

## Data Audit Framework practical exercise

Audit scenario.....	3
Overview of audit activities.....	4
Description of TFTS.....	5
Confidentiality agreement.....	6
Audit scoping interview.....	7
TFTS questionnaire results.....	8
Data interview: Sam Lever.....	12
Data interview: Christie Rohan.....	14

**Please note this exercise is entirely fictitious**

## Question sheet

1. How would you let staff in TFTS know about the audit? What would you tell them to encourage participation?

2. What will Glasgow Uni and TFTS gain from the audit - what is the business case?

3. What kind of classification would be appropriate given the impetus behind the audit and expected future use of the audit results?

4. How many data assets have you recorded for TFTS? Which of these would you like to find out more about and why?

5. What data management issues have you identified and what recommendations would you make to minimise these in the future?

## **Audit scenario**

Imagine you are a records manager working in the Data Protection office at the University of Glasgow. The University has recently secured a grant from the JISC to enhance its Institutional Repository to accommodate research data as well as publications. In future it will accept a variety of data types including numerical or textual records, image collections and audiovisual material.

The University has matched JISC's grant to extend the project over four years. This allows a rolling programme of development and data collection. The first year will focus on expanding capacity and skills to facilitate data deposit. A trial deposit period is planned at the start of year two to allow the new systems to be tested before wider collection begins. The trial will be limited to materials created with University funding. Subject-focused collection drives will then be run across all Faculties with particular emphasis on research data that can not be deposited elsewhere, for example in a discipline-based data centre or publisher's repository.

The University wants to find out what quantity and type of research data is currently held within departments to scope development work and facilitate later collection drives. Several data audits are proposed to collect this information. As you have experience of conducting information audits you have been asked to lead this programme. Six audits will be completed in a variety of contexts across the University. For the purposes of this exercise we will focus on an audit in the Department of Theatre Film and Television Studies (TFTS) which is underway.

Using the materials provided in this pack, complete an audit record for TFTS in the DAF online tool and answer the questions provided.

## Overview of audit activities

Date(s)	Activity	Heading
3/11/08 to 7/11/08	Initial research conducted into departmental holdings and approach to data curation	TFTS overview Confidentiality agreement Scoping interview
10/11/08 to 21/11/08	Questionnaire circulated to staff	Questionnaire results
24/11/08	Interview with Sam Lever, librarian	Data interview: Sam Lever
27/11/08	Interview with Christie Rohan, researcher	Data interview: Christie Rohan

## Description of TFTS

The following details were taken from departmental website

Theatre Film and Television Studies (TFTS) at the University of Glasgow is one of the most well-established departments for these subjects in the UK. We have a strong team of over twenty teaching and research staff who specialise in diverse areas of performance and media arts. Several taught courses are offered. Through these students build up critical knowledge, creative approaches and confidence in their own skills. An excellent subject library situated on-site provides access to web resources, course literature, a large video collection and theatre production materials. The Department places a strong emphasis on research and benefits greatly from financial support offered by previous alumni, most notably Dame Margaret McArthur. Several widely renowned research projects stem from McArthur Fellowships as well as from other research council funding secured by TFTS. These include *Changing Approaches to Theatrical Tuition*, *Glasgow Cinemas* and *Shakespearean performance database*.

TFTS is a thriving hub of performance and media arts research and teaching. If you would like to become part of our community, either as a member of staff or prospective student please get in touch with the Secretary, Jeanette Barrie, on the details below.

Theatre Film and Television Studies Department,  
University of Glasgow,  
The Gilmorehill Centre,  
University Avenue,  
Glasgow G12 8QQ

tel: +44 (0)141-330 2678

fax: +44 (0)141-330 1084

email: [info@tfts.arts.gla.ac.uk](mailto:info@tfts.arts.gla.ac.uk)

## Confidentiality agreement

### Data Audit Framework Statement of Confidentiality

This document represents a confirmation of the University of Glasgow's commitment to refrain from any public disclosure of documentation, interview material or any related conclusions collected, considered or obtained as part of the data audit in the Department of Theatre Film and Television taking place in Nov-Dec 2008. This statement of confidentiality applies to all content collected during and prior to the on-site audit activities, and to any conclusions formed on the basis of that content.

All staff from the University of Glasgow Institutional Repository and Data Protection office are party to this agreement, and bound by the commitment to nondisclosure. In the event of additional University of Glasgow staff or associates participating in the audit activities they too will be bound by the terms of this document.

Information to remain confidential includes, but is not limited to:

- content of data assets;
- know-how;
- system documentation ;
- organisational or staffing information;
- financial information;
- verbal representations;
- business strategies;
- results of other audits or organisational appraisals.

The agreement to non-disclosure continues in perpetuity.

The results of the audit will be used internally to conceive a confidential audit report to be presented to the Department of Theatre Film and Television Studies and staff involved in the Institutional Repository development project. TFTS may subsequently use this document in any way it wishes.

S. Dunbar  
Signed

1<sup>st</sup> November 2008  
Date

SIMON DUNBAR (Director, DP and FoI Office)  
Name in block capitals

## **Audit scoping interview**

### **Notes from meeting held with James Ferguson, Head of TFTS, on 5<sup>th</sup> November 08**

The meeting began with a discussion on what the University hoped to achieve. The concern that key research materials were left within departments and not properly curated was noted. Plans to enhance the IR were outlined and an explanation given as to how the audit in TFTS would help suggest what kinds of research materials were likely to be deposited in the future. The focus on research outputs from University-funded projects or those without a central place of deposit was noted.

#### The research context in TFTS

The Department has a rich and varied research programme. TFTS is very close knit and monthly staff meetings are held to provide updates on progress and discuss avenues for research. These meetings help to decide what proposals to put forward for University funds earmarked for performance research, such as the McArthur Fellowships. Staff often work collaboratively, and also with students. Indeed a large number of doctoral students come direct from the undergraduate courses. Additional research funding has been secured from AHRC, ESRC & the Arts Council.

#### Existing data management practices

Methods of working vary between staff members. Some choose to store materials on the shared drive (normally when collaborating) but others keep it locally as they feel this gives better control and security. No policy on creating and preserving research data is in place at present. Best practice guides from the Performance Data Archive (PDA) are used on research council projects and some staff have drawn these processes into their everyday work. At present though, approaches are ad hoc.

The librarian, Sam Lever, unofficially keeps tabs on research through staff meetings. He tracks outputs to see if they could become library resources. When research data is deposited with archives he makes sure a copy is accessible from here – normally as a web resource. On research council projects part of the original project budget is set aside for archiving. Internal arrangements are not as formal - part of the library budget is set aside for acquiring and maintaining new resources so research data is kept accessible in this way. The Department intends to devise a data policy and formalise the process of tracking and maintaining research.

#### Audit practicalities

James agreed for TFTS staff to contribute to the audit and suggested that Sam Lever would be the best point of contact. As it is term time, a questionnaire would make it easier for staff to take part. This could be followed up by interviewing a few key staff such as Sam, the librarian, and Christie, Head of Theatre. Christie has run several research projects and supervises many of the doctoral students.

TFTS would like to use the audit results to develop a data policy. It was agreed the core element set would be sufficient for this and University needs. Any information on data quantities or projected future growth will be added where possible to help the Institutional Repository with capacity planning.

## TFTS questionnaire results

11 questionnaire responses were received. Responses to precursor questions are summarised and the table lists research assets staff noted.

**1. Are you involved in creating or using research data? – please explain**

7 respondents created and used research data - 4 just used resources. Comments noted around 60% of research is University funded with other projects supported by research councils or arts organisations. Most comments focused on locally created resources that were used but 3 people spoke about community-led reference collections and holdings in data archives that are crucial to their research.

**2. What types and quantity of research data do you create or use?**

Common data types were databases, digital image collections and audiovisual materials. In terms of quantity most responses were either 1-50GB or in the case of audiovisual collections c.1TB. Of the respondents who created research data, 3 were unsure how much they had.

**3. Who is responsible for managing your research data?**

The primary response was 'me'. In the case of AHRC grants the Performance Data Archive was noted and ESRC grants were deposited with the UKDA. 2 people noted material was passed to the librarian and some reference was made to University support for backup.

**4. Do you follow any policies or have a data management plan in place?**

Most researchers noted the lack of formal University policies but mentioned practices they follow. These ranged from keeping all material in one place (the hard drive of their office pc), using a version control table on shared reports, Dublin Core metadata for digitisation projects, and agreeing file naming conventions and structures at the outset of collaborative projects. Some people found it difficult to work with others due to differences in working practices. Data management plans were only created when required (i.e. for AHRC and ESRC grants) – Sam and the data archives assisted in the creation of these.

**5. What happens in terms of backup? – how often, on what media, and by whom**

5 people were unaware what was in place in terms of backup. Others were aware networked storage was backed up centrally but only one could provide details (done weekly by local Arts Faculty Support). Two respondents had semi-regular processes for data held locally or on external devices. One person said they made sure data created remotely was updated to the server asap to avoid version issues.

**6. Is your data accessible – can it be reused?**

Several respondents noted rights issues that prevented data from being reused or meant it was only available internally or for education. The librarian noted a particular issue with legacy collections – often the correct agreements were not made at the time of capture and can't be renegotiated. Public access and reuse is common for research council funded projects. University research tends to be internal.

**7. Any other comments**

Comments included:

- I'm really pleased we've been selected as a pilot. Staff here create lots of research data that's underexploited.
- I can see the benefits of effective data management from AHRC projects I've run but to be honest I just want to focus on the research, not have to become an archivist as well.
- Any changes that come out of this need to respect our working practices – it's not always appropriate to make data accessible or to expose our research methods.

**Audit Form 2: Inventory of data assets**

Name of data asset	Asset description	Asset Manager(s)	Reference	General comments
Changing Approaches to Theatrical Tuition	A database listing theatre lecturers and tutors that have worked in TFTS since the 1980s along with their teaching methods. The study will lead to a conference and publications on how teaching practices have evolved and reflecting broader societal change.	Christie Rohan Sam Lever	In progress - not widely accessible yet. Copy held on library server:  L:/data/master/catt	A McArthur funded project. A follow-up AHRC proposal is planned.
Glasgow Cinemas	A database of cinema screenings in the 1920s and 1930s in Glasgow. It records what films screened, where, for how long, and attendance figures.	UKDA	<a href="http://www.ukda.ac.uk">www.ukda.ac.uk</a>  I think library links to it	ESRC funded so already deposited with UKDA
Theatre photos	A library resource on student plays. Boxes of photos by year to browse. Also has an online catalogue to search by production, student name, date and keywords. Some video too I think...	Departmental librarian	Library	n/a
Shakespearean performance database	A database listing performances of Shakespeare's works in Stratford. Details venue, dates, cast and production company. Often includes images. There are also qualitative outputs such as interview transcripts and video recordings included where possible.	Christie Rohan PDA	<a href="http://www.pda.ac.uk/shsp">www.pda.ac.uk/shsp</a>  My original research is at: H:/research/AHRC/Shakespeare	Created by Christie Rohan with AHRC funding.

Student productions audiovisual collection	Collection of photographs and video of student productions staged in TFTS since 1979. Held in Departmental library. Consists of a mix of over 6500 analogue and digital images and 80 hours of video. Video mostly from last 6-8 years.	Sam Lever	<a href="http://www.libcat.gla.ac.uk/4235">www.libcat.gla.ac.uk/4235</a>	Continually being added to by students. Used on regular basis for teaching and as source of inspiration.
TV oral history project	Oral history recordings on the impact of TV in Glasgow. Available internally.	David Mann	Library	Longitudinal study run by last senior lecturer in TV.
Losing the Act	Video material from within the department and online is being collated to assess how performance is recorded and what is lost in this process.	Mark Arnott	Need to contact Mark for details – abstract not yet online	A PhD project
Attitudes to Shakespeare	A monograph by Christie Rohan on the popularity of different Shakespearean works over the past decades.	Harper Collins?	<a href="http://www.harpercollins.com/Title/pages_objid=39627">www.harpercollins.com/Title/pages_objid=39627</a> Main library has a copy	n/a
Moving History	Digitised film collections. Footage owned and created by the department is maintained by Scottish Screen Archive and was included in this national digitisation project. Available online.	Scottish Screen PDA	<a href="http://www.pda.ac.uk/movhist">www.pda.ac.uk/movhist</a>	HLF funded. The Scottish collection includes footage of University staff and students in the 1890s
Student research outputs	Various images, databases and videos are being created and used by doctoral students. Lots of these are funded by the University so we should keep them.	Several students	Best to ask James Ferguson or secretary for list of student projects	n/a

## **Data interview: Sam Lever**

Interview arranged by email. Follow up on [sl2z@tfts.arts.gla.ac.uk](mailto:sl2z@tfts.arts.gla.ac.uk) or x3549

### **What role do you have in terms of the Departmental research data?**

Well, as librarian I need to make sure staff and students have access to resources. There is a lot of useful research going on in the Department but the outputs aren't always shared so I assumed an unofficial role to track projects and encourage people to place data in the library at the end. We don't really have the resources to manage this properly but it does make it accessible. Students are sometimes employed part-time to make collections more usable, for example by cataloguing.

### **Have you encountered any data issues?**

Several! We don't have any formal guidance or policies so things aren't always done in the best way. We've had issues making data available due to rights, and poor naming conventions and capture processes can make the cataloguing I do with the students quite time consuming. There's a fair amount of research that falls under the radar too. A number of researchers aren't aware of how their data is placed at risk. They think that if it's on their pc it will be safe – somebody can always get it back for them. It's only when they've lost material or been forced to think about curation – for example on AHRC or ESRC projects – that this mindset changes.

The greatest challenge for me is consistency, especially for shared collections like the student theatre videos and images. We try to have naming rules and procedures for data capture as this is being added to every year but people do things their own way. There can also be authority control issues. As we need to tidy up new material and add it to overall catalogue, access isn't provided for a month or so. Some researchers keep their own copy and use that in the meantime, which can confuse things.

### **What are the main research assets you are responsible for?**

Generally we're just involved in maintaining collections but with student theatre we have more of a hand in creation. I'm trying to get involved in student research too as a lot of data is created but long-term curation falls to the student and supervisor.

#### Student theatre audiovisual collection

This is the most used and regularly updated collection. I'd say it's accessed at least once a day, probably because it's so varied – you can find Beckett, Shakespeare, Chekov and students' own material. It ranges from the early days of the department to now. It's something the department is renowned for. There has been a long standing agreement that all rights are waived so staff and students can access and reuse this material in any way they want. That's quite rare. The collection is added to about 5 or 6 times a year. Students capture the events. Photos used to be black and white, occasionally we have negatives on file but mostly just prints. Now we use a digital camera – I try to make sure we create TIFF files but often I get JPEGs back. The video is much better. The technician is really strict about videoing performance and makes sure the image and sound quality is really high. Videos are either MPEG2 or AVI and any audio is WAV. It's a large and growing collection too – c. 1TB.

### TV oral histories

I mention this project as I'm semi-responsible for it. Dave Mann who's done all the interviews and recordings has recently retired so the Head of TFTS has asked me to look after the material. It was a pet project for Dave – he started it as a student and kept on adding material when he could by interviewing the general public. There are a few reuse issues with it. The collection dates from early 1960s-2006 and early agreements aren't always clear on what's permissible in terms of access. Dave administered this on a case by case basis. It was only a couple a people a year so not too onerous. It's likely the early ones will remain closed, which is a shame as the topics have broad use. He asked how TV affected home life, what people watched, who had access to TV, whether they watched in groups etc. It's taking a while to work out his filing; the collection is quite cumbersome to access. Most of the older recordings on tape were digitised with a Uni grant in 2004 to WAV for preservation and MP3 for access. I've kept these on the local server ever since.

### **What happens in terms of managing / archiving these?**

Any digital resources in the library are held on the local server and backed up by Arts Faculty Support every week. We try to keep on top of things that may be lost by digitising. I keep master copies of all records in a closed area on the server and only let people use access copies. The physical resources like the photos are kept locked away. I let users browse them and borrow if absolutely necessary – they should be kept in the library though. Some researchers are defensive about 'their collections' wanting to control access and make changes as and when. Once it comes to the library though it's fixed; any new versions or additions are properly recorded.

### **What would you like to see happen in the future in terms of data management?**

The main recommendation is to formalise procedures so we stop running into issues such as data loss and irretrievability. I'm glad the IR is being developed as we can't really preserve research data – we just make it accessible where we can. Creating resources properly in the first place would save a lot of effort later on.

## Data interview: Christie Rohan

Interview arranged by email. Follow up on [cr8q@tfts.arts.gla.ac.uk](mailto:cr8q@tfts.arts.gla.ac.uk) or x2758

### What role do you have in terms of the Departmental research data?

First and foremost I'm a researcher so my role is as a creator. As Head of Theatre I have a role to play in teaching too but my focus is on staff and student research. I use other people's data extensively and have run several collaborative projects. As the Principal Investigator on these I was responsible for data management so had to decide what standards and approaches we'd adopt. Creating data that would be accepted for deposit by PDA was quite a challenge initially.

### Have you encountered any data issues?

Embarrassingly I have actually lost data before. A few years back I kept the only recording of a student play on disk and it became damaged and wouldn't play. It wasn't backed up either, so that's it gone forever. I follow procedure now! I've come across other issues when working with colleagues. It can be hard at times to find their resources as they're named and filed in a different way. It's easier once resources are in the library and standardised. I tend to follow the procedures I use on AHRC projects in all research now as a little effort early on really does pay off.

### What are the main research assets you are responsible for?

#### Shakespearean performance database

This database was created over four years from 1998-2002 from information in local archives and old newspapers. It covers all productions of Shakespeare's plays in Stratford from 1970-2000 and gives various details such as the play, the cast, how it was staged, the run, audience and any recordings of it. I collected the data to analyse attitudes towards Shakespeare and track popularity of plays and staging methods. I published the results in 2004. I was given loads of additional material such as photos, videos, ephemera... and interviewed several artists. These were digitised and added where possible but not all made it into the collection as this wasn't part of planned work. These are still here in my office in boxes or on my personal networked storage space. The database was deposited with PDA in 2003. I created it in Microsoft Access but they wanted a csv file for preservation so I exported the data. It's not updated now but people use it regularly – apparently 5,000 unique users a month. It was all cleared for educational access only.

#### Changing approaches to theatrical tuition

For the past two years I've been working with colleagues here – Simon Black and Linda Pitt – to assess how approaches to teaching theatre have changed since the 1980s. We've used the student productions av collection, departmental play scripts and personal teaching collections extensively. Most department staff have all agreed to share details and are happy for the final database to be publicly accessible – some will need to be restricted though. We created a database in the same way as on the Shakespearean performance project. The research collection phase finished in June 2008 and we've since been analysing the data for publication. Sam is maintaining it for us. A conference is planned for September 2009. The research will form part of an AHRC proposal to run a comparative study at the University of Edinburgh and RSAMD so we can assess Scottish approaches more widely.

### PhD students

I also do a lot of work with students, especially for the student productions av collection but I imagine Sam has told you about that as he's ultimately responsible. When students create data or collect together materials to assess such as in Mark's project "Losing the Act" I make sure they think about what will happen about long-term preservation. Mark is lucky as the Arts Council provides a repository and fund data training on their Fellowships. The University provides a number of scholarships too but there isn't such a robust process for curation as yet. Sam does what he can through the library but isn't always able to help. The new IR will definitely be of use.

### **What happens in terms of managing / archiving these?**

I try to make sure somebody more qualified than me takes over the data in the long-term. I'm not really interested in doing the curation – I'd rather focus on research. Having the library here is great as we've been able to ask Sam's advice and get him to maintain shared resources like the theatrical tuition database. The Shakespeare collection is kept by PDA and the things I have here are on the networked drive so that should be sufficient for backup. I tend to keep research data on my personal space so nobody else can access or change it.

### **What would you like to see happen in the future in terms of data management?**

One thing that concerns me is equipping students properly. It took me a while to understand the concept of digital preservation. If we taught student research practices and noted data issues this would help. I was concerned that University data was not being managed so I'm glad about this project. I do wonder though how blurred boundaries will be managed. Our theatrical tuition project is McArthur funded but the follow-up may be supported by AHRC and deposited with PDA. We need to make sure links are maintained between related datasets if so.