

# Semantic Web

RDF(S)

# Storia del Semantic Web

- ▶ Il Web è stato “inventato” (tra gli altri) da **Tim Berners-Lee**, un fisico che lavora al CERN
- ▶ La sua visione originale del Web era assai più ambiziosa di quella che è la realtà del Web (sintattico) esistente:



“... a goal of the Web was that, if the interaction between person and hypertext could be so intuitive that the **machine-readable** information space gave an accurate representation of the state of people's thoughts, interactions, and work patterns, then **machine analysis** could become a very powerful management tool, seeing patterns in our work and facilitating our working together through the typical problems which beset the management of large organizations.”

# Storia (continua)

- ▶ TBL (e altri) da allora stanno lavorando alla realizzazione di quella visione, che è divenuta nota come **Semantic Web**
  - Es., articolo nell'edizione di maggio 2001 di Scientific American

Scientific American, May 2001:



PHOTO CREDIT HERE

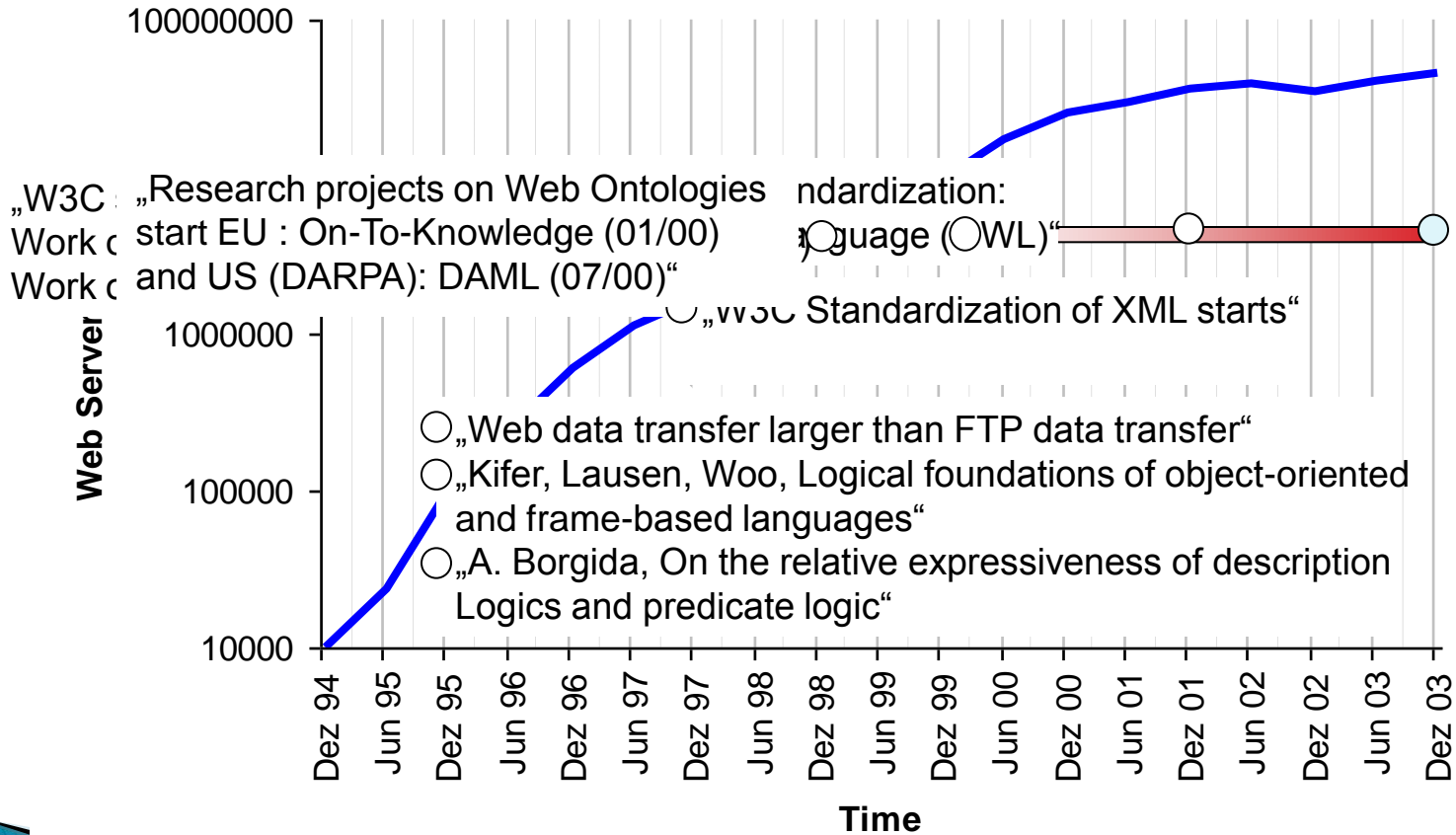
# THE SEMANTIC WEB

A new form of Web content  
that is meaningful to computers  
will unleash a revolution of new abilities

by  
TIM BERNERS-LEE,  
JAMES HENDLER and  
ORA LASSILA

# ... Semantic Web HISTORY

10.2.2004. Semantic Web = **Web** + **Database technology**  
**Resource Description Framework (RDF)**  
**Web Ontology Language (OWL)**  
 become **W3C recommendations**



[ Source: <http://www.zakon.org/robert/internet/timeline/> ]

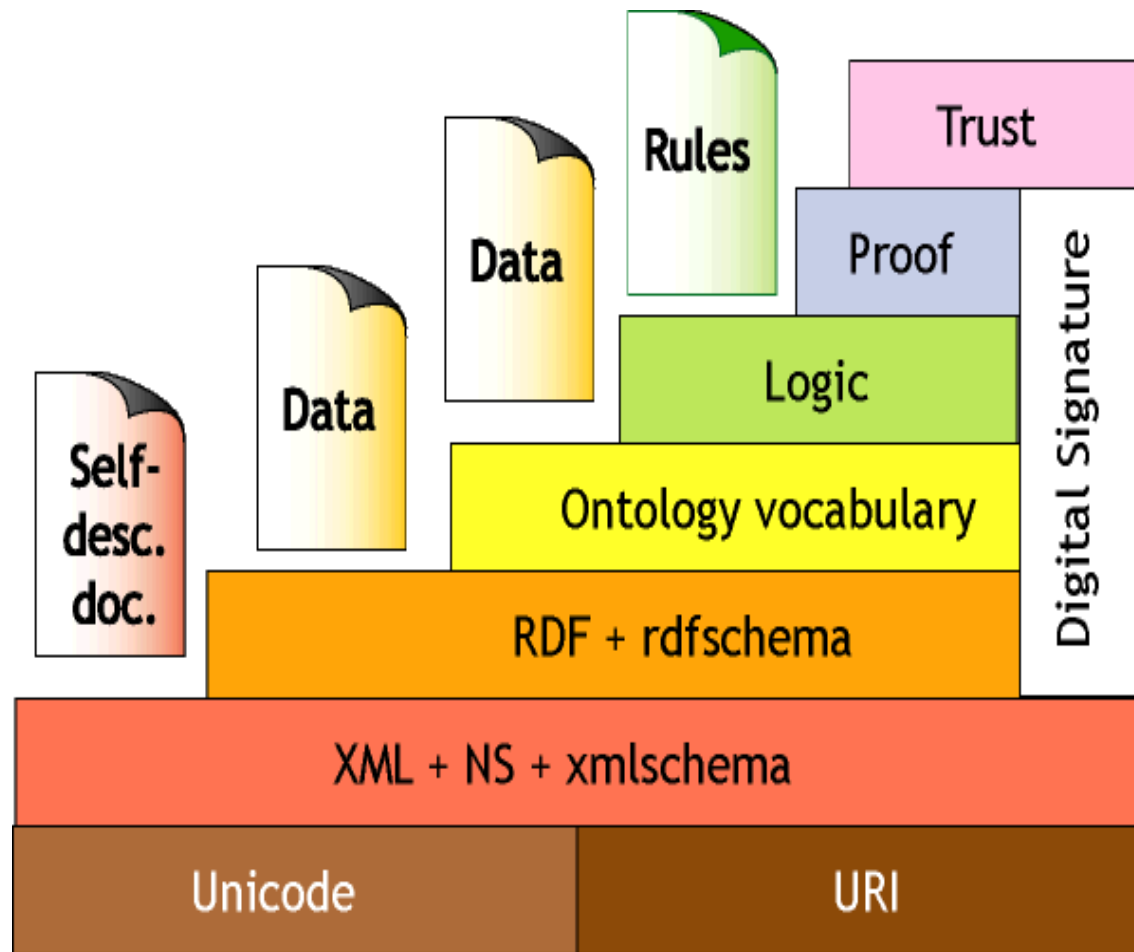
# Definizione di Semantic Web

„The **Semantic Web** is an extension of the current web in which information is given well-defined **meaning**, better enabling computers and people to **work in co-operation**.“

[Berners-Lee et al., 2001]

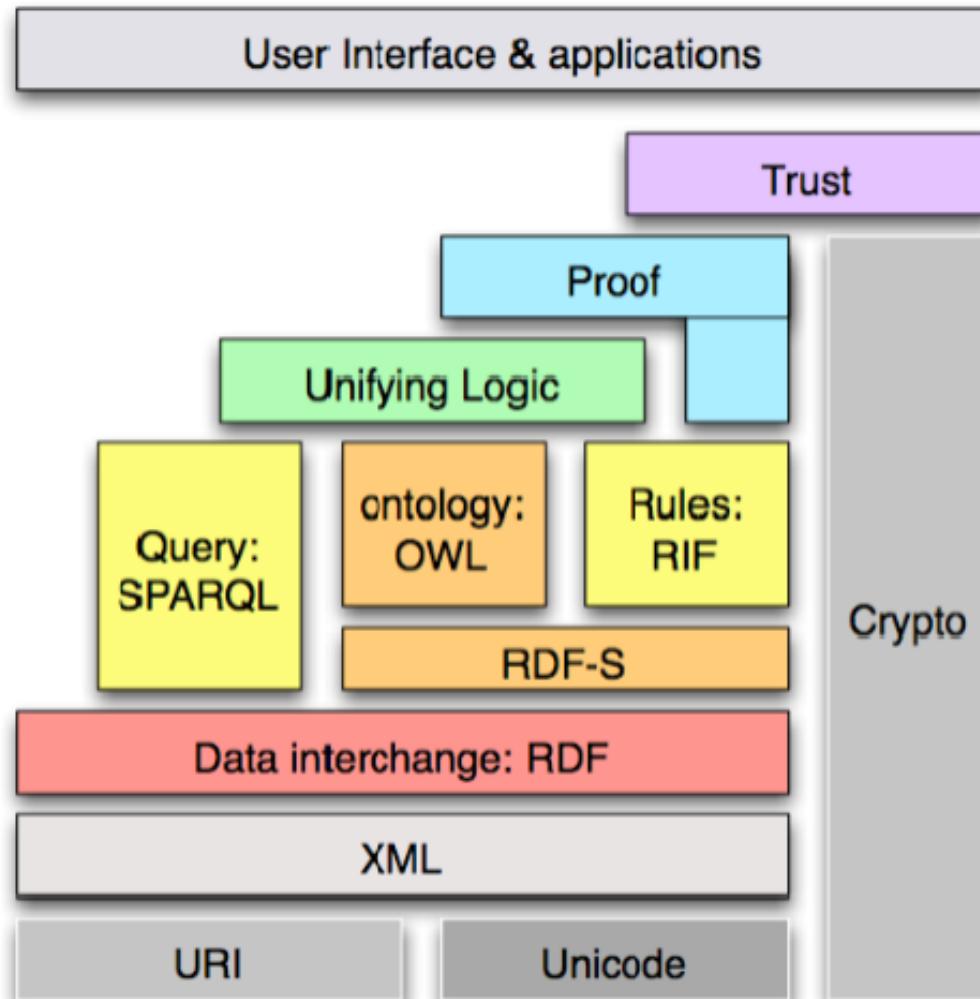


# La visione del Semantic Web ("torta a strati" versione 2000)





# La visione del Semantic Web ("torta a strati" versione 2006)





# Dove siamo oggi: il Web Sintattico



http://www2002.org

## WWW 2002

THE ELEVENTH INTERNATIONAL WORLD WIDE WEB CONFERENCE

Sheraton Waikiki Hotel  
Honolulu, Hawaii, USA  
7-11 May 2002

HAWAII

LOCATION. 5 DAYS. LEARN. INTERACT.

Registered participants coming from:  
Australia - Canada - Chile - Denmark - France - Germany - Ghana - Hong Kong - India - Italy - Ireland - Japan - Malta - New Zealand - The Netherlands - Norway - Singapore - Switzerland - The United States - Vietnam - Zambia

REGISTER NOW

On 7-11 May 2002, Honolulu, Hawaii will provide the backdrop for The Eleventh International World Wide Web Conference. This prestigious series of the International World Wide Web Conference Committee (IW3C2) attracts participants from around the world, and it provides a public forum for the World Wide Web Consortium (W3C) through the annual W3C track.

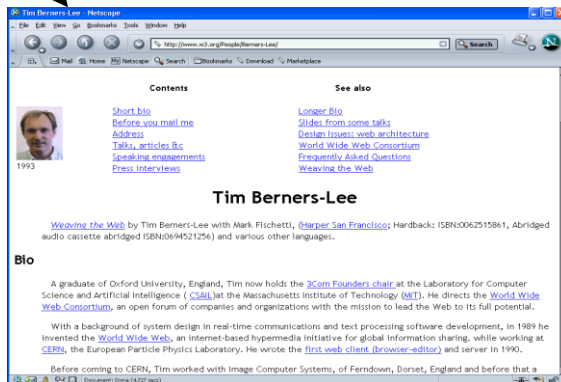
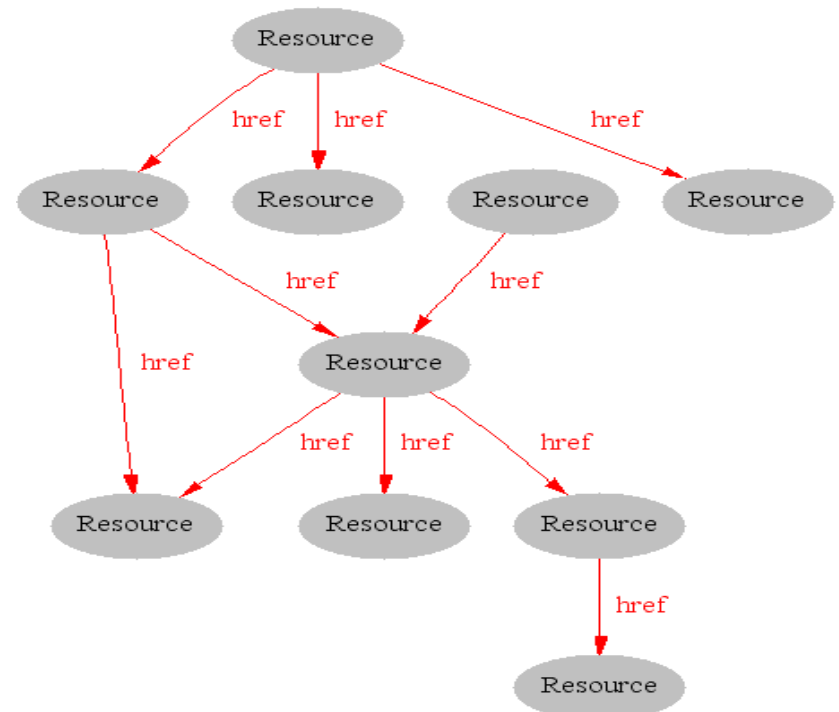
The conference is being organized by the International World Wide Web Conference Committee (IW3C2), the University of Hawaii and the Pacific Telecommunications Council (PTC).

FEATURED SPEAKERS (CONFIRMED)

Tim Berners-Lee, inventor of the World Wide Web and Director of the W3C who now holds the Scott Founders chair at the Laboratory for Computer Science (LCS) at the Massachusetts Institute of Technology (MIT).

Richard A. Dainoff, vice president and chief technology officer for Hewlett-Packard Company.

Ian Foster, guru of "Grid Computing", associate



Tim Berners-Lee - Message

Contents

- Short bio
- Before you mail me
- Address
- Tabla, articles &c
- Speaking engagements
- Press interviews

See also

- Longer Bio
- Slides from some talks
- Design Issues, web architecture
- World Wide Web Consortium
- Frequently Asked Questions
- Weaving the Web

### Tim Berners-Lee

[View the Web](#) by Tim Berners-Lee with Mark Fischetti, [Chaper San Francisco](#); Hardback: ISBN0062515861, Abridged audio cassette abridged ISBN0694521256 and various other languages.

**Bio**

A graduate of Oxford University, England, Tim now holds the [3Com Founders chair](#) at the Laboratory for Computer Science and Artificial Intelligence (LCSAIL) at the Massachusetts Institute of Technology (MIT). He directs the [World Wide Web Consortium](#), an open forum of companies and organizations with the mission to lead the Web to its full potential.

With a background of system design in real-time communications and text processing software development, in 1989 he invented the [World Wide Web](#), an internet-based hypermedia initiative for global information sharing, while working at CERN, the European Particle Physics Laboratory. He wrote the [first web client \(Browser-editor\)](#) and server in 1990.

Before coming to CERN, Tim worked with Image Computer Systems, of Ferndown, Dorset, England and before that a

# Il Web Sintattico è...

- ▶ **Una biblioteca digitale ipermediale**
  - Una biblioteca di documenti (pagine web) interconnessi da una rete di connessioni (links)
- ▶ **Un database, una piattaforma applicativa**
  - Un portale applicativo comune, accessibile tramite pagine web e che presenta i risultati come pagine web
- ▶ **Una piattaforma di distribuzione multimediale**
  - Canali radio locali in streaming in tutto il mondo! Trailers, videoclip, news immediatamente disponibili ovunque.
- ▶ **Uno schema di denominazione**
  - Definisce un'identità univoca per i documenti

# Ovvero il Web Sintattico è ...

- ▶ Un luogo in cui
  - I computer si occupano della presentazione delle informazioni (facile) e
  - Le persone si occupano di pensare e interpretare le informazioni (difficile).
- ▶ *Ma perché non far fare ai computer la maggior parte del lavoro difficile?*

# Lavoro difficile con il Web

## Sintattico...

- ▶ Interrogazioni complesse che coinvolgono **conoscenza di fondo**
  - Trova informazioni sugli “animali che usano il sonar ma non sono pipistrelli, delfini o balene”
- ▶ Localizzazione di informazioni in **data repositories**
  - Organizzazione di **es. Gufo di campagna**
  - Beni e servizi **rapporto qualità/prezzo**
  - Risultati di es **ul genoma umano**
- ▶ Delega di comp **si ad “agenti” web**
  - Prenotami un **er il prossimo weekend,**  
in un luogo c **oppo lontano, in cui si**  
parli italiano c



# Qual'è il Problema?

- Consideriamo una pagina web tipica:

The image shows a screenshot of the WWW 2002 website. At the top, there is a blue banner with the text 'WWW 2002' in large yellow letters. Below this, the text reads 'THE ELEVENTH INTERNATIONAL WORLD WIDE WEB CONFERENCE' and 'Sheraton Waikiki Hotel, Honolulu, Hawaii, USA, 7-11 May 2002'. A logo on the left shows a stylized figure holding a globe, with 'WWW 2002 HAWAII' below it. On the right, there is a logo for 'CONFERENCE ORGANIZERS' with the text 'International World Wide Web Conference Committee'. Below the banner, the text '1 LOCATION. 5 DAYS. LEARN. INTERACT.' is displayed. A sidebar on the left contains a list of links: 'Conference Proceedings', 'Call for Participation', 'Program', 'Registration Information', 'Hotel Accommodation', 'Conference Committee', 'Sponsorship/Exhibition Opportunities', 'Volunteer Information', 'Information about Hawaii', and 'Previous & Future WWW Conferences'. The main content area features a section titled 'Registered participants coming from:' followed by a list of countries: 'Australia · Canada · Chile · Denmark · France · Germany · Ghana · Hong Kong · India · Italy · Ireland · Japan · Malta · New Zealand · The Netherlands · Norway · Singapore · Switzerland · The United States · Vietnam · Zambia'. Below this list is a blue button labeled 'REGISTER NOW'. Further down, there is a section titled 'On 7-11 May 2002, Honolulu, Hawaii will provide the backdrop for The Eleventh International World Wide Web Conference. This prestigious series of the International World Wide Web Conference Committee (I<sup>3</sup>W<sup>2</sup>) attracts participants from around the world, and it provides a public forum for the World Wide Web Consortium (W3C) through the annual W3C track.' Below this, it states 'The conference is being organized by the International World Wide Web Conference Committee (I<sup>3</sup>W<sup>2</sup>), the University of Hawaii and the Pacific Telecommunications Council (PTC)'. At the bottom, there is a section titled 'FEATURED SPEAKERS (CONFIRMED)' with two columns of speaker information. The first column lists 'Tim Berners-Lee, inventor of the World Wide Web and Director of the W3C who now holds the 3Com Founders chair at the Laboratory for Computer Science (LCS) at the Massachusetts Institute of Technology (MIT)'. The second column lists 'Richard A. DeMillo, vice president and chief technology officer for Hewlett-Packard Company.' Below these, there are small portraits and names of other speakers, including 'Ian Foster, guru of "Grid Computing", associate' and 'McArthur Prize Winner'.

- Il Markup consiste di:
  - Visualizzazione dell'informazione (es. dimensione e colore del font)
  - Hyper-link a contenuti collegati
- Il contenuto semantico è accessibile agli esseri umani ma non (facilmente) ai computer...



# Che informazione possiamo vedere...

WWW2002

The eleventh international world wide web conference

Sheraton waikiki hotel, Honolulu, hawaii, USA

7-11 may 2002, 1 location 5 days learn interact

Registered participants coming from

australia, canada, chile denmark, france, germany, ghana, hong kong,, norway, singapore, switzerland, the united kingdom, the united states, vietnam, zaire

Register now

On the 7<sup>th</sup> May Honolulu will provide the backdrop of the eleventh international world wide web conference. This prestigious event..

Speakers confirmed

Tim berners-lee

Tim is the well known inventor of the Web, ...

Ian Foster

Ian is the pioneer of the Grid, the next generation internet ...





# XML

*Markup definito dall'utente e specifico per un dominio*

HTML:

```
<H1>Knowledge Management</H1>
  <UL>
    <LI>Manager: John Davies
    <LI>Project: SEKT  </UL>
```

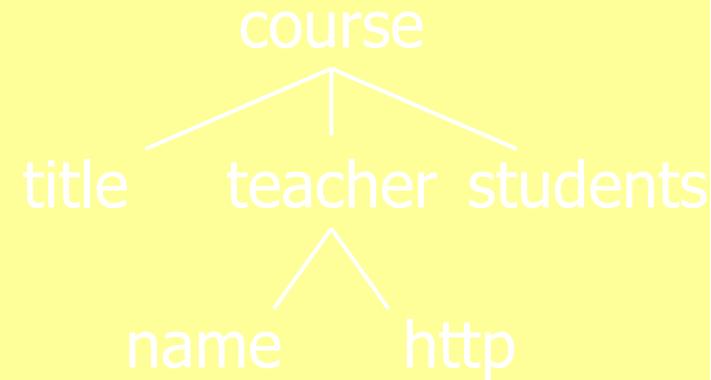
XML:

```
<research-topic>
  <title>Knowledge Management</title>
  <manager>John Davies</manager>
  <project>SEKT</project>
</research-topic>
```

# Documento XML = albero etichettato

- nodo = etichetta + contenuto

```
<course date="...">  
  <title>...</title>  
  <teacher>...</teacher>  
    <name>...</name>  
    <http>...</http>  
  <students>...</students>  
</course>
```



- DTD o XML-Schema: semplici grammatiche per definire la struttura degli alberi corretti

# Soluzione: XML con tag “significativi”?

<name>☯☯☯📁📁📁

✿⚡Ⓜ Ⓜ●Ⓜ ✨Ⓜ■◆⚡ ⋈■◆Ⓜ□■🔍◆⋈□■🔍● ◆□□●⚡ ◆⋈⚡Ⓜ  
◆Ⓜ👉Ⓜ□■</name>

<location>💧⚡Ⓜ□🔍◆□■ ◆🔍⋈&⋈&⋈ ⚡□◆Ⓜ●

👉□■□●◆●◆📁 ⚡🔍◆🔍⋈(⋈)📁 +💧👉</location>

<date>📁📁📁📁📁 ○🔍📁 📁📁📁📁</date>

<slogan>📁 ●□Ⓜ🔍◆⋈□■ 📁 ⚡🔍◆ ●Ⓜ🔍□■  
⋈■◆Ⓜ□🔍◆</slogan>

<participants>✿Ⓜ 📂⋈◆◆Ⓜ□Ⓜ⚡ □🔍□◆⋈Ⓜ⋈□🔍◆◆ Ⓜ□○⋈■📂  
⋈□□○

🔍◆◆□🔍●⋈🔍📁 Ⓜ🔍📁🔍🔍🔍🔍 Ⓜ⚡⋈●Ⓜ ⚡Ⓜ■○🔍□&📁  
⋈□🔍ⓂⓂ📁 📂Ⓜ□○🔍📁📁 📂⚡🔍🔍🔍🔍 ⚡□■📂 &□■📂  
⋈■⚡⋈🔍📁 ⋈□Ⓜ●🔍📁⚡ ⋈◆🔍●📁📁 ⋈🔍🔍🔍🔍 ○🔍●◆🔍📁  
■Ⓜ◆ ⋈Ⓜ🔍●🔍📁⚡ ◆⚡Ⓜ ■Ⓜ◆⚡Ⓜ□●🔍📁⚡📁 ■□□◆🔍📁  
◆⋈■📂🔍□□□Ⓜ📁 ◆◆⋈◆⋈Ⓜ□●🔍📁⚡ ◆⚡Ⓜ ◆■⋈◆Ⓜ⚡  
&⋈■📂⚡□○📁 ◆⚡Ⓜ ◆■⋈◆Ⓜ⚡ ◆◆🔍◆Ⓜ◆📁 ✨⋈Ⓜ◆■🔍○📁  
⋈🔍⋈□Ⓜ</participants>



# Limitazioni nell'uso di XML per il markup semantico

XML di per sé non ha nessun supporto per:

- ▶ **Vocabolario** ontologico specifico per un dominio
    - Quali parole dovremmo usare per descrivere un insieme di concetti?
  - ▶ **Primitive di modellazione** ontologiche
    - Come possiamo combinare tali concetti, es. “automobile è a-kind-of (subclass-of) veicolo”
- ⇒ richiede un accordo preventivo su vocabolario e primitive

Uso fattibile solo in caso di collaborazione stretta

- Agenti in una comunità piccola e stabile
- Pagine in una intranet piccola e stabile
- .. **Non adatta a risorse Web da condividere**



# Limitazioni del Web odierno



L'orientamento è machine-to-human,  
e non machine-to-machine

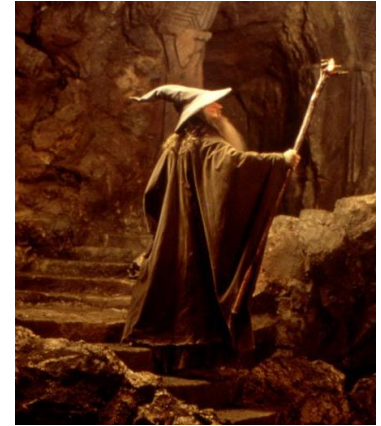
# XML è un primo passo

- ▶ Markup Semantico
  - HTML ⇒ presentazione
  - XML ⇒ significato
- ▶ Metadati
  - All'interno dei documenti, non fra documenti
  - *prescrittivi*, non *descrittivi*
  - Nessuna predisposizione o supporto per vocabolario e primitive di modellazione
- ▶ RDF è il passo successivo



# Resource Description Framework (RDF)

- ▶ Uno standard del W3C
- ▶ Descrive relazioni tra risorse
- ▶ Consiste di triple o “affermazioni”:
  - <oggetto, proprietà, oggetto>
  - <Tolkien, wrote, The Lord of the Rings>
  
- ▶ RDFS estende RDF con un “vocabolario ontologico” standard:
  - Class, Property
  - Type, subClassOf
  - domain, range



# Cos'è RDF ?

## ■ RDF è un modello dei dati

- Il modello è neutro rispetto al dominio e all'applicazione e pronto per l'internazionalizzazione
- il modello può essere visto come un grafo direzionato etichettato o come un modello dei dati object-oriented (oggetto/attributo/valore)

## ■ Il modello dei dati RDF è uno strato concettuale astratto indipendente da XML

- Di conseguenza, XML è una sintassi di trasferimento, non una componente di RDF
- I dati RDF possono anche non comparire mai in formato XML

# Il modello RDF

- Una "affermazione" RDF consiste di  
risorse (= nodi, identificate da URI)  
che hanno proprietà  
che assumono valori (= nodi, stringhe)

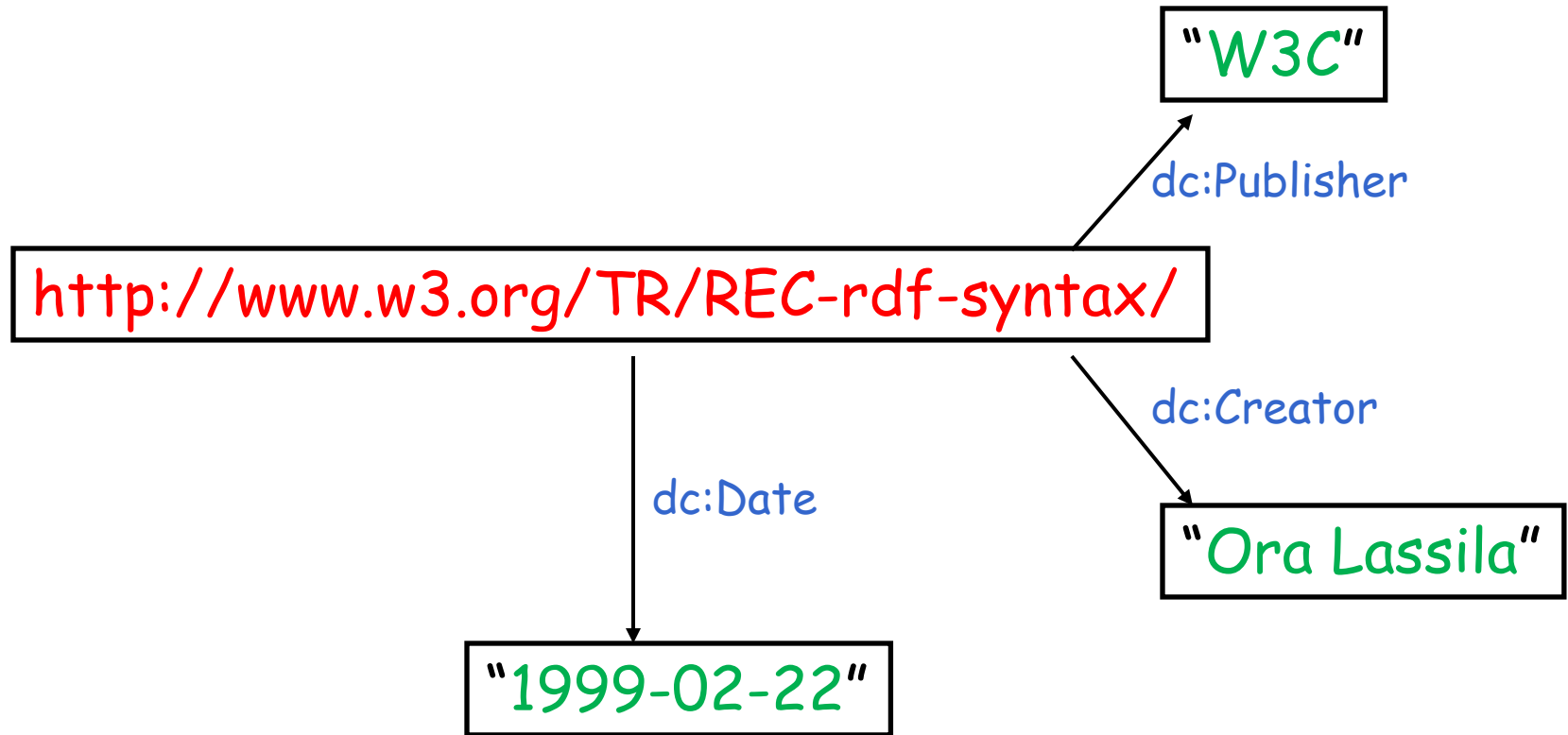
= soggetto  
= predicato  
= oggetto



"<http://www.w3.org/TR/REC-rdf-syntax/> ha come autore Ora Lassila"



# Il Modello RDF – Esempio

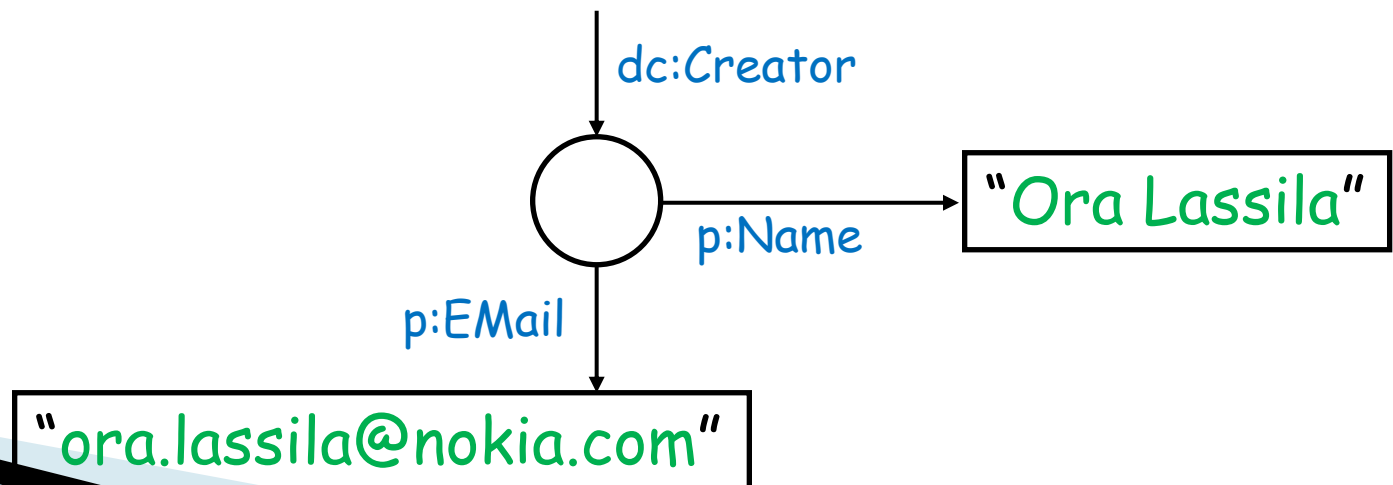


# Valori Complessi

- Finora abbiamo considerato proprietà aventi come valori stringhe
- Anche un nodo del grafo (corrispondente a una risorsa) può essere il valore di una proprietà
  - Sono possibili strutture ad albero o grafo arbitrariamente complesse
  - sintatticamente, i valori possono essere inglobati (lessicalmente in-line) o referenziati (linked)

<http://www.w3.org/TR/REC-rdf-syntax/>

- Esempio:



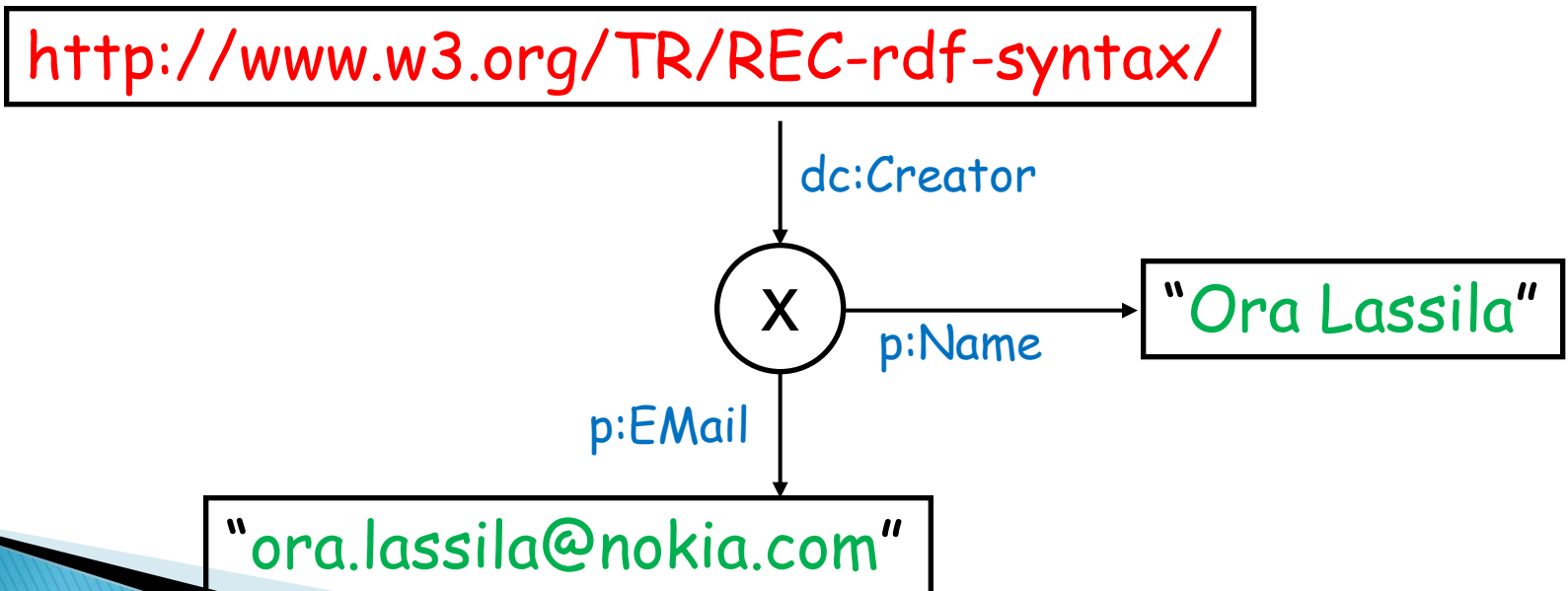
# Valori Complessi (continua)

- Triple corrispondenti

{ `http://www.w3.org/TR/PR-rdf-syntax/`, `dc:Creator`, `x` }

{ `x`, `p:Name`, `"Ora Lassila"` }

{ `x`, `p:EMail`, `"ora.lassila@nokia.com"` }



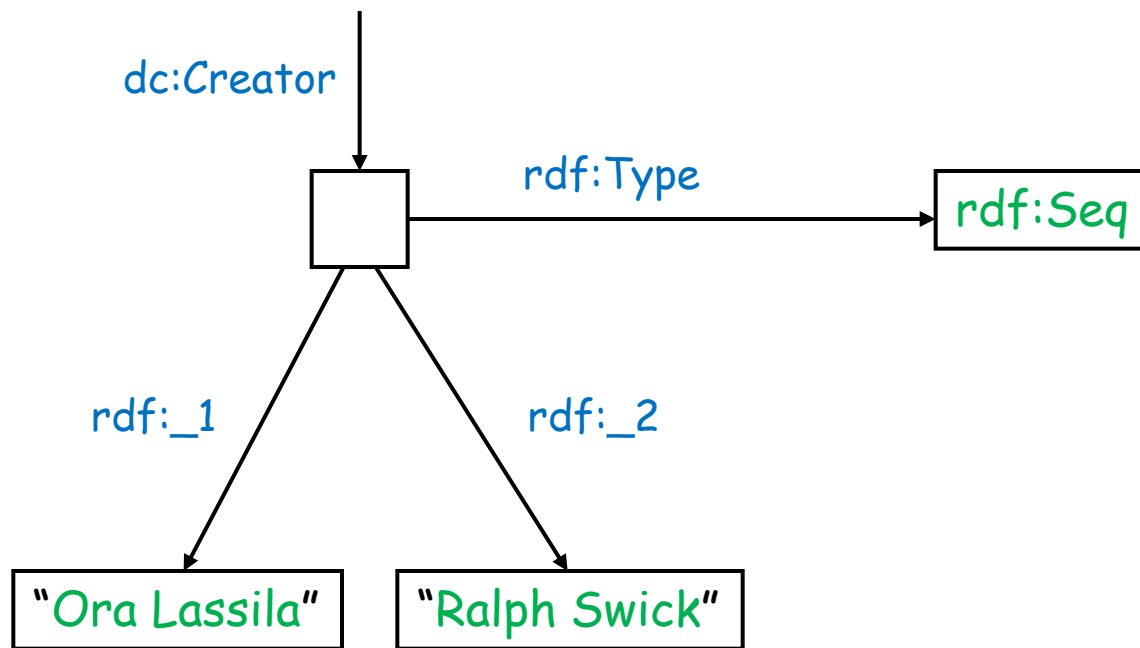
# Contenitori (Containers)

- I Containers sono collezioni
  - Consentono di raggruppare risorse (o valori letterali)
- E' possibile fare affermazioni sul container (come un tutto) o individualmente sui suoi membri
- Esistono diversi tipi di container
  - bag – collezione non ordinata
  - seq – collezione ordinata (= "sequenza")
  - alt – rappresenta alternative
- E' anche possibile creare collezioni basate su pattern di URI
  - Per esempio, tutti i file in un particolare sito web
- Sono consentiti valori duplicati
  - Non c'è nessun meccanismo per imporre vincoli di unicità



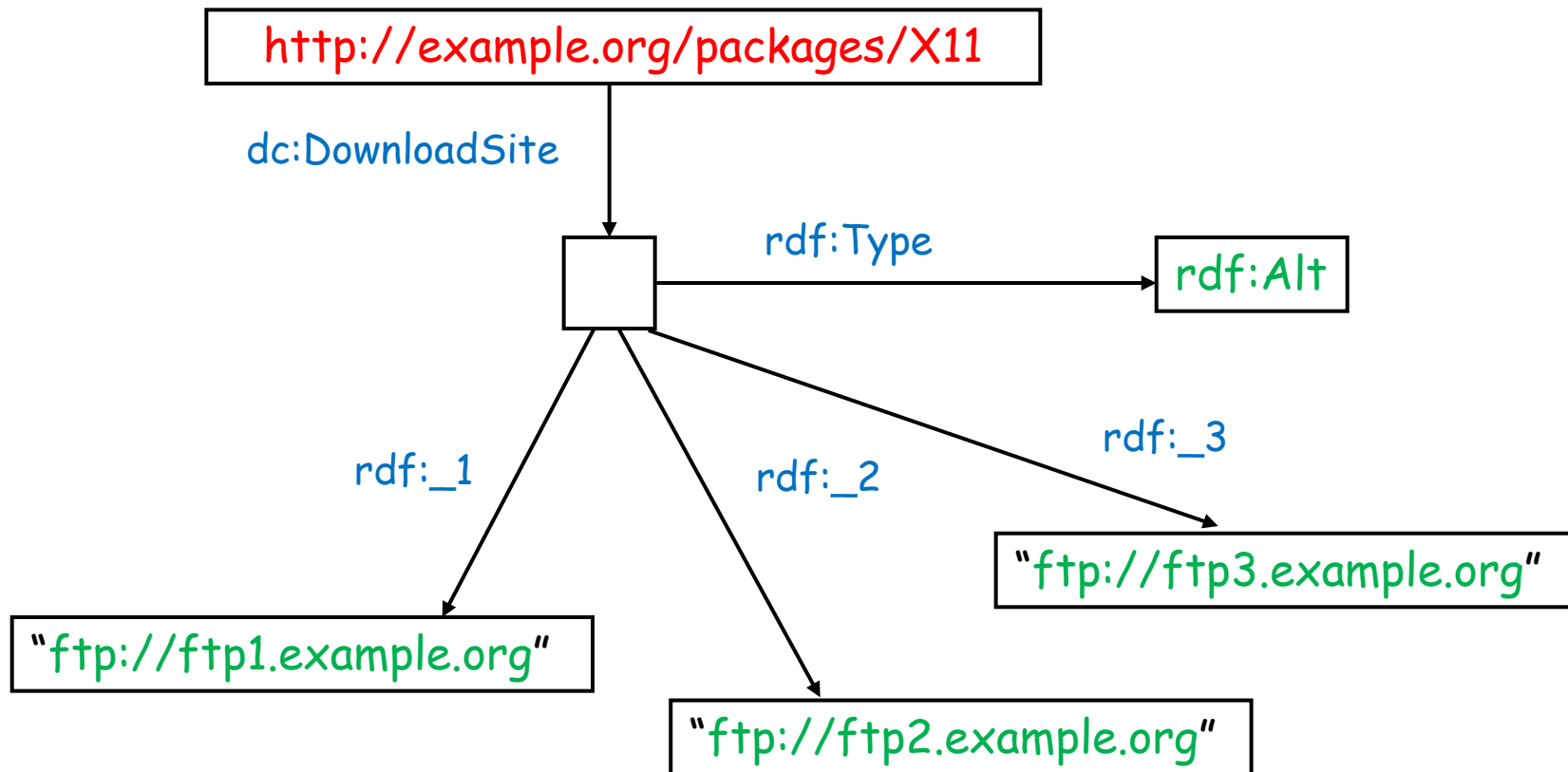
# Containers – Esempio di Seq

<http://www.w3.org/TR/REC-rdf-syntax>



Le proprietà `rdf:Type`  
e, per i tipi container, le proprietà `rdf:_1`, `rdf:_2`, ...  
sono predefinite (si può usare anche `rdf:li`)

# Containers – Esempio di ALt

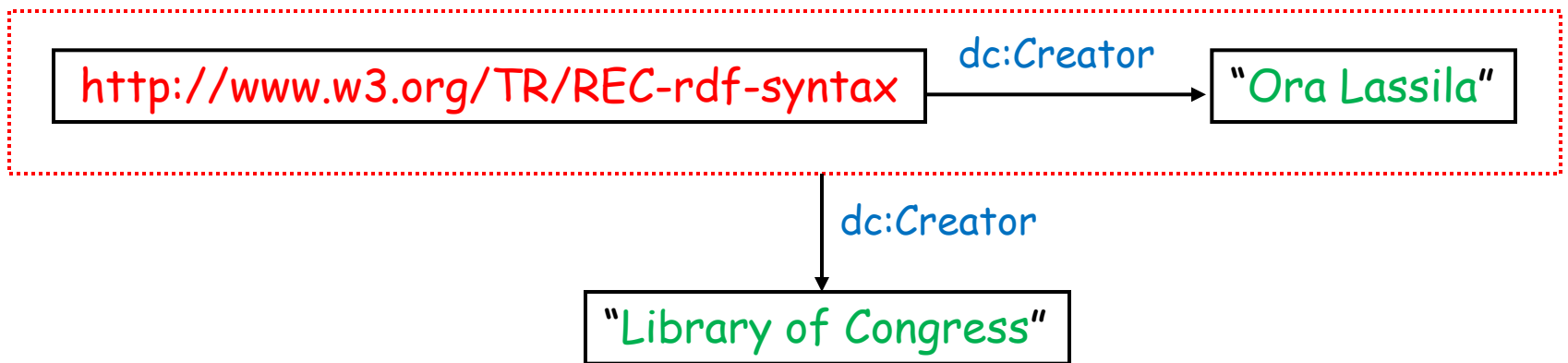


# Affermazioni di ordine superiore

- Si possono fare affermazioni RDF su di altre affermazioni RDF
  - esempio: “Ralph crede che il web contenga un miliardo di documenti”
- Affermazioni di ordine superiore
  - Ci consentono di esprimere credenze (e altre modalità: possibilità, desiderabilità...)
  - Sono importanti per i modelli di trust, firma digitale, etc.
  - e per esprimere metadati sui metadati
  - Si rappresentano modellando RDF usando RDF

# Reificazione

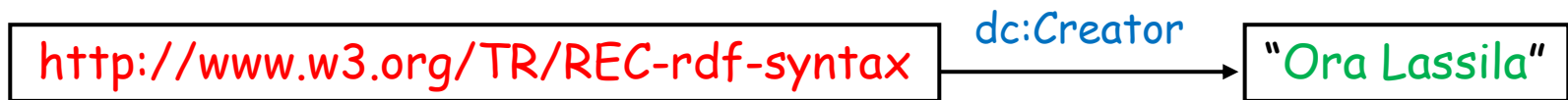
- RDF non è realmente un linguaggio del secondo ordine
- Ma fornisce un vocabolario di predicati predefiniti per la reificazione



- Il rettangolo rosso tratteggiato corrisponde alle seguenti affermazioni
  - { x, **rdf:predicate**, "dc:creator" }
  - { x, **rdf:subject**, "http://www.w3.org/TR/REC-rdf-syntax" }
  - { x, **rdf:object**, "Ora Lassila" }
  - { x, **rdf:type**, "**rdf:statement**" }

# RDF/XML

- XML, al pari di altri formalismi sintattici (N3, Turtle, N-Triples, RDFa, RXX...), può essere usato per codificare un grafo RDF in un formato leggibile dalle macchine
- Ad esempio la tripla RDF:



- Si codifica in XML come:

```
> <rdf:Description
>     rdf:about="http://www.w3.org/TR/REC-rdf-syntax">
>     <dc:Creator>Ora Lassila</dc:Creator>
> </rdf:Description>
```

# RDF/XML (2)

- La Serializzazione in formato XML prevede anche la dichiarazione in modo standard dei namespace per il linguaggio RDF e per i vocabolari usati dall'utente
- Completando l'esempio di prima, il file XML diventa:

```
▶ <?xml version="1.0"?>
▶ <rdf:RDF
▶   xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
▶   xmlns:dc="http://purl.org/dc/elements/1.1/">
▶ <rdf:Description
▶   rdf:about="http://www.w3.org/TR/REC-rdf-syntax">
▶   <dc:Creator>Ora Lassila</dc:Creator>
▶ </rdf:Description>
▶ </rdf:RDF>
```

# RDF/XML - Semplificazioni

- Più proprietà possono essere codificate per la stessa risorsa
- Riferimenti URI sono codificati come attributi
- Ai valori letterali può essere attribuito un tipo (gli stessi definiti in XML-Schema)
- ...
- Esempio:


```
▶ <rdf:Description
    rdf:about="http://www.w3.org/TR/REC-rdf-syntax">
▶   <dc:Creator>Ora Lassila</dc:Creator>
▶   <dc:Publisher>W3C</dc:Publisher>
▶   <p:EMail
    rdf:resource="mailto:ora.lassila@nokia.com"/>
▶   <dc>Date
    rdf:datatype="http://...#date">1999-02-22</dc>Date>
▶ </rdf:Description>
```



# RDFS – RDF Schema

- ▶ Analogo al type system di un linguaggio object-oriented, ma costruito sopra RDF
- ▶ Definisce un piccolo vocabolario per RDF:
  - rdfs:Class, rdfs:subClassOf
  - rdfs:Resource, rdfs:Literal
  - rdfs:Property, rdfs:subPropertyOf
  - rdfs:range, rdfs:domain
  - rdfs:label, rdfs:comment, rdfs:seeAlso
- ▶ Le applicazioni di RDF Schema sono per lo più rimandate all'uso in OWL, che ne arricchisce le funzionalità

# RDF Schema *vs* XML Schema

- ▶ Entrambi servono a definire vocabolari
  - ▶ Un documento XML Schema ci fornisce i dettagli sintattici
  - ▶ Un documento RDF Schema ci fornisce il significato
  - ▶ Un documento OWL (più avanti) ci consente di catturare un significato più ricco
- 

# Sintassi di RDF Schema in XML

```
<rdf:Description ID="MotorVehicle">  
  <rdf:type resource="http://www.w3.org/...#Class"/>  
  <rdfs:subClassOf rdf:resource="http://www.w3.org/...#Resource"/>  
</rdf:Description>
```

```
<rdf:Description ID="Truck">  
  <rdf:type resource="http://www.w3.org/...#Class"/>  
  <rdfs:subClassOf rdf:resource="#MotorVehicle"/>  
</rdf:Description>
```

```
<rdf:Description ID="registeredTo">  
  <rdf:type resource="http://www.w3.org/...#Property"/>  
  <rdfs:domain rdf:resource="#MotorVehicle"/>  
  <rdfs:range rdf:resource="#Person"/>  
</rdf:Description>
```

```
<rdf:Description ID="ownedBy">  
  <rdf:type resource="http://www.w3.org/...#Property"/>  
  <rdfs:subPropertyOf rdf:resource="#registeredTo"/>  
</rdf:Description>
```

# RDF e RDFS

- ▶ RDFS definisce una semplice ontologia:
  - Le classi di interesse e le loro proprietà e relazioni
  - Su quali concetti vogliamo ragionare e come sono tra di loro collegati
  - *Es: Ci sono autori, e gli autori scrivono libri*
- ▶ RDF definisce le istanze di queste classi e delle loro proprietà:
  - Mark Twain è un autore
  - Mark Twain ha scritto “Le Avventure di Tom Sawyer”
  - “Le Avventure di Tom Sawyer” è un libro
- ▶ Notazione:  $\text{RDF(S)} = \text{RDF} + \text{RDFS}$

# RDF

“<http://www.famouswriters.org/twain/mark>” **hasName** “Mark Twain”

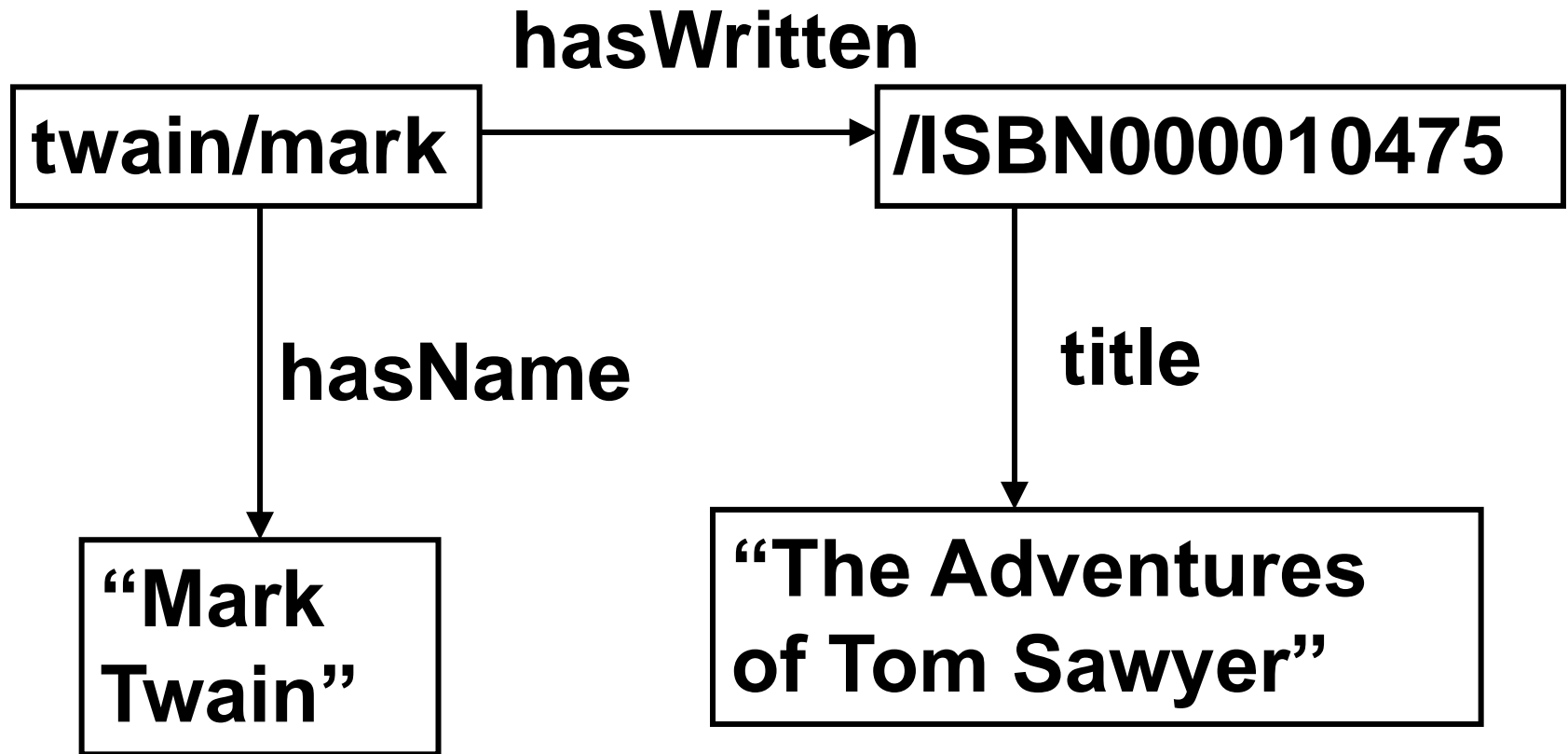
“<http://www.famouswriters.org/twain/mark>”  
**hasWritten** “<http://www.books.org/ISBN00001047582>”

“<http://www.books.org/ISBN00001047582>”  
**title** “The Adventures of Tom Sawyer”

## Versione XML:

```
<rdf:Description rdf:about="http://www.famouswriters.org/twain/mark">  
  <ex:hasName>Mark Twain</ex:hasName>  
  <ex:hasWritten rdf:resource="http://www.books.org/ISBN00001047"/>  
</rdf:Description>  
<rdf:Description rdf:about="http://www.books.org/ISBN00001047582">  
  <ex:title>The Adventures of Tom Sawyer</ex:title>  
</rdf:Description>
```

# Il grafo RDF di esempio



# RDF(S) definizioni

FamousWriter subclassof Writer

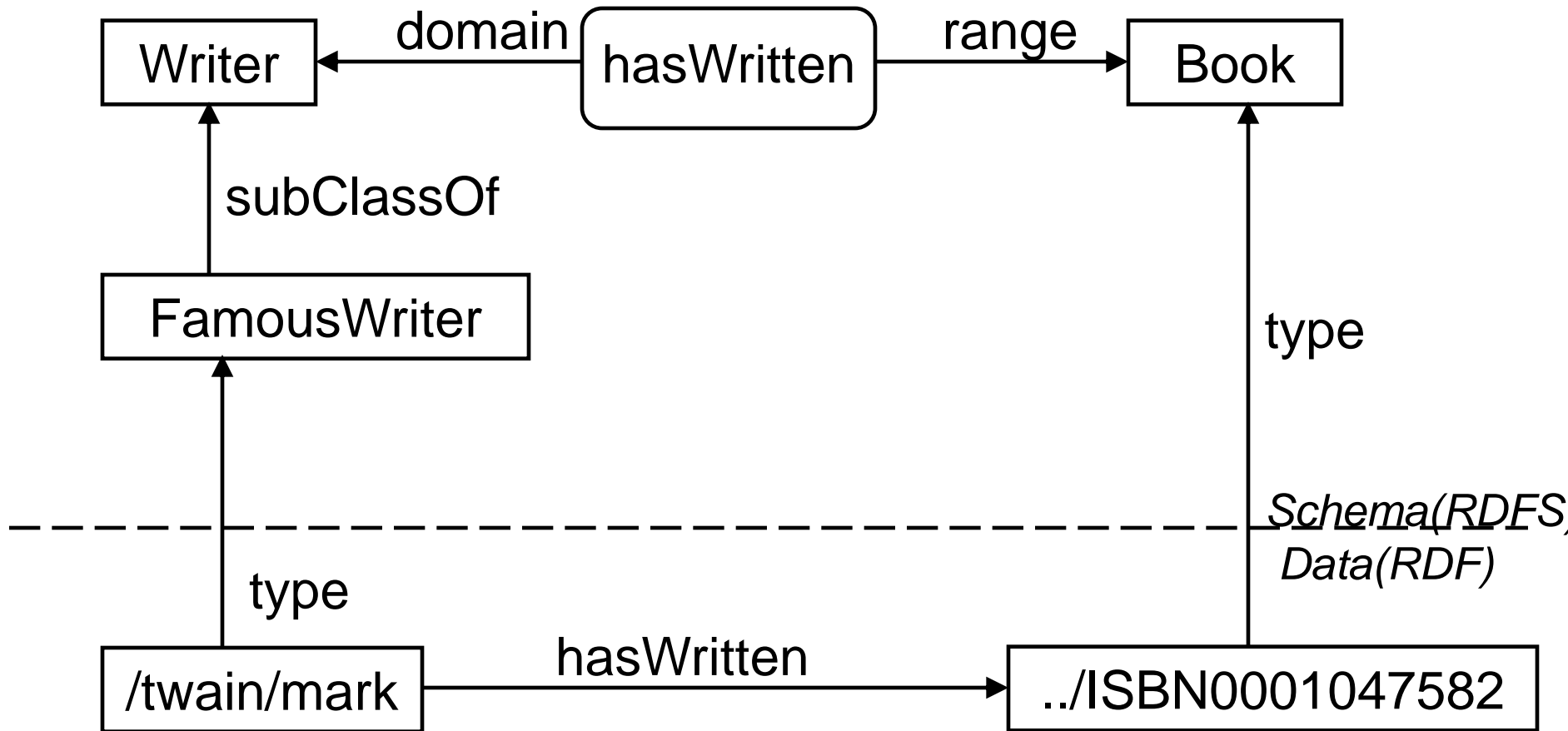
“<http://www.famouswriters.org/twain/mark>”  
type “<http://example.org/schema#FamousWriter>”

## Versione XML:

```
<rdf:Description ID="FamousWriter"  
  <rdf:type resource="http://www.w3.org/...#Class"/>  
  <rdf:subClassof rdf:resource="Writer"/>  
</rdf:Description>  
<rdf:Description rdf:about="http://www.famouswriters.org/twain/mark">  
  <rdf:type resource="http://example.org/schema#FamousWriter"/>  
</rdf:Description>
```

# Lo Schema RDF di esempio

Annotazione di risorse WWW e link *semantici*





# Conclusioni su RDF(S)

- ▶ Il primo passo avanti dopo XML:
  - (limitato) **supporto ontologico** di primitive di modellazione
  - Possibilità di definire **vocabolari**
- ▶ Ma non è ancora sufficiente:
  - **nessun** significato descritto precisamente
  - **nessun** modello di inferenza