JHOVE Tutorial

Carl Wilson, Technical Lead
Open Preservation Foundation
iPres 2018 Boston, Massachusetts
The Open Preservation Foundation

• A not for profit, global membership association providing stewardship of open-source tools for the digital preservation community.

• Founded in 2010 to sustain the results of the EU PLANETS project.

• The OPF provides stewardship for JHOVE, veraPDF, FIDO and more.

• Carl is the Technical Lead at the Open Preservation Foundation
Agenda

• A brief overview and history of JHOVE
• Demonstration of installation and running JHOVE
• A look at JHOVE results and how to interpret them
• How you can contribute to JHOVE
• A look at the Foundation’s plans for JHOVE
What is JHOVE and a little history

Purpose & Background
What is JHOVE?

JSTOR/Harvard Object Validation Environment
An EXTENSIBLE software framework for format:
- identification,
- validation, and
- characterisation.
JHOVE provides some answers for the following questions regarding a digital object:

- What format is it?
- The object is of format: F, is it?
- The object is of format: F, what are its properties?
JHOVE is available in three forms:

• Command line Tool - used by researchers and practitioners / embedded in DP workflows.

• Desktop GUI - file at a time interface, aimed at general users wanting to try JHOVE.

• JHOVE API - used by developers embedding JHOVE in their own projects.
JHOVE Origins

Versions 1.1 (April, 2008) - 1.6 (January, 2011) developed by:

• Harvard University Library &
• JSTOR; with funding from the
• Andrew W. Mellon Foundation

and made available under the GNU Lesser Public License (LGPL).
Versions 1.7 (August, 2012) - 1.11 (September, 2013) maintained and released independently by Gary McGath.

Around May 2014 Gary put out a call for somebody to take over the maintenance of JHOVE......
In March/April 2015 the Open Preservation Foundation took over the mantle of JHOVE maintainer from Gary following discussions with Gary, Harvard and JSTOR (Portico).
The OPF Inheritance

JHOVE assets OPF is now responsible for:

• code base (GitHub);
• 15 format modules;
• website http://jhove.openpreservation.org/;
• 16 downloadable release packages;
• mailing lists;
• issue tracker with 78 open issues, 70 closed; and
• 11 open pull requests, 203 closed on GitHub
Installing & using JHOVE, a demonstration

Getting Started with JHOVE
Analysing the JHOVE results

What’s JHOVE Telling Me
Anyone can contribute and help

Improving JHOVE
Log an Issue

Use the GitHub Issue tracker:
https://github.com/openpreserve/jhove/issues

• If there’s something wrong.
• If you have an idea about something that can be improved.
• If you’d like to see a new feature.
You can also submit a contribution as a GitHub Pull Request: https://github.com/openpreserve/jhove/pulls

• This can be a fix for the code base
• It could be improvements to the documentation or website
• It could even be a whole new module
A look at the JHOVE roadmap and questions

The Future of JHOVE
The Future of JHOVE

- Better process for module versioning
- Assign IDs for error messages
- Instrumentation/monitoring/measurement
- JHOVE webservice REST API
- Standardise underlying reporting mechanism
- Java version increment
- Updating JHOVE test corpus / error reporting coverage
- Strategies for user updates and major changes (pre major changes in 1.24)