



ESCAPE

European Science Cluster of Astronomy &
Particle physics ESFRI research Infrastructures

Metadata for software: Beyond citations

Mark Kettenis
JIVE

Software metadata standards

Datacite

Zenodo

SWH

Codemeta

All these to focus on
referencing software



What software can I use to process this data?



Data formats: input and output

- Choose software based on data formats they operate on
- Distinguish between input and output
 - Input: Can I use this software to process this data?
 - Output: Can this software produce...?
- Science platform (ESCAPE WP5) can use this to offer choices
- Not all communities use well-defined/standardized data formats?



Data formats: media types

- RFC6838 “Media Types”

- Used to be known as MIME types

- Registered types can be found at:

- <https://www.iana.org/assignments/media-types/media-types.xhtml>

- Has a namespace for “unregistered” formats

- Example

- application/fits, image/fits (RFC4047)

- Already used in many VO protocols!



How can I run this software?

(without spending a day to install it)



- Source code
 - Build dependencies
 - Compiler, code generators, build system
 - Runtime dependencies
 - OS, libraries, CLI tools
- Traditional packages
 - Package format
 - OS-dependent
 - Runtime dependencies
 - Other packages needed
- Containers
 - Container format
 - Needs to be specific: “dockerfile” or “dockerimage” instead of just “docker”



Multiple distribution methods are possible!

Jupyter notebooks?

For OSSR we need to concentrate on a subset



What computing resources do I need to run this software?



- CPU type
 - Highly optimized binary may not run on all CPUs
- Parallelized
 - Provision multiple cores?
- Cluster-aware
 - Provision multiple machines?
- GPU
 - Needs GPU hardware specifics?

Not always dependent on the software:

- Number of cores
- Memory requirements

Essential for provisioning
resources by science
platform

Minimal viable
configuration?



I need your help!

We need a Vocabulary for data formats:

- Survey of data formats in use by partners in ESCAPE
- And assign media types to them
- Potentially publish as IVOA note

