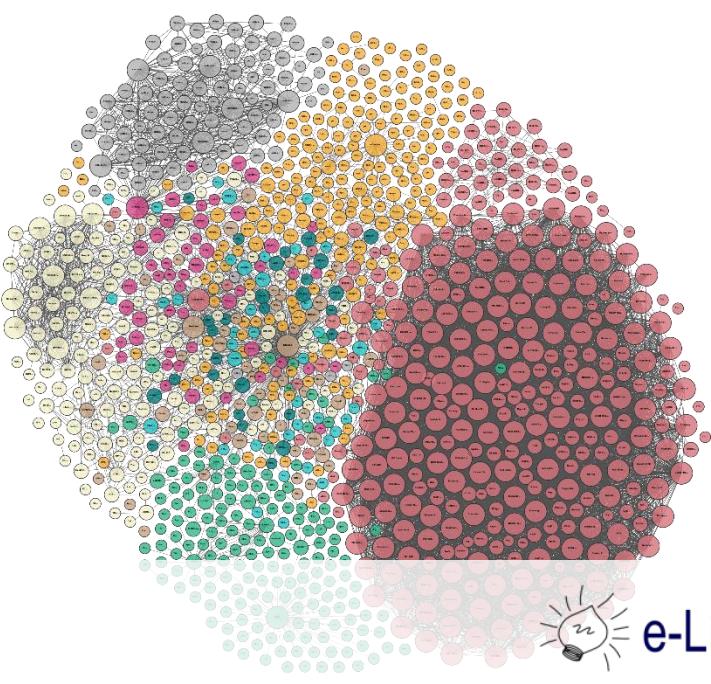
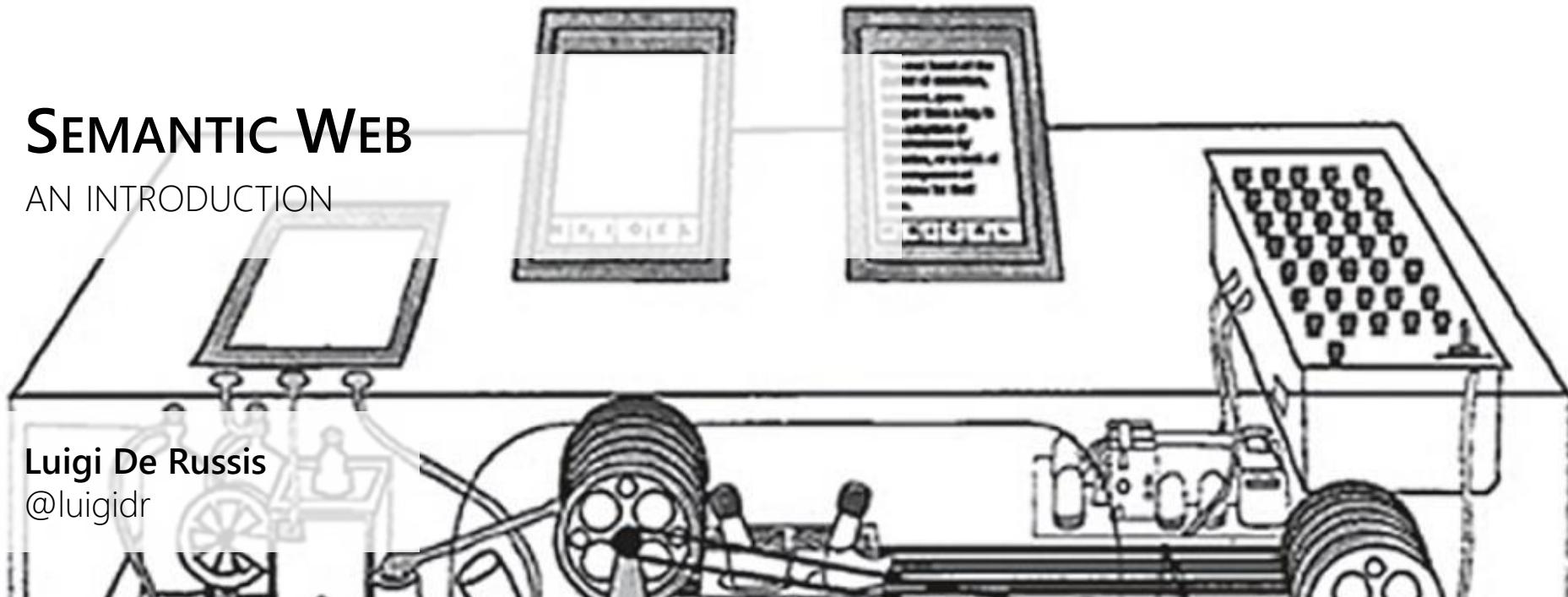
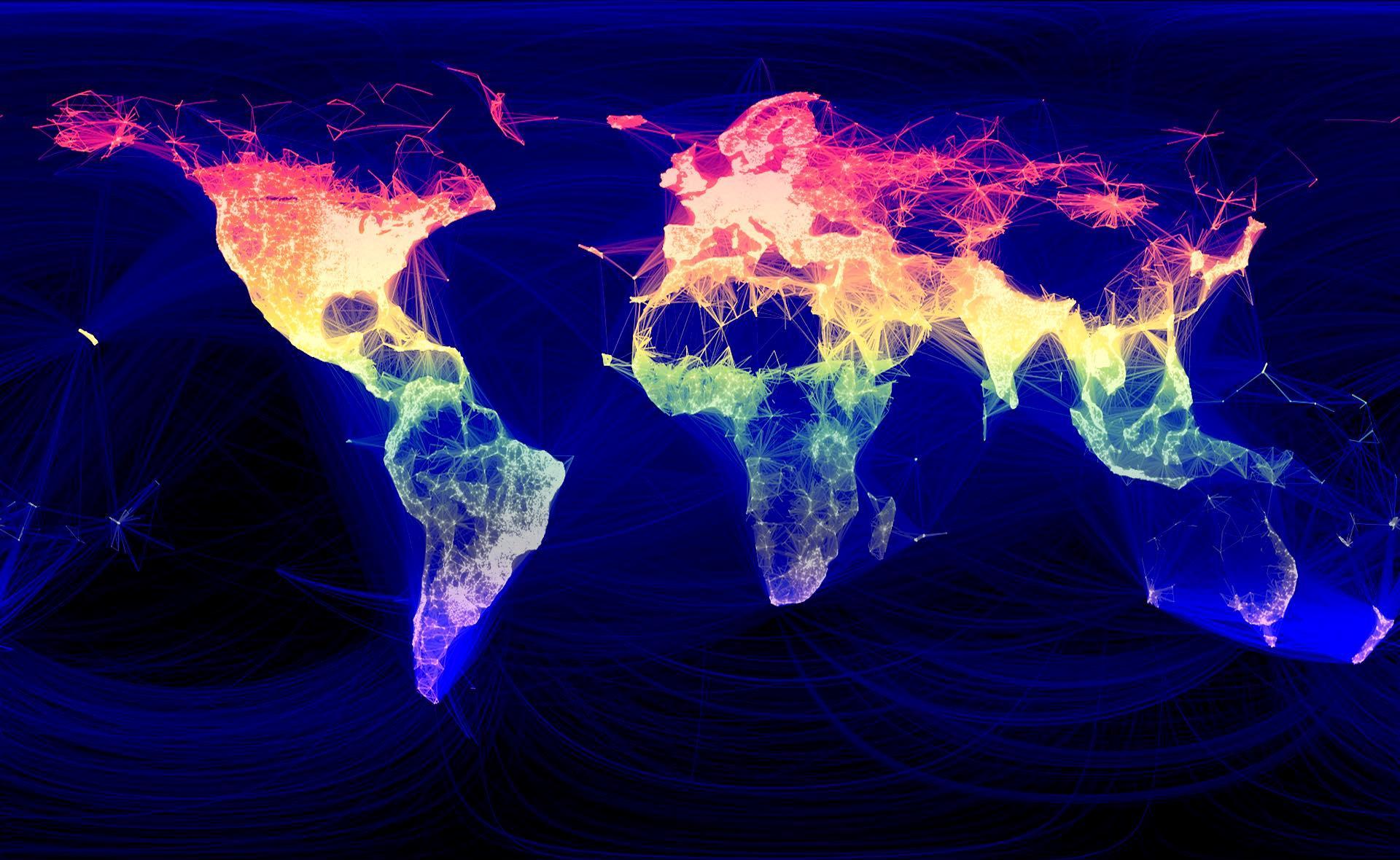


SEMANTIC WEB

AN INTRODUCTION

Luigi De Russis
@luigidr



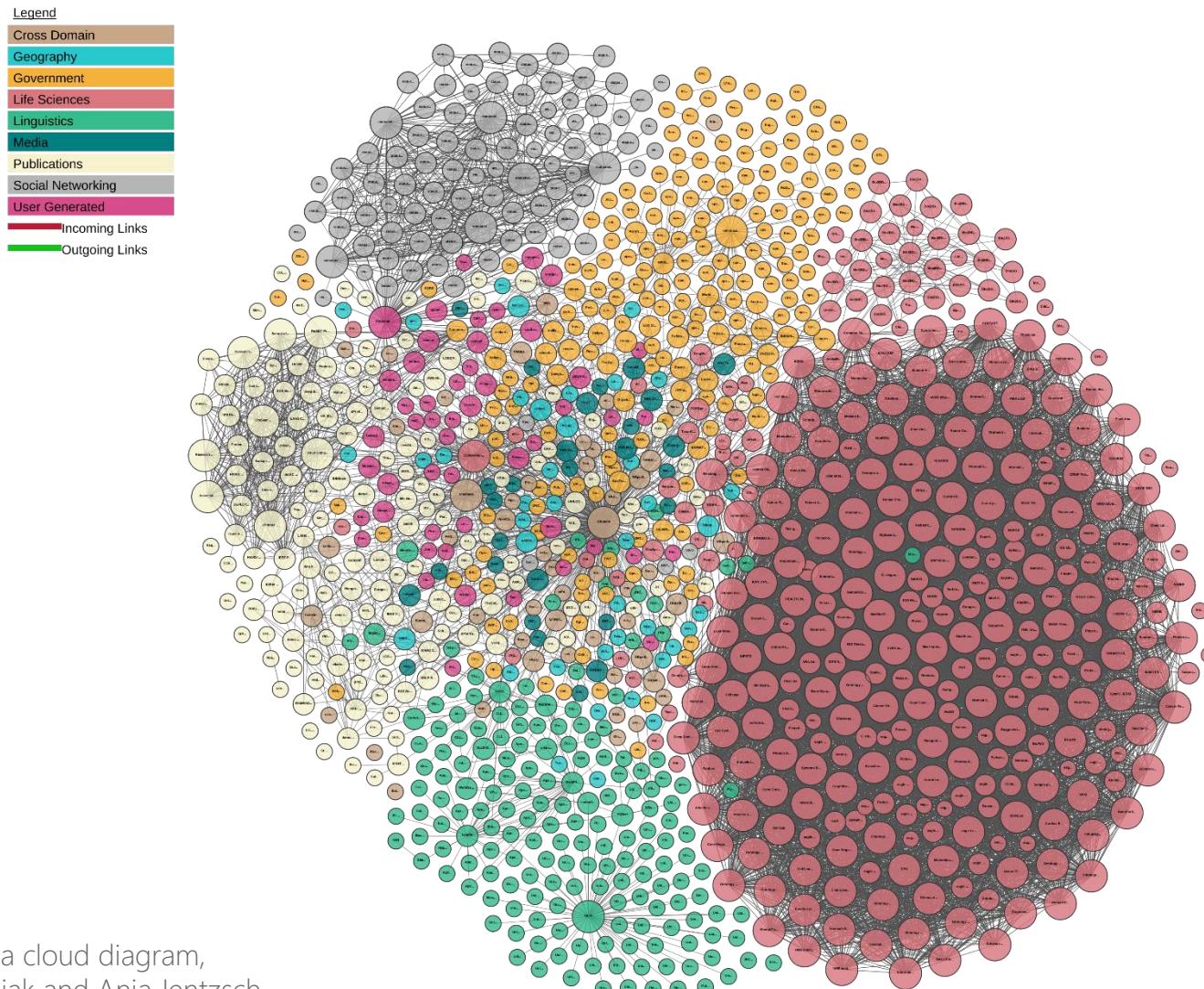


THE WEB IS A **WEB OF DOCUMENT** FOR PEOPLE,
NOT FOR MACHINES

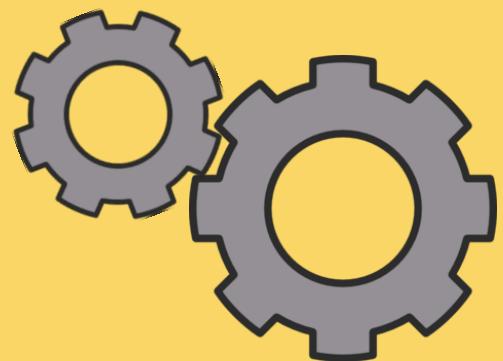
THE WEB IS A WEB OF DOCUMENT



THE SEMANTIC WEB IS A WEB OF DATA



THINK



HOW TO GET DATA FROM THE WEB?

AN EXAMPLE

We would like to create an online catalog of all the Computer Engineering degrees at university level

HOW TO GET DATA FROM THE WEB?

AN EXAMPLE

We would like to create an online catalog of all the Computer Engineering degrees at university level

The screenshot shows the official website of the Politecnico di Torino. At the top, there's a blue header bar with the university's logo and name. Below it, a white sidebar on the left contains links for "Offera formativa" (with a description), "Apply@polito" (with a link to the application form), "Orientamento" (with a link to orientation information), and "News - eventi - avvisi". The main content area features a large banner for the "POLY2NUC" course, which includes a deadline of 31.07.2017.

This screenshot shows the website of the Politecnico di Milano. It features a similar header with the university's logo and name. A prominent banner in the center promotes the "POLY2NUC" course, listing dates from 26.05.2017 to 05.06.2017. Below the banner, there's a section for the "SCUOLA DI ARCHITETTURA" (School of Architecture).

The screenshot shows the website of Sapienza Università di Roma. The header includes the university's logo and name, along with links for various university departments. A large red banner in the center encourages donations, stating "Il tuo 5 diventa 1000" (Your 5 becomes 1000) and "DONA IL 5 PER MILLE ALLA SAPIENZA". Below the banner, there are three student portraits with captions: "Iscrizioni 2017-2018 aggiornamenti e novità", "#SonoUnDonatore", and "Prendi buoni voti, pagherai meno tasse!".

HOW TO GET DATA FROM THE WEB?



Data is locked in
"data islands"



Limited or no
access to this
data

{JSON}



DATA ON THE WEB IS NOT ENOUGH

WE NEED A PROPER INFRASTRUCTURE

DATA SHOULD BE **AVAILABLE** ON THE WEB

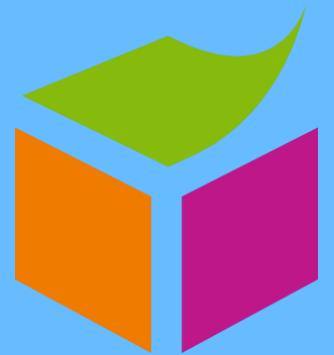
ACCESSIBLE AND STRUCTURED VIA STANDARD WEB TECHNOLOGIES
NOT CONTROLLED BY APPLICATIONS, ONLY

DATA SHOULD BE **INTERLINKED** OVER THE WEB

I.E., DATA CAN BE INTEGRATED OVER THE WEB

THIS IS WHERE THE SEMANTIC WEB COMES IN

DEFINITIONS





To a computer, the Web is a flat, boring world, devoid of meaning. This is a pity, as in fact documents on the Web describe real objects and imaginary concepts. [...]

Adding semantics to the Web involves two things: allowing documents which have information in **machine-readable** forms, and allowing links to be created with **relationship values**. Only when we have this extra level of semantics we will be able to use computer power to help us exploit the information to a greater extent than our own reading.

TIM BERNERS-LEE, 1994

THE WEB, 1994 C.A.



Yellow Pages · People Search · City Maps · News Headlines · Stock Quotes · Sports Scores

Search Options

- Arts - - Humanities, Photography, Architecture, ...
- Business and Economy [Xtra] - - Directory, Investments, Classifieds, ...
- Computers and Internet [Xtra] - - Internet, WWW, Software, Multimedia, ...
- Education - - Universities, K-12, Courses, ...
- Entertainment [Xtra] - - TV, Movies, Music, Magazines, ...
- Government - - Politics [Xtra], Agencies, Law, Military, ...
- Health [Xtra] - - Medicine, Drugs, Diseases, Fitness, ...
- News [Xtra] - - World [Xtra], Daily, Current Events, ...
- Recreation and Sports [Xtra] - - Sports, Games, Travel, Autos, Outdoors, ...

NEW!

Search White House Press Releases, Radio Addresses, Photos and Web Pages

- To search White House Press Releases, Radio Addresses, Photos and Web Pages, enter a TERM or PHRASE in the box below which describes your topic of interest (for example, "social security benefits for retired people").

TERM/PHRASE

- Adjust the START and END dates to limit your search to a specific timeframe. Select from the ITEMS list the number of documents you would like to receive each, then indicate the order in which your results will appear. By "DATE" will return the most recent documents first. By "RELEVANCE" will return the most relevant documents first.

START DATE

 01 19 1993

ITEMS RETURNED FROM EACH CATEGORY

 10

© DATE RELEVANCE

[Reset this search form](#) [Submit this search form](#) [Start with a new search form](#)



Welcome to Apple



Introducing CyberDrive

Register today for a free CD-ROM.

EMIATE 300

Mobile,
Affordable,
& Smart



MOVIES
FROM MARS

QuickTime VR
Takes You Out
of this World

What's Hot

Preorder Mac OS 8

Now you can [preorder Mac OS 8](#), described by Macworld as "the most comprehensive update to the Mac OS in years, sporting a bold new

Be the First to Know

Learn about new Macintosh software releases the moment they become available. Check [Hot Mac Products](#) to hear about programs like Speed

No GOOGLE, WIKIPEDIA, AMAZON,
... YET!



I have a dream for the Web [in which computers] become capable of analyzing all the data on the Web – the content, links, and transaction between people and computers.

A “Semantic Web”, which should make this possible, has yet to emerge, but when it does, the day-to-day mechanisms of trade, bureaucracy and our daily lives will be handled by machines talking to machines. The “intelligent agents” people have touted for ages will finally materialize.

TIM BERNERS-LEE, 1999

THE SEMANTIC WEB IS A WEB OF DATA

THE SEMANTIC WEB **IS** THE WEB

SAME BASE TECHNOLOGIES, EVOLUTIONARY, DECENTRALIZED

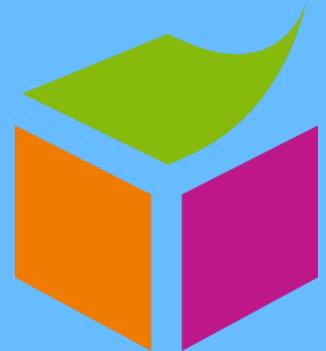
IT IS ABOUT COMMON FORMATS

FOR INTEGRATION AND COMBINATION OF DATA DRAWN FROM
DIVERSE SOURCES

IT IS ABOUT A LANGUAGE

FOR RECORDING HOW THE DATA RELATES TO REAL WORLD OBJECTS

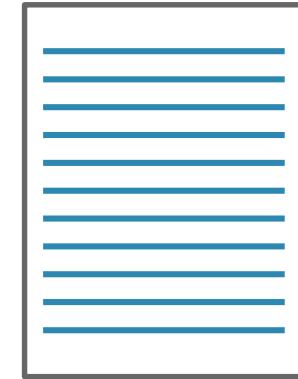
FUNDAMENTALS



RESOURCE AND DESCRIPTION

RESOURCE

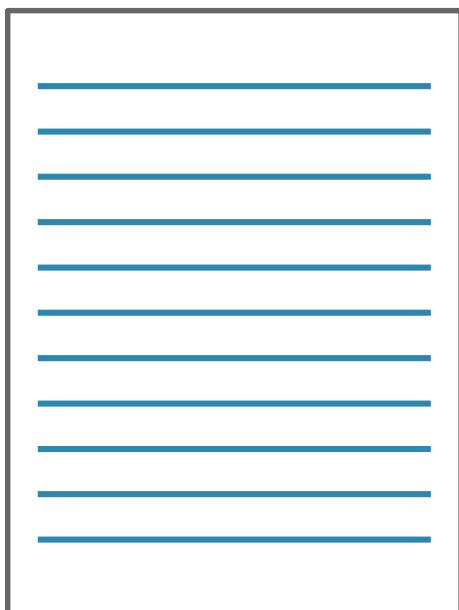
EVERY DOCUMENT "REACHABLE" ON THE WEB
NO MATTER THE CONTENT, FORMAT, LANGUAGE, ETC.



RESOURCE DESCRIPTION

INDEPENDENT FROM THE FORMAT
STANDARD LANGUAGE (BASED ON METADATA)

RESOURCES



DESCRIPTION



DESCRIPTION



Title
Author
Date
Topic
Quality



Title
Author
Date
Topic

















OK, IT'S TIME TO COOK!

URIs ↗

RDF ↗

↗ SPARQL

↗ RDFS, OWL

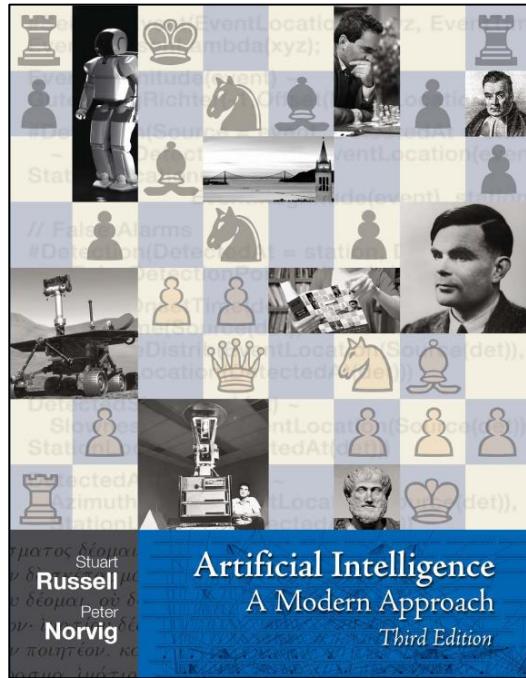
↗ OWL, SWRL

MODELING DATA



EXAMPLE: BOOKSTORE

REPRESENT THE FOLLOWING DATA ABOUT THE AI BOOK AS A SET OF RELATIONS



TITLE: "ARTIFICIAL INTELLIGENCE: A MODERN APPROACH"

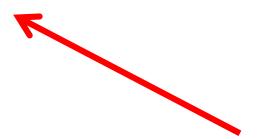
AUTHOR: RUSSEL, STUART AND NORVIG, PETER

PUBLISHER: PRENTICE HALL

ISBN: 978-0136042594

EXAMPLE: BOOKSTORE

`http://...isbn/978013604
2594`



Resource

EXAMPLE: BOOKSTORE

http://...isbn/978013604
2594

Artificial Intelligence: A
Modern Approach

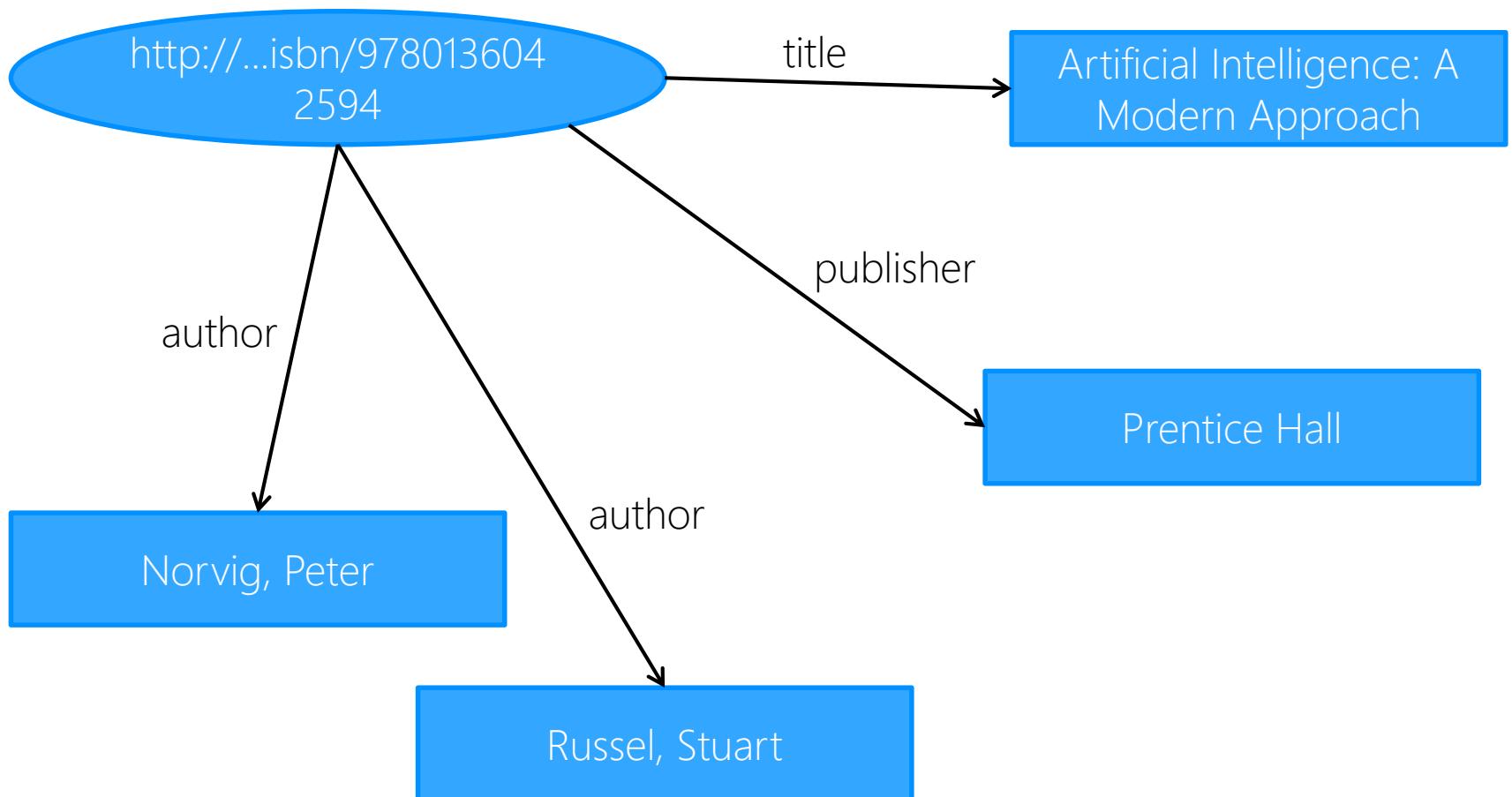
Literal



EXAMPLE: BOOKSTORE



EXAMPLE: BOOKSTORE



RDF: RESOURCE DESCRIPTION FRAMEWORK

STRUCTURED IN STATEMENTS

http://...isbn/978013604

2594

SUBJECT

A RESOURCE (URI)

PREDICATE

A VERB, PROPERTY, OR RELATIONSHIP

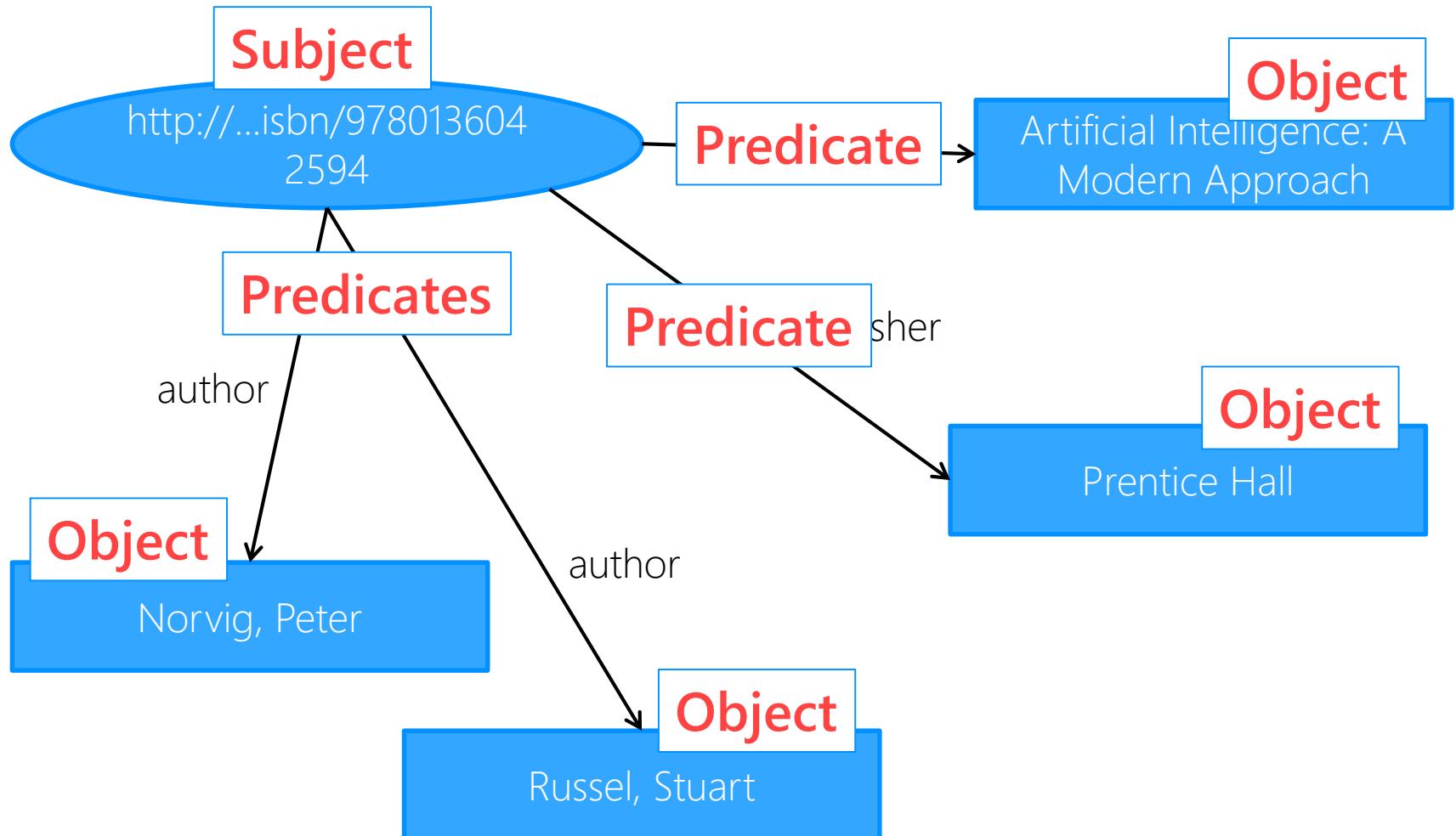
author

Norvig, Peter

OBJECT

A RESOURCE OR A LITERAL STRING

EXAMPLE: BOOKSTORE



EXAMPLE: BOOKSTORE



RDF IN XML SYNTAX

```
<rdf:RDF xmlns:rdf="http://www.w3.org/.../22-rdf-syntax-ns#">
  <rdf:Description about="http://... isbn/9780136042594">
    <title>Artificial Intelligence: A Modern Approach</title>
  </rdf:Description>
</RDF>
```

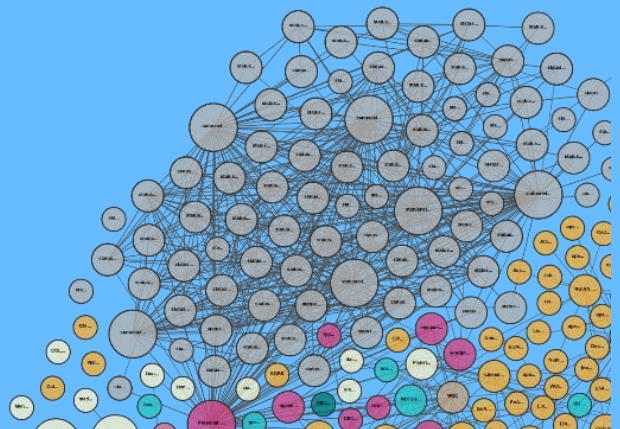
EXAMPLE: BOOKSTORE



RDF IN TURTLE

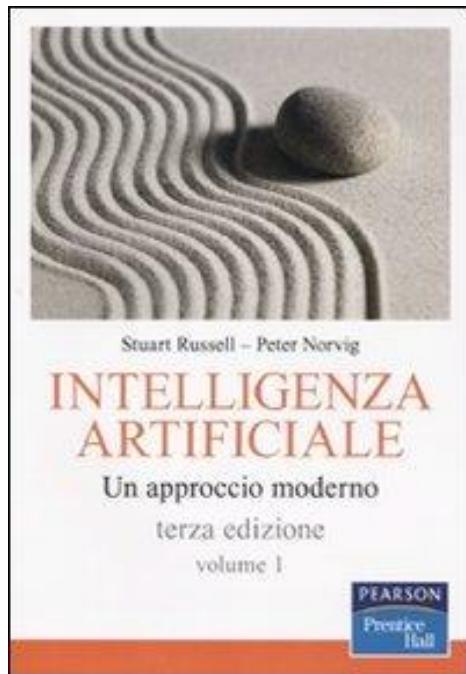
```
<http://... isbn/9780136042594>
    title "Artificial Intelligence: A Modern Approach"
```

LINKING DATA



EXAMPLE: BOOKSTORE

REPRESENT THE FOLLOWING DATA ABOUT THE ITALIAN TRANSLATION OF THE AI BOOK AS A SET OF **RELATIONS**



TITLE: "INTELLIGENZA ARTIFICIALE. UN APPROCCIO MODERNO"

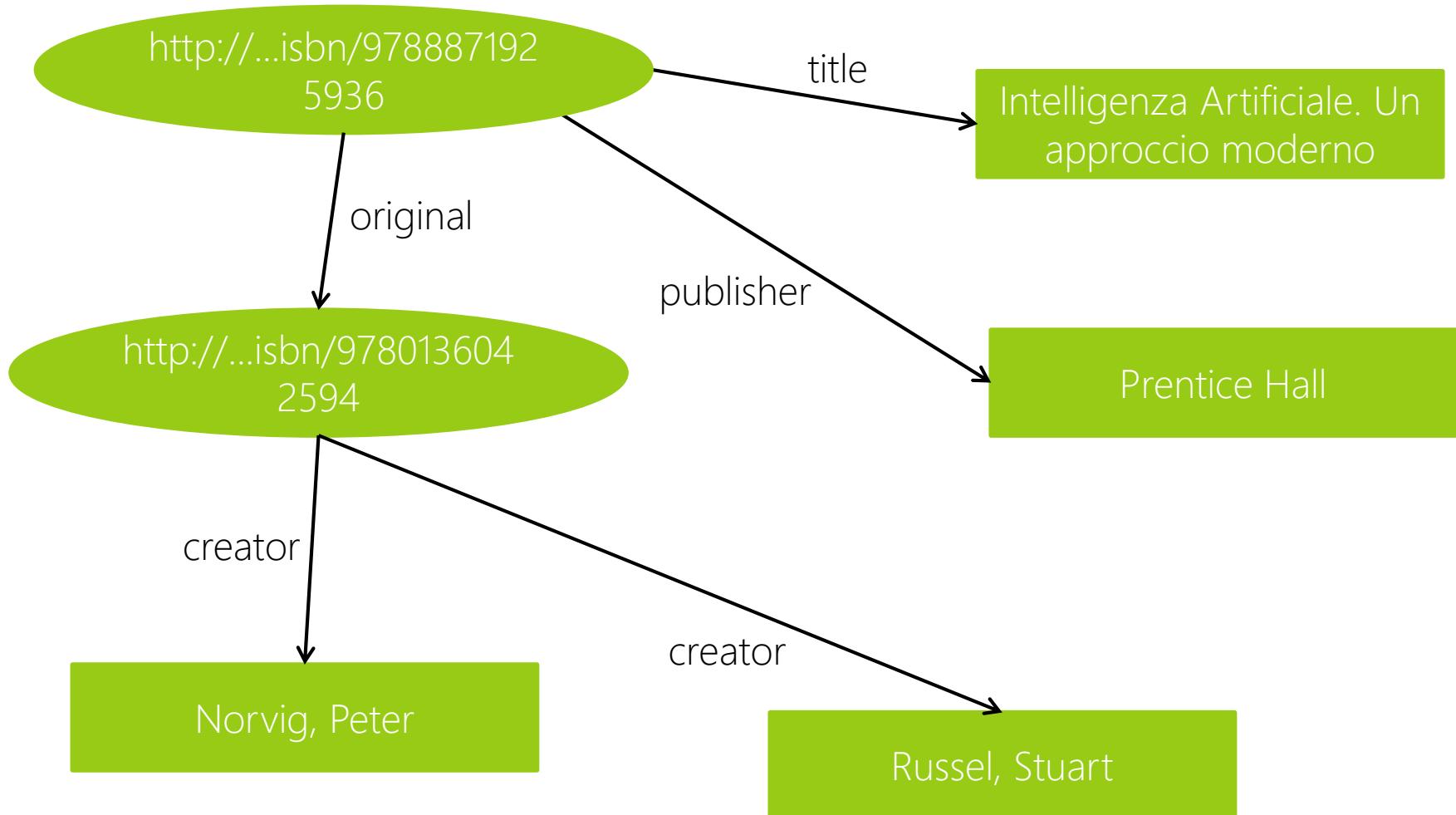
AUTHOR: RUSSEL, STUART AND NORVIG, PETER

PUBLISHER: PRENTICE HALL

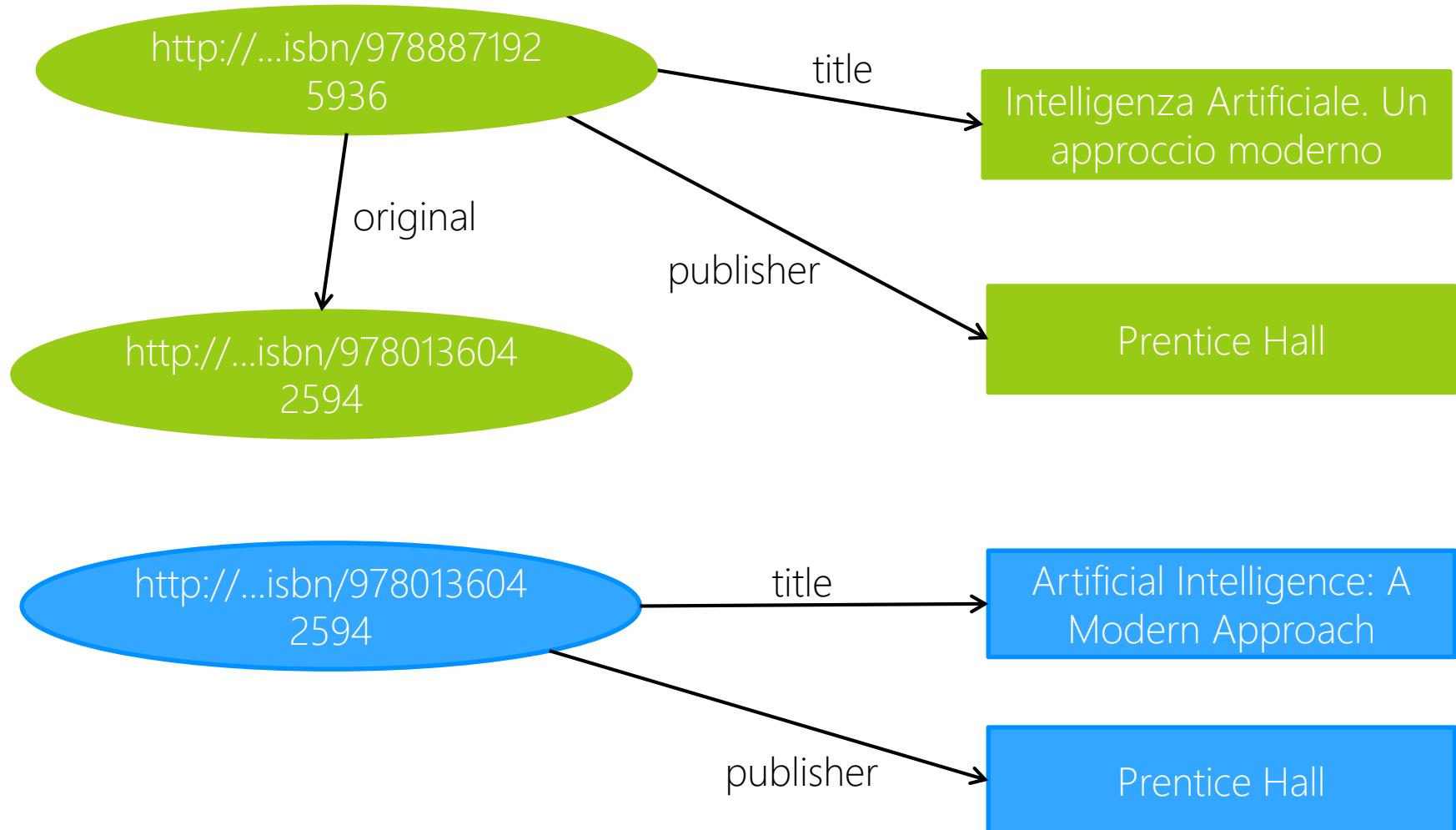
ISBN: 978-8871925936

ORIGINAL ISBN: 978-0136042594

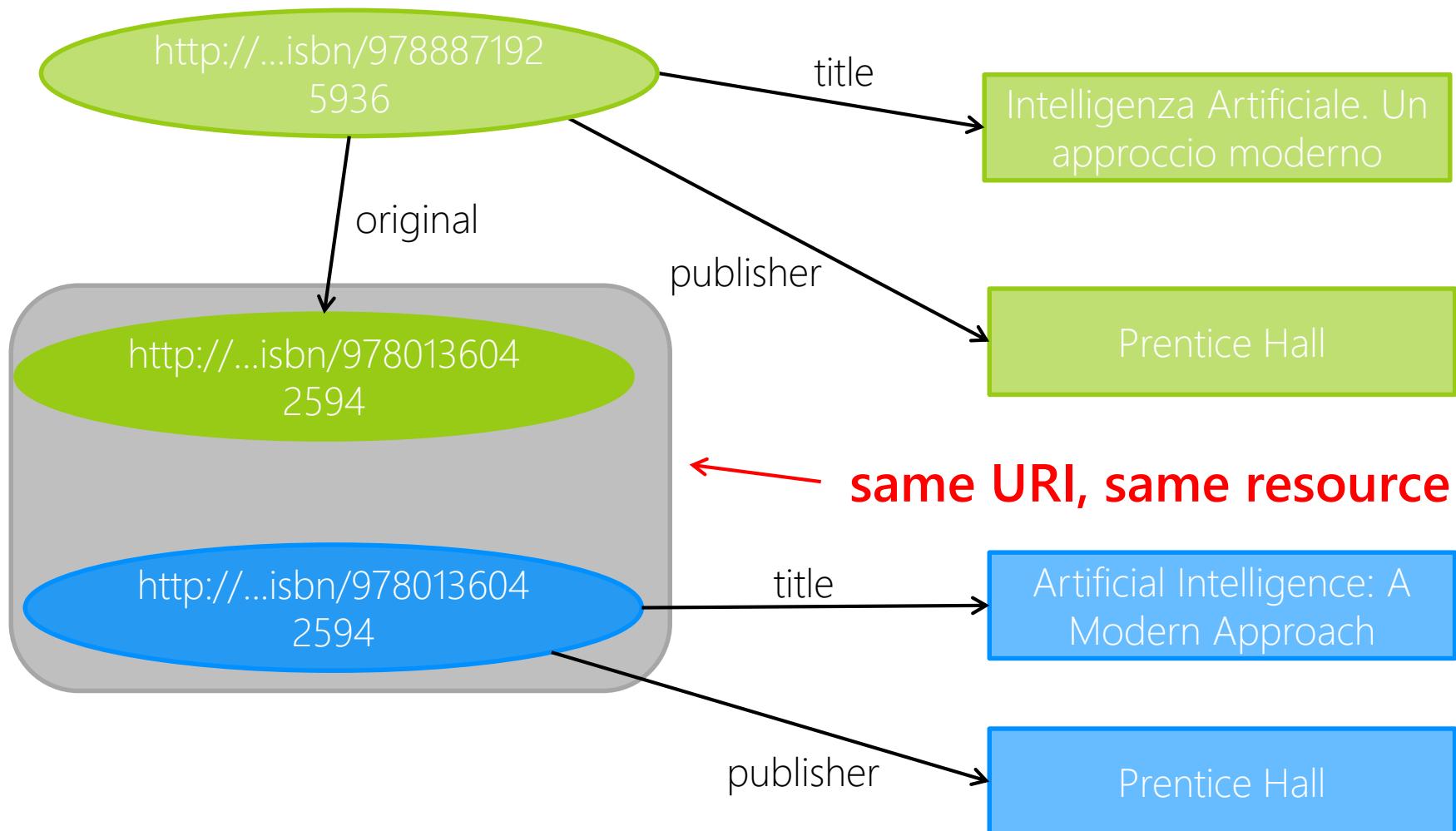
EXAMPLE: BOOKSTORE



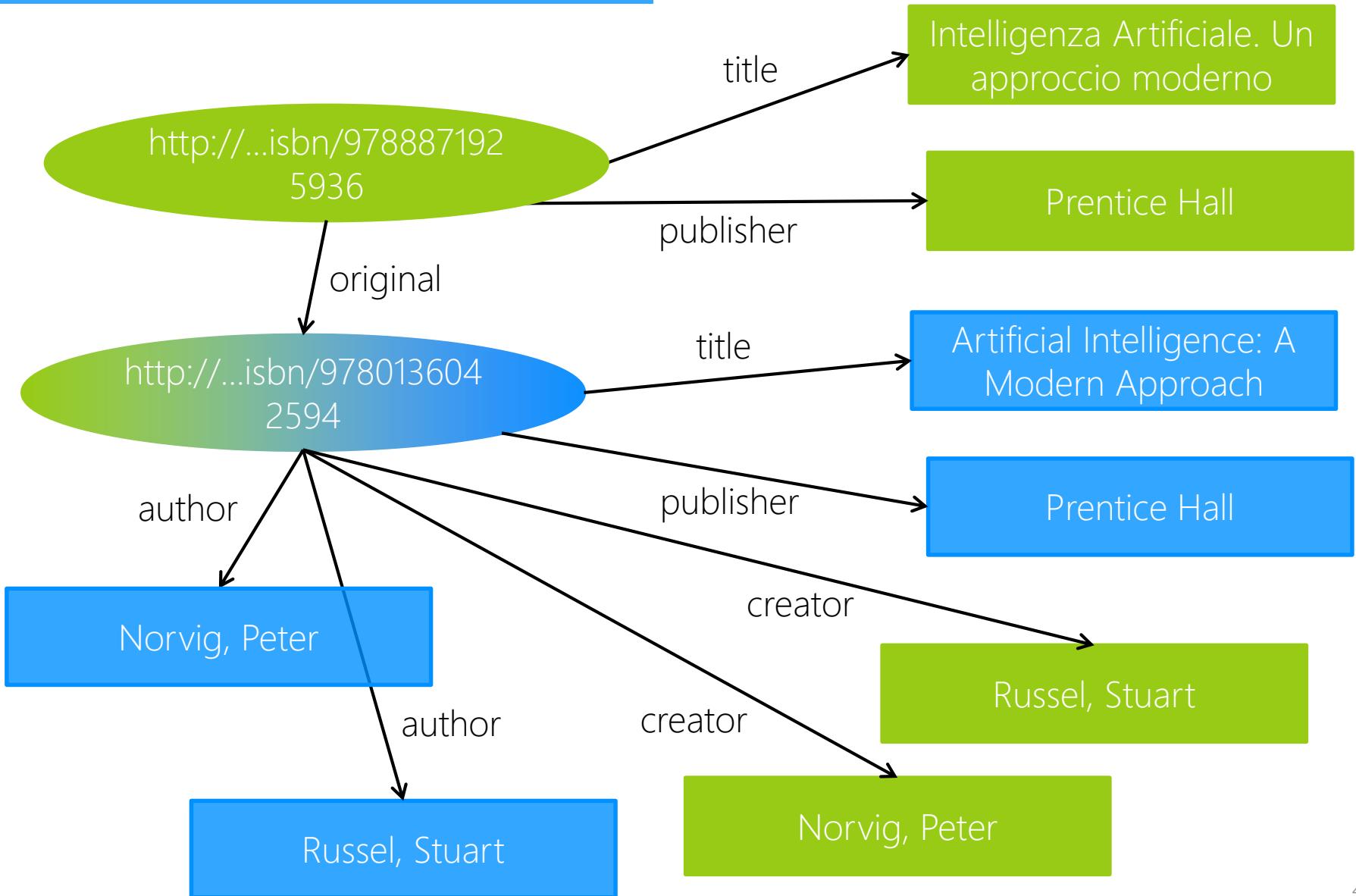
EXAMPLE: BOOKSTORE



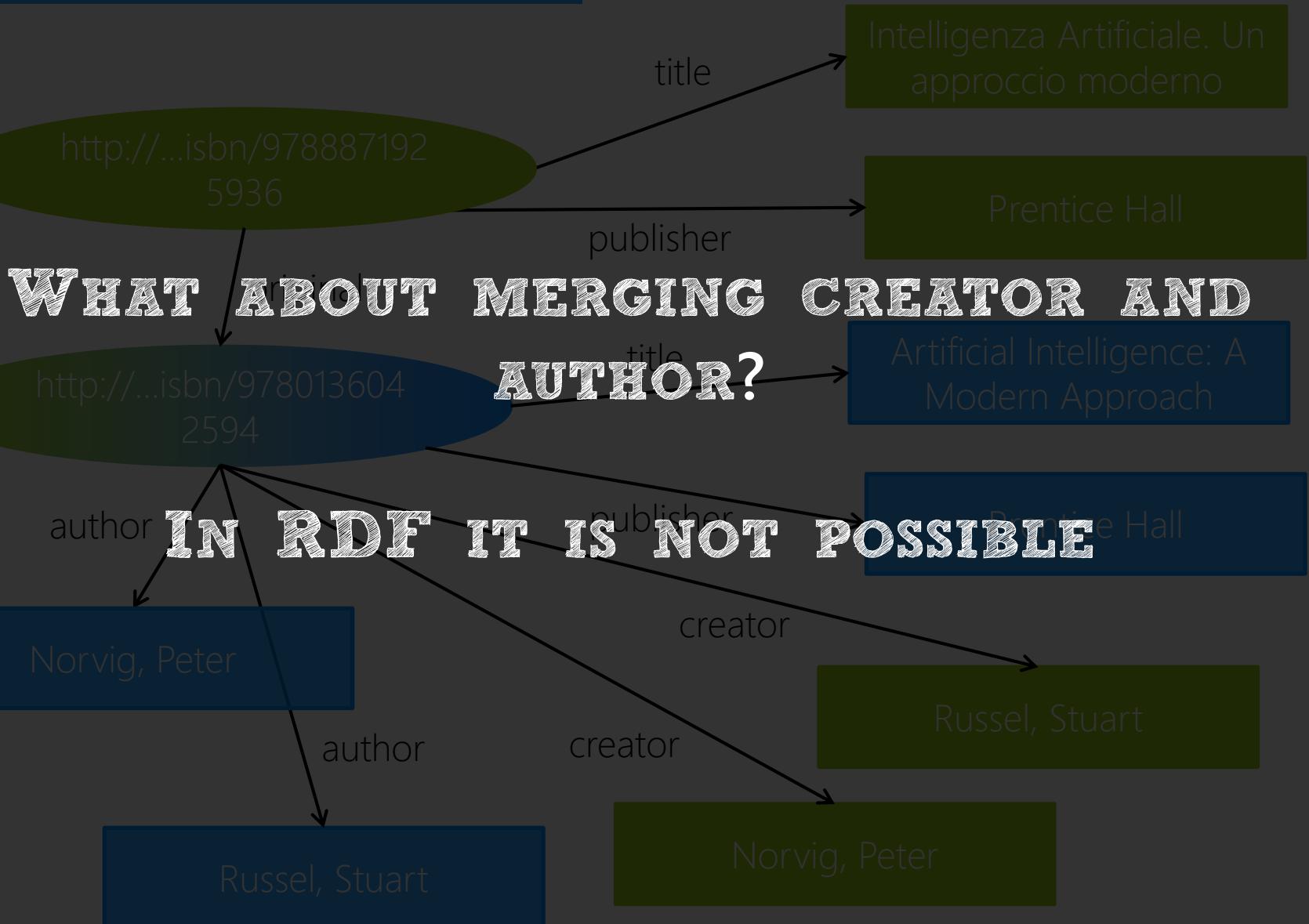
EXAMPLE: BOOKSTORE



EXAMPLE: BOOKSTORE



EXAMPLE: BOOKSTORE



PROBLEM: FIELD NAMES ARE ARBITRARY

SYNONYMS

AUTHOR OR CREATOR OR MAKER OR
CONTRIBUTOR OR...

SINGULAR OR PLURAL

AUTHOR OR AUTHORS

SOLUTION: STANDARDS

GENERAL OR DOMAIN-SPECIFIC

PROBLEM: FIELD VALUES ARE ARBITRARY

VALUE TYPE

STRING, DATE, INTEGER...

VALUE FORMAT

"NORVIG, PETER" OR " NORVIG, P." OR
"PETER NORVIG" OR...

VALUE RESTRICTIONS

ONE VALUE OR MULTIPLE VALUES (HOW
MANY?)

SOLUTIONS:

STANDARDS

CONTROLLED VOCABULARY

SEMANTICALLY RICH

DESCRIPTIONS

DUBLIN CORE



Dublin Core Elements

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

GENERAL VOCABULARY

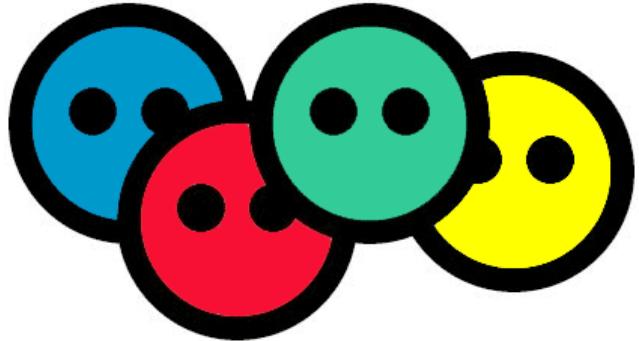
DUBLIN CORE METADATA INITIATIVE (DCMI)

[HTTP://DUBLINCORE.ORG](http://dublincore.org)

BUILDING BLOCKS TO DEFINE METADATA FOR THE SEMANTIC WEB

DEFINE TITLE, CONTRIBUTOR, PUBLISHER, LICENSE, DATE, LANGUAGE, ETC.

FRIEND OF A FRIEND (FOAF)



GENERAL ONTOLOGY

DESCRIBE PERSONS, THEIR ACTIVITIES AND THEIR RELATIONS TO OTHER PEOPLE AND OBJECTS

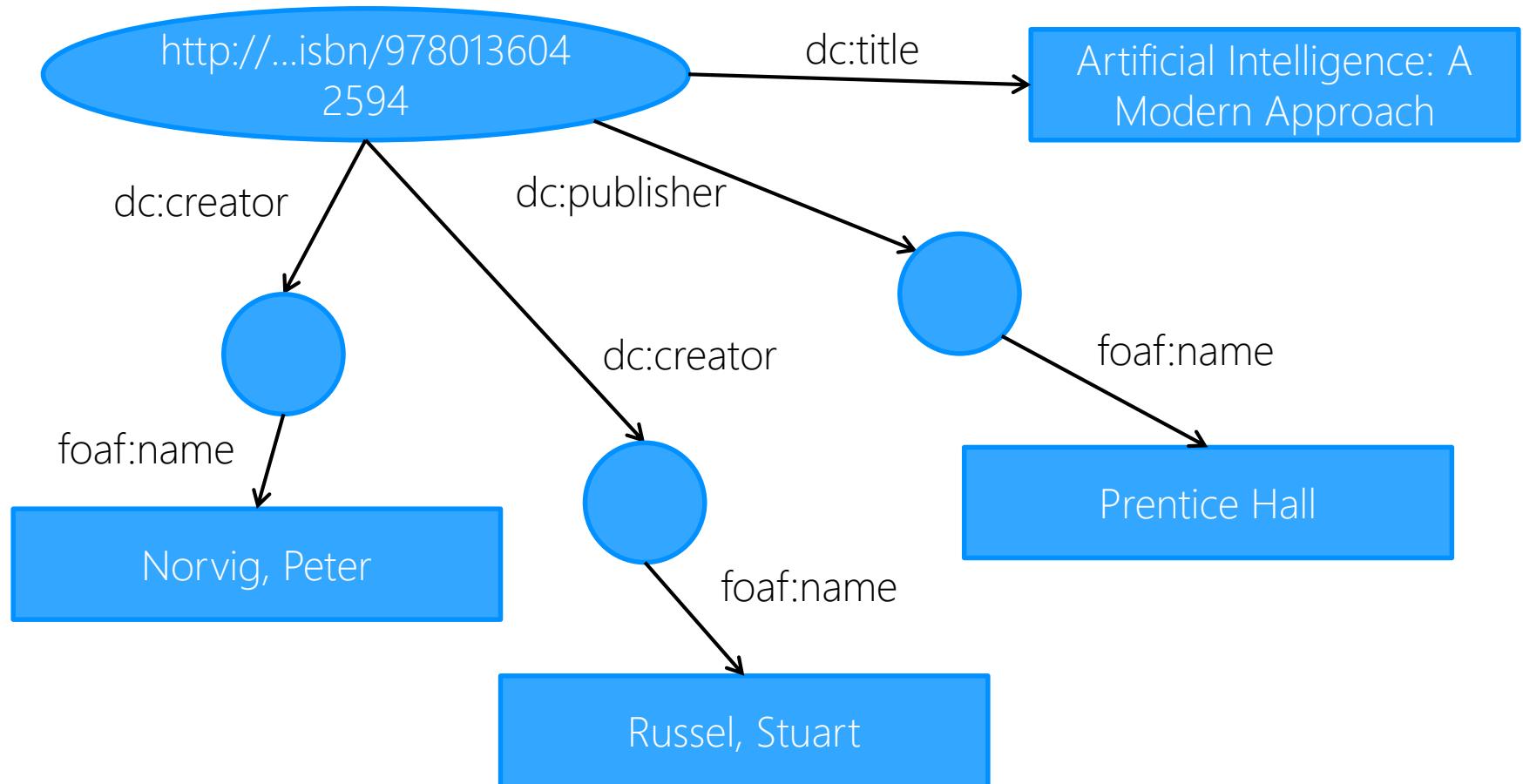
[HTTP://WWW.FOAF-PROJECT.ORG](http://www.foaf-project.org)

**BUILDING BLOCKS TO DEFINE STRUCTURED
RELATIONS BETWEEN PEOPLE**

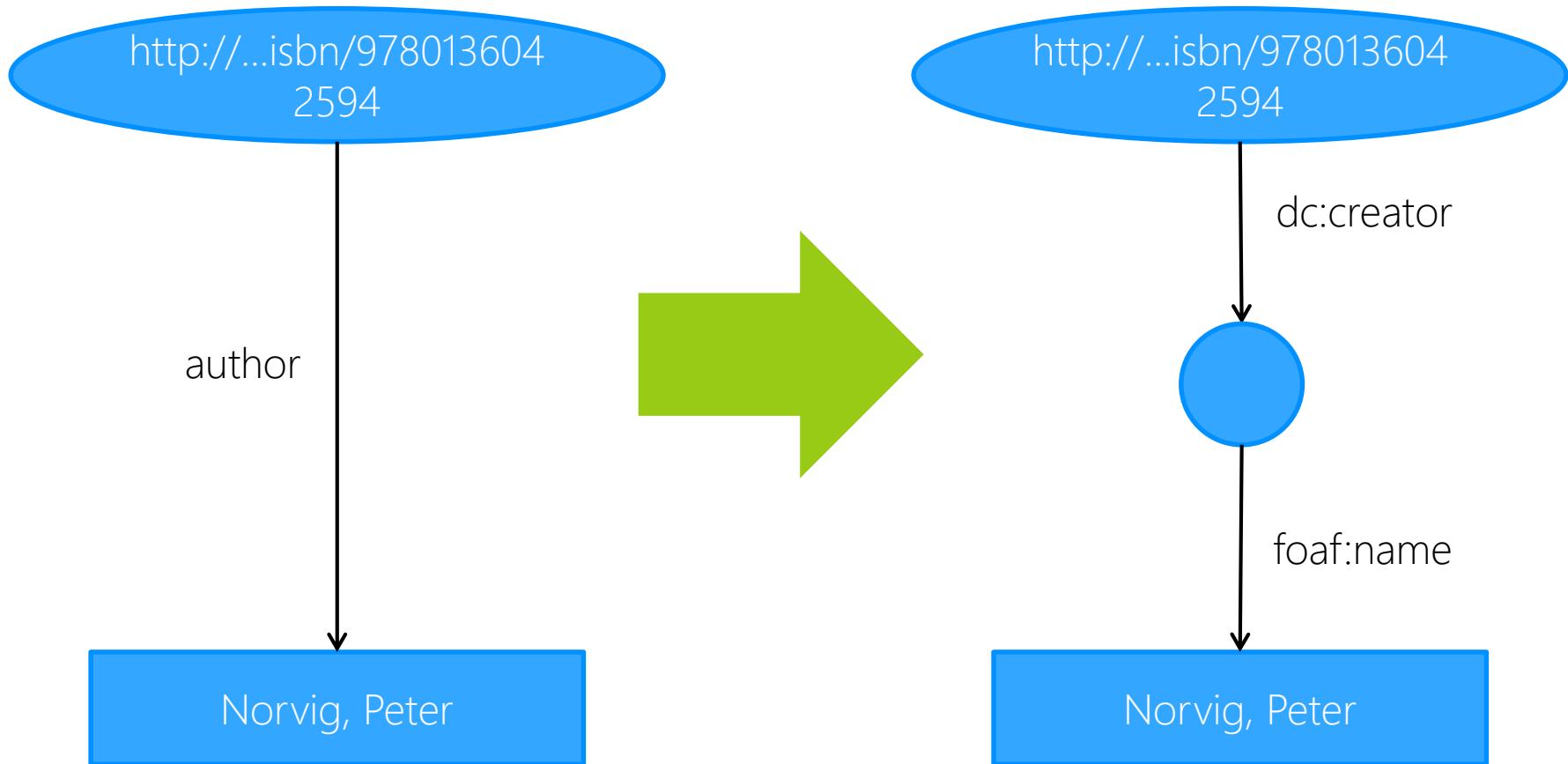
DEFINE NAME, FAMILYNAME, GIVENNAME, KNOWS, AGE, NICK, ETC.

EXAMPLE: BOOKSTORE

foaf: <http://xmlns.com/foaf/spec>
dc: <http://purl.org/dc/terms>



WHY?



RDF SCHEMA



RDF SCHEMA (RDFS)

SCHEMA

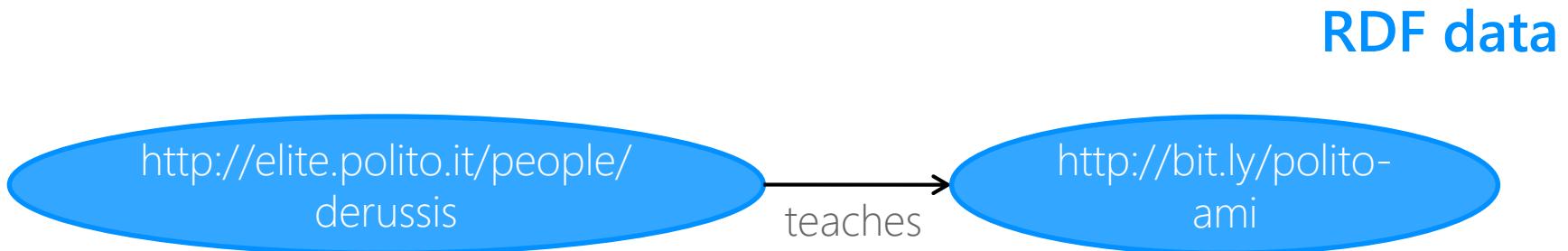
DEFINITION OF THE NODES AND PREDICATES USED IN A RDF DOCUMENT

RDFS CAN BE SEE AS A "META-MODEL"

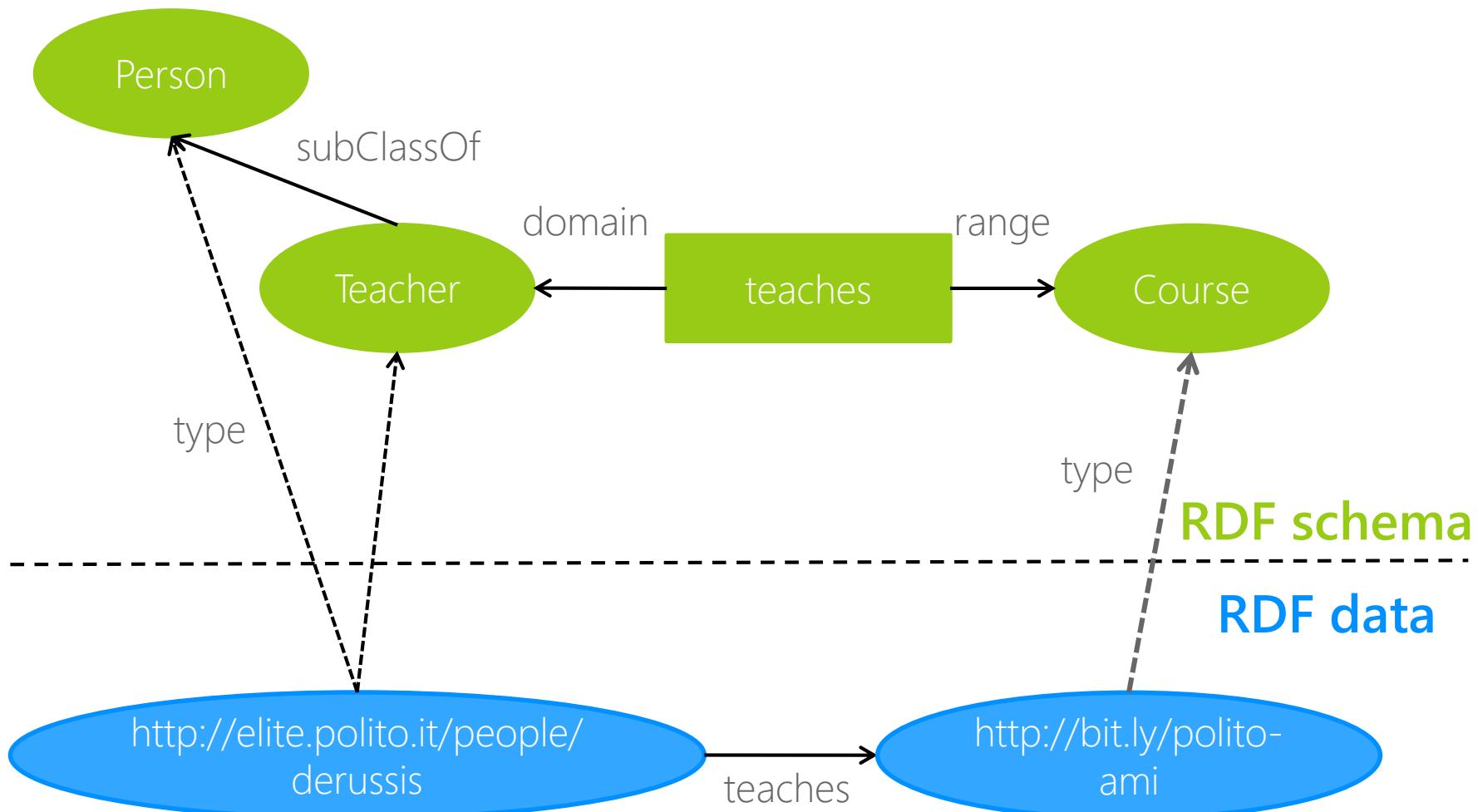
DOMAIN AND RANGE

RDFS DESCRIBES PROPERTIES IN TERMS OF CLASSES OF RESOURCE
TO WHICH THEY APPLY
(FROM A "DOMAIN" TO A "RANGE")

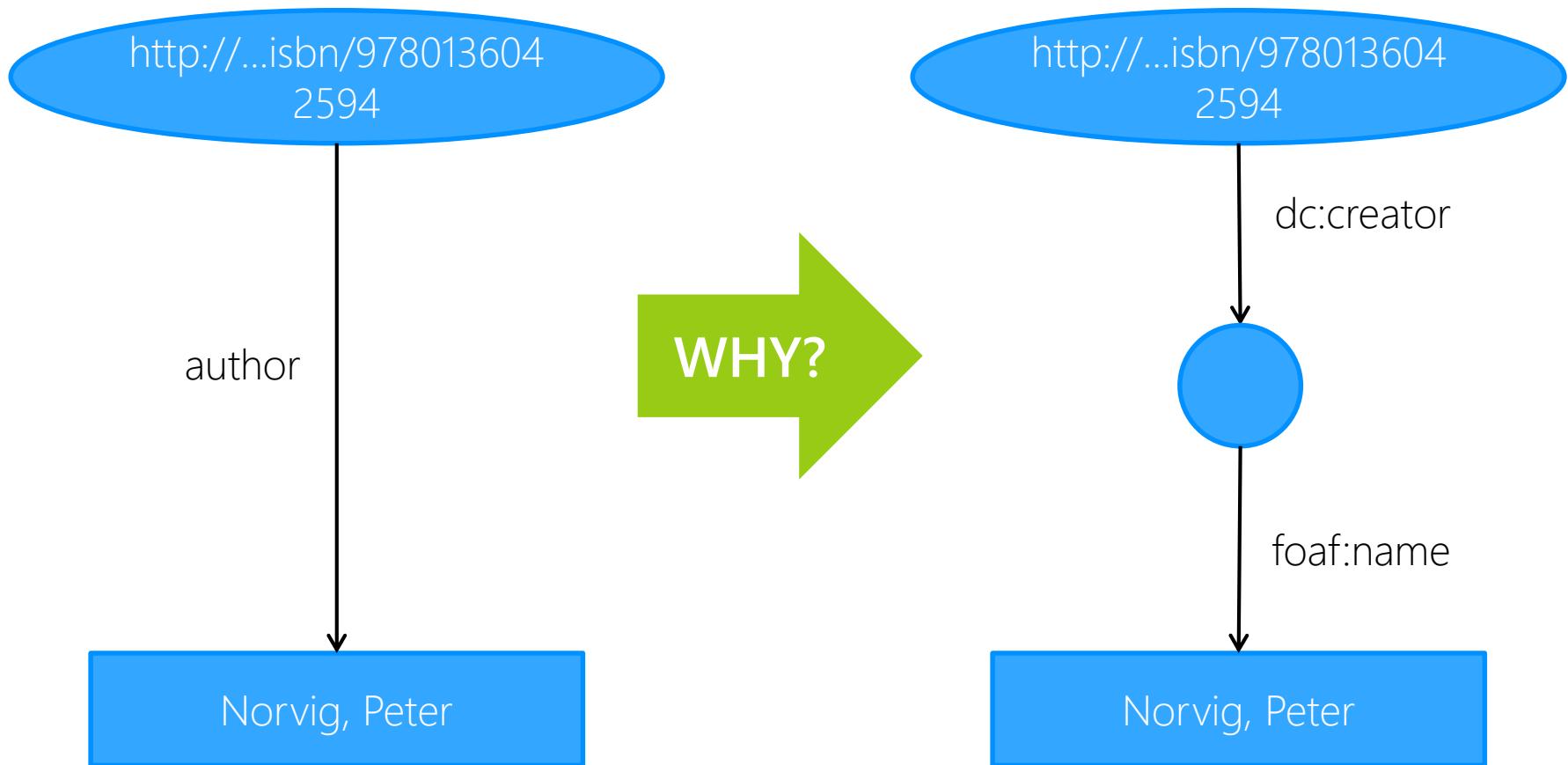
EXAMPLE



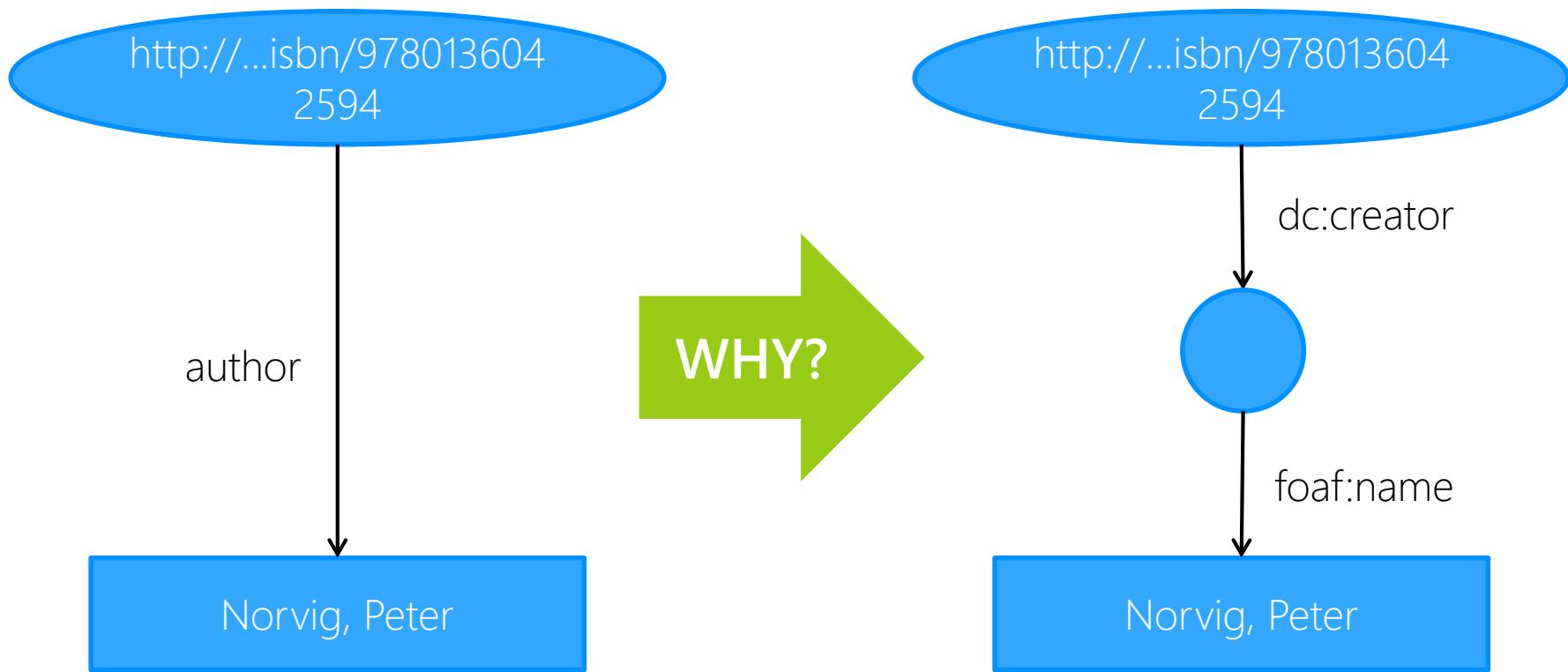
EXAMPLE



BACK TO THE BOOKSTORE EXAMPLE...



BACK TO THE BOOKSTORE EXAMPLE...



DC:CREATOR HAS RANGE AGENT, I.E., A CLASS (RESOURCE), NOT A LITERAL:
WE USE AN **ANONYMOUS CLASS** FOR THIS SCOPE.
FINALLY, FOAF:NAME HAS RANGE RDFS:LITERAL.

RDFS EXPRESSIVITY

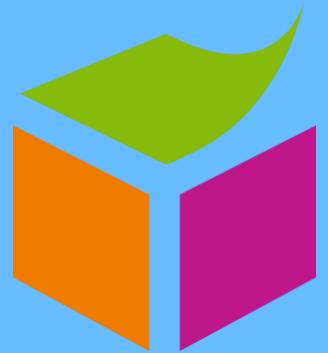
SIMPLE RELATIONSHIP BETWEEN THINGS

RDFS PROVIDES A VOCABULARY TO EXPRESS RELATIONSHIP BETWEEN THINGS (E.G., `SUBCLASSOF` OR `TYPE`)

CANNOT DESCRIBE COMPLEX RELATIONSHIP

RDFS CANNOT DESCRIBE DATA IN TERMS OF SET OF OPERATIONS (E.G., `UNIONOF`),
EQUIVALENCE (E.G., `SAMEAS`)
OR CARDINALITY (E.G. `ALLVALUEFROM`)

OWL



WEB ONTOLOGY LANGUAGE (OWL)

KNOWLEDGE REPRESENTATION LANGUAGE

DESIGNED TO FORMULATE, EXCHANGE AND REASON WITH KNOWLEDGE
ABOUT A **DOMAIN OF INTEREST**

CURRENTLY, AT VERSION 2

ONTOLOGY

IT IS THE **RESULT** OF THE MODELING PROCESS

AN ONTOLOGY IS AN **EXPLICIT DESCRIPTION** OF A DOMAIN
IN TERMS OF CONCEPTS, PROPERTIES AND ATTRIBUTES OF CONCEPTS,
CONSTRAINTS ON PROPERTIES AND ATTRIBUTES, AND INDIVIDUALS

AN ONTOLOGY DEFINES A COMMON VOCABULARY
AND A SHARED UNDERSTANDING

INDIVIDUALS, CLASSES, AND PROPERTIES

EXAMPLES

"POLITECNICO DI TORINO IS A UNIVERSITY"

"POLITECNICO DI TORINO HAS A PROFESSOR NAMED ELIO PICCOLO"

"POLITECNICO DI TORINO" IS AN OBJECT: AN **INDIVIDUAL** IN OWL2

"UNIVERSITY" IS A CATEGORY: A **CLASS** IN OWL2

"HAS A PROFESSOR" IS A RELATION: A **PROPERTY** IN OWL2

"ELIO PICCOLO" IS AN **INDIVIDUAL**, TOO

OBJECT VS. DATA PROPERTIES

OBJECT PROPERTIES

RELATE OBJECTS TO OBJECTS

"POLITECNICO DI TORINO HAS A PROFESSOR NAMED ELIO PICCOLO"

DATA PROPERTIES

ASSIGN DATA VALUES TO OBJECTS

"THE PROFESSOR IS NAMED ELIO PICCOLO"

WEB ONTOLOGY LANGUAGE

EXPRESSIVITY

DESIGNED TO REPRESENT RICH AND COMPLEX KNOWLEDGE
ABOUT THINGS, GROUP OF THINGS, AND THEIR RELATIONSHIPS

LOGIC-BASED

KNOWLEDGE EXPRESSED IN OWL CAN BE **REASONED** WITH A COMPUTER PROGRAM
TO VERIFY ITS CONSISTENCY OR TO MAKE IMPLICIT KNOWLEDGE EXPLICIT

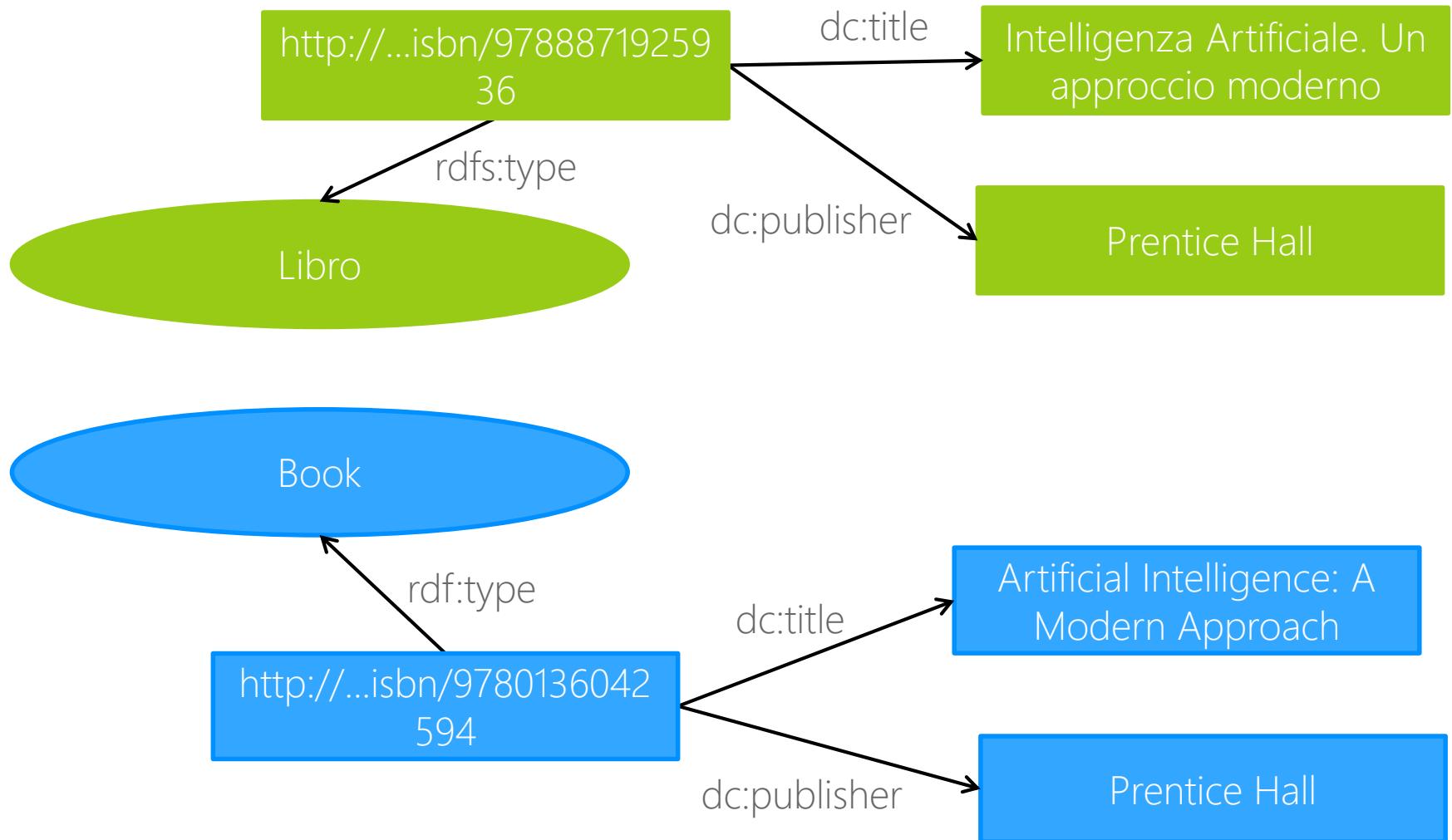
LINKED

ONTOLOGIES IN OWL CAN BE PUBLISHED ON THE WEB AND MY
REFER TO OR **BE REFERRED** FROM OTHER OWL ONTOLOGIES

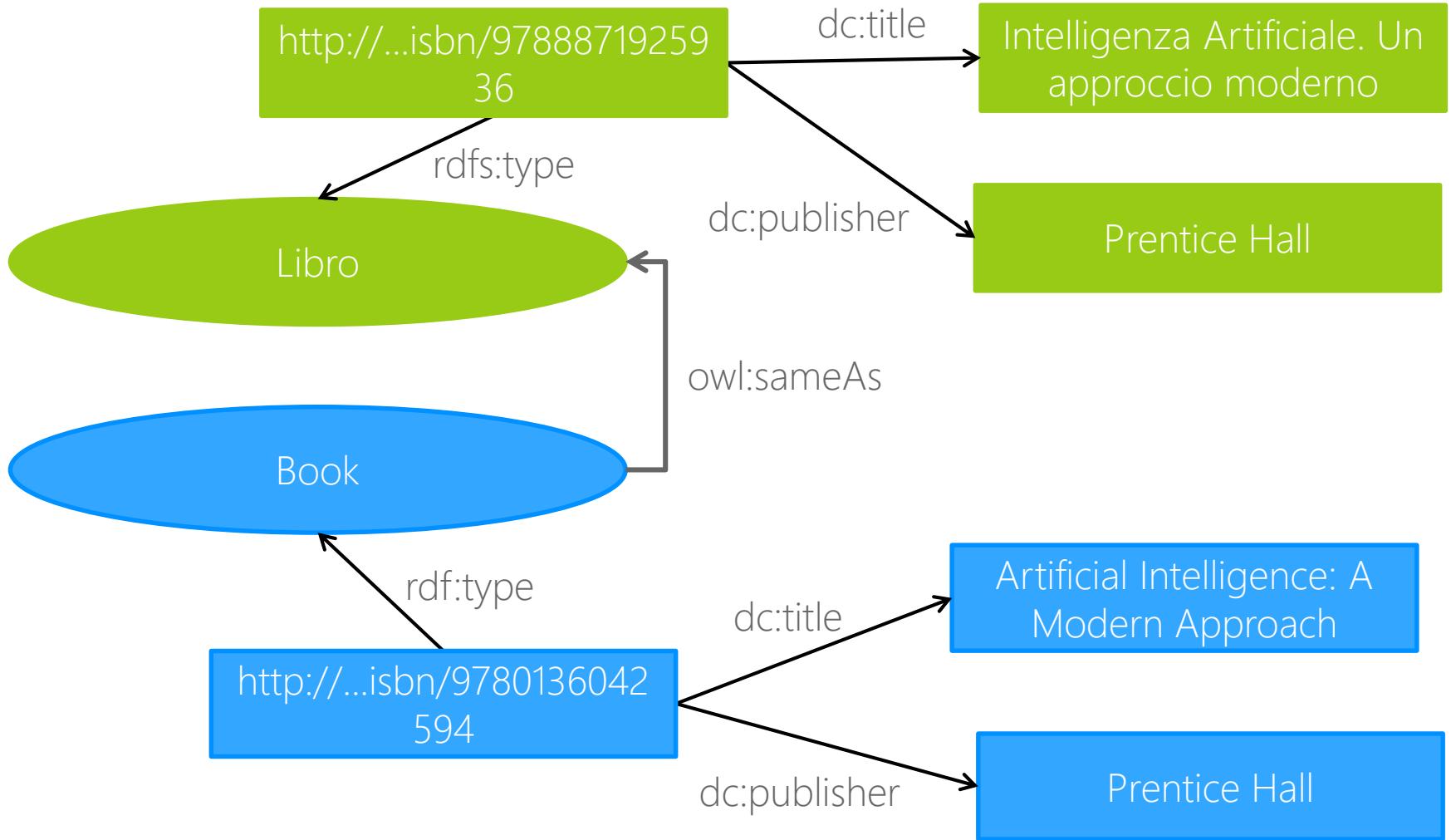
CHOOSE THE **SYNTAX** YOU LIKE

VARIOUS SYNTAXES ARE AVAILABLE FOR OWL, FOR DIFFERENT PURPOSES
(RDF/XML, TURTLE, MANCHESTER, ETC.)

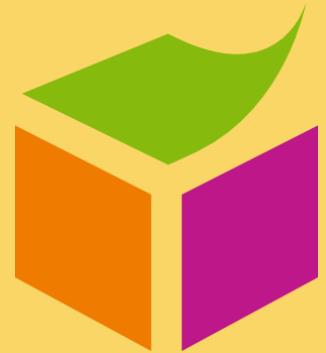
EXAMPLE: BOOKSTORE



EXAMPLE: BOOKSTORE



HANDS ON OWL2



TOOLS FOR OWL

EDITORS

MOST COMMON EDITOR: PROTÉGÉ 5

OTHER TOOLS: TOPBRAID COMPOSER, NEON TOOLKIT

SPECIAL PURPOSES APPS, ESP. FOR LIGHT-WEIGHT ONTOLOGIES
(E.G., FOAF EDITORS)

(<http://semanticweb.org/wiki/Editors>)



REASONERS

OWL DL: PELLET 2.0, HERMIT, FACT++, RACERPRO

OWL EL: CEL, SHER, SNOROCKET, ELLY

OWL RL: OWLIM, JENA, ORACLE OWL REASONER

OWL QL: OWLGRES, QUONTO, QUILL

(<http://semanticweb.org/wiki/Reasoners>)

ONTOLOGY ENGINEERING

A.K.A. THE PROCESS OF BUILDING AND MAINTAINING AN ONTOLOGY

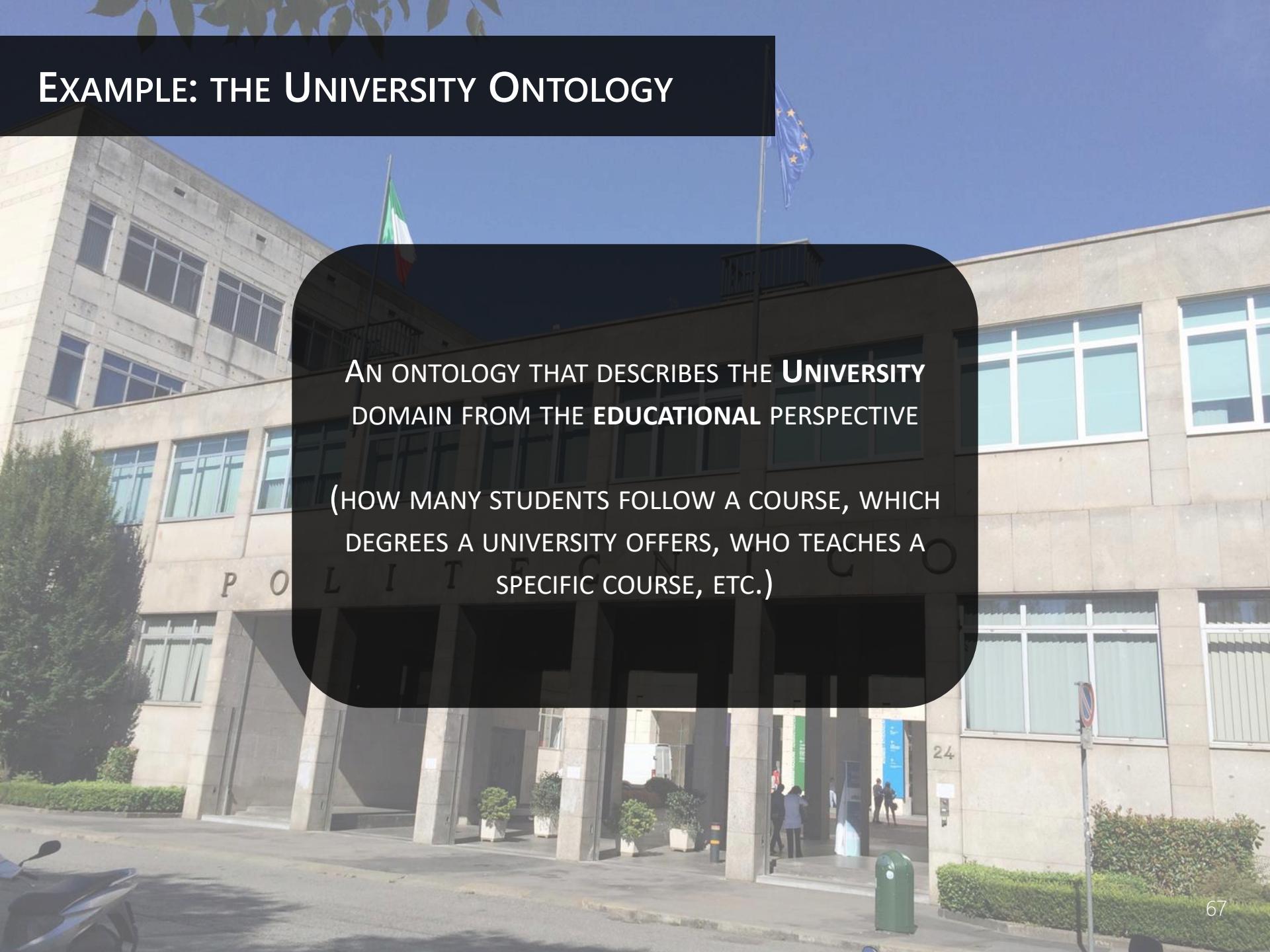
WHY DEVELOP AN ONTOLOGY?

1. TO SHARE **COMMON UNDERSTANDING** OF THE STRUCTURE OF INFORMATION, AMONG PEOPLE AND SOFTWARE ARTIFACTS
2. TO ENABLE **REUSE** OF DOMAIN KNOWLEDGE (STANDARDS)
3. TO MAKE DOMAIN ASSUMPTIONS **EXPLICIT**
4. TO **SEPARATE** DOMAIN KNOWLEDGE FROM THE OPERATIONAL KNOWLEDGE

How?



EXAMPLE: THE UNIVERSITY ONTOLOGY



AN ONTOLOGY THAT DESCRIBES THE **UNIVERSITY**
DOMAIN FROM THE **EDUCATIONAL PERSPECTIVE**

(HOW MANY STUDENTS FOLLOW A COURSE, WHICH
DEGREES A UNIVERSITY OFFERS, WHO TEACHES A
SPECIFIC COURSE, ETC.)

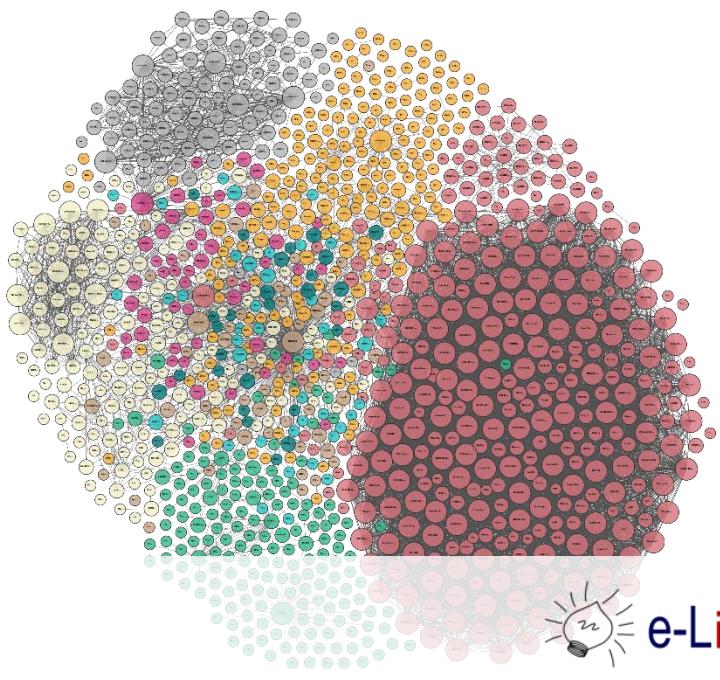
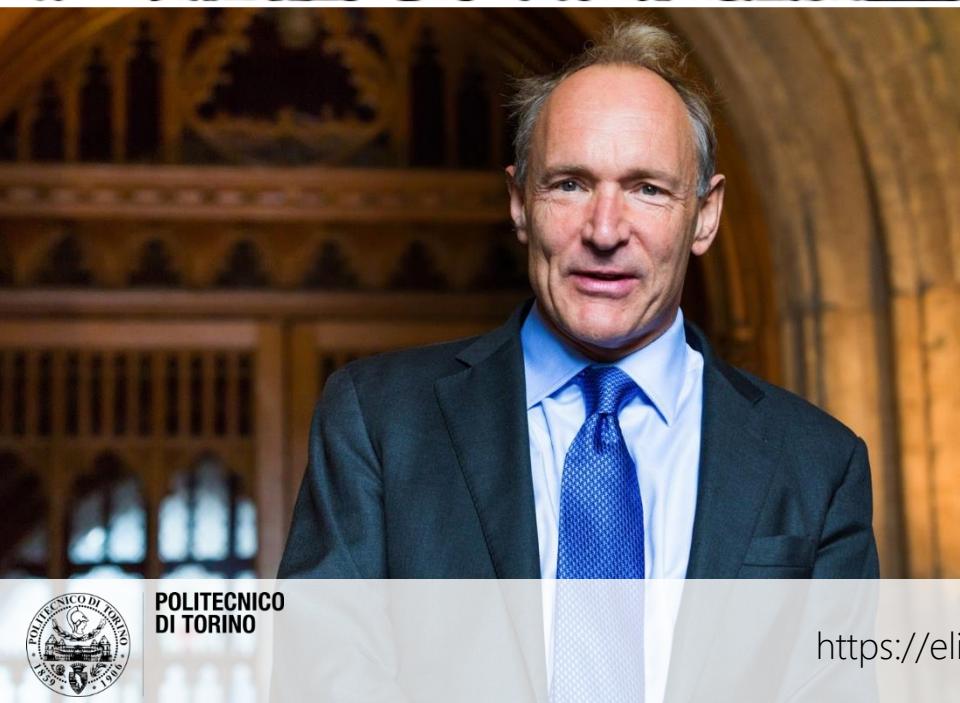
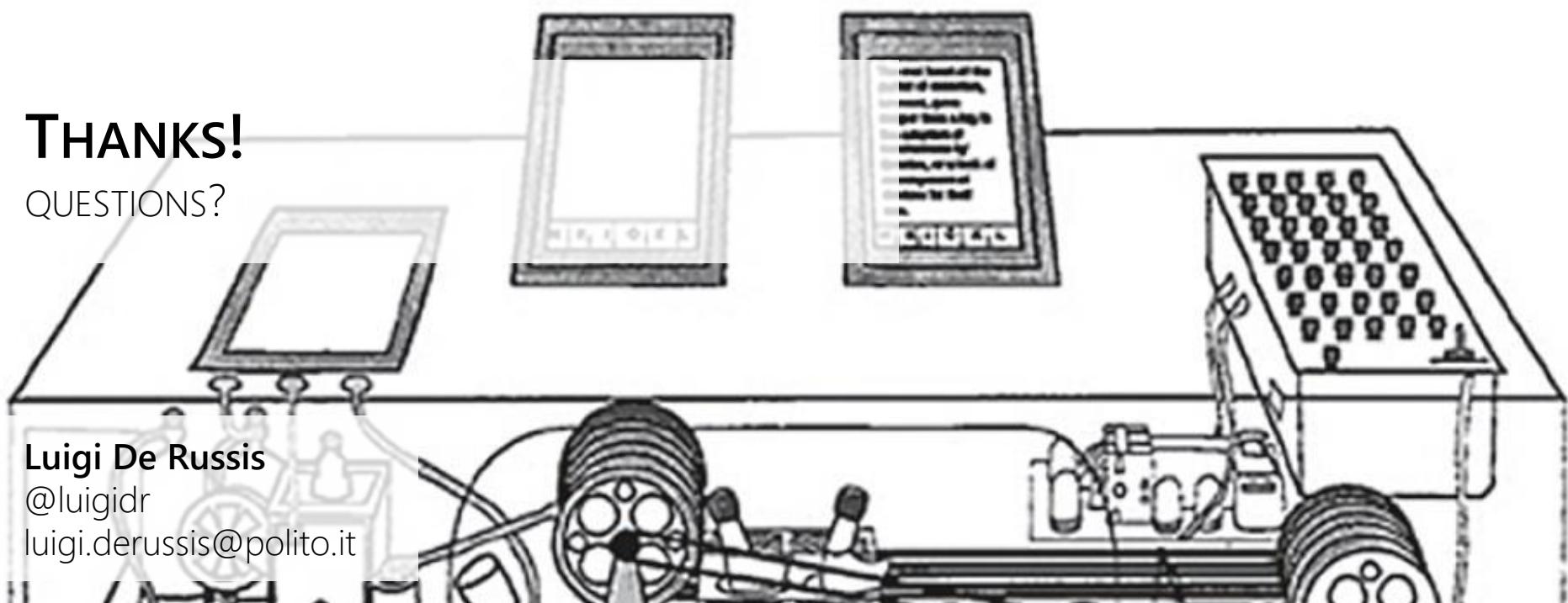
THANKS!

QUESTIONS?

Luigi De Russis

@luigidr

luigi.derussis@polito.it



REFERENCES

Semantic Web standards: <http://w3c.org/standards/semanticweb>

Semantic Web Wiki: <http://semanticweb.org>

Semantic Web FAQ: <http://www.w3c.org/2001/sw/SW-FAQ>

Book: A Semantic Web Primer (<http://www.semanticwebprimer.org>)

Book: Semantic Web Programming (<http://semwebprogramming.org>)

Semantic Web course @PoliTo: <https://elite.polito.it/teaching/current-courses/360-01rrdiu-semantic-web>

LICENSE

This work is licensed under the Creative Commons "Attribution-NonCommercial-ShareAlike Unported (CC BY-NC-SA 4.0)" License.

You are free:

- to **Share** - to copy, distribute and transmit the work
- to **Remix** - to adapt the work

Under the following conditions:

-  – **Attribution** - You must attribute the work in the manner specified by the author or licensor (but not in any way that suggests that they endorse you or your use of the work).
-  – **Noncommercial** - You may not use this work for commercial purposes.
-  – **Share Alike** - If you alter, transform, or build upon this work, you may distribute the resulting work only under the same or similar license to this one.

To view a copy of this license, visit

<https://creativecommons.org/licenses/by-nc-sa/4.0/>