

Subgroup discussion

Rob Grim

Alison Heatherington

Sebastian Rahtz

Dirk Roorda

What is MIXED?

Our recommendations for what MIXED should concentrate on:

Approaches:

- Starting point for re-use
- Metadata standards – adding technical metadata at point of conversion
- Digital documentation on quality and validity of conversions

Goals:

- Data conversion service on a file basis, fits into wider workflow
- Framework for managing conversion
- Definer of output formats

Condition:

- Web service, which can be called, independent of workflow. So it can fit into any workflow.
- Interaction with data type services

Approach: Focus on re-use

- Re-use different to dissemination
- Dissemination is future use in a way the data producer intended.
- How does re-use take place? Not just format but context (metadata must be right)
- Can be re-used in other ways. MIXED is stronger on re-use than dissemination, as some functionality gets ignored.
- The place of MIXED is in management phase rather than ingest

Approach: multiple manifestations

- Create entities in such a way that they can be combined in future in useful ways
- MIXED enhances possibilities of re-use. Doesn't solve all Digital Preservation problems.
- 3 manifestations:
 - Original
 - Presentation
 - Re-usable version

Metadata

- Software obsolescence less of a problem than metadata issues
- Good information management important, creation and retention of metadata
- Make sure data is right (preprocessing) – more than half the battle.
- 3 ways of producing metadata:
 - Metadata production,
 - Metadata repositories,
 - Metadata registries.

Approaches: Documentation

- Documentation – open and accessible, including source code. Describe conversion process to allow users to carry out own validity checks.
- Must be machine parseable

Technical considerations:

- Validity checks – round trip processes
- Test suites – accessible to users
- Submit content to UDFR
- Use UDFR/PRONOM to assist with gaps analysis

Future goals for MIXED

- Establish position in broader community and leverage activity
- Test case, community involvement
- Gaps and opportunities analysis, so know where to focus effort – which data types to look at
- More granularity / standardisation of data types.
- Building trust by proving quality of conversions
- Collaboration – Planets, PRONOM, UDFR etc