





From Life on the Shelves to Digital Shelf-Life: Representing Journalism as an Historical Artefact in the Digital Domain

Dr Suzanne Paylor and Dr James Mussell ncse [www.ncse.kcl.ac.uk] s.paylor@bbk.ac.uk, j.mussell@bbk.ac.uk



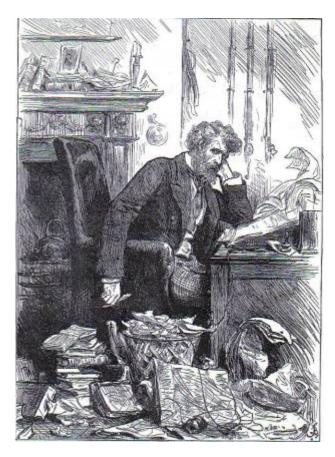


Project deliverables:

- Full-text digital edition of six nineteenth-century serials freely available via the web
- Users navigate facsimile images of pages
- Searching text strings on OCR text
- Searching by bibliographic, generic and thematic metadata
- Contextual and scholarly materials







"The Newspaper Editor," *Illustrated London News*, 11 September 1880.

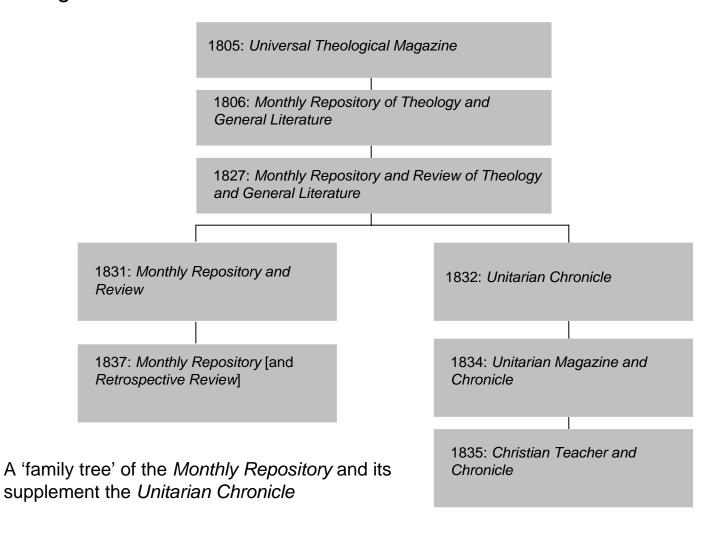
Research objectives:

- Compilation of most complete edition possible from public repositories and private collections
- Innovative approach to digital editing
 - New model of scholarly editing in the digital domain
 - Implementation of model through new approaches to digitisation and processing and development of hybrid software and bespoke interface
 - New methods of text processing via text mining and data analysis
 - New collaborative working model



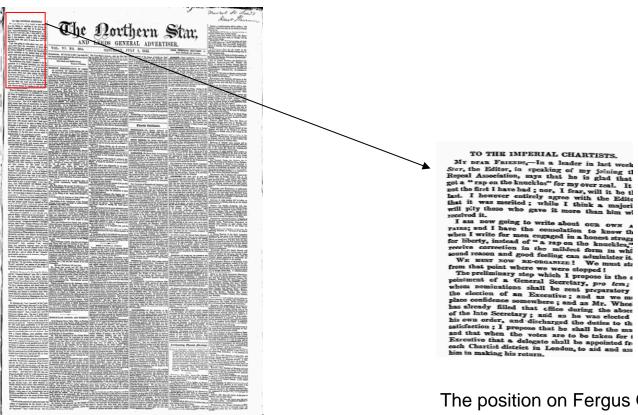


The Challenge of Scale





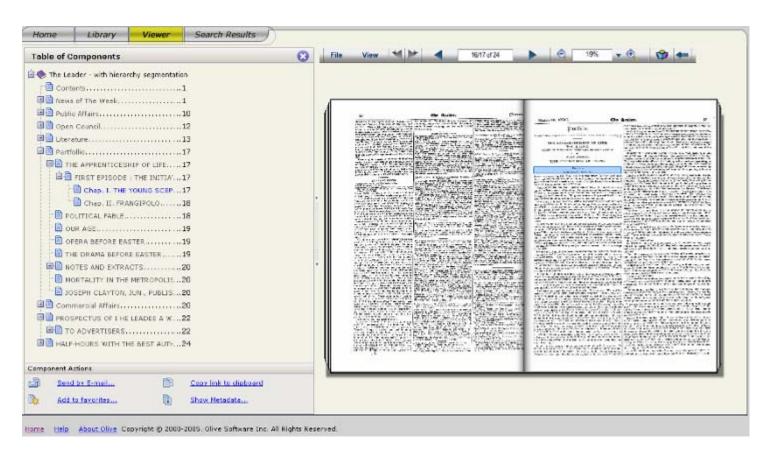
The Challenge of Depth: Layout



The position on Fergus O'Connor's weekly address to readers signals it as important content

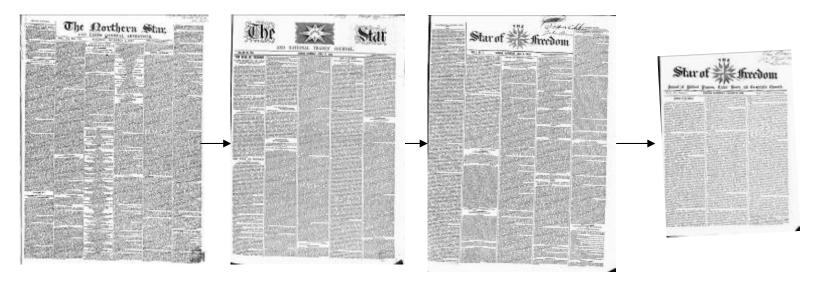


The Challenge of Depth: structure





Challenge of Depth: Mobility of form in the Northern Star



Changes of format and paper size in the Northern Star over its run.













Left: two images of an illustration in the *Tomahawk* taken from microfilm. Right: an image of the same page generated digitally directly from the hard copy.

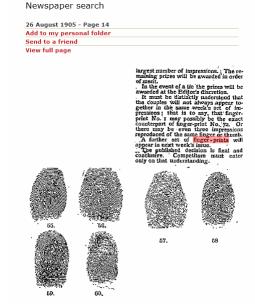








User's path to content of Penny Illustrated Newspaper via search interface





Edition architecture within Olive application:

edition > title > volume > number > department > item

e.g.

ncse > Leader > 7 > Leader, no. 319, 3 May 1856, edition 1 > 'Review of the Week' > 'Untitled Article'.

Although the hierarchy runs throughout the edition, the first four levels are displayed in a folder tree, and the final two on the page itself.





Edition > Title > Volume > Number > department > item

Edition

- Title

- Volume #1

Front Matter

- Number #1 (+ edition)

- wrapper

- number contents

dept #1

item #1item #2

- etc...

dept #2

dept #3

etc...

- Number #2

Etc.

- Index

End Matter

- Volume #2

Etc...