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

**ncse titles:**

*Monthly Repository* (1806-1838)
*Publishers’ Circular* (1880-1890)
*Tomahawk* (1867-1870)
*English Woman’s Journal* (1858-1864)
*Leader* (1850-1859)
*Northern Star* (1838-1852)
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

Distribution of 98,565 pages in ncse edition
The Challenge of Scale

A ‘family tree’ of the Monthly Repository and its supplement the Unitarian Chronicle
The position on Fergus O'Connor's weekly address to readers signals it as important content.
The Challenge of Depth: Typography

A picture of a tomahawk which almost always precedes the department called ‘The Week’.

This is the department header, signaled by the double line which precedes it.

This is the item within “The Week” It is separated from the department heading by a single line and from the next item with the same line type.
The Challenge of Depth: Temporality

Adverting pages from a heavily illustrated Christmas number of the Publisher's Circular

Northern Star, first and second edition, 5, 20 August 1842, p. 1. The first edition was published wanting the illustration due to the late arrival of the train carrying its plate.
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.
An archive model digitisation

In this archive display of search results isolate content from its context and some of the content of the segment (see highlight left) is lost. On selecting full page view users can only zoom in on segments of content rather than elements of form.

Source: British Library digitisation of the Penny Illustrated News:

http://www.collectbritain.co.uk/system/paper/
Mastheads from the six ncse titles and a British Museum Stamp for the back of an illustration plate in the *Tomahawk*
Representation: Edition structure

Edition
- Title
- Volume #1
  - Front Matter
  - Number #1 (edition town/country/#)
    - wrapper
    - number contents
      - department #1
        - item #1
        - item #2
        - etc...
      - department #2
      - department #3
        etc...
- Number #2
  Etc.
- Index
- End Matter
- Volume #2
  Etc...
Bibliographical Control and Editorial Choices: Representing material ideals

The volume title page for 1881 appearing in the advertising section of the January 16 1882 issue
Bibliographical Control and Editorial Choices: Representing material ideals

Pages from the Tomahawk Almanac 1869
Bibliographical Control and Editorial Choices: Representing material ideals

Portraits of prominent Chartists issued with the *Northern Star*
Bibliographical Control and Editorial Choices: Representing material ideals

Representation of department>item hierarchy in table of components
Bibliographical Control and Editorial Choices: Representing material ideals

The English Woman's Journal and Tomahawk which do not have a clear department-item structure
Bibliographical Control and Editorial Choices: Representing material ideals

Fold out illustrations in the *Tomahawk*
Representation: Metadata schema

**ncse Essential Metadata Fields**

**A. Number Level Metadata**
- Inputted Metadata:
  - A1. Periodical Title
  - A2. Date day
  - A3. Date month
  - A4. Date year
  - A5. Page span
  - A6. Source
  - A7. Series
  - A8. Volume
  - A9. Number
  - A10. Edition
  - A11. Bibliographical Location (number)

**B. Department Level Metadata**
- Inherited Metadata (from number):
  - A1. Periodical Title
  - A2. Date day
  - A3. Date month
  - A4. Date year
  - A5. Source
  - A7. Series
  - A8. Volume
  - A9. Number
  - A10. Edition
  - A11. Bibliographical Location (number)
- Inputted Metadata:
  - B1. Department Title
  - B2. Department Generic Metadata

**C. Item Level Metadata**
- Inherited Metadata (from department):
  - A1. Periodical Title
  - A2. Date day
  - A3. Date month
  - A4. Date year
  - A5. Source
  - A7. Series
  - A8. Volume
  - A9. Number
  - A10. Edition
  - A11. Bibliographical Location (number)
- Inputted Metadata:
  - B1. Department Title
  - B2. Department Generic Metadata
  - C1. Page number
  - C2. Page Span
  - C3. Bibliographic Location (item)
  - C4. Item Generic Metadata
  - C5. Image Description
Bibliographical Control and Editorial Choices: identifying thematic content

Police report from the Northern Star, May 2, 1846 concerning a violent disturbance in a pub

Commentary piece on the effects of drink of the working classes in the English Woman’s Journal, 1858, p. 306
Bibliographical Control and Editorial Choices: Representing thematic content

Detail from ncse’s working concept map
Editorial Challenges – representing form
Rich page content:

(Pictures)

Left to right:

a printed finger post from the *Monthly Repository* in 1814 and a hand-drawn finger post from a margin in the same volume; a finger post from an advert in the *Publishers’ Circular* in 1880 (“all orders”); and the heading to “Open Council,” the correspondence department from the *Leader*

Fold out illustrations in the *Tomahawk*

Advertising pages from the *Publisher's Circular*