Preserving Computer-Aided Design
Digital Preservation Coalition Report

Alex Ball
17 October 2013

Abstract

The Digital Preservation Coalition Technology Watch report ‘Preserving Computer-Aided Design’ explains the particular challenges of working with CAD models, and sets out some of the standards and techniques that are used to help preserve them. In this talk, Alex Ball relates these challenges and practices to particular use cases of reference, rationale and reuse.

Contents

1. Introduction ................................................................. 1
2. Preserving for reference ................................................... 3
3. Preserving for rationale .................................................... 5
4. Preserving for reuse ....................................................... 7
5. Preservation approaches .................................................. 8
6. Final thoughts ............................................................. 10

1 Introduction

I work for the Digital Curation Centre. We are a collaboration between the Universities of Edinburgh, Bath and Glasgow . . . funded as a centre of expertise helping the UK Higher Education sector. While these days we concentrate mainly on research data management practice, we remain interested in the full range of digital curation activities including digital preservation.

I was first introduced to the challenges of preserving CAD in the KIM Project . . . EPSRC Grand Challenge . . . ESRC funding . . . 11 different universities . . . aerospace, defence and construction. Looked at challenges of firms providing products as services . . . information organisation, selection policies for data, and finding ways of learning design lessons from in-service data and knowledge.
Subsequently I co-wrote a report on data curation in engineering disciplines as part of the DCC SCARP (Sharing, Curation, Reuse and Preservation) project . . . involved in two Jisc-funded projects looking at research data management in the University of Bath’s Mechanical Engineering Department.

¶ So when the DPC asked me to write a Technology Watch report on Preserving CAD it was an honour . . . opportunity to look at CAD uses outside the mechanical and aerospace engineering context where I’d done most work.

When you start looking, you find CAD being used in all sorts of different places.¶

- Originally an electronic pencil to help create 2D design drawings of products . . .
- . . . and Floor/site plans of buildings.
- Archaeologists found CAD models convenient for Archaeological site records
- But they really came into their own when they started being used for 3D product models . . .
- . . . and 3D architectural models
- If you’ve ever seen an archaeology documentary on Channel 4 you will have seen CAD used to create 3D impressions/reconstructions of ruined buildings
- CAD is also used in the entertainment industry for creating Virtual worlds for computer games . . .
- . . . and 3D animations for films.

§ And when you come back to them later, you tend to be interested in doing one of three things (these are the 3 ‘R’s proposed by NIST initiative ‘Sustaining Engineering Informatics’): looking up what they say (Reference), asking why the model turned out the way it did (Rationale), or adapting it for a new purpose (Reuse). ¹

¶ Why should we be interested in preserving these things? Well hospitals, schools, skyscrapers, and so on . . . all built to last . . . with care can last centuries. But the CAD systems they are designed on go obsolete very quickly . . . 10 years.

So far we’ve been managing with paper plans and physical models, but as things get more complicated, as shapes become more organic, that is going to change. The engineering sector has already gone through this. Aircraft carriers and locomotives may not be in active service for centuries, but they still last a long time. We’re already at the point where printouts are not adequately expressive, plus

- you can’t redesign from them,
- you can’t manufacture from them because the machining robots expect 3D CAD models, and
- printouts tell you very little about why something was designed a particular way.
Figure 1: CAD models are used by more than just CAD systems

That kind of information is locked away in things like the design history, which only the CAD system can tell you. And the CAD system won’t be around in ten years.

Actually, it’s not just the CAD system we need to worry about because CAD models are used as input for a variety of other systems (Figure 1) . . . I’ve already mentioned manufacturing systems; also FEA, GISes and CGI renderers and ray tracers. With all these systems relying on CAD models, it is important to consider how we can preserve them.

I will say right at the outset that there are no easy answers and no complete solutions. What we have to do is identify the aspects of a model we are most interested in preserving and concentrate on that.

2 Preserving for reference

There might be situations where you are only interested in the final form of the model, in what we call the shape data. Here are a few examples: ¶

Companies like BAE Systems, Rolls Royce and Balfour Beatty are now providing their products as services, so it’s in their interests to make them easy to maintain. If their maintenance crews can tell them exactly where and what the issues are, they can improve future designs.

It’s standard practice to show a customer an impression of a design for approval. The 2D ones we’re used to are good, but you can often give a better impression in 3D.

On big engineering projects you often find different companies working on different bits. The lead contractor has to make sure all the bits join up, but the companies don’t

---

want to give each other full access to their own designs in case they steal them, so they just share the shape data.

If all you need is the shape data, then probably the safest thing to do is export the CAD model to a lightweight visualisation format. These tend to be very simple, so there’s little to go wrong, and also compact so they don’t take up much space. X3D is an ISO standard, while U3D is an ECMA standard; these are on the safer end of the spectrum in terms of long term support. The others [HSF, DWF, DWFX, 3D XML] are examples of published proprietary specifications, which aren’t quite as secure.

§ When products are in service a long time, there’s always a danger of running out of spare parts, so it’s important to be able to make some more. For this you need more shape data plus Product Manufacturing Information, like the geometric dimensions and tolerances. There are lightweight formats that can encode that as well. JT is an ISO standard commonly used in aerospace and the automotive industry. Less common but still interesting is PRC, which is used in 3D PDFs and is almost an ISO standard. IGES was the first CAD exchange standard ever devised. It has now been withdrawn but it is widely supported, though systems do seem to pick and choose which parts of the specification they support.

I have a couple of real-life examples of this use case. ¶ Here (Figure 2) is a model of the rear case of a television. The manufacturer, Philips, would never need to edit it, but they might reuse it for a lot of different models. So JT was the perfect archival format for them. They could manufacture from the JT file, and use it for reference when designing a new set of innards. It’s presented here in a collaboration and archiving system developed for them by the European SHAMAN Project.

Another firm that will remain nameless needed to recover the designs for an old product of theirs. By that point the original CAD files were utterly unreadable, but it turned out they had converted them to IGES at some point to show them to a third party. The firm were able to salvage the models from the IGES files and manufacture spare parts from them.
3 Preserving for rationale

Sometimes the ‘what’ isn’t enough; you also need to know the ‘why’. ¶

If a plane crashes, the investigators have to determine if the plane was sabotaged, poorly maintained, or if there was a fundamental flaw in the design. If it was a flaw, they need to know how it was introduced. The same sort of thing happens with damaged buildings.

A rather more common scenario is tied to reuse. There was a company that was scaling up operations, so decided to set up more factories. Obeying the maxim of ‘if it ain’t broke, don’t fix it’, they simply transplanted the machine, cabling and pipework design from their first factory to their new ones. But the buildings themselves weren’t identical. The new ones didn’t have so many internal pillars for a start, so when everything was installed they found they had pipes taking odd routes to avoid pillars that weren’t there. The lesson there is if you are reusing a design, you need to know which requirement or constraint prompted a given decision.

From an academic perspective it can be interesting to know how a design was built up. Did an architect start from the outside and work in, or from the inside a work out?

1. Insert cylinder $l = 20 \ r = 1.0$

   Change cylinder $l = 40 \ r = 0.5$

2. Insert sprocket $r = 3.0$

3. Fit sprocket to cylinder

4. Group cylinder and sprocket

5. Scale group by $1.75 \times$

   . . .

Figure 3: Construction history modelling

¶ One source of evidence is the construction history of the CAD model (Figure 3). Construction history modelling is a bit like the ultimate undo button, in that it records all the steps that went into the design. You can also insert extra steps and see the effect it would have had.

¶ Perhaps even more useful is feature-based modelling. Features explicitly record the engineering significance of certain shapes in the design. So a curved blend between two surfaces might anticipate a manufacturing process or remove a concentration of stress; the feature semantics would tell you which the designer intended. Incidentally, this sort of thing can cause conflicts. Take this shape (Figure 4).

The designer might need this marked up as rib features applied to a base surface in order to calculate the optimum distribution. The manufacturing engineer, though, might need it marked up as cavities milled out of a thick surface. It’s not usually possible to have both in the same model.
The lightweight models I mentioned just now won’t contain that sort of information. You would either need the native file formats used by the CAD systems themselves, or a full exchange format such as DXF or STEP.

DXF has a published specification and is supposed to be able to express all the information encoded by AutoCAD’s native DWG format. By all accounts, though, recent specifications have not been detailed enough for full support to be implemented in third party tools.

STEP is the Standard for the Exchange of Product Model Data, or ISO 10303. It has been in continuous development for almost 30 years, and at 600 parts is the largest of all the ISO standards. There are various ways of expressing CAD data in STEP. Most CAD systems support AP 203 Edition 1, which is quite basic. But Edition 2 is more expressive, and AP 214 extends support in a different direction. And soon there will be AP 242 that supersedes both of them. Some vendors are already working on supporting it through the CAx Implementer Forum, which bodes well for the future.

STEP is also the basis of other international standards such as the Industry Foundation Classes for Building Information Models, though I think it’s fair to say that it is still early days for IFC.

I should point out that there are other ways of recording rational outside of CAD models. In the KIM Project, the team at Cambridge developed a tool called DRed, which built on the graphical issue-based information system. Here (Figure 5) is a DRed map showing how a particular design issue was solved. The map is hyperlinked so you can jump back and forth to other maps, reports, spreadsheets, and designs.
4 Preserving for reuse

When a customer’s requirements change, this usually means the design needs to be tweaked. In addition to the techniques I’ve already mentioned, CAD systems have some other capabilities that can make editing an existing design easier.

![Gear wheel and pulley](proc1.jpg)  
**Figure 6:** Parametric modelling was used here to adjust the design and position of a gear wheel and pulley to accommodate a shorter, thicker drive shaft.

Parametric modelling helps to automate the editing process (Figure 6). Aspects of the design are made to depend on parameters, and constraints are added so the computer knows how to adjust things should those parameters change.

![Procedural textures](proc2.jpg)  
**Figure 7:** Procedural modelling was used to generate the city (left) and to texture and displace the bricks (right).

Procedural modelling (Figure 7) is where algorithms are used not only to automate repetitive tasks but to give surfaces fine texture, generate realistic snowdrifts or extrapolate whole cityscapes from a single building...it is used extensively for constructing virtual worlds. It is now starting to be used in design work for creating intelligent parametric features, generating intricate structural patterns and finding optimal forms in architecture.

![CAD system examples](proc1.jpg)  
**Figure 8:** CAD systems provide mechanisms for including the same part many times in an assembly (Figure 8), so if you need to edit it, you only have to do it once. The same mechanism can be used to make a design modular, so parts can be swapped in and out as optional extras.

Some CAD systems have a library of standard parts like doors, windows, and staircases, preloaded with feature semantics and parametric properties. If these parts are hard-coded into the software, it can be very difficult to represent them reliably in a different system.
Alternatively, the parts might be saved as separate CAD files and included in an assembly by reference. This is more portable, but requires keeping track of multiple files and preserving the links between them . . . relative path names, IDs managed by a central register . . . In architecture and construction, a Building Information Model (BIM) system . . . in industrial engineering, this would be done by the Product Lifecycle Management (PLM) system.

These systems also manage relationships between CAD files and other data and models, and handle access control. Preservation systems might need to be able to do the same thing themselves.

This idea of reusing parts in assemblies is commonly supported, though as I mentioned the part libraries will differ between systems. Visualisation formats typically lose parametric information. STEP can support parametric models but as far as I know hasn’t standardised procedural modelling yet. Indeed, because such a variety of programming languages are used, there is hardly any interoperability in this area at the moment.

5 Preservation approaches

So, that’s a look at some of the types of information found in CAD models and how widely they are typically supported. So what are our preservation options?

Preserve the original CAD model

- Implies preserving software through emulation.

Pros preserves maximum information; easier to guarantee provenance.

Cons need to preserve expertise in the system; need an amenable software licence; hard to maintain integration with current systems.

Good for reference purposes, but not for future reuse.
Rolling format migrations

- Migration to newer format versions or new CAD systems.
  
  **Pros** models usable by current designers and software.
  
  **Cons** cost of validating each migration; incremental data loss/corruption.

  **Good for** models in active development/use, but not for long-term archiving.

Normalisation

- Migration to (a) STEP/IFC (b) a visualisation format.
  
  **Pros** only two migrations needed, so limited data loss/corruption; back-up in case a migration goes wrong.
  
  **Cons** cost of validating each migration.

  **Good for** long-term archiving and reuse.

Validation

- Creating a set of validation properties at the time of creation, as a benchmark for testing future interpretations of the model.

- Need to choose properties to test that reflect what is important to preserve. *LOTAR Project suggests using a cloud of points for marking the vertices, edges and surfaces of the geometry.*

  **Pros** provides confidence that any errors/corruption will be caught.
  
  **Cons** requires access to original CAD system.

  **Good for** long-term archiving and reuse.

Supporting documentation

- Annotations preserving vulnerable semantics
- Standards and conventions used, e.g. US National CAD Standard
- Client’s specification
- Rationale models
- Process models
- Fully commented procedural scripts
6 Final thoughts

So, in conclusion then, CAD files can be used for many different purposes, so it is important for an archive to establish why models are being kept and use that as the basis for deciding a preservation strategy.

If particular properties are important, create validation tests that can prove whether they have survived.

There is information that cannot be exported from native CAD formats, so they may as well be preserved for as long as they are readable, but since they will be unreadable one day, they should also normalized to a full exchange standard like STEP, and to a visualisation format for backup.

It is important not to neglect supporting documentation, which can be vital for future users to be able to understand the models.

And if we are serious about preserving CAD model we need to impress on regulatory bodies and CAD vendors how important support for standards is. The information locked up in CAD models is too valuable for its survival to be left to the whims of single corporation whose interest lie elsewhere. It needs to be put into open, standard formats where it can be exchanged and preserved for the benefit of society as a whole.

Alex Ball. DCC/UKOLN, University of Bath. http://alexball.me.uk/

Except where otherwise stated, this work is licensed under the Creative Commons Attribution 4.0 International licence: http://creativecommons.org/licenses/by/4.0/

The DCC is supported by Jisc.
For more information, please visit http://www.dcc.ac.uk/

Preserving CAD DPC Technology Watch Report: http://dx.doi.org/10.7207/twr13-02