A Digital Forensics Approach to Collecting Born Digital Archives at the British Library

Callum McKean, Digital Lead Curator, Contemporary Archives and Manuscripts
Defining Digital Forensics

Digital forensics is a branch of forensic science that focuses on identifying, acquiring, processing, analysing, and reporting on data stored electronically.

The main goal of digital forensics is to extract data from the electronic evidence, process it into actionable intelligence and present the findings for prosecution. All processes utilize sound forensic techniques to ensure the findings are admissible in court.

Interpol, 2024.
What do we gain?

• It allows us to store data independently of vulnerable carriers

• It allows us to make the most authentic copies possible

• It ensures that accurate metadata is retained

• It allows us to document and recreate our actions

• It does not foreclose future research possibilities.
How do we implement this approach?
Appraisal

• We conduct a rapid appraisal of hard-drives, USB sticks and computers using a **SPEKTOR Ultra** digital forensics machine.

• We inspect other forms of physical media with the Depositor, looking for labels which may give a clue to their content

• We ask Depositors to fill out a questionnaire in order to make a record of their hardware and software use
SPEKTOR Ultra Digital Forensics machine in Collector Mode for rapid appraisal

Amstrad Floppy Disks from the Will Self Archive (Add MS 89203) with handwritten labels
Capture and Preservation

- After acquisition, flash storage and computers are forensically captured as .e01 files using a **SPEKTOR Ultra** digital forensics machine.

- Magnetic carriers (such as Floppy Disks) are captured at bit-level using **Kryoflux**.

- For some carriers, we cannot make forensic copies (zip disks, Amstrad disks etc.) For these we copy the files.

- These preservation copies are ingested into the Library’s **Minimum Preservation Tool** (for backup and regular fixity checking).

- Carriers are photographed, labelled and moved to long-term storage.
Kryoflux creating bit-level captures of 3.5inch floppy disks from the Wendy Cope Archive (Add MS 89108)

Intermedia and Windows XP machine used for non-forensic data extraction from other magnetic media
Why?

• We can not solve every problem immediately

• We do not know what research will look like in the future

• We do not know what the technological landscape will look like in the future

• We do not know what archival practice will look like in the future
Access – Genetic Criticism
Access – Data Analytics

Redacted network visualisation of the Harold Pinter e-mail archive created in Gephi as part of the project ‘Data Analytics and Network Visualisation for Hybrid Correspondence Archives’. Latest re-usable code available via GitHub.

Graph showing the time distribution of drafts of Andrea Levy’s novel *Small Island* (2004) created as part of the *Writers Digital Lives* PhD Placement project by Ariel Li.
Access - Emulation

Emulating the Windows 95 computing environment used by John Maynard Smith (Add MS 86569)
Thank you