

# Digital preservation at Wellcome

Alex Chan ~ [a.chan@wellcome.ac.uk](mailto:a.chan@wellcome.ac.uk) ~ they/them  
Senior software developer at Wellcome Collection

Image: The Anatomical Theatre at Cambridge.  
Published for R. Ackermann's *History of Cambridge*  
Wellcome Collection, CC BY.





# Digital preservation at Wellcome

**Alex Chan ~ a.chan@wellcome.ac.uk ~ they/them  
Senior software developer at Wellcome Collection**

Image: The Anatomical Theatre at Cambridge.  
Published for R. Ackermann's *History of Cambridge*  
Wellcome Collection, CC BY.







Wellcome Collection is a free museum and library that aims to challenge how we all think and feel about health





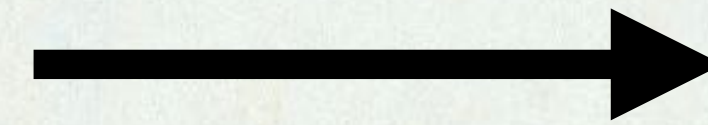




Production



Storage

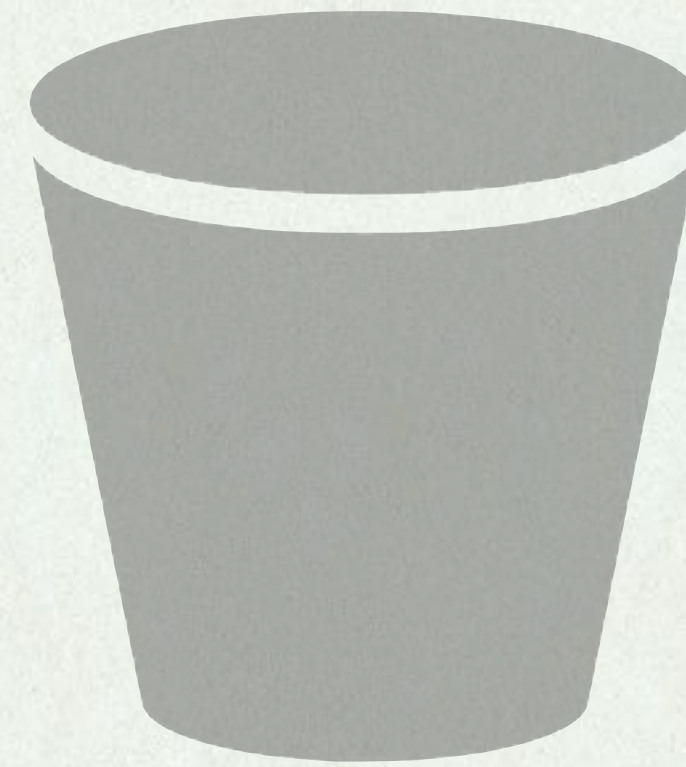
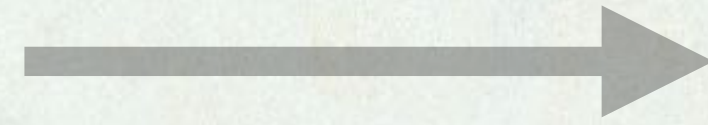


Retrieval

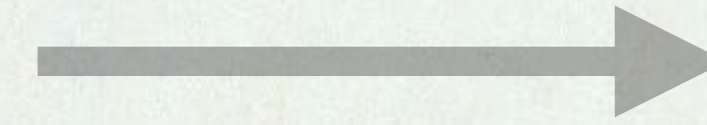




Production



Storage



Retrieval

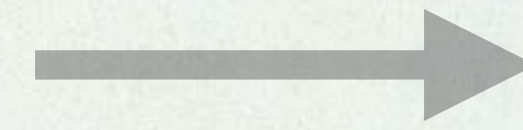


## **digitised assets**

planned digitisation/  
ad-hoc request



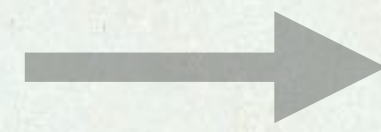
photography



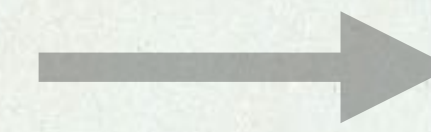
ingested with  
Goobi

## **born-digital assets**

files transferred  
to Wellcome



files appraised  
by archivists

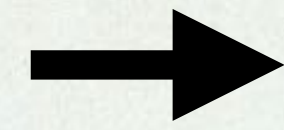


catalogued in  
Archivematica

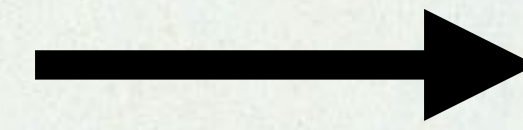


## **digitised assets**

planned digitisation/  
ad-hoc request



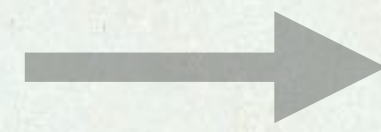
photography



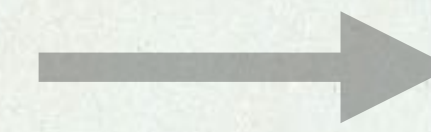
ingested with  
Goobi

## **born-digital assets**

files transferred  
to Wellcome



files appraised  
by archivists

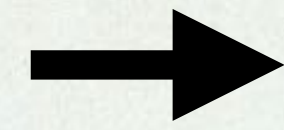


catalogued in  
Archivematica

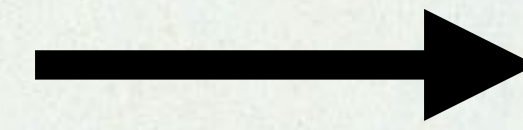


## **digitised assets**

planned digitisation/  
ad-hoc request



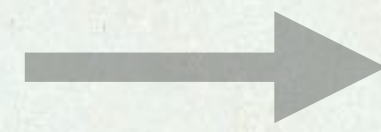
photography



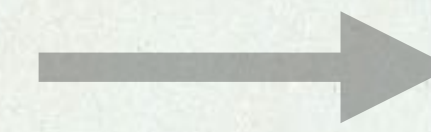
ingested with  
Goobi

## **born-digital assets**

files transferred  
to Wellcome



files appraised  
by archivists



catalogued in  
Archivematica







# Goobi Intranda

Home page My tasks Workflow Administration Controlling ? 13:50:22 Donnerstag, 21. Mai 2015 Steffen Hankiewicz

## Processes

Home page > Processes

Processes

Hits: 48

ID	Process title	Status	Project	Action
13	> alphrudeh_PPN591416573		Manuscripts	
14	> alsdehou_FPN560091583		Manuscripts	
113	> austPrIn_000145107		Manuscripts	
35	> bergsphi_62501/145		Manuscripts	
23	> bianjani_618586231		Manuscripts	
16	> biannoti_PPN523090374		Manuscripts	
21	> hojaanat_620650583		Manuscripts	
38	> clainelv_578780275_0001		Manuscripts	
19	> clainelv_PPN578780275_0001		Manuscripts	
42	> dissinsis_626739478		Manuscripts	

<< < Previous page Page 1 of 5 Next page >>

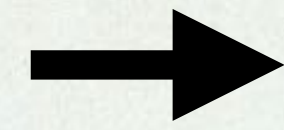
### Possible actions

- Export metadata for DMS
- Set status of process up?
- Set status of process down?
- Execute GoobiScript
- Export search result
- Calculate number of metadata and images
- Statistical evaluation

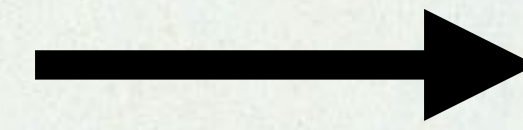


## **digitised assets**

planned digitisation/  
ad-hoc request



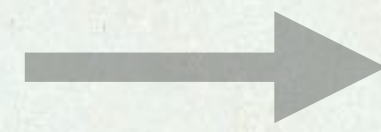
photography



ingested with  
Goobi

## **born-digital assets**

files transferred  
to Wellcome



files appraised  
by archivists



catalogued in  
Archivematica



## digitised assets



## born-digital assets





# Archivematica

## Artefactual

The screenshot shows the Archivematica web interface. At the top, there is a navigation bar with the following items: Archivematica, Transfer (5), Backlog, Appraisal, Ingest (1), Archival storage, Preservation planning, Access, Administration, admin, and Connected. Below the navigation bar, there is a form for creating a new transfer. The form includes a dropdown menu for 'Transfer type' set to 'Standard', and three input fields for 'Transfer name', 'Accession no.', and 'Access system ID'. There are 'Browse' and 'Start transfer' buttons. A checkbox for 'Approve automatically' is checked. Below the form, there is a table of transfers with the following columns: Transfer, UUID, and Transfer start time. The table contains three rows of transfer data, each with a list of microservices that have been executed for that transfer.

Transfer	UUID	Transfer start time
PPTHW_2466	f1058692-18b1-4221-8a15-a1653a8a4f43	2019-10-02 15:43
<ul style="list-style-type: none"><li>Microservice: Create SIP from Transfer</li><li>Microservice: Complete transfer</li><li>Microservice: Examine contents</li><li>Microservice: Validation</li><li>Microservice: Parse external files</li><li>Microservice: Generate METS.xml document</li><li>Microservice: Characterize and extract metadata</li><li>Microservice: Extract packages</li><li>Microservice: Identify file format</li><li>Microservice: Clean up names</li><li>Microservice: Generate transfer structure report</li><li>Microservice: Scan for viruses</li><li>Microservice: Quarantine</li><li>Microservice: Verify transfer checksums</li><li>Microservice: Reformat metadata files</li><li>Microservice: Assign file UUIDs and checksums</li><li>Microservice: Include default Transfer processingMCP.xml</li><li>Microservice: Rename with transfer UUID</li><li>Microservice: Verify transfer compliance</li><li>Microservice: Approve transfer</li></ul>		
PPTHW_2466	41099e05-63b7-4d22-9579-7490aac8992c	2019-10-02 15:43
<ul style="list-style-type: none"><li>Microservice: Failed transfer</li><li>Microservice: Approve transfer</li></ul>		
b30011826	c52062ee-4594-4af2-9671-1a91fe36d353	2019-06-08 09:26
<ul style="list-style-type: none"><li>Microservice: Create SIP from Transfer</li><li>Microservice: Complete transfer</li><li>Microservice: Examine contents</li><li>Microservice: Validation</li><li>Microservice: Parse external files</li><li>Microservice: Generate METS.xml document</li><li>Microservice: Characterize and extract metadata</li><li>Microservice: Extract packages</li><li>Microservice: Identify file format</li><li>Microservice: Clean up names</li></ul>		



## **digitised assets**



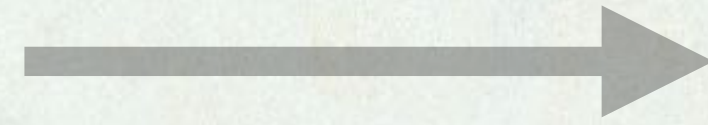
## **born-digital assets**



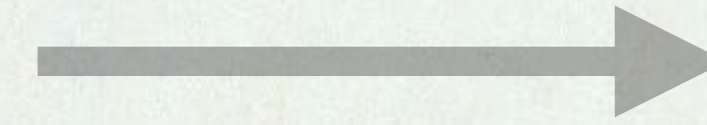




Production

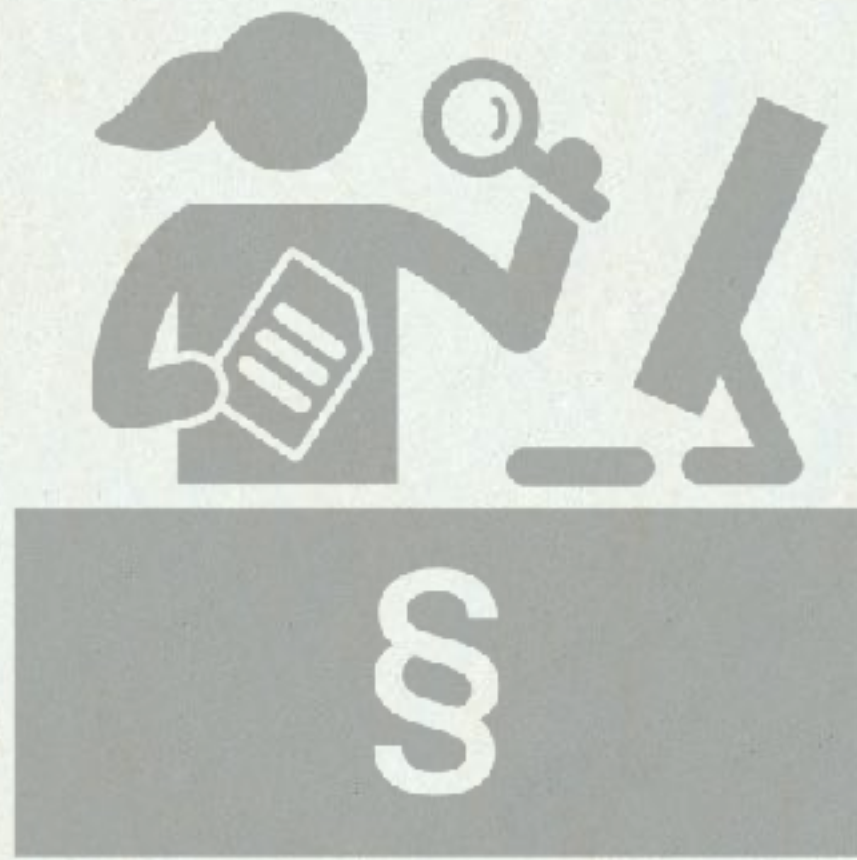


Storage

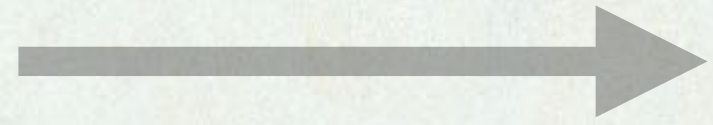


Retrieval

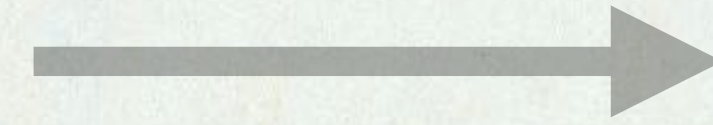




Production



Storage



Retrieval

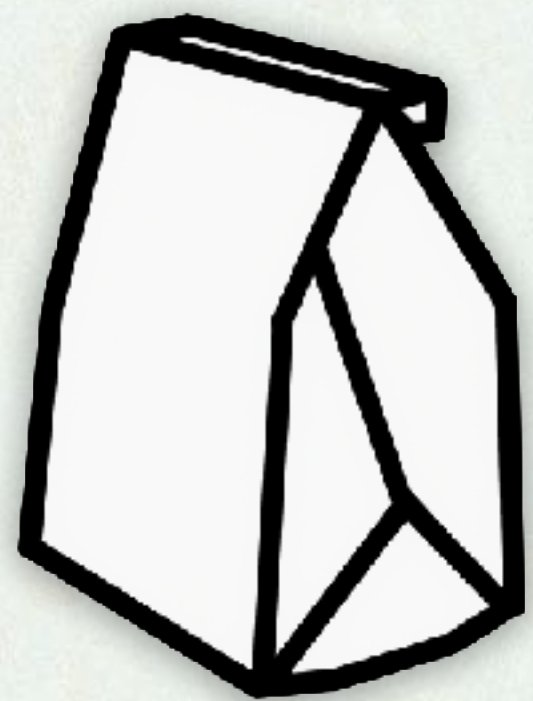




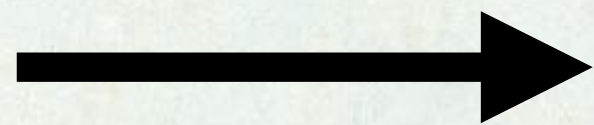
**Preservica**

Active digital preservation

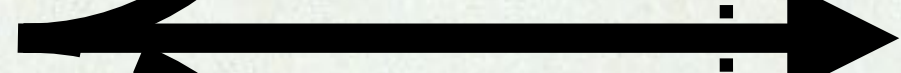




BagIt



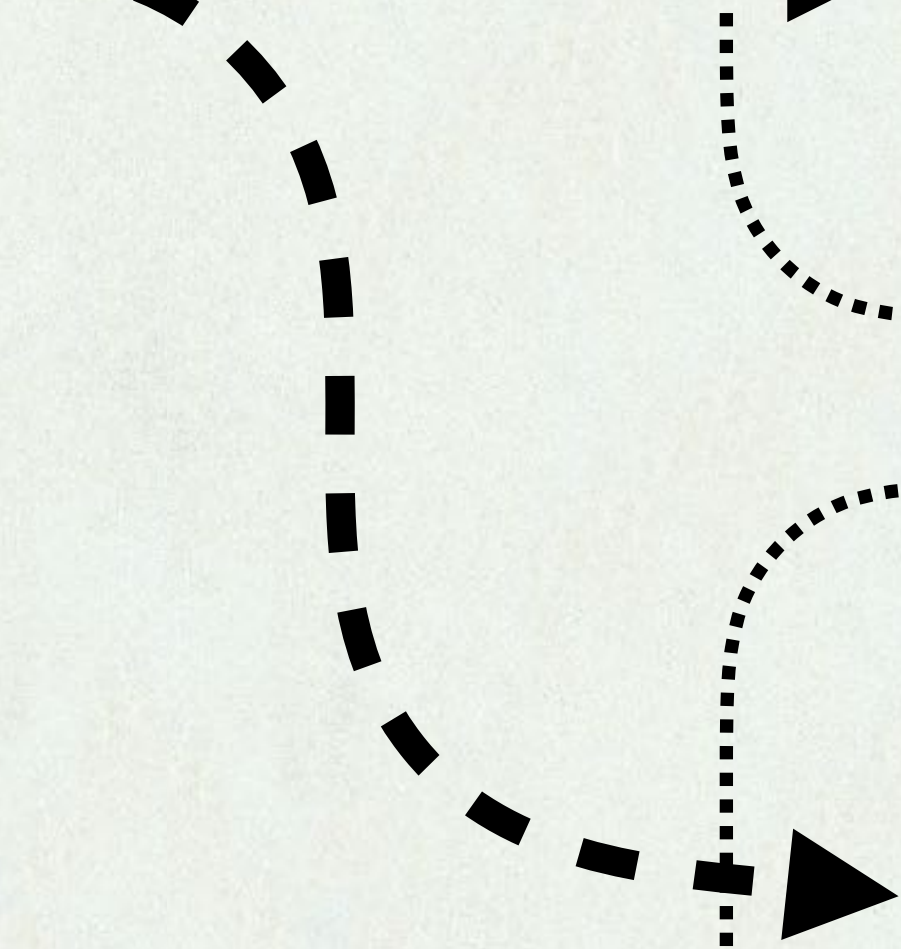
Storage  
service



S3



S3 Glacier



Azure







Storage

Open-source storage service

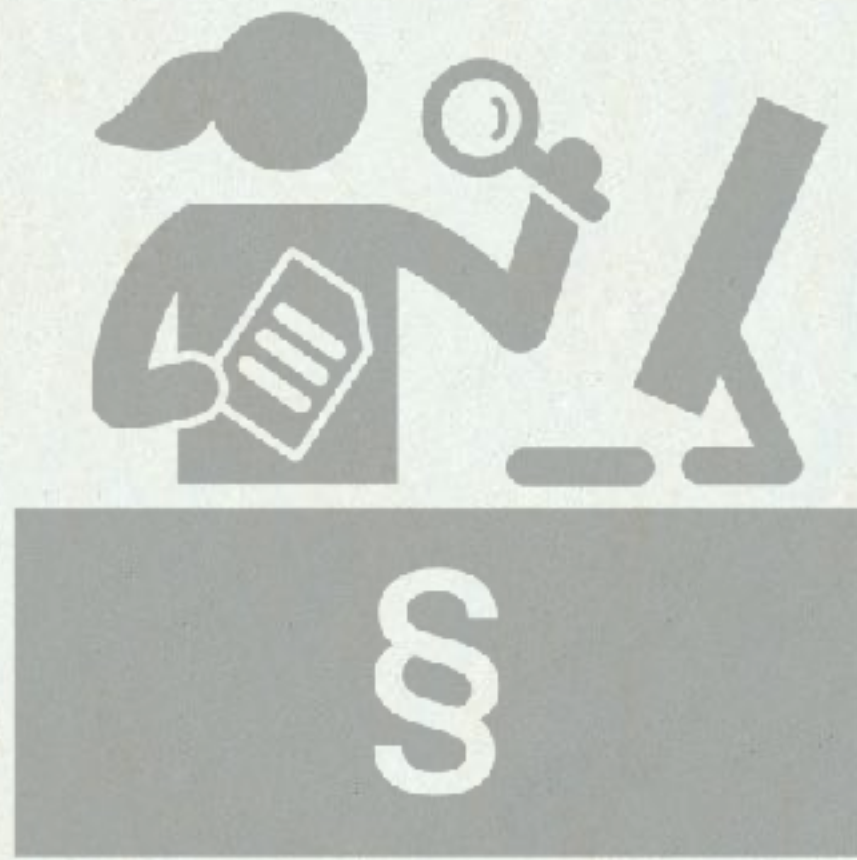
Bagit bags for ingest

Multiple layers of verification

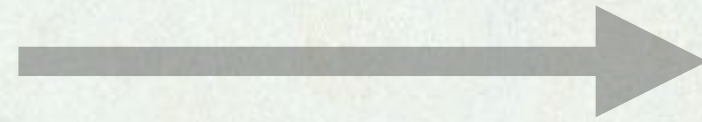
Running in AWS

Human-readable storage in  
AWS (Dublin) / Azure (Amsterdam)





Production

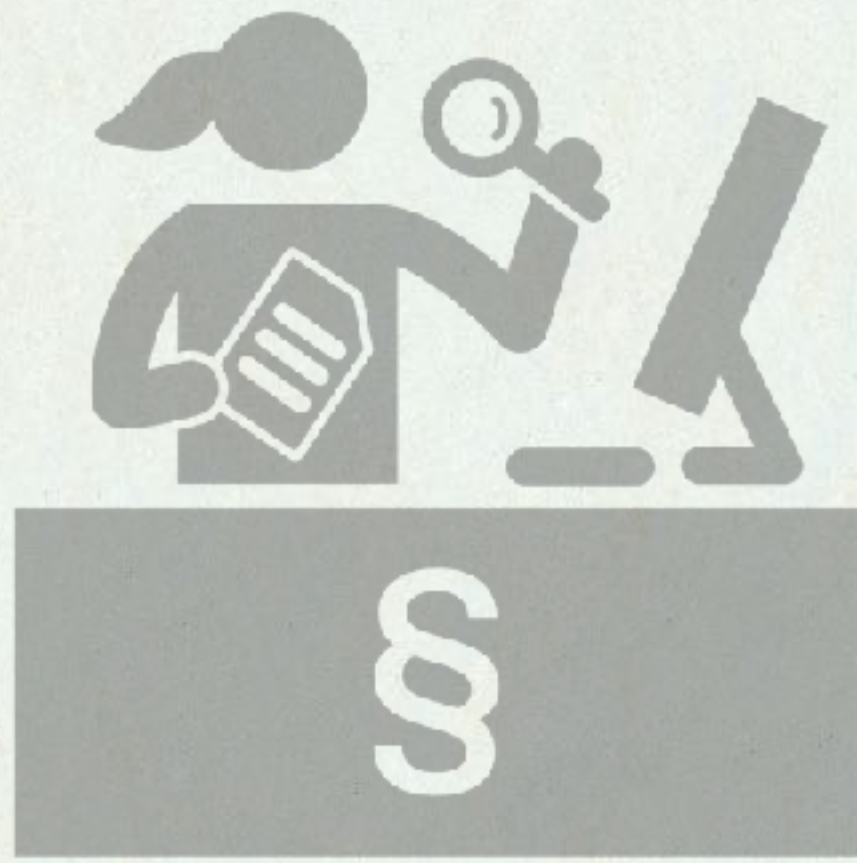


Storage

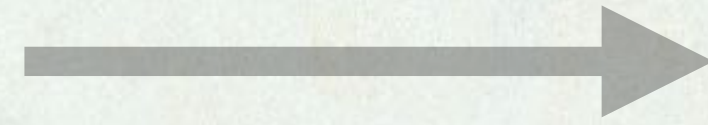


Retrieval

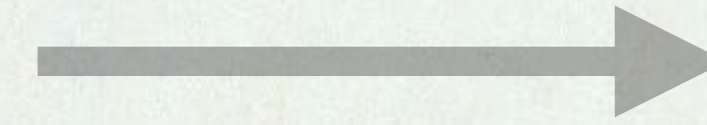




Production

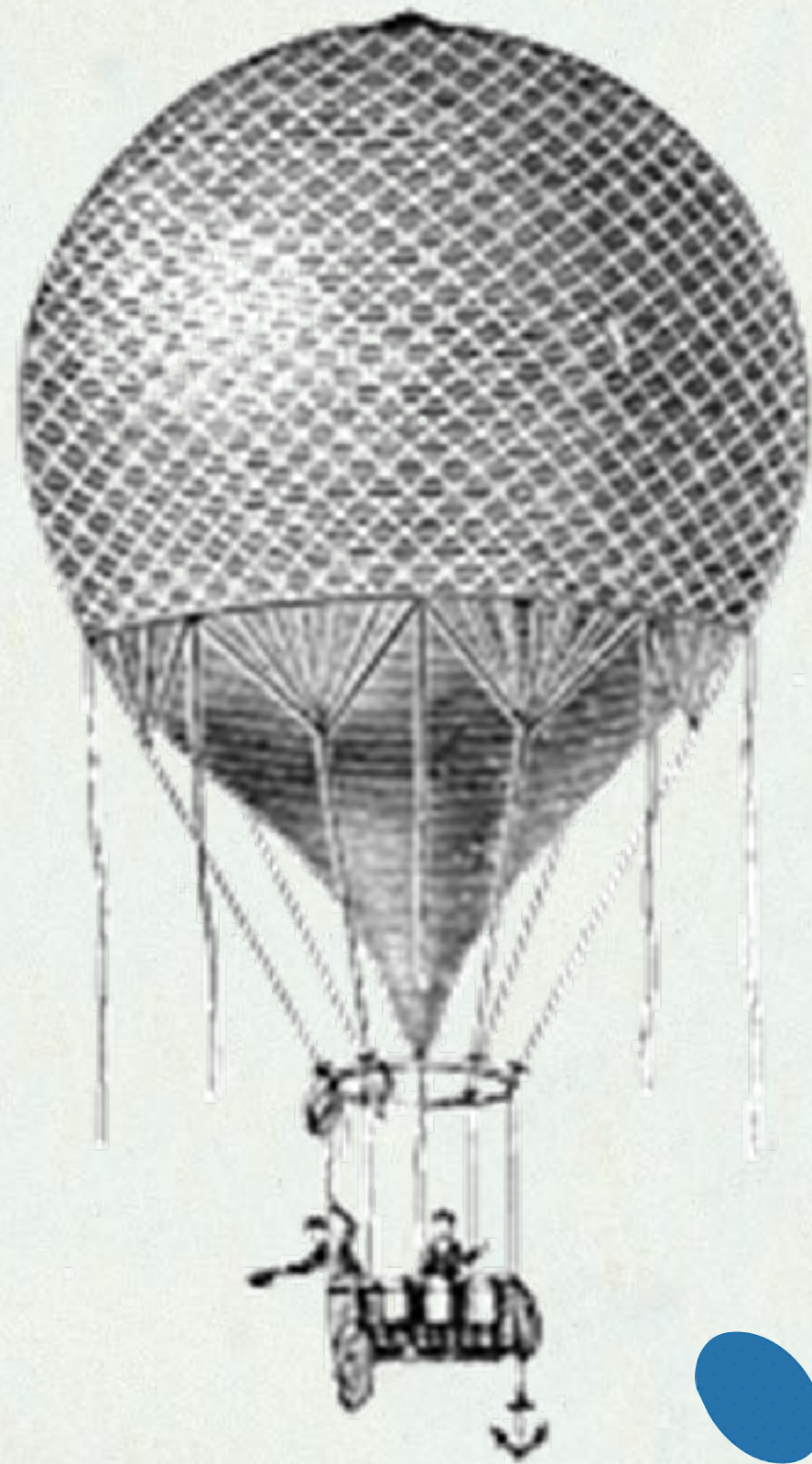


Storage



Retrieval





# DLCS/DDS

Made by Digirati

IIF Presentation, Image  
and Auth APIs





THE SLENDER LORIS, IN WAKING AND SLEEPING POSTURE.

# Loris

Open-source  
IIIF Image API server





**unified catalogue search** 





marie curie

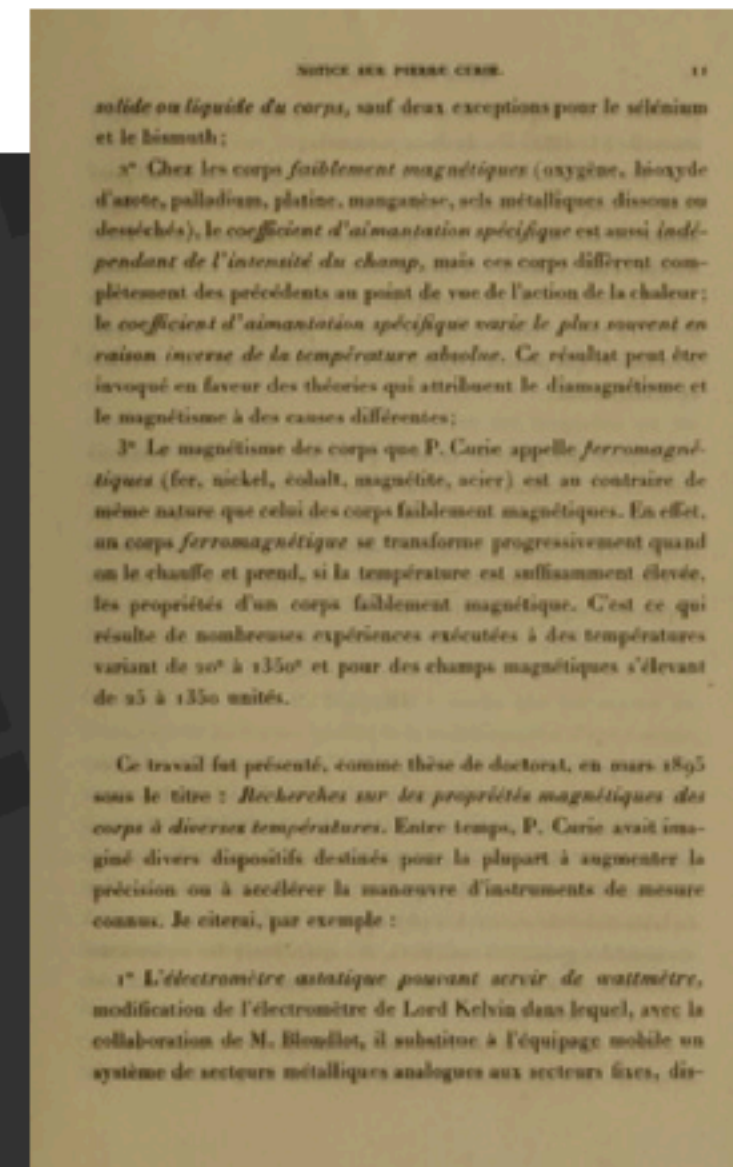
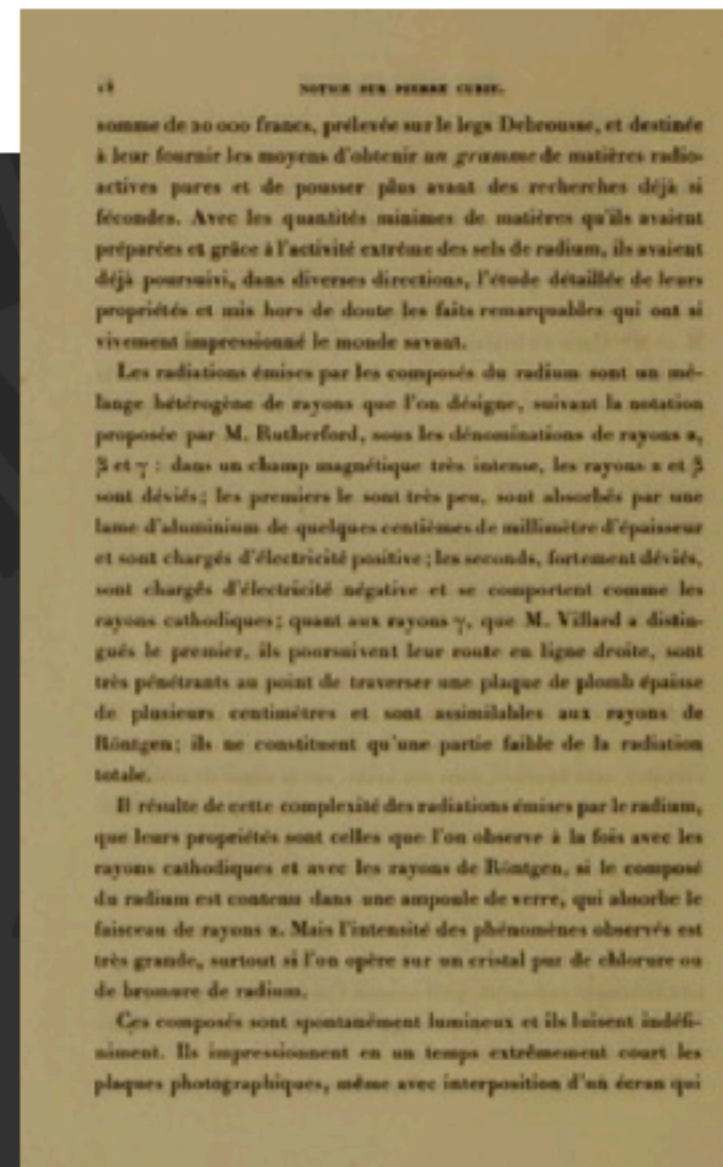
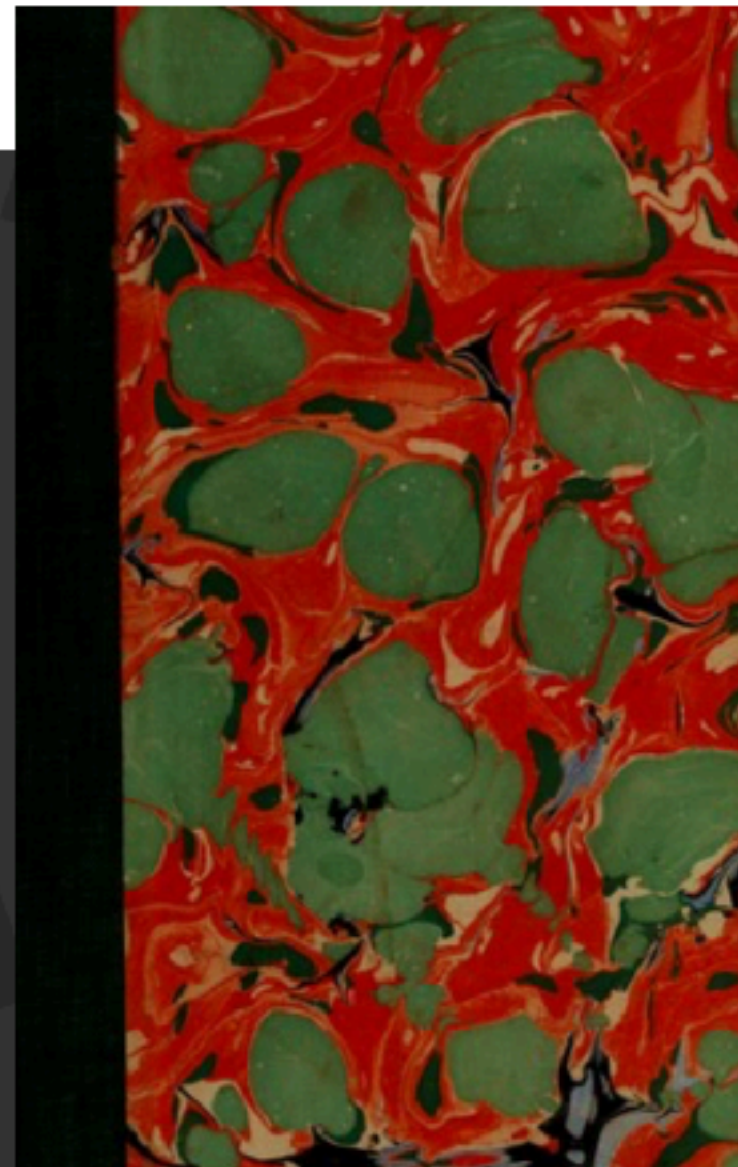
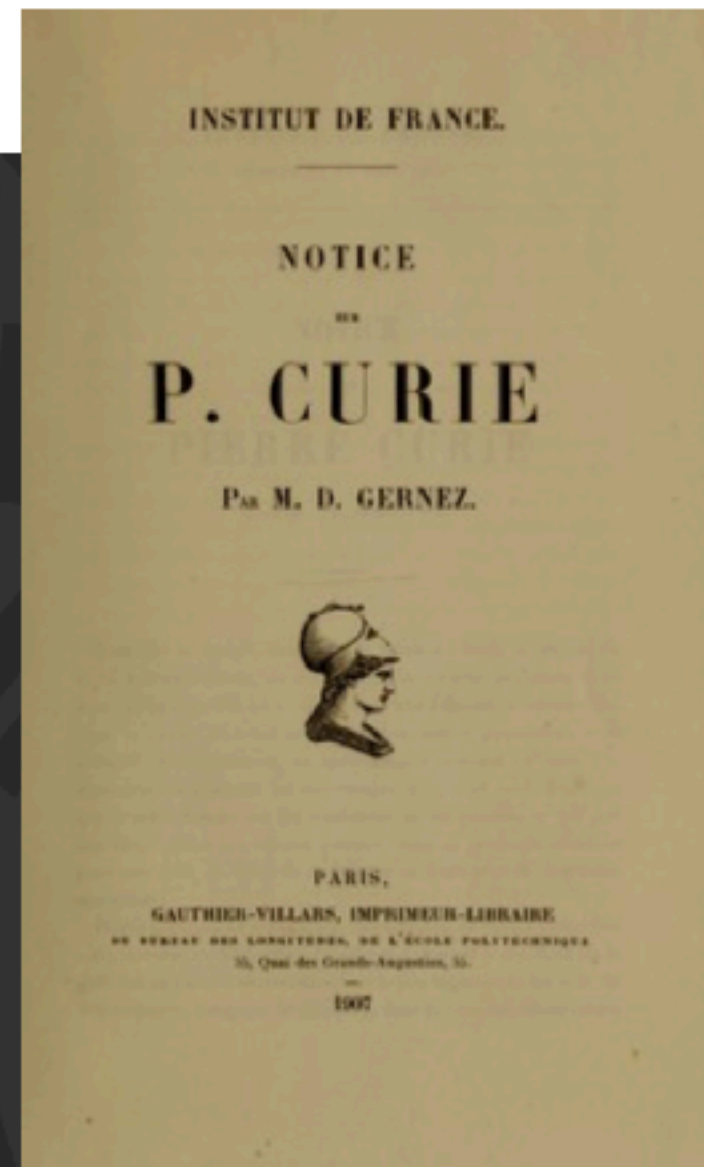


Search results

E-books

## Notice sur P. Curie

Gernez, D. 1834-1910. Date 1907



36 images

Download

Public Domain

[Can I use this?](#)





Retrieval

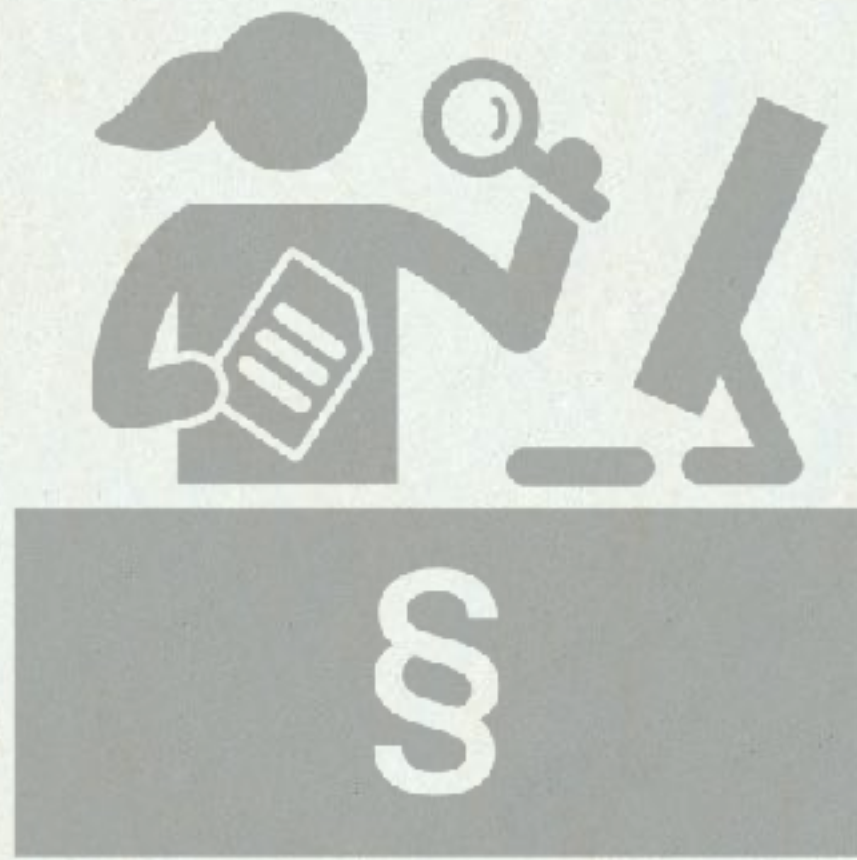
DLCS: IIF Presentation, Image and Auth APIs

Loris: IIF Image API

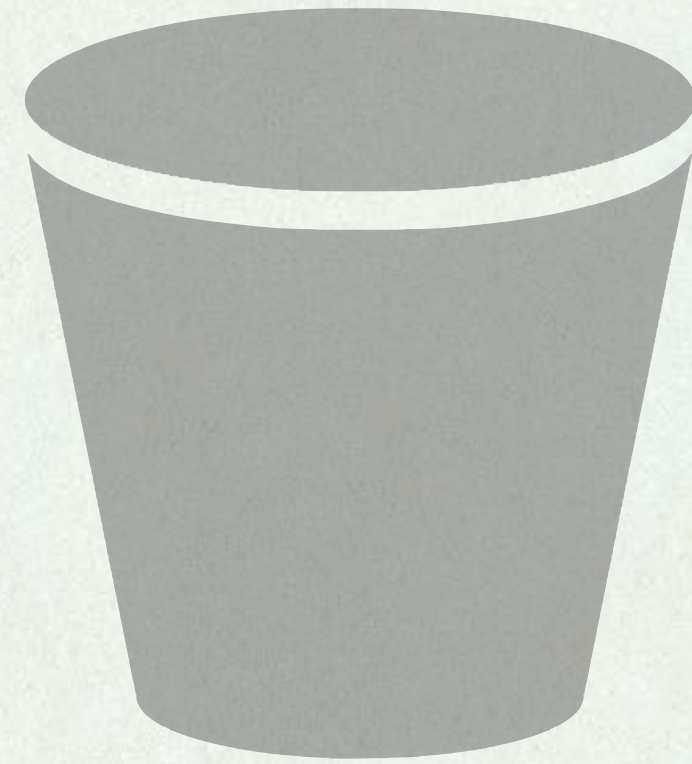
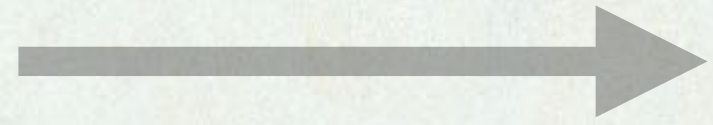
Unified catalogue search and API

New Wellcome Collection site

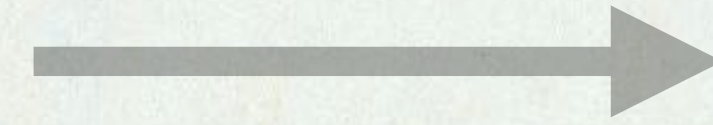




Production



Storage



Retrieval

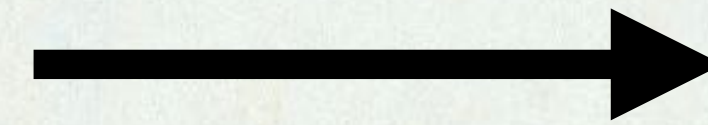




Production



Storage



Retrieval



# Read more:

Email me:

**[a.chan@wellcome.ac.uk](mailto:a.chan@wellcome.ac.uk)**

Browse our catalogue:

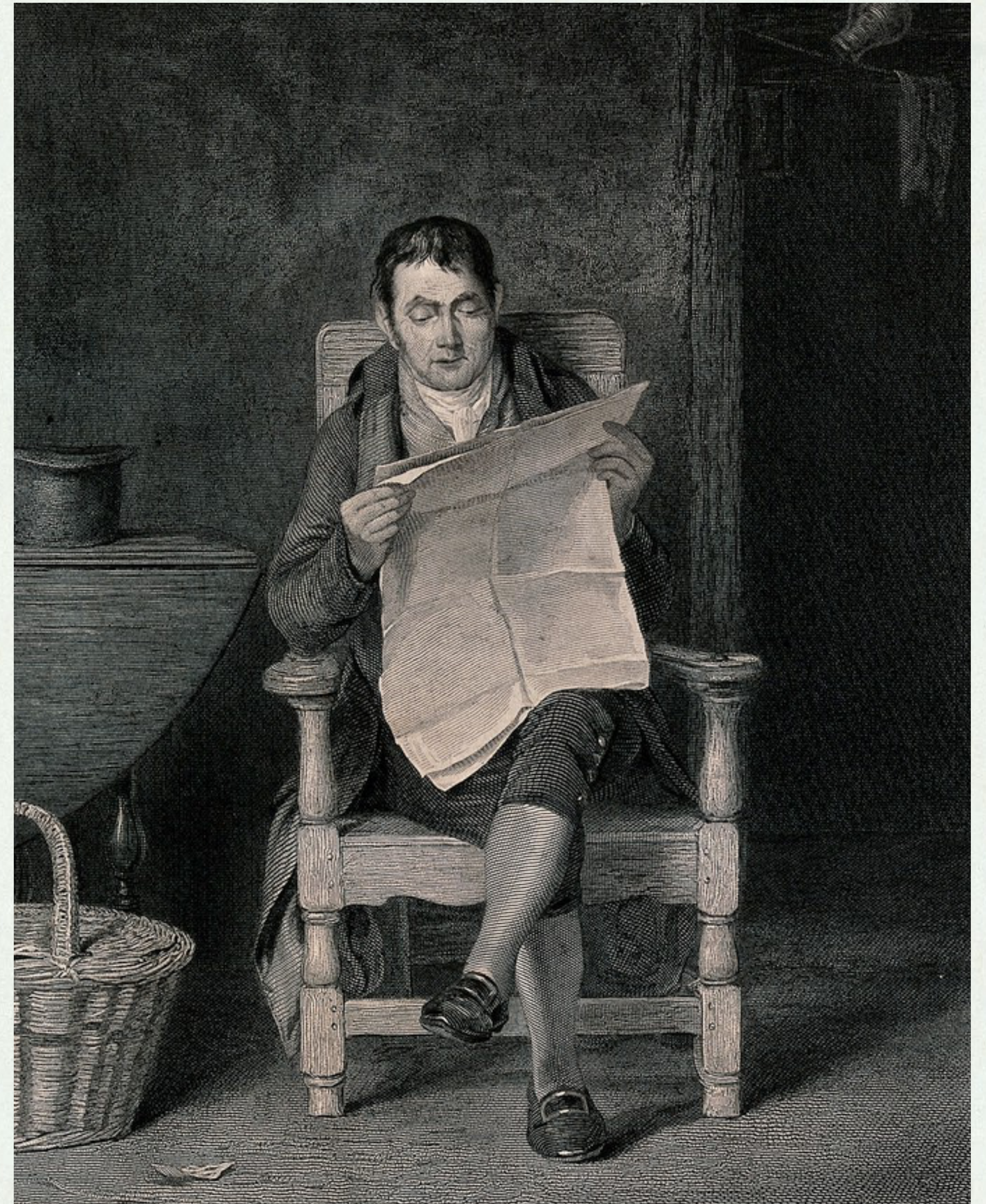
**[wellcomecollection.org/works](https://wellcomecollection.org/works)**

Get our code:

**[github.com/wellcometrust/platform](https://github.com/wellcometrust/platform)**

Read about our work (and get slides):

**[stacks.wellcomecollection.org](https://stacks.wellcomecollection.org)**





# Digital preservation at Wellcome

**Alex Chan ~ a.chan@wellcome.ac.uk ~ they/them  
Senior software developer at Wellcome Collection**

Image: The Anatomical Theatre at Cambridge.  
Published for R. Ackermann's *History of Cambridge*  
Wellcome Collection, CC BY.





# Image credits

Slides 1, 2, 31:

The Anatomical Theatre at Cambridge. Published for R. Ackermann's History of Cambridge.

Credit: Wellcome Collection, CC BY. <https://wellcomecollection.org/works/h76dp2d4>

Slide 3:

Building exterior of Wellcome Collection, 183 Euston Road.

Internal ref C0098841. Credit: Ben Gilbert, Wellcome Collection.

Slide 4, left (shelves):

Library spaces in Wellcome Library. Internal ref C0098841. Credit: Ben Gilbert, Wellcome Collection.

Slide 4, centre (staircase):

The Reading Room in Wellcome Collection. Credit: Ben Gilbert, Wellcome Collection, CC BY-NC.

Slide 4, right (sculpture):

*Being Human* gallery in Wellcome Collection. Credit: Steven Poock, Wellcome Collection, CC BY-NC.

Slides 5, 6, 16, 17, 21, 22, 28, 29 (Production):

Legal archivist icon, by Miroslav Kurdov. From The Noun Project, CC BY.

<https://thenounproject.com/term/legal-archivist/1823613/>

Slides 5, 6, 16, 17, 19, 20, 21, 22, 28, 29 (Bucket):

S3 bucket from AWS Simple Icons. Public domain.

[https://commons.wikimedia.org/wiki/File:AWS\\_Simple\\_Icons\\_Storage\\_Amazon\\_S3.svg](https://commons.wikimedia.org/wiki/File:AWS_Simple_Icons_Storage_Amazon_S3.svg)

Slides 5, 6, 16, 17, 21, 22, 27, 28, 29 (Retrieval):

Legal counsel icon, by Miroslav Kurdov. From The Noun Project, CC BY.

<https://thenounproject.com/term/legal-counsel/1824436/>

Slide 10:

Steven Pocock photographing artworks for facsimile digital copies.

Internal ref EP 000099. Credit: Thomas Farnetti, Wellcome Collection.

Slide 11:

Screenshot of Goobi, from Intranada GmBH.

<https://www.intranada.com/en/digiverso/goobi/goobi-overview/>

Slide 13:

Screenshot of Wellcome's instance of Archivematica. Credit: Wellcome Collection, CC BY.

Slide 18:

Preservica logo.

Slide 19 (bag):

Paper bag icon, by Dorian Dance. From The Noun Project, CC BY.

<https://thenounproject.com/term/paper-bag/28579/>

Slide 19 (gears):

Process icon, by Alice Design. From The Noun Project, CC BY.

<https://thenounproject.com/term/process/2473979/>

Slide 23:

Balloon icon from the DLCS website. <https://digirati.com/media/1107/dlcslogo.png>

Slide 24:

*An illustration of two Lorises. The Slender Loris, in waking and sleeping posture. Taken from Royal Natural History volume 1, page 230. Public domain.*

<https://commons.wikimedia.org/wiki/File:SlenderLorisLyd2.png>

Slide 26:

Screenshot of "Notice sur P. Curie", Wellcome Collection site. <https://wellcomecollection.org/works/ptnfbubj>

Slide 30:

A seated man reading a newspaper in a kitchen (?), a basket by his feet. Engraving by C.W. Sharpe after T.S.

Goode. Credit: Wellcome Collection, CC BY. <https://wellcomecollection.org/works/xsx2wvfb>