§Related software

In the initial analysis, this element was intended to include information on software resources (algorithms, scripts, libraries, applications, etc.) that are related to the dataset in the way specified by a role. See also “Citations”.

Questions:
What exactly is the type of software we are talking about? (operating system, hardware specific software, processing data, applications …) - Software that is related to the creation of this data set -> but then, once it comes to dependencies, we need to go further down the chain - that’s a never ending story -> do containers provide a solution?
What do we do with authors of versions where they just forked of some major based software_ -> even for a fork, the authors of the software being used (at certain time) should be known

---
Additional questions: is this the software used to analyze/model the data? Open the data file?

Problems might occur when data was generated using a pipeline/scientific workflow. That will mean a few software combined together and order of software is important. Should the exact computation path be described on metadata level or just information about used software?

Just to start collecting relevant descriptors:

Some standard elements:

- Name of the software
- Description
- Author(s)
- Type of software (program, library, Web Service …)
- Version/Release
- State
- License -> Availability
- Documentation -> Related publication
- Dependencies
- Source -> Location
- Repository
- Unique identifier [Should identifier be considered per software no matter of version/release or should it be considered new version/release mean

new id. There is also a case when someone use software from repository with some commits not yet released/versioned but with some commits relative to previous version made. If software is hosted on git based repo git commit ID (SHA-1 Hash) could be considered as unique ID]
- Form that the identifier provides (source code, container, binary …)

Concerning processed datasets, other things are also important:

- Environment (operating system, compilers, library version …)
- Parameter settings (default values being used; special selections being chosen)
- Code basis

Potential input: the Python community uses what they call Trove classifiers to categorize software: https://pypi.python.org/pypi?%3Aaction=list_classifiers It's quite simple but at the same time it seems to serve that community nicely. Some parts are specific to the Python ecosystem but some are more generic. At the Python Package Index (PyPI; https://pypi.python.org/pypi) it can be seen in action.

Ted Habermann -
This page makes it clear that we need a standard model for citations that can be used in many ways (see Citations).

Mappings

Please give below:
- The name of your data archive, repository, catalogue, etc.
- For your native or internal metadata scheme (or one you normally use), please list the elements that correspond to this high-level element. For each one:
    - Give the name of the element
    - Give any constraints on the value (controlled vocabulary, syntax/encoding)
    - Highlight any divergence in the semantics of the element from the discussion above.

Eagle-i (www.eagle-i.net; search.eagle-i.net/model)
Software data model:
https://search.eagle-i.net/model/#t=http://purl.obolibrary.org/obo/ERO_0000071&of=score

Software
Definition:
Digitally stored data such as computer programs and other kinds of information read and written by computers.

Annotations: ClassGroup_InstanceCreate; ClassGroup_PrimaryResourceType
URI: http://purl.obolibrary.org/obo/ERO_0000071
Properties:
* indicates a required property
***anything not a string is another resource type in the eagle-i ontology
Access Restriction(s) string
Additional Name string
Algorithm used Algorithm
Coded in Programming language
Contact Person
Data Input Data
Data Output Data
Developed by Organization; Person
Exchange facilitator string
Funded by Organization
Manufacturer Organization; Person
Operating System string
Part of Collection Resource Collection
Related data collection method Data collection method
Related grant number string
Related Publication or Documentation Document
Related study design Quantitative study design
Related Technique Technique
Resource Description string
Software license Software license
*Software name String
*Software typeSoftware type
Software purpose Software purpose
Used by Organization
Version string
Website(s) string