Document Distribution Note
Release to Members: Immediate
Release to Public: 18/04/2014

Archiving the Digital: Current Efforts to Preserve Digital Design Records

Royal Institute of British Architects, London 17-18th October 2013

About the event

The Royal Institute of British Architects invited WK to participate in a 2 day workshop on their premises in London to raise awareness and survey the state of the art for digital preservation in architecture.

These notes are intended to provide an informal briefing for members of the DPC not able to attend in person. They only represent the sessions that WK was able to attend was able to make notes. For an authoritative and comprehensive report, readers are encouraged to contact the organisers or speakers directly. A number of DPC members were present in their own right – Alex Ball (DCC), Mike Pearson and Ifor Ap Daffydd (NLW), Emily Nimmo (RCAHMS), Paul Conway (KCL), Matthew Addis (personal member), Ruggero Lancia (HATII), Shona Robertson (Parliamentary Archives) Keith May (English Heritage) and others.

Introduction

Kurt Helfrich (RIBA) – RIBA and Digital Records

It's been recognised for a while that architectural archives can't just collect paper and/ or models. CAD has been used by architects since at least 1961. AutoCAD was released in 1982 and agencies like Aberrant Architects aim to be entirely digital, and mostly on remote platforms of one form or another. Archives are still hybrid in nature, but RIBA's challenge recently have been to make sense of how to receive the archives – King's Cross, Pater Noster Square and the British Library are all large and high profile projects which RIBA has experience with recently – but in each case the digital archives have been problematic for different reasons. Experience shows that it is more useful to collect through the lifetime of a project than to wait to the end.

Session 1:

Maygene Daniels (National Gallery of Art, Washington DC) Archiving Design Records in the OAIS Model

Experience of the architectural archives at the National Gallery of Art is perhaps incomplete but they are of general use. Everything is a state of the art but these case studies show what can be done. OAIS provides a conceptual over-view of what is required for long-term preservation. The repository of NGA at a simple level is a series of files but it's the services that support it and the ingest processes that manage the flow of information into the repository matter. Rules like sticking to specific file formats and avoiding proprietary software matter. Naming conventions and evidential metadata matter a great deal. Quality assurance and vigilance are constantly necessary so at NGA all work is cross-checked with two staff working in pairs. On completion of ingest the archive take

responsibility for the ownership of the archive. Multiple copies of files (and multiple versions) can be ingested for different functions – access, preservation and so forth. Not everything will go into an archive, but a records centre has been created for files that we want to keep but can't or don't want to promise. The context of NGA's work on digital preservation has been a large architectural programme of its own meaning there is a rich seam of data over which they have more control than one might have routinely. A system called E-Builder provides a basic information and records management functions for this project and PDF/A has been chosen as a format for preserving records that come out of it. Graphic documents are largely created with CAD but E-builder requires that any documents it stores or transfers are PDFs – so all of the details and functionality one might get with CAD are already lost within the 'live' records management system. The architects were delighted that they only had to provide digital copies of files. The record is huge and there is still need of appraisal and better tools for access – but it points to what is needed in the future.

Alex Ball (DCC) Preserving Computer-Aided Design

CAD is used across a range of disciplines – engineering, archaeology and so forth. Digital records are generally better because – especially in engineering – they generally are inputs to other systems and these systems will not be able to deal with paper records. There are different ways of approaching the problem of preservation depending on the use case but a lot of it boils down to what formats we want to use the records for. STEP, for example, is very thorough but it's also very hard work and few vendors will support it. DXF is less impressive but widely used. There are also other ways to record change histories and rationales which are not exactly preserving CAD but which are incredibly useful as documentation and may well be useful in their own right. Again for uses like procedural modelling or visualisation it might be more useful to preserve the procedural logic rather than the output, while libraries of components might be more useful for highly complex files. Preservation systems may need to replicate these kinds of functions or develop work-rounds or documentation for them. Migration/ emulation / normalisation each approach has its own pros and cons. Validation approaches exist to capture the success or otherwise of migration – such as the point-cloud checking system recommended by LOTAR. Supporting documentation is also vital and there are standards here depending on what is being described. Need to impress on regulators and vendors the value and importance of long term preservation.

Ines Zalduendo (Harvard University, USA) Collecting and archiving design materials in an evolving born digital landscape

Architectural practice works in a global and inter-disciplinary environment. Expanded practices and is also associated with a rapid growth of records and a step change in the risk of loss. Archivists are therefore feeling overwhelmed – and even architects who use CAD feel like newcomers to a foreign land. Collaboration across systems is also like collaboration across time. Harvard university repository is helping make practical guidance for preservation. Currently it does not receive CAD formats so these are not accepted except in really narrow terms: however it does accept PDFs and Text formats as well as images so there is some scope for inclusion of things like PDF. This narrows considerably the value of the items saved but it is progress on simply losing the files. A new generation of repository is imminent and planning is underway for how it might support CAD better. Pragmatic decisions characterise the approach. A range of projects are worth consideration:

- Schiff Foundation report which surveyed the long-term preservation of CAD
- AHDS CAD Preservation Manual by Keith Westcott
- ISO Standard 82045 2005, a standard specifically for architecture, engineering and related to preserve drawings including CAD formats and
- FAÇADE (2006-9) at MIT led by Mackenzie Smith
- NARA Report on 3d content categories 2008
- Geomapp (NDIPP) project 2007-11 report studied long term preservation of geospatial data
- LoC study on formats for the long term. Not specifically about CAD but the principles are the same.
- DURAark – EC funded project which is current. Includes a range of institutes in preservation of 3d data. It includes LTU and University of Bonn…

David Peycere (Cite d'architecture et du Patrimoine, Paris) Gau:di programme and what is needed next

The Gau:di project was an EU project which brought together a series of European initiatives (including RIBA) to look at emerging practice in architecture including in archiving. The project produced a series of recommendations for firms about their practice. The project finished in around 2009 after 2 projects over 6 years. The work was relatively simple at the time but was in its own way ground-breaking. The project team was then dispersed but the expertise and experience has not been wasted. For example the Patrimoine du Paris draws on this expertise to facilitate its modern collecting practices. It has been deployed recently through the collection of relatively simple 2d CAD based drawings which are conventional in style, in Arc+ format and AutoCAD. This has caused them to address a range of issues which are familiar to others in the room – should we migrate or maintain the original, how should we alter the cataloguing of files, how to provide access online, and to what should access be provided …. Inventories of files will be available online but access to the file is only possible on location. Processes and automation have been tested, for example to open and validate each file. This should have a significant impact on working time. Need to engage earlier in their career – as digital records.

Discussion (Chaired by Nathaniel Parks, Art Institute of Chicago)

- Need to engage with creators earlier. How do we do that? Harvard planning office is also the commissioning agency so there is a possibility to insert standards at the point of commissioning rather than at the end. But if that's not the case then we depend on the practices and the skills of the staff. Architects are getting better at their work and insisting on standards and this is probably the way forward. This is analogous to the pressures being brought to bear on researchers in higher education and the way that early career researchers are being trained.
- How long will things accumulate before we have actual solutions – such as in Harvard. AutoCad files cause real problems and the reality is that the repository is not able to handle them. Descriptive metadata is available – though 2nd best.

- RIBA has written guidelines for the archivists working with practitioners which includes help with digital. Strong themes are selection and appraisal. This is a traditional archival skill and remains important – but it creates tensions around liabilities. Legal complexity.
- Use case? How did you map against users?

Session 2

Martien de Vletter (Canadian Centre for Architecture) Born Digital Archives and the Archaeology of the Digital

Telling how many people are working on the problem of digital preservation around the world simultaneously. The Archaeology of the Digital is a big project for CCA, investigating how digital revolution has change architectural design and thinking in the last 30 years. A series of projects will be highlighted as if through oral history which will be presented as an exhibition and online, but with archives compiled for the long term. 25 projects have been selected by a curator to highlight this, 4 in the first set, six in the second set and so on. It's surprisingly paper based. The exhibition will travel. The curators sit and talk with the architects first and visit them in their offices and discuss the exhibition. It's noticeable how projects involve so many different people – designers, students, technicians and architects. In some cases there is almost nothing left – which is why it's called archaeology. The fragments of projects are unpredictable and it's clear that not all the data is still available and even when it does survive there's no guarantee that data can be extracted. That is a creative process and the research archive will be assembled for the future even if it can't be accessed just now. Contracts and deeds of gifts for digital archives seem a lot more complicated than for the paper ones. Negotiating with architects for their files have been interesting and a bit random. The whole process has been sketched out and there's a lot of learning on the way: but it's like driving with a map on your lap…

Stuart Jeffrey (GSA) Emerging Challenges in Heritage 3D Modelling

Significant challenge in this space is that the word archive is sometimes confused with backing-up: they are not the same thing but many people think they are doing archiving. Other issue of data bloat – preservation is not cheap and volumes create problems but the impression is that storage is cheap. The cost of preservation is fundamentally about labour costs – and this needs to be represented.

Stuart Chalmers (Building Research Establishment) Emerging Challenges: BIM

Building information modelling is soon to be a mandatory feature of all public projects in the UK from 2015. It's like the 3rd generation of CAD: instead of thinking about design of just the building we consider a range of additional information components about building performance. BIM has been recommended and developed across a range of building and construction sector standards. There is a building data dictionary and ways to exchange information which are becoming mandatory. Would be really interesting to map BIM and preservation metadata in the emerging and unclear areas of BIM level 3 which is actually about managing the metadata associated with BIM. Emerging challenges for BIM include looking at how we make it open data. Funding has been given

for the BIM gateway project to try this, using classification systems open and available in as many formats as usable. www.cpic.org.uk is the result.

Cristina Durand (Amsterdam City Archives) Understanding the Use of Collections to Prioritise Tasks

Session 4: Business Models

Matthew Collins (University of York) Caroline Checkley-Scott (University of Manchester) and Stephen Milner (Manchester University) What do you call a conservator standing between a scientist and a scholar? A translator

We all have less money and we all have to demonstrate impact in new ways. We need to concentrate on the material nature of collections. Academic researchers are mostly small business men looking to raise funds for research. Lots of promise of new techniques and technologies but too little integration in practice: the promise is only ever realised in good times but the researchers also have a habit of moving on and then forgetting what has been achieved. AHRC recognises that there is a significant potential for collaboration between the science and the humanities. Multiple competing demands and hard to make a case for culture let alone conservation. Cost intensive work of conservation is not attractive and is not easily made subject to market forces. Waste from the conservation process is more interesting than you might think – so instead of the scientists simply coming along and doing science, the waste products of conservation make it possible for conservators to 'do more'. Much of the analysis is done electronically so can be distributed.

Christina Duffy (BL) From meteorites to Magna Carta a scientists journey into collection care

Physics and art history are an unusual combination but necessary for each other - but most entrants into the profession come from an arts background. Three types of imaging at the BL – multispectral imaging which helps identify the compounds which have been used in ink, while it also allows you to spot changes in documents where elements have been erased. Reflectance transformation imaging allows you to automatically re-light an image therefore allowing the 3d topography of apparently flat documents to be immediately visible. Digital microscopy allows really high quality measurements of surfaces, the interaction of fibre and ink, the cracking of dried inks and also in predicting where wear and tear are likely. You can see individual tool marks in book binding which in turn helps interpret how a document was constructed. The images are very popular with the public and they are useful in themselves.

About this document

Version 1	Written at conference	14/10/2013	WK
Version 2	Distributed		DPC members