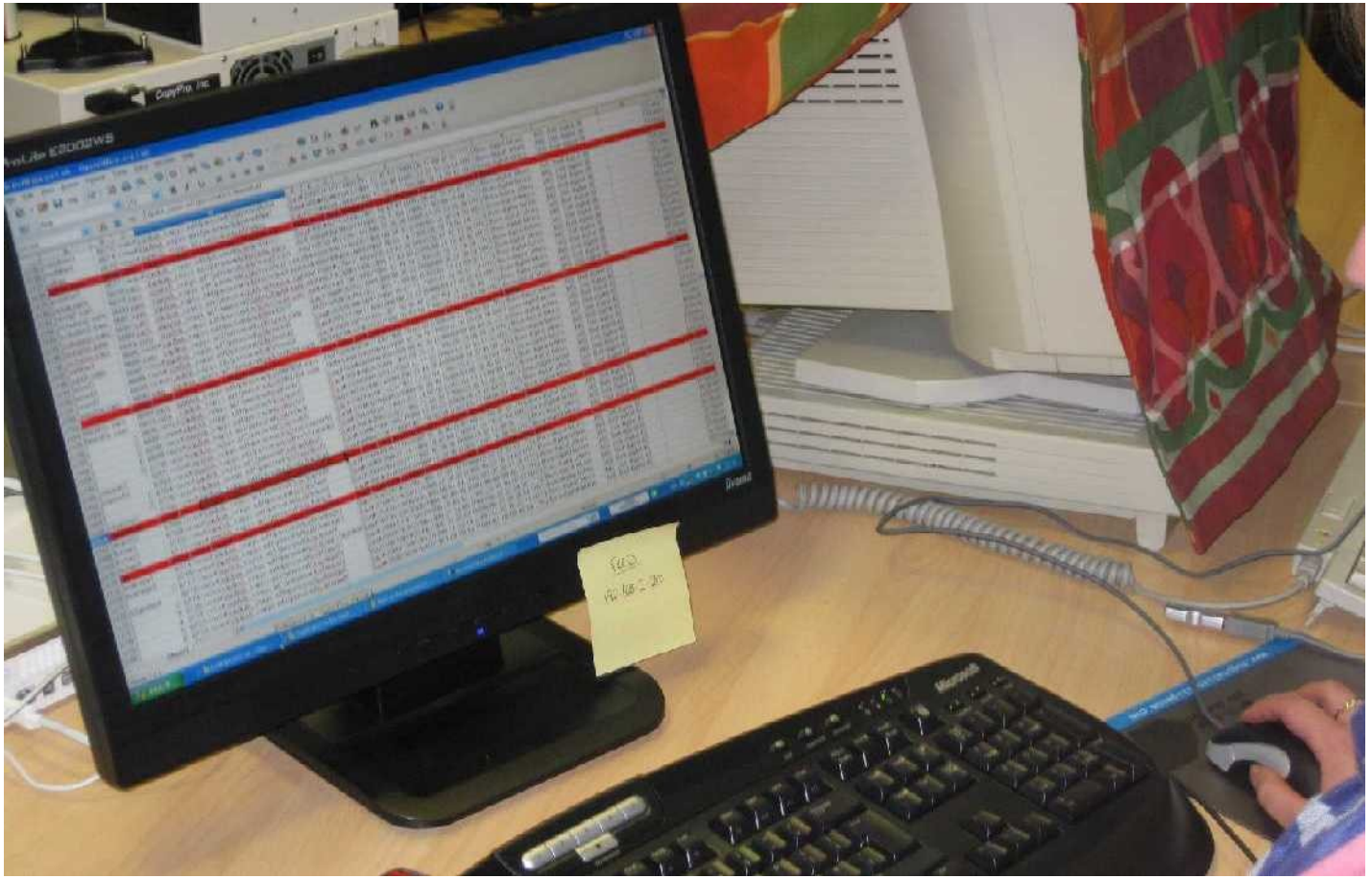

Several legacy WP formats in use

# Appraisal - Barbara Castle



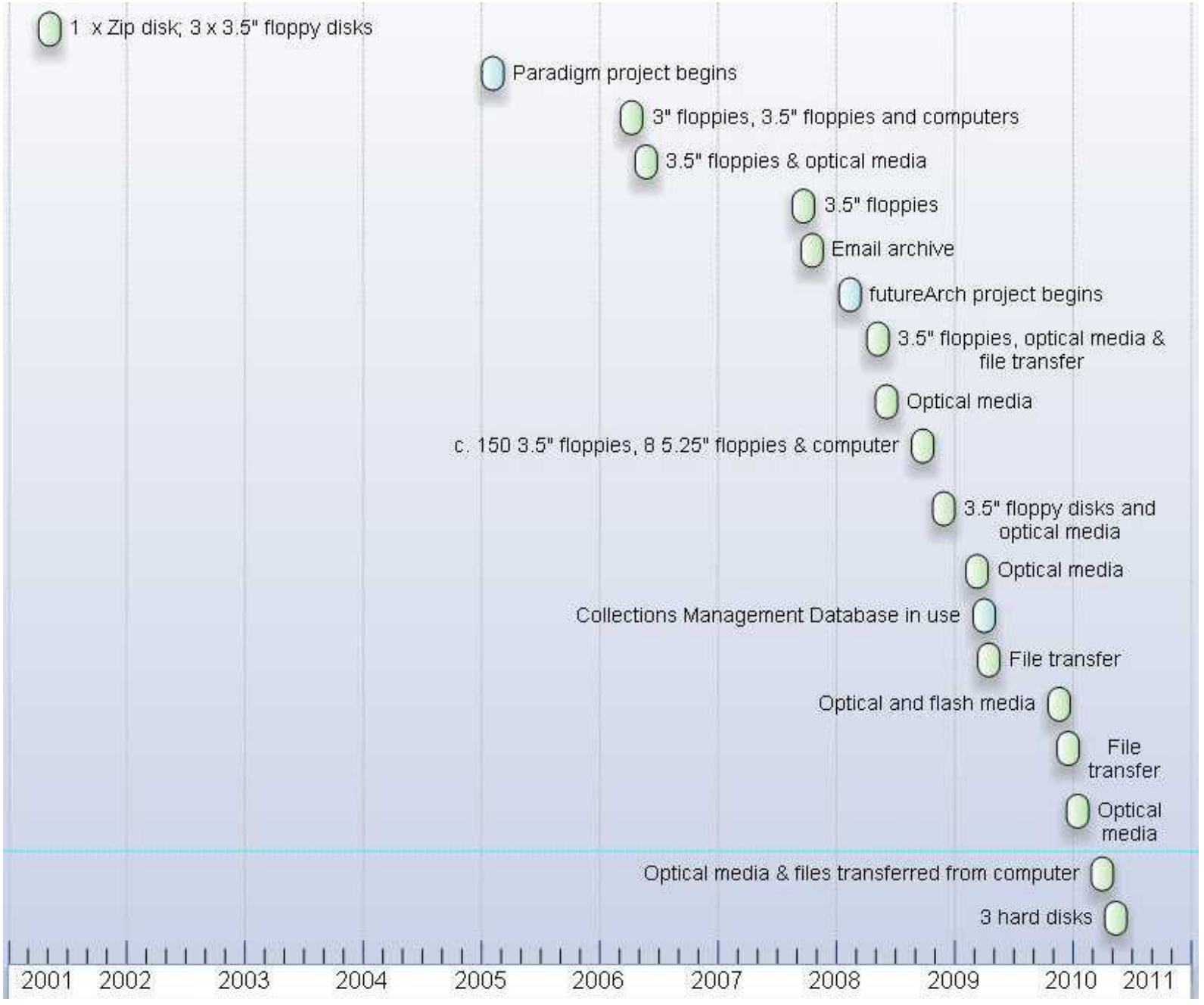

tex

# Digital object metadata



Metadata export From FTK used by cataloguer

# Growth of born-digital



# Many b-d colls in process

- Various phases:
  - Separation
  - Capture
  - Appraisal
  - Cataloguing
  - Preparation for publishing
- Supporting a number of project cataloguers/curatorial reviewers
- Scalability issues with current set-up

# Questions?

**Email me:**
susan.thomas@bodley.ox.ac.uk

**More online...**
Bodleian Electronic Archives & Manuscripts
http://www.ouls.ox.ac.uk/beam
futureArch project blog
http://futurearchives.blogspot.com/