Vocabolari RDF

Dublin Core, FOAF, Schema.org

Dublin Core

Metadati e archivi

- In un <mark>archivio</mark>, i documenti sono descritte in base alle loro proprietà: i metadati
  - Per es. autore, titolo, anno, ecc.

- I documenti/opere (risorse per RDF) sono descritti in un <mark>catalogo</mark>

- Il catalogo è un insieme di '<mark>schede</mark>' che descrivono le risorse

From Catalog Records to Linked Data

- «The Dublin Core, a set of fifteen generic, widely used elements -- Creator, Contributor, Publisher, Title, Date, Language, Format, Subject, Description, Identifier, Relation, Source, Type, Coverage, and Rights -- was first drafted at a 1995 meeting in Dublin, Ohio, initially to facilitate information discovery on an explosively growing Web by embedding simple, card-catalog-like metadata in its pages»

- Evoluzioni successive:
    1. Metadati in formato XML con mapping tra schemi diversi
    2. Formato RDF con apporti da altri vocabolari (es., BIBO, the Bibliographic Ontology)
    3. Linked Data: diffusione pervasiva dei termini di Dublin Core, supporto per l'estrazione della conoscenza

https://www.dublincore.org/resources/metadata-basics/

Identikit di Dublin Core

E' uno schema di metadati

Per annotare risorse testuali e multimediali

Nasce per il Web

E' lo schema di metadati più semplice

Standard ISO 15836

IETF RFC 5013

Ha lo scopo esplicito di permettere il reperimento di risorse:

"Organizations, even small, possess information worth millions, but cannot get it to the right people"

DM Direct Special Report, 2006

Dublin Core

Dublin Core Metadata Element Set (DCMES)

- L'insieme base di 15 descrittori e il loro significato
- http://dublincore.org/documents/dces/

DCTerms

- I termini del vocabolario Dublin Core
- Classi e proprietà con namespace dc:
  http://purl.org/dc/elements/1.1/
  http://purl.org/dc/elements/1.1/creator
- http://dublincore.org/documents/dcmi-terms/
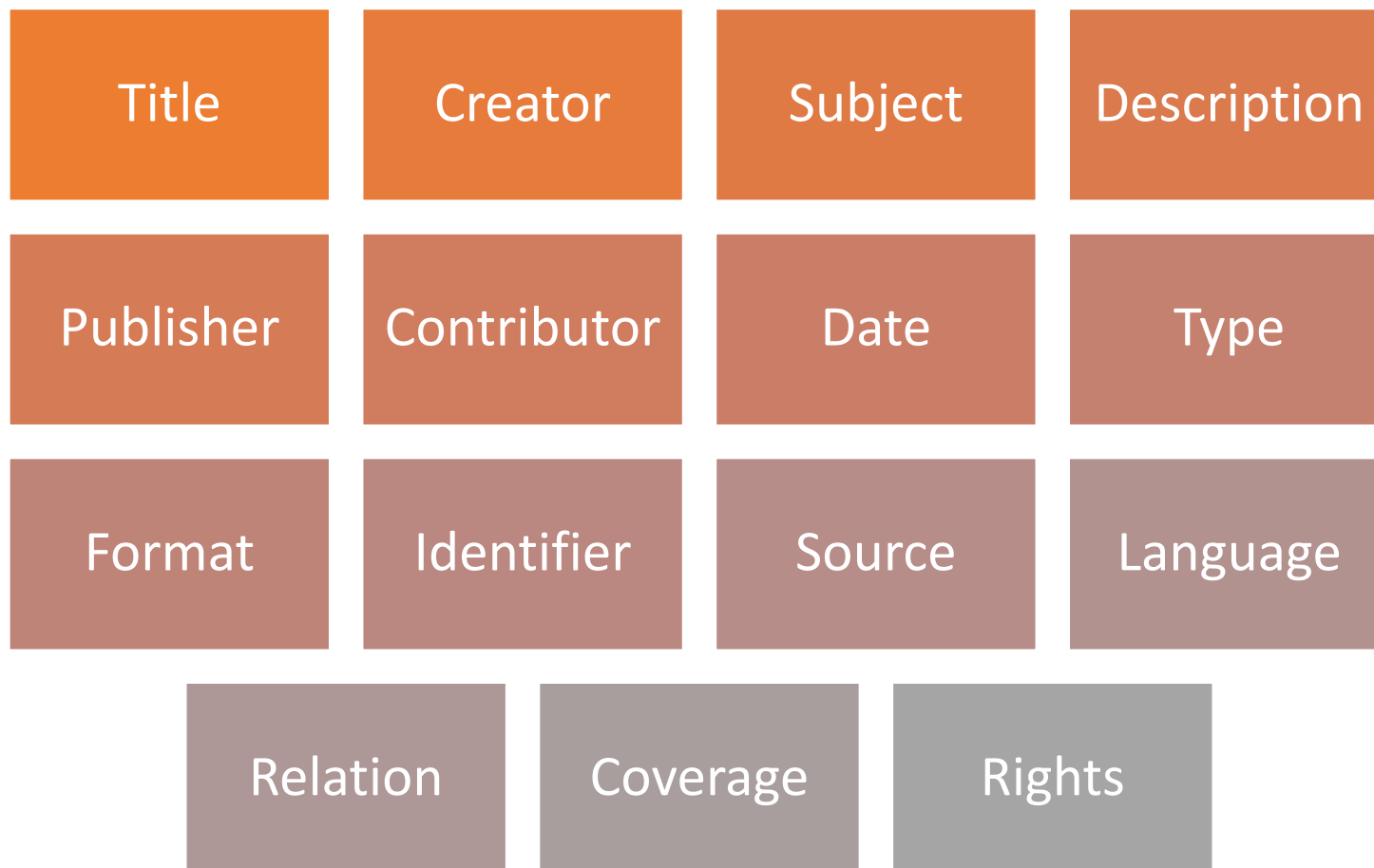
Cosa si può descrivere

DC descrive risorse: "anything that has identity"

- RFC2396 (URI)
- "a resource will typically be an information or service resource, but may be applied more broadly"

Tipi di risorsa:

- Collection, Dataset, Event, Image (Still or Moving), Interactive Resource, Service, Software, Sound, Text, Physical Object.

Elementi di Dublin Core Metadata Element Set (MES)

Title	Creator	Subject	Description
Publisher	Contributor	Date	Type
Format	Identifier	Source	Language
Relation	Coverage	Rights	

dc:creator

Term Name: creator	
URI:	http://purl.org/dc/elements/1.1/creator
Label:	Creator
Definition:	An entity primarily responsible for making the resource.
Comment:	Examples of a Creator include a person, an organization, or a service. Typically, the name of a Creator should be used to indicate the entity.

- «Domains and ranges specify what kind of described resources and value resources are associated with a given property.»

- «Domains and ranges express the meanings implicit in natural-language definitions in an explicit form that is usable for the automatic processing of logical inferences.»

- «When a given property is encountered, an inferencing application may use information about the domains and ranges assigned to a property in order to make inferences about the resources described thereby.»

ragionamento RDF: https://www.w3.org/TR/rdf11-mt/

Specifica dei <mark>termini</mark>

Name:	A token appended to the URI of a DCMI namespace to create the URI of the term.
Label:	The human-readable label assigned to the term.
URI:	The Uniform Resource Identifier used to uniquely identify a term.
Definition:	A statement that represents the concept and essential nature of the term.
Type of Term:	The type of term as described in the DCMI Abstract Model [DCAM].

Comment:	Additional information about the term or its application.
See:	Authoritative documentation related to the term.
References:	A resource referenced in the Definition or Comment.
Refines:	A Property of which the described term is a Sub-Property.
Broader Than:	A Class of which the described term is a Super-Class.
Narrower Than:	A Class of which the described term is a Sub-Class.
Has Domain:	A Class of which a resource described by the term is an Instance.
Has Range:	A Class of which a value described by the term is an Instance.
Member Of:	An enumerated set of resources (Vocabulary Encoding Scheme) of which the term is a Member.
Instance Of:	A Class of which the described term is an instance.
Version:	A specific historical description of a term.
Equivalent Property:	A Property to which the described term is equivalent.

Term definition: creator (property)

Term Name: creator **More details**	
URI:	http://purl.org/dc/terms/creator
Label:	Creator
Definition:	An entity primarily responsible for making the resource.
Comment:	Examples of a Creator include a person, an organization, or a service.
Type of Term:	http://www.w3.org/1999/02/22-rdf-syntax-ns#Property
Has Range:	http://purl.org/dc/terms/Agent
Refines:	•http://purl.org/dc/elements/1.1/creator •http://purl.org/dc/terms/contributor

http://purl.org/dc/terms/creator

Term definition: Agent (class)

Term Name: Agent	
URI:	http://purl.org/dc/terms/Agent
Label:	Agent
Definition:	A resource that acts or has the power to act.
Comment:	Examples of Agent include person, organization, and software agent.
Type of Term:	Class
Instance Of:	http://purl.org/dc/terms/AgentClass
Version:	http://dublincore.org/usage/terms/history/#Agent-001

Pubblicazione dei metadati in RDF

ex:myFurniture dcterms:title "Alvar Aalto Chair No. 66" .

ex:mySong dcterms:title "Autumn Leaves", "The Dead Leaves" .

ex:myPainting dcterms:title
      "La Joconde"@fre , "Mona Lisa"@eng , "La Gioconda"@ita .

ex:myDatabase dcterms:title "Data from my last evaluation" ;
     dcterms:accessRights [rdfs:label "My colleagues only"] .

ex:myBook dcterms:creator _shakespearesName .
  _:shakespearesName foaf:familyName "Shakespeare" ; foaf:givenName "William" .

https://www.dublincore.org/resources/userguide/publishing_metadata/

Europeana

- https://www.europeana.eu

- Archivio europeo dei beni culturali

- I dati sono descritti in un formato basato su Dublin Core

- La semantica con cui sono descritti i dati permette di creare collezioni in modo semi-automatico

OAI-PMH

- The Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) is a low-barrier mechanism for repository **interoperability**.

- Data Providers are repositories that **expose** structured metadata via OAI-PMH.

- Service Providers then make OAI-PMH service requests to **harvest** that metadata.

- OAI-PMH is a set of six verbs or **services** that are invoked within HTTP.

- https://www.openarchives.org/OAI/openarchivesprotocol.html

Interrogare Europeana (esempio)

https://www.europeana.eu

title	date	creator
Coplas flamencas de autor	2004	Picasso, Pablo (1881-1973)
Peinture contemporaine dans les collections vaudoises	1971	Picasso, Pablo
Primer ciclo de 5 conciertos	2008	Forum Musical; Picasso, Pablo, 1881-1973
Primer ciclo de 5 conciertos	1966-03-04	Forum Musical; Picasso, Pablo, 1881-1973
Pablo Picasso	unknown	Picasso, Pablo

La query effettuata si avvale dei termini DC:

SELECT DISTINCT ?title ?date ?creator
WHERE { ?opera **dc:title** ?title; **dc:creator** ?creator; **dc:date** ?date
FILTER (regex(?creator, 'Picasso'))

Friend of a Friend

FOAF

FOAF CORE

- Agent
- Person
- name
- title
- img
- depiction (depicts)
- familyName
- givenName
- knows
- based_near
- age
- made (maker)
- primaryTopic (primaryTopicOf)
- Project
- Organization
- Group
- member
- Document
- Image

SOCIAL WEB

- nick
- mbox
- homepage
- weblog
- openid
- jabberID
- mbox_sha1sum
- interest
- topic_interest
- topic (page)
- workplaceHomepage
- workInfoHomepage
- schoolHomepage
- publications
- currentProject
- pastProject
- account
- OnlineAccount
- accountName
- accountServiceHomepage
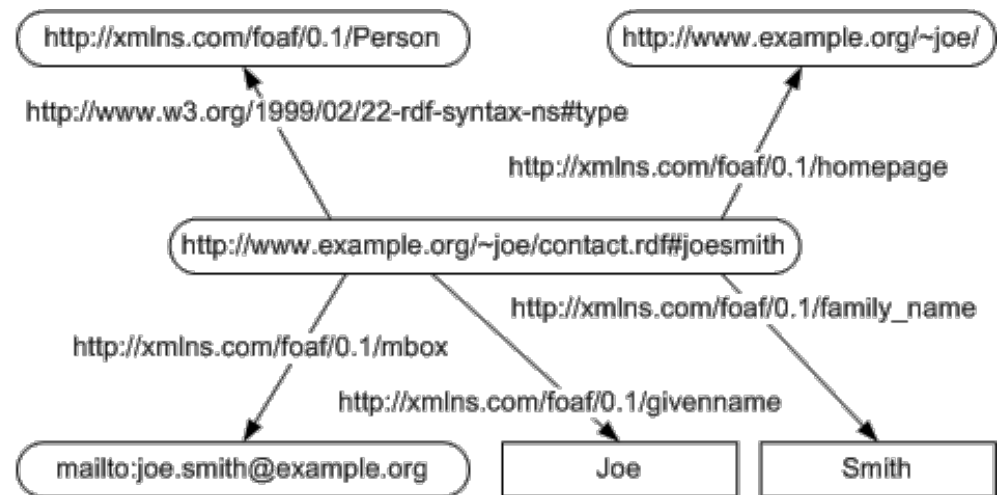- PersonalProfileDocument
- tipjar
- sha1
- thumbnail
- logo

Example class: Agent

Agent - An agent (eg. person, group, software or physical artifact).

Status:	stable
Properties include:	gender yahooChatID account birthday icqChatID aimChatID jabberID made mbox interest tipjar skypeIDtopic_interest age mbox_sha1sum status msnChatID openid holdsAccount weblog
Used with:	maker member
Has Subclass	Group Person Organization

Friend of a Friend

- Documentazione:
  http://xmlns.com/foaf/spec/

- Namespace:
  http://xmlns.com/foaf/0.1/

Schema.org

schema

Schema.org

- «Schema.org is a collaborative, community activity with a mission to create, maintain, and promote schemas for structured data on the Internet, on web pages, in email messages, and beyond.»

- «Founded by Google, Microsoft, Yahoo and Yandex, Schema.org vocabularies are developed by an open [community](#) process»

Schemi

Types

- Il vocabolario è suddiviso in un insieme di schemi (o «types»), organizzati in una tassonomia

- Ereditarietà

- Thing
  - Action
  - CreativeWork
  - Event
  - Intangible
  - Organization
  - …

Eterogeneità

- Organization
  - Airline
  - Consortium
  - Corporation
  - EducationalOrganization
    - CollegeOrUniversity
    - ElementarySchool
    - HighSchool
    - MiddleSchool
    - Preschool
    - School
  - LocalBusiness
    - AnimalShelter
    - ArchiveOrganization
    - AutomotiveBusiness
    - …

Esempio: Movie

Property	Expected Type	Description
Properties from Movie		
actor	Person	An actor, e.g. in tv, radio, movie, video games etc., or in an event. Actors can be associated with individual items or with a series, episode, clip. Supersedes actors.
countryOfOrigin	Country	The country of the principal offices of the production company or individual responsible for the movie or program.
director	Person	A director of e.g. tv, radio, movie, video gaming etc. content, or of an event. Directors can be associated with individual items or with a series, episode, clip. Supersedes directors.
duration	Duration	The duration of the item (movie, audio recording, event, etc.) in ISO 8601 date format.
musicBy	MusicGroup or Person	The composer of the soundtrack.
productionCompany	Organization	The production company or studio responsible for the item e.g. series, video game, episode etc.
subtitleLanguage	Language or Text	Languages in which subtitles/captions are available, in IETF BCP 47 standard format.
trailer	VideoObject	The trailer of a movie or tv/radio series, season, episode, etc.

Anche proprietà ereditate (es. da CreativeWork)