SPRUCE Mashup – London 2

Laser scanning and 3D data
Collections

- "3D Modelling Data for Terracotta Roundels" (Historic Royal Palaces)

- "Laser Scanning data of Gabo sculptures" (Tate Collection)
Short version

- Go read the Archaeological Data Services' "Laser Scanning for Archaeology: A Guide to Good Practice"

- (And if you're a DPC member, read the notes from a recent event, "3D Laser Scanning: Seeking a New Standard in Documentation")
Issues

"What have we got, and what does it do?"

(Essentially: "appraisal")
What is laser-scanning/3D-modelling?

- Scanning
- "Registering"/processing
- Modelling
- Rendering
What happens?

- **Scanning**: produces "raw" laser scan data, can be proprietary, or plain-text
- **Registration/other processing**: produces "point-cloud" data, often in a plain-text format
- **Modelling**: produces a 3d model, in an often-"standard" (if proprietary) format
  (and/or may produce an "object mesh" object -- again, often in a plain-text format)
- **Rendering**: produces images, videos, VR, etc. (usual formats)
Scanning data
Point cloud
Object/surface mesh
Rendered image
Rendered image
Example file types/counts from one HRP roundel (at various resolutions/scale):

- 5 obj
- 4 tif
- 4 rtf
- 4 ply
- 4 exe
- 3 txt
- 3 mov
- 1 wrl
- 1 swatch
- 1 mayaSwatches
- 1 fbx
- 1 dmg
- 1 asc

(In total, so far, about 9.6 Gb, for three separate scanning events; but that’s just one roundel out of twelve.)
Example file types/counts for one Tate sculpture (at single resolution/scale)

- 496 pqk
- 478 mask
- 478 hist
- 248 pf
- 230 bre
- 21 jpg
- 4 db
- 4 avi
- 3 archive
- 2 txt
- 2 3dm
- 1 vlt
- 1 pwk
- 1 dwg
What have we worked out?

- **bre** -- laser scan data (?) -- proprietary format
- **pqk** -- laser scan data (?) -- proprietary format?
- **mask/hist** -- ??? (seems to belong with laser scan data)
- **asc/txt** -- ASCII text -- point clouds or scan data, generally
- **obj/ply** -- object meshes (mostly plain text) -- contain surface/object meshes (i.e., "joined-up" point-clouds)
- **3dm** -- "3d model", represents the object as a collection of 3d geometry shapes (can be opened, and possibly modified by a number of tools)
- **fbx/wrl** -- also 3d models, but of a different format
What have we worked out (2)?

- dwg -- "drawing" -- 3d model in proprietary AutoDesk format
- pwk -- "Polyworks" -- 3d model in proprietary Polyworks format
- avi -- AVI video -- rendered video of object
- jpg/tif -- Images -- rendered images of object
- mov -- video/QTVR -- rendered video, or (QuickTime) virtual-reality object
- mayaSwatches/swatch -- suspected rendering assets (images/textures/etc.)
- (archive/dmg/exe/vlt/rtf -- miscellaneous cruft)
What HAVEN’T we worked out?

Exactly what do we preserve? (See ADS best practice guidance -- especially regarding metadata for preservation) But still not completely clear.

- Images/videos/VR are all (in-theory) derivable from the 3d models (in the same way JPGs are derived from TIFFs), but does depend on relevant software to perform this task

- Point-clouds are easily preserved (just plain text), but may be results of pre-processing work, which may or may not need to be preserved in its own right

- 3d models and object meshes are also important (since there may be pre-processing work here, too), but often highly-tied to the software that created them

- For laser-scanning/point-clouds/etc., recent DPC event notes comment on new file formats ("E57") which are intended to be more interoperable, but weren't designed with preservation in mind, and still very much in development/controversial/etc.
What are the take-aways?

- Seriously, go read the ADS guide to best-practice.

- This is a novel, still-evolving space; probably need some in-house documentation/expertise, as well as clear and open relationships with vendors who can be relied upon to provide the "right" data.

- Very much *not* a "solved problem“ (yet).

- Shared learning/community/etc. should help out somewhat.