

*Everything you never wanted to know about data  
catalogs, but had to learn anyway*

# Table of Contents





*DRAFT*

This document is only a rough draft.

---

**Preamble**

**About this book**

**About the author**

# Introduction to RDF

# Linked data and the semantic web

semantic

Resource Description Framework

## Statements and triples

*statements*

*triples*

<subject> <predicate> <object>

*Example 1. Example of statements*

```
<John> <buys an> <apple>  
<Jane> <is born in> <Paris>
```

*everything*

September 2015

John has been working for ACME Corp since

```
<John> <is an employee in> <Contract>  
<ACME Corp> <is an employer in> <Contract>  
<Contract> <starts in> "9/2015"
```

subject

object



# Identifiers

*link* *identify*

## URI and URLs

NOTE	<a href="#">example.com</a>			
NOTE	<a href="http://example.com">http://example.com</a>	<a href="https://example.com">https://example.com</a>	s	not

## Dereferencable URIs

		Accept	
NOTE	<i>have</i>		

## Literals

*literal*

## Language tags and data types

*or*

## ... and beyond that

*graph* *context*

---

# Modelling

## Classes and properties

Document Person

NOTE

## Vocabularies and ontologies

*vocabulary*

*ontology*

Organization

Document

## Reusing vocabularies

*find*

## RDF Schema

### Describing classes

### Describing properties

Domain Range

---



# Validation

Paris      Jane lives in Paris      Jane lives in London  
London

## SHACL

## File formats

*serialized*

*Example 2. Set of statements*

```
<vCard> <is a> <Standard>  
<vCard> <has label> "Ontology for vCard"@en  
<vCard> <is published on> "22 May 2014"
```

## N-Triples

grep

*Example 3. N-Triples file*

```
<http://www.w3.org/2006/vcard/ns#> <http://www.w3.org/1999/02/22-rdf-syntax-  
ns#type> <http://purl.org/dc/terms/Standard> .  
<http://www.w3.org/2006/vcard/ns#> <http://www.w3.org/1999/02/22-rdf-syntax-  
ns#label> "Ontology for vCard"@en .
```

```
<http://www.w3.org/2006/vcard/ns#> <http://purl.org/dc/terms/issued> "2014-05-14"^^<http://www.w3.org/2001/XMLSchema#date> .
```

## Turtle

prettier

a

*Example 4. Turtle file*

```
@prefix dct: <http://purl.org/dc/terms/> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .

<http://www.w3.org/2006/vcard/ns#>:
  a dct:Standard ;
  rdf:label "Ontology for vCard"@en ;
  dct:issued "2014-05-14"^^xsd:date .
```

PREFIX

@prefix

## RDF/XML

*Example 5. RDF/XML file*

```
<?xml version="1.0" encoding="utf-8" ?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dct="http://purl.org/dc/terms/">

  <dct:Standard rdf:about="http://www.w3.org/2006/vcard/ns#">
```

```
<rdf:label xml:lang="en">Ontology for vCard</rdf:label>
< dct:issued rdf:datatype="http://www.w3.org/2001/XMLSchema#date">2014-05-
14</dct:issued>
</dct:Standard>

</rdf:RDF>
```

## **Other**

## **JSON-LD**

## **Embedding in webpages**

### **RDFa in HTML**

## **JSON-LD in HTML**

---

# Storing RDF: Triple stores

*triple stores*

# Querying and updating with SPARQL



# Data Catalog (DCAT)

# DCAT

title simplified model

Catalog - Dataset

Dataset - Distribution

## DCAT 2

## DCAT 3

---

# DCAT-AP

## Catalog

**Describing the catalog**

## Dataset

**Describing the dataset**

dct:title

dct:description

dcat:keyword dcat:theme

**Licenses and rights**

## Distribution

**Accessing and downloading**

dcat:accessURL dcat:downloadURL

## Dataservice

---

# Various specializations

**BRegDCAT-AP**

**CiteDCAT-AP**

**DCAT-AP-JRC**

**DCAT-AP HvD**

**EPOS-DCAT-AP**

**GeoDCAT-AP**

**GNSS-DCAT-AP**

**HealthDCAT-AP**

**MLDCAT-AP**

---

**MobilityDCAT-AP**

**StatDCAT-AP**

# National variants

**Belgium**

**Finland**

**Germany: DCAT-AP.DE**

**Italy: DCAT-AP\_IT**

**The Netherlands: DCAT-AP-DONL**

**Norway: DCAT-AP-NO**

**Poland: DCAT-AP-PL**

**Spain: DCAT-AP-Spain**

**Sweden: DCAT-AP-SE**

**Switzerland: DCAT-AP CH**

---

# Vocabularies

# ADMS

Full name	
Namespace	
Prefix	
See also	
Classes	
Properties	

---



# DCTERMS

Full name	
Namespace	
Prefix	
See also	
Classes	
Properties	

	created	issued	modified
title			description

``creator`, `contributor`, `rightsHolder``

accessRights, `license` rightsLicenseDocumentRightsStatement

conformsTo Standard

accrualPeriodicity  
Frequency

NOTE

# DQV

Full name	
Namespace	
Prefix	
See also	

# FOAF

Full name	
Namespace	
Prefix	
See also	
Classes	
Properties	

Person

Organization

Agent

# GeoSPARQL

Full name	
Namespace	
Prefix	
See also	
Data types	

---

# LOCN

Full name	
Namespace	
Prefix	
See also	
Classes	
Properties	

---

**ORG**

# OWL

Full name	
Namespace	
Prefix	
See also	
Classes	
Properties	

`owl:sameAs`

`skos:exactMatch`

# PROV

Full name	
Namespace	
Prefix	
See also	
Properties	

---



# Schema

Full name	
Namespace	
Prefix	
See also	
Properties	

`schema:startDate`      `schema:endDate`  
`dcat:startDate`and `dcat:endDate`  
  
`dcat:`

# SKOS

Full name	
Namespace	
Prefix	
See also	
Classes	
Properties	

*term*

skos:Concept

skos:broader

skos:narrower

## Labels

skos:prefLabel

skos:altLabel

## Mapping to other thesauri

skos:broadMatch

skos:narrowMatch

skos:closeMatch

skos:exactMatch

# VCARD

Full name	
Namespace	
Prefix	
See also	
Classes	
Properties	

---

# XSD

Full name	
Namespace	
Prefix	
See also	
Data types	

anyURI

dc:landingPage

resource

literal

integer

int

bigint

decimal

float

double

date

dateTime

duration

# Appendix A: More on identifiers

URN, URI, IRI

Persistent URLs (PURL)

Skolem URI

Persons

Organizations

File formats

dct:format

dcat:mediaType dcat:compressFormat dcat:packageFormat

Contribut

---

## Geographic areas

# **Appendix B: Common mistakes in DCAT-AP**

**Not using the EU Publications Office controlled lists**

**Incorrect date or duration formats**

**Mixing contact point IDs and organization IDs**

**Mixing IDs and URLs**

**Not setting explicit language tags**

# Appendix C: Tools



# **RDF libraries**

**Java**

**Javascript / NodeJS**

**Python, PHP**

---

# Triplestores

## Triple

# Other tools