Dataset Usage Vocabulary

Bernadette Farias Lóscio[1], Eric Stephan[2], Sumit Purohit[2]

[1] Centro de Informática, Universidade Federal de Pernambuco

[2] Pacific Northwest National Laboratory
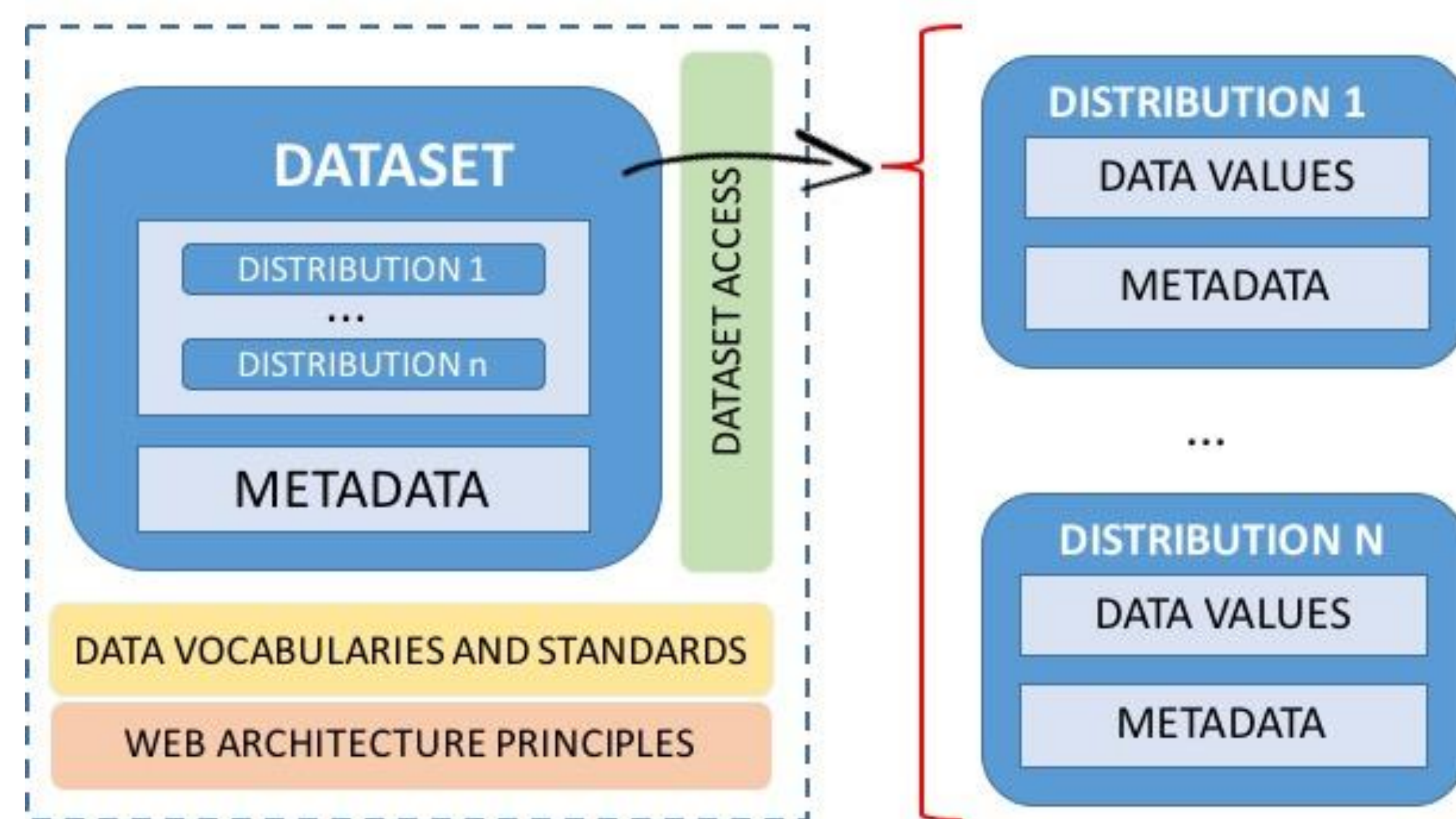
Pacific Northwest
NATIONAL LABORATORY

Proudly Operated by Battelle Since 1965

Objective

As part of the activities of the W3C Data on the Web Best Practices Working Group, the Dataset Usage Vocabulary was developed to address aspects related to the consumption of published Web datasets.
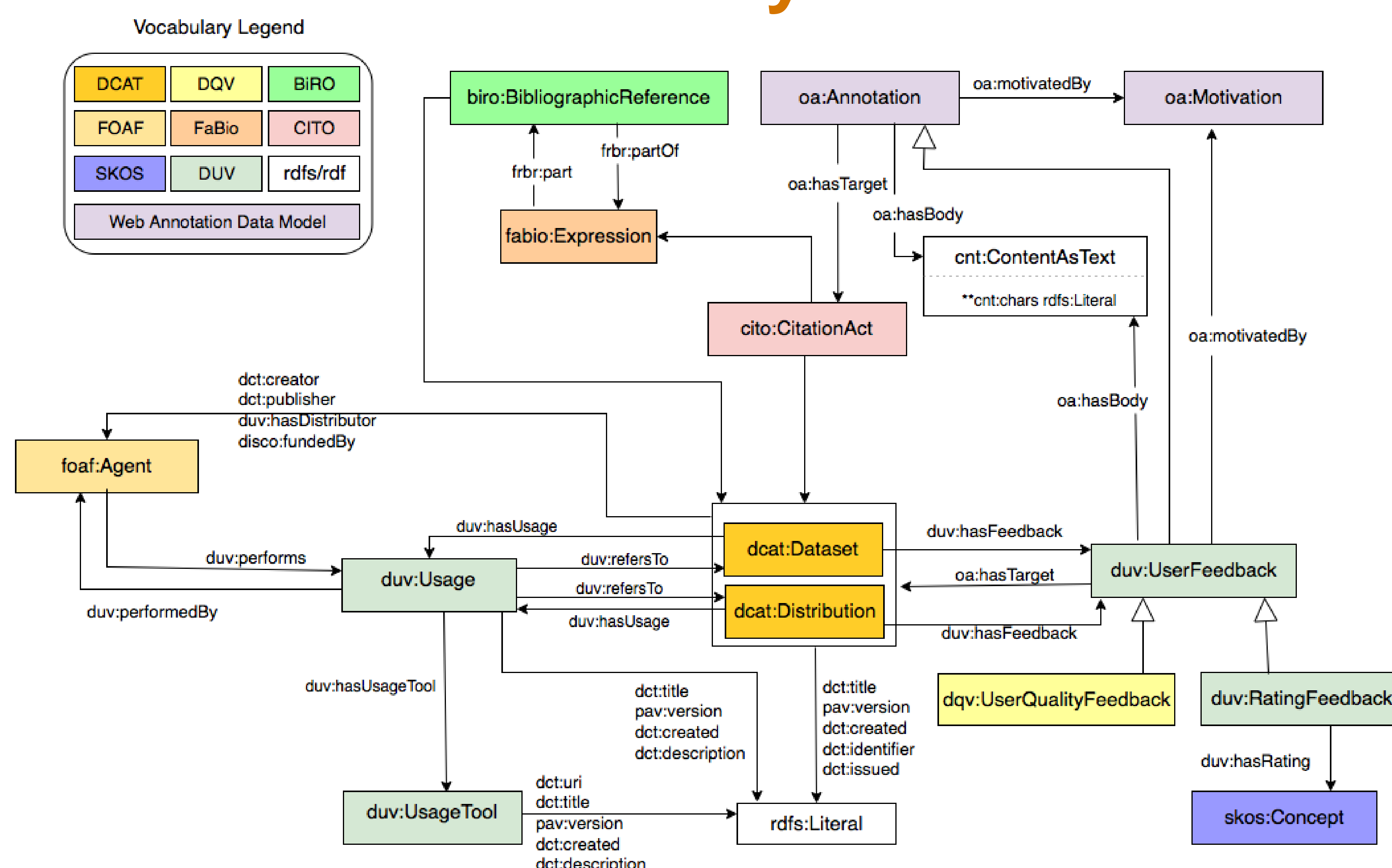
Who is using the dataset?

USAGE	CITATION	FEEDBACK
How can data be used?	*How to cite a dataset?*	*What do users think about the data?*

Context

Motivation

Question	Real-world Challenge
Can the publisher suggest appropriate tools that consumers can use to interact with the datasets?	When data is published there is no available guidance suggested tools to interact with the data sources.
What authoritative source asserted facts about a dataset?	There is no consistent way of citing data sources
As a consumer I've discovered a new soil moisture dataset, can I trust the data providers?	Citing scientific resources thoroughly, viewing community feedback.
What errors have been spotted and corrected in government open data?	To promote public trust transparency through the ability to openly convey and revise data when errors are found.
When I reuse someone else's data how do I formally acknowledge the data producers?	Those who reuse data but were not their publishers don't know how to formally acknowledge their originators.
How do consumers rate or provide feedback to published experimental results?	When data is published on the Web there is no mechanism for users to rate and review data.
Do publishers know what kinds of consumers are using their open experimental results?	Publishers are not aware of consumers using their data.

Vocabulary Overview

Example

```
ex:timetable-001
  a dcat:Dataset;
  dct:title "Bus timetable of MyCity"^^xsd:string;
  prism:doi "10.3456/4567.21"^^xsd:string ;
  pav:version "series-1.2"^^xsd:string;
  dct:issued "2015-MAY-05"^^xsd:date;
  dct:created "2015-MAY-01"^^xsd:date;
  dct:creator <http://example.org/transport-agency/contact> .
```

DCAT Dataset metadata including basic electronic citation criteria.

```
ex:timetable-memorandum-reference
  a biro:BibliographicReference;
  dct:bibliographicCitation
  "Costello, E. Mayor (2015). City Timetable Schedule Change
  Memorandum Effective  January 1, 2015. Sept 15, 2014.
  DOI:0.3456/4567.21"^^xsd:string  .
```

Memorandum bibliographic reference

```
ex:policy-expression
  a  fabio:Policy  ;
  frbr:part ex:timetable_memorandum-reference .
```

Using faBIO to characterize referenced material as a policy.

```
ex:memo-cites-timetable a CitationAct ;
  cito:hasCitingEntity policy-expression ;
  cito:hasCitedEntity ex:timetable-001 .
```

Using CiTO to express that the memorandum cites dataset.

```
ex:comment1Content
  a cnt:ContentAsText ;
  cnt:chars "This timetable is missing route 3"^^xsd:string .

ex:comment1 a duv:UserFeedback ;
  oa:hasBody ex:comment1Content ;
  oa:hasTarget ex:timetable-001;
  dct:creator ex:localresident .
```

Consumer feedback is Is provided through the Web Annotation (OA). Oa:Motivation can be used to provide rationale for annotation.

https://www.w3.org/TR/vocab-duv/