Having your cake and eating it too: JSON-LD as an RDF serialization format

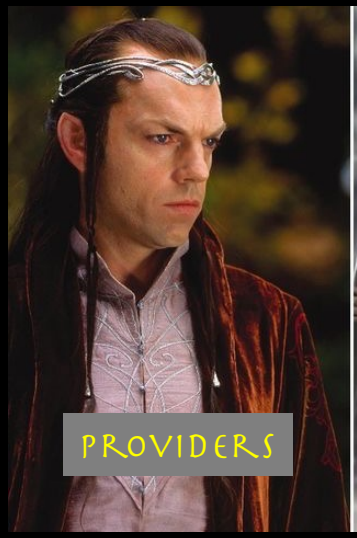
Presenter: Steve Baskauf https://orcid.org/0000-0003-4365-3135
steve.baskauf@vanderbilt.edu
https://doi.org/10.3897/biss.5.74266

Biodiversity Information Standards TDWG

Jean & Alexander Heard LIBRARIES

VANDERBILT UNIVERSITY

THE THREE RINGS OF POWER



TABLES

JSON

RDF

"The three rings, fairest of all, the Elf-lords hid from him, and his hand never touched them or sullied them."

ONE RING TO RULE THEM ALL

JSON-LD

"The Enemy still lacks one thing to give him strength and knowledge to beat down all resistance. … He lacks the One Ring."

- The Fellowship of the Ring, Chapter 2.

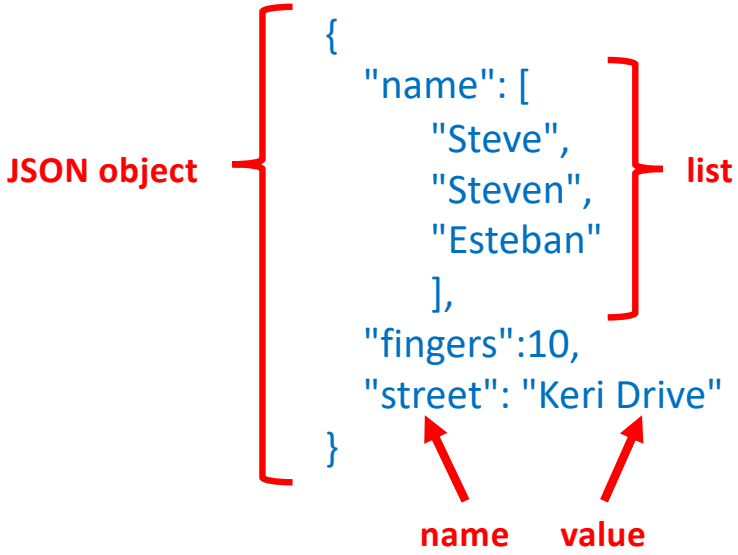
TABLES

JSON

RDF

Background

What is JSON?



- JSON = **J**ava**S**cript **o**bject **n**otation.
- Simple, structured data exchange format.
- Most common data transmission format for APIs.
- Consumed by most programming languages and web pages.

Reference: https://www.json.org/

Reference: https://www.w3.org/TR/json-ld11/

What is JSON-LD?

- JSON-based serialization for Linked Data.
- Valid JSON
- 100% compatible with existing JSON tools.
- Preferred format for structured metadata in web pages.
- A W3C standard and a serialization of RDF.

```
{
  "@context": {
    "dcterms": "http://purl.org/dc/terms/",
    "dcterms:type": {"@type": "@id"}
  },
  "@graph": [
    {
    "@id": "https://www.inaturalist.org/observations/82716069",
    "@type": "http://purl.org/dc/dcmitype/Sound",
    "dcterms:title": "Northern Cardinal song"
  }
  ]
}
```

JSON-LD

=

```
@prefix dcterms: <http://purl.org/dc/terms/> .
<https://www.inaturalist.org/observations/82716069>
   a <http://purl.org/dc/dcmitype/Sound> ;
   dcterms:title "Northern Cardinal song".
```

RDF Turtle

=

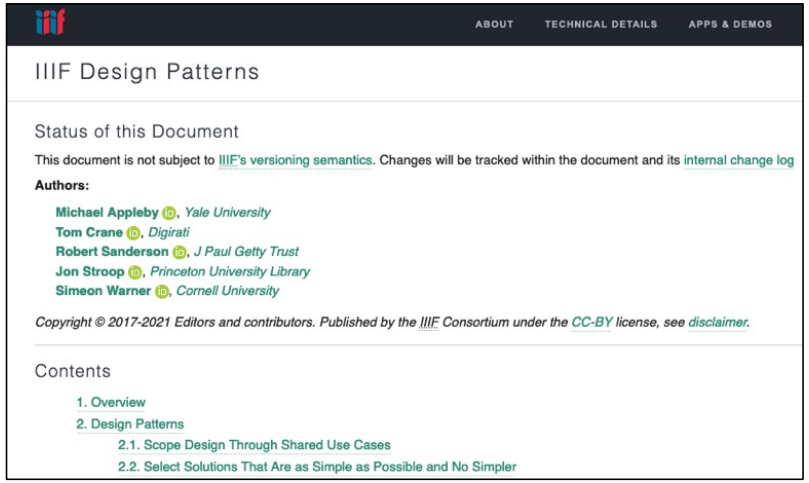
```
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
      xmlns:dcterms="http://purl.org/dc/terms/">
 <rdf:Description rdf:about="https://www.inaturalist.org/observations/82716069">
  <rdf:type rdf:resource="http://purl.org/dc/dcmitype/Sound"/>
  <dcterms:title>Northern Cardinal song</dc:title>
 </rdf:Description>
</rdf:RDF>
```

RDF/XML

Reference: https://iiif.io/api/annex/notes/design_patterns/

What does "narrow design pattern" mean?

Anamsiddiqui125, CC BY-SA 4.0, via Wikimedia Commons



- Inspired by IIIF Design Patterns document.

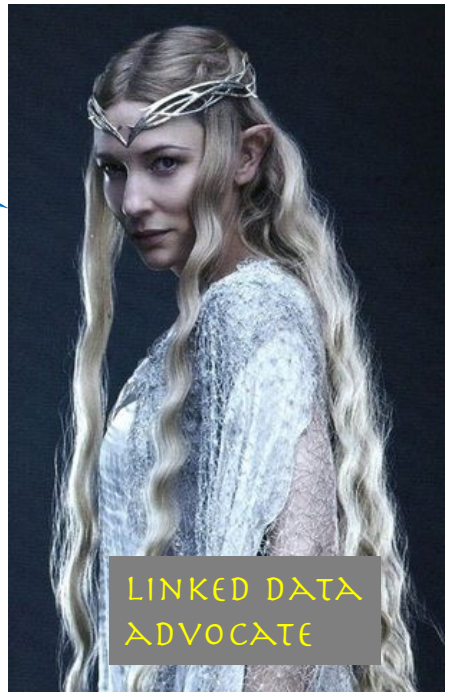
Reference: https://iiif.io/api/annex/notes/design_patterns/

What does "narrow design pattern" mean?

- Section 2.7: "IIF specifications … should not require an RDF based development stack to implement, but it must be possible to implement using one."


```
{
   "@context": {
    "dcterms": "http://purl.org/dc/terms/",
    "dcterms:type": {"@type": "@id"}
   },
   "@graph": [
    {
    "@id":
"https://www.inaturalist.org/observations/82716069",
    "@type": "http://purl.org/dc/dcmitype/Sound",
    "dcterms:title": "Northern Cardinal song"
   }
   ]
}
```

JSON-LD = RDF



Reference: https://iiif.io/api/annex/notes/design_patterns/

What does "narrow design pattern" mean?

- Section 2.8: "Developers must be able to treat the representation as plain JSON, with a predictable structure."

```
{
   "@context": {
    "dcterms": "http://purl.org/dc/terms/",
    "dcterms:type": {"@type": "@id"}
   },
   "@graph": [
    {
    "@id":
"https://www.inaturalist.org/observations/82716069",
    "@type": "http://purl.org/dc/dcmitype/Sound",
    "dcterms:title": "Northern Cardinal song"
   }
   ]
}
```

JSON-LD = JSON

Reference: https://iiif.io/api/annex/notes/design_patterns/

What does "narrow design pattern" mean?

- Section 3: describes restrictions for predictable structure and familiar patterns.

Wow! That looks something like a row from my table !!!

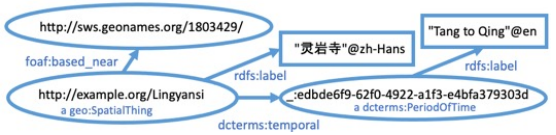
...
"dcterms:title": "Northern Cardinal song",
"dcterms:identifier": "https://www.inaturalist.org/observations/82716069",
"dcterms:type": "http://purl.org/dc/dcmitype/Sound",
"dc:rights": "(c) 2021 cecildev8n5",
"xmp:CreateDate": "2021-06-12T14:08:10.3-04:00",
...
"dc:format": "audio/mp4",
"ac:mediaDuration": 30.186,
"ac:mediaSpeed": 1.0
...

JSON-LD = TABULAR STRUCTURE-LIKE

Transformation between formats

- A narrow JSON-LD design pattern facilitates:
  - having cake (Linked Data) and eating it (consumable JSON)
  - lossless and self-describing transformation between formats

RDF graph



TRIPLE STORE

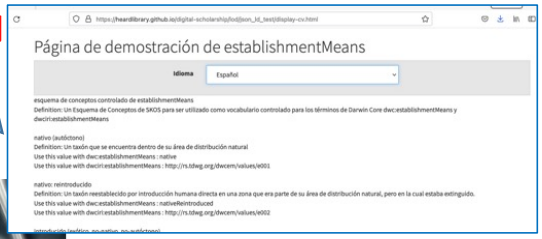
Image from Clipart Panda

conversion script

Image from Clipart Kid

SPARQL Update (have cake)

application consumes JSON (eat cake)



WEB PAGE

E	F	G	H
dcterm:description	dwc:scientificName	dwc:identifiedBy	dwc:dateIdentified
song of Northern Cardinal	Cardinalis cardinalis	cecildev8n5 | Caleb Helsel	2021-06-12
cicada song	Neotibicen linnei	Steven J. Baskauf	2021-06-14
song of Northern Cardinal	Cardinalis cardinalis	cecildev8n5 | Caleb Helsel	2021-06-12
song of Northern Cardinal	Cardinalis cardinalis	cecildev8n5 | Caleb Helsel	2021-06-12

database table

```
{
  "@context": {
    "dcterms": "http://purl.org/dc/terms/",
    "dcterms:type": {"@type": "@id"}
  },
  "@graph": [
    {
    "@id": "https://www.inaturalist.org/observations/82716069",
    "@type": "http://purl.org/dc/dcmitype/Sound",
    "dcterms:title": "Northern Cardinal song"
    }
  ]
}
```

JSON-LD

Design patterns for TDWG

Overarching principles

Reference: http://rs.tdwg.org/dwc/terms/simple/

1. The "Simple Darwin Core" principle:
   - The main JSON object should be mappable to a row in a table.
   - The names (i.e. keys) of the name/value pairs should be the "term names" of the properties denoted by the column headers.
   - The values of the name/value pairs should be the data values in the row representing the main object.

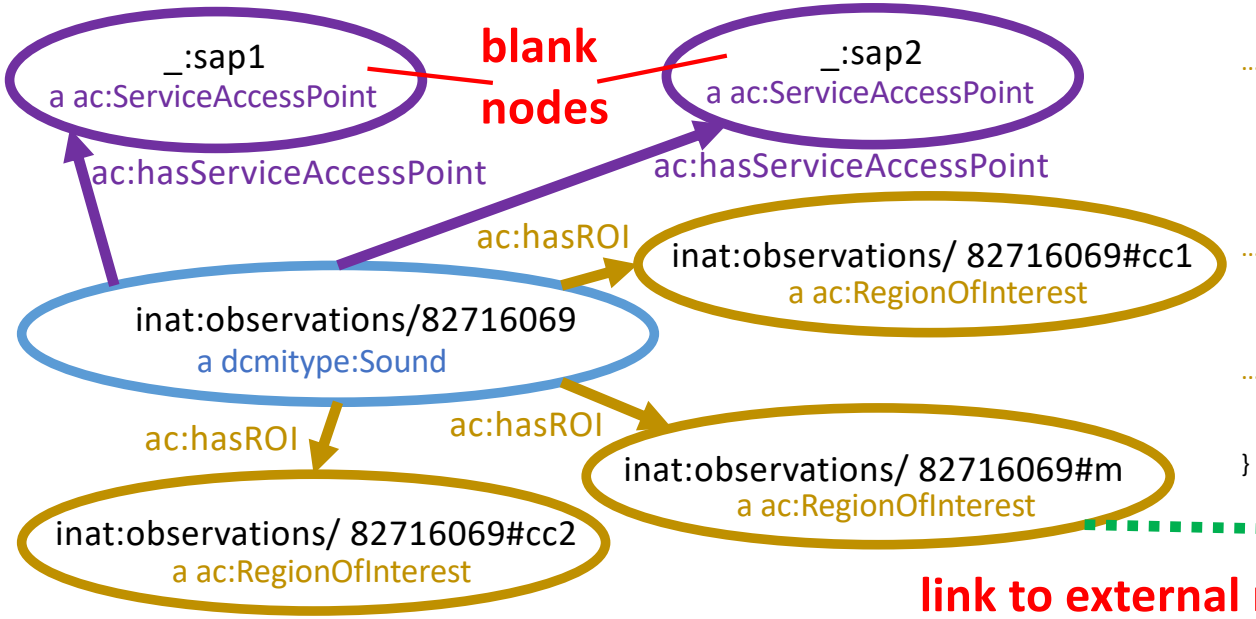

```
{
  "@id": "https://www.inaturalist.org/observations/82716069",
  "@type": "http://purl.org/dc/dcmitype/Sound",
  "dcterms:title": "Northern Cardinal (Cardinalis cardinalis) on June 12, 2021 at 02:08 PM by cecildev8n5",
  "dcterms:identifier": "https://www.inaturalist.org/observations/82716069",
  "dcterms:type": "http://purl.org/dc/dcmitype/Sound",
  "ac:metadataLanguage": "http://id.loc.gov/vocabulary/iso639-2/eng",
  "dc:rights": "(c) 2021 cecildev8n5",
  "xmp:CreateDate": "2021-06-12T14:08:10.3-04:00"
}
```


	A	B	C	D	E	F
1	dcterms:title	dcterms:identifier	dcterms:type	ac:metadataLanguage	dc:rights	xmp:CreateDate
2	Northern Cardinal (Cardinalis cardinalis	https://www.inaturalist.org/observations/82716069	http://purl.org/dc/dcmitype/Sound	http://id.loc.gov/vocabulary/iso639-2/eng	(c) 2021 cecildev8n5	2021-06-12T14:08:10.3-04:00

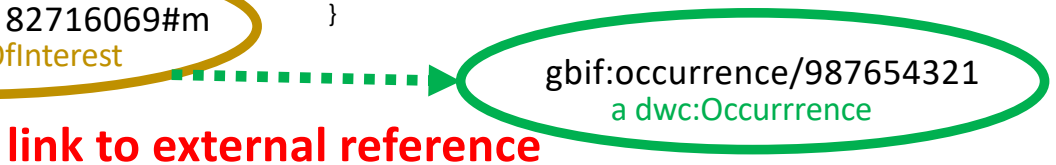
Overarching principles

Reference: http://rs.tdwg.org/dwc/terms/guides/text/

2. The "star schema" principle:
   - No more than one level of nesting
   - Facilitate 1:many relationships to unidentified objects
   - Other links by reference.




```
{
  "@id": "https://www.inaturalist.org/observations/82716069",
  "@type": "http://purl.org/dc/dcmitype/Sound",
... more metadata
  "ac:hasServiceAccessPoint": [
    {
... anonymous SAP 1
    },
    {
... anonymous SAP 2
    }
  ],
  "ac:hasROI": [
    {
      "@id": "https://www.inaturalist.org/observations/82716069#cc1",
... more metadata
    },
    {
      "@id": "https://www.inaturalist.org/observations/82716069#m",
      "ac:associatedObservationReference":
                    "https://www.gbif.org/occurrence/987654321",
... more metadata
    },
    {
      "@id": "https://www.inaturalist.org/observations/82716069#cc2",
... more metadata
    }
  ]
}
```

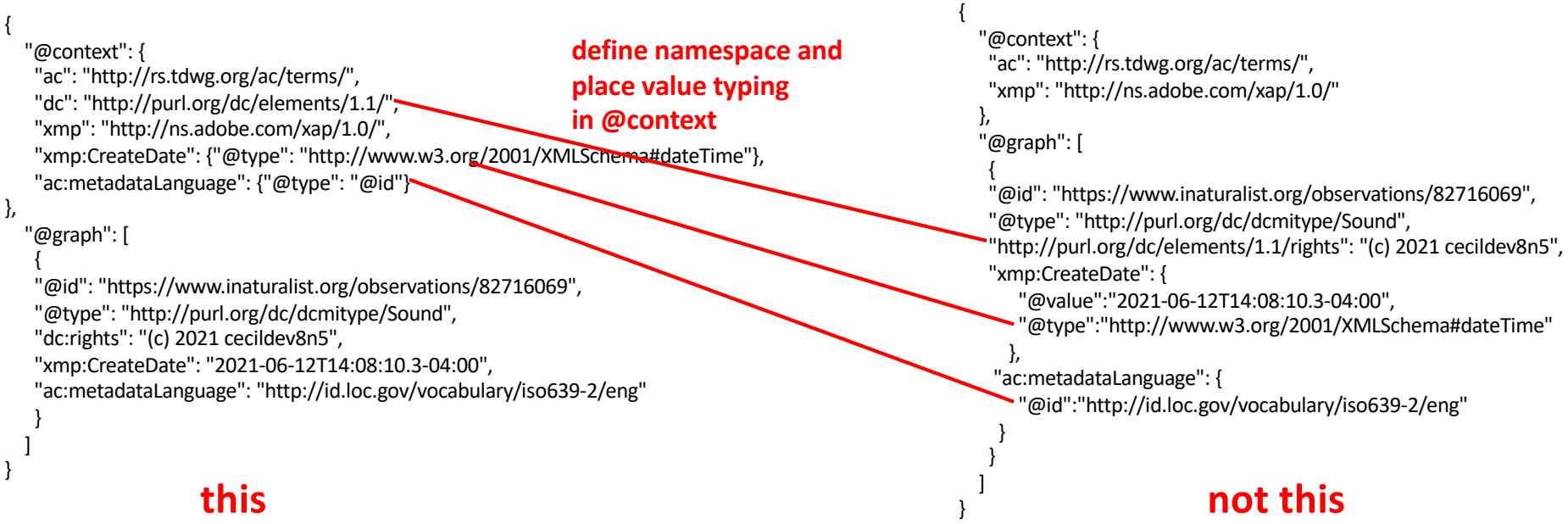


link to external reference

Overarching principles

3. Simple but self-describing:

- Value structure should be simple.
- Semantics should be described in the @context section.
- The @context section should be included in the document.



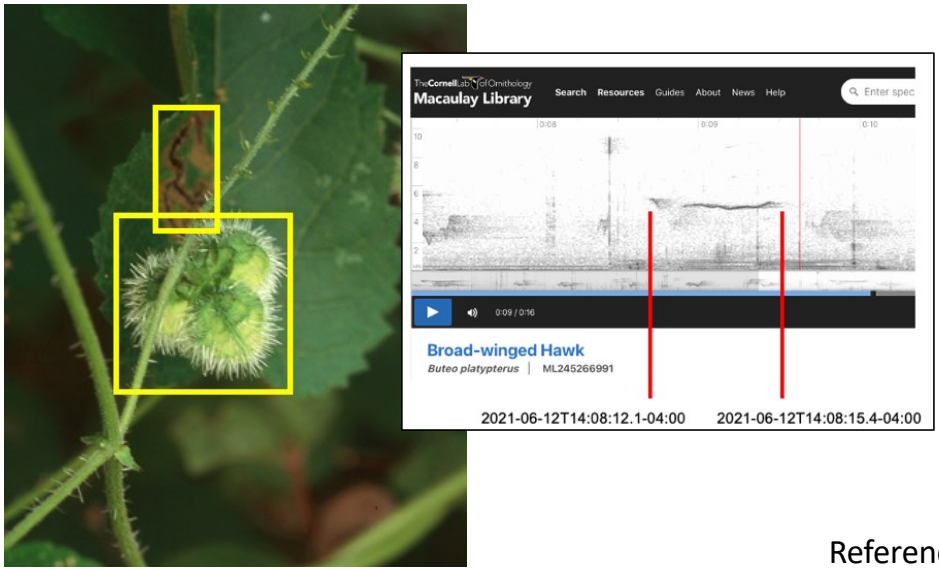
Design pattern details in a 11-point document

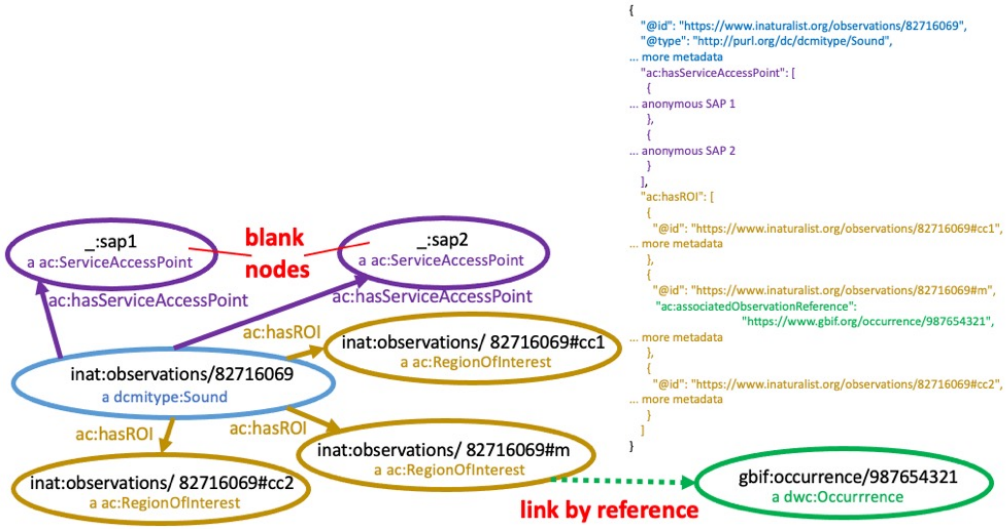
- [bundled with this PDF]

Three examples

1. Audubon Core Regions of Interest (ROIs)

- Refer to specific parts of media items.
- 1:many relationship suitable for design pattern
- "Recipes" document with examples: https://github.com/tdwg/ac/blob/master/roi-recipes.md





Reference: http://rs.tdwg.org/ac/doc/termlist/#711-region-of-interest-vocabulary

2. Multilingual controlled vocabularies

F1:F1048576 term_deprecated

	A	B	C	D	E	F	G	H	I	J	K	L	
1	document_modified	term_localName	term_isDefinedBy	term_created	term_modified	term_deprecated	replaces_term	replaces	replaces	label	definition	usage	notes
2	2020-10-17T11:10:09-04:00	e	http://rs.tdwg.org/dwcem/values/	2020-10-13	2020-10-13					establishmentMeans controlled concep	A SKOS Concept Scheme to be used as a controlled vocabulary for the Darwin Core terms	dwc:	establishr
3	2020-10-17T11:10:09-04:00	e001	http://rs.tdwg.org/dwcem/values/	2020-10-13	2020-10-13					native (indigenous)	A taxon occurring within its natural range		What is
4	2020-10-17T11:10:09-04:00	e002	http://rs.tdwg.org/dwcem/values/	2020-10-13	2020-10-13					native: reintroduced	A taxon re-established by direct introduction by humans into an area which was once part of its n		Where a
5	2020-10-17T11:10:09-04:00	e003	http://rs.tdwg.org/dwcem/values/	2020-10-13	2020-10-13					introduced (alien, exotic, non-native, no	Establishment of a taxon by human agency into an area that is not part of its natural range		Organis
6	2020-10-17T11:10:09-04:00	e004	http://rs.tdwg.org/dwcem/values/	2020-10-13	2020-10-13					introduced: assisted colonisation	Establishment of a taxon specifically with the intention of creating a self-sustaining wild populatio		In the e
7	2020-10-17T11:10:09-04:00	e005	http://rs.tdwg.org/dwcem/values/	2020-10-13	2020-10-13					vagrant (casual)	The temporary occurrence of a taxon far outside its natural or migratory range.		Natural
8	2020-10-17T11:10:09-04:00	e006	http://rs.tdwg.org/dwcem/values/	2020-10-13	2020-10-13					uncertain (unknown, cryptogenic)	The origin of the occurrence of the taxon in an area is obscure		When th
9													

+

	A	B	C	D	E	F	G
1	term_localN	label_en	label_es	label_nl	definition_en	definition_es	definition_nl
2	e	establishmentMeans controlled conc	esquema de conceptos controlado de establ	establishmentMeans gecontroleerd	A SKOS Concept Scheme to be used as a cor	Un Esquema de Conceptos de SKOS para ser uti	Een SKOS Conceptschema te gebruiken als een gecontolee
3	e001	native (indigenous)	nativo (autóctono)	inheems	A taxon occurring within its natural range	Un taxón que se encuentra dentro de su área de d	Een taxon dat voorkomt binnen zijn natuurlijk verspreidingsg
4	e002	native: reintroduced	nativo: reintroducido	inheems: geherintroduceerd	A taxon re-established by direct introduction by	Un taxón reestablecido por introducción humana d	Een taxon dat zich opnieuw vestigde door directe introductie
5	e003	introduced (alien, exotic, non-native,	introducido (exótico, no-nativo, no-autóctono	geïntroduceerd (uitheems, niet-inhe	Establishment of a taxon by human agency int	Establecimiento de un taxón por acción humana e	Vestiging van een taxon door menselijk handelen in een gebi
6	e004	introduced: assisted colonisation	introducido: colonización asistida	geïntroduceerd: geassisteerde kolo	Establishment of a taxon specifically with the a	Establecimiento de un taxón específicamente cor	Vestiging van een taxon specifiek met het doel om een zich
7	e005	vagrant (casual)	vagabundo (incidental)	zwervend (incidenteel)	The temporary occurrence of a taxon far outsid	La presencia temporaria de un taxón muy por afue	Het tijdelijk voorkomen van een taxon ver buiten zijn natuurlij
8	e006	uncertain (unknown, cryptogenic)	incierto (desconocido, criptogénico)	onzeker (onbekend, cryptogeen)	The origin of the occurrence of the taxon in an	El origen de la presencia de un taxón en una zona	De oorsprong van het voorkomen van het taxon in een gebie
9							
10							
11							

normative definition +

non-normative translations ->

SKOS JSON-LD

```
{
  "@context": {
    "rdf": "http://www.w3.org/1999/02/22-rdf-syntax-ns#",
    "rdfs": "http://www.w3.org/2000/01/rdf-schema#",
    "skos": "http://www.w3.org/2004/02/skos/core#",
    "xsd": "http://www.w3.org/2001/XMLSchema#"
  },
  "@graph": [
    {
      "@id": "http://rs.tdwg.org/dwcem/values/e",
      "@type": "http://www.w3.org/2004/02/skos/core#ConceptScheme",
      "skos:prefLabel": [
        {
          "@language": "en",
          "@value": "establishmentMeans controlled concept scheme"
        },
        {
          "@language": "es",
          "@value": "esquema de conceptos controlado de establishmentMeans"
        },
        {
          "@language": "nl",
          "@value": "establishmentMeans gecontroleerd conceptschema"
        }
      ],
      "skos:definition": [
        {
          "@language": "en",
          "@value": "A SKOS Concept Scheme to be used as a controlled vocabulary for the Darwin Core t
        },
        {
          "@language": "es",
          "@value": "Un Esquema de Conceptos de SKOS para ser utilizado como vocabulario controlado pa
dwciri:establishmentMeans"
        },
        {
          "@language": "nl",
          "@value": "Een SKOS Conceptschema te gebruiken als een gecontoleerde woordenschat voor de Da
        }
      ]
    },
    {
      "@id": "http://rs.tdwg.org/dwcem/values/e001",
      "@type": "http://www.w3.org/2004/02/skos/core#Concept",
      "rdf:value": "native",
```


Translation workshop Thursday!

Conversion script: https://github.com/tdwg/rs.tdwg.org/blob/gh-pages/cvJson/build_json_ld_for_controlled_vocabularies.ipynb

2. Multilingual controlled vocabularies

- Web application: https://tdwg.github.io/rs.tdwg.org/cvJson/display-cv.html
- Triplestore/example SPARQL query: https://gist.github.com/baskaufs/916a5210e8df75a41174dffdf9573d7d

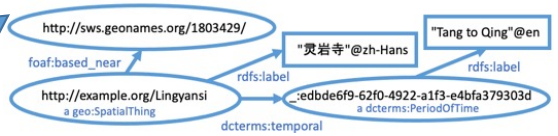
Image from Clipart Panda

E	F	G	H
dcterm:description	dwc:scientificName	dwc:identifiedBy	dwc:dateIdentified
song of Northern Cardinal	Cardinalis cardinalis	cecildev8n5 | Caleb Helsel	2021-06-12
cicada song	Neotibicen linnei	Steven J. Baskauf	2021-06-14
song of Northern Cardinal	Cardinalis cardinalis	cecildev8n5 | Caleb Helsel	2021-06-12
song of Northern Cardinal	Cardinalis cardinalis	cecildev8n5 | Caleb Helsel	2021-06-12

conversion script

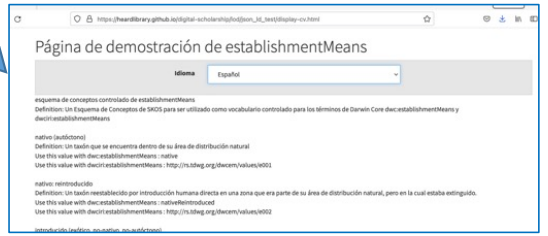
Image from Clipart Kid

load in triple store



TRIPLE STORE

application consumes JSON



WEB PAGE

```
{
  "@context": {
    "dcterms": "http://purl.org/dc/terms/",
    "dcterms:type": {"@type": "@id"}
  },
  "@graph": [
    {
    "@id": "https://www.inaturalist.org/observations/82716069",
    "@type": "http://purl.org/dc/dcmitype/Sound",
    "dcterms:title": "Northern Cardinal song"
    }
  ]
}
```

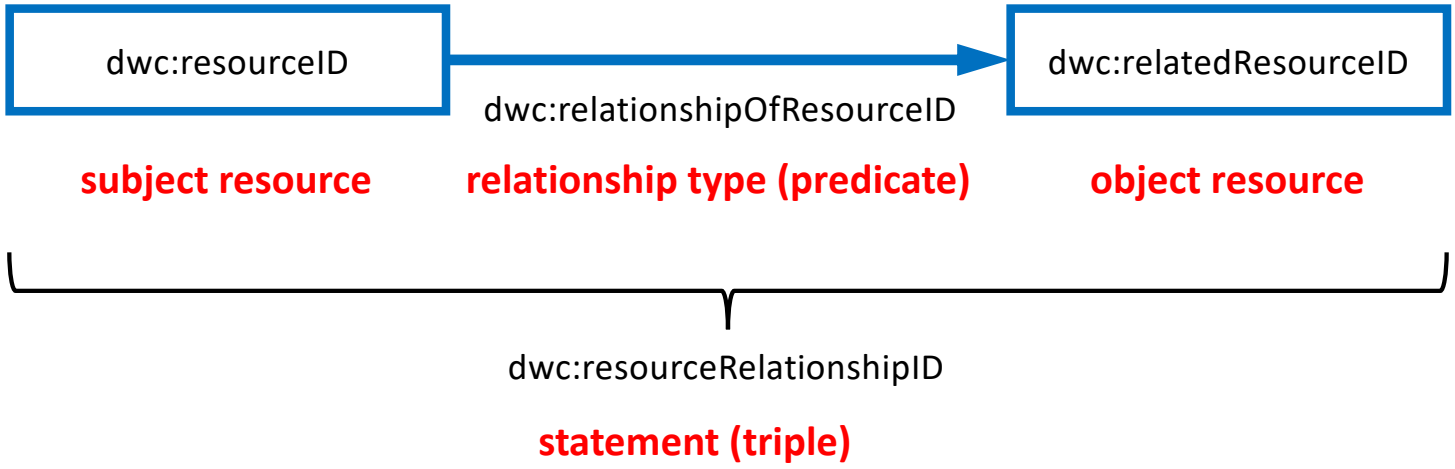
References: http://rs.tdwg.org/dwc/doc/em/ and http://rs.tdwg.org/ac/doc/format/

Reference: https://dwc.tdwg.org/terms/#resourcerelationship
Example data: https://gist.github.com/baskaufs/639d05dd485564a125feaf1b6adaac17

3.Darwin Core ResourceRelationship class: DATABASE

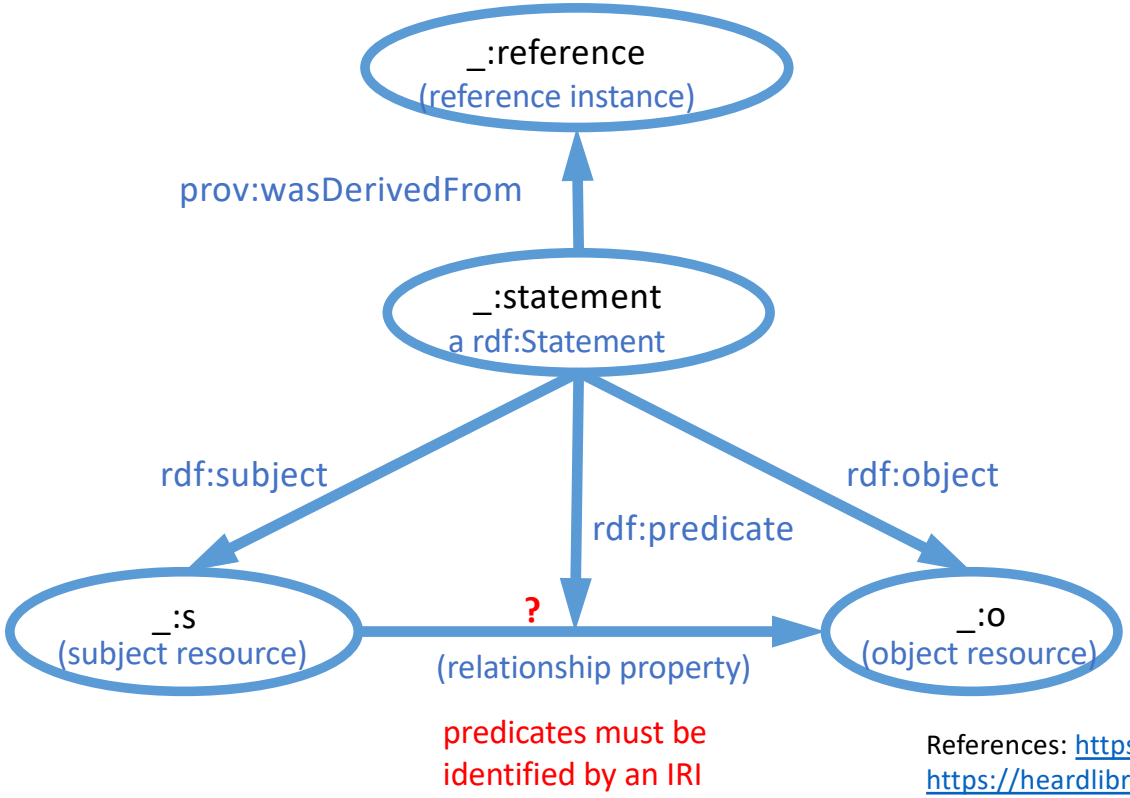
- Newly clarified definitions and new term relationshipOfResourceID
- Used for tabular representation of statements

	A	B	C	D	E	F	G
1	dwc:resourceRelationshipID	dwc:relationshipEstablishedDate	dwc:relationshipRemarks	dwc:resourceID	dwc:relationshipOfResourceID	dwc:relatedResourceID	dwc:relationshipAccordingTo
2	04b16710-b09c-11e8-96f8-529269fb1459	1963-03-08T14:07:00-0600	pollinator captured in the act	f809b9e0-b09b-11e8-96f8-529269fb1459	http://purl.obolibrary.org/obo/RO_0002456	dc609808-b09b-11e8-96f8-529269fb1459	Julie Woodruff



Darwin Core ResourceRelationship class: TRIPLE STORE

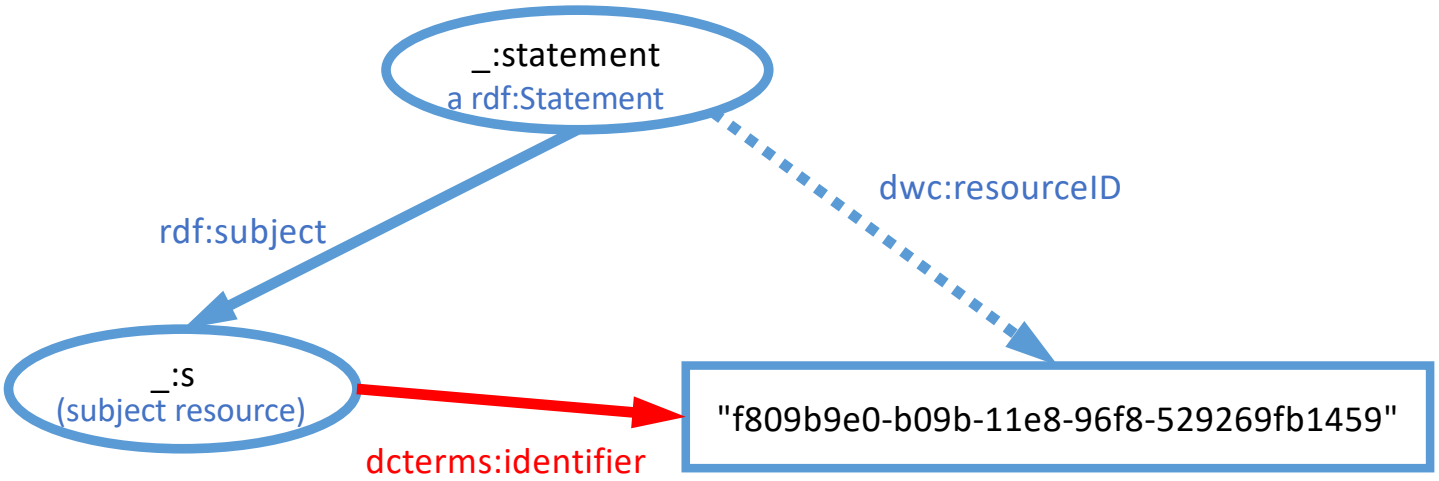
- RDF reification vocabulary model for describing a triple + Wikibase model for references



References: https://www.w3.org/TR/rdf-schema/#ch_reificationvocab
https://heardlibrary.github.io/digital-scholarship/lod/wikibase/#references

Darwin Core ResourceRelationship class: TRIPLE STORE

- dwc:resourceID = "An identifier for the resource that is the subject of the relationship."
- Similar model for rdf:predicate and rdf:object



Darwin Core ResourceRelationship class: API

- JSON-LD maps ResourceRelationship terms to RDF graph
- Generally follows the same design pattern as before.


```
{
  "@context": {
    "rdf": "http://www.w3.org/1999/02/22-rdf-syntax-ns#",
    "dwc": "http://rs.tdwg.org/dwc/terms/",
    "dcterms": "http://purl.org/dc/terms/",
    "dwc:relationshipEstablishedDate": {"@type": "http://www.w3.org/2001/XMLSchema#dateTime"}
  },
  "@graph": [
    {
    "@type": "http://www.w3.org/1999/02/22-rdf-syntax-ns#Statement",
    "dcterms:identifier": "04b16710-b09c-11e8-96f8-529269fb1459",
    "dwc:relationshipEstablishedDate": "1963-03-08T14:07-0600",
    "dwc:relationshipRemarks": "pollinator captured in the act",
    "rdf:subject": {"dcterms:identifier": "f809b9e0-b09b-11e8-96f8-529269fb1459"},
    "rdf:predicate": {"dcterms:identifier": "http://purl.obolibrary.org/obo/RO_0002456"},
    "rdf:object": {"dcterms:identifier": "dc609808-b09b-11e8-96f8-529269fb1459"},
    "prov:wasDerivedFrom": {"rdfs:label": "Julie Woodruff"}
    }
  ]
}
```

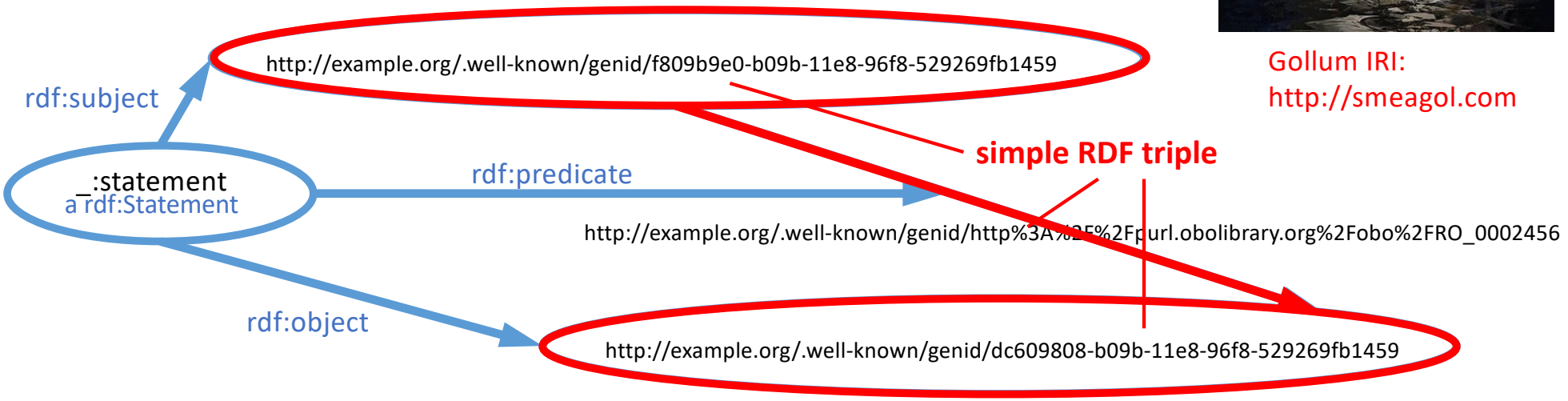
Data at: https://gist.github.com/baskaufs/f8f8ddc3b2161beea789492e350ec374

Darwin Core ResourceRelationship class: TRIPLE STORE

- Use SPARQL CONSTRUCT to generate **Skolem IRIs** from IDs to replace blank nodes and generate a predicate IRI.
- Resulting **RDF triples** easily queried.

```
prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
prefix dcterms: <http://purl.org/dc/terms/>
construct {?subject ?predicate ?object}
from <https://resourcerelationship>
where {
 bind("http://example.org/.well-known/genid/" as ?skolem)
 ?statement rdf:subject/dcterms:identifier ?subjectID.
 ?statement rdf:predicate/dcterms:identifier ?predicateID.
 ?statement rdf:object/dcterms:identifier ?objectID.
 bind(iri(concat(?skolem, encode_for_uri(?subjectID))) as ?subject)
 bind(iri(concat(?skolem, encode_for_uri(?predicateID))) as ?predicate)
 bind(iri(concat(?skolem, encode_for_uri(?objectID))) as ?object)
 }
```

Gollum IRI:
http://smeagol.com



Reference: https://www.w3.org/TR/rdf11-concepts/#section-skolemization
Example query: https://gist.github.com/baskaufs/379a694e1f4e235e3261972995fdfbf0

Image: Frédéric Bennett (Benef), CC BY-SA 4.0, via Wikimedia Commons

Conclusions

Spoiler alert !

THE HAPPY ENDING...

JSON-LD

LINKED DATA ADVOCATES

PROVIDERS

DEVELOPERS

Wait a minute…

TDWG TAG ?

JSON-LD

In our story Sauron is the good guy!

Do we really want interoperability for all parts of our community?