Stephen Gallagher:

lessons learnt (so far)

from a hybrid literary collection

Simon Wilson, Senior Archivist
Overview

• Stephen Gallagher

• Description & arrangement of a hybrid collection

• Lessons learnt (so far)
Stephen Gallagher
Stephen Gallagher

• Novelist, screen-writer and director

• Novels - *Chimera, Oktober, Bedlam Detective*

• TV includes *Doctor Who, BUGS, Rosemary & Thyme, Eleventh Hour, Crusoe, Silent Witness*
Stephen Gallagher

2005 – deposited 42 boxes (7m) of paper archives including:
- notebooks
- outlines
- drafts
- treatments
- shooting scripts etc

Collection gives an excellent insight into the literary process
Literary collections @ Hull

Modern literary material already at Hull includes Philip Larkin, Stevie Smith, Alan Plater and John Godber
Born-digital archives

- 2010 – approached about depositing born-digital archives as part of the AIMS project

- Brought material on external hard drive for us to copy 14,320 files (13.6 GB) including drafts, correspondence etc also saved webpages – that were sources of ideas or inspiration

- Further b-d material in 2012 and 2013 now have over 25,000 files (39.2 GB)

Every outline, script and novel draft has flown back and forth without ever existing as hard copy until (in the case of the scripts) printed and handed to the actors
Audit

Collect metadata about each and every file

Use *DROID* to collect technical metadata for each file – file type, date created, file size, checksum (digital signature) etc.
### Audit

#### Cataloguing Creativity

15 Nov 2013

---

<table>
<thead>
<tr>
<th>Filename</th>
<th>Date created</th>
<th>Date last accessed</th>
<th>File extension</th>
<th>File Size</th>
<th>Checksum</th>
<th>Filepath</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bookplate.pub</td>
<td>18/02/2009 20:27</td>
<td>08/01/2009 17:55</td>
<td>pub</td>
<td>26,112</td>
<td>B6881A1D799335CDA4A989C9F4C878FE</td>
<td>D:\born digital\Stephen Gallagher\Bookplate.pub</td>
</tr>
<tr>
<td>Compslip.pub</td>
<td>04/06/2008 07:42</td>
<td>08/01/2009 17:56</td>
<td>pub</td>
<td>20,592</td>
<td>8A8A49791BF0AA0EA883978BA8B02040D48</td>
<td>D:\born digital\Stephen Gallagher\Compslip.pub</td>
</tr>
<tr>
<td>A Year In Oktober.doc</td>
<td>04/06/2008 07:42</td>
<td>10/10/1998 17:32</td>
<td>doc</td>
<td>56,296</td>
<td>AD6645728F635CE31C66431C5C0F0</td>
<td>D:\born digital\Stephen Gallagher\Articles\A Year In Oktober.doc</td>
</tr>
<tr>
<td>American notes 2007.doc</td>
<td>04/06/2008 07:42</td>
<td>13/07/2007 12:15</td>
<td>doc</td>
<td>72,704</td>
<td>7F5C633491159AD5C5D7EAA31575ECC</td>
<td>D:\born digital\Stephen Gallagher\Articles\American notes 2007.doc</td>
</tr>
<tr>
<td>ACQUIRED.doc</td>
<td>04/06/2008 07:42</td>
<td>16/08/2006 01:17</td>
<td>doc</td>
<td>33,948</td>
<td>6D82B2A4C4750B30C2D713561003353E</td>
<td>D:\born digital\Stephen Gallagher\Articles\ACQUIRED.doc</td>
</tr>
</tbody>
</table>

---

<table>
<thead>
<tr>
<th>Folder Name</th>
<th>Filepath</th>
<th>Number of Sub Folders</th>
<th>Number of Files</th>
<th>Folder Size (kb)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Runaways Project</td>
<td>D:\born digital\Stephen Gallagher\UBTV DEVELOPMENT 2009\Runaways Project\</td>
<td>0</td>
<td>18</td>
<td>5,093,923</td>
</tr>
<tr>
<td>Train Wreck Jane</td>
<td>D:\born digital\Stephen Gallagher\UBTV DEVELOPMENT 2009\Train Wreck Jane\</td>
<td>0</td>
<td>1</td>
<td>32,768</td>
</tr>
<tr>
<td>My Radio</td>
<td>D:\born digital\Stephen Gallagher\My Radio\</td>
<td>10</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Gallagher - A Resistance To Pressure 256 kbps</td>
<td>D:\born digital\Stephen Gallagher\My Radio\Gallagher - A Resistance To Pressure 256 kbps\</td>
<td>0</td>
<td>1</td>
<td>165,969,792</td>
</tr>
<tr>
<td>Gallagher - An Alternative To Suicide 96 kbps</td>
<td>D:\born digital\Stephen Gallagher\My Radio\Gallagher - An Alternative To Suicide 96 kbps\</td>
<td>0</td>
<td>1</td>
<td>61,613,349</td>
</tr>
</tbody>
</table>
Description & arrangement of a hybrid collection
Description

Create a single integrated catalogue
  - same creator – only the format is different

Series level (eg a folder of files relating to a specific piece of work) – appears to offer a suitable return for effort

Use GLAM cataloguing guidelines, thesaurus and glossary to establish consistency in terminology used
Can you describe a file without looking at it?
- some files created in *FinalDraft* software
- managed to get access via our Drama Dept – save as PDF

Born-digital material – greater effort to investigate a file
(similar to material in another language...)

Risk management - if you don’t check every file
for sensitive information
Arrangement

Talking to depositor - each novel/script is seen as a distinct project

Structure for the collection – not based on physical format & must allow for accruals

Based on physical output (radio, novel etc) each ‘project’ becomes a series

Reflects his way of working & the archival principles of control – creator, archivist and user can all understand
Access

Currently not providing access to born-digital material

Discussing access with Stephen Gallagher
- especially regarding un-published material

Do expect to make some material available online
- via the University’s digital repository (Fedora)

Hope to pilot a sample of b-d content in 2014
- show that content can be managed via user permissions – key feature of the repository
What we have achieved

Stephen Gallagher collection
- a draft boxlist for the paper material
- notes (in Excel) for some of the b-d folders
- draft structure in CALM
- identified a solution for handling FinalDraft files

Born-digital archives
- identified ‘series’ as our default level of cataloguing b-d
- workflow, procedures and documentation

Looking to use checksum (digital signature) to identify duplicate material in subsequent accruals of b-d material
Processing plan

Processing Plan
AccNo: 2010/15  Ref: U DGA

OVERVIEW

Collection Title: Stephen Gallagher
Creator / Depositor: Stephen Gallagher

Related Material at UHA:
- Paper archives: already deposited
  - 2000/10 (42 boxes) — mainly paper with a few boxes of publications, copies of DVDs etc
  - 2005/14 (13 boxes) — further publications (foreign editions etc) and production material
Not tackled — blog / website (possibly recommend the British Library Web Archive) and email

Brief Description of the material:
Material relates to his writing (short-stories, novels, radio and screen) including research process, drafts etc. Also material relating to his blog / website with some publicity/promotional material. There are only isolated email messages (no mailboxes).

Extent: 13.6 GB  No of files: 14,320

Comments re extent:
There are also 39 3" Amstrad discs

ARCHIVAL DESCRIPTION

Proposed level of archival description to be applied:
- Primarily at series level

Justification:
Stephen Gallagher considers each piece of work as a discrete project. Interest in the material is likely to be on two accounts:
- Writing process following a particular story from idea through research, drafts, pitching and completion (whether publication of novel or filming of screenplay etc)
- A particular place of work

This means that if describing the project we do not necessarily need to describe particular content

Cataloguing Priority for this accession:

<table>
<thead>
<tr>
<th>Priority Score:</th>
<th>18 / 24</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Research potential</td>
<td>3</td>
</tr>
<tr>
<td>2. MHC specialist area</td>
<td>3</td>
</tr>
<tr>
<td>3. Toxicity/time crucial</td>
<td>1</td>
</tr>
<tr>
<td>4. High teaching potential</td>
<td>2</td>
</tr>
<tr>
<td>5. Education potential</td>
<td>2</td>
</tr>
<tr>
<td>6. Community/outreach potential</td>
<td>1</td>
</tr>
<tr>
<td>7. Summary list is sufficient</td>
<td>3</td>
</tr>
<tr>
<td>8. Complexity of cataloguing</td>
<td>3</td>
</tr>
</tbody>
</table>

APPRAISAL

Is appraisal necessary? Yes  No  N/A
Potential for appraisal?
Initial investigations identified very little material that could or should be appraised

ARRANGEMENT

Integrate with existing arrangement? Yes  No  N/A
Does the current arrangement include bad material? Yes  No  N/A

Justification:
There is considerable overlap between paper and born-digital material

Potential arrangement issues:
- Paper files being catalogued at file level — need to consider implications for discovery & access
- To not try to describe each born-digital item. Include an overview of born-digital material within the series description

Any restricted / sensitive content:
- Some personal material (eg references for third parties) that should be closed
- Suggest that most recent work (last 2 years) should be closed (discuss this with SG)
- Research folder (1235 files in 87 folders, 24.1MB) material is mostly saved web pages — need to consider arrangement/access issues
- MyWeb folder (44 files, 1.85GB) recorded broadcasts can be included in the archives but are subject to copyright so should not be made available online via repository

PRESERVATION

Media Issues:
- Main body of material was selected by SG and transferred via external hard drive
- There are 39 3" Amstrad discs that cannot be read with current hardware

Content Issues:
- 291 files in .docx format (*.dot) sent Mary-Jane Dickinson (Drama) to use their copy of Fontcross. Other material (other files) sent to Chris MacDonald. 291 files (June/July 2011) and created PDF copies for access
- How to present the old website content to users as web page (via a web browser etc) rather as individual unlinked pages

Proposed preservation actions:
- Import the Fontcross PDFs and attach to the original *.dot file

Plan produced by: Simon Wilson  Date: 13th Sept 2011

Suggested Review Date:
Todo list

Liaising with ICT colleagues on a technical solution to ingest 10,000+ files into the University's digital repository

Looking to test ArchiveSphere (Penn State) as a possible tool to access and manage born-digital content in Fedora

Discuss with colleagues in History and English questions and concerns regarding access to born-digital material

Demonstrate to depositors we can manage (ie allow or restrict) access to content via permissions
Lessons learnt (so far)
Lessons learnt

Technical up-skilling associated with b-d

Software – DROID, FTK Imager, Fedora

Hardware – write blockers, forensic workstation

Process & workflow surprisingly very similar

Processing plan – identify issues for appraisal, arrangement and preservation as part of the initial audit and assessment
Conclusion

Description and appraisal remain key irrespective of format **BUT** the sheer volume of born-digital material is scary

Every collection is different

We (generally) don’t yet know what our user expects/wants from us regarding born-digital material

Same person (if possible) to work on paper & born-digital material, move from a specialist skill to ‘business as usual’
Contact details

Simon Wilson
Tel 01482 317506
Email s.wilson@hull.ac.uk
www.hullhistorycentre.org.uk

HHC blog - http://hullhistorycentre.blogspot.com/

AIMS White Paper –
http://www.digitalcurationsservices.org/aims/whitepaper/