

Stephen Gallagher: lessons learnt (so far) from a hybrid literary collection

Simon Wilson, Senior Archivist

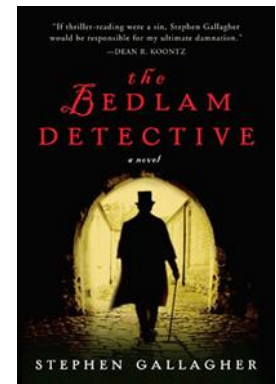
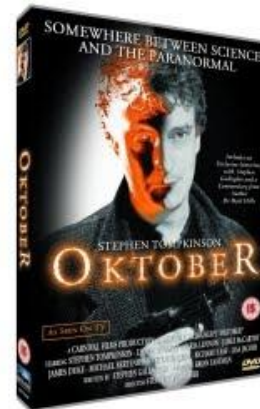
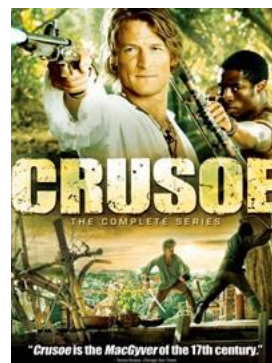
Overview

- 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*

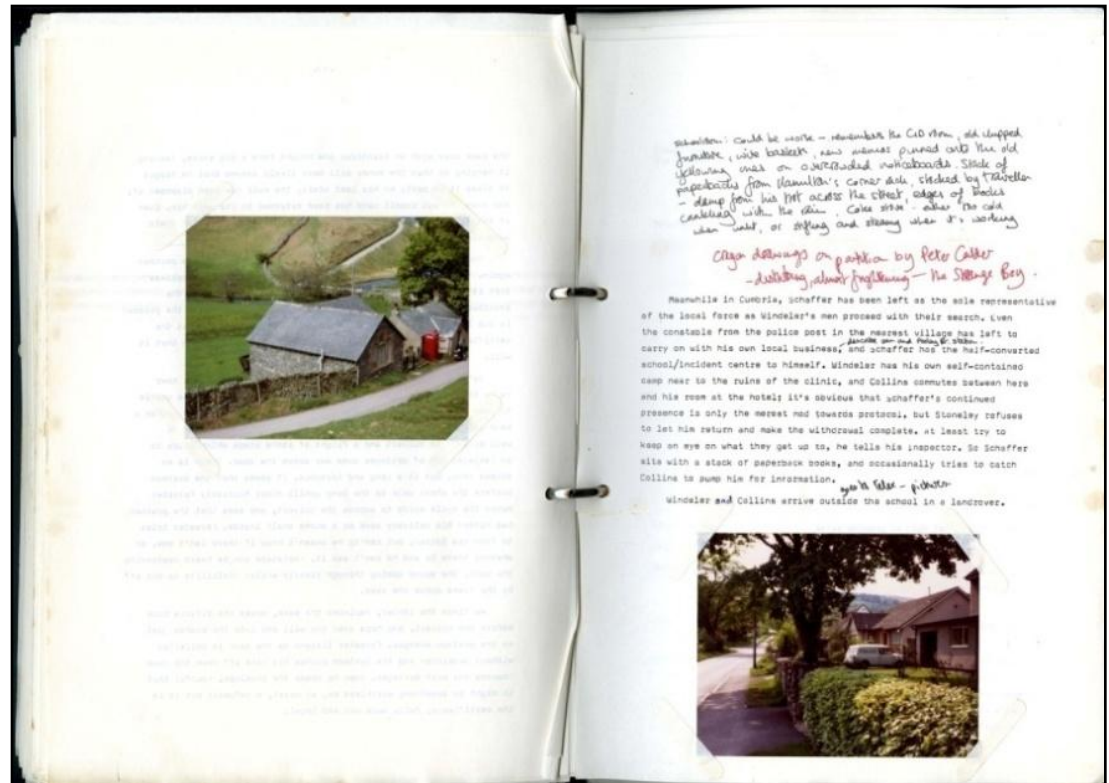


Stephen Gallagher

2005 – deposited 42 boxes (7m) of paper archives including:

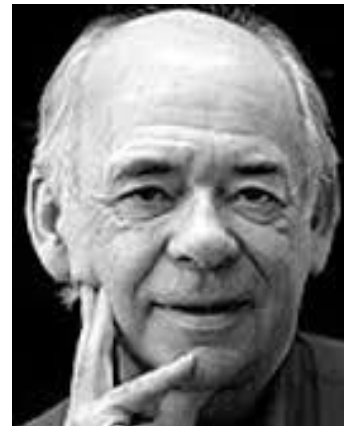
- notebooks
- outlines
- drafts
- treatments
- shooting scripts etc

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 (39.2GB)

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

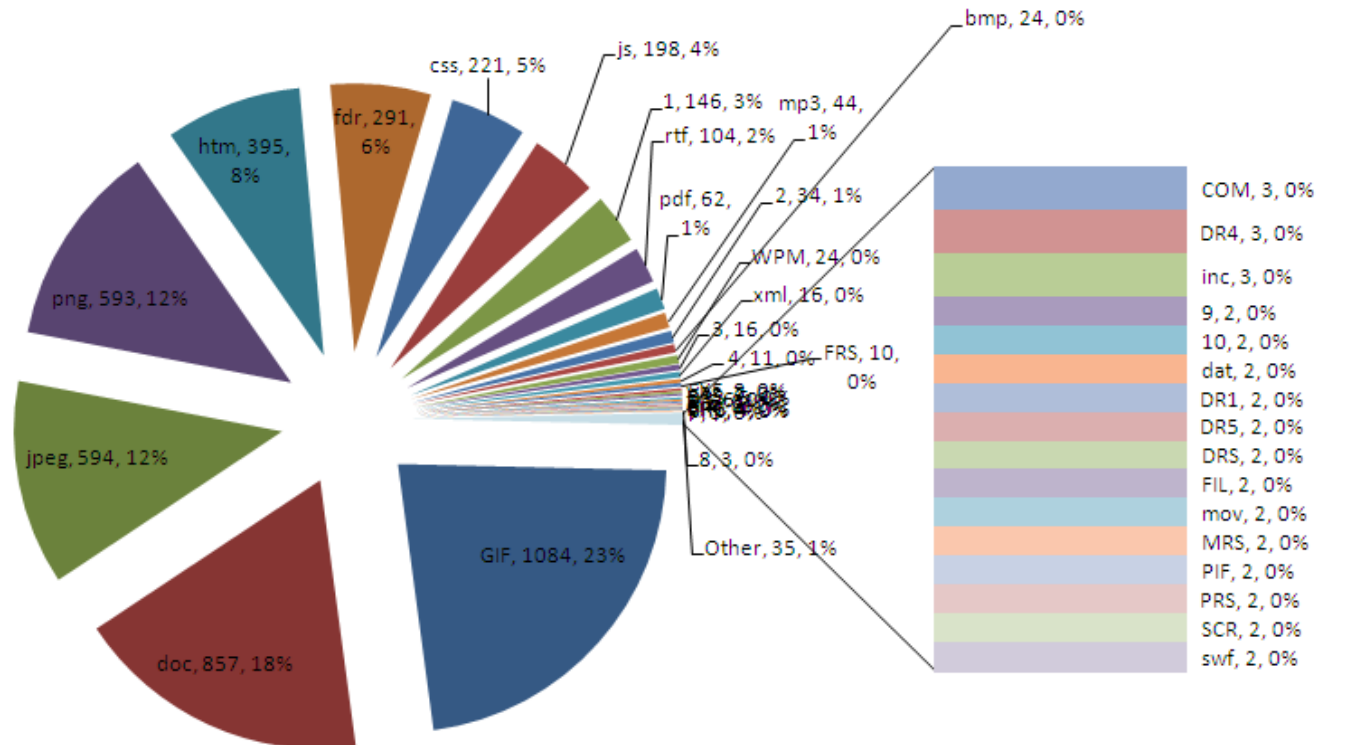
2010-15 DirPrnInfo and catalogue notes.xls [Compatibility Mode] - Microsoft Excel

Filename	Date created	Date last accessed	File extension	File Size	Checksum	Filepath
Bookplate.pub	18/02/2009 20:27	08/01/2009 17:55	pub	26,112	B6B81A1D799335CDAA4989C9F4CB78FE	D:\born digital\Stephen Gallagher\Bookplate.pub
CompSlip.pub	24/06/2008 23:42	08/01/2009 17:56	pub	20,992	868A497B18FAAA988597B48300240D48	D:\born digital\Stephen Gallagher\CompSlip.pub
Letterhead template.doc	24/06/2008 23:42	06/09/2007 19:14	doc	21,504	40265233F00648F21F0386D99A4774BF	D:\born digital\Stephen Gallagher\Letterhead template.doc
A Year in Oktober.doc	24/06/2008 23:42	16/10/1998 17:32	doc	55,296	AD0E4497263F85CE3C1569CA35C61FF0	D:\born digital\Stephen Gallagher\Articles\A Year in Oktober.doc
Abner Stein Books CV April 2006.doc	24/06/2008 23:42	17/04/2006 15:14	doc	32,768	B95AC91D8884EB21145E9C53BA6032AA	D:\born digital\Stephen Gallagher\Articles\Abner Stein Books CV April 2006.doc
Agency CV November 2005.doc	24/06/2008 23:42	22/11/2005 01:28	doc	40,960	CB63602C9473330A4C39B88585411015A	D:\born digital\Stephen Gallagher\Articles\Agency CV November 2005.doc
American notes 2007.doc	24/06/2008 23:42	13/07/2007 12:15	doc	72,704	F5C662A91F11EA9FC65DF8AA31575EEC	D:\born digital\Stephen Gallagher\Articles\American notes 2007.doc
ARCHIVET.doc	24/06/2008 23:42	16/09/2001 21:17	doc	34,304	849B48A65E4756C65E9E732B123DCE3E	D:\born digital\Stephen Gallagher\Articles\ARCHIVET.doc

	A	B	C	D	E
1	Folder Name	Filepath	Number of Sub Folders	Number of Files	Folder Size (kb)
11	Runaways Project	D:\born digital\Stephen Gallagher\JBTV DEVELOPMENT 2009\Runaways Project\	0	18	5,093,923
12	Train Wreck Jane	D:\born digital\Stephen Gallagher\JBTV DEVELOPMENT 2009\Train Wreck Jane\	0	1	32,768
13	My Radio	D:\born digital\Stephen Gallagher\My Radio\	10	0	0
14	Gallagher - A Resistance To Pressure 256 kbps	D:\born digital\Stephen Gallagher\My Radio\Gallagher - A Resistance To Pressure 256 kbps\	0	1	165,969,792
15	Gallagher - An Alternative To Suicide 96 kbps	D:\born digital\Stephen Gallagher\My Radio\Gallagher - An Alternative To Suicide 96 kbps\	0	1	61,613,349

File types

Frequency of file types



- GIF
- doc
- jpeg
- png
- htm
- fdr
- css
- js
- 1
- rtf
- pdf
- mp3
- 2
- bmp
- WPM
- 3
- xml
- 4
- FRS
- EXE
- VRS
- DR3
- txt
- 5
- eml
- 6
- DR2
- php
- 7
- 8
- COM
- DR4
- inc
- 9
- 10
- dat
- DR1
- DR5
- DRS
- FIL
- mov
- MRS
- PIF
- PRS
- SCR
- swf

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

Description

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

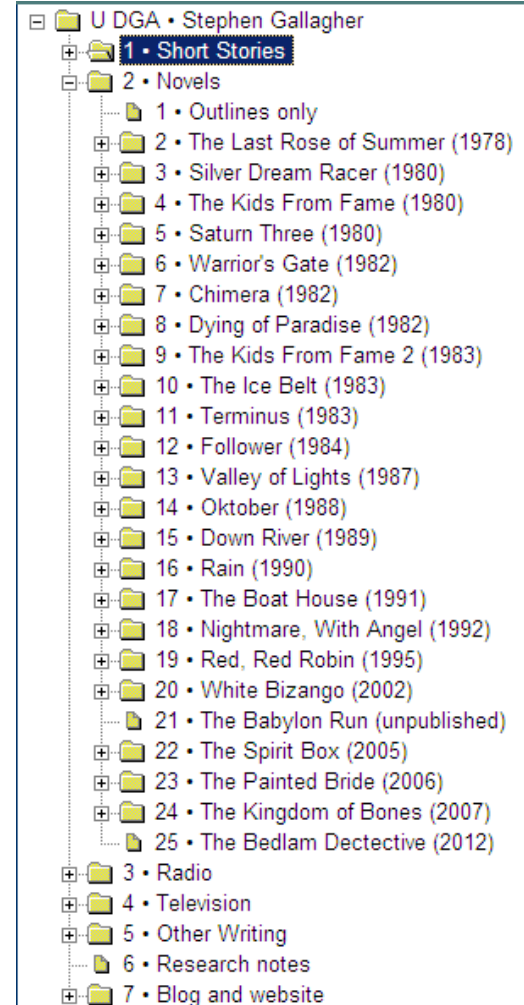
Arrangement

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

Acc No: 2010/15 Ref: U DGA

OVERVIEW			
Collection Title:		Stephen Gallagher	
Creator / Depositor:		Stephen Gallagher	
Related Material at HUA:			
Paper archives already deposited			
<ul style="list-style-type: none"> - 2008/10 (42 boxes) – mainly paper with a few boxes of publications, copies of DVDs etc - 2010/14 (12 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:			
<ul style="list-style-type: none"> • 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:			
<ul style="list-style-type: none"> - 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 piece of work 			
This means that if describe the project we do not necessarily need to describe particular content			
Cataloguing Priority for this accession:			Priority Score:
1. Research potential	3	5. Education potential	2
2. HHC specialist area	3	6. Community/outreach potential	1
3. Topicality / time crucial	1	7. Summary list is sufficient	3
4. UoH teaching potential	2	8. Complexity of cataloguing	3
			18 / 24
Scoring: 3 = high, 2 = medium, 1 = low 0 = no potential			

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 b-d material?	Yes- No -N/A
Justification:	
There is considerable overlap between paper and born-digital material	
Potential arrangement issues?	
<ul style="list-style-type: none"> • Paper files being catalogued at file level – need to consider implications for discovery & access • To not try to describe each born-digital item but include an overview of born-digital material within the series description 	
Any restricted / sensitive content?	
<ul style="list-style-type: none"> • Some personal material (eg references for 3rd parties) that should be closed • Suggest that most recent work (i.e. last x years) should be closed [discuss this with SG] • <u>ResearchDocs</u> folder (1226 files in 87 folders, 14.5MB) material is mostly saved web-pages – need to consider arrangement /access issues • <u>MyRadio</u> 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:	
<ul style="list-style-type: none"> • 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:	
<ul style="list-style-type: none"> • 291 files in <u>FinalDraft</u> format (*.fdr) contact Mary-Jane Dickenson (Drama) to use their copy of <u>FinalDraft</u> – looked at files (June/July 2011) and created PDF copies for access • How to present the old website content to users as web pages (via a web browser etc) rather as individual unlinked pages 	
Proposed preservation actions:	
Import the <u>FinalDraft</u> PDFs and attach to the original *.fdr file	
Plan produced by:	Simon Wilson
Date:	13 th 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

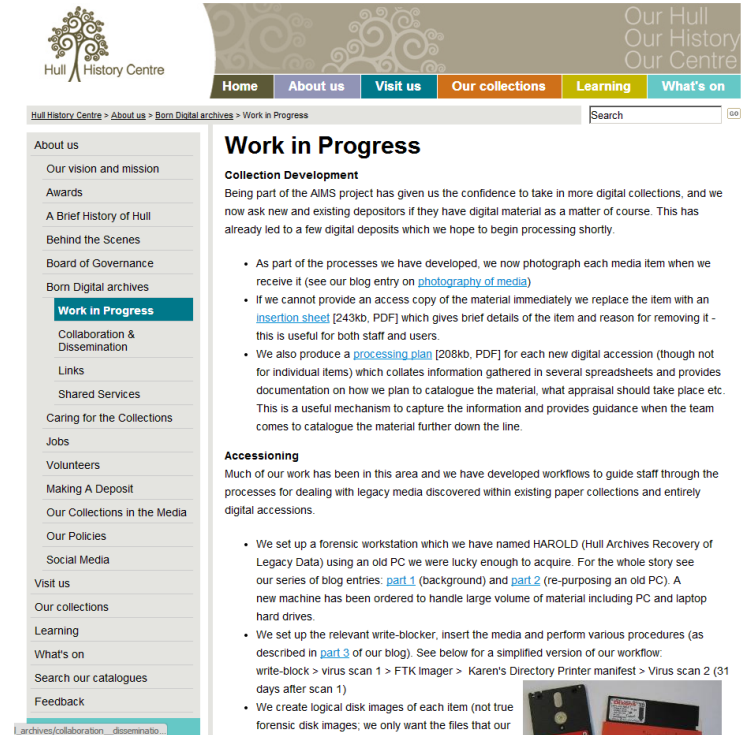
Email s.wilson@hull.ac.uk

www.hullhistorycentre.org.uk

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

AIMS White Paper –

<http://www.digitalcurationervices.org/aims/whitepaper/>



The screenshot shows the Hull History Centre website. The header includes the Hull History Centre logo and navigation links: Home, About us, Visit us, Our collections, Learning, and What's on. The main content area is titled 'Work in Progress' and is divided into two sections: 'Collection Development' and 'Accessioning'. The 'Collection Development' section discusses the AMS project and lists several bullet points regarding digital collections, including photography of media, insertion sheets, and processing plans. The 'Accessioning' section describes workflows for dealing with legacy media and lists bullet points about forensic workstations (HAROLD) and write-blockers. A small image of a forensic workstation is visible at the bottom right of the screenshot.