

Guidelines for Resource ID Formation

Last Update: September 24, 2009

A Resource ID is a Universal Resource Identifier (URI) with the basic formation of:

scheme://authority/path

Each described resource in a system must have a unique identifier. Resources described with in accordance with the SPASE data model will always have a scheme of "spase".

To aid in generating unique identifiers the following conventions should be used for defining the path portion of the URI. This convention uses information about the resource to create an informative and unique URI. The full formation rule for a resource identifier is:

spase://NameAuthority/ResourceType/Project/Observatory/InstrumentType/Cadence

where

NameAuthority: Is the accepted name or acronym for the domain of the resource. This is commonly the acronym for the virtual observatory. For example, for resources named by the Virtual Magnetospheric Observatory the *NameAuthority* is "VMO". Authority names must be unique within the scheme. SPASE maintains a list of authority names for the "spase" scheme.

ResourceType: The type or class of resource. For example, for a NumericalData resource the *ResourceType* is "NumericalData".

Project: The name of the top level observatory, observatory group, project or agency. For example, the project IGPP/LANL for a NumericalData resource the *Project* is "IGPP/LANL". The project field can be repeated as necessary to represent organizational hierarchy. The *Project* may be omitted if not relevant.

Observatory: The name of the observatory which hosts the instrument. For example, an observatory in the IGPP/LANL project is "CRT".

InstrumentType: The type of instrument used in the observation. This should match the allowed values for the InstrumentType element in the SPASE data model. For example, a magnetometer would have the value of "Magnetometer". For complex instrument suites, it is recommended to add an extra level in the Instrument Type to represent the suite. For example, .../SuiteName/Subinstrument/...

Cadence: The nominal cadence of the observed values expressed as an ISO 8601 duration. For example, data with a 1 second resolution would have a *cadence* of "PT1S". Commas in fractional times must be replaced with a decimal point (see note under "Character Limitations")

Special Note: The *cadence* can be replaced with any value which reflects the organization of the underlying data provided that the URI remains unique. For example, the year of observation (i.e. 2008) or other similar multi-part value (i.e. 2008/October).

The described template for Resource Identifiers is applied up to the appropriate level. For example, Resource Identifiers for Observatory resources stop at *Observatory* and for Instrument resources stop at *Instrument*.

Character Limitations

The Augmented Backus–Naur Form (ABNF) for a Resource Identifier is:

```
ALPHA = A-Z / a-z
DIGIT = 0-9
segment = ALPHA / DIGIT / "-" / "." / "_"
scheme = segment
authority = segment
path = segment *( "/" segment)

URI = scheme "://" authority "/" path
```

This is consistent with that for a URI, but is more constrained. The restriction on allowed characters affects ISO 8601 durations. With ISO 8601 durations "," notation at the start of the fractional part of a value. For a Resource Identifier the "," must be replaced with a "." notation. For example, the ISO 8601 representation for a 1.5 second duration is "PT1,5S" which is represented in a Resource Identifier as "PT1.5S".

A Full Example

A Resource Identifier for NumericalData from the magnetometer at the CRT observatory in the IGPP/LANL project which has an observatory cadence of 1 second would be:

```
spase://VMO/NumericalData/IGPPLANL/CRT/Magnetometer/PT1S
```

Special Cases

There are two special cases of Resource Identifiers which have slightly different formation rules.

Person Resources

A person resource is assigned a Resource Identifier where the portion after the ResourceType is the name of the person represented in the following format:

```
firstName [" MiddleInitial] "." LastName
```

Using a "." to separate names allows for hyphenated names to properly represented.

So "John W. Smith" would have a Person Resource Identifier assigned by the VMO naming authority of:

```
spase://VMO/Person/John.W.Smith
```

When name conflicts occur, a unique identifier can be created by eliminating the middle initial or adding a sequence number to the end of the name. For example, another John W. Smith may be assigned the resource identifier of:

spase://VMO/Person/John.W.Smith-2

Granule Resources

Granule resources are augmentations to a parent resource and preserve this pedigree in the assigned Resource Identifier by appending a node to the Resource Identifier of the Parent. For example, if a Granule with the name of "2008" is part of the NumericalData resource with the resource identifier of:

spase://VMO/NumericalData/IGPPLANL/CRT/Magnetometer/PT1S

it would be assigned a Resource Identifier of:

spase://VMO/NumericalData/IGPPLANL/CRT/Magnetometer/PT1S/2008

References

RFC3986, Uniform Resource Identifier (URI): Generic Syntax, January 2005,
<http://tools.ietf.org/html/rfc3986>

ISO 8601 – 2004(E), Data elements and interchange formats — Information interchange — Representation of dates and times, Third edition, 2004-12-01.

History

Original: October 3, 2008 – Todd King

Updated: September 24, 2009 – Todd King