

DIGITAL TOOLS FOR HUMANISTS
SUMMER SCHOOL 2023,
“Laboratorio di Cultura Digitale”, **University of Pisa**

Accessing and transforming historic media

Friday, 16 June 2023

PART I: Digital Challenges & Digital Forensics



Dr Seamus Ross,
Professor, Faculty of Information, University of Toronto

- All these Powerpoint Slides contain copyright material. The copyright in the lecture as a whole rests with the instructor. In many cases the copyright in the material within the lecture belongs to the instructor. In other cases it belongs to the University. In some cases the slides contain material the copyright of which belongs to other individuals, institutions, or entities. Where this is the case this material is *used* under one of the exemptions of the Canadian Copyright Act (<https://laws-lois.justice.gc.ca/eng/acts/C-42/index.html> (Links to an external site.)), such as, but not limited to, Fair Dealing (Sections 29, 29.1 and 29.2), and Education Exemptions (Sections 30.04, 29.4(1), 29.4(2), 29.5, 29.6, and 29.7). Where material is used under such exemptions as Copyright Act Sections 29.1 and 29.2 the source of the work and the author are clearly stated on the relevant slides.
- Provision of access to these Powerpoint Slides themselves is done under the exemption granted in Section 30.01 (Communication by Telecommunication). This exemption requires that you delete these slides within 30 days of the end of the course.

➤ Welcome and Introduction

➤ Who am I

➤ Overview of the day

- Lectures in Morning
- Interactive Activities & Experimentation in Afternoon

Timetable & What we will cover

- 09:00 – 10:30 Digital Resources, Preservation, and Forensics
- 10:30 – 11:00 Break
- 11:00 – 12:30 Explore Digital Materials
- 12:30 – 14:00 Lunch
- 14:00 – 15:30 Web Archiving and Web Archives
- 15:30 – 16:00 Break
- 16:00 – 17:00 Story-telling with Web Archives
- 17:00 – 17:30 Discussion

PRESERVATION ---

- ❖ Without preservation, vast amounts of data are at risk of being lost or simply disappearing as platforms evolve or change, potentially depriving historians and researchers of important sources of information and distinct perspectives.
- ❖ Preserving content not merely a technical issue of archiving
- ❖ We have an ethical responsibility of safeguarding tangible cultural heritage of the digital era.
- ❖ Access is hugely technically challenging – and we are just coming to terms with that.

- ❖ So we will talk about preservation and then ago digital forensics.

Williams v. Sprint/United Mgmt. Co.

230 F.R.D. 640 (D. Kan. 2005)
Decided Sep 29, 2005

Andrew H. McCue, Martin M. Meyers, The Meyers Law Firm, LC, Dennis E. Egan, Stephen J. Dennis, Bert S. Braud, The Popham Law Firm, P.C., Kansas City, MO, Daniel B. Kohrman, Laurie A. McCann, Thomas W. Osborne, AARP Foundation Litigation, Washington, DC, Kenneth B. McClain, Humphrey, Farrington & McClain, Gene P. Graham, Jr., Deborah J. Blakely, White, Allinder, Graham & Buckley LLC, Independence, MO, Dirk L. Hubbard, John M. Klamann, Klamann & Hubbard, P.A., Overland Park, KS, for Plaintiffs.

Michael H. Witt, pro se.

Sandra M. Cuskaden, pro se.

Maxine L. Coffey, pro se.

Chris R. Pace, Jill S. Ferrel, Stephany J. Newport, Overland Park, KS, Christine F. Miller, Harry B. Wilson, Jr., James F. Monafó, Joseph H. Guffey, Michael F. Jones, Tamara M. Spicer, Husch & Eppenberger, LLC, St. Louis, MO, David A. Schatz, Kara Marie Dorssom, David M. Eisenberg, John J. Yates, Patrick F. Hulla, Philip R. Dupont, Husch & Eppenberger, LLC, Kansas City, MO, for Defendant.

⁶⁴¹ *640 *641

MEMORANDUM AND ORDER

WAXSE, United States Magistrate Judge.

Plaintiff Shirley Williams filed this suit on

during a reduction-in-force (RIF). Currently, 1727 plaintiffs remain in the case out of the 2354 plaintiffs who opted into this provisionally certified collective action pursuant to 29 U.S.C. § 216(b). The parties are presently engaged in discovery concerning the merits of Plaintiffs' pattern and practice allegations. This matter is presently before the Court on Defendant's Response to the Court's July 12, 2005 Order (doc. 3037), which ordered Defendant to show cause why it should not produce electronic Microsoft Excel spreadsheets in the manner in which they were maintained and why it should not be sanctioned for "scrubbing" the metadata and locking certain data on the electronic spreadsheets prior to producing them to Plaintiffs without either the agreement⁶⁴² of the parties or the approval of the Court.

I. Background Information

Plaintiff Williams commenced this action in April 2003, and, to date, the docket reflects that over 3300 pleadings and orders have been filed. The case is assigned to Chief Judge John W. Lungstrum but is referred to the undersigned Magistrate Judge for pretrial proceedings, including discovery. Due to the highly contentious nature of this litigation, the Magistrate Judge has conducted discovery conferences twice a month since March 2005 to resolve discovery issues identified by the parties. One of the ongoing discovery disputes has been Defendant's



A Public Record at Risk: The Dire State of News Archiving in the Digital Age

By Sharon Ringel and Angela Woodall
MARCH 28, 2019



SHARE ON TWITTER



SHARE ON FACEBOOK



EMAIL THIS STORY

[Executive Summary](#) | [Introduction](#) | [Methodology](#) | [Perceptions of News Preservation](#) | [The Intricacy of Archiving Digital News](#) | [Approaches to Preservation](#) | [Conclusions](#) | [Appendix: Additional Resources](#) | [Acknowledgments](#) | [Citations](#)

EXECUTIVE SUMMARY

This research report explores archiving practices and policies across newspapers, magazines, wire services, and digital-only news producers, with the aim of identifying the current state of archiving and potential strategies for preserving content in an age of digital distribution. Between March 2018 and January 2019, we conducted interviews with 48 individuals from 30 news organizations and preservation initiatives.

What we found was that the majority of news outlets had not given any thought to even basic strategies for preserving their digital

ABOUT THE TOW CENTER FOR DIGITAL JOURNALISM

The Tow Center for Digital Journalism at Columbia Graduate School of Journalism is a research center exploring the ways in which technology is changing journalism, its practice and its consumption – as we seek new ways to judge the reliability, standards and credibility of information online.

TOW REPORTS

FRIDAY, SEPTEMBER 4TH, 2014

[Guide to Native Advertising](#)

Ava Strah

Sharon Ringel and Angela Woodall, 2019 (March 28), *A Public Record at Risk: The Dire State of News Archiving in the Digital Age*, https://www.cjr.org/tow_center_reports/the-dire-state-of-news-archiving-in-the-digital-age.php

CONCLUSION

Preservation is a multi-pronged process that technology can assist. But ultimately, maintaining news for the future depends on deliberate practices that involve planning around tasks such as migrating content to new formats, assigning consistent metadata, and indexing. Like most media organizations, the individuals interviewed for this report care about maintaining access to the news. But they are at a loss for what to do and may doubt their ability to prioritize preservation.

HOW IS THE DIGITAL DECADENCE

The statement said Bethell had used his official email account as well as his private email account to send and receive emails relevant to the contracts, and that he had also used his mobile phone for SMS and WhatsApp messages. But it said Bethell had confirmed that about six months ago his phone was broken and replaced and that his new phone did not contain the phone data.

Government lawyers revealed Bethell had not been issued with a “preservation notice” requiring him to save documents because ministers’ official correspondence was routinely saved as a matter of course. However, this did not cover government business conducted by private means.

Bethell is already under investigation by the Information Commissioner’s Office (ICO) over the use of private emails for government business, prompted by revelations that his former boss **Matt Hancock** was using a private account at the height of the pandemic.

News Opinion Sport Culture Lifestyle More

UK ► UK politics Education Media Society Law Scotland Wales Northern Ireland

Health policy


Covid contracts: minister replaced phone before it could be searched

Government expected to disclose James Bethell's correspondence relating to award of £85m of contracts for Covid tests

- Coronavirus - latest updates
- See all our coronavirus coverage

Rowena Mason Deputy political editor

Wed 4 Aug 2021 13:11 BST



▲ A government lawyer's witness statement said Bethell replaced his phone in early 2021 and it may no longer be possible to retrieve information about dealings with Abingdon. Photograph: Roger Harris/UK Parliament


Labour has called for an inquiry into the use of WhatsApp within the government, after it emerged a health minister replaced his mobile phone before it could be searched for information relevant to £85m of deals that are subject to a legal challenge.

James Bethell, who oversaw the award of Covid contracts, is one of those under scrutiny over the way deals for personal protective equipment (PPE) and tests were allocated at the height of the pandemic.

As part of legal proceedings issued by the Good Law Project, the government is expected to disclose Lord Bethell's correspondence including by email, WhatsApp and SMS relating to the award of £85m of contracts for antibody tests to Abingdon Health.

The secretary of state has a responsibility to preserve and search documents for information relevant to the case from the point at which judicial review proceedings were issued in late 2020, under the government's "duty of candour".

Advertisement



Webinar: Fast-Tracking COVID-19 Research

<https://www.theguardian.com/politics/2021/aug/04/covid-contracts-minister-lord-bethell-replaced-phone-before-it-could-be-searched>

News Opinion Sport Culture Lifestyle More

UK politics Education Media Society Law Scotland Wales Northern Ireland


Politics

UK government admits ministers can use self-deleting messages

Exclusive: Civil servants also able to delete messages instantly, as fears grow about accountability

non-ideologue Legal & HR correspondent

13 Jun 2021 15:00 BST



▲ Transparency campaigners say it is vital to preserve details of government decision-making. Illustration: David Greig

Ministers and civil servants are allowed to set messages to delete instantly, the government has admitted, amplifying concerns about its transparency and accountability.

The confirmation comes as concerns grow that self-destructing messages are being used to avoid scrutiny of decision-making processes, including on key issues such as the government's coronavirus response.


A letter from the Department for Digital, Culture, Media and Sport (DCMS) sent to the *Citizens*, a non-profit organisation, in response to a freedom of information request and seen by the Guardian, says: "Instant messaging (through Google Workspace) may be used in preference to email for routine communications where there is no need to retain a record of the communication."

"Chat messages are retained for 90 days to provide staff with the opportunity to record any substantive conversations, after which time they are permanently deleted. Users can also switch history off, meaning messages will be deleted once a chat session has finished."

The letter says that the use of other instant messaging platforms is managed through DCMS's use of collaboration tools guidance, which was also provided but contains no reference to - or restriction on - self-destructing message services.

Transparency campaigners have expressed alarm at a culture of "government by WhatsApp". The *Citizens* has threatened legal action, saying use of such functions makes it impossible to carry out required legal checks about whether a message should be archived for posterity. Information that could be useful to a public inquiry, or otherwise fall within the scope of an FOI request, may be lost as a result.

Advertisement



Pay Less. Get More. Learn More

<https://www.theguardian.com/politics/2021/jun/13/uk-government-admits-ministers-can-use-self-deleting-messages>

European Commission defiant over Von der Leyen's Pfizer texts

EU executive defends its right not to keep records of president's messages after rebuke from ombudsman



Ursula von der Leyen at the G7 summit in Germany this week. Photograph: Ludovic Marin/AP/Getty Images

The European Commission has said it cannot and does not need to find text messages that its president, **Ursula von der Leyen**, exchanged with the boss of Pfizer at the height of the pandemic, fuelling its dispute with the EU's internal watchdog.

The commission's defence of its right not to keep records of Von der Leyen's text messages was published on Wednesday by the EU's official watchdog, the European Ombudsman, which conducted an initial investigation after a complaint about transparency.

In a sharp rebuke issued in January, the ombudsman, Emily O'Reilly, **accused** the EU executive of **maladministration**. She said text messages concerning EU policies and decisions should be treated as documents subject to EU transparency rules.

<https://www.theguardian.com/world/2022/jun/29/european-commission-defiant-over-von-der-leyens-pfizer-texts>

Decision on the European Commission's refusal of public access to text messages exchanged between the Commission President and the CEO of a pharmaceutical company on the purchase of a COVID 19 vaccine (case 1316/2021/MIG)

DECISION

CASE 1316/2021/MIG - OPENED ON Thursday | 16 September 2021 - **RECOMMENDATION ON** Wednesday | 26 January 2022 - **DECISION ON** Tuesday | 12 July 2022 - **INSTITUTION CONCERNED** European Commission (Maladministration found)

<https://www.ombudsman.europa.eu/en/decision/en/158295>

EU watchdog: Failure to find VDL-Pfizer texts is 'wake-up call'

The European Ombudsman issued recommendations for preserving texts after so-called 'leakgate'



EU Ombudsman Emily O'Reilly

July 12, 2022 | 12:47 PM

en/decision/158295

The EU's institutional watchdog again checked the European Commission Thursday for failing to provide or even seriously search for text messages that President Ursula von der Leyen allegedly exchanged with a Pfizer executive.

In reflecting on earlier findings of maladministration, European Ombudsman Emily O'Reilly

Last month, the Commission replied that such “short-lived, ephemeral documents are not kept.” At the same time, the Commission acknowledged that these texts could fall under the definition of documents that need to be kept.

<https://www.politico.eu/article/eu-watchdog-european-ombudsman-emily-oreilly-failure-to-find-vdl-pfizer-texts-is-wake-up-call/>

Regulation 1049/2001, which sets out the public's right to access EU documents, defines a document as *“any content whatever its medium (written on paper or stored in electronic form or as a sound, visual or audiovisual recording) concerning a matter relating to the policies, activities and decisions falling within the institution's sphere of responsibility”*.

The recommendations ([see full list here](#)) say that:

- Work-related text and instant messages should be recognised as EU documents.
- Technological solutions should be put in place to enable the easy recording of such messages.
- Staff should have clear guidance on how such messages should be recorded.
- Requests for public access to documents that could cover text messages should be dealt with in a way that considers all locations where such messages might be stored.

<https://www.ombudsman.europa.eu/en/press-release/en/158303>

Print subscriptions | Sign in | Search jobs | Search | International edition

Support the Guardian
Fund independent journalism with €5 per month
Support us →

The Guardian

News | **Opinion** | Sport | Culture | Lifestyle | More


The Guardian view | Columnists | Cartoons | Opinion videos | Letters

Opinion
Coronavirus

This article is more than 3 months old

What do Matt Hancock's WhatsApp messages show? Not what the Telegraph wants us to see

Devi Sridhar




Advertisement

Wed 1 Mar 2023 13:38 GMT

The messages confirm what advisers like me knew: that during the pandemic, ministers weren't following the science' at all

Prof Devi Sridhar is chair of global public health at the University of Edinburgh

291



<https://www.theguardian.com/commentisfree/2023/mar/01/matt-hancock-whatsapp-messages-telegraph-covid-pandemic>

Boris Johnson says he has handed over Covid WhatsApps

1 June



Coronavirus public inquiry



Mr Johnson's spokesperson said the cabinet office has had access to the unredacted documents for "months"

Former prime minister Boris Johnson says he has given the UK government all the WhatsApp messages and notebooks demanded by the Covid-19 Inquiry.

Mr Johnson is urging the government to hand the material to the inquiry in full without redactions.

The inquiry, which begins public hearings in two weeks, is investigating how ministers handled the pandemic.

The government has so far refused to hand over material it does not consider relevant.

<https://www.bbc.com/news/uk-politics-65770586>


NEWS CULTURE MUSIC PODCASTS & SHOWS SEARCH

NATIONAL

What the subpoena for the Secret Service's erased texts means for the Jan. 6 probe

Updated July 17, 2022 11:01 AM ET

JULIANA KRIE



Chairman Bennie Thompson, D-Miss., leads as the House select committee investigating the Jan. 6 attack on the U.S. Capitol in a hearing at the Capitol on Tuesday.

The Secret Service has been subpoenaed in the ongoing probe into the riot on the U.S. Capitol on Jan. 6, 2021, a move that a former federal prosecutor calls aggressive and significant.

The House select committee leading the investigation is asking the federal agency to turn over reportedly deleted text messages from the days surrounding the attack as well as any relevant action reports. The Secret Service has until Tuesday to produce agents' phone records, that some believe may shed light on President Donald Trump's actions during the riot.

<https://www.npr.org/2022/07/16/1111857502/jan-6-panel-subpoenas-secret-service-erased-texts>


NEWS CULTURE MUSIC PODCASTS & SHOWS SEARCH

POLITICS

The National Archives is looking into reports that the Secret Service deleted texts

Updated July 19, 2022 6:27 PM ET

WASHINGTON DESK



The chief records officer of the U.S. government has asked the Secret Service to determine whether any of its text messages on Jan. 6 and Jan. 8, 2021, were improperly deleted.

Reports that the Secret Service deleted text messages related to the Jan. 6, 2021 Capitol attack have caught the attention of the chief records officer of the U.S. Government.

That officer, Laurence Brewer, said in a letter to the Department of Homeland Security on Tuesday that the National Archives and Records Administration "has become aware of the potential unauthorized deletion of United States Secret Service (Secret Service) text messages" that were dated Jan. 5 and Jan. 6, 2021.

ud-2022-0054-dhs-uss-opens (Contributed by NPR Politics (NPR))

Office of the Chief Records Officer of the U.S. Government

<https://www.npr.org/2022/07/19/112288183/secret-service-deleted-texts-national-archives-letter>

What might happen if the Secret Service doesn't comply

If the Secret Service is unable to turn over the deleted messages, the next major question will be if that's intentional, according to Ankush Khardori, a former federal prosecutor.

"There's a big factual difference between the inadvertent loss of communications and a deliberate effort to delete these communications," Khardori told NPR. "Really what you would want to know is what are the Secret Service's record keeping rules, regulations and protocols, did anyone run afoul of them and in the worst case scenario, did someone make a deliberate effort to destroy these communications."

To get to the bottom of what happened, he said Congress may launch an investigation into the Secret Service's record keeping system or call for members of the Secret Service to testify.



Khardori said it's too soon to tell but the Jan. 6 committee's probe into an agency of the executive branch is significant.

"It's not that unusual for Congress to seek information from the executive branch, including through subpoenas, but this is different because it's more public, more assertive, more aggressive and it suggests concern among at least some members of the committee that the Secret Service hasn't been forthright with their answers," he said.

The next Jan. 6 committee hearing is scheduled at 8 p.m. ET Thursday, with a specific focus on Trump's failure to act to stop the insurrection.




OFFICE OF INSPECTOR GENERAL
Department of Homeland Security
Washington, DC 20528 / www.oig.dhs.gov

July 13, 2022

The Honorable Gary C. Peters, Chairman
The Honorable Rob Portman, Ranking Member
Committee on Homeland Security and Governmental Affairs
United States Senate
Washington, DC 20510

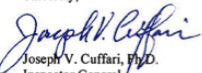
The Honorable Bennie G. Thompson, Chairman
The Honorable John Katko, Ranking Member
Committee on Homeland Security
U.S. House of Representatives
Washington, DC 20515

Dear Chairman Peters, Ranking Member Portman, Chairman Thompson, and Ranking Member Katko:

The Inspector General Act provides that an Inspector General shall have timely access to all Department records and other materials related to Department programs overseen by the Inspector General. I am writing to offer a briefing regarding ongoing records access issues that we have been experiencing with the Department of Homeland Security. There are two specific issues that we would like to address with your respective committees.

First, the Department notified us that many U.S. Secret Service (USSS) text messages, from January 5 and 6, 2021, were erased as part of a device-replacement program. The USSS erased those text messages *after* OIG requested records of electronic communications from the USSS, as part of our evaluation of events at the Capitol on January 6. Second, DHS personnel have repeatedly told OIG inspectors that they were not permitted to provide records directly to OIG and that such records had to first undergo review by DHS attorneys. This review led to weeks-long delays in OIG obtaining records and created confusion over whether all records had been produced.

I will make myself and my staff available at your convenience. Please call me with questions, or your staff may contact Kristen Fredricks, Chief of Staff, at (202) 981-6000.

Sincerely,

Joseph V. Cuffari, Ph.D.
Inspector General

<https://www.documentcloud.org/documents/22087607-dhs-oig-letter-to-hsgac-chs-071322>

RECORDS FOR PRESERVATION

Archives increasingly acquire records in digital form:

- Email
- Databases
- Documents
- Images
- Audio recordings
- Sensor data
- Social Media (e.g., Twitter, Youtube, Facebook, Instagram)
- Games
- Software
- Algorithms

SALMAN RUSHDIE FINDING AID

WOODRUFF

BUSINESS

HEALTH SCIENCES

LAW

MARBL

OXFORD COLLEGE

THEOLOGY

EMORY



EMORY
LIBRARIES &
INFORMATION
TECHNOLOGY

EmoryFindingAids

► SCHOOLS ► LIBRARIES ► RESOURCES

HOME

ADVANCED SEARCH

REQUEST MATERIALS

CONTRIBUTORS

SEARCH TIPS

FINDINGAID FAQ

FEEDBACK

Search EmoryFindingAids

only collections with digital resources

SEARCH ►

Search for items in the container list of **Salman Rushdie papers, 1947-2008**

only digital resources

SEARCH THIS FINDING AID ►

Send us your feedback.

FEEDBACK ►

REQUEST FROM THIS COLLECTION

RUSHDIE, SALMAN.

SALMAN RUSHDIE PAPERS, 1947-2012

Emory University
Stuart A. Rose Manuscript, Archives, and Rare Book Library
Atlanta, GA 30322
404-727-6887
rose.library@emory.edu
Permanent link: <http://pid.emory.edu/ark:/25593/8zv36>

[Printable PDF](#)

TABLE OF CONTENTS

- > [Descriptive Summary](#)
- > [Administrative Information](#)
- > [Collection Description](#)
- > [Selected Search Terms](#)

DESCRIPTION OF SERIES

- > [Series 1: Journals, appointment books, and notebooks, 1974-2003](#)
- > [Series 2: Writings by Rushdie, 1964-2006](#)
- > [Series 3: Writing by others, 1983-2005](#)
- > [Series 4: Correspondence, 1974-2006](#)
- > [Series 5: Personal papers, 1964-2005](#)
- > [Series 6: Subject files, 1976-2006](#)
- > [Series 7: Photographs, circa 1947-2006](#)
- > [Series 8: Printed material, 1980-2008](#)
- > [Series 9: Memorabilia, 1982-1999](#)
- > [Series 10: Audiovisual, 1981-2008](#)
- > [Series 11: Born digital materials](#)

<https://findingaids.library.emory.edu/documents/rushdie1000/>

Visit: <https://findingaids.library.emory.edu/documents/rushdie1000/>

DESCRIPTION OF SERIES

- › [Series 1: Journals, appointment books, and notebooks, 1974-2003](#)
- › [Series 2: Writings by Rushdie, 1964-2006](#)
 - › [Subseries 2.1: Fiction, 1973-2006](#)
 - › [Subseries 2.2: Nonfiction, 1981-2002](#)
 - › [Subseries 2.3: Scripts, 1984-2002](#)
 - › [Subseries 2.4: Other writings, 1964-2002](#)
- › [Series 3: Writing by others, 1983-2005](#)
 - › [Subseries 3.1: Writings about Rushdie, 1983-2004](#)
 - › [Subseries 3.2: Other writings, 1983-2005](#)
- › [Series 4: Correspondence, 1974-2006](#)
- › [Series 5: Personal papers, 1964-2005](#)
 - › [Subseries 5.1: Financial records, 1974-2005](#)
 - › [Subseries 5.2: Legal papers, 1976-2004](#)
 - › [Subseries 5.3: Other personal papers, 1964-2005](#)
 - › [Subseries 5.4: Family papers, 1984-2004](#)
- › [Series 6: Subject files, 1976-2006](#)
- › [Series 7: Photographs, circa 1947-2006](#)
 - › [Subseries 7.1: Salman Rushdie, circa 1960-2006](#)
 - › [Subseries 7.2: Other people and places, circa 1980-2000](#)
 - › [Subseries 7.3: Slides and negatives, 1972-1996](#)
 - › [Subseries 7.4: Family photographs, circa 1947- circa 2000](#)
- › [Series 8: Printed material, 1980-2008](#)
 - › [Subseries 8.1: Printed material by Rushdie, 1980-2005](#)
 - › [Subseries 8.2: Printed material about Rushdie, 1975-2008](#)
 - › [Subseries 8.3: General printed material, 1982-2005](#)
- › [Series 9: Memorabilia, 1982-1999](#)
- › [Series 10: Audiovisual, 1981-2008](#)
 - › [Subseries 10.1: Audio recordings, 1986-2005](#)
 - › [Subseries 10.2: Video recordings, 1981-2008](#)
- › [Series 11: Born digital materials](#)
- › [Series 12: Unprocessed additions](#)
- › [Series 13: Additions received from Elizabeth West, 1988-2012 \(bulk 1988-2000\)](#)

RUSHDIE, SALMAN.

SALMAN RUSHDIE PAPERS > BORN DIGITAL MATERIALS

 [Printable PDF](#)

+ TABLE OF CONTENTS

+ DESCRIPTION OF SERIES

Search EmoryFindingAids

only collections with digital resources

SEARCH ▶

Search for items in the container list of **Salman Rushdie papers, 1947-2008**

only digital resources

SEARCH THIS FINDING AID ▶

Send us your feedback.

FEEDBACK ▶

REQUEST FROM THIS COLLECTION

SERIES 11

BORN DIGITAL MATERIALS

Scope and Content Note

The computers and related devices include: one Macintosh Performa 5400/180, one Macintosh PowerBook 5300c, two Macintosh PowerBook G3 models, and one SmartDisk FWFL60 FireLite 60GB 2.5" FireWire Portable Hard Drive. At present, only the Macintosh Performa 5400/180 has been processed.

Restrictions on Access

Use of the original computers and related devices is restricted. Access to the digital files is only available in the Stuart A. Rose Manuscript, Archives, and Rare Book Library (the Rose Library).

Researchers must contact the Rose Library in advance for access to unprocessed born digital materials in this collection. Collection restrictions, copyright limitations, or technical complications may hinder the Rose Library's ability to provide access to unprocessed born digital materials.

Box	Folder	Content
RRL		Access copies of processed born digital material [Reading room access ONLY]
-	-	Macintosh Performa 5400/180 [Original RESTRICTED]
-	-	Macintosh Powerbook 5300c [Original RESTRICTED]
-	-	Macintosh Powerbook G3 [1] [Original RESTRICTED]
-	-	Macintosh Powerbook G3 [2] [Original RESTRICTED]
-	-	SmartDisk FWFL60 FireLite 60GB 2.5" FireWire Portable Hard Drive [Original RESTRICTED]
-	-	"CBS news-productions, 'Salman Rushdie biography,' producer: Shimkin, 8 documents, Microsoft Word and Pilotware," 3.5" floppy disk [Original RESTRICTED]



Hard to believe, but the old IBM 029 is fully restored and back in action! This demo shows manual punching, program-controlled punching, fast duplication, and interpreting, which are the main features of the 029. Everything works - almost. It still did miss an A in the interpret run in this video, and occasionally misfeeds a card, which I'll fix later. But it is completely usable now.

Full 029 keypunch story in video series here:

Part 0: <https://youtu.be/Ey0F0mq0Nys> (arrival and overview)

Part 1: <https://youtu.be/6XpR4cwCado> (first power up)

Part 2: https://youtu.be/b5P5bBv_wNk (first mechanical motions)

Part 3: <https://youtu.be/zlOx0ZaJQqY> (start debugging)

Part 4: <https://youtu.be/hNhM3kjrYgl> (cam restoration)

Part 5: <https://youtu.be/U2YCMEm9Gck> (it punches again!)

Part 6: https://youtu.be/FkM_FRNXqU (keyboard repair)

Part 7: https://youtu.be/2_A_PfLSYOM (print head repair)

Part 8: <https://youtu.be/lojweEfXlXg> (programming drum repair)

Part 9: <https://youtu.be/YnnGbcM-H8c> (full working demo)

1964 IBM 029 Keypunch Card Punching Demonstration

1964 IBM 029 Keypunch Card Punching Demonstration,
<https://www.youtube.com/watch?v=YnnGbcM-H8c>

PUNCH CARD READER IN ACTION
[HTTPS://WWW.YOUTUBE.COM/WATCH?V=YLKBYMKP6P0](https://www.youtube.com/watch?v=YLKBYMKP6P0)

Retrieved 25 Nov 2019

Computer Punch Cards - Historical Overview - (4 Oct 2015),
<https://www.youtube.com/watch?v=YXE6HjN8heg>

Formats	
WFS	85307
CSV	81744
WMS	78041
Plain text	54259
HTML	30870
PDF	21416
json	20868
ZIP	20776
Excel XLS	19528
Excel XLSX	19277
xml	12703
WMS	12653
WFS	10533
Esri Shape	8090
TSV	6696
KML	6067

Atom Feed	5463
gml	4750
Karte	4512
Webanwendung	4511
Provisional data	4082
GNU zip	3597
Word DOC	2576
GeoJSON	2390
Word DOCX	2037
JPEG	1836
Diverse	1743
SERVICE	1150
ODS	980
xlsx	835
esri rest	748
.xlsx	728
ArcGIS Map Service	688

ArcGIS Map Preview	652
download	603
KMZ	598
MULTIFORMAT	565
csv	541
Nedladdning	504
application/json	454
xls	435
PNG	431
.xls	407
MDB	390
text/csv	372
TIFF	347
excel (.xlsx)	339
dxf	337
shape	312
view	312

European Data Portal, File Formats (2020)

0101011101100101011011000110001101101111011011010
1100101 0111010001101111 011101000110100001100101
01010000010010010101001101000001
010100110111010101101101011011010110010101110010
010100110110001101101000011011110110111101101100
0110111101101110
0101010001101111011011110110110001110011
0110100101101110 011101000110100001100101
0100010001101001011001110110100101110100011000010
1101100
0100100001110101011011010110000101101110011010010
111010001101001011001010111001100101110

DIGITAL CONTENT AND ECONOMIC VALUE



Digital technologies such as social media, sensors, digital imagery, monitoring devices, business processes are creating a **tidal wave of digital materials to be ingested** much of which has archival and also commercial potential and some of which needs to be curated by memory institutions—new technologies, processes, and cultural attitudes are needed to make these options possible.



These digital materials are creating vast resources which have near term economic value and longer term social, cultural and longitudinal economic potential.

DIGITAL CURATION AND RESILIENCE

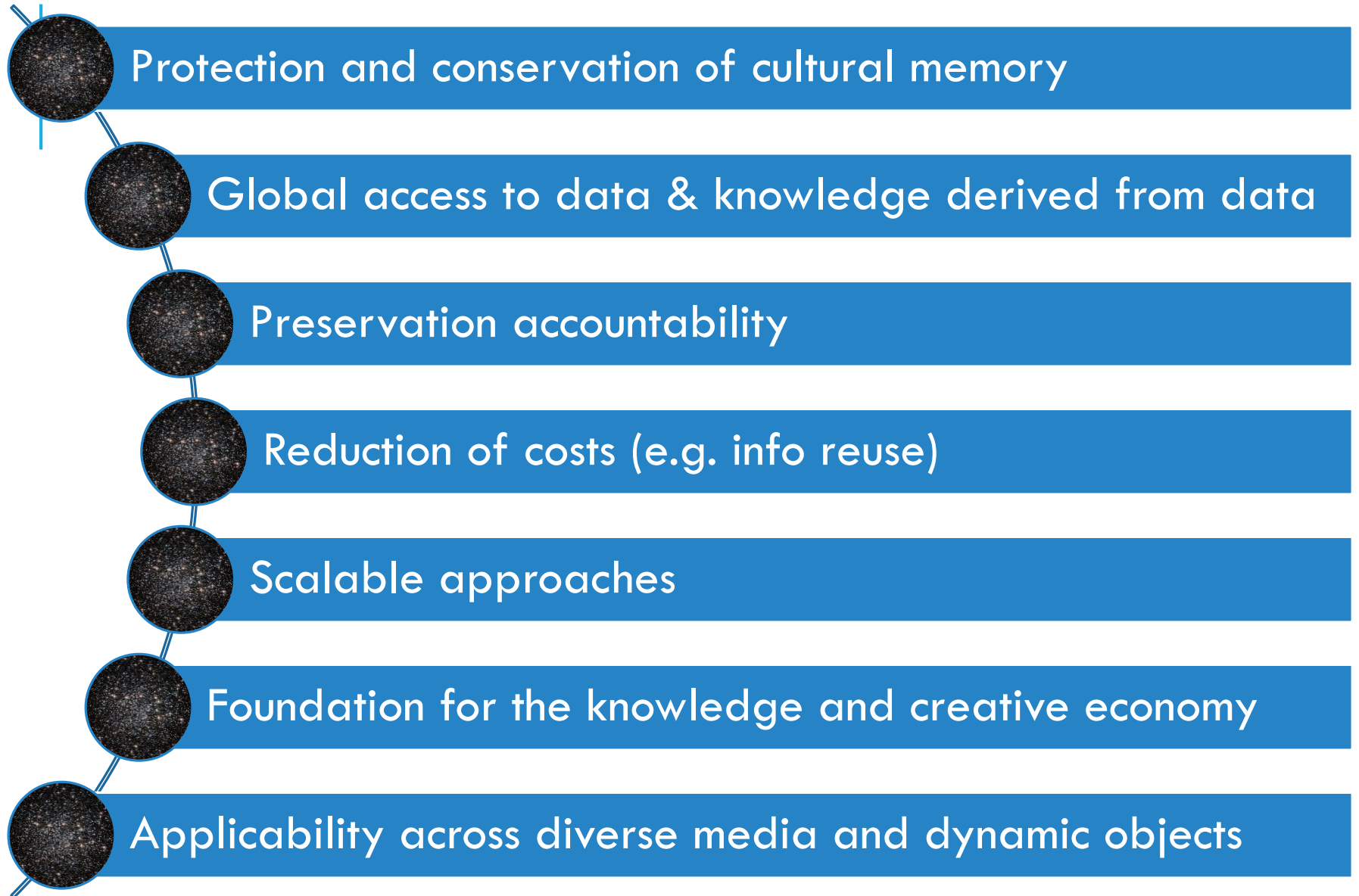
“[d]igital curation [...] is about **maintaining**, and **adding value** to, a trusted body of digital information for current and future use” by adopting a **lifecycle** approach [22] and by foregrounding the need for “subject description and linkage to discipline-based ontologies [...] descriptive information that allows re-analysis of datasets of scientific, scholarly significance,” and other as a prerequisite to ensuring future “fitness for purpose”.

- Our goal is to rollout data management methods and infrastructure that will impart **resilience** to digital objects in the face of changing technologies.

From **Panos Constantopoulos and Costis Dallas, "Aspects of a digital curation agenda for cultural heritage"**
http://www.academia.edu/931035/Aspects_of_a_Digital_Curation_Agenda_for_Cultural_Heritage

[22] is M. Pennock, "Digital curation: a life-cycle approach to managing and preserving usable digital information," *Library and Archives Journal*, vol. 1, January 2007 2007. [Online] Available: http://www.ukoln.ac.uk/ukoln/staff/m.pennock/publications/docs/lib-arch_curation.pdf

EXPECTATIONS FOR DIGITAL CURATION

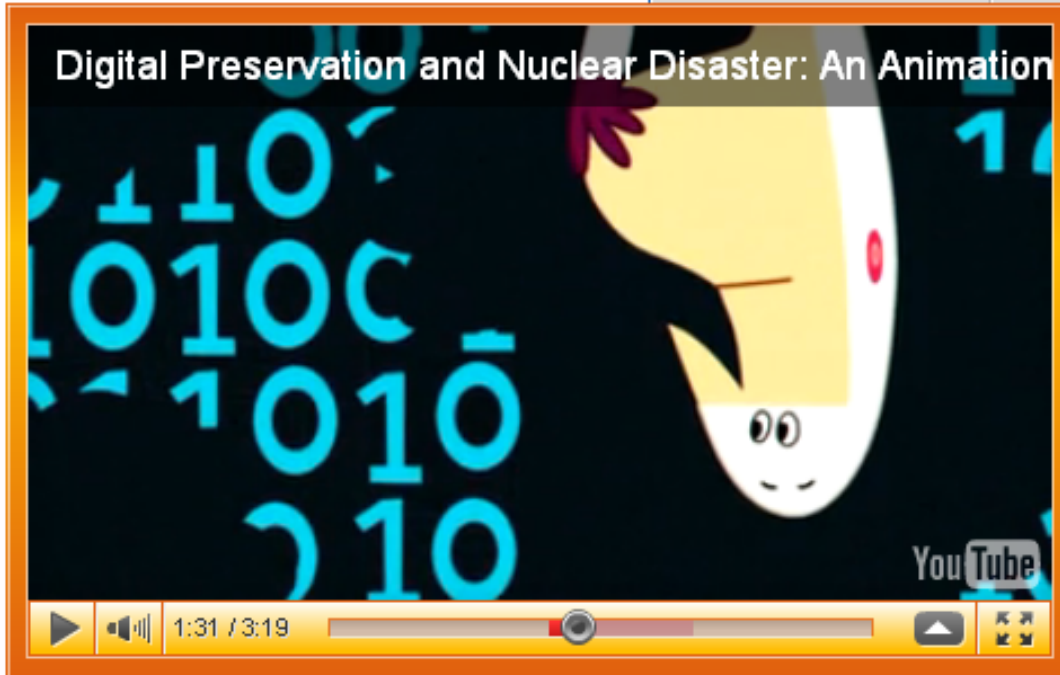


Digital Curation Communication

Digital Preservation and Nuclear Disaster: An Animation

Posted on 6th May 2009

Download This Video



see on youtube.com

A screenshot of a YouTube channel page for the user 'wepreserve'. The channel name is 'wepreserve' and it was created on 01 květen, 2009. The channel description states: "Team Digital Preservation saves the world from nuclear disaster caused by the work of Team Chaos." Below the description is the channel's URL: "http://www.youtube.com/watch?v=pbBa6Oam7-w". The page shows a video player with a shield-shaped logo that says "TEAM DIGITAL PRESERVATION" and "10101010". The video has 31 ratings and 11,485 views. There are social media sharing options for Facebook and Twitter. A list of recommended videos is shown on the right side of the page.

WePreserve (DigitalPreservationEurope), 2009, *Digital Preservation and Nuclear Disaster: An Animation*, <https://www.youtube.com/watch?v=pbBa6Oam7-w>

SO WHAT WE KNOW IS THAT

Long term access to digital materials is not inherently guaranteed, it is prone to risks

- Some technological.
- Some social.
- Some organisational.
- And some cultural.

Actual risks can be assessed and measured—actual risks can be managed.

- Risk Identification
- Risk Management

OBSTACLES TO ACCESSING SURVIVING DIGITAL RESOURCES

Loss of functionality of access devices (e.g. lack of drivers or interface functionality)

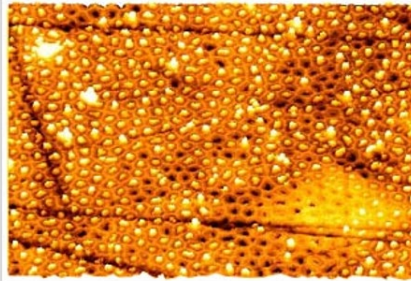
Media degradation (e.g. temp & hum, disaster, manufacturer defects)

Loss of manipulation capabilities (e.g. hardware, software, applications)

Loss of presentation capabilities

Weak links in creation chain (capture, manipulation, storage, dissemination)

Example: What can go wrong with Magnetic Media?



Magnetic Particles in tape surface

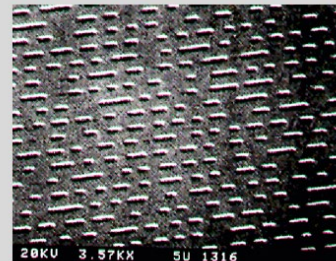
Image © Park Scientific Instruments,
<http://shell7.ba.best.com/~wwwpark/appnotes>

- ❖ Hydrolysis
- ❖ Binder breakdown
- ❖ Particle breakdown
- ❖ Loss of lubricant
- ❖ Deformation
- ❖ “I dropped it” I really did

Example: What can go wrong with CDs?

<i>Environmental Impacts</i>	<i>Handling Impacts</i>	<i>Mechanical Impacts</i>
<i>corrosive gases</i>	<i>shocks</i>	<i>degradation of hardware</i>
<i>temperature humidity</i>	<i>abrasions</i>	
<i>Exposure to UV Light</i>	<i>scratches</i>	

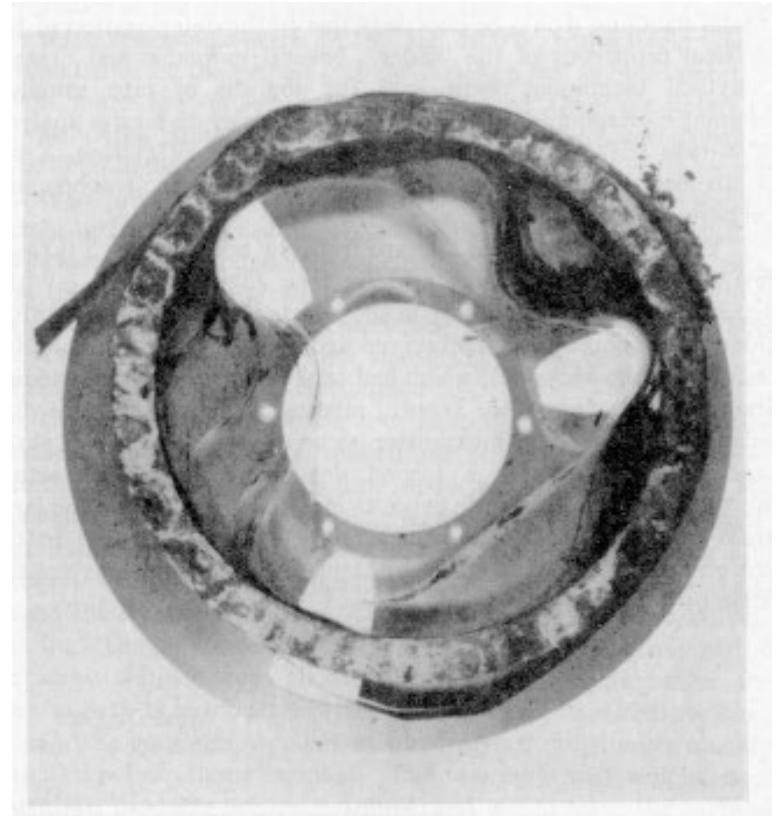
Is a CD an Archival Medium?



© Hewlett-Packard, 1999



THE CHALLENGER DISASTER

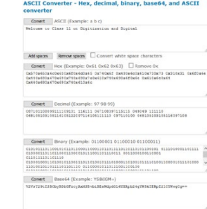


Images © NASA & IBM



Early Computers (stock footage / archival footage)

EARLY COMPUTERS (STOCK FOOTAGE / ARCHIVAL FOOTAGE), (26 MAY 2011),
[HTTPS://WWW.YOUTUBE.COM/WATCH?V=R9LNJOA8WQ8](https://www.youtube.com/watch?v=R9LNJOA8WQ8)



ASCII Converter - Hex, decimal, binary, base64, and ASCII converter

Convert ASCII (Example: a b c)

Welcome to the PISA Summer School on Tools in the Digital Humanities.

Add spaces **Remove spaces** Convert white space characters

Convert Hex (Example: 0x61 0x62 0x63) Remove 0x

0x57 0x65 0x63 0x6f 0x6d 0x65 0x74 0x6f 0x74 0x68 0x65
0x50 0x49 0x53 0x41 0x53 0x75 0x6d 0x6d 0x65 0x72
0x53 0x63 0x68 0x6f 0x6f 0x6c 0x6f 0x6e 0x54 0x6f 0x6f 0x6c 0x73
0x69 0x6e 0x74 0x68 0x65 0x44 0x69 0x67 0x69 0x74 0x61 0x6c

Convert Decimal (Example: 97 98 99)

087 101 108 099 111 109 101 116 111 116 104 101 080 073 083 065
083 117 109 109 101 114 083 099 104 111 111 108 111 110
084 111 111 108 115 105 110 116 104 101 068 105 103 105 116 097 108
072 117 109 097 110 105 116 105 101 115 046

Convert Binary (Example: 01100001 01100010 01100011)

010001000110100101100111011010010111010001100001011011
00
010010000111010101101101011000010110111001101001011101
0001101001011001010111001100101110

Convert Base64 (Example: YSBiIGM=)

V2VsY29tZSB0byB0aGUgUElTQSBTdW1tZXIgaU2Nob29sIG9uIFRvb2xzI
GluIHRoZSBEaWdpdGFsIEh1bWVuaXRpZXMu

<https://www.branah.com/ascii-converter>

HIGH-LEVEL PRESERVATION VIEW



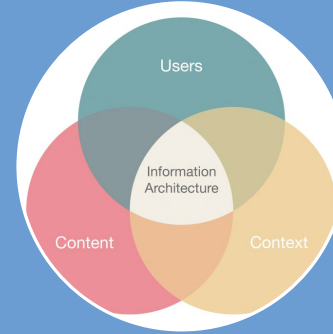
Bit Stream

- (01100101101010010)



Information Content

- (e.g. images, sounds, text)



Context of Information

- (e.g. information architecture, linkages, interrelatedness)



Experience

- (e.g., speed, layout, quality of display device, input device characteristics)



KEEP IN MIND.....PERFORMANCE

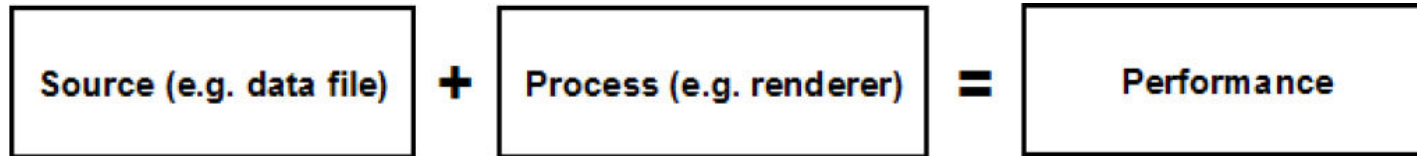


Figure 2.1: National Archives of Australia's Performance Model

Multi-layered performance and semantic intelligibility any many different layers

A relationship between the user of an object and the object itself which is “brokered by software and hardware” (NAA, 2002)

- Example
 - Data represented as magnetic charges on media
 - Interpreted as 1's and 0's and presented as a sequence which to the Operating System appears as a file.
 - File presented to application which performs it.
- **Performance is nuanced and dependent upon a variety of successful performances**

source

process

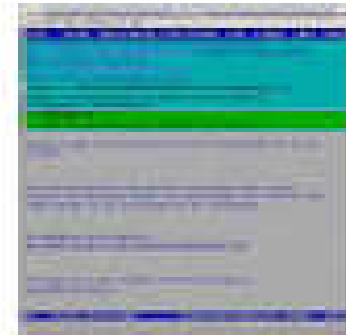
performance



Email
Data Object



Thunderbird
Linux OS



MS Outlook 2007
Microsoft Windows XP

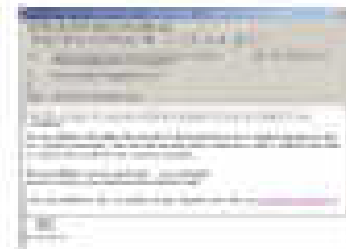


Figure 5 Application of the Performance Model to emails

G. Knight, and L. Montague, 2009, “ InSPECT: Final Report”, <http://www.significantproperties.org.uk/inspect-finalreport.pdf>, p.27

DIGITAL ENTITIES HAVE



Syntax



Semantics

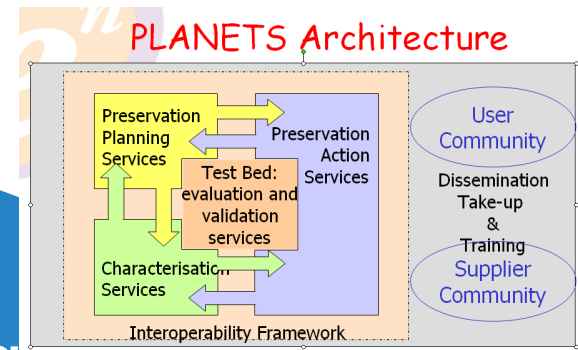


Pragmatics



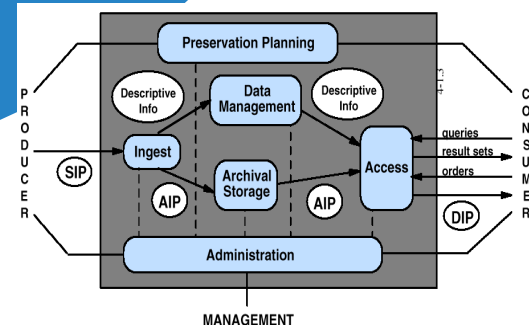


Digital Curation Centre: <http://www.dcc.ac.uk/resources/curation-lifecycle-model>



Abstraction and Modelling provides a mechanism to improve understanding and communication

- Planets Project – Preservation Model
- OAIS Digital Preservation Model
- DCC UK: Digital Curation Lifecycle Model



OAIS Model & Example

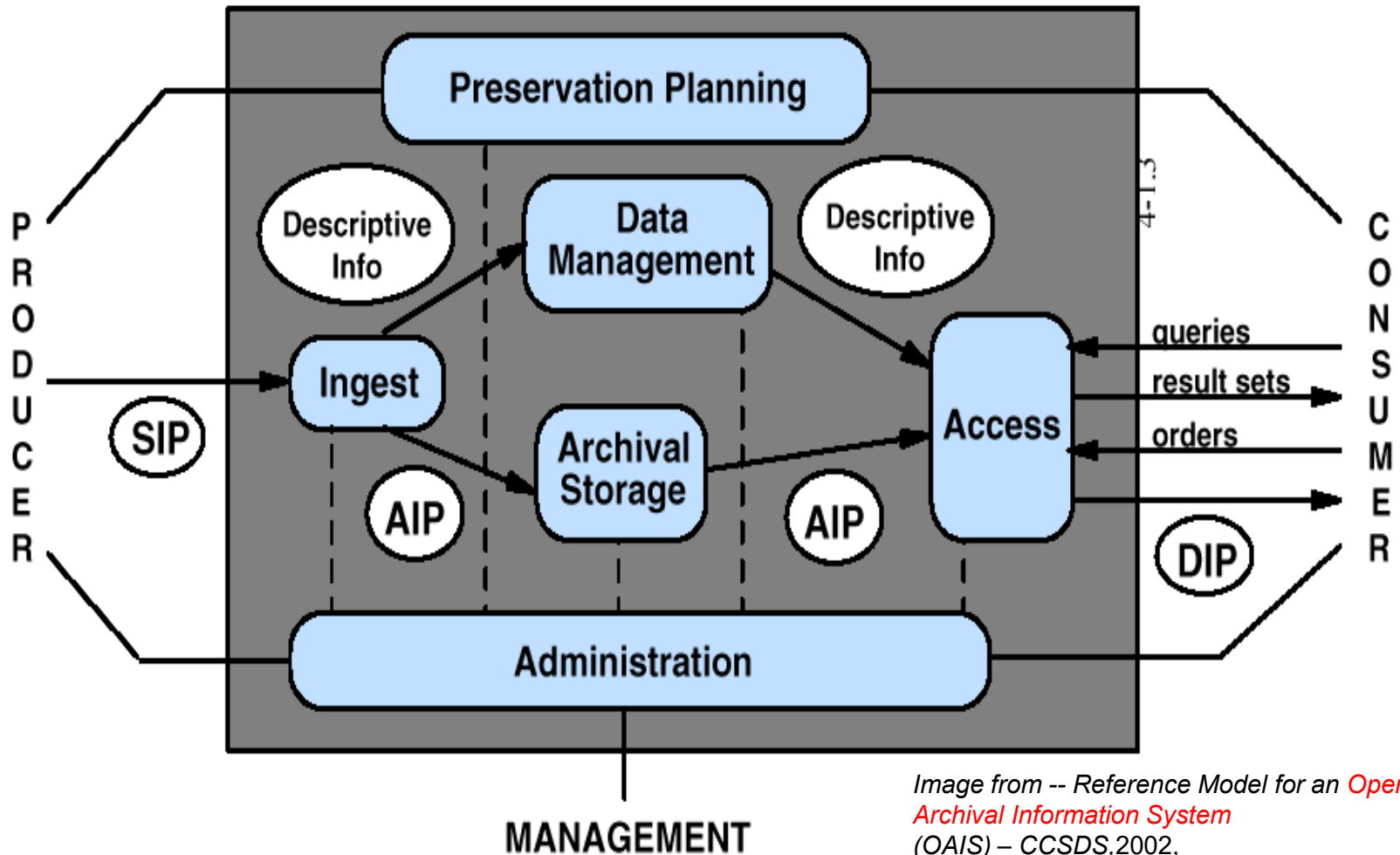
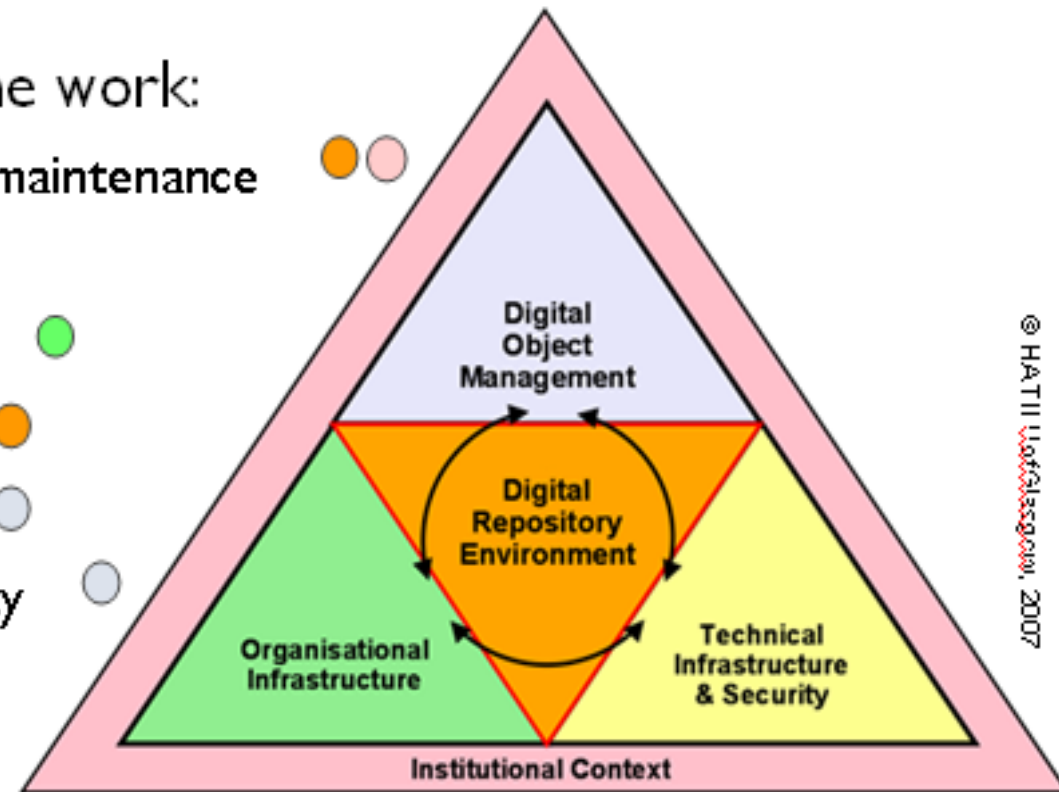


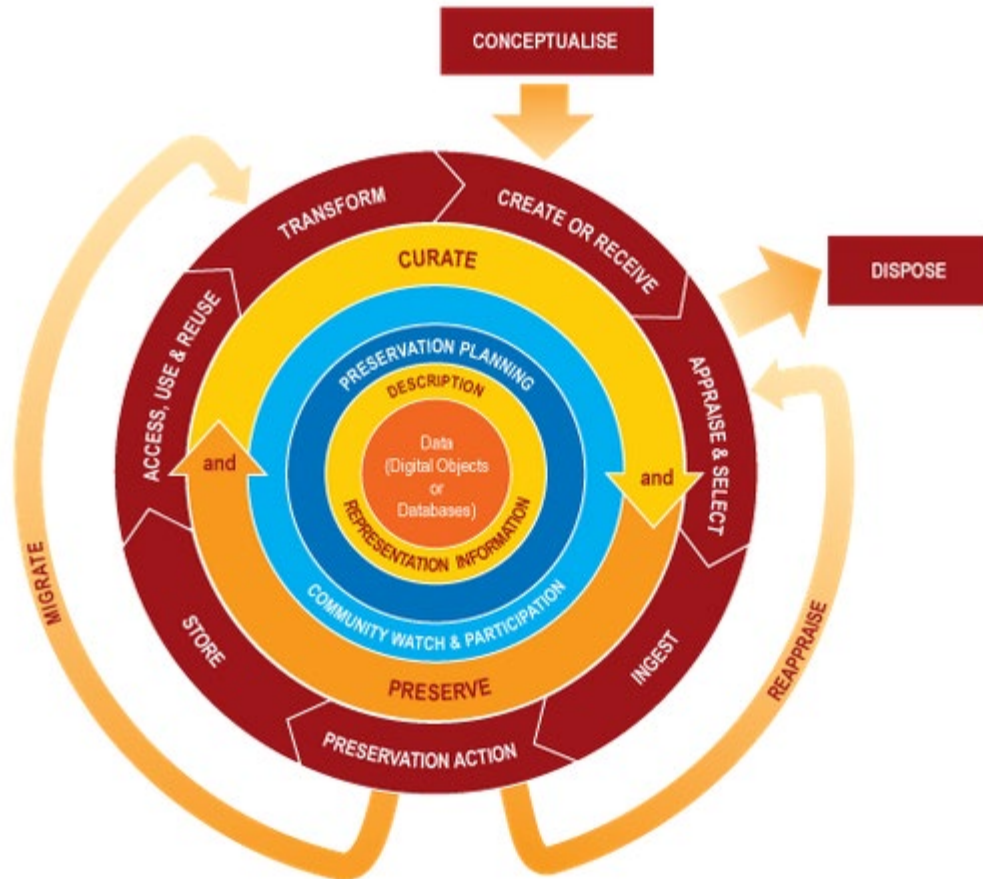
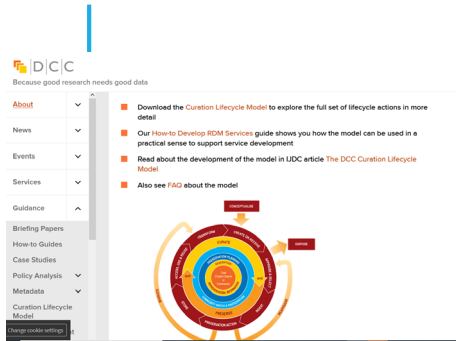
Image from -- Reference Model for an *Open Archival Information System* (OAIS) – CCSDS,2002,
<http://www.ccsds.org/documents/650x0b1.pdf>

Digital preservation repository core criteria

- An intellectual context for the work:
 - Commitment to digital object maintenance
 - Organisational fitness
 - Legal & regulatory legitimacy
 - Effective & efficient policies
 - Acquisition & ingest criteria
 - Integrity, authenticity & usability
 - Audit trail and metadata
 - Dissemination
 - Preservation planning & action
 - Adequate technical infrastructure



Digital Curation Lifecycle: DCC UK



<https://www.dcc.ac.uk/guidance/curation-lifecycle-model>

The Curation Lifecycle

The DCC Curation Lifecycle Model provides a graphical high level overview of the stages required for successful curation and preservation of data from initial conceptualisation or receipt. The model can be used to plan activities within an organisation or consortium to ensure that all necessary stages are undertaken, each in the correct sequence. The model enables granular functionality to be mapped against it; to define roles and responsibilities, and build a framework of standards and technologies to implement. It can help with the process of identifying additional steps which may be required, or actions which are not required by certain situations or disciplines, and ensuring that processes and policies are adequately documented.

Data (Digital Objects or Databases)

Data, any information in binary digital form, is at the centre of the Curation Lifecycle. This includes:

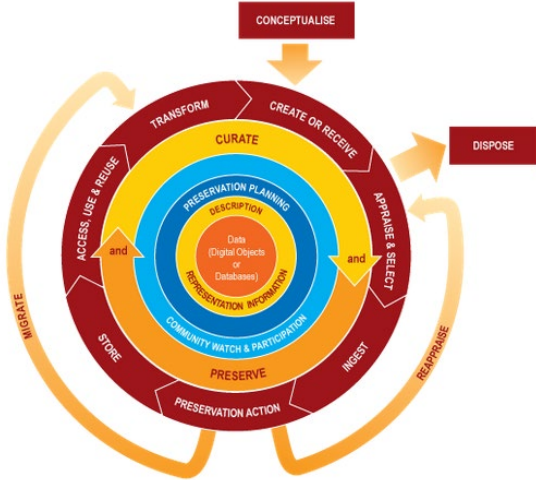
- Digital Objects**
 - Simple Digital Objects are discrete digital items; such as textual files, images or sound files, along with their related identifiers and metadata.
 - Complex Digital Objects are discrete digital objects, made by combining a number of other digital objects, such as websites.

Databases Structured collections of records or data stored in a computer system.

Full Lifecycle Actions

- Description and Representation Information** Assign administrative, descriptive, technical, structural and preservation metadata, using appropriate standards, to ensure adequate description and control over the long-term. Collect and assign representation information required to understand and render both the digital material and the associated metadata.
- Preservation Planning** Plan for preservation throughout the curation lifecycle of digital material. This would include plans for management and administration of all curation lifecycle actions.
- Community Watch and Participation** Maintain a watch on appropriate community activities, and participate in the development of shared standards, tools and suitable software.
- Curate and Preserve** Be aware of, and undertake management and administrative actions planned to promote curation and preservation throughout the curation lifecycle.

<https://www.dcc.ac.uk/guidance/curation-lifecycle-model>



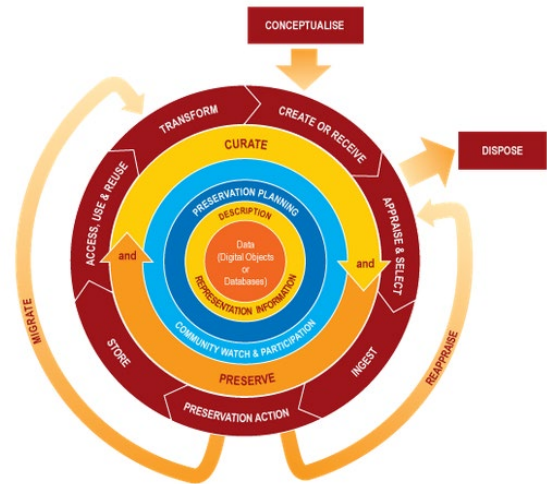
Sequential Actions

Conceptualise	Conceive and plan the creation of data, including capture method and storage options.
Create or Receive	Create data including administrative, descriptive, structural and technical metadata. Preservation metadata may also be added at the time of creation. Receive data, in accordance with documented collecting policies, from data creators, other archives, repositories or data centres, and if required assign appropriate metadata.
Appraise and Select	Evaluate data and select for long-term curation and preservation. Adhere to documented guidance, policies or legal requirements.
Ingest	Transfer data to an archive, repository, data centre or other custodian. Adhere to documented guidance, policies or legal requirements.
Preservation Action	Undertake actions to ensure long-term preservation and retention of the authoritative nature of data. Preservation actions should ensure that data remains authentic, reliable and usable while maintaining its integrity. Actions include data cleaning, validation, assigning preservation metadata, assigning representation information and ensuring acceptable data structures or file formats.
Store	Store the data in a secure manner adhering to relevant standards.
Access, Use and Reuse	Ensure that data is accessible to both designated users and reusers, on a day-to-day basis. This may be in the form of publicly available published information. Robust access controls and authentication procedures may be applicable.
Transform	Create new data from the original, for example <ul style="list-style-type: none"> - By migration into a different format. - By creating a subset, by selection or query, to create newly derived results, perhaps for publication.

Occasional Actions

Dispose	Dispose of data, which has not been selected for long-term curation and preservation in accordance with documented policies, guidance or legal requirements. Typically data may be transferred to another archive, repository, data centre or other custodian. In some instances data is destroyed. The data's nature may, for legal reasons, necessitate secure destruction.
Reappraise	Return data which fails validation procedures for further appraisal and reselection.
Migrate	Migrate data to a different format. This may be done to accord with the storage environment or to ensure the data's immunity from hardware or software obsolescence.

<https://www.dcc.ac.uk/guidance/curation-lifecycle-model>



Digital Curation/Preservation Tools

Metadata Extraction	Tools that support the extraction of metadata from files.
Metadata Processing	Tools that support the processing or management of metadata.
Multi Format Rendering	Tools that support the rendering of a cross section of file format or content categories.
OCR	Tools that support the generation of text from bitmap images, otherwise known as Optical Character Recognition
Organisational Audit	Tools that that enable an audit of an organisation's capability with respect to preservation, typically relating to a maturity model
Persistent Identification	Tools that support the unique and persistent identification of files or intellectual entities.
Personal Archiving	Tools that support the preservation and archiving of data relating to individuals.
Planning	Tools that support the planning of preservation activities.
Policy	Tools that support the development and management of digital preservation policy.
Preservation System	Digital repository applications that typically perform a number of functions across the digital lifecycle such as ingest, storage, preservation action and access.
Quality Assurance	Tools that support quality checking of digital resources, identifying damaged, incomplete or low quality data. Typically used to identify damage introduced via processes such as format migration or digitisation.
Redaction	Tools that support the removal of selected information from digital files. Typically used for removal of sensitive information like telephone or credit card numbers from personal archives before providing access to users.
Rendering	Tools that support the rendering of digital resources so they can be viewed, printed, or otherwise accessed by users.
Repair	Tools that support the repair of damaged or corrupted data.
Secure Deletion	Tools that support deletion of data in a way that cannot be reversed, typically to avoid third parties stealing sensitive information from decommissioned or recycled hardware.
Service	Tools that operate as a remote, perhaps cloud based, service
Storage	Tools that support the storage of digital resources, possibly in multiple locations to avoid loss of data due to hardware or other failures.
Transfer	Tools that support transfer of packaged digital resources from one organization to another.
Validation	Tools that support the validation of digital files, typically against a file format specification.

https://coptr.digipres.org/index.php/Tool_Functions

Digital Curation/Preservation Tools

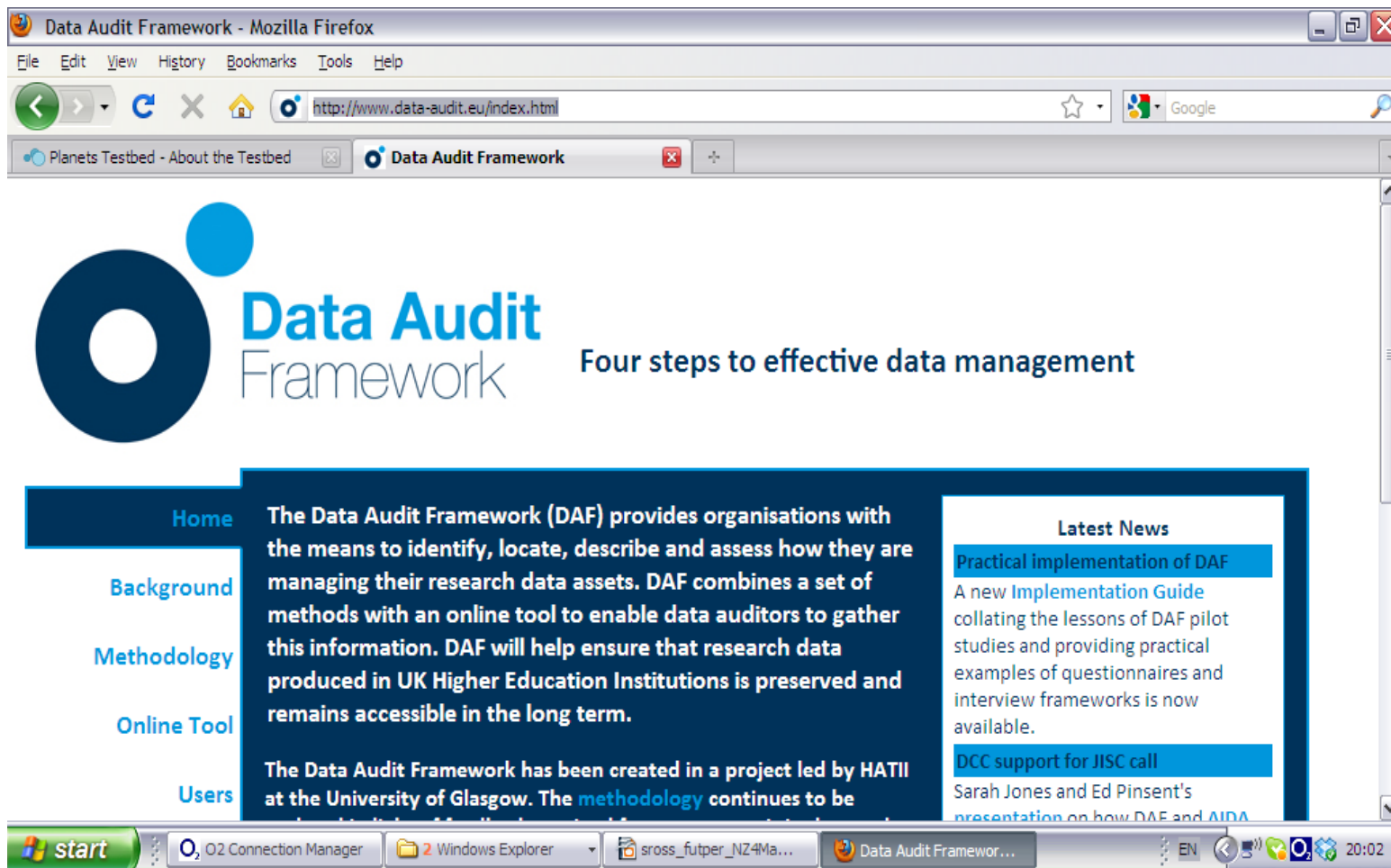
Organisational Audit

Tools for this function

Tool	Purpose
CARDIO	CARDIO is a benchmarking tool for data management strategy development
CTS (Core Trust Seal)	CoreTrustSeal offers to any interested data repository, certification based on the DSA–WDS Core Trustworthy Data Repositories Requirements catalogue and procedures
DMAOnline (Data Management Administration Online)	Provides a single dashboard view of how various departments contribute to RDM activities and how an institution is performing in terms of its compliance with policies
DPC RAM (Rapid Assessment Model)	A maturity modelling tool that has been designed to enable rapid benchmarking of an organization's digital preservation capability.
DPCMM (Digital Preservation Capability Maturity Model)	Maturity / gap analysis model for digital preservation
DRAMBORA	DRAMBORA offers a quantifiable insight into the severity of risks faced by repositories right now, and an effective means for reporting these.
Data Asset Framework	The Data Asset Framework (formerly the Data Audit Framework) provides organisations with the means to identify, locate, describe and assess how they are managing their research data assets.
Embedding Repositories Self-Assessment Tool	Embedding Repositories Self-Assessment Tool is comprised of a series of questions designed to quantify the degree that a digital repository is 'embedded' within its institution – the extent to which both the organisation's research and its administrative culture recognise the repository's value and take full advantage of its capacity.
NDSA Levels of Preservation	The "Levels of Digital Preservation" are a tiered set of recommendations for how organizations should begin to build or enhance their digital preservation activities.
OPD for RDM	An RDF based list of basic RDM infrastructure components to make this infrastructure more visible and easier to identify
RMCAS	RMCAS is an assessment tool for organisations wishing to map their current records management infrastructure against community best-practice.

https://coptr.digipres.org/index.php/Tool_Functions

Knowing What Data you have



The screenshot shows a Mozilla Firefox browser window displaying the Data Audit Framework website. The browser's address bar shows the URL <http://www.data-audit.eu/index.html>. The website features a large blue circular logo on the left, followed by the text "Data Audit Framework" and the tagline "Four steps to effective data management". Below this, there is a navigation menu with links for "Home", "Background", "Methodology", "Online Tool", and "Users". The main content area contains a paragraph describing the Data Audit Framework (DAF) and its purpose, along with a "Latest News" section listing recent updates.

Data Audit Framework
Four steps to effective data management

Home The Data Audit Framework (DAF) provides organisations with the means to identify, locate, describe and assess how they are managing their research data assets. DAF combines a set of methods with an online tool to enable data auditors to gather this information. DAF will help ensure that research data produced in UK Higher Education Institutions is preserved and remains accessible in the long term.

Background

Methodology

Online Tool

Users The Data Audit Framework has been created in a project led by HATI at the University of Glasgow. The methodology continues to be

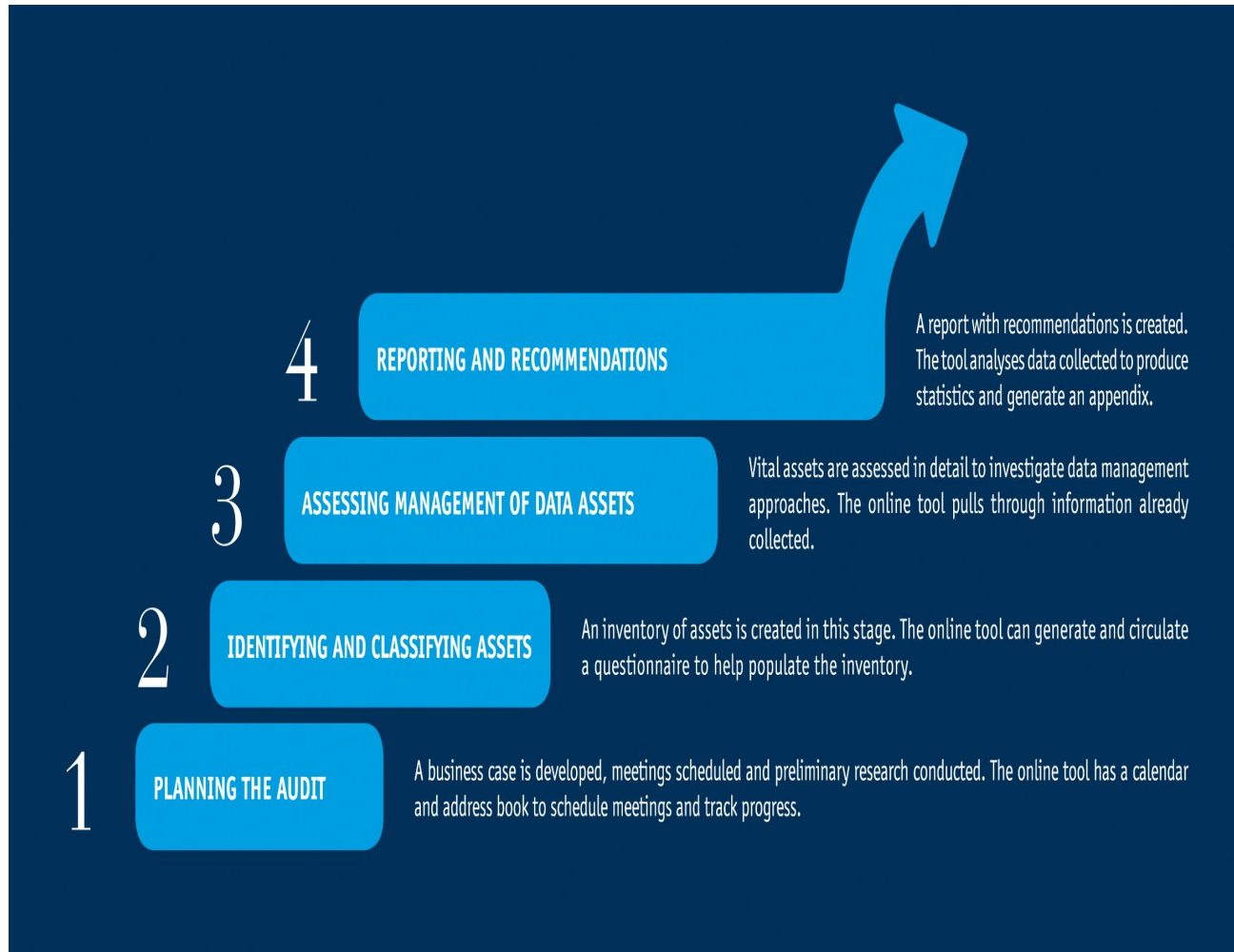
Latest News

Practical implementation of DAF
A new [Implementation Guide](#) collating the lessons of DAF pilot studies and providing practical examples of questionnaires and interview frameworks is now available.

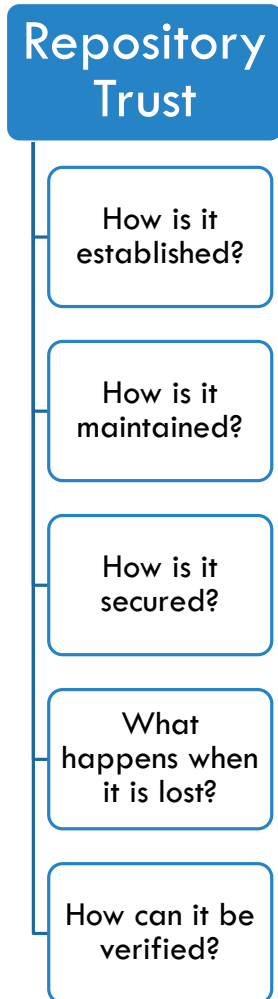
DCC support for JISC call
Sarah Jones and Ed Pinsent's [presentation on how DAF and AIDA](#)

[HTTP://WWW.DATA-AUDIT.EU/INDEX.HTML](http://www.data-audit.eu/index.html)

Conducting a Data Audit



How do we know that our data and other digital assets are secure and can be returned as authentic entities?



Linking Risk and Trust

Are repositories capable of:

- identifying and prioritising the risks that impede their activities?
- managing the risks to mitigate the likelihood of their occurrence?
- establishing effective contingencies to alleviate the effects of the risks that occur?

If so, then they are likely to engender a trustworthy status – if they can demonstrate these capabilities

STARTING POINT FOR VALIDATING TRUST

Independent measuring of repositories, including archives, capability to do their job is essential to users

Taken as axiomatic that audit is a mechanism for establishing the trustworthiness of a repository

We seek to develop the debate on the evidence required for objective and transparent assessment of the effectiveness of digital archives/repositories



DIPLOMATIC PRINCIPLES

Who? Who were the Actors? (Quis?)

What? (Quid?)

In what manner? form, formulae, style (Quomodo?)

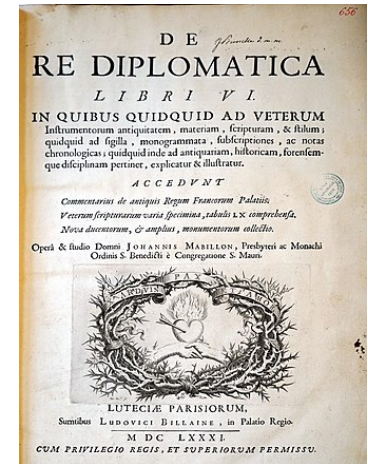
What support, aid or help? (e.g. drafting, engrossing, and ratification) (Quibus Auxiliis?)

Why? What is its purpose? (Cur?)

Where? (Ubi?)

When? (Quando?)

L.E. Boyle, 'Diplomatics,' in J.M. Powell (ed.), *Medieval Studies: An Introduction*, (Syracuse, NY: Syracuse University Press, 1976), 69-101



Left image from De Re Diplomatica (1681), Titre de 1681, [bibliothèque Carnegie \(Reims\)](#)

https://upload.wikimedia.org/wikipedia/commons/thumb/4/43/De_re_diplomatica_17765.jpg/800px-De_re_diplomatica_17765.jpg

CONCEPTS OF DIPLOMATICS (SEE BOYLE) ALSO BECOMES

A tool to read the digital archive itself

- What was the motivation?
- What was the intended purpose?

Appraisal, Arrangement, and Description are all subject to analysis:

- who, what, in what manner (e.g. what standards, how), with what support, aid or help, why (e.g. what purpose), where, and when



RÉSUMÉ Il y a quinze ans, Elizabeth Diamond décrivait l'archiviste comme un scientifique médico-légal. Depuis quelques années, plusieurs auteurs dans le domaine de l'archivistique ont qualifié les professionnels responsables de la préservation des documents numériques de conservateurs de confiance (« trusted keepers »), ou de gardiens (« custodians »). Sans doute, dans l'environnement numérique, on fait de plus en plus appel aux professionnels de l'information pour évaluer et préserver l'authenticité des documents dont ils sont responsables, et pour agir en tant que tierce parties neutres. Mais sont-ils qualifiés pour remplir ce rôle? Cet article tente d'identifier les connaissances que doit avoir le professionnel d'information de confiance pour être capable d'évaluer la véridité (« trustworthiness ») des documents numériques et pour assurer que leur authenticité puisse être démontrée, au besoin, à n'importe quel point dans leur cycle de vie. Pour ce faire, l'article présente des concepts développés par le projet InterPARES dans le domaine de la diplomatie des documents numériques; il compare ceux-ci aux concepts pertinents dérivés d'une discipline relativement nouvelle, le numérique médico-légal (« digital forensics »); il discute des méthodologies dont se servent les deux disciplines, et il propose des domaines qui pourraient être explorés conjointement par les experts en diplomatie et en numérique médico-légal afin de développer un corpus de savoir intégré que l'on pourrait nommer la science médico-légale des documents numériques (« Digital Records Forensics »).

ABSTRACT Fifteen years ago, Elizabeth Diamond described the archivist as a forensic scientist. In the past few years, several archival writers have referred to professionals responsible for keeping digital records as trusted keepers or custodians. Undoubtedly, in the digital environment, record professionals are increasingly called to assess and preserve the authenticity of the records they are responsible for, and to act as neutral third parties. But, are they qualified to fulfill this role? This article aims to begin identifying the body of knowledge that a trusted record professional needs in order to assess the trustworthiness of digital records and ensure that their continuing

* I dedicate this article to the memory of Elizabeth Diamond who encouraged and inspired me when I was learning to be a Canadian archivist. The fact that it took me fifteen years to truly understand her call and bring it to fruition shows her foresight and imagination.

ARCHIVARIA 68 (Fall 2009) 39–66

Archivaria, The Journal of the Association of Canadian Archivists – All rights reserved

Regardless of the fact that, in digital forensics, references to authenticity appear to focus on the data or content in the record rather than on its formal aspects, like diplomatics, the importance of protecting both the documentary and digital presentation of a record for purposes of authentication is implicit in the discussion of digital forensics practices. For example, Ghirardini and Faggioli state that, although conversion of digital evidence to forms and formats different from the original is a process useful to its accessibility and analysis, it “modifies its nature.” This implies that converted records cannot be used as evidence and must always be accompanied by the records in the original presentation.⁵² Although these authors write in the context of a civil law system, which does not consider records hearsay and rules them admissible if proven authentic, the issue they identify also relates to the “best evidence” requirement for admissibility in a common law context, according to which evidence must be submitted in the most authoritative status of transmission, which is the original or an authenticated copy of the original when the former is not accessible.⁵³ Indeed,

Duranti, Luciana. 2009, “From Digital Diplomats to Digital Records Forensics.” *Archivaria*, vol. 68, no. 68, 2009, pp. 39–66 and p 19

examination to develop a theory of what should exist. Ideally, a digital records forensics would articulate:

- how a variety of digital systems should be designed to create and maintain trustworthy digital records that can be regarded as material evidence of facts and acts, serving at the same time transparency, accountability, and users' needs;
- how the authenticity of digital records can be verified when its presumption is weak;
- how the records of the creators should be reliably extracted from the systems in which they reside, and maintained in long-term storage either with the creator or with the preserver in such a way that their authenticity can be presumed;
- how records should be authentically reproduced in the course of their long-term preservation;
- how the features of the records, the actions conducted over them, and the changes caused by such actions should be documented;
- how the records submitted to court as evidence should be kept after the conclusion of court proceedings for as long as needed, so that they remain

Duranti, Luciana. 2009, "From Digital Diplomats to Digital Records Forensics." *Archivaria*, vol. 68, no. 68, 2009, p., 27

5.4.1 Admissibility of Electronic Evidence

The criteria for the admissibility of electronic evidence may differ from jurisdiction to jurisdiction. Generally, the Examiner should consider the following criteria when evaluating electronic evidence for trial:

General Criteria for the Admissibility of Electronic Evidence	
Authenticity	The evidence must establish facts in a way that cannot be disputed and be representative of its original state.
Completeness	The analysis of, or any opinion based on, the evidence must tell the whole story and not be tailored to match a more favourable or desired perspective.
Reliability	There must be nothing about the way in which the evidence was collected and subsequently handled that may cast doubt on its authenticity or veracity.
Convincing	The evidence must be persuasive as to the facts it represents, and must be able to convince the stakeholder of the truth in court.
Proportionality	The methods used to gather the evidence must be fair and proportionate to the interests of justice: the prejudice (i.e. the level of intrusion or coercion) caused to the rights of any party should not outweigh the probative value of the evidence (i.e. its value as proof).

Table 11. General Criteria for the Admissibility of Electronic Evidence

Global Guidelines for Digital Forensics Laboratories

file:///C:/Users/Administrator/Downloads/INTERPOL_DFL_GlobalGuidelinesDigitalForensics.pdf, p 52

DIGITAL FORENSICS

Digital forensics is the process of “collecting, analyzing, and preserving digital data for the purpose of identifying, detecting, and preventing cybercrime.”

DIGITAL FORENSICS IN THE ARCHIVE

- digital forensics can help ensure the authenticity and integrity of the data being examined.
- digital archives may contain large amounts of data that can be difficult to sort through and verify without the tools and techniques used in digital forensics.
- digital forensics can also help researchers identify gaps or discrepancies in the data that may be indicative of tampering or other types of manipulation.
- By using digital forensics to validate the data in digital archives, researchers can rely on the information with greater confidence and trust that the data represents what it claims to represent.
- digital forensics can also help preserve the digital archives themselves, ensuring that the data remains accessible and usable for future research projects.

CORE DIGITAL FORENSIC TOOLS

1. **EnCase Forensic:** a widely used commercial digital forensics tool. It provides comprehensive features for acquiring, analyzing, and reporting on evidence from a variety of digital sources, including computers, mobile devices, and network storage.
2. **AccessData Forensic Toolkit (FTK):** FTK is another popular commercial tool used for digital forensics investigations. It offers advanced capabilities for acquiring, processing, and examining digital evidence from diverse sources, including disk images, email archives, and databases.
3. **Autopsy:** Autopsy is an open-source digital forensics platform widely used by both professionals and enthusiasts. It provides a user-friendly interface and features for analyzing disk images, file system metadata, and extracting artifacts such as deleted files, internet history, and email data.
4. **Volatility:** Volatility is an open-source memory forensics framework. It allows forensic investigators to extract and analyze volatile memory (RAM) from a running system. This tool is useful for detecting and investigating advanced malware, examining network connections, and recovering encryption keys.
5. **Wireshark:** Wireshark is a powerful open-source network protocol analyzer. It captures and analyzes network traffic, allowing investigators to examine network communications, identify suspicious activities, and uncover evidence related to cyberattacks or unauthorized access.
6. **Cellebrite UFED:** Cellebrite UFED is a commercial tool widely used in mobile device forensics. It supports the extraction and analysis of data from a broad range of mobile devices, including smartphones, tablets, and GPS units. It can recover deleted data, extract call logs, messages, and app data, and perform advanced searches.
7. **X-Ways Forensics:** X-Ways Forensics is a comprehensive and efficient commercial forensic tool. It offers features for acquiring, analyzing, and reporting on digital evidence. It can handle large datasets, has advanced searching capabilities, and supports a wide range of file systems and storage media.
8. **Oxygen Forensic Detective:** Oxygen Forensic Detective is a commercial tool designed for mobile device forensics. It provides powerful capabilities for extracting, analyzing, and visualizing data from smartphones, tablets, and cloud services. It supports multiple mobile operating systems and includes advanced features for decoding app data, analyzing social media activities, and recovering deleted data.

WHAT KIND OF TOOLS DO

- ❖ Acquisition and preservation of entire hard drives or individual digital media, ensuring the integrity and authenticity of the original content.
- ❖ Need tools to understand the organization and context of digital materials.
- ❖ Need to know the types of files present within a digital collection, which is crucial for preservation planning and ensuring long-term access to the content.
- ❖ Need to be able to extract metadata as it aids in the organization, management, and documentation of digital materials.
- ❖ Need to identify and redact sensitive data: email addresses, credit card numbers, and other personally identifiable information (PII).
- ❖ Tools are needed to manage the chain of custody, recording findings, and generating reports for investigative or archival purposes.
- ❖ Needed to support digital objects to preservation repositories, ensuring that born-digital materials are appropriately managed for long-term access and integrity.

BITCURATOR

BitCurator is a digital forensics software environment specifically designed for the curation and analysis of born-digital materials in libraries, archives, and museums. It provides a set of integrated open-source tools and workflows to aid in the management and preservation of digital assets.

BitCurator provides a user-friendly interface and a range of automated and manual tools that are specifically tailored for the needs of digital archivists, librarians, and curators. It enables the processing, management, and preservation of digital materials, while also facilitating the investigation and analysis of digital artifacts for forensic purposes.



BitCurator's Open Source Approach: An Interview With Cal Lee

December 2, 2013

Posted by: [Trevor Owens](#)

Share this post:



Cal Lee, Associate Professor at the School of Information and Library Science at the University of North Carolina at Chapel Hill

Open source software is playing an important role in digital stewardship. In an effort to better understand the role open source software is playing, the [NDSA infrastructure working group](#) is reaching out to folks working on a range of open source projects. Our goal is to develop a better understanding of their work and how they are thinking about the role of open source software in digital preservation in general.

<https://blogs.loc.gov/thesignal/2013/12/bitcurators-open-source-approach-an-interview-with-cal-lee/>

Home

Kam Woods edited this page on Apr 16 · 13 revisions

The BitCurator Access project developed tools to help libraries, archives, and museums provide web-based and local access to born-digital materials held on disk images. BitCurator Access tools simplify access to raw and forensically-packaged disk images, allowing users to incorporate these objects into access environments while preserving original order and relevant environmental context. Using open source digital forensics software libraries, these tools enable detailed analysis of file and file system provenance, quality and accessibility of files, metadata in files and the file system, and residual or hidden data.

BitCurator Access focused on four areas of interest related to accessing born-digital collections:

- Web-based access to raw and forensically packaged disk images
- Redaction of file items, metadata and hidden data from disk images
- OS and executable virtualization for legacy disk images
- Transforming and using digital forensics metadata in collecting environments

BitCurator Access Webtools

The [bitcurator-access-webtools](#) project is a [Flask](#) application that allows users to browse file systems in raw and forensically packaged disk images within a web browser. The application can parse raw and E01-packaged images containing FAT16, FAT32, NTFS, HFS+, and EXT 2/3/4 file systems, and allows users to navigate the file system contents, download individual files, and search the contents within a simple web interface.

For more information on the design of the application, along with instructions on how to obtain and build the software, see

Pages 4

BitCurator Access Webtools Downloads

[Download bitcurator-access-webtools](#)

[Read the bitcurator-access-webtools Quick Start Guide](#)

BitCurator Access Redaction Downloads

[Download bitcurator-access-redaction](#)

[Read the bitcurator-access-redaction Quick Start Guide](#)

Clone this wiki locally

<https://github.com/BitCurator/bi>



<https://github.com/BitCurator/bitcurator-access/wiki>

BitCurator / bitcurator-access-webtools Public Notifications Fo

<> Code Issues Pull requests 2 Actions Projects Wiki Security Insights

main 10 branches 52 tags Go to file Code

kamwoods Updated README with EOL notice f9f561f on Apr 11 629 commits

bcaw	Package and syntax updates	3 years ago
conf	FEAT - Custom MIME mapping	5 years ago
disk-images	Added a sampler image	5 years ago
docs	Updated documentation and version, cleaned up attic	5 years ago
externals	Updated documentation and version, cleaned up attic	5 years ago
provision	Package and syntax updates	3 years ago
scripts	FEAT - Compound partition IDs	5 years ago
tests	FEAT - Group image handling	6 years ago
.gitignore	FIX - Tidied warnings	6 years ago
LICENSE	Added full text of GPLv3	6 years ago
README.md	Updated README with EOL notice	2 months ago
Vagrantfile	Package and syntax updates	3 years ago
bcaw.service	Minor cleanup	5 years ago
bcaw_celery_task.py	Minor cleanup	5 years ago
nginx_config	Mods to fix service	6 years ago

About
Tools to browse disk images and file system metadata in a web service
github.com/BitCurator/bitcurator-access-...
python flask web forensics
disk-image
Readme
GPL-3.0 license
19 stars
7 watching
6 forks
Report repository

Releases
52 tags

Packages
No packages published

Contributors 4

<https://github.com/BitCurator/bitcurator-access-webtools>



BitCurator

BitCurator project repositories. See <https://bitcurator.github.io/> for a top level description of how these repos are organized and what they contain.

9 followers <https://bitcurator.github.io/> [@bitcurator](#) bitcurator@gmail.com

[Overview](#) [Repositories 14](#) [Discussions](#) [Projects](#) [Packages](#) [People 3](#)

Pinned

[bitcurator-distro](#) Public
BitCurator Environment: Using, building, and maintaining BitCurator
★ 46

[bitcurator-access](#) Public archive
BitCurator Access: Portal repository for the BitCurator Access tools
★ 6

[bitcurator-salt](#) Public
● SaltStack ★ 2 🍴 5

[bitcurator-nlp](#) Public archive
BitCurator NLP: Portal repository for the BitCurator NLP tools
★ 17 🍴 1

[bitcurator-cli](#) Public
● JavaScript 🍴 1

[bitcurator-access-webtools](#) Public
Tools to browse disk images and file system metadata in a web service
● Python ★ 19 🍴 6

People



Top languages

● Python ● Jav
● SaltStack

Most used topics

[disk-image](#) [py](#)

Repositories

🔍 Find a repository...

Type ▾

Language ▾

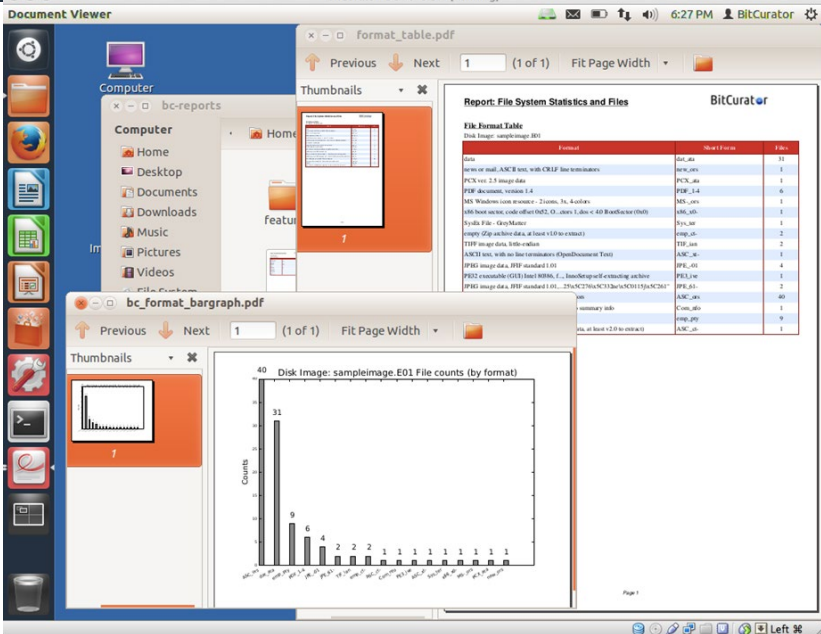
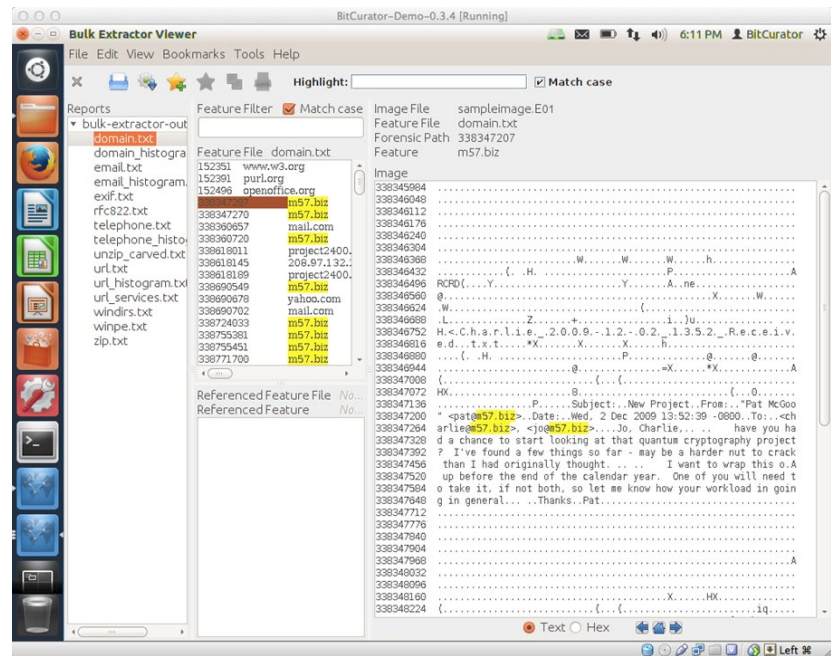
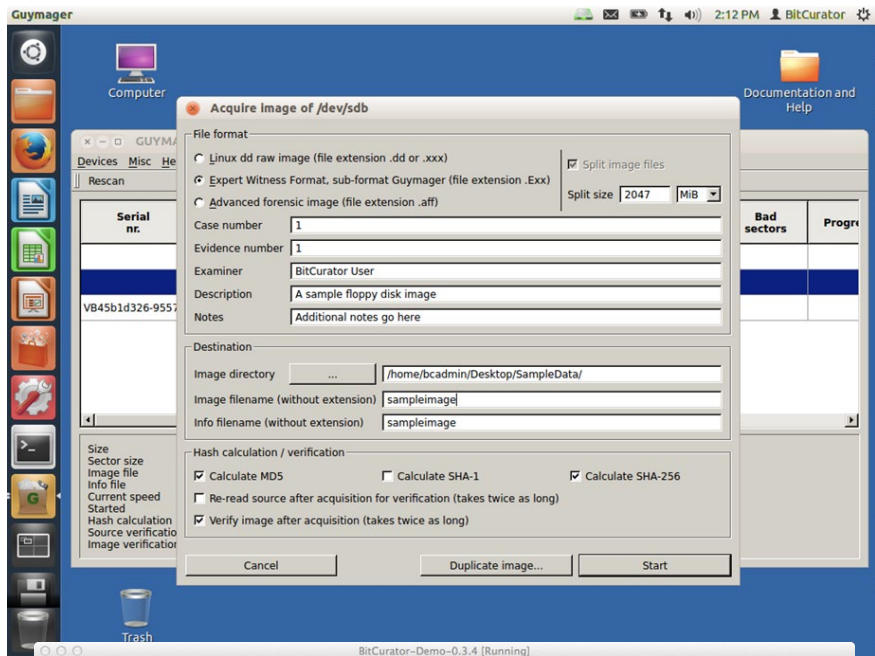
Sort ▾

[bitcurator-access-webtools](#) Public
Tools to browse disk images and file system metadata in a web service
● Python ★ 19 📄 GPL-3.0 🍴 6 🔄 0 📄 2 Updated 4 days ago

<https://github.com/BitCurator>

BITCURATOR

- 1. Disk Imaging:** BitCurator supports the creation of forensic disk images.
- 2. File System Analysis:** The software enables the exploration and analysis of file systems, providing detailed information about file attributes, metadata, and directory structures.
- 3. File Format Identification:** BitCurator includes tools for identifying and validating file formats.
- 4. Metadata Extraction:** The software allows for the extraction of metadata from digital files, including technical metadata such as file size, creation dates, and checksums.
- 5. Bulk Extractor:** BitCurator incorporates Bulk Extractor, a tool used for scanning and extracting specific types of information from large volumes of digital data.
- 6. Reporting and Documentation:** BitCurator offers reporting capabilities to document and record the analysis and preservation activities performed on digital materials.
- 7. Integration with Digital Preservation Systems:** BitCurator is designed to work in conjunction with digital preservation systems and workflows.



<https://blogs.loc.gov/thesignal/2013/12/bitcurators-open-source-approach-an-interview-with-cal-lee/>

END-OF-LIFE (EOL) NOTICE

This research software has reached end-of-life. The code in this repository is no longer actively maintained or supported.

About

The **BitCurator Access Webtools** project allows users to browse file systems contained within disk images using a web browser. It is intended to support access requirements in libraries, archives, and museums preserving born-digital materials extracted from source media as raw or forensically-packaged disk images.

The service uses open source libraries and toolkits including The Sleuth Kit, PyTSK, and the Flask web microservices framework. It uses PyLucene along with format-specific text-extraction tools to index the contents of files contained in disk images, allowing users to search for relevant content without individually inspecting files.

This repository includes a simple build script that deploys the web service as in a VirtualBox VM using Vagrant. It includes several sample images (in the "disk-images" directory) to get you started.

Find out more at <https://github.com/BitCurator/bitcurator-access/wiki>

Getting started

This software uses Vagrant to provision a virtual machine in which **bitcurator-access-webtools** runs. To start, make sure you have VirtualBox and Vagrant installed on your Windows, Mac, or Linux host:

- <http://www.virtualbox.org/>
- <https://vagrantup.com>

Download the latest release (.zip or .tar.gz file) from <https://github.com/BitCurator/bitcurator-access-webtools/releases> and extract the contents. In a terminal, change into the extracted bitcurator-access-webtools-x-x-xx directory (using the release numbers for your release in place of the x's), and make sure the associated Vagrant box (bentu/ubuntu-18.04) is added:

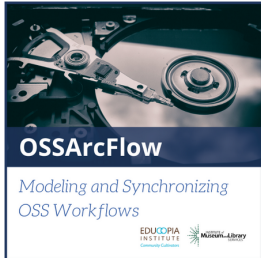
<https://bitcuratorconsortium.org/>

Category Tag [Reset](#) Show per page Search:

DATE	TITLE	SUMMARY	BCC AUTHOR
2015-01-07	A Content Model for Disk Images	A presentation of how to model the data collected with BitCurator This content is shared for reuse with a Creative ...	Matthew Farrell, Duke University Libraries
2019-10-24	A Pinch of Salt: Creating Customized BitCurator builds	Have you ever wanted to customize the BitCurator environment? Thanks to decisions made by the BitCurator developers, modifications can be ...	David Cirella
2019-10-24	Adding "Why" Questions to the BitCurator QuickStart Guide to Build a Comprehensive Graduate Archival Teaching/Learning Module	The BitCurator QuickStart Guide, designed as a software installation and application manual, provides step-by-step instructions for practitioners to download and ...	Jane Zhang
2022-05-26	Advanced Digital Forensics Slides	Description These are the slides from Cal Lee and Kam Woods's "Advanced Digital Forensics" SAA class. There are a number ...	Cal Lee, Kam Woods
2016-11-29	Advanced Topics Webinar – BCA Webtools	This is a recording of a BitCurator Consortium webinar on BitCurator Access WebTools, the third of a series of webinars ...	Cal Lee, Kam Woods, BitCurator Consortium
2017-02-16	Advanced Topics Webinar – bulk_extractor Beyond the Basics	This is a recording of a BitCurator Consortium members-only webinar on bulk_extractor: Beyond the Basics. This webinar focused on advanced ...	Michael Olson, Sandy Ortiz
2016-02-12	Advanced Topics Webinar – Kryoflux	This webinar's focus is on the features, uses, and application of the Kryoflux hardware interface and software tool, including the ...	Dorothy Waugh, Matthew Farrell, Walker Sampson
2016-06-02	Advanced Topics Webinar – Scripting in BitCurator	This webinar's focus is on scripting within the BitCurator environment and is led by Dianne Dietrich (Cornell University) and Jarrett ...	Dianne Dietrich, Jarrett Drake
2017-04-28	All Together Now: Introducing the KryoFlux User Guide	The KryoFlux, a floppy disk controller card developed by the Software Preservation Society, has become the de facto standard for ...	Dorothy Waugh, Shira Peltzman, Jennifer Allen
2017-04-28	Automated Processing of Disk Images and Directories in BitCurator	As a means to more efficiently process large-scale digital archives and with inspiration from Jess Whyte's scripting work at the ...	Tessa Walsh

<https://bitcuratorconsortium.org/resources/>

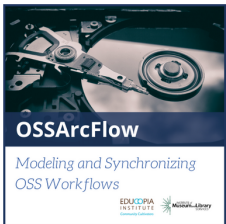
POSTED: AUGUST 18, 2020



OSSArcFlow

The OSSArcFlow project began in 2017 and ended in June 2020. OSSArcFlow was supported by a grant from the Institute of Museum and Library Services. OSSArcFlow project personnel worked with partner institutions to document and analyze born-digital workflows, and created an implementation guide and videos to support workflow documentation and analysis.

For more information, visit the [OSSArcFlow page](#) on the [Educo피아 website](#).



2017-2020

OSSArcFlow

Investigating, Synchronizing, and Modeling a Range of Archival Workflows for Born-Digital Content

Quick Links to Project Deliverables

1. [Digital Dossiers \(All Partners\)](#)
2. [As-Is Workflows \(All Partners\)](#)
3. [Recorded In-Person Partner Meeting Sessions \(YouTube\)](#)
 - [Opening remarks](#)
 - [Panel 1](#)
 - [Panel 2](#)
 - [Panel 3 \(Part One and Part Two\)](#)
4. [Guide to Documenting Born-Digital Archival Workflows](#)
5. [Video Learning Modules \(YouTube\)](#)
 - [Learning Module 1: Common Steps in OSS Born-Digital](#)

Principal Investigator(s):

Katherine Skinner
Sam Meister
Cal Lee

Project Manager(s):

Jessica Meyerson
Alex Chassanoff
Hannah Wang

<https://bitcuratorconsortium.org/ossarcflow/>

OSSArcFlow

Guide to Documenting Born-Digital Archival Workflows

AUTHORS

Alexandra Chassanoff and Colin Post

EDITORS

Katherine Skinner (Lead Editor), Jessica Farrell, Brandon Locke, Caitlin Perry (Copyeditor), Kari Smith, Hannah Wang

CONTRIBUTORS

Christopher A. Lee, Sam Meister, Jessica Meyerson, Andrew Rabkin, Yinglong Zhang

DESIGNER

Hannah Ballard



https://educopia.org/wp-content/uploads/2020/06/OSSArcFlow_Guide_FINAL-1.pdf

<https://educopia.org/ossarcflow/>

STEPS

1. Gather information before acquisition
2. Transfer materials to institution
3. Create disk image
4. Run virus checks
5. File identification & format characterization
6. Check file integrity & ensure fixity
7. Create accession record
8. Analyze & identify sensitive content
9. Analyze forensic/technical metadata
10. Create/extract digital object metadata
11. Assemble AIP
12. Assemble DIP
13. Transfer AIP to preservation environment

6. Check file integrity & ensure fixity

Fixity measures can be used to verify that there have been no undocumented changes to digital objects.

A commonly used example of a fixity measure is a cryptographic hash (also called a checksum). In this step, an institution uses a tool that calculates a numerical value (checksum) for each digital file, and then outputs and/or stores that value. In the future, the institution can use the same tool,

or another tool that uses the same algorithm, to calculate a numerical value for each digital file, and then compare the new value to the old. If the values do not match, then further action may be required to investigate what caused the change, restore the file, document the change, or other next steps that vary based on local institutional practice.

Checking file integrity using the checksum may happen as an institution migrates the digital file from one storage location to another, or as it replicates a file (e.g., for preservation storage), or on a fixed schedule to identify file degradation or compromises. It may also be helpful to provide researchers with the checksum of a file along with the file when it is available for download, or to include the checksum in the descriptive metadata, in order to demonstrate that there was no corruption in file transfer.

CASE STUDY: DUKE UNIVERSITY

"Fixity is monitored by a homegrown system called File Tracker, and replicated copies are created in house, with copies remaining onsite online, copies offline and off site (but local to NC) and another copy offline that we intend to send to a geographically separate site but have not identified the site yet." (2018)

Implementation recommendations

- There are a few different algorithms for creating cryptographic hashes, and these can be distinguished by their names, for example MD5 or SHA-512. Some institutions go through the process of researching different algorithms to inform their decision regarding which algorithm to use.

18

https://educopia.org/wp-content/uploads/2020/06/OSSArcFlow_Guide_FINAL-1.pdf

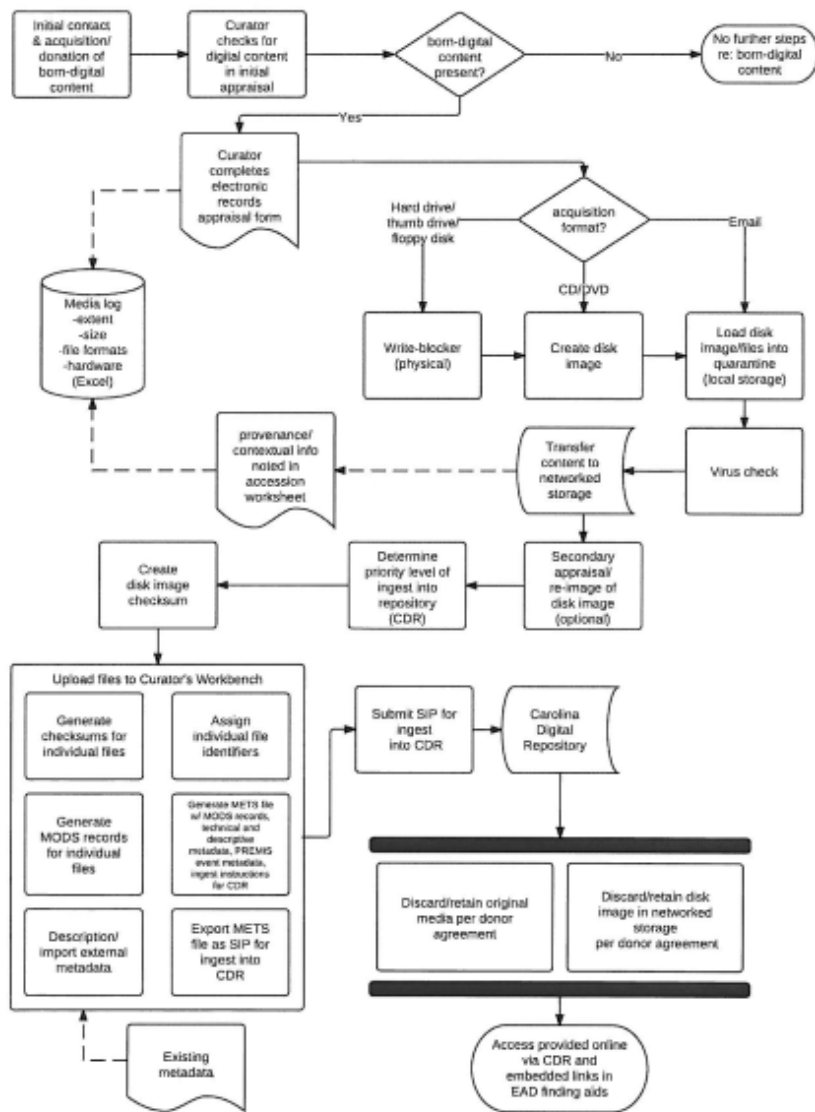


Figure 1: UNC-Chapel Hill born-digital workflow

Gengenbach, M., Chassanoff, A., & Olsen, P. 2012. Integrating digital forensics into born-digital workflows: The BitCurator project. American Society for Information Science and Technology. Meeting. Proceedings of the ... ASIST Annual Meeting, 49(1), 1–4.

<https://doi.org/10.1002/meet.14504901343>

STEPS IN DIGITAL FORENSICS (PART 1)

Identification and Planning:

- Determine the scope and objectives of the investigation.

- Identify the digital devices or storage media to be examined.

- Assess the resources, tools, and expertise required for the investigation.

Acquisition:

- Safely acquire a forensic image or a copy of the original digital media.

- Use specialized tools to ensure the integrity and authenticity of the acquired data.

- Document the acquisition process, including relevant metadata and timestamps.

Preservation:

- Take measures to ensure the preservation and protection of the acquired data.

- Use write-blocking techniques to prevent accidental modification of the evidence.

- Store the acquired data in a secure and controlled environment.

STEPS IN DIGITAL FORENSICS (PART 2)

Examination and Analysis:

- Analyze the acquired data to identify relevant files, folders, and artifacts.

- Use forensic tools and techniques to recover deleted or hidden data.

- Perform keyword searches, hash analysis, metadata examination, and timeline analysis.

- Identify and extract potential evidence, such as documents, emails, chat logs, images, etc.

Reconstruction and Interpretation:

- Reconstruct the sequence of events and activities related to the case.

- Analyze the recovered evidence in the context of the investigation.

- Interpret the findings and draw conclusions based on the available evidence.

Reporting:

- Document the entire investigation process, including the steps taken and the tools used.

- Prepare a detailed report that presents the findings, analysis, and interpretations.

- Include relevant information such as metadata, timestamps, and file hashes.

- Ensure that the report is clear, concise, and suitable for legal proceedings if necessary.

STEPS IN DIGITAL FORENSICS (PART 3)

- Presentation and Testimony:
 - Present the findings and conclusions to relevant stakeholders, such as investigators, legal teams, or courts
 - Provide expert testimony if required, explaining the methodology, analysis process, and findings.
- Archiving and Retention:
 - Safely archive and retain the acquired evidence, ensuring its integrity and availability for future reference.
 - Follow established guidelines and legal requirements for evidence retention.



National Archives of Ireland Reading Room
(c) National Archives of Ireland.



The Central Reading Room in the Archives and Library Building, Yad Vashem Archives © Yad Vashem



Vatican Apostolic Archives, ©
<https://www.vaticannews.va/en/vatican-city/news/2021-03/papal-archives-vatican-open-world-sergio-pagano.html>

BitCurator Project

Written by Sarah Lake, Junior Digital Archivist
September 2019



Context and Goals

Since August 2019, John Richan, Digital Archivist and Sarah Lake, Junior Digital Archivist have been working on a project to bridge gaps in digital archives processing workflows using open-source digital forensics applications found in the BitCurator Environment. This project, which is funded in part by a grant from the Young Canada Works Program, is part of the development of RMA's Digital Preservation Lab which is set to launch in 2020.

About BitCurator

From the [BitCurator Consortium](#) website:

The BitCurator Environment is a Ubuntu-derived Linux distribution geared towards the needs of archivists and librarians. It includes a suite of open source digital forensics and data analysis tools to help collecting institutions process born-digital materials. BitCurator supports positive digital preservation outcomes using software and practices adopted from the digital forensics community.

- In the BitCurator Environment you can:
 - Create forensic disk images: Disk images packaged with metadata about devices, file systems, and the creation process.
 - Analyze files and file systems: View details on file system contents from a wide variety of file systems.

BITCURATOR USERS FORUM Virtual Conference took place Oct. 12 - 14, 2021

John Richan (RMA Digital Archivist) and Sarah Lake (CU Digital Preservation Librarian) presented *BitCurator from Scratch: Internship to Production Environment Implementation at Concordia University* on Oct. 13th, Session Three.

[View their presentation](#)

DIGITAL ARCHIVES
PRESERVATION STRATEGY

ARCHIVES & SPECIAL
COLLECTIONS
SHARED CATALOGUE

<https://www.concordia.ca/offices/archives/digital-preservation-bitcurator.html>



INTERPOL

GLOBAL GUIDELINES FOR DIGITAL FORENSICS LABORATORIES

INTERPOL For official use only

May 2019

Global Guidelines for Digital Forensics Laboratories

file:///C:/Users/Administrator/Downloads/INTERPOL_DFL_GlobalGuidelinesDigitalForensics.pdf

The DFL must also consider having a Case Management System for its operation. This system will hold the database of cases, exhibits, Examiner's name and forensic results. Minimum entries for a Case Management System are as follows:

- Date, time and person delivering and receiving the exhibit to the DFL.
- Unique exhibit reference number.
- Unique case reference number.
- Requester's name and contact details.
- Type of crime and related act.
- The names of DFL staff that have had contact with the exhibits.
- The narrative for the case request.
- Time factors – such as delivery dates and anticipated court dates.
- Analysis process – Imaging, examination, extraction, calculated hash values, etc.
- Exact date and time of analysis conducted, as well as the Examiner's name.
- Details of quality assurance by colleagues and managers.
- Result of the analysis.
- Record of communication with the Requester.

All this information is vitally important to show continuity, credibility and verification of actions and evidence. The DFL can create the system internally, purchase an off-the-shelf solution, or hire a programmer to develop the system.

The following figure shows the laboratory analysis model:



Figure 5. Digital Forensics Laboratory Analysis Model

Global Guidelines for Digital Forensics Laboratories

file:///C:/Users/Administrator/Downloads/INTERPOL_DFL_GlobalGuidelinesDigitalForensics.pdf, p 25 and p 30

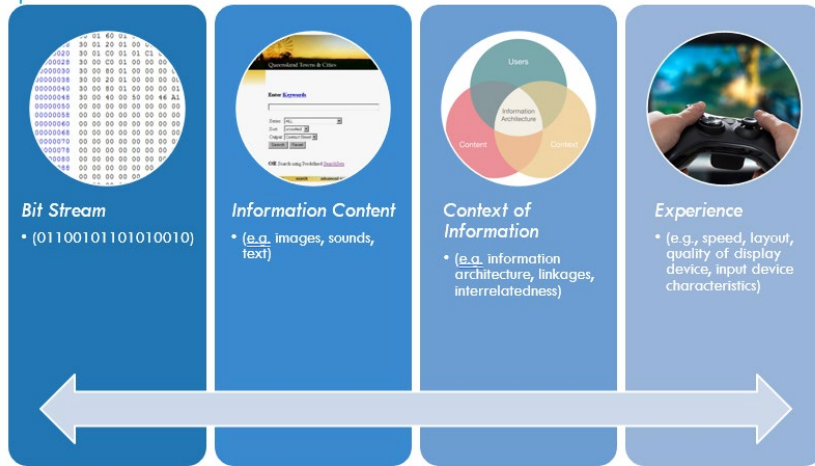
	Dead Acquisition	Live Acquisition
What?	<p>Dead acquisition is conducted on a dead system. A dead system is a system that is not running; turned off, with no power.</p> <p>When the system is dead, volatile data in temporary storage areas such as RAM memory, running processes, cache or active application dialogues on a computer will no longer be available.</p>	<p>Live acquisition is conducted on a live system. A live system is a system that is up and running where information may be altered as data is continuously being processed.</p> <p>Because of the rich evidentiary value that could be discovered in a live system, switching it off may cause loss of volatile data, such as data stored on the cloud, encrypted data, running process, network connected and mounted file system.</p>
How?	<p>The process of conducting dead acquisition is straightforward as it is normally done automatically using forensic equipment.</p> <p>The hard disk must first be taken out of the computer before connecting it to the equipment, if possible.</p> <p>In some cases netbook computers or devices with soldered solid state drive storage cannot be extracted in dead acquisition. Other methods to perform extraction in such cases, like booting the system with a live CD/USB, should be considered.</p>	<p>Data on a system have different levels of volatility. These data will be lost if the system is switched off or rebooted. Whenever the Examiner acquires live data, it is sensible to collect from the most volatile data to the least volatile.</p> <p>The typical level of volatility, from the most to least volatile is as follows:</p> <ul style="list-style-type: none"> • Memory • Swap File • Network Processes • System Processes • File System Information
When?	<p>Dead acquisition is conducted when:</p> <ul style="list-style-type: none"> • System is switched off • Deleted data is more important than volatile data 	<p>Live acquisition is conducted when:</p> <ul style="list-style-type: none"> • The system is business-critical and cannot be shut down • Volatile data are more important than deleted data

The common process for conducting data acquisition is illustrated in the following figure:

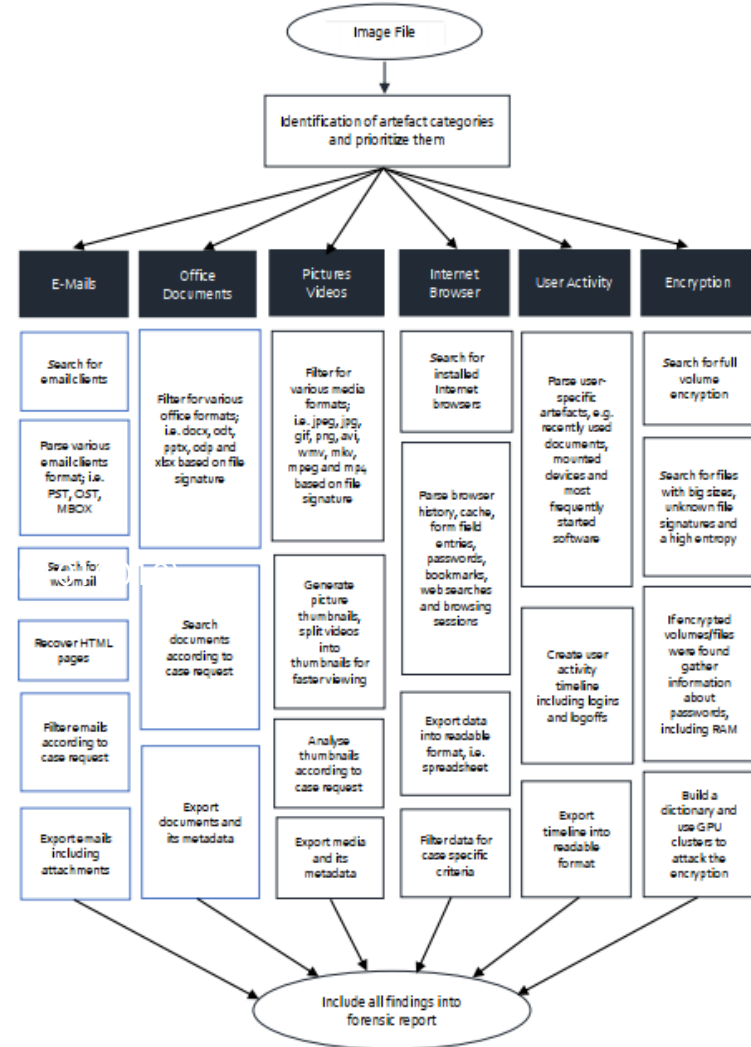


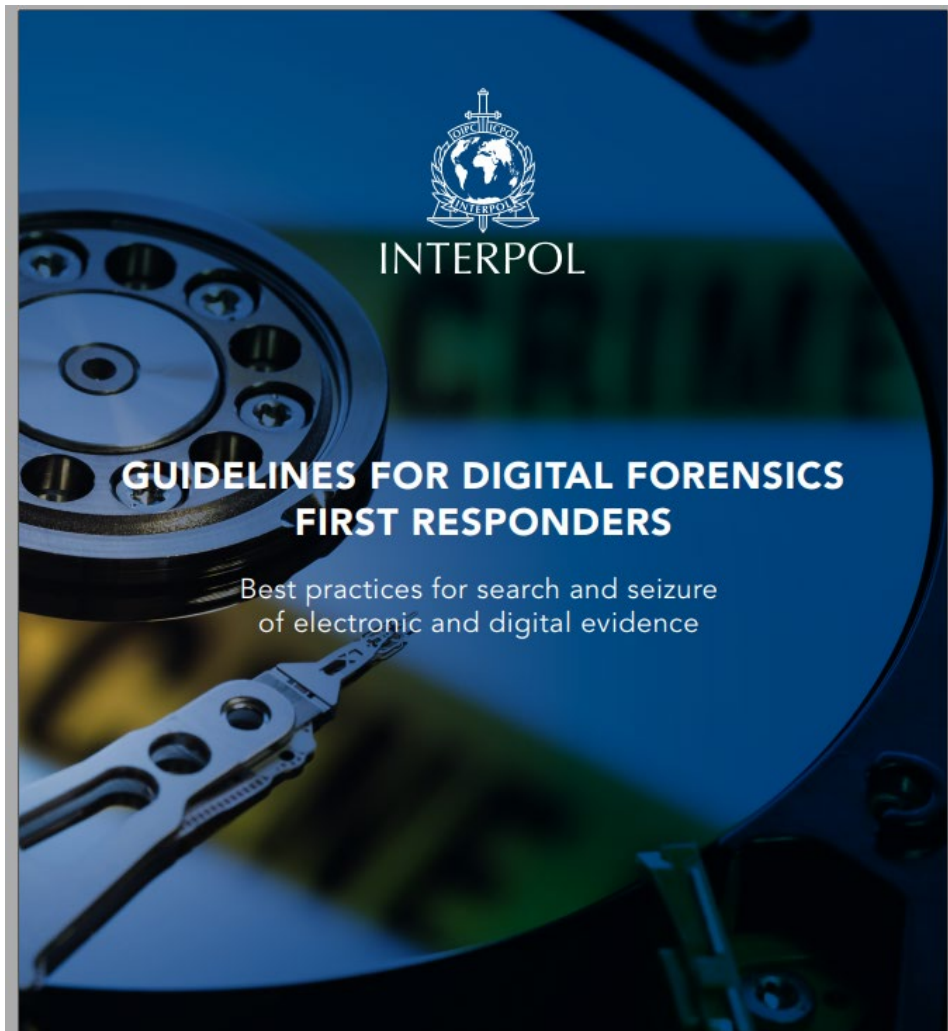
Figure 6. Data Acquisition Process on Computer

Table 6. Method for conducting acquisition - Dead Acquisition and Live Acquisition



APPENDIX I: PROCESS OF ANALYSING EXHIBIT'S ARTEFACTS





Interpol, 2021, Guidelines for Digital Forensics First Responders: Best practices for search and seizure of electronic and digital evidence, file:///C:/Users/Administrator/Downloads/Guidelines_to_Digital_Forensics_First_Responders_V7.pdf

If a specific tool like FiRST is not available, you could consider the list shown below based on the list created by Kuhlee and Völzow (*Computer Forensik Hacks*, O'Reilly, ISBN 978-3-86899-121-5), aimed at facilitating the choice of the most appropriate tool to capture specific fragments of volatile data.

Volatile Fragment	Windows tools	Linux tools
RAM content	Dumpit, Winen, Mdd, FTK Imager	dd, fmem
Routing table, ARP cache, Kernel statistics	Route PRINT, arp -a, netstat	netstat -r -n route arp -a
DNS cache	Ipconfig/displaydns	mdc dumpdb (if installed)
List of running processes	PsList, ListDLLs, CurrProcess, tasklist	ps -ef, lsof
Active network connections		netstat -a, ifconfig
Programs and services using the network	sc queryex, netstat -ab	netstat -tunp
Open files	Handle, PsFile, Openfiles, net file	lsof, fuser
Network shares	Net share, Dumpsec	showmount -e, showmount -a smbclient -L
Open ports	OpenPorts, ports, netstat -an	netstat -an, lsof
Connected users	Psloggedon, whoami, ntlast, netusers /l	w, who -T, last
Encrypted archives	Manage-bde (Bitlocker), efsinfo (EFS)	mount -v, ls /media

01¥01010101010101\$01010101010101€01010101010101¥0101010101

Active network shares	Fsinfo, reg (mounted Devices)	mount -v, ls /media
Remote accesses and network monitoring	Psloglist	/etc/syslog.conf Port UDP 514
System and network configuration	Systeminfo, msinfo32, ipconfig /all	ifconfig -a netstat -in
Storage devices	Reg (Mounted Devices), Net share, netstat -a	mount -v, ls /media
Date and time	Time /T, date /T, uptime	time, date, uptime
Environment variables	Cmd /c set	env, set
Clipboard	pclip	
Disk content	FTK Imager, EnCase, Tableau Imager	dc3dd, ewfacquire, Guymager

Many of these tools are available on the Microsoft Sysinternals website or in the official Linux repositories.

Interpol, 2021, Guidelines for Digital Forensics First Responders: Best practices for search and seizure of electronic and digital evidence,
file:///C:/Users/Administrator/Downloads/Guidelines_to_Digital_Forensics_First_Responder_s_V7.pdf

NEXT STEPS

- ❖ Develop policy frameworks and best-practice agreements for donor relations, liability, workflows, and researcher access
- ❖ Develop regional networks for collaboration
- ❖ Define requirements for and develop new tools
- ❖ Aid in articulating a scholarly research agenda
- ❖ Collect more stories and case studies
- ❖ Facilitate training
- ❖ Encourage cross-publication of research literature and crosspromotion of professional events
- ❖ Pursue terminology mapping



RÉSUMÉ Une grande partie des discussions récentes dans le domaine des études archivistiques au sujet de la justice sociale ont adopté un cadre légaliste axé sur les droits pour définir le rôle des documents, des centres d'archives et des archivistes tant dans les questions de violations des droits humains que pour tenir les individus et les gouvernements responsables quant aux questions des droits humains de base, tels le droit à la vie, à la vie privée et à la liberté d'expression. Pourtant, depuis des décennies les écrits scientifiques féministes ont mis en doute l'universalité d'un cadre axé sur les droits, affirmant plutôt que l'éthique de la sollicitude est un modèle plus inclusif et plus pertinent pour envisager et mettre en place une société plus juste. Cet article propose le changement du modèle théorique dont se servent les archivistes et les spécialistes en études archivistiques pour répondre aux questions de justice sociale – remplaçant celui basé sur les droits individuels par celui basé sur l'éthique féministe. Dans l'approche d'éthique féministe, les archivistes sont perçus comme gardiens responsables, liés aux créateurs de documents, aux sujets, aux utilisateurs et aux communautés grâce à un réseau de liens de responsabilités qui sont mutuellement affectifs. Cet article propose quatre changements inter-reliés dans ces rapports archivistiques, basés sur une empathie radicale.

ABSTRACT Much recent discussion about social justice in archival studies has assumed a legalistic, rights-based framework to delineate the role of records, archives, and archivists in both the violation of human rights and in holding individuals and governments accountable for basic human rights, such as the right to life, privacy, and freedom of expression. Yet decades of feminist scholarship have called into question the universality of a rights-based framework, arguing instead that an ethics of care is a more inclusive and apt model for envisioning and enacting a more just society. This article proposes a shift in the theoretical model used by archivists and archival studies scholars to address social justice concerns – from that based on individual rights to a model based on feminist ethics. In a feminist ethics approach, archivists are seen as caregivers, bound to records creators, subjects, users, and communities through a web of mutual affective responsibility. This article proposes four interrelated shifts in these archival relationships, based on radical empathy.

Ethics of Care

- ❖ “...decades of feminist scholarship have called into question the universality of a rights-based framework, arguing instead that an ethics of care is a more inclusive and apt model for envisioning and enacting a more just society” (Caswell & Cifor, 2016, p 24)
- ❖ “Affective responsibilities should be marked by radical empathy” (ibid., 25)
- ❖ “...archivist has an affective responsibility to responsibly emphasize with each of the stakeholders...” (ibid., 41)

Michelle Caswell and Marika Cifor, 2016, “From Human Rights to Feminist Ethics: Radical Empathy in the Archives,” *Archivaria* 81 (Spring), pp., 23--43.

AFFECTIVE RESPONSIBILITY, THE RELATIONSHIP:

“between the archivist and record creator” --
“acknowledge ethical bond, but also hinges an ethical orientation on it” (ibid., 33)

“between archivist and the subject of the records” –
“responsibility to those about whom the records are created” (ibid., 36)

“between the archivist and the user” – “sometimes allowing for affect can be as simple as giving the user space and time to feel” (ibid., 37-38)

“between the archivist and the larger community” – “to the future” (ibid., 39-41).



THE EXERCISES — PURPOSE AND GOALS

USEFUL RESOURCES

Matthew G. Kirschenbaum, Richard Ovenden, and, Gabriela Redwine, 2010, Digital Forensics and Born-Digital Content in Cultural Heritage Collections, Council on Library and Information Resources, Washington, D.C., <https://www.clir.org/pubs/reports/pub149/>

Global Guidelines for Digital Forensics Laboratories
file:///C:/Users/Administrator/Downloads/INTERPOL_DFL_GlobalGuidelinesDigitalForensics.pdf

Interpol, 2021, Guidelines for Digital Forensics First Responders: Best practices for search and seizure of electronic and digital evidence,
file:///C:/Users/Administrator/Downloads/Guidelines_to_Digital_Forensics_First_Responders_V7.pdf