Digital Preservation Interoperability through Preservation Actions Registries

Matthew Addis, Justin Simpson, Jon Tilbury, Jack O'Sullivan, Paul Stokes, Carl Wilson

JISC

Arkivum, Arefactual, Preservica, Open Preservation Foundation

http://parcore.org/presentations/

Agenda

- Background & Motivation
- Project outcomes
- Proof of concept implementation
- Next Steps

Background: The problem

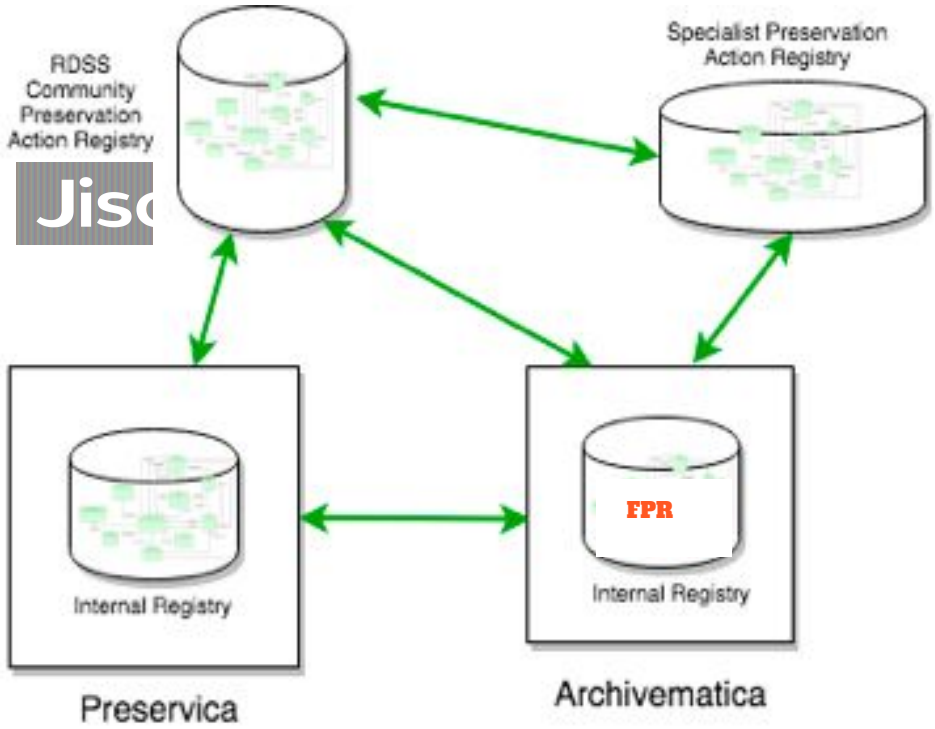
- Users want the best advice, wherever it comes from
  - Identification, property extraction, validation, migration, rendering, tools
- Multiple parallel initiatives research and advise on best practice
  - Products such as Preservica & Archivematica
  - Practitioners
  - Academics
  - Specialists
- but they don't talk to each other effectively

Background: Motivation and Objectives

- Want to
  - Improve the quality and ease of use of advice sent to practitioners as soon as it is available
  - Improve research cooperation and reduce repetition
- Expected outcome
  - Provide a mechanism to exchange information between all parties regardless of which system they use
- Exclusions
  - Protocols for prioritising and authorising which advice applies to which user / system / intent
  - One registry to rule them all

Background: Jisc RDSS Project

Development of a multi-vendor shared services platform drove discussions of interoperability of format policies (i.e. “preservation actions”) between preservation systems.



Background: Project Conception

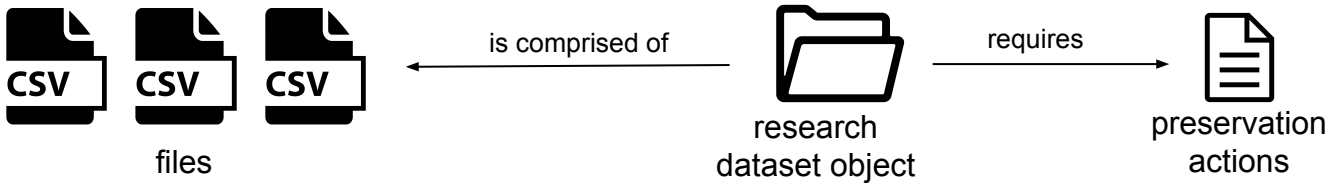
A JISC funded project to initiate the process to deliver benefits to RDSS users

Arkivum, Preservica and Artefactual as RDSS product suppliers

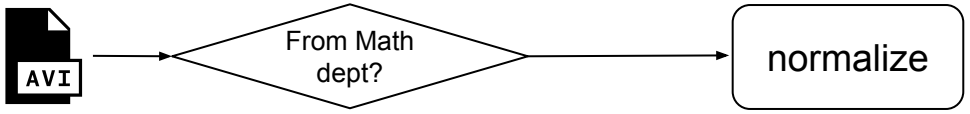
Open Preservation Foundation as respected independent shared DP technology supplier

Background: Motivations

1) Preservation is not just about file formats, it's about intellectual entities/objects

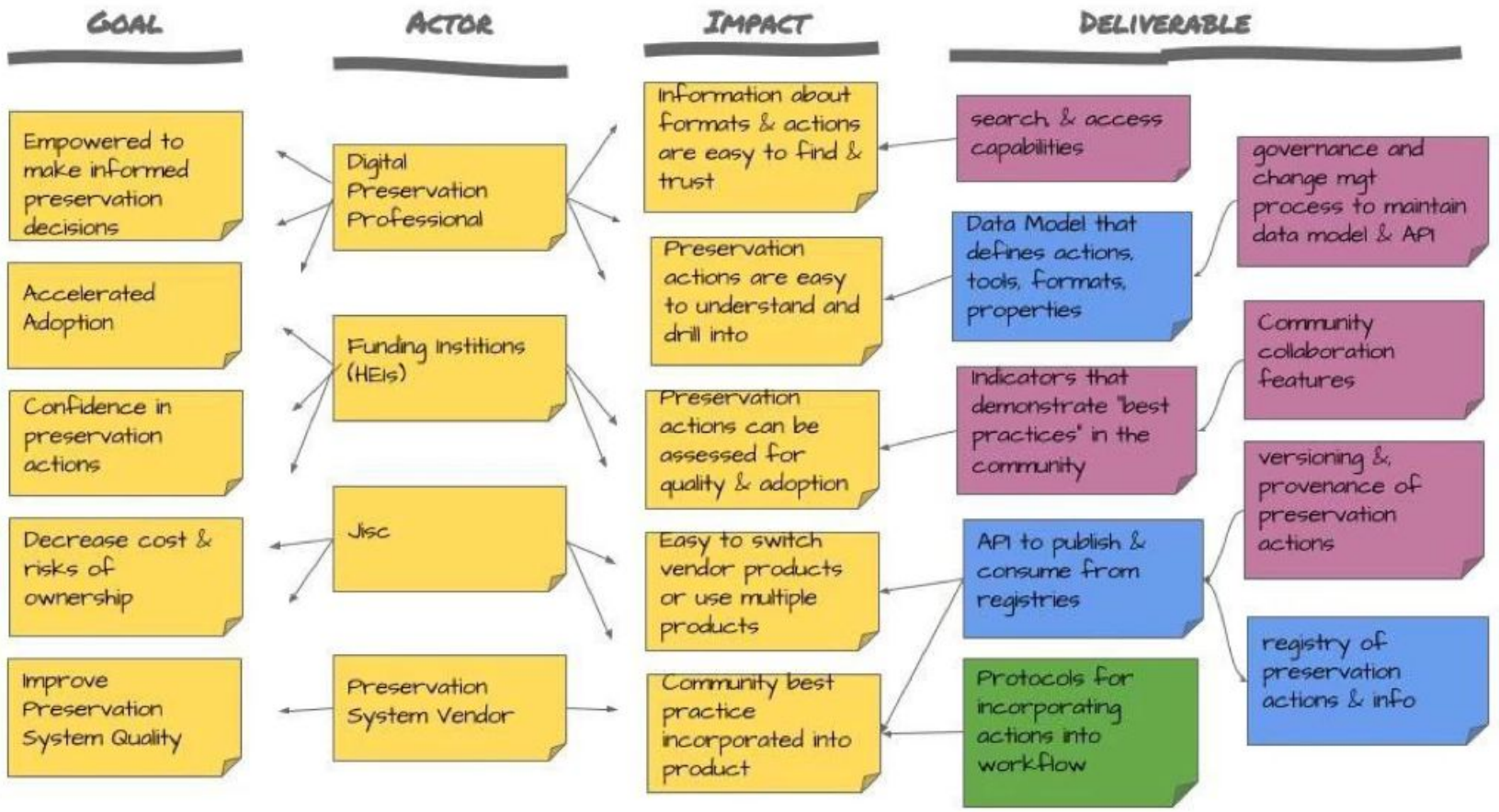


2) We need a way to define/describe context - why is this action being taken? what is the business rule?



3) Reinventing the wheel - preservation actions are not portable across systems (e.g. Archivematica, Preservica, others)

Background: PAR Impacts



Open Preservation Foundation

arkivum Bringing archived data to life

Preservica Active digital preservation

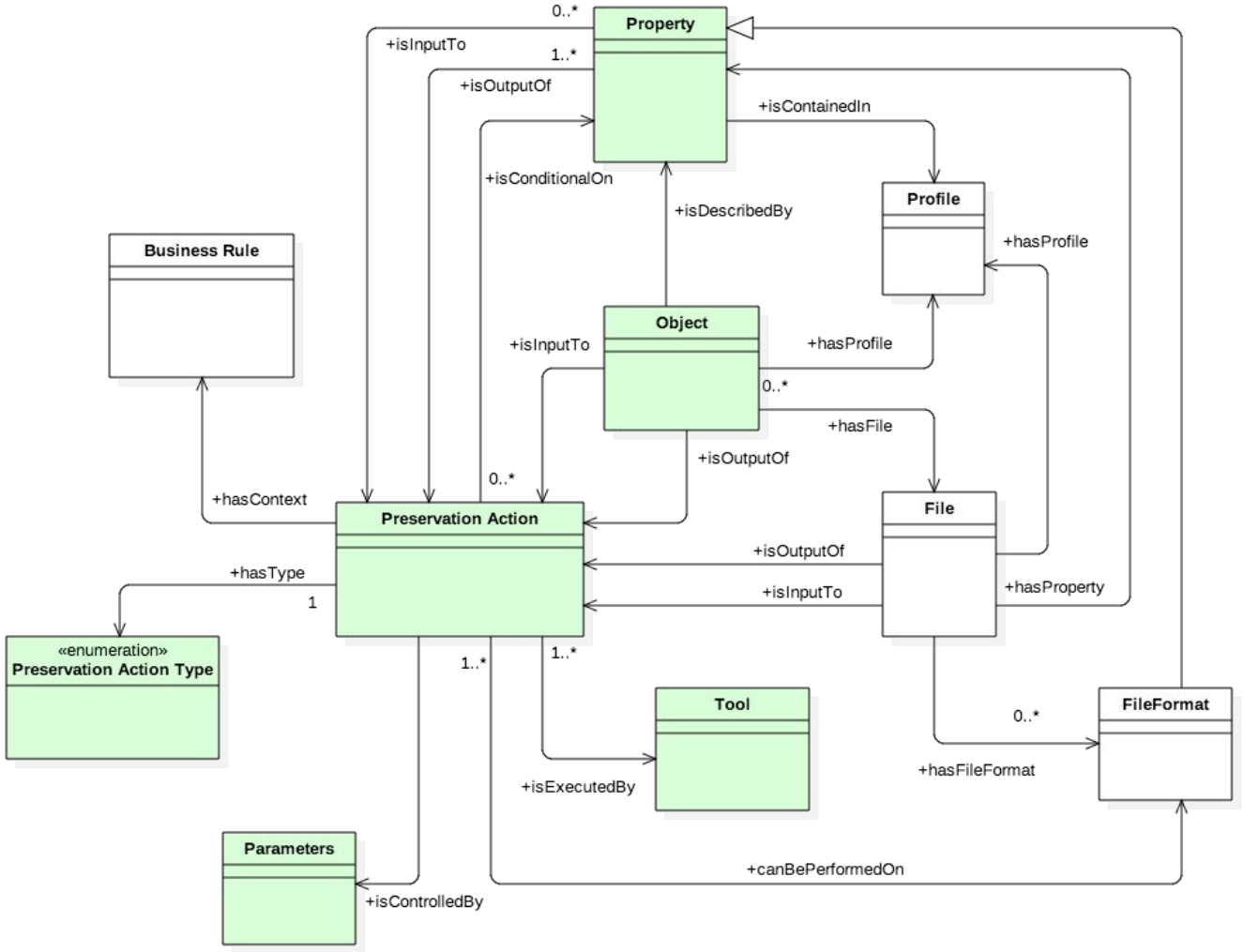
artefactual

Jisc

What have we produced and why?

Conceptual Model	• Common framework for everyone to work to • Something to argue about and agree upon! • Interlingua between preservation systems
Json Schemas	• Formal definition of the PAR model • Machine readable, used in API payloads • Used to test and validate interoperability
API	• Common interface for preservation systems • Well defined way to exchange information
Executable DP Actions	• Cross-platform way to deploy/run tools • Unambiguous and vendor independent
Proof of Concept	• Prove PAR is possible! • Not just a talking shop or paper exercise • Reference implementation to share

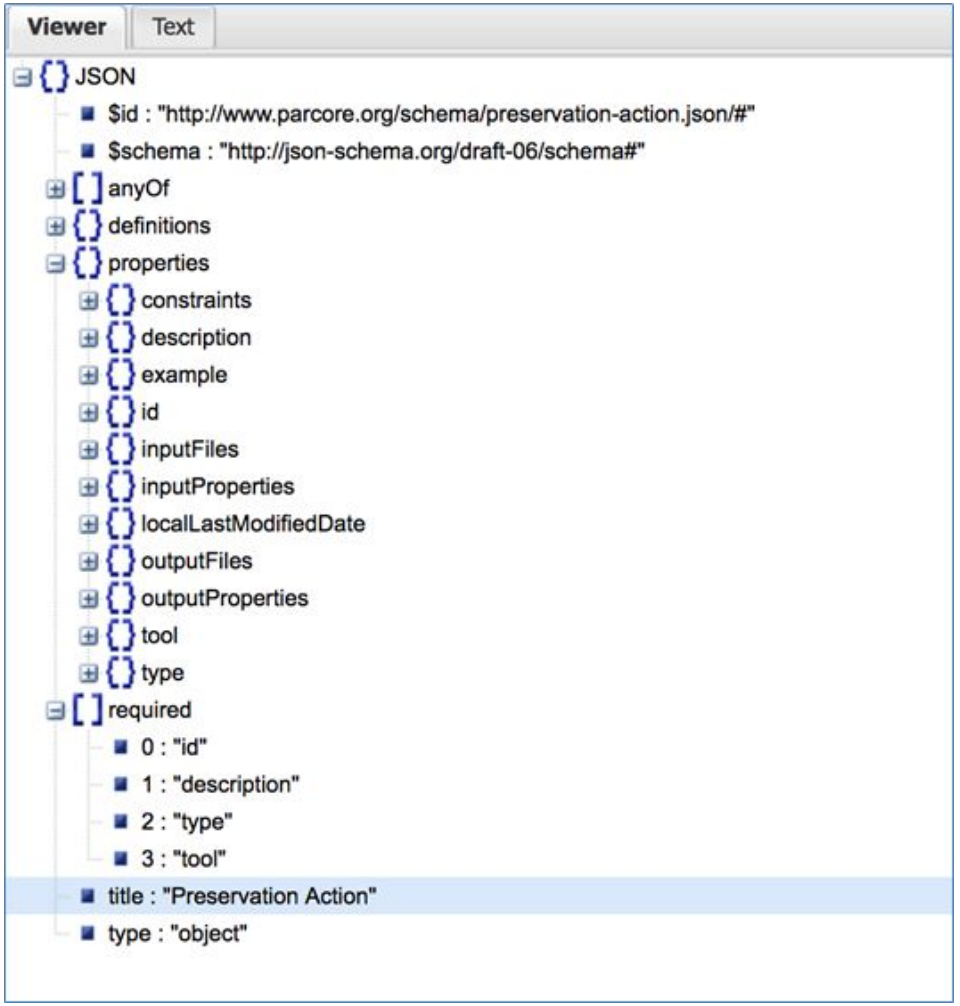
PAR Conceptual Model



https://doi.org/10.6084/m9.figshare.6628418

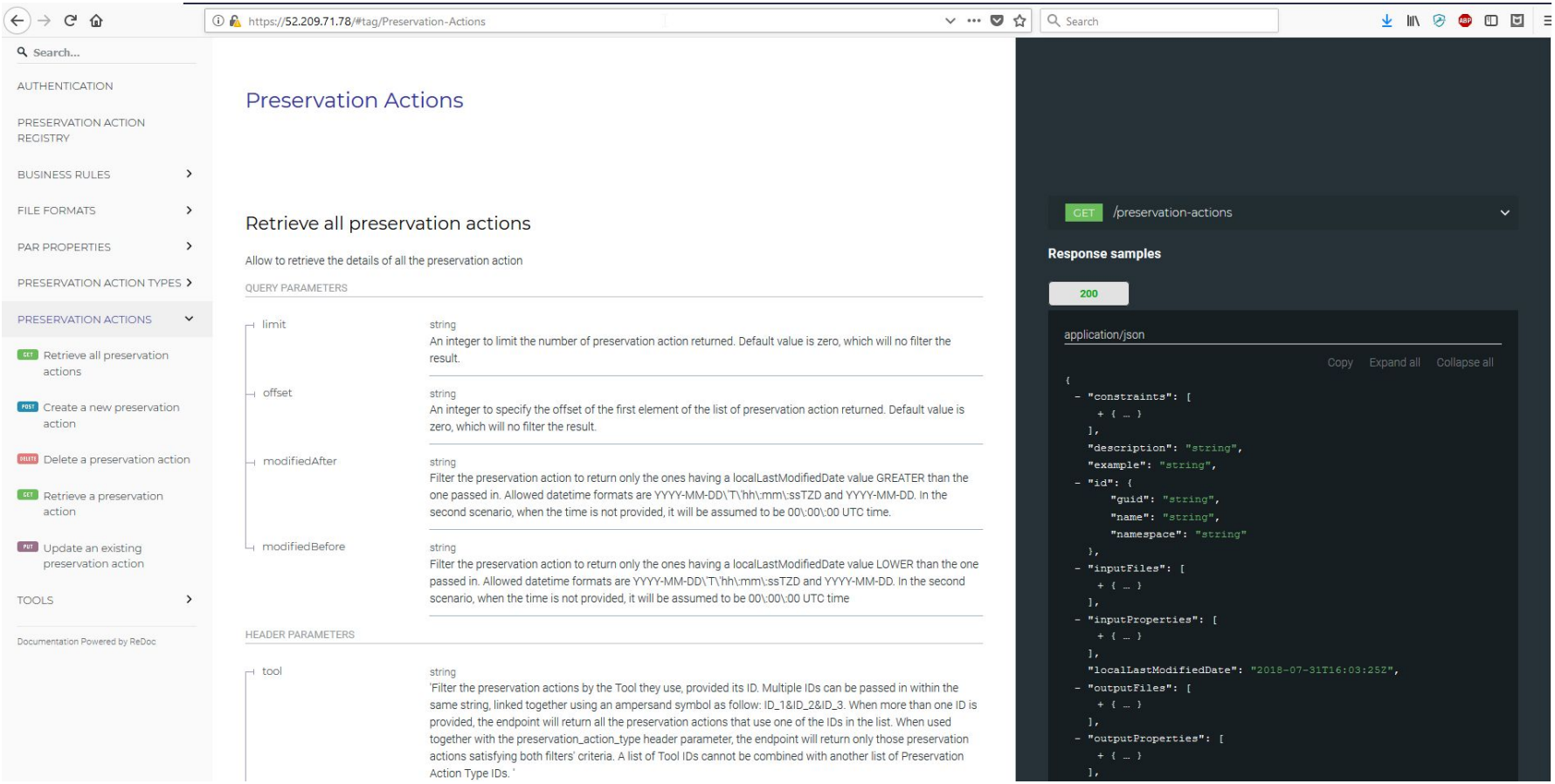
JSON schemas

- Tool
- Action
- Action Type
- Format
- Property
- Business Rule



https://github.com/JiscRDSS/rdss-par/tree/master/schemas

APIs



https://github.com/JiscRDSS/rds-par/tree/master/api

Open Preservation Foundation

arkivum Bringing archived data to life

Preservica Active digital preservation

artefactual®

Jisc

Executable Tool Definitions

- Machine readable spec for running a tool
  - Tool command line
  - Parameters and flags
  - Inputs and outputs
  - Pre and post processing


```
[job mediaInfo2.cwl] completed success
{
    "width": "1280",
    "bitrate": "748253",
    "height": "720"
}
Final process status is success
```

Property extraction

Fixity check

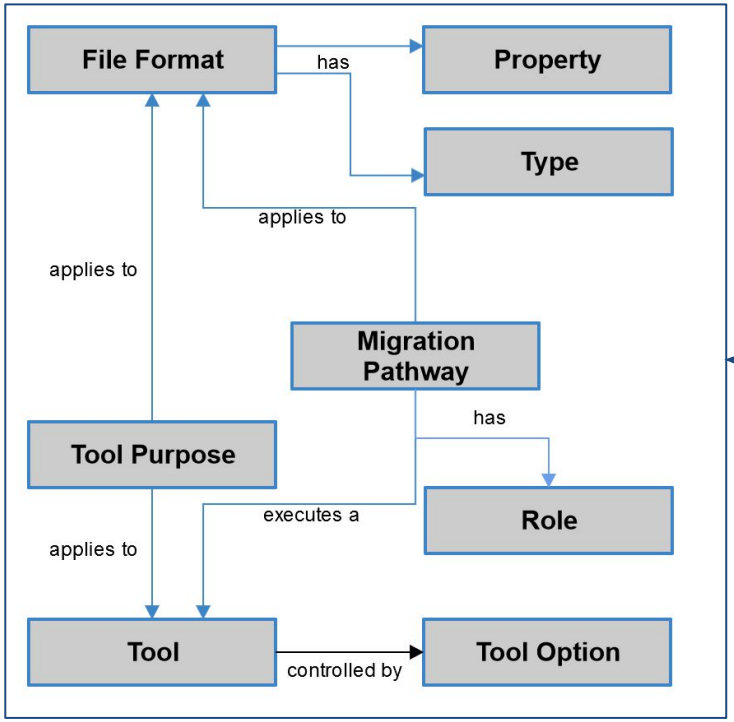
```
[job md5check2.cwl] completed success
{
    "fixity_report": "PASS"
}
Final process status is success
```

https://github.com/JiscRDSS/rdss-par/tree/master/examples/cwl

Open Preservation Foundation · arkivum Bringing archived data to life · Preservica Active digital preservation · artefactual · Jisc

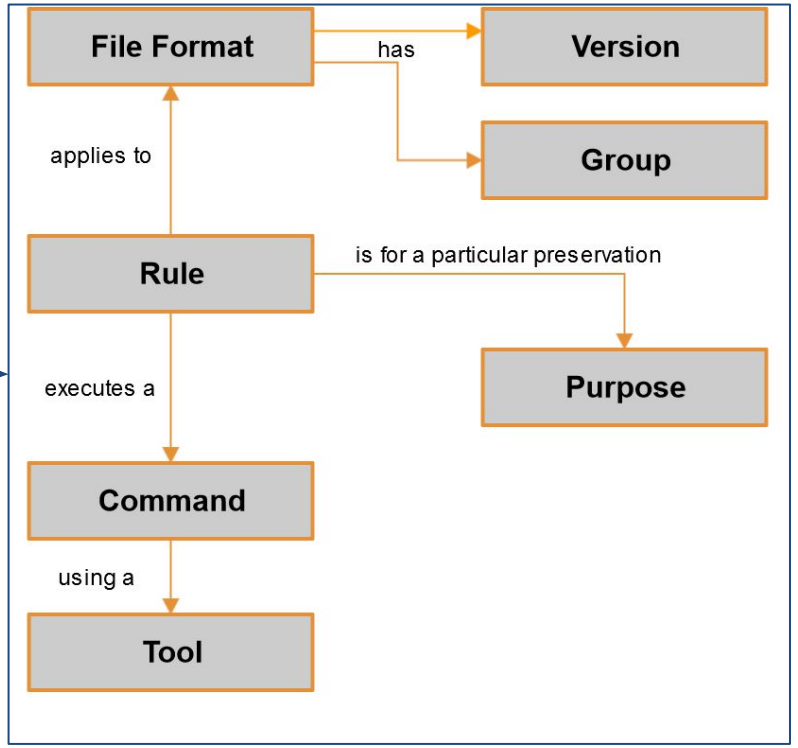
Registry (In)compatibility

Preservica Registry

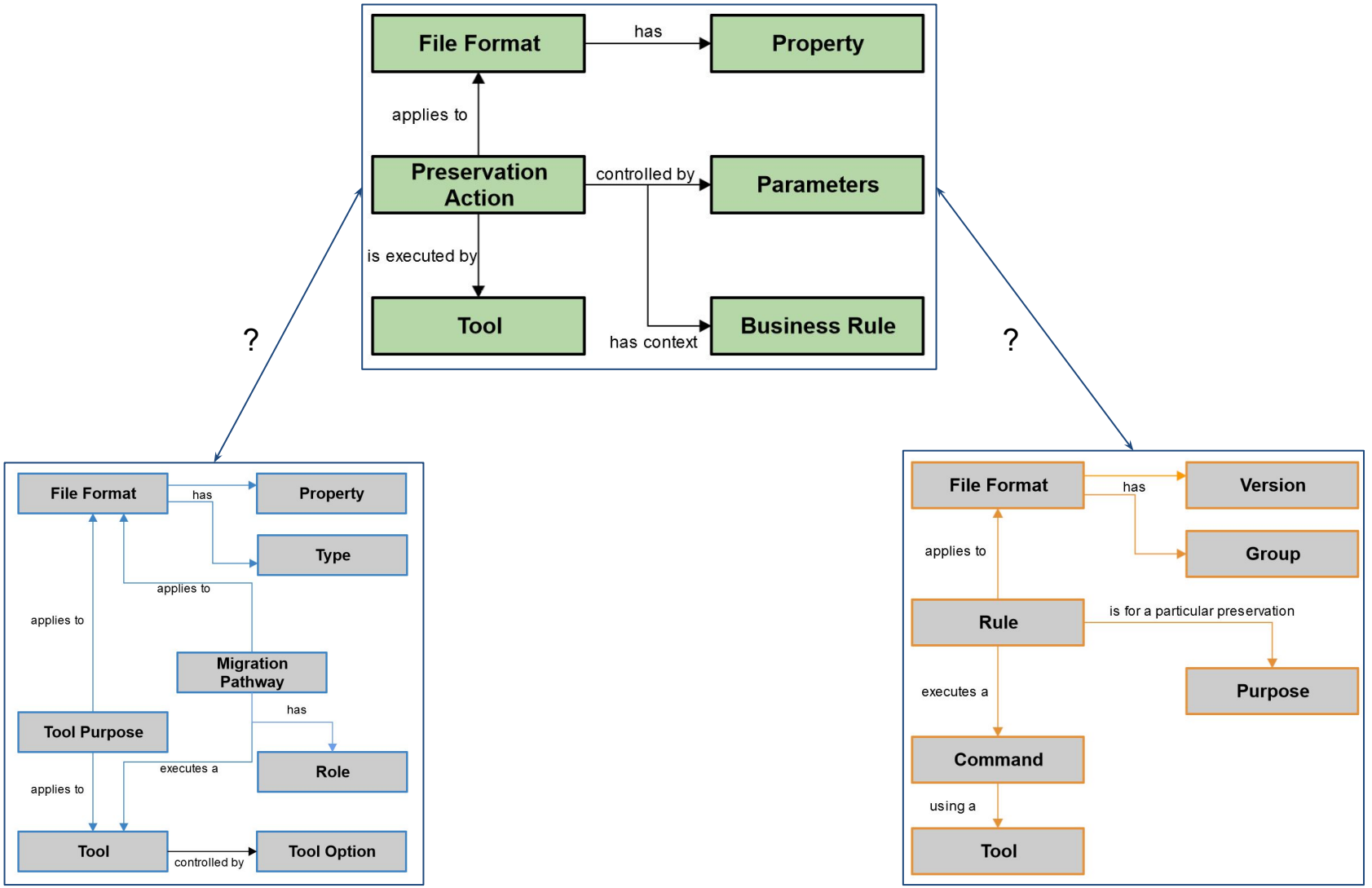


?

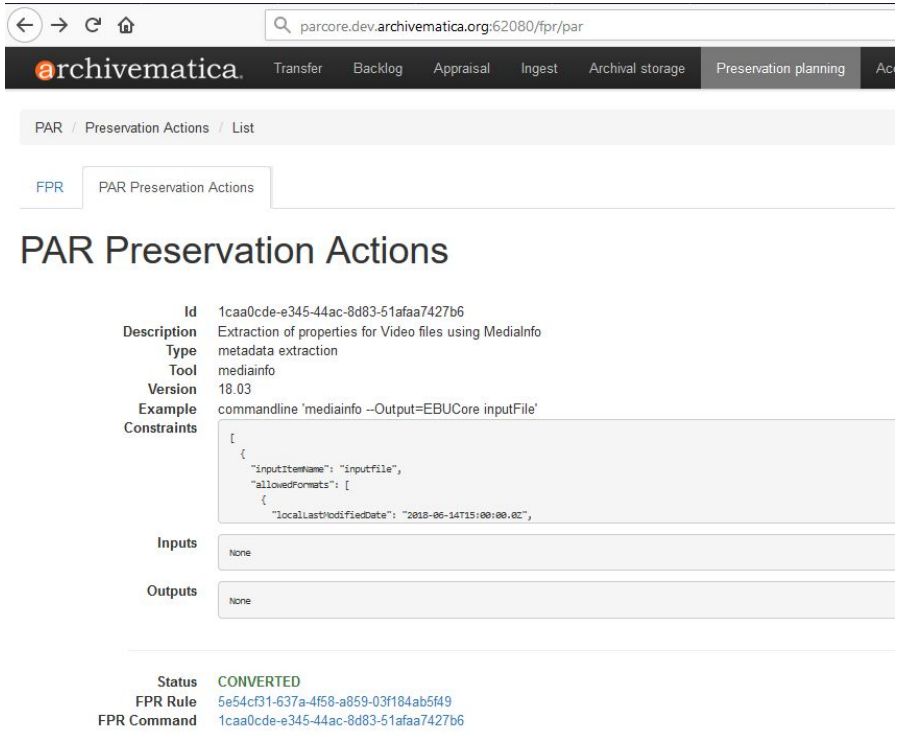
Archivematica FPR

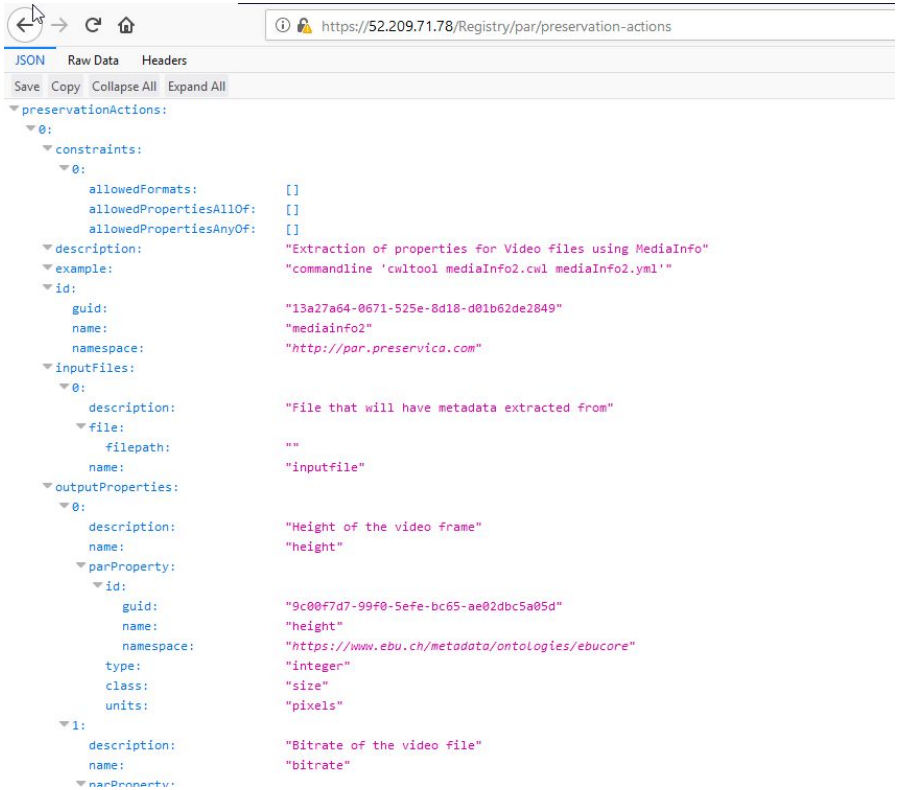


Common Language

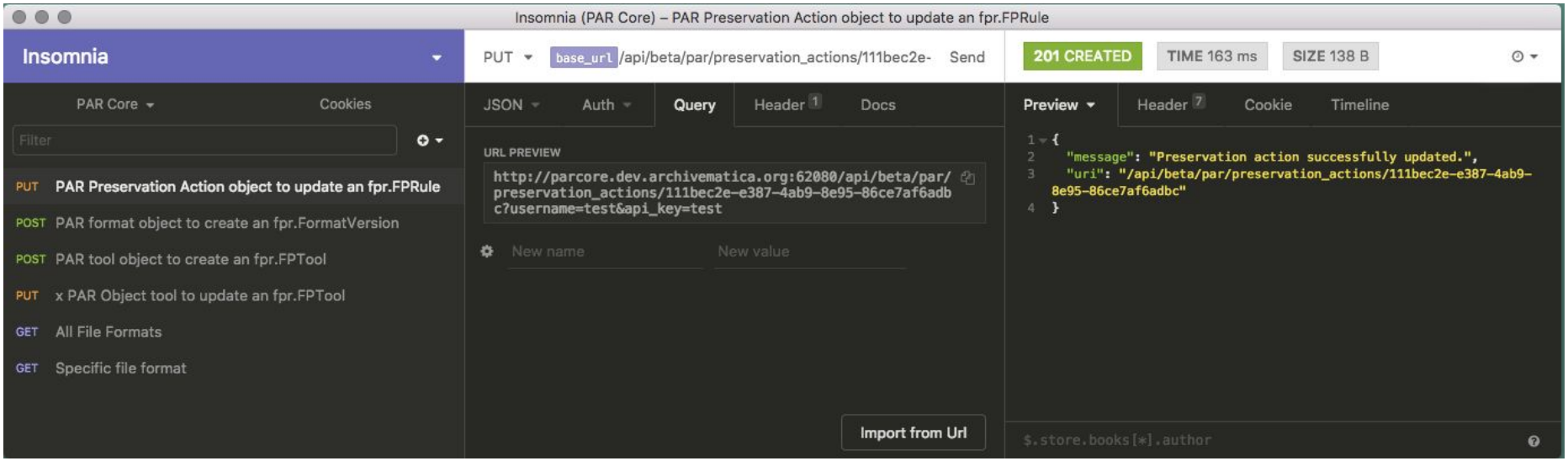


API Exposure





Working API Examples



Demo API Servers:

http://parcore.dev.archivematica.org:62080 (user: test, apikey: test)
http://52.209.71.78/Registry/par (user: test, password: test)

Next steps

- Real use cases demonstrating real benefits
- Consortium of funders
- OPF coordination
- More members to provide API endpoints
- More sync tools to exchange information between systems
- Ability to act on information exchanged

Resources

- Project pages
  - http://www.parcore.org/
- Github repo
  - https://github.com/JiscRDSS/rdss-par/
- iPRES paper
  - https://doi.org/10.6084/m9.figshare.6628418
- DPC blog post
  - https://www.dpconline.org/blog/a-new-era-in-collaboration-in-digital-preservation-research
- Project announcement and contacts
  - http://openpreservation.org/news/arkivum-artefactual-the-open-preservation-foundation-and-preservica-collaborate-on-new-jisc-initiative-for-sharing-preservation-action-best-practice/

Open Preservation Foundation arkivum Bringing archived data to life Preservica Active digital preservation artefactual Jisc

Interested?

This could be the first step in a global exchange of best practice between all stakeholders in DP

We want you to make sure this is truly global.

We'll be at après-iPRES
Thursday 27 September 1-5pm
Kotzen Room, Lefavour Hall at Simmons College

Vote for PAR to find out more!

Contact: info@parcore.org

Open Preservation Foundation arkivum Bringing archived data to life Preservica Active digital preservation artefactual Jisc