ARCmagazine

www.archives.org.uk

The need is GROWING
DIG FOR VICTORY STILL

Shakespeare birthplace trust

Menu

Discover Shakespeare

Everything ⌄   Search the collection...   🔍

Archives and Technology Issue

Discover Shakespeare Page 17

Archives & Records Association
UK & Ireland

CONFERENCE

2017

Keep up to date with all the news and build up to this year's conference at http://conference.archives.org.uk

HILTON MANCHESTER DEANSGATE
30th August - 1st September 2017

ARCmagazine

Contents

Welcome to **ARC Magazine** July 2017

Welcome to the July edition of *ARC*. This month's Section for Archives and Technology issue contains a diverse range of articles, all of which reflect the experimental and frequently pioneering work currently taking place across the sector.

Archivists, record managers and conservators are welcoming opportunities and challenges of our digital society with confidence and enthusiasm. These trends are reflected in the activities undertaken by repositories across the UK and Ireland – some of which feature in this issue – and in the work of the Section for Archives and Technology.

The Section aims to build a community of practice and research in archives, records and technology, to provide a forum for discussion and support, and works towards best-practice in the curation of digitised and born-digital materials.

To find out more about their activities, visit the ARA website:

http://www.archives.org.uk/about/sections-interest-groups/archives-a-technology.html

Finally, thank you to all the authors who have given up their valuable time to share their expert knowledge, and to Ann MacDonald for sourcing the articles included in this issue.

Ellie Pridgeon
ARC **Editor**

opening lines

Once again it was an immense pleasure to host these events in conjunction with ARA and IRMS and, if you joined us for one of these sessions, thank you for your contribution. I would also like to thank all our guest speakers who shared their experiences and best-practice recommendations.

If you'd like to learn more about how preservation and access can create value across your organisation then come to Manchester at the end of August for the ARA's annual Conference, where there will be a three-day 'digital stream' of presentations and practical sessions to bring you up to date on all things digital.

As the Preservica team will also be there, I hope to see you then!

David Portman
Preservica

It's hard to believe another year has passed, and with it comes the end of a highly rated series of practical digital preservation events.

Taking place in London and Edinburgh, this year's successful programme focused on supporting Archivists and Records Managers to raise awareness with their key business decision makers and information technologists of the importance of protecting long-term and permanent digital information and the value that can be delivered to the organisation in doing so.

I was fortunate enough to take part in the programme this year, providing me, along with attendees, a unique opportunity to hear from a number of highly-respected industry experts and digital preservation practitioners who lent their expertise to the programme, including William Kilbride of the Digital Preservation Coalition and Jason R. Baron of DrinkerBiddle LLP.

Inspiring digital preservation project stories and accomplishments were shared by Tina Staples & James Mortlock at HSBC and Emily Nimmo & Lesley Ferguson at Historic Environment Scotland.

Correction

In the June edition of ARC, in Charlotte Dennard's excellent piece on GEC Traction (pp36-37), we mistakenly referenced the copyright of two images as 'Alstom/GE.' The copyright references should have read 'With the kind permission of ALSTOM UK Holdings Limited'. Our apologies to Charlotte and her colleagues and to ALSTOM UK Holdings. We have amended the digital version of the June edition accordingly.

The Editors

Continuing Professional Development (CPD) News
(Incorporating Registration News)

ARA CPD

Three big bits of news on ARA CPD this month as we move ever close towards launch at Conference in less than three months.

We are delighted to welcome **Chris Sheridan** as the new ARA CPD Programme Manager. Chris has two decades of experience in working with CPD programmes, including with membership organisations. He started formally on the job the week of 19 June but will very soon be connecting with many members. His early mission (now that he has chosen to accept it), of course, will be to review the Programme design and get things tickety-boo for launch. In amongst all that will be briefing and introductory training sessions on ARA CPD during the autumn. The Board and staff of the Association and the CPD Steering Group (and sub-groups) will all be on hand to help him.

While we are on the subject of launch, we will soon put out a formal call for people willing to act as mentors to all levels of the new ARA CPD qualifications: Foundation, Registration and Fellowship. We know that time is increasingly at a premium in everyone's working lives. But do please consider taking part as a mentor. It's one way of maintaining your own CPD, if you enrol, or staying in touch with the profession (eg, perhaps if you are approaching retirement) or simply widening your network of contacts. With our new online system, *Mahara*, gone are the days of shunting papers backwards and forwards in the mail. Hopefully, this should mean that any mentoring time that you are able to devote will be light on bureaucracy.

Finally, quick feedback from the ARA Board meeting in Stirling on 12 June. The Board has given its final approval to the current design/principles of ARA CPD. This will enable us to crunch the numbers and identify the **fees** for taking part. The CEO (John Chambers) will present options to the next Board meeting on 29 August, in time for Conference and the launch of the programme. ARA CPD has to break even to be viable, so the options will address ways of achieving this, while also taking account of the financial pressures that many members face. We are all keenly aware of this and the need to get the balance right.

Registration

With just three months to go before the final deadline for submissions under the 'old style' scheme, many of you will be adding the final touches to your portfolios. Last month, the need for good evidence was discussed. However, whilst it is important, evidence is only one component of the portfolio, the bulk of which will be made up of the 'learning outcome forms' (LOFs).

Candidates should now have a clear idea about which LOFs will be included in their portfolios. The final decision over LOFs provides a good opportunity to think about the portfolio as a whole. Are there areas where you can cross-reference one LOF with another? Did the undertaking of a period of study or training contribute to the successful completion of a work achievement? Certainly, some of these ideas should be included in the 'follow up / continued benefit' section as reflections on your various activities. However, taking an overview across the whole portfolio allows a candidate to demonstrate that they are a well-rounded professional. After all, our professional skills and activities don't exist in silos, separated from one another. As we develop in one area, so we improve in others too, and this fact will be reflected in a good portfolio.

Contacts

General Registration Scheme Enquiries:
registrar@archives.org.uk
Registration Scheme Events Enquiries:
regschemeevents@archives.org.uk
Registration Scheme Admin and Bursaries:
regschemeadmin@archives.org.uk
Registration Scheme Communications Officer:
regschemecomms@archives.org.uk
Registration Scheme Mentor Queries and Advice:
regschemementors@archives.org.uk

Jon Elliott
Head of Public Affairs, ARA
Richard Wragg
Communications Officer, Registration Scheme

> *Don't forget: Candidates have 3 months to submit their portfolio under the existing Registration Scheme.*

A basic guide to ARA's new CPD programme: part four

Fellowship – a New CPD Pathway and ARA Qualification

In the last month's issue of ARC, we gave you an insight into one of ARA CPD's innovations: the new Foundation level of professional qualification, for members entering the profession without a formal academic qualification in record-keeping. This month, we will take a look at a new CPD pathway, **Fellowship.**

The original CPD survey of ARA members in 2010 revealed a strong wish for a higher level of assessed CPD, one that would follow on from the already well-established Registration Scheme. Members achieving Registered status under the soon-to-end Registration Scheme lamented the fact that once they had 'Registered', 'that was it.' They said that they wanted to be able to set their sights higher and have a framework for demonstrating their personal and professional development at a more advanced level.

After looking carefully at a range of options, the CPD Steering Group and Pilot Group came up with a robust but flexible CPD offer for ARA members with a significant record of achievement in the sector. A sub-group headed by Gillian Sheldrick and Heather Romaine then reviewed the requirements to make sure they work for professionals working across archives, records management and conservation at a senior level. The result is the new Fellowship qualification.

The Fundamentals

The current ARA CPD Programme Criteria – including for Fellowship - are available on the CPD pages of the website (www.archives.org.uk/cpd) They come with the usual caveat: all criteria are subject to ongoing and constant revisions as you would expect from a programme that is new, driven by participants themselves and that needs to be responsive to participants' experiences. But we can summarise the basic concept of Fellowship – and the logic behind it - as follows.

Anyone wishing to enrol in ARA CPD with a view to assessment for Fellowship will of course need to be a member of the ARA. S/he will have secured Registered status under the current scheme that is ending or (in due course) the new Registration qualification under ARA CPD. Members of ARA and to have held Registration for nine years become eligible to be assessed for Fellowship. But they will need to have enrolled in ARA CPD and revalidated their Registration qualification/status at least once through the new programme.

The Logic

Why are we doing it this way? Members told us that they wanted a CPD Programme based on merit and that represented a clear break with the time (now long since past) when Registered status was a matter of simply paying a fee. Therefore, making 'revalidation' of one's Registered status or Registration qualification before becoming eligible for Fellowship assessment ensures a level playing field for all members.

In addition, members have been telling us that they do not want CPD to be about 'seniority' or long service or being an 'insider' in the ARA or the wider profession, but achievement in one's chosen field or fields. Some eventual Fellows will no doubt be people with a record of achievement at a senior management level and have had extensive and varied careers, but the point is that you will not have to 'fit' such a mould to be eligible for Fellowship.

Interested?

If you are keen to work towards your ARA CPD Fellowship qualification following the launch of the new Programme at Conference in August, you should enrol just like anyone else. If you have been Registered under the current scheme for at least nine years, you may in theory be able to revalidate at Registered level as early as the spring of 2018 (a 'light-touch' assessment compared to usual Registration) and then move onto fellowship assessment in autumn 2018.

But, in reality, unless you have been maintaining evidence of your CPD towards the new Fellowship level for some years in preparation for the launch of ARA CPD, it will make sense for most people to plan for an assessment later and build their evidence more gradually. Fellowship – quite rightly – is designed to be the pinnacle of the profession.

The bar will be set high.

The Process

Fellowship assessment will follow the same format as for Foundation and Registration. All levels of ARA CPD afre based on the same competency framework (updated version on the CPD website pages), submitting an electronic portfolio of write-ups and a CV. To achieve the Fellowship qualification, you will need to pass six competencies from at levels 4 and 5 (the highest level of marking).

Candidates for Fellowship will also need to write two **additional supporting statements**, one about their CPD development and one about how they have shared their experience and expertise with others – ie, what they have given back to the profession. Furthermore, candidates will attend a 'professional review' to explore their portfolio in more detail.

And sticking with the theme of merit, successful Fellowship candidates will not escape revalidation: they will be required to do so every 5 years (again, a 'light-touch' process demonstrating that they are maintaining evidence of their CPD, not a full repeat of Fellowship assessment).

What, in the end – you may ask – do successful applicants get for all this endeavour and effort? The ARA will issue **new post-nominals** for the Fellowship qualification (**FARA**). Over time, we anticipate that the Fellowship will become as prominent in the recordkeeping world as it is in so many other professions.

New CPD Logo

> "Fellowship – quite rightly – is designed to be the pinnacle of the record-keeping profession. The bar will be set high."

ARA Conference 2017 Update

Elizabeth Denham: Final Confirmed Keynote Speaker

We have our final confirmed Conference keynote speaker. **Elizabeth Denham** has been the UK's Information Commissioner since mid-2016. A native of Canada, Ms Denham is a former Information and Privacy Commissioner for British Columbia and Assistant Privacy Commissioner of Canada.

Since her appointment as UK ICO, based just outside Manchester, Ms Denham has focused extensively on the pro-active disclosure of records, on data protection in the social media space and public access rights.

Ms Denham will deliver the keynote address – the first item on the second morning of Conference - on **31 August**. She will be among fellow professionals. She has a BA in History and a Master's degree in Archival and Information Science from UBC, so is conversant with both records management and archival issues from a practical and operational perspective.

At Conference, Ms Denham will set out her views on the new **EU General Data Protection Regulation (GDPR), Freedom of Information and nuisance marketing**, among many others. There will be time for questions.

So, if you or your organisation want the latest in UK Government thinking on any of these issues from someone who is a leading thinker in her own right and who understands the world you work in, sign up to attend Conference now, via the new website pages: conference.archives.org.uk   And **remember: your three-day conference fee includes comfortable, on-site accommodation and all meals**.

We are delighted that all three keynote speakers – Dr Alan Billings, Dr Neil Brodie and Ms Denham - have agreed to join us at Conference. Collectively,

Conference keynote speaker. Elizabeth Denham

they offer a full-spectrum coverage of leading-edge challenges facing practitioners in the archives, records-management and conservation world; and they will complement what is already an impressive three-day programme.

To register to attend Conference, visit: http://conference.archives.org.uk/

John Chambers
CEO, ARA

> 66 *ARA Conference: Get up to Speed with GDPR and Freedom of Information* 99

Collecting Matters

Archive Service Accreditation has recently had its fourth birthday, and with over 70 accredited archive services across the UK, the time is right to carry out a 'light touch' refresh of the programme.

The refresh will look at clarifying areas, at the supporting documentation, and at how archive services have used Archive Service Accreditation. It will also be a chance to check in to ensure barriers to participating in Accreditation are as low as possible.

The review of documentation will be an opportunity to include increased expectations concerning the management of born-digital collections. At their annual meeting in May 2017, the Archive Service Accreditation Committee approved changes piloted in 2016-17 to be shared with the sector, and then incorporated into the programme.

These changes reflect and support the progress that archive services are making in digital preservation. As well as asking archive services to comment on how they meet the National Digital Stewardship Alliance (NDSA) 'levels of preservation' – a simple risk-based maturity model. Applicants will demonstrate their activity in digital collecting, access, storage provision and sustainable budgets, and provide evidence of staff with sufficient skills and knowledge for this critical developing work area.

The programme will re-emphasise the importance of policies, plans and procedures reflecting the needs of digital as well as analogue holdings.

The National Archives' (TNA) website has further information about this for archive services which will be applying for Archive Service Accreditation after April 2018, when the changes become effective.

For the first time, archive services holding only born-digital collections will be eligible to apply for Archive Service Accreditation too.

Melinda Haunton and Jane Shillaker
The National Archives (UK)

Website: www.nationalarchives.gov.uk/accreditation
Email: accreditation@nationalarchives.gsi.gov.uk

Welcome to Archives and Technology special issue of ARC

An increasingly digital society offers a bright and relevant landscape for archives and records. The diversity of articles that have been submitted for this special issue prove how versatile and adaptable the archive profession is to new technologies: whether it is developing podcasts for mobile platforms, or meeting the complex requirements of mass-scale digital preservation.

In 2017, the Section for Archives and Technology (SAT) decided that we were going to be very positively digital. SAT sponsored the Digital Preservation Coalition's annual student conference in January because we understand how much new professionals are valuing their digital toolkit. At our 'digital ambition' event held in Hull – the UK City of Culture – in 2017, we had a range of presentations covering cataloguing systems, collaborative digital preservation initiative, and digital skills. We have continued this 'digital ambition' theme in this special issue, and the articles which were submitted were exactly the inspiration for which we were looking.

From The National Archives, Mark Bell and Sonia Ranade share some of their key questions, processes and solutions for curating a vast and varied public digital archive. At the Royal College of Nursing, Elisabeth Thurlow explores solutions to managing a digital repository, tackling digital preservation challenges one pilot project at a time. This is a theme that appears throughout the issue: that collection managers are trying, testing, developing, and learning.

Amy Hurst and Jennifer Reid from the Shakespeare Birthplace Trust describe the journey from a digital vision into an integrated collections management system, and resulting in the Discover Shakespeare user interface. At the Electricity Supply Board in Ireland, Kirsten Mulrennan considers holistic approaches to delivering content through digital media – from websites to infographics. Underpinning archival digital ambition is the usability and accessibility of material, a key concept that runs through the special issue.

Digital culture has dramatically transformed over the past few years, and it is very exciting to see the digital opportunities that are being seized by professionals across the archive sector. Jenny Bunn considers the requirement for sector-wide dialogue on archival descriptive standards with the release of the International Council on Archives' (ICA) 'records in contexts' standard. The launch of an updated and refreshed Archives Hub is covered by Jane Stevenson, who explains how developments have improved data accuracy.

Archivists are welcoming digital challenges as a playground that is both experimental and innovative. Elspeth Millar from the British Stand-Up Comedy Archive at the University of Kent describes public engagement through podcasts, and Louise Pichel discusses her investigation of running AtoM on credit-card sized Raspberry Pi computers. Celebrating Hull City of Culture, Simon Wilson considers the opportunity for The National Archives to capture the impact of 'Hull 2017', and the digital legacy that will arise from City of Culture status.

Further to this, all of the articles for this issue convey a joy and confidence in managing the digital landscape.

Ann MacDonald
ARA Section for Archives and Technology

Digital records – as safe as paper?

It all started with a business case. In February 2011, the Public Record Office of Northern Ireland's (PRONI) business case for a digital preservation solution was approved, and a formal project launched the following month. Although the high-level plan suggested that all the steps along the way had been identified, the reality was that the team was engaging in research from the outset. A number of high-level decisions were made early on in the project:

- We would build our system in line with (our interpretation of) Open Archival Information Systems (OAIS) reference model
- Digital records would be catalogued in PRONI's master catalogue alongside the more traditional records, rather than being accessed via a separate catalogue or system
- We would use existing tools where possible – after all there is no point reinventing the wheel
- We would use csv files rather than XML for metadata (due to a lack of expertise in XML)
- We would create a standard structure for every accession

The first action we carry out on an accession is to load it onto our Quarantine server, and virus-check it using Symantec Endpoint Protection. Everything will be subject to a separate virus check using Sophos Endpoint Security.

We then create an inventory of the transfer using TreeSize Pro. This creates a listing of every folder and digital record, and generates a MD5 checksum for each record.

It became clear to us very early on that, since data could come to us in almost any structure, we would have to 'prepare' the data before it could be considered a Submission Information Package (SIP). This involves a number of tasks, and requires substantial manual intervention:

- Restructuring the data to conform with our standard structure
- Checking and – if necessary – amending path / filename lengths
- Generating a unique ID for every digital object
- Adding additional, mandatory metadata, especially access restrictions and rights information
- Identifying file format information using DROID, and adding this to the metadata files

The accession is now packaged using Bagger, which also creates a MD5 checksum, and we compare this against the one generated by TreeSize Pro. Finally, we add the DROID csv file into the package. This is our SIP, which can now be transferred to the data preparation system for further processing.

An example of an inventory of records

The main processes here include:
- Checking the checksums
- Updating our file format register
- Creating a metadata csv file for every digital object
- Creating an export file for transfer to our master catalogue system
- Creating any access formats that may be required

The file format information created by DROID is checked against our file format register to see if the accession contains any new file formats. If there are new formats, the process stops, and we have to manually update the register with the preservation and access formats we plan to use for these formats.

This is what we regard as an Archival Information Package (AIP), and is ready for the final stage of its journey, at least for now. This final stage sees:

- The transfer of the accession to our archival storage system
- Overnight replication of the accession to another server in a physically separate data centre
- Metadata extracted and stored in a SQL database
- Checksums checked.

The records that are 'open' (available to the public) will be copied to our access filestores and accessed via our on-site catalogue – CALM – and our public catalogue – eCatalogue. We regard this as our dissemination information package.

This also sees the start of a new process as checksums will have to be checked and the number of records in each accession compared on an ongoing basis. In addition, we will have to monitor file formats regularly with a view to identifying potential preservation interventions. In this way, we hope to keep our digital records (at least) as safe as paper! The steps taken to date have, just like the original business case, proved to be an initial starting point. The system continues to evolve with implementation and practical application, and I imagine that this will continue for some time to come.

Hugh Campbell
PRONI

Hull: City of Culture on the horizon

In November 2013, Hull was announced as UK City of Culture for 2017. The scheme, echoing that of European City of Culture, seeks to use culture as the focus for regeneration. The award does not come with its own dedicated funds, but with the expectation of approaches to national funding agencies.

A strong element of Hull's submission was the fact that it had secured £18m from local organisations and businesses (this fundraising has continued and is now over £32m), with the simple premise that the more funding that is secured, the more that can be done. Unlike the Olympics, there is no blueprint that is simply moved from city to city. City of Culture is what the city makes of it – and that includes the interpretation of what culture is!

As funding has been secured or committed and projects taken shape, it has been necessary to maintain a very high level of confidentiality. This did lead to lots of criticism that 'nothing was happening', but releasing the programme in stages has given organisers more time to plan, and has created a sense of anticipation.

University of Hull Archives
The University of Hull Archives will have several bites of the cherry. As a partner in the Hull History Centre, we are actively involved in a number of exhibitions and events –

Some iconic artefacts will be added to the archive, including the bicycle used to deliver the City of Culture bid in September 2013. Painted by a local artist, it features many references to the city's past. Photograph copyright Patrick Mateer.

including Hull's role in the outbreak of the Civil War in 1642, and Philip Larkin. In addition, we have a number of opportunities to support activities being developed across the university, including a programme relating to Hull authors that overlapped heavily with the university's strong literary holdings. We also supported the Antony Minghella retrospective season of films with a small exhibition from the archives.

Once in a lifetime opportunity
City of Culture represents a tremendous opportunity to do so much more than normal, not only in terms of our work, but also to develop partnerships and contribute to initiatives across the city. At the outset, organisers hoped the year would attract an additional one million visitors to Hull. The History Centre has noticed a healthy increase in the number of people visiting to look at our exhibitions,

but also to use the archives. In our planning, we have sought to create a programme of exhibitions and events that will appeal to existing and new audiences alike. We currently estimate that 2017 will represent the equivalent of four to five 'normal years', and need to look at how we can maintain the momentum and the growing expectations of our users in 2018 and beyond.

City of Culture archive
We have been working closely with the City of Culture team to 'archive City of Culture'. If you say it quickly it does not sound so bad...but we are expecting to receive a huge born-digital archive of files that will allow everybody to study the how, the why and what of it all.

The City of Culture archive was launched in Hull on 11 April. Jeff James, Chief Executive and Keeper at The

Jeff James (The National Archives), Martin Green (Hull 2017) and Glenn Burgess (University of Hull) at the launch of the Digital Archive. Photograph copyright Patrick Mateer.

National Archives (UK), who was in Hull to present Hull History Centre with its archive accreditation certificate and to help launch the archive, said: "The archive seeks to capture a pivotal point in Hull's history and will become a key part of the collective memory of the city. I'm looking forward to seeing how it matures and informs future cities of culture."

Martin Green, Chief Executive of Hull 2017, said: "The Hull 2017 archives will tell many tales from the creative minds behind the spectacular performances and events to the people of Hull who experienced it. It will be an important record of our incredible year which will inspire creativity and innovation for years to come."

We are currently pulling the pieces of this vast and ambitious jigsaw puzzle together. We are keen to engage with colleagues about some of the digital preservation and engagement challenges we face once these have become

more apparent to us! We will provide regular updates via various social media channels. So please like us, follow us and retweet us and visit the History Centre's blog:

http://hullhistorycentre.blogspot.co.uk

If you are passing by, do stop and pop in. The city will be having a year-long CultureFest, and the History Centre will be right in the middle of it all. If we survive the experience, we will share it in 2018.
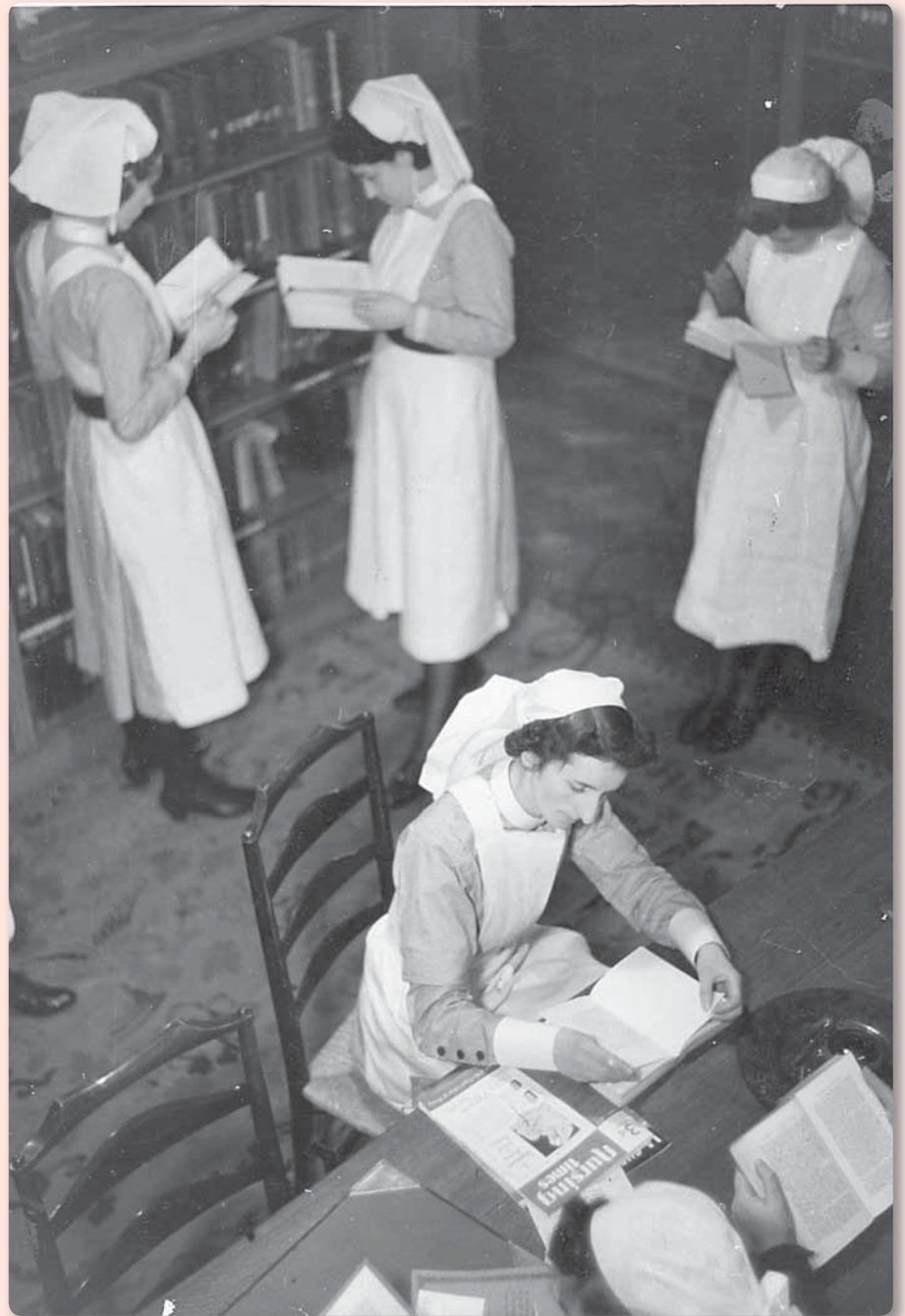
Simon Wilson
University of Hull Archives

Voice of nursing: improving patient care

Last year, the Royal College of Nursing celebrated its centenary, raising questions about what would survive for the celebration of our bicentenary. Inspired to find a solution to secure its digital future, the College initiated a strategy, including a new archive post and a digital repository from January 2017. In the first of our digital preservation pilots, we are focusing on born-digital publications created by the College, which include clinical guidance. Once ingested into our digital repository, we will trial and explore how we ensure long-term access to key resources, and appropriate quality assurance.

The College was founded in 1916, in the middle of the First World War, as a professional organisation for trained nurses. The College founders set out to improve the skills and knowledge of nurses through education for the benefit of the patient. Today, the College is a professional membership organisation and trade union supporting over 435,000 nurses and health care workers, and producing and delivering information for members and the public that is evidence-based, accurate, and up-to-date.

To safeguard our information which is increasingly being created or born digitally, the College purchased a hosted digital repository system – Preservica Cloud Edition – in December 2016, and I joined the team

Nurses studying in the Royal College of Nursing Library. c.1945. Photograph copyright Royal College of Nursing.

as their first digital archivist in January 2017. Whilst the College has employed a professional archivist since 1986, this is the first dedicated digital archives role.

During 2017, we are carrying out a series of small pilot projects targeting key departments. In the first, we are working with our communications department, focusing on capturing a sample of Royal College of Nursing publications. Our publications cover a wide range of topics, from clinical guidance to advice on bullying. They

support best-practice in nursing, which in turn supports best-practice in patient care. However, best-practice changes over time, and our publications track the changes which took place and when.

Our initial sample – College publications produced in 2010 - contains over 100 publications, mainly saved as PDF documents. We are using this sample to establish the process for transferring these digital resources to the digital repository. Following this, we will expand the pilot to

capture publications produced from 2011 onwards, and our earlier surviving digital publications.

We are in the early stages of our digital repository project, but, as a membership organisation, we are already thinking about how we will provide access for our members. Access to our current born-digital publications is provided via the College's website, but a recent move to a new website has highlighted the loss of access to older publications, some of which are still useful in practice. To support greater access, we have also been researching the potential use of the 'universal access module' offered by Preservica, and its ability to meet the expectations of our members and the public.

We want to ensure future access to these key resources for future generations of nurses, researchers and fellow clinicians outside nursing. However, due to the nature of our publications, we must ensure appropriate quality assurance when providing access to historical healthcare guidance alongside contemporary guidance. The publications we archive must indicate whether or not they are current clinical practice if made available digitally. Moving forward, working closely with our communications colleagues, we envision that publications older than five years – and those marked as replaced, superseded or obsolete – will be taken down from the College website and added to the digital repository. They will be available via the library catalogue, but this will point towards them being held in the digital repository.

By safeguarding our digital records for the future, we will ensure that the nursing resources we create and provide online will remain available for future generations. Actively managing our digital legacy will allow us to continue to capture our corporate memory as the voice of nursing.

Elisabeth Thurlow
Royal College of Nursing

Campaigners demonstrating for nurses' pay. 2014. Photograph copyright Royal College of Nursing.

A health care assistant viewing Royal College of Nursing webpages. 2015. Photograph copyright Royal College of Nursing.

Discover Shakespeare: bringing together the archive, library and museum collections at the Shakespeare Birthplace Trust

At the Shakespeare Birthplace Trust, we care for and provide access to an internationally significant collection, which underpins our vision of leading the world's enjoyment and understanding of Shakespeare's works, life and times.

Discover Shakespeare homepage. Photograph copyright Shakespeare Birthplace Trust.

Our collection comprises archive records, library books and museum objects. The archive includes the records of local institutions and families, dating from the 12th century to the present day. It provides the key sources for Shakespeare's biography and the development of his home town and locale, including local government records, and the archive of the Royal Shakespeare Company. A sizeable research library and early printed book collection includes Shakespeare's own works, source books, and work by his literary peers and influences. Our museum collection of world-class early-modern material culture engages audiences with Shakespeare's world, whilst art and Shakespeareana tells the enduring story of

Shakespeare's cultural impact. This extremely diverse collection is of interest to a wide range of audiences, and one of our chief challenges is meeting the needs of all our users, as well as the specialist management requirements of this wide range of material.

The Shakespeare Birthplace Trust was established in 1847. Originally built as a heritage conservation trust, it soon began collecting archive, library and museum material. Initially, there was little distinction made between these categories, but as the collections grew, so did the professionalisation of the curatorial care. The material was managed by separate members of staff, and ultimately by separate departments. In recent times, the catalogue and collections management software was divided between CALM (archives), Talis (library) and Modes (museums).

Each of the curatorial teams had an affinity to 'their' system, but there were barriers to our knowledge across the collections. Whilst always apparent, these barriers became increasingly noticeable as more cross-collection projects were undertaken, and as researchers increasingly

suffered confusion and frustration with curatorial distinctions which were perceived as hindering ready access.
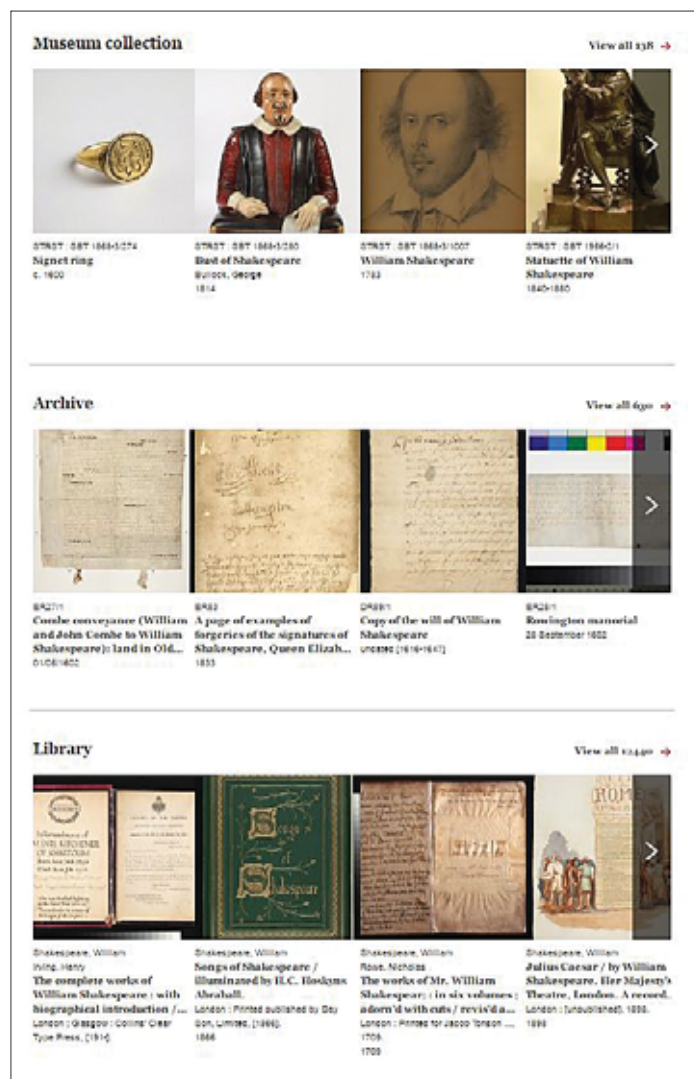
We envisioned an integrated system that would provide one point of access for users, and support increasing integration between the archive, library and museum teams in managing the collection. This approach would pool our resources, allow us to work more efficiently, share and learn from best practice across the team, and spark more creative ideas for engaging and building our audiences. We wanted to create a single point of entry to our collection that was visual and accessible to non-specialists and specialists alike, that would allow us to show off the breadth and variety of the collection.

We investigated several software options and discussed our requirements. We brought in an external consultant to help us to understand how we saw the collection internally, and how users wanted to interrogate our information. This research was crucial to the ultimate success of the project, but it was also a gruelling experience, as professionally-informed positions across the disciplines were analysed from differing perspectives. It became apparent that no single system on offer would immediately satisfy all legitimate requirements, but given the trust and mutual understanding fostered during this process, we felt confident in adopting an incremental approach. We therefore chose a flexible system that could – in time – be developed to meet all our needs and aspirations for professionally robust, integrated collections management services. Once settled on a way forward, we embarked on the task of bringing our three sets of data together, and agreeing joint systems and configurations. This has been both challenging and enlightening.

In the new Qi system, we have our own modules for each specialist curatorial area:

www.keepthinking.it/qi

These modules meet the standards dictated by the discipline, but we share modules for activities such as exhibitions, events and loans, as well as authority-controlled lists, such as people and organisations, and place records. We have all been surprised and heartened by how this change has impacted positively on our working relationships.

Discover Shakespeare 'everything' search return. Photograph copyright Shakespeare Birthplace Trust.

We are much more aware of the impact our work is having on each other, and we are already feeling the positive effect of more efficient working. For our users, they have the option to search 'everything' (library, archive and museum data), or still go to the separate datasets.

There have been a number of challenges in driving this project forward. Key challenges include legacy data and data improvement, and agreeing configuration and new processes across the disciplines. This has involved bringing all staff and our core / regular users along with the project, whilst also managing expectations of the pace of change. The flexibility of the system means we can look at making changes to the configuration to directly address concerns. The consensus of users has been positive, as revealed by our recent Public Services Quality Group (PSQG) survey results, where satisfaction regarding 'usability of the online catalogue' has increased by 36% since launch of the new system.

> 66 *This extremely diverse collection is of interest to a wide range of audiences, and one of our chief challenges is meeting the needs of all our users, as well as the specialist management requirements of this wide range of material.* 99

We are currently finalising a behind-the-scenes configuration, which provides a strong foundation for further improving the user experience. For example, we are reviewing online presentation, including list and extended tree view for archives – something that was not possible within the constraints of the project's first phase. We are also excited by the possibility of integrating the subject index cards from our reading room into Qi, possibly using keyword functionality, which would provide a form of 'guided' remote access. We are also considering legacy data improvement projects with cloud-based Qi, giving us new opportunities for approval workflows for remote volunteers. We have also been able to develop our first cross-collection online exhibitions, powered directly by Qi and integrated with the core catalogue.

We are still in the midst of this project, so it is hard to fully reflect. However, we think we have all been surprised by how transformative the ongoing process is. We have found that by sharing and learning about each other's collections and curatorial practices, we have been able to come up with more creative ideas to solve problems, and find new ways of sharing our collection with our audiences.

To explore Discover Shakespeare see:

http://collections.shakespeare.org.uk

Amy Hurst and Jennifer Reid
Shakespeare Birthplace Trust

Update: Descriptive Standards Roundtable

In 2013, the Section for Archives and Technology launched the Descriptive Standards Roundtable to describe its ongoing attempts to provoke discussion and debate around cataloguing standards and practice, since then, members of the profession have been brought together under this umbrella on a number of occasions, with attention coming to focus on the issues and challenges that are arising as cataloguing practice evolves in the face of born digital material.

With this focus in mind, attempts were made in 2016 to address the cataloguing of born digital material head on. These attempts coincided with a special issue of Archives and Records, which contained case studies of existing practice, as well as more reflective and theoretical pieces around the topic. At a United Kingdom Archives Discovery (UKAD) workshop in London on 18 March 2016, a speed writing event was held to collectively author a set of 'Best Guess Guidelines for Cataloguing Born Digital Material'. This document, based on ISAD(G), was subsequently published on the ARA website and opened to all for comment and consultation. The results of this exercise were discussed at a workshop at the ARA Conference in the late summer, and the intention was that subsequent to this event a second edition of the guidelines should be produced.

What the workshop proved, however, was that it is not easy to reach consensus, and that achieving the right balance between best guess and guideline, or localised and standardised practice. is never going to be simple. With this in mind, the question of what to do with the guidelines remained unanswered – and it is still undecided. As a snapshot of the thinking around this issue in 2016, as well as a record of some further (and often differing) thoughts

on that thinking, the guidelines do still have merit. They can be found at:

www.archives.org.uk/about/community/groups/viewgroup/50-best-guess-guidelines-for-cataloguing-born-digital-material.html

The future of the guidelines, however, remains uncertain, and we would welcome any suggestions on that score. Please email:

j.bunn@ucl.ac.uk

In the autumn, this uncertainty was increased with the publication of a new draft International Council on Archives (ICA) standard for archival description, namely Records in Contexts (RiC). The period allowed for consultation was fairly short, and so attention shifted from the 'Best Guess Guidelines' to facilitating a collective response to the draft. An event was arranged for November 2016, and was attended by around 20 individuals. A number of fruitful discussions took place, and a response was subsequently drawn up and submitted to the Experts Group. It can be found here:

www.archives.org.uk/about/sections-interest-groups/archives-a-technology/news-and-events-sp-1847188959.html

What the publication of the draft RiC standard shows clearly is that the descriptive standards landscape is changing, and it will be more important than ever for everyone to take part in the conversation going forward.

Jenny Bunn
University College, London

A History of Comedy in Several Objects

A History of Comedy in Several Objects – affectionately known as histcompod – is a new audio podcast from the British Stand-Up Comedy Archive (BSUCA), a group of collections held within the University of Kent's Special Collections & Archives.

BSUCA has been in existence since 2013, when two large collections relating to stand-up comedy were deposited. We then began to actively collect archival material relating to stand-up comedy in 2015, through a one-year funded BSUCA collection development, digitisation, and access project. The BSUCA now contains material from over 30 depositors, including comedians, comedy promoters, comedy producers, comedy venues, as well as academics studying popular and comic performance.

In late 2016, Oliver Double, Director of the Popular and Comic Performance research centre at the University of Kent, and Elspeth Millar, Digital Archivist, submitted an application to the University of Kent's Public Engagement with Research Fund, a fund specifically to support and encourage high-quality, innovative and effective public engagement activities based on the research undertaken at the university. Our application was for funding for a podcast that would focus on the

Squeezy stress hand grenade deposited by Mark Thomas as part of his collection. Photograph copyright University of Kent.

A History of Comedy in Several Objects logo. Photograph copyright University of Kent.

BSUCA collections, and the academic study and research of stand-up comedy being undertaken at the University of Kent. Podcasts are video or audio files that can be downloaded from the internet and played 'on the go' (on your mp3 player, your phone, in your car). They are often similar in structure and format to a radio programme, being generally between 15 minutes and 60 minutes per episode. The popularity of the podcast formats means that many broadcast radio programmes are also often now released as podcasts for people to download and listen to at a time convenient to them.

We were successful in our application to the Public Engagement with Research Fund, and began recording episodes of histcompod in January 2017. At the time of writing, six episodes have been released, out of a planned 30 episodes.

The premise of histcompod is that we feature a different record or object from the BSUCA collections each episode. We describe the object or record, place it within the context of the collection from which it is from, and actively interrogate it to find out what we can learn about the art and craft of stand-up comedy, or what we can learn about the recent history of the form. We also provide supplementary sources – such as links to other

Alternative Cabaret flyer. Photograph copyright University of Kent.

relevant sources in the BSUCA collections, links in our blog posts to further readings we have been doing in preparation for the podcast, and audio interviews with depositors – some of which are already archival recordings, and some are interviews we are going out to record especially for the podcast. Some of the records we have so far looked at include: a flyer from a collective of performers called Alternative Cabaret (one of the key groups in the early alternative cabaret / comedy scene in the late 1970s and early 1980s), a letter written in 1983 to some of those active in the alternative cabaret scene about the formation of a comedy trade union, set lists (on postcards) from Robin Ince, and a squeezy stress hand grenade deposited by Mark Thomas as part of his collection.

We wanted to do a podcast as it seemed a wonderful way to introduce new communities to the BSUCA collections, and to offer an opportunity for those communities to get involved and submit their own content. It was also fairly easy to accomplish in terms of technology.

Histcompod has proved to be an engaging, informative and entertaining way to introduce potential audiences to the British Stand-Up Comedy Archive collections, and to the academic study of stand-up comedy at the University of Kent, particularly those who perhaps would not visit an academic archive or library. The short and snappy nature of podcasts means that it has been a useful way to engage with people without taking up too much time (for example, people can listen whilst commuting or exercising). In other words, listeners do not need to go out of their way to listen to our podcast. It also makes the archive collections much more accessible as it delivers content straight to people and takes the BSUCA collections – out of the reading room in effect. Although the podcast is audio only, we have also been able to show images of some of the items we discuss via our Flickr site, our blog and our Twitter account.

A podcast also provides a way to introduce listeners to archival material more generally, and to institutions that look after archival material. It has offered us the opportunity to demonstrate to our listeners how we acquire material, how we process and preserve material, and how we provide access to material. Yet it also offers listeners the opportunity to think about how archival material and other historical sources should be approached, and what questions to think about when using primary sources. Who created the record and when? What was the purpose of the record? What other sources should be consulted alongside this record?

> *Perhaps most importantly, histcompod has provided us with an opportunity to directly engage with people and get people to submit their own content and ideas by email or by social media.*

The technology we needed to provide a good podcast has been readily available. As part of our application for funding, we did include budget for a dedicated recorder and lapel mics, and for a 35-hour project assistant post to help with editing. Otherwise we have used free software for editing (Audacity), and for publicity and promotion (Twitter, Facebook, WordPress, Flickr), and a minimal cost site for hosting the podcast (Jellycast). The podcast is then accessible through the iTunes and Acast platforms.

Perhaps most importantly, histcompod has provided us with an opportunity to directly engage with people and get people to submit their own content and ideas by email or by social media. Histcompod invites listeners to get involved in the podcast, and suggests ways of doing this, such as:
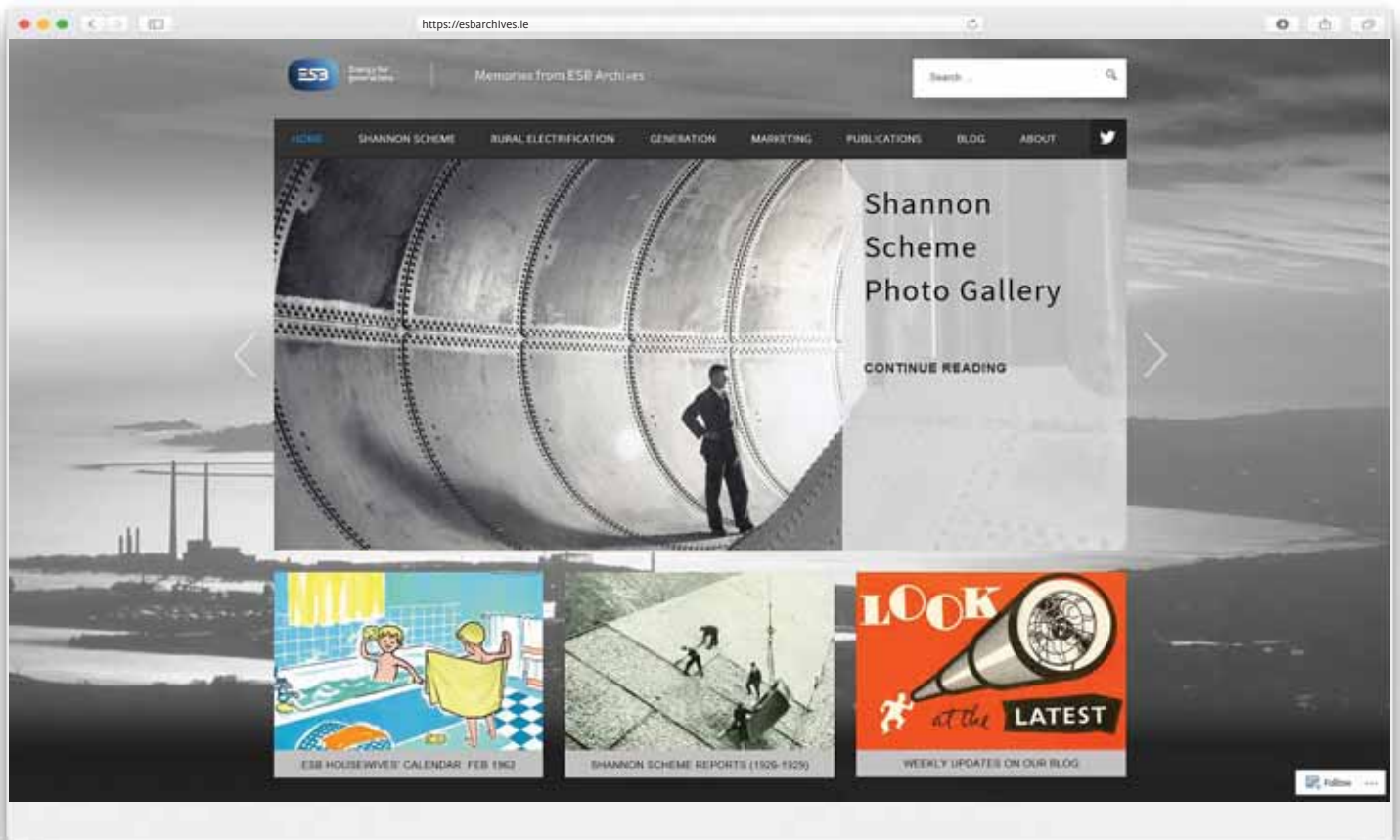
- Searching the online catalogue, and nominating a record as the feature for one episode
- Making an appointment to look at material from the BSUCA in person by visiting the reading room, and submitting their thoughts (in written or audio form) about the item
- Recording their own version of our theme tune and sending it to us to use in a future episode (no one has yet done this…)

To find out more about the collection and project see:

http://blogs.kent.ac.uk/standupcomedyarchive
https://www.acast.com/ahistoryofcomedypodcast

Twitter:
www.twitter.com/histcompod
www.twitter.com/unikentstandup

Elspeth Millar
University of Kent

ESB Archives website homepage. Photograph copyright ESB.

Digital ambition re-energises the archive of the Electricity Supply Board, Ireland

The Electricity Supply Board (ESB) Archives was founded in 1991 to collect, preserve and provide access to the records of Ireland's leading energy provider. In 2015, the archive was redeveloped with a digital focus aimed at exploiting the content of its rich collections, for the benefit of ESB staff and the Irish public.
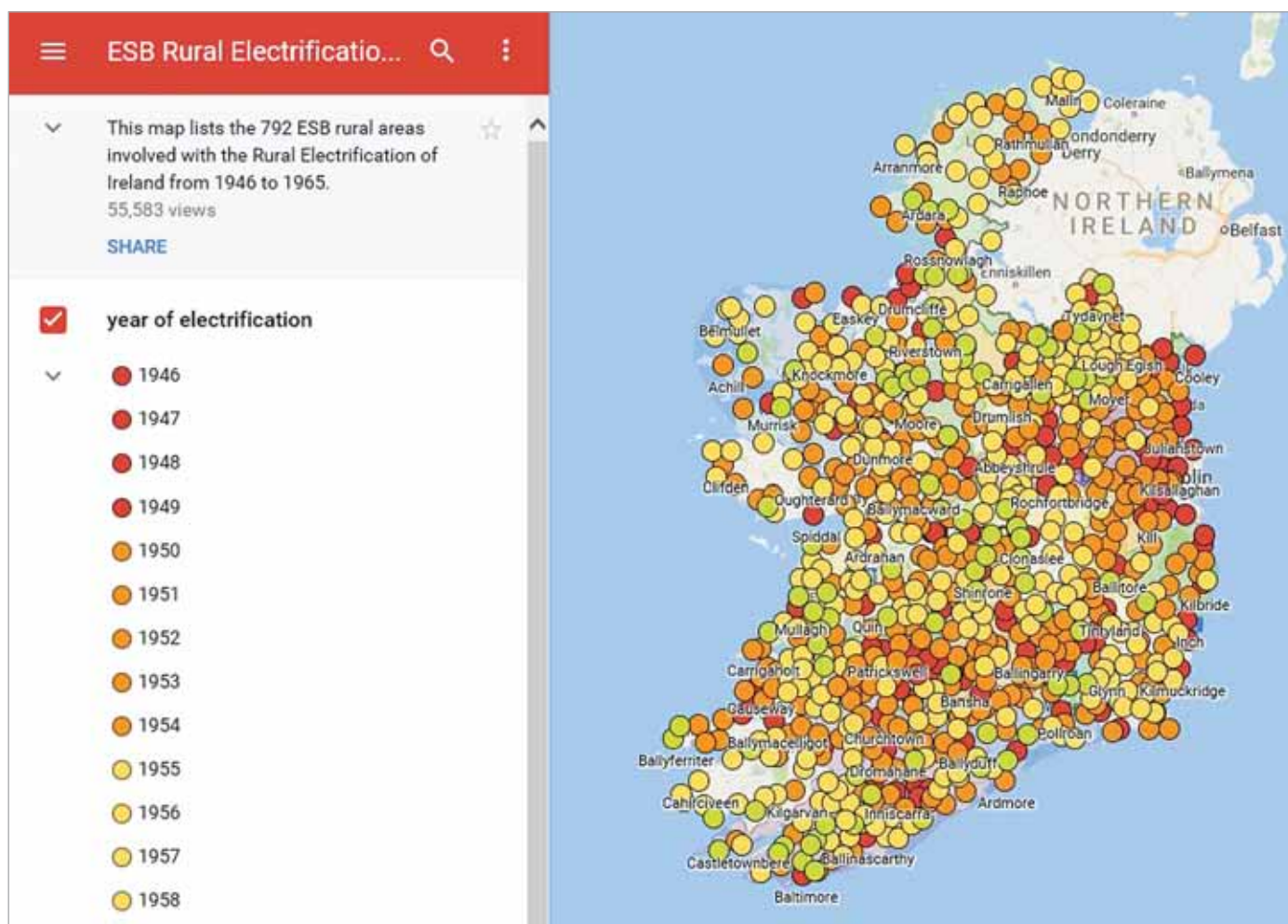
ESB's tagline 'Energy for generations' invokes both the past and future of the company, and 'going digital' offers the perfect opportunity to open up its archive collections in new and innovative ways. It was important to the ESB

Archives team from the beginning that the project utilised as much open-source technology as possible. This gave full editorial and design control to the archives staff, and avoided platform constraints. Moreover, it increased efficiency, allowing the archive to support the everyday business functions of ESB by quickly providing both evidential and entertaining content.

In April 2016, ESB archives launched its new website using free website and blog creator WordPress, and an associated Twitter account:

https://esbarchives.ie
Twitter: @esbarchives

By the end of 2016, we had over 300,000 views from all over the world. The website was designed to share exciting audio-visual content and high-level context relating to our most popular collections. We update our home page and blog regularly with targeted and topical
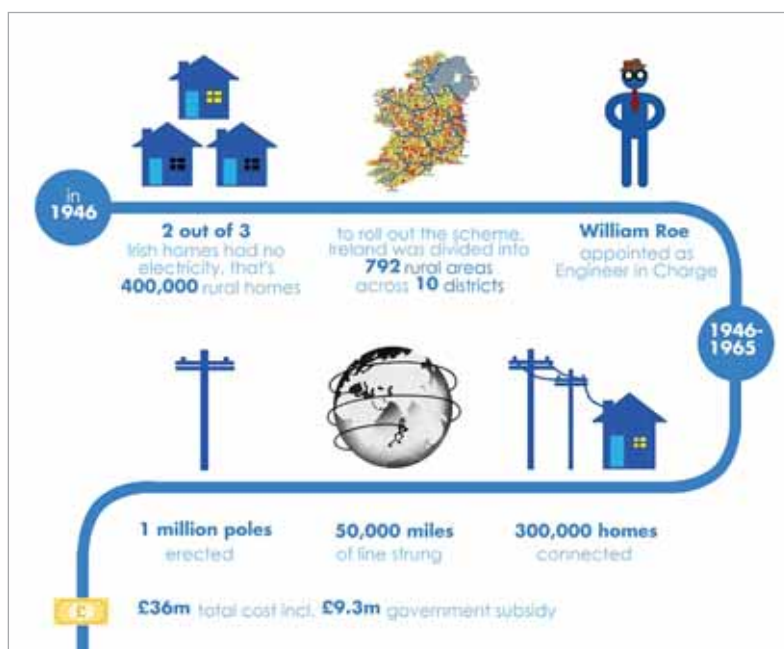
Rural Electrification Scheme map. Photograph copyright ESB.

material which has been recently uncovered or digitised. Our 'publications' tab allows users to access a variety of resources, such as every ESB annual report from 1928. Our website keeps our archive current, and has inspired members of the public and retired ESB staff to share their ESB story.

Beyond just presenting our newly-digitised content to the public online, we have used additional open-source resources to bring this content to life. One of our key collections relates to the Rural Electrification Scheme, when electricity was rolled out to 400,000 homes across 792 rural areas of Ireland (1946-65). We compiled a simple Excel spreadsheet relating to each of these areas, detailing not only where each was located and when work began and finished, but also the number of poles erected, the length of line strung, and the number of consumers connected. We uploaded this spreadsheet to Google Maps to create an interactive map of Ireland, where our users can search for their local area, and click through to read further details on our website.

Through the use of Piktochart, we have created a series of infographics which help break down some

Infographic extract. Photograph copyright ESB.

of the more complex facts and figures arising from our collections – such as the volume of materials used in the construction of the hydroelectric Shannon Scheme (1926-29), or how the first phrase of rural electrification was organised. These infographics are created with our user's needs in mind, and help

66 *Digital has offered us new avenues of in-reach, allowing us to further educate staff about the value of their archive, and moreover, to build on their sense of pride in their company's heritage.* 99

Rural electrification. Photograph copyright ESB.

us create creative content for our website to engage all user types, from academics, to casual users, to school children. We aim to take this inclusive approach further by launching educational packs for schools online in the coming year.

We have shone our digital focus internally too. All blog posts and tweets are promoted on our internal staff networking site Yammer, making sure we reach staff who do not monitor our external social media sites. We recently opened our archive of company newspapers and magazines to staff, using an 'out-of-the-box' SharePoint package. There are over 1,000 issues available, dating from 1947 to the present. Each issue is keyword searchable, and staff can easily search or browse by title, decade, year and month. Issues can also be downloaded for personal reference. While this SharePoint option is not perfect – we would ideally like more customised search functions and displays – it has given us the opportunity to allow ESB's 7,000 staff to access these records quickly and easily from wherever they are based in Ireland and beyond. To support this, we have used Piktochart to create a handy 'how-to' search guide for staff. In this way,

digital has offered us new avenues of in-reach, allowing us to further educate staff about the value of their archive, and, moreover, to build on their sense of pride in their company's heritage.

Our digital ambition over the past two years has earned us two ESB awards for 'Innovation in Corporate Social Responsibility' and 'Going the Extra Mile', and brought us to the finals of the Irish Blog Awards, the Eir Spider Awards, and the Digital Media Awards. Aside from the activities outlined above, we have digitised over 3,000 items, and are busy working behind the scenes to catalogue and prepare a selection of material for future online release. We are determined to utilise our archive to its full potential, which means embracing the past and the future in equal measure.

Kirsten Mulrennan
ESB Archives

AtoM on Raspberry Pi. Photograph copyright Louise Pichel.

AtoM archival description software on a Raspberry Pi

As an archivist who has always had an interest in anything to do with computers, the emergence of the Raspberry Pi in 2012 was very exciting. The credit card-sized computer was developed with the intention of providing an affordable platform for children to learn to code, but it also offers opportunities for anyone willing to experiment.

I already had experience of installing AtoM in a Windows environment, so I was curious as to how easy it would be to install and run on a Raspberry Pi. Given how small, cheap and portable the Pi is, I could see a few possibilities for using them in an archival context, perhaps running 'plug and play' installations of AtoM, or as a low-risk testing environment for those considering larger scale rollouts of the software.

I was drawn to the fact that the Raspberry Pi runs on Linux, an open source alternative to Windows that is free to install. Unlike Windows, which uses graphical interfaces for installing software and other administrator level operations, Linux encourages you to use text-based commands in a window, known as a 'terminal', to get things done. While this can feel a little daunting, the lack of graphical 'bloat' results in a satisfyingly speedy experience when completing tasks such as installing software, even on a machine as relatively underpowered as the Raspberry Pi. Linux also offers the user freedom to pick and choose elements of software to install,

making it a sensible choice for older or generally underpowered systems.

I also found this modular approach useful for understanding exactly what was installed and how elements worked together. This was helpful when trying to troubleshoot problems, as I could narrow down the cause of any issues. With a little extra knowledge, using commands in the terminal also makes it easier to run small bits of code to automate routine tasks such as file conversions or re-naming, which I can see being useful in a digital asset management context.

I successfully installed AtoM on both a Raspberry Pi 2 and 3, but – unsurprisingly – both ran slowly when compared to installation on a standard PC. However, the rate of development of Raspberry Pi computers suggests that it will not be too long before the overall experience is speedier, offering potential for use as a portable, low-cost collections management solution for demonstration purposes.

For more information on AtoM and the Raspberry Pi see:

www.accesstomemory.org/en
www.raspberrypi.org

Louise Pichel

The new Archives Hub: making pipe dreams a reality

I am delighted to say that the new Archives Hub website went live in December 2016:

https://archiveshub.jisc.ac.uk

However, is is much more than simply a new website – it is the result of a long process to develop a whole new approach to the ingest, processing and display of archive descriptions.
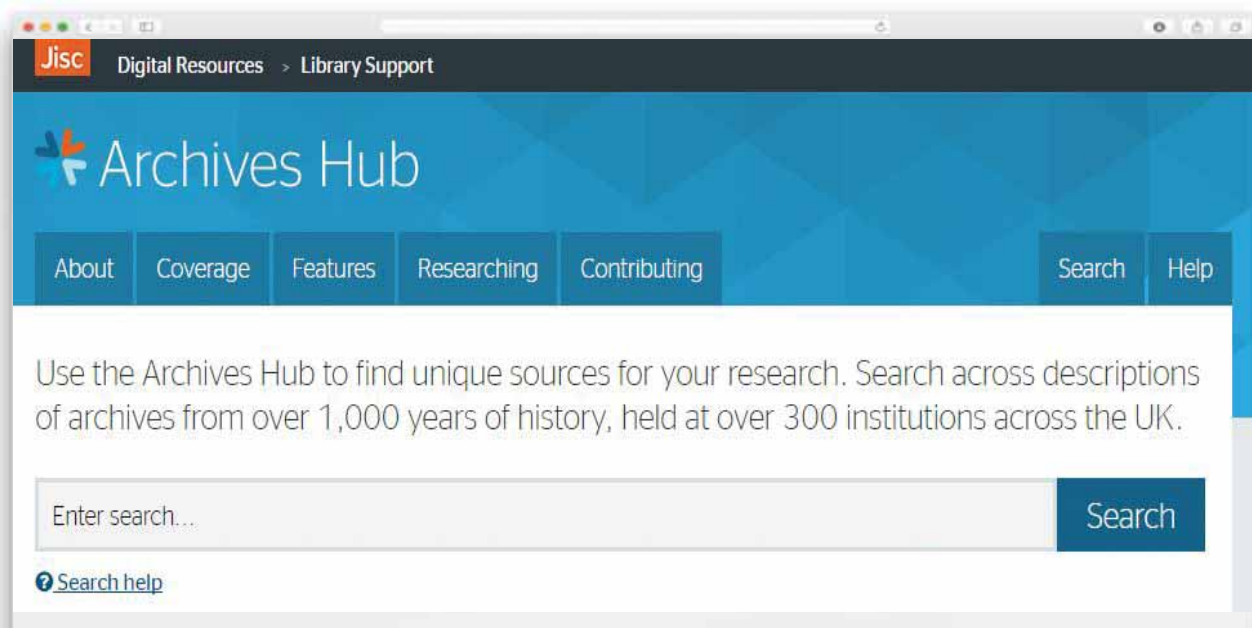
The Archives Hub is run by Jisc, a not-for-profit provider of services for research and education, and it is important to demonstrate that the archives aggregation service we provide is highly valuable and well-used, and worth its place within the Jisc portfolio. We needed to implement a system that was efficient and flexible, to harness the latest technology, and be able to meet changing needs. We wanted to create a service where the data processing – which is at the heart of an aggregation service – was more rigorous, more automated, and more consistent. This would allow us to develop new functionality for users, and also to free up more of our time to concentrate on adding value through innovation, in order to keep the service current and relevant. Our very ambitious goal was to improve the consistency of the data we already have, by which I mean reducing variation and errors – where appropriate – as well as applying the same principles to new data.

Our approach to data normalisation (or standardisation) was to create 'pipelines', which are repeatable transformation processes, to assess and potentially change the markup or content of the description. Pipelines allow for version control, because revisions of the descriptions go through exactly the same process every time (or they are knowingly changed for a defined purpose). This allows for a very flexible approach, because we can customise the pipelines over time.

The pipelines are XSLT files that transform parts of a description. For example, ensuring the repository name is consistent, removing empty tags, fixing incorrect web links, adding ISO language codes, adding agreed information about access conditions, checking for non-unique references. Many of these things can be flagged and corrected. If they cannot be dealt with in an automated way, we liaise with the repository concerned, and decide upon the best way to fix the issues. We have a default pipeline that works for many of our contributors, and is the standard processing for all those using our Encoded Archival Description (EAD) editor to catalogue.

However, there is never a one-size-fits-all approach. Even for those archives using the same archival management system, the exports vary substantially. For example, for descriptions created in Axiell's Calm system, some repositories use the Calm reference, and others use the

The Archives Hub website

alternative reference that they create themselves. We can implement the appropriate pipeline to ensure that we display the right reference for them. The implementation of pipelines is a highly-skilled and complex process, and it has necessitated a great deal of clarity about what the Hub requirements are for data – what we can accept, what we need to change, and what we reject. We have had many, many conversations about almost all aspects of description, trying to find the right middle way between practicality and the quality that consistency brings to the data.

We wanted to go further with this idea of automated processing. Not only would we introduce a customisable processing framework, but we would also give contributors the ability to add and delete descriptions themselves. We had always had the ambition to enable contributors to administer their own data, but our old system was not suitable for this approach. With the pipelines set up, we can give contributors access to an interface where they can view their descriptions, upload, and delete them. We are confident that most descriptions will be accepted, and will go live almost instantly. However, if there is an error in the processing, the contributor can either choose to analyse and amend the description themselves, or they can ask the Hub team to do it. We simply check every week to see whether there are descriptions that have not been accepted by the system, and then take appropriate action.

The work on the data processing is largely unseen, although many of our contributors can attest to the work that the team has put into this because we have cajoled many of them during the process of cleaning the old data, when automated fixing has not been possible, and we have had to resort to manual fixing. The system itself –

which uses a technology called Elastic – was developed by Knowledge Integration, and has some very important advantages. One of these is the architecture. The data model and indexing structure has a significant impact on system performance and response times. In the new system, each EAD record is broken down into many separate sections, which are each indexed separately, so that the specific section in which there is a search match can be delivered immediately to the user. So, if a single lower level component matches a search, it is delivered immediately, and other components are delivered as they are requested. This means most searches take less than one second.

The website application, developed by Gooii, can be seen here:

https://archiveshub.jisc.ac.uk

This is the visible end of a great deal of hard work. It reflects a new approach whereby archive descriptions are not alone at the top of the tree, but they sit alongside themed collection descriptions and repository descriptions. Whilst the new Records in Context (RiC) standard was not published during development, our approach is similar to this type of model. It means, for example, that we can implement a planned 'people and organisations' section as an extension of the current site, with no change in the underlying architecture. The site has a fresh new look and feel, with more emphasis on filtering of results, a 'type ahead' function that suggests possible matches, an expandable hit list, a new tree structure with a 'mini map' to give an immediate indication of overall extent, and a 'search within' that highlights search terms throughout a large multi-level description. The site is a responsive design that is mobile friendly, which is just as well, with something like

66 The site is a responsive design that is mobile friendly, which is just as well, with something like one third of our use now coming from mobiles and tablets. 99

one third of our use now coming from mobiles and tablets. We have recently implemented an interactive map to show where collections are held, and we are working on our name authority descriptions, which will open the door to more opportunities for greater context and connectivity.
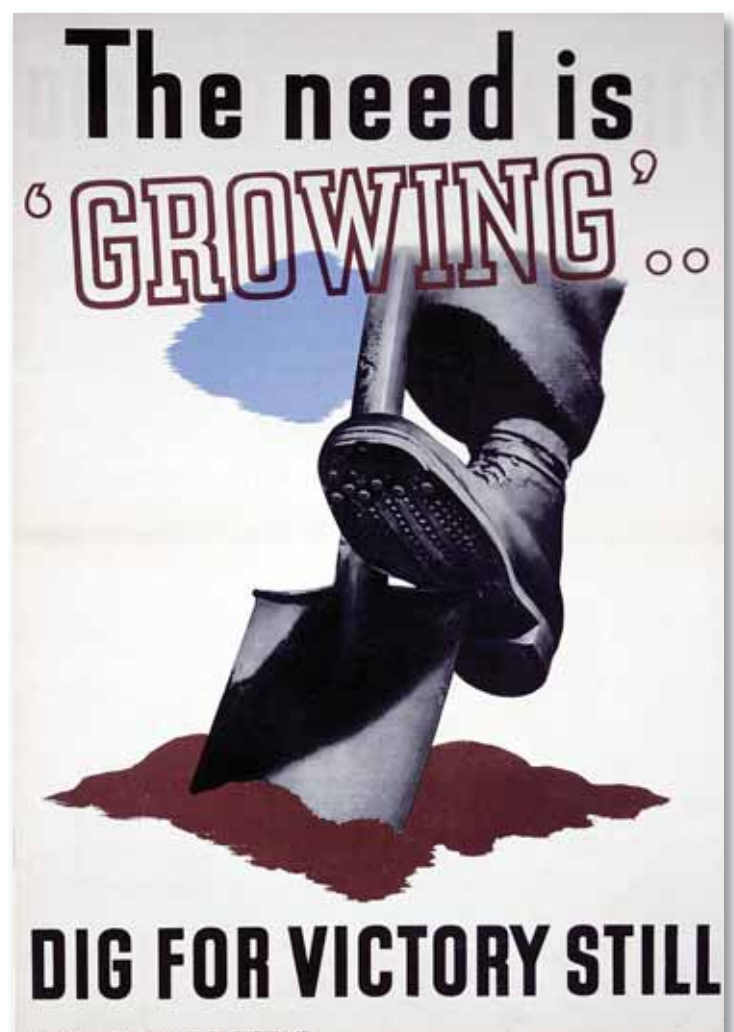
Jane Stevenson
The Archives Hub

Digital technologies for a digital National Archives

The National Archives (TNA) collects and preserves the public record for the benefit of the nation, safeguarding and making it available as an evidence base for research. Since the turn of the millennium, the rise of digital collections has seen one of the world's oldest and largest archives become one of the world's largest repositories of public data too.

Each digital accession brings a greater diversity of formats. This year we have received high-definition video, executable code, and a virtual reality model. Unlike paper records, which are easily described and stored following archival practice developed in the 1920s, the scale, complexity and diversity of digital records cannot be accommodated using traditional approaches. We must invent new practice with supporting tools and technology. Here, we explore some challenges for digital archives, highlighting new technical approaches.

For paper files, the first step is the selection of records worthy of permanent preservation, followed by a review for sensitive content – a painstaking process involving reading every page and making expert judgements. This model is unsustainable in the digital world, which is typified by huge volumes and little structure (sometimes referred to as the 'digital heap'). Imagine reviewing the entire contents of a government email system, or

Digging into the digital heap. Ministry of Information and Central Office of Information: posters and publications, TNA. Photograph copyright TNA.

watching a year's worth of Supreme Court proceedings on video! We are trialling machine-learning approaches to sort the heap into categories, and help identify sensitive or valuable content. To fill the gap until this research delivers a usable tool, we have experimented with commercial data mining and e-discovery software that can make sense of unstructured data. The aim of this is not to replace human judgement with machines,
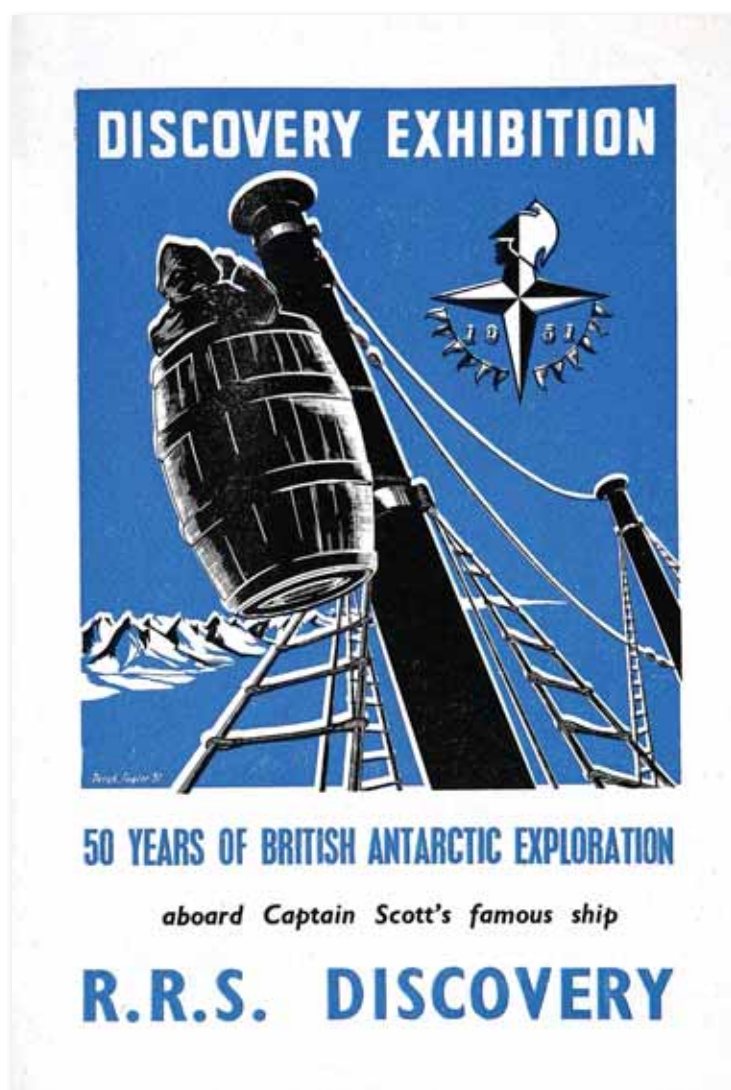
but to assist skilled reviewers and use their limited time wisely.

Collecting the digital record is complex. Government decision-making is increasingly supported by algorithms, and it is not clear how we can secure a transparent record of this process. Is the machine learning model the record? Or should the data that trained the model be archived too? How will we preserve systems that learn and evolve?

Preservation has its challenges too – the first of which is storing all that data. Not only have numbers of documents exploded since the dawn of the digital age, but the advent of rich formats such as high-definition video has seen storage requirements grow exponentially. Just as paper documents must be held in the right conditions, so too digital files must be kept moving to avoid 'bit rot'. At scale, this operation is computationally expensive, so we are exploring risk-aware, statistical methods for optimising storage. Our PRONOM file format registry is used across the world, but cannot keep pace with the ever-expanding list of formats. We are developing automated pattern-matching approaches to create new signatures for file characterisation. Finally, there is the issue of trust. How can we guarantee that the document we archive today is the same one we release to the public in 20 or even 100 years' time? A new project will experiment with distributed ledger technology to develop methods for publishing cryptographic hashes of digital material as a promise to the future that the record has not been changed.

While digital records pose challenges for collection and preservation, they offer new opportunities for access. Once we recognise that the records are also data, we open up the potential to work at scale using computational techniques to trace patterns in the records, recombine them to tell new stories, or bring them to life through visualisation.

Our access projects include a data infrastructure to make UK legislation available as a resource for research. This work applied techniques such as natural language processing and structure-aware search. Our 'Traces through Time' project developed probabilistic techniques for linking the people who appear in our records, creating new routes for navigation across our collections and those of other archives. For the first time, the links we publish are not a curated assertion of authoritative truth, but rather an expression of confidence. Instead of trying to cleanse and standardise

A different kind of voyage of Discovery. Records of the Festival of Britain Office, TNA. Photograph copyright TNA.

data, our techniques quantify and work around the 'fuzziness' and uncertainty found in historical records. Now, instead of a catalogue of boxes of paper, connected only by their archival context, we can create a connected web of people, places, events, and ideas. We still need to explore methods for making this accessible to new and old audiences, using modern visualisation and data mining techniques to deliver new insights.

This portfolio of active data science projects has been developed in collaboration with records creators, academic partners, and the archives sector. The tools and techniques we're experimenting with are a mixture of 'tried and tested' approaches and cutting-edge research challenges. We will need a whole range of approaches to address the problems of collecting the right data, safeguarding it appropriately, and enabling citizens to work with our public digital archives.

Mark Bell and Sonia Ranade
The National Archives (UK)

DON'T RISK IT!

Know Your Records

Phase two Training:
Risk, Internal Advocacy and the Records Manager

First session: Section for Records Management and Information Governance AGM, London, 24 July

Further Sessions for Members Elsewhere in the UK & Republic of Ireland to be announced soon.

Contact: ARA's Head of Public Affairs if you'd Like to Host a Session in your Nation/Region.

Are Your Organisation's records safe? Are you RISKING IT?

ARCHI'VE EXPLORED

Explore Your Archive 2017: Launch Week, 18-26 November

Order Campaign Badges and Pencils now from the ARA office in Taunton.
Social Media Campaign Details and Revamped Website to Follow Soon.

ARCHI'VE PRESERVED AND CONSERVED

Mini-campaign for Conservators: ArchI've Preserved and Conserved
Launch Date: 27 June 2017 at Houses of Parliament, London

Free attendance for ARA Members Interested in Conservation
Contact: ARA Head of Public Affairs for more details

ARA

Archives & Records
Association
UK & Ireland

CONFERENCE

2017

HILTON MANCHESTER DEANSGATE
30TH AUGUST - 1ST SEPTEMBER 2017