

## Ngôn ngữ miêu tả RDF

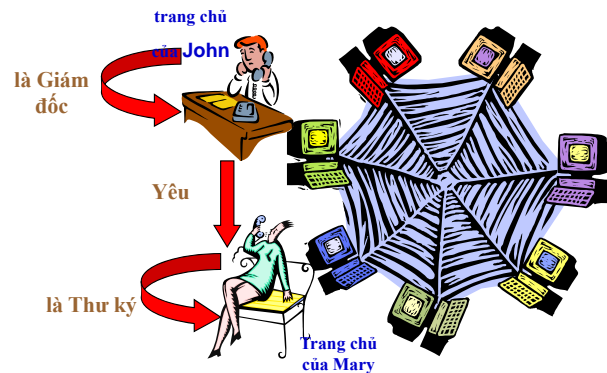
### Chủ đề chính

- RDF
- RDFS

2

### Ứng dụng Web thông minh

- Mô tả
  - thuộc tính ngữ nghĩa của một tài nguyên Web
  - quan hệ ngữ nghĩa giữa chúng



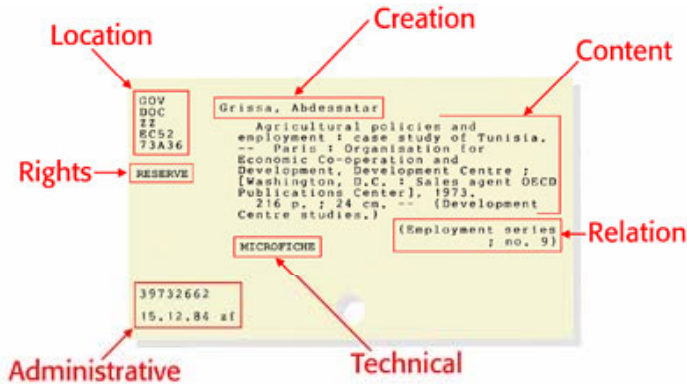
3

### Siêu dữ liệu

- Ví dụ 1
  - Cổ vật – Địa điểm khai quật
  - Người – Ngày sinh
- Ví dụ 2
  - Bài báo khoa học – ngày xuất bản
  - Bản đồ - Tỷ lệ
  - File audio - tần số lấy mẫu
  - Ảnh số - máy ảnh
  - Mục dữ liệu trong CSDL - Người nhập liệu
  - Trang Web - chủ đề

4

## Siêu dữ liệu



5

## RDF - Resource Description Framework

- Công nghệ nền móng xử lý siêu dữ liệu Web;
- Giúp tạo ra sự liên tác giữa các ứng dụng trên Web trao đổi thông tin diễn dịch được bởi máy tính
- Ngôn ngữ biểu diễn thông tin về một
  - TÀI NGUYÊN (resource) Web
  - thông tin về một thứ gì đó có thể được định danh trên Web, ngay cả khi nó không thể được lấy về trực tiếp từ Web

6

## Mục đích

- Tạo ra một mô hình dữ liệu đơn giản
- biểu diễn ngữ nghĩa hình thức – có thể suy diễn
- Sử dụng các từ vựng dựa URI – cú pháp XML
- Cho phép sử dụng các kiểu dữ liệu trong lược đồ XML
- Cho phép tạo ra thông tin về bất cứ tài nguyên nào

7

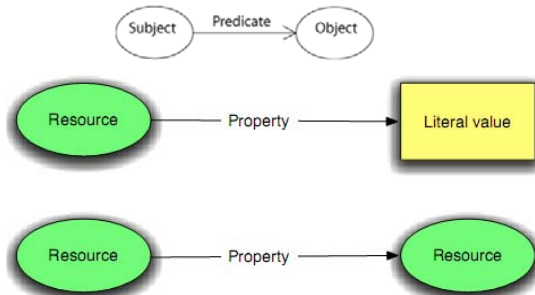
## Nguyên lý cơ bản

- Phân tách rạch ròi
  - Cấu trúc của Mô hình (đồ thị RDF)
  - Ngữ nghĩa diễn dịch
  - Cú pháp cụ thể (XML, TN, N3, ...)
- Hai kiểu dữ liệu duy nhất
  - URI/URIref:
  - Literal
    - String hoặc kiểu dữ liệu XSD

8

## Mô hình dữ liệu RDF

- Bộ ba: subject, predicate, object
- Một diễn tả: Tập các bộ ba



9

## Các đối tượng chính

- Tài nguyên
  - Subject của một tuyên bố RDF là một tài nguyên
- Thuộc tính
  - Predicate của một tuyên bố RDF là thuộc tính của một tài nguyên
- Phát biểu
  - Object : giá trị của thuộc tính đó

10

## Ràng buộc

- Subject chỉ có thể là
  - tham chiếu URI – node trắng
- Predicate
  - tham chiếu URI
- Object
  - tham chiếu URI, hằng giá trị, node trắng

11

## Tài nguyên và URI

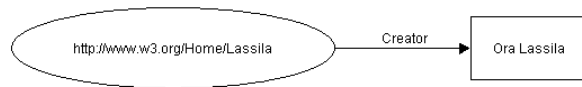
- Tài nguyên là bất cứ thứ gì có định danh ID
- *Uniform Resource Identifiers (URI)* – công cụ đơn giản cho phép xác định 1 tài nguyên

12

## Ví dụ về tuyên bố RDF

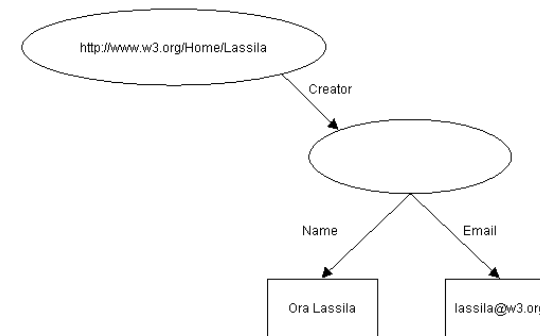
*Ora Lassila is the creator of the resource  
http://www.w3.org/Home/Lassila.*

Subject (resource)	http://www.w3.org/Home/Lassila
Predicate (property)	Creator
Object (literal)	"Ora Lassila"



## Thuộc tính có giá trị cấu trúc

*The individual whose name is Ora Lassila, email <lassila@w3.org>, is the creator of  
http://www.w3.org/Home/Lassila.*



## Thuộc tính có giá trị cấu trúc

Subject (resource)	http://www.w3.org/Home/Lassila
Predicate (property)	Creator
Object (literal)	SOMETHING

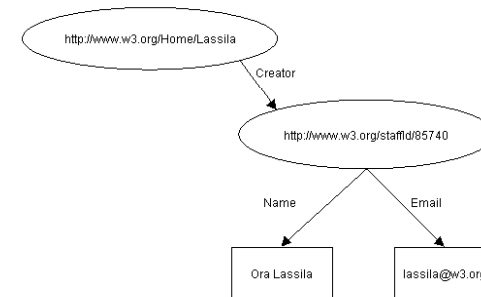
  

Subject (resource)	SOMETHING
Predicate (property)	Name
Object (literal)	"Ora Lassila"

Subject (resource)	SOMETHING
Predicate (property)	Email
Object (literal)	lassila@w3.org

## Thuộc tính có giá trị cấu trúc

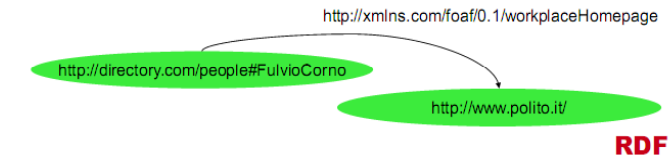


The individual referred to by employee id 85740 is named Ora Lassila and has the email address lassila@w3.org. The resource <http://www.w3.org/Home/Lassila> was created by this individual.

## Thuộc tính có giá trị cấu trúc

Subject (resource)	http://www.w3.org/Home/Lassila
Predicate (property)	Creator
Object (resource)	http://www.w3.org/staffid/85740
Subject (resource)	http://www.w3.org/staffid/85740
Predicate (property)	Name
Object (literal)	"Ora Lassila"
Subject (resource)	http://www.w3.org/staffid/85740
Predicate (property)	Email
Object (literal)	lassila@w3.org

## Thông tin trong bộ ba



CompanyHomePage

PersonID	Homepage
FulvioCorno	http://www.polito.it/

**Relational database**

**First order  
logic predicate**

```
HasCompanyHomePage(
  'FulvioCorno',
  'http://www.polito.it/' );
```

18

## Sự khác nhau

- Các bảng trong CSDL Quan hệ:
  - Số cột bất kỳ
- Vị từ trong logic bậc một:
  - Nhiều đối số
- Bộ ba RDF:
  - Một Subject, 1 Object

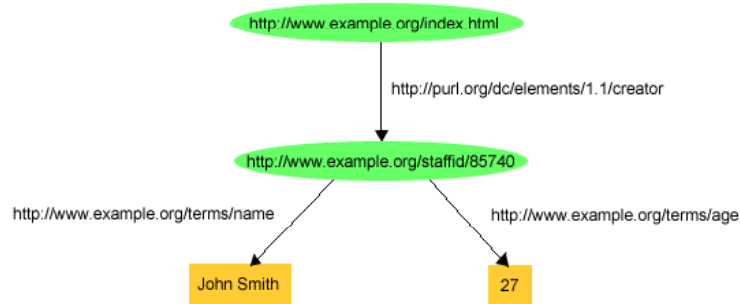
19

## Ví dụ



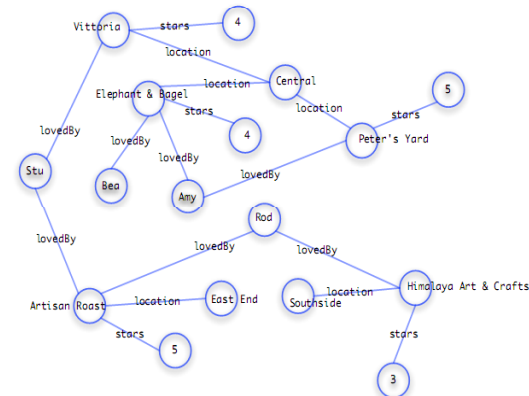
20

## Ví dụ



## RDF Graph

- 1 Triple – bộ ba
- Tuyên bố RDF – nhiều bộ ba: Một đồ thị



22

## URI

- Biểu diễn 1 nút (subject hoặc object)
  - Các thẻ hiện: Eric Miller, xác định bởi <http://www.w3.org/People/EM/contact#me>
  - Loại thực thể (lớp): Person, xác định bởi <http://www.w3.org/2000/10/swap/pim/contact#Person>
  - giá trị thuộc tính: <mailto:em@w3.org>
- Biểu diễn vị từ Predicates
  - thuộc tính: mailbox
    - <http://www.w3.org/2000/10/swap/pim/contact#mailbox>

23

## URI > URL

- URL = uniform resource locator
  - xác định – lấy về tài nguyên web
- URI = uniform resource identifier
  - Tổng quát hơn
  - Xác định các tài nguyên không có địa chỉ mạng
  - Các tổ chức – cá nhân có thể tự tạo ra các URI và dùng chúng để xác định các sự vật (cụ thể hoặc trừu tượng)

24

## URI Ref , URI Fragment

- Một tham chiếu URI là một URI có thể có định danh bộ phận ở cuối:

scheme authority path fragment  
 http ://www.inf.ed.ac.uk /ontology#annotation

http://www.example.com/index.html  
 http://www.example.com/index.html#Section2

25

## Tiếp

- URIref = URI reference
- → nhiều tài nguyên khác nhau
  - Tham chiếu URI của một file RDF với nhiều định nghĩa
  - Để xác định một "bộ phận" của URI, sử dụng #
    - Ví dụ: http://example.org/index#person

26

## Cú pháp RDF/XML

```
<?xml version="1.0"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:contact="http://www.w3.org/2000/10/swap/pim/contact#">
  <contact:Person rdf:about="http://www.w3.org/People/EM/contact#me">
    <contact:fullName>Eric Miller</contact:fullName>
    <contact:mailbox rdf:resource="mailto:em@w3.org"/>
    <contact:personalTitle>Dr.</contact:personalTitle>
  </contact:Person>
</rdf:RDF>
```

27

## Cú pháp RDF/XML

```
<?xml version="1.0"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:contact="http://www.w3.org/2000/10/swap/pim/contact#">
  <contact:Person rdf:about="http://www.w3.org/People/EM/contact#me">
    <contact:fullName>Eric Miller</contact:fullName>
    <contact:mailbox rdf:resource="mailto:em@w3.org"/>
    <contact:personalTitle>Dr.</contact:personalTitle>
  </contact:Person>
</rdf:RDF>
```

Name space shortcut.  
 Equivalent to  
 http://www.w3.org/2000/10/swap/pim/contact#fullName

28

## Namespace

- Tránh phải sử dụng URI đầy đủ
- XML Qualified Names (Qnames- tiếp đầu ngữ)
  - Gắn một tiền tố (prefix) với một URI;

Prefix	Namespace URI
edstaff	http://www.ed.ac.uk/staffid#
infcourses	http://www.inf.ed.ac.uk/teaching/courses/
dc	http://purl.org/dc/elements/1.1/
dbpedia	http://dbpedia.org/resource/

```
edstaff:9888
infcourses:masws
dc:creator
dbpedia:In_the_Heart_of_the_Moon
```

29

## Một số QName thông dụng

```
@prefix rdf: http://www.w3.org/1999/02/22-rdf-syntaxns#
@prefix rdfs: http://www.w3.org/2000/01/rdf-schema#
@prefix dc: <http://purl.org/dc/elements/1.1/>
@prefix ex: <http://www.example.org/>
@prefix exterms: <http://www.example.org/terms/ >
@prefix exstaff: <http://www.example.org/staffid/ >
@prefix ex2: <http://www.domain2.example.org/>
```

30

## Ví dụ

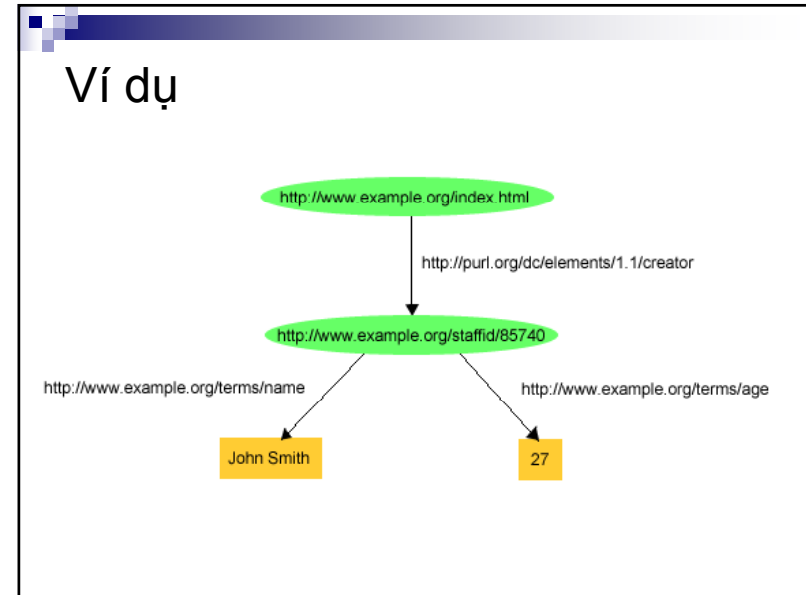
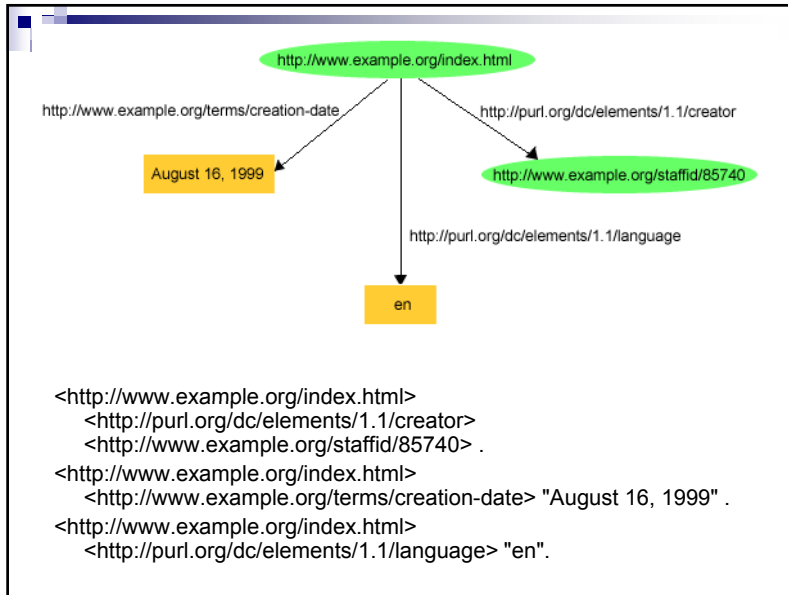
*Ora Lassila is the creator of the resource*  
*http://www.w3.org/Home/Lassila.*

```
<rdf:RDF>
  <rdf:Description about=
    "http://www.w3.org/Home/Lassila">
    <s:Creator>Ora Lassila</s:Creator>
  </rdf:Description>
</rdf:RDF>
```

```
xmlns:s="http://description.org/schema/"
```

- http://www.example.org/index.html has a creation-date whose value is August 16, 1999.
- http://www.example.org/index.html has a language whose value is English.





## RDF containers

- Sử dụng khi giá trị của một thuộc tính là tập hợp
  - ☐ bag
  - ☐ sequence
  - ☐ alternative

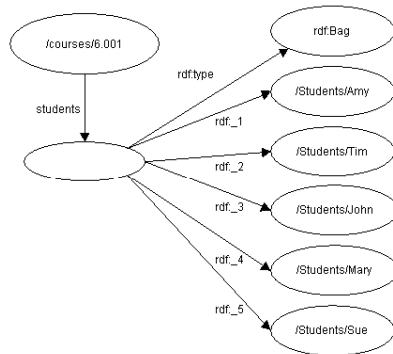
35

## Bag.

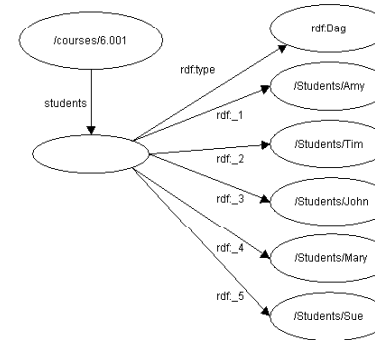
- Danh sách không sắp xếp các tài nguyên hoặc hằng giá trị.
- Dùng để khai báo một thuộc tính đa giá trị và thứ tự các giá trị không mang ý nghĩa.
- Cho phép các giá trị trùng nhau.

## Ví dụ

Các sinh viên tham gia khóa học số 6.001 là Amy, Tim, John, Mary, và Sue



## Ví dụ



## Ví dụ

Các sinh viên tham gia khóa học số 6.001 là Amy, Tim, John, Mary, và Sue

```
<rdf:RDF>
  <rdf:Description about="http://mycollege.edu/courses/6.001">
    <s:students>
      <rdf:Bag>
        <rdf:li resource="http://mycollege.edu/students/Amy"/>
        <rdf:li resource="http://mycollege.edu/students/Tim"/>
        <rdf:li resource="http://mycollege.edu/students/John"/>
        <rdf:li resource="http://mycollege.edu/students/Mary"/>
        <rdf:li resource="http://mycollege.edu/students/Sue"/>
      </rdf:Bag>
    </s:students>
  </rdf:Description>
</rdf:RDF>
```

## Sequence

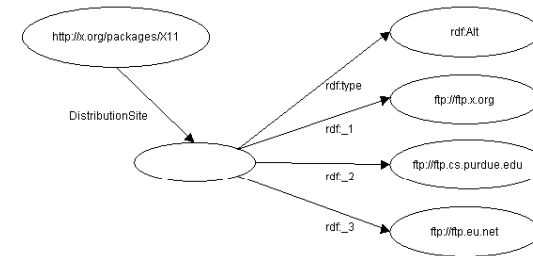
- Danh sách có thứ tự
- Thứ tự các giá trị của 1 thuộc tính mang ý nghĩa.
- Ví dụ thứ tự từ điển.
- Cho phép giá trị trùng nhau.

## Alternative

- Danh sách tài nguyên và hằng giá trị thay thế cho một giá trị thuộc tính.
- Chỉ chọn một trong danh sách các giá trị.

## Ví dụ

Source code của X11 có thể được tìm thấy tại *ftp.x.org*, *ftp.cs.purdue.edu*, hoặc *ftp.eu.net*



## Ví dụ

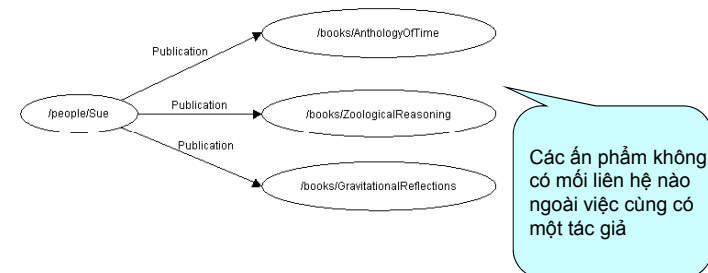
Source code của X11 có thể được tìm thấy tại *ftp.x.org*, *ftp.cs.purdue.edu*, hoặc *ftp.eu.net*

```

<rdf:RDF>
  <rdf:Description about="http://x.org/packages/X11">
    <s:DistributionSite>
      <rdf:Alt>
        <rdf:li resource="ftp://ftp.x.org"/>
        <rdf:li resource="ftp://ftp.cs.purdue.edu"/>
        <rdf:li resource="ftp://ftp.eu.net"/>
      </rdf:Alt>
    </s:DistributionSite>
  </rdf:Description>
</rdf:RDF>
  
```

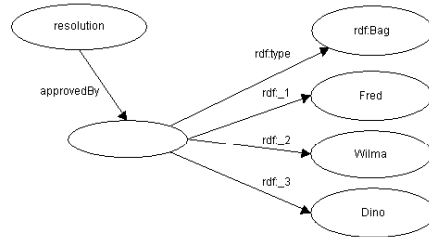
## Thuộc tính bộ chứa vs lập thuộc tính

Sue đã viết các ấn phẩm "Anthology of Time", "Zoological Reasoning", "Gravitational Reflections".



## Tiếp

Hội đồng gồm Fred, Wilma, Dino đồng ý với nghị quyết.



Ba người như một chủ thể.  
Không cần thiết phân tách

## Tuyên bố về tuyên bố khác

### ■ Tuyên bố:

*"Ora Lassila is the creator of the resource <http://www.w3.org/Home/Lassila>".*

### ■ RDF coi đây như một sự thật (fact). Xét tuyên bố:

*"Ralph Swick says that Ora Lassila is the creator of the resource <http://www.w3.org/Home/Lassila>"*

### ■ Không có nghĩa nói về chủ thể Lassila mà là về tuyên bố của Ralph

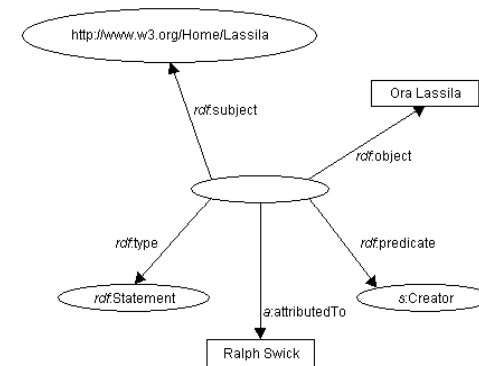
## Tuyên bố về tuyên bố

### ■ Sử dụng các thuộc tính sau:

- **subject**
  - Chỉ định tài nguyên được mô tả bởi tuyên bố. Giá trị của thuộc tính subject là tài nguyên mà tuyên bố nói về (e.g., <http://www.w3.org/Home/Lassila>).
- **predicate**
  - Chỉ định thuộc tính ban đầu
- **object**
  - Giá trị
- **type**
  - Giá trị của thuộc tính này mô tả kiểu của tài nguyên mới tạo ra. **RDF:Statement**.

## Tiếp

*"Ralph Swick says that Ora Lassila is the creator of the resource <http://www.w3.org/Home/Lassila>"*



## Tiếp

"Ralph Swick says that Ora Lassila is the creator of the resource  
<http://www.w3.org/Home/Lassila>"

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/TR/1999/PR-rdf-syntax-19990105#"
  xmlns:a="http://description.org/schema/">
  <rdf:Description>
    <rdf:subject resource="http://www.w3.org/Home/Lassila" />
    <rdf:predicate resource="http://description.org/schema#Creator" />
    <rdf:object>Ora Lassila</rdf:object>
    <rdf:type resource="http://www.w3.org/TR/1999/PR-rdf-syntax-19990105#Statement" />
    <a:attributedTo>Ralph Swick</a:attributedTo>
  </rdf:Description>
</rdf:RDF>
```

## Bài tập: diễn tả bằng RDF

- Oracle Corporation (NASDAQ: ORCL) and Sun Microsystems (NASDAQ: JAVA) announced today they have entered into a definitive agreement under which Oracle will acquire Sun common stock for \$9.50 per share in cash.
- Sun Microsystems, Inc. (NASDAQ: JAVA) develops the technologies that power the global marketplace. Sun can be found in more than 100 countries and on the Web at <http://www.sun.com>.
- Oracle (NASDAQ: ORCL) is the world's largest enterprise software company. For more information about Oracle, please visit our Web site at <http://www.oracle.com>.

50

## Từ vựng RDF

- Tập các URI (không phải các từ)
- Mỗi tổ chức định nghĩa từ vựng riêng cho mình sử dụng các tiền tố URI riêng.
- ví dụ: Dublin Core (dc:title, dc:creator, dc:date, ...).
- RDF không phân tích URIs và không diễn dịch cho bất cứ tiền tố chung nào.
- URI thường chỉ đến vị trí mà nội dung thông tin về tài nguyên có thể được tìm thấy.

51

## FOAF

- Dự án FOAF (<http://www.foaf-project.org/>): định nghĩa thuật ngữ (bởi RDF) cho các trang homepage cá nhân, tổ chức, cty mà máy tính có thể hiểu..
- Tập trung mô tả người, liên kết giữa người và mọi thứ khác trên Web:
  - tạo ra tài liệu,
  - tham dự các cuộc họp,
  - có trong các bức ảnh, vv
- FOAF Vocabulary: <http://xmlns.com/foaf/0.1/>

```
foaf:name
foaf:knows
foaf:homepage
foaf:weblog
foaf:mbox
```

52

## Bài tập tìm hiểu

### ■ Dublin Core

- đặc tả: <http://dublincore.org/documents/dces/>
- Namespace:  
xmlns:dc="http://purl.org/dc/elements/1.1/"

### ■ FOAF

- đặc tả: <http://xmlns.com/foaf/spec/>
- Namespace:  
xmlns:foaf="http://xmlns.com/foaf/0.1/"

53

## Bài tập tìm hiểu

### ■ Recent Dublin Core enhancement: DCMI Metadata Terms

- Specification:  
<http://dublincore.org/documents/dcmi-terms/>
- Namespace:  
xmlns:dcterms=<http://purl.org/dc/terms/>

### ■ RSS 1.0

- [http://en.wikipedia.org/wiki/RSS \(file format\)](http://en.wikipedia.org/wiki/RSS_(file_format))

54

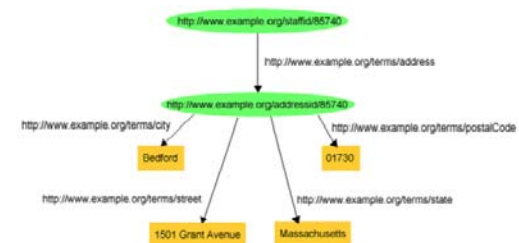
## Nút trắng

### ■ Đôi khi trong quá trình chi tiết hóa việc mô tả các quan hệ

- Thêm nút mới
- Nếu nút này không bao giờ được tham chiếu tới ngoài đồ thị RDF hiện tại → không nhất thiết phải có tên.

55

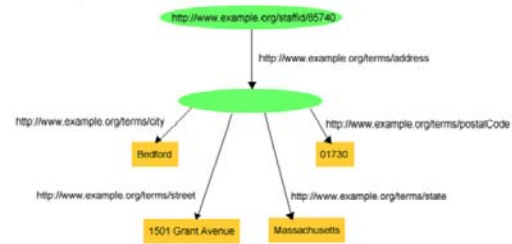
## Ví dụ



exstaff:85740	exterm:address	exaddressid:85740 .
exaddressid:85740	exterm:street	"1501 Grant Avenue" .
exaddressid:85740	exterm:city	"Bedford" .
exaddressid:85740	exterm:state	"Massachusetts" .
exaddressid:85740	exterm:postalCode	"01730" .

56

## Ví dụ



exstaff:85740	exterm:address	_joaddress.
_joaddress	exterm:street	"1501 Grant Avenue" .
_joaddress	exterm:city	"Bedford" .
_joaddress	exterm:state	"Massachusetts" .
_joaddress	exterm:postalCode	"01730" .

57