Is the Screensaver a Significant Property?

Defining a Metadata Model and Cataloging Workflow for Software and Computing Environments

Seth Anderson, Yale University Library
Best Practices Exchange 2019
Columbus, OH
May 1, 2019

Yale
Project Goal

Deploy and scale infrastructure and services for software emulation, including distributed management, sharing, documentation/discovery, and access.
What is Emulation-as-a-Service?
Distributed Mgmt

- A network of distributed nodes, each contributing to the EaaSI service and the software development roadmap.
Sharing

- In-network sharing of software images and configured environments.
- Yale University Library will contribute at least 3000 pre-configured software applications running in configured environments.
Access

- Emulated CD-ROM environment sharing service
- Virtual Reading Rooms Service
- Scientific Software Portal
- API to automatically render objects in original software via emulation
Documentation/Discovery

- Incorporating services developed by Wikidata for Digital Preservation
- Comprehensive, open, machine-readable documentation
- Defining profile for description of software and computer environments
Metadata for software and computing emulation
Our needs

1. Discovery
2. Provenance
3. Administration
4. Automation
Existing “Schema”

- Trustworthy Online Technical Environment Metadata (TOTEM) Database
- GAme MEtadata and Ctiation Project (GAMECIP)
- Video Game Metadata Schema
- Wikidata for Digital Preservation
Metadata
Model/Schema
WORK

EXPRESSION

MANIFESTATION

EaaSI – Software Hierarchy
<table>
<thead>
<tr>
<th>Category</th>
<th>Description</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOFTWARE VERSION</td>
<td>An expression of a software application which provides a single coherent presentation with a well defined functionality and behaviour.</td>
<td>Word 2000 (version 9.0)</td>
</tr>
<tr>
<td>SOFTWARE OBJECT</td>
<td>Application and system software that is available for installation/configuration in an environment.</td>
<td>Word 2000 (version 9.0) from the NSRL collection</td>
</tr>
<tr>
<td>OBJECT FILES</td>
<td>Files necessary for installation and/or operation of a software application within a computing environment.</td>
<td>Disk images of installation media</td>
</tr>
</tbody>
</table>
CONFIGURED SOFTWARE

Application or operating system software installed and configured for operation within an emulated computing environment.

SOFTWARE ENVIRONMENT

Configured software components—operating systems, drivers, libraries, applications, etc.—of computing environment as recorded on environment disk image.

ENVIRONMENT IMAGE

Recorded contents of computing environment’s software configuration.
CONFIGURED MACHINE

Emulator configuration settings that replicate the hardware of a physical computer system.

EMULATOR

Emulation software used to manifest hardware configuration settings in the computing environment.
**COMPUTING ENVIRONMENT**

Combination of software environment, configured machine, and metadata presented to users of the EaaSI system.

**OBJECT ENVIRONMENT**

Computing environment with associated digital object(s) for rendering and interaction.
EaaSI – Objects

Computing Environment is part of Object Environment is part of Digital Object
DIGITAL OBJECT

Information or works from institutional collections intended for representation and interpretation within an emulated Environment.

OBJECT ENVIRONMENT

Computing environment with associated digital object(s) for rendering and interaction.
Data Capture Methods
Challenges

- Controlling data quality
- Edge cases
- Domain specific metadata
- Ease of use
- User fatigue
Next Steps

- Continue schema definition
- EaaSI UI updates
- Explore edge cases
A Very Special Thanks to our Funders...
Thanks!

You can find me at
- @sth_rbrrt & seth.r.anderson@yale.edu

And learn more at
- softwarepreservationnetwork.org/eaasi
- And on Twitter at #eaasi