

# Lecture 2

## Knowledge Graphs

COMP 474/6741, Winter 2021

### [Introduction](#)

- Motivation
- Why Knowledge Graphs?

### [History](#)

- Knowledge Representation
- Semantic Web
- Knowledge Graphs

### [RDF](#)

- Introduction
- RDF Triples
- Literals
- Blank Nodes
- DBpedia
- Namespaces
- Serialization
- Programming

### [Data Integration](#) [Example](#)

### [Conclusions](#)

- Architecture
- Examples

### [Notes and Further](#) [Reading](#)

René Witte  
Department of Computer Science  
and Software Engineering  
Concordia University

## 1 Introduction

### Introduction

Motivation

Why Knowledge Graphs?

## 2 History

### History

Knowledge Representation

Semantic Web

Knowledge Graphs

## 3 The Resource Description Framework (RDF)

### RDF

Introduction

RDF Triples

Literals

Blank Nodes

DBpedia

Namespaces

Serialization

Programming

## 4 Example: Data Integration with Knowledge Graphs

### Data Integration Example

## 5 Conclusions

### Conclusions

Architecture

Examples

## 6 Notes and Further Reading

### Notes and Further Reading

## Introduction

Motivation

Why Knowledge Graphs?

## History

Knowledge Representation

Semantic Web

Knowledge Graphs

## RDF

Introduction

RDF Triples

Literals

Blank Nodes

DBpedia

Namespaces

Serialization

Programming

## Data Integration

### Example

## Conclusions

Architecture

Examples

## Notes and Further

### Reading

## Slides Credit

- Includes slides from Jay Pujara & Sameer Singh, *Mining Knowledge Graphs from Text*, <https://kgtutorial.github.io/>
- Includes slides by Ivan Herman, W3C [Her]

## 1 Introduction

Motivation

Why Knowledge Graphs?

## 2 History

## 3 The Resource Description Framework (RDF)

## 4 Example: Data Integration with Knowledge Graphs

## 5 Conclusions

## 6 Notes and Further Reading

### Introduction

#### Motivation

Why Knowledge Graphs?

### History

Knowledge Representation

Semantic Web

Knowledge Graphs

### RDF

Introduction

RDF Triples

Literals

Blank Nodes

DBpedia

Namespaces

Serialization

Programming

### Data Integration

#### Example

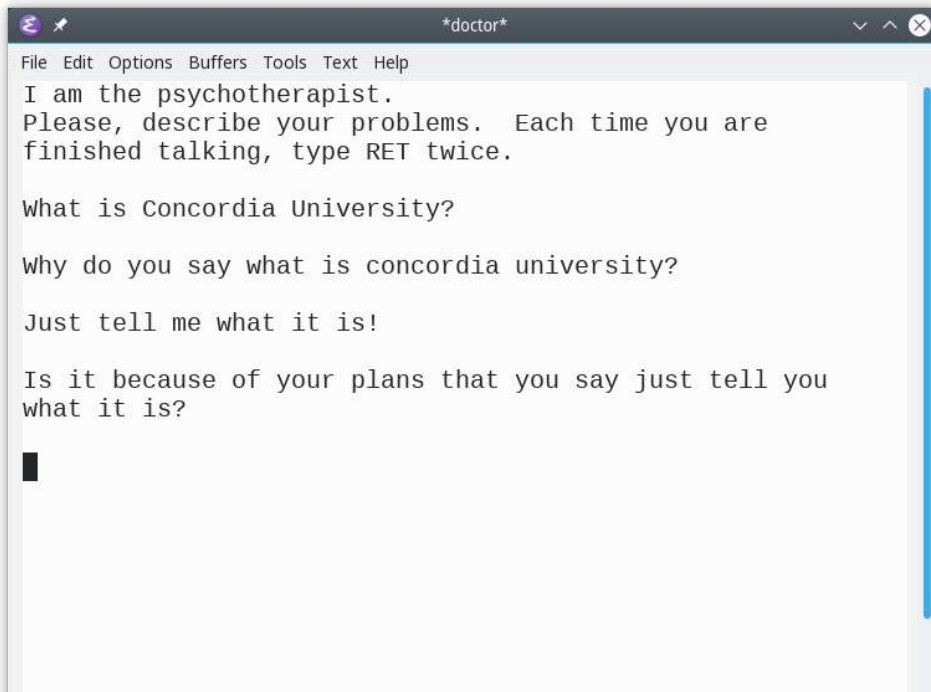
### Conclusions

Architecture

Examples

### Notes and Further

#### Reading

[Introduction](#)[Motivation](#)[Why Knowledge Graphs?](#)[History](#)[Knowledge Representation](#)[Semantic Web](#)[Knowledge Graphs](#)[RDF](#)[Introduction](#)[RDF Triples](#)[Literals](#)[Blank Nodes](#)[DBpedia](#)[Namespaces](#)[Serialization](#)[Programming](#)[Data Integration](#)[Example](#)[Conclusions](#)[Architecture](#)[Examples](#)[Notes and Further](#)[Reading](#)A screenshot of a terminal window titled '\*doctor\*' showing the Eliza chat program. The window has a menu bar with 'File', 'Edit', 'Options', 'Buffers', 'Tools', 'Text', and 'Help'. The text inside the window shows a conversation where the user asks about Concordia University and the program responds with generic psychotherapist-like statements. A cursor is visible at the bottom left.

```
*doctor*
File Edit Options Buffers Tools Text Help
I am the psychotherapist.
Please, describe your problems. Each time you are
finished talking, type RET twice.

What is Concordia University?

Why do you say what is concordia university?

Just tell me what it is!

Is it because of your plans that you say just tell you
what it is?

█
```



## Introduction

### Motivation

Why Knowledge Graphs?

### History

Knowledge Representation

Semantic Web

Knowledge Graphs

### RDF

Introduction

RDF Triples

Literals

Blank Nodes

DBpedia

Namespaces

Serialization

Programming

### Data Integration

#### Example

### Conclusions

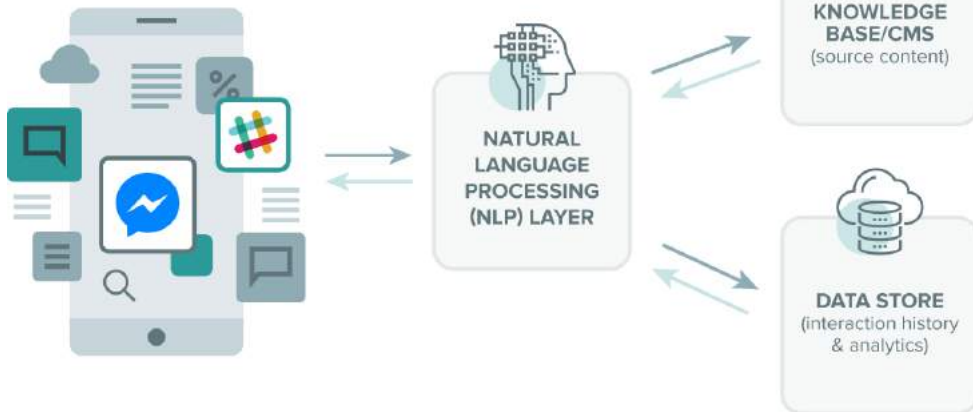
Architecture

Examples

### Notes and Further

#### Reading

# Generic Assistant Architecture



René Witte



## Introduction

### Motivation

Why Knowledge Graphs?

### History

Knowledge Representation

Semantic Web

Knowledge Graphs

### RDF

Introduction

RDF Triples

Literals

Blank Nodes

DBpedia

Namespaces

Serialization

Programming

### Data Integration

#### Example

### Conclusions

Architecture

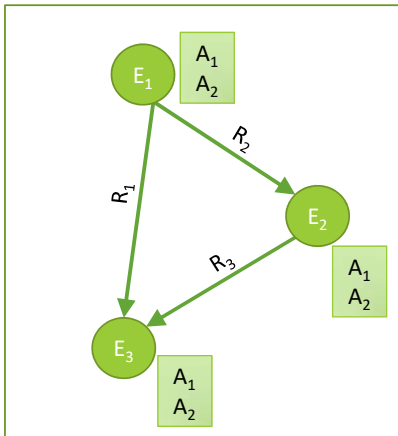
Examples

### Notes and Further

#### Reading

# What is a knowledge graph?

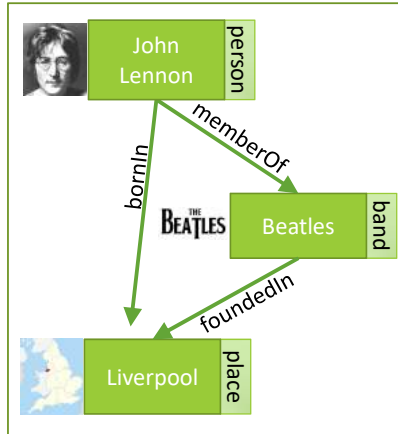
- Knowledge in graph form!
- Captures entities, attributes, and relationships
- Nodes are entities
- Nodes are labeled with attributes (e.g., types)
- Typed edges between two nodes capture a relationship between entities





# Example knowledge graph

- Knowledge in graph form!
- Captures entities, attributes, and relationships
- Nodes are entities
- Nodes are labeled with attributes (e.g., types)
- Typed edges between two nodes capture a relationship between entities



# Why knowledge graphs?

---

- Humans:
  - Combat information overload
  - Explore via intuitive structure
  - Tool for supporting knowledge-driven tasks
- AIs:
  - Key ingredient for many AI tasks
  - Bridge from data to human semantics
  - Use decades of work on graph analysis

[Introduction](#)[Motivation](#)[Why Knowledge Graphs?](#)[History](#)[Knowledge Representation](#)[Semantic Web](#)[Knowledge Graphs](#)[RDF](#)[Introduction](#)[RDF Triples](#)[Literals](#)[Blank Nodes](#)[DBpedia](#)[Namespaces](#)[Serialization](#)[Programming](#)[Data Integration](#)[Example](#)[Conclusions](#)[Architecture](#)[Examples](#)[Notes and Further](#)[Reading](#)

# Applications 1: QA/Agents

## Introduction

Motivation

Why Knowledge Graphs?

## History

Knowledge Representation

Semantic Web

Knowledge Graphs

## RDF

Introduction

RDF Triples

Literals

Blank Nodes

DBpedia

Namespaces

Serialization

Programming

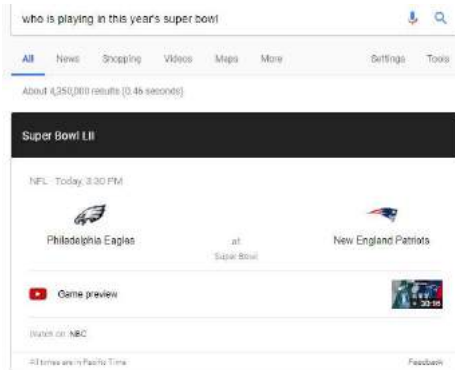
## Data Integration Example

## Conclusions

Architecture

Examples

## Notes and Further Reading



# Applications 2: Decision Support

## Introduction

Motivation

Why Knowledge Graphs?

## History

Knowledge Representation

Semantic Web

Knowledge Graphs

## RDF

Introduction

RDF Triples

Literals

Blank Nodes

DBpedia

Namespaces

Serialization

Programming

## Data Integration

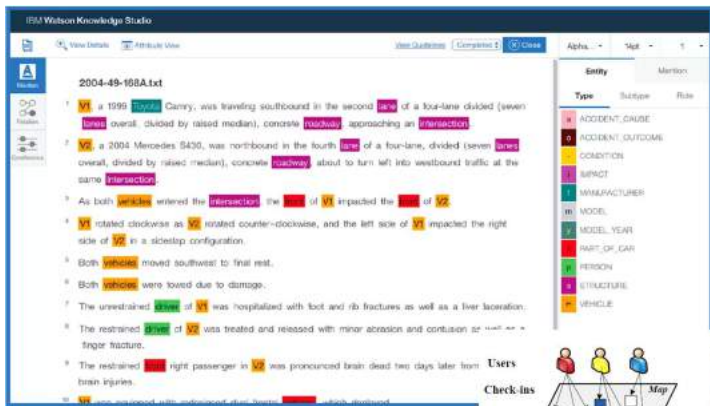
### Example

## Conclusions

Architecture

Examples

## Notes and Further Reading

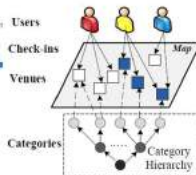


IBM Watson Knowledge Studio

2004-49-168A.txt

- 1 Y1 is a 1999 Toyota Camry, was traveling southbound in the second lane of a four-lane divided (seven lanes overall, divided by raised median), concrete roadway approaching an intersection.
- 2 Y2, a 2004 Mercedes S430, was northbound in the fourth lane of a four-lane, divided (seven lanes overall, divided by raised median), concrete roadway about to turn left into westbound traffic at the same intersection.
- 3 As both vehicles entered the intersection, the rear of Y1 impacted the rear of Y2.
- 4 Y1 rotated clockwise as Y2 rotated counter-clockwise, and the left side of Y1 impacted the right side of Y2 in a sidestep configuration.
- 5 Both vehicles moved southwest to final rest.
- 6 Both vehicles were towed due to damage.
- 7 The unrestrained driver of Y1 was hospitalized with foot and rib fractures as well as a liver laceration.
- 8 The restrained driver of Y2 was treated and released with minor abrasion and contusion to right hand and finger fracture.
- 9 The restrained right passenger in Y2 was pronounced brain dead two days later from brain injuries.
- 10 Y2 was equipped with advanced driver's assist system which deactivated.

Entity		Mention
Type	Subtype	Role
ACCIDENT_CAUSE		
ACCIDENT_OUTCOME		
CONDITION		
IMPACT		
MANUFACTURER		
MODEL		
MODEL_YEAR		
PART_OF_CAR		
PERSON		
STRUCTURE		
VEHICLE		



# Applications 3: Fueling Discovery

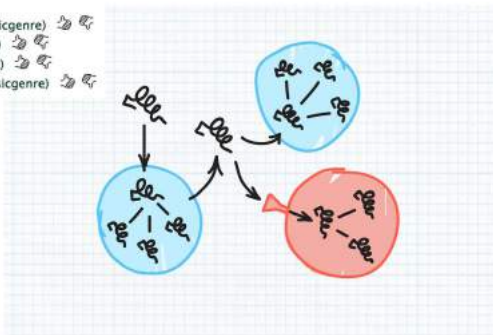
## beatles (musicartist)

literal strings: BEATLES, Beatles, beatles

### Help NELL Learn!

NELL wants to know if these be  
If they are or ever were, click thumbs-up. Or

- beatles is a musical artist 👍👎
- beatles is a musician in the genre classic pop (musicgenre) 👍👎
- beatles is a musician in the genre pop (musicgenre) 👍👎
- beatles is a musician in the genre rock (musicgenre) 👍👎
- beatles is a musician in the genre classic rock (musicgenre) 👍👎



## Introduction

Motivation

Why Knowledge Graphs?

## History

Knowledge Representation

Semantic Web

Knowledge Graphs

## RDF

Introduction

RDF Triples

Literals

Blank Nodes

DBpedia

Namespaces

Serialization

Programming

## Data Integration Example

## Conclusions

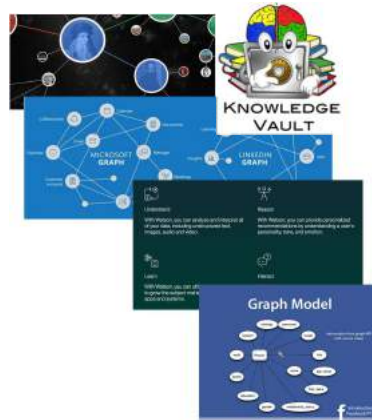
Architecture

Examples

## Notes and Further Reading

# Knowledge Graphs & Industry

- Google Knowledge Graph
  - Google Knowledge Vault
- Amazon Product Graph
- Facebook Graph API
- IBM Watson
- Microsoft Satori
  - Project Hanover/Literome
- LinkedIn Knowledge Graph
- Yandex Object Answer
- Diffbot, GraphIQ, Maana, ParseHub, Reactor Labs, SpazioDati



## Introduction

Motivation

Why Knowledge Graphs?

## History

Knowledge Representation

Semantic Web

Knowledge Graphs

## RDF

Introduction

RDF Triples

Literals

Blank Nodes

DBpedia

Namespaces

Serialization

Programming

## Data Integration Example

## Conclusions


Architecture

Examples

## Notes and Further Reading

# Where do knowledge graphs come from?

- Structured Text
  - Wikipedia Infoboxes, tables, databases, social nets

The Beatles			
 <p>The "Fab Four" Beatles lineup in 1964. Clockwise from top left: John Lennon, Paul McCartney, Ringo Starr and George Harrison</p>			
Background Information			
Origin	Liverpool, England, United Kingdom		
Genres	Rock · pop		
Years active	1960–1970		
Labels	EMI · PolyGram · Swan · Vee · United Artists		
Associated acts	The Quarry · Preston · P		
Website	thebeatles.com		
Past members	John Lennon Paul McCartney George Harrison Ringo Starr		

© National Geographic, Getty, Liverpool			
Mon 30th	07:18	07:09	10:36
Tue 31st	07:55	07:43	10:14
Wed 1st	08:33	08:21	10:53
Thu 2nd	09:11	08:59	11:31
Fri 3rd	09:49	09:37	12:09
Sat 4th	10:27	10:15	12:47
Sun 5th	11:05	10:53	13:25
Mon 6th	11:43	11:31	14:03
Tue 7th	12:21	12:09	14:41
Wed 8th	12:59	12:47	15:19
Thu 9th	13:37	13:25	15:57
Fri 10th	14:15	14:03	16:35
Sat 11th	14:53	14:41	17:13
Sun 12th	15:31	15:19	17:51
Mon 13th	16:09	15:57	18:29
Tue 14th	16:47	16:35	19:07
Wed 15th	17:25	17:13	19:45
Thu 16th	18:03	17:51	20:23
Fri 17th	18:41	18:29	21:01
Sat 18th	19:19	19:07	21:39
Sun 19th	19:57	19:45	22:17
Mon 20th	20:35	20:23	22:55
Tue 21st	21:13	21:01	23:33
Wed 22nd	21:51	21:39	24:11
Thu 23rd	22:29	22:17	24:49
Fri 24th	23:07	22:55	25:27
Sat 25th	23:45	23:33	26:05
Sun 26th	24:23	24:11	26:43
Mon 27th	25:01	24:49	27:21
Tue 28th	25:39	25:27	27:59
Wed 29th	26:17	26:05	28:37
Thu 30th	26:55	26:43	29:15
Fri 31st	27:33	27:21	29:53

## Introduction

Motivation

Why Knowledge Graphs?

## History

Knowledge Representation

Semantic Web

Knowledge Graphs

## RDF

Introduction

RDF Triples

Literals

Blank Nodes

DBpedia

Namespaces

Serialization

Programming

## Data Integration Example

## Conclusions

Architecture

Examples

## Notes and Further Reading

## Introduction

## Why Knowledge Graphs?

## History

## Semantic Web

## Knowledge Graphs

## RDF

## Introduction

## RDF Triples

### Literals

### Blank Nodes

DBpedia

## Namespaces

## Serialization

## Programming

## Data Integration Example

## Conclusions

## Architecture

### Examples

## Notes and Further Reading

- **Structured Text**
  - Wikipedia Infoboxes, tables, databases, social nets
- **Unstructured Text**
  - WWW, news, social media, reference articles

## Beatles last live performance

Published: Thursday, January 26th 2017, 5:24 am PST

Updated: Monday, January 30th 2017, 4:06 am PST

Written by Jim Eftink, Producer      CONNECT



(KPVS) - How about a little Beatles history.

It was on this date in 1969, the band performed their last live public performance.

### Allan Williams, First Manager of the Beatles, Dies at 86

(Source: Stock news by ALAN KIRBY, THE 11, 1994)





# Where do knowledge graphs come from?

- Structured Text
  - Wikipedia Infoboxes, tables, databases, social nets
- Unstructured Text
  - WWW, news, social media, reference articles
- Images



## Introduction

Motivation

Why Knowledge Graphs?

## History

Knowledge Representation

Semantic Web

Knowledge Graphs

## RDF

Introduction

RDF Triples

Literals

Blank Nodes

DBpedia

Namespaces

Serialization

Programming

## Data Integration Example

## Conclusions

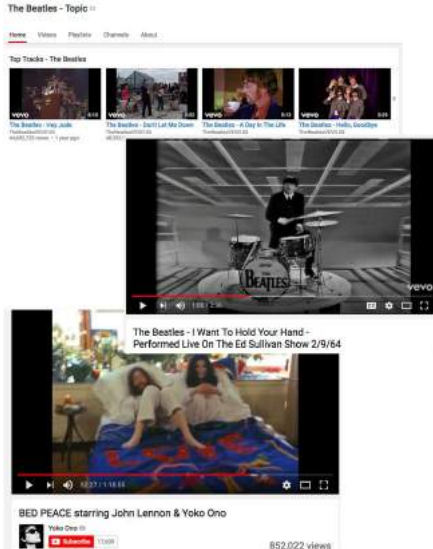
Architecture

Examples

## Notes and Further Reading

# Where do knowledge graphs come from?

- Structured Text
  - Wikipedia Infoboxes, tables, databases, social nets
- Unstructured Text
  - WWW, news, social media, reference articles
- Images
- Video
  - YouTube, video feeds



## Introduction

Motivation

Why Knowledge Graphs?

## History

Knowledge Representation

Semantic Web

Knowledge Graphs

## RDF

Introduction

RDF Triples

Literals

Blank Nodes

DBpedia

Namespaces

Serialization

Programming

## Data Integration

### Example

## Conclusions

Architecture

Examples

## Notes and Further

### Reading

## 1 Introduction

### Introduction

Motivation  
Why Knowledge Graphs?

## 2 History

Knowledge Representation  
Semantic Web  
Knowledge Graphs

### History

Knowledge Representation  
Semantic Web  
Knowledge Graphs

## 3 The Resource Description Framework (RDF)

### RDF

Introduction  
RDF Triples  
Literals  
Blank Nodes  
DBpedia  
Namespaces  
Serialization  
Programming

## 4 Example: Data Integration with Knowledge Graphs

### Data Integration Example

## 5 Conclusions

### Conclusions

Architecture  
Examples

## 6 Notes and Further Reading

### Notes and Further Reading

# History of Knowledge Representation (KR)

René Witte



## Introduction

Motivation  
Why Knowledge Graphs?

## History

### Knowledge Representation

Semantic Web  
Knowledge Graphs

## RDF

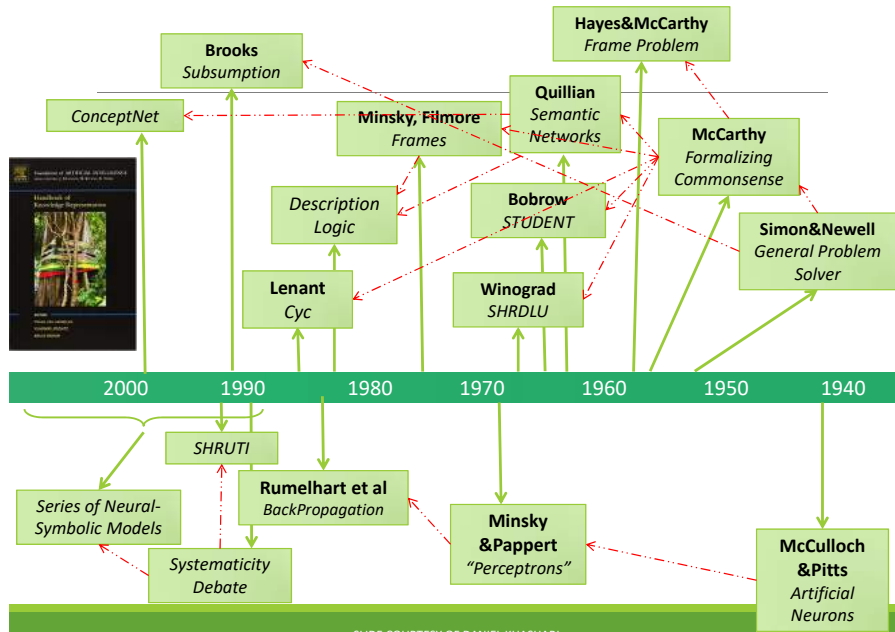
Introduction  
RDF Triples  
Literals  
Blank Nodes  
DBpedia  
Namespaces  
Serialization  
Programming

## Data Integration Example

## Conclusions

Architecture  
Examples

## Notes and Further Reading



# Knowledge Representation

---

- Decades of research into knowledge representation
- Most knowledge graph implementations use RDF triples
  - `<rdf:subject, rdf:predicate, rdf:object> : r(s,p,o)`
  - Temporal scoping, reification, and skolemization...
- ABox (assertions) versus TBox (terminology)
- Common ontological primitives
  - `rdfs:domain`, `rdfs:range`, `rdf:type`, `rdfs:subClassOf`, `rdfs:subPropertyOf`, ...
  - `owl:inverseOf`, `owl:TransitiveProperty`, `owl:FunctionalProperty`, ...

## Introduction

Motivation

Why Knowledge Graphs?

## History

Knowledge Representation

Semantic Web

Knowledge Graphs

## RDF

Introduction

RDF Triples

Literals

Blank Nodes

DBpedia

Namespaces

Serialization

Programming

## Data Integration

### Example

## Conclusions

Architecture

Examples

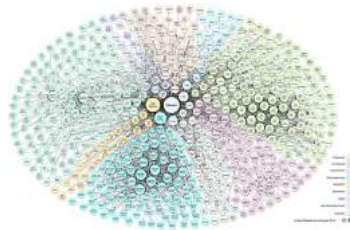
## Notes and Further

### Reading

# Semantic Web

---

- Standards for defining and exchanging knowledge
  - RDF, RDFa, JSON-LD, schema.org
  - RDFS, OWL, SKOS, FOAF
- Annotated data provide critical resource for automation
- Major weakness: annotate everything?



## Introduction

Motivation

Why Knowledge Graphs?

## History

Knowledge Representation

Semantic Web

Knowledge Graphs

## RDF

Introduction

RDF Triples

Literals

Blank Nodes

DBpedia

Namespaces

Serialization

Programming

## Data Integration

### Example

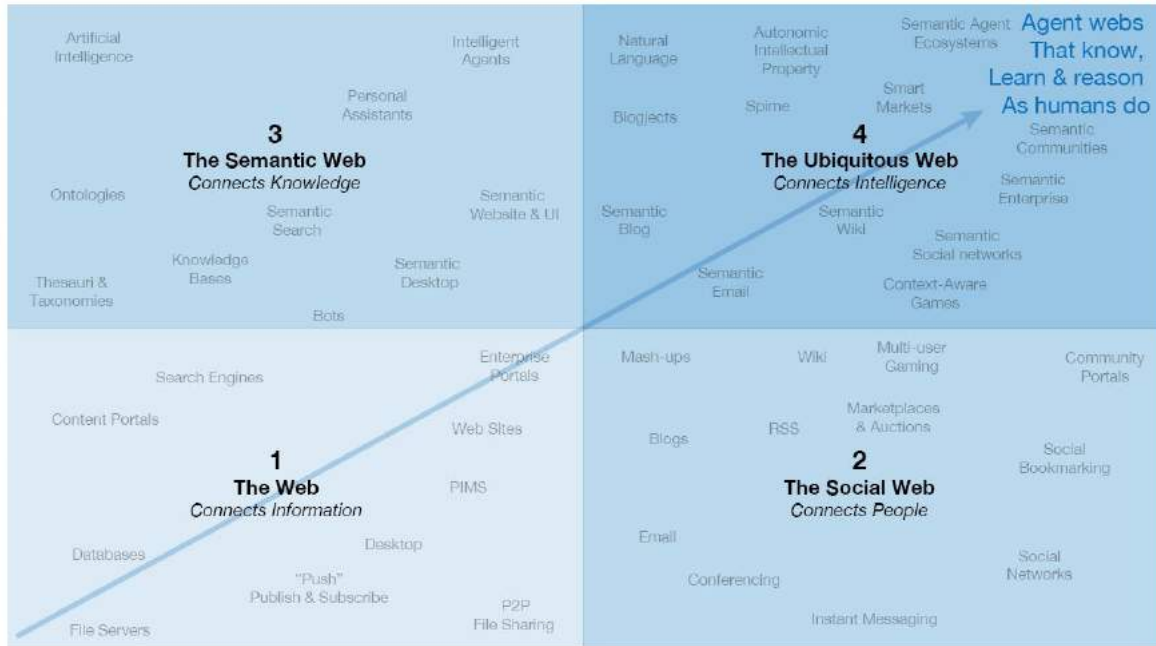
## Conclusions

Architecture

Examples

## Notes and Further

### Reading



Increasing Social Connectivity

# Apple's "Knowledge Navigator" Vision (1987)

René Witte



## Introduction

- Motivation
- Why Knowledge Graphs?

## History

- Knowledge Representation
- Semantic Web**
- Knowledge Graphs

## RDF

- Introduction
- RDF Triples
- Literals
- Blank Nodes
- DBpedia
- Namespaces
- Serialization
- Programming

## Data Integration Example

## Conclusions

- Architecture
- Examples

## Notes and Further Reading

<https://www.youtube.com/watch?v=umJsITGzXd0>



## From 1950–2020...

- Concepts have been around for a long time (Semantic Networks, Frames, Description Logic, ...)

## 1980s/90s

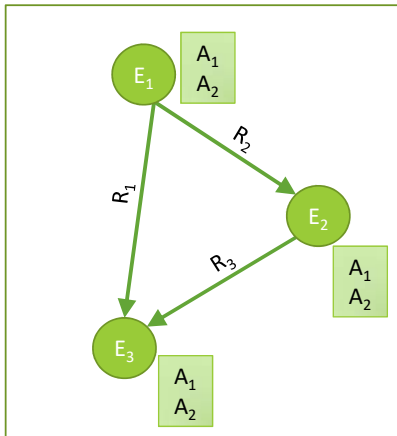
- AI/IS systems suffer from the *Knowledge Acquisition Bottleneck*
- One of the reasons for the *AI Winter* at that time

## Technology

- Open standards, based on W3C recommendations, e.g., [RDF](#)
- Proprietary products, e.g., [Neo4J](#) or [Oracle Spatial and Graph](#)
- We now have substantial [knowledge bases](#) available, both proprietary (e.g., Facebook Graph Search, Google Knowledge Graph) and open access (e.g., Wikidata, DBpedia, YAGO)

# What is a knowledge graph?

- Knowledge in graph form!
- Captures entities, attributes, and relationships
- Nodes are entities
- Nodes are labeled with attributes (e.g., types)
- Typed edges between two nodes capture a relationship between entities



## Introduction

Motivation

Why Knowledge Graphs?

## History

Knowledge Representation

Semantic Web

Knowledge Graphs

## RDF

Introduction

RDF Triples

Literals

Blank Nodes

DBpedia

Namespaces

Serialization

Programming

## Data Integration Example

## Conclusions

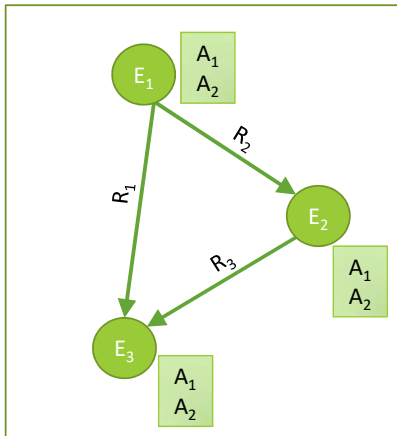
Architecture

Examples

## Notes and Further Reading

# Basic problems

- **Who** are the entities (nodes) in the graph?
- **What** are their attributes and types (labels)?
- **How** are they related (edges)?



## Introduction

Motivation

Why Knowledge Graphs?

## History

Knowledge Representation

Semantic Web

Knowledge Graphs

## RDF

Introduction

RDF Triples

Literals

Blank Nodes

DBpedia

Namespaces

Serialization

Programming

## Data Integration Example

## Conclusions

Architecture

Examples

## Notes and Further Reading

## 1 Introduction

## 2 History

## 3 The Resource Description Framework (RDF)

Introduction

RDF Triples

Literals

Blank Nodes

DBpedia

Namespaces

Serialization

Programming

## 4 Example: Data Integration with Knowledge Graphs

## 5 Conclusions

## 6 Notes and Further Reading

### Introduction

Motivation

Why Knowledge Graphs?

### History

Knowledge Representation

Semantic Web

Knowledge Graphs

### RDF

Introduction

RDF Triples

Literals

Blank Nodes

DBpedia

Namespaces

Serialization

Programming

### Data Integration

#### Example

### Conclusions

Architecture

Examples

### Notes and Further

#### Reading



## Introduction

Motivation

Why Knowledge Graphs?

## History

Knowledge Representation

Semantic Web

Knowledge Graphs

## RDF

Introduction

RDF Triples

Literals

Blank Nodes

DBpedia

Namespaces

Serialization

Programming

## Data Integration Example

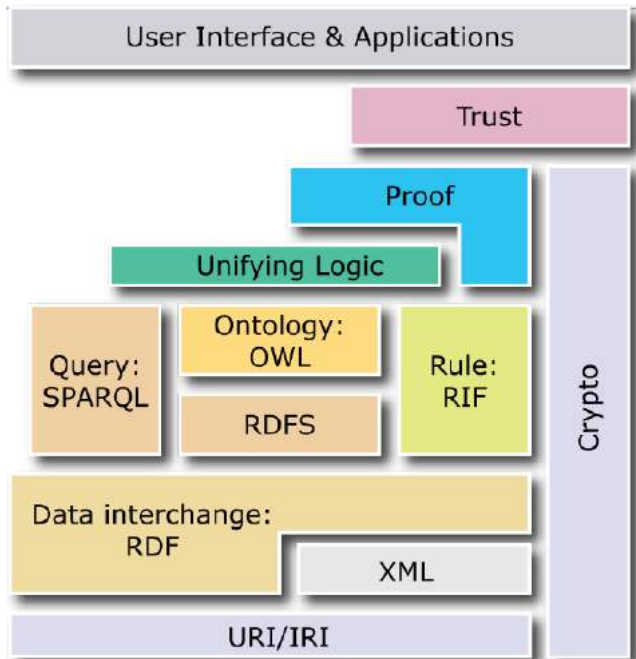
## Conclusions

Architecture

Examples

## Notes and Further Reading

# The W3C “Layer Cake”



René Witte



## Introduction

- Motivation
- Why Knowledge Graphs?

## History

- Knowledge Representation
- Semantic Web
- Knowledge Graphs

## RDF

### Introduction

- RDF Triples
- Literals
- Blank Nodes
- DBpedia
- Namespaces
- Serialization
- Programming

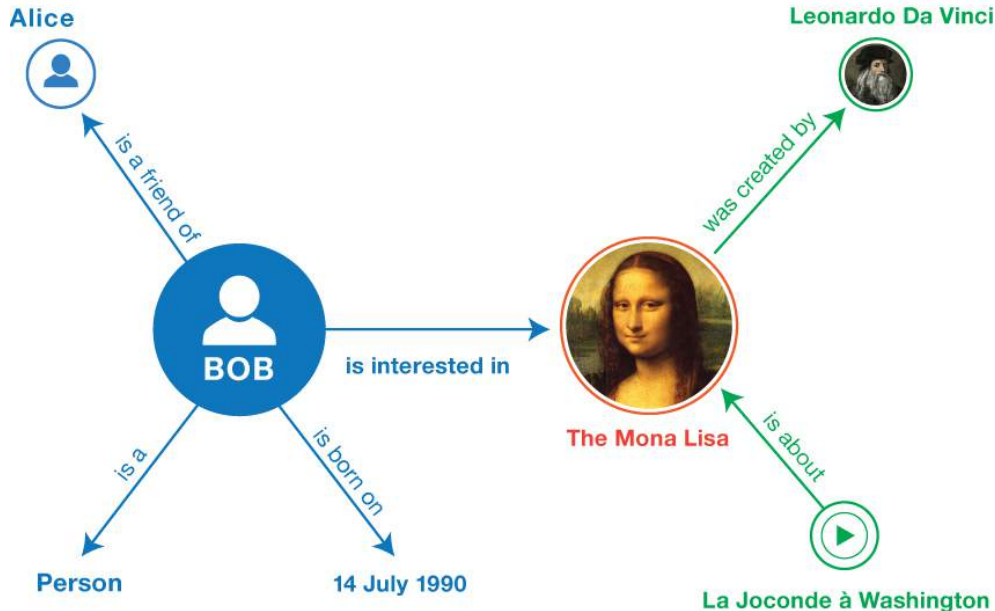
## Data Integration Example

## Conclusions

- Architecture
- Examples

## Notes and Further Reading

# Knowledge as Graphs



René Witte



## Introduction

- Motivation
- Why Knowledge Graphs?

## History

- Knowledge Representation
- Semantic Web
- Knowledge Graphs

## RDF

### Introduction

- RDF Triples
- Literals
- Blank Nodes
- DBpedia
- Namespaces
- Serialization
- Programming

## Data Integration Example

## Conclusions

- Architecture
- Examples

## Notes and Further Reading

→ **Worksheet #1: Tasks 1 & 2**

<https://www.w3.org/TR/rdf11-primer/>

## Representation of Knowledge Graphs

In a system, we represent graphs in form of **triples**:

`<subject> <predicate> <object>`

(The *predicate* is sometimes called *property*.)

## Examples

`<Bob> <is a> <person>.`

`<Bob> <is a friend of> <Alice>.`

`<Bob> <is born on> <the 14th of July 1990>.`

`<Bob> <is interested in> <the Mona Lisa>.`

`<the Mona Lisa> <was created by> <Leonardo da Vinci>.`

→ **Worksheet #1: Tasks 3 & 4**



## Introduction

Motivation

Why Knowledge Graphs?

## History

Knowledge Representation

Semantic Web

Knowledge Graphs

## RDF

### Introduction

RDF Triples

Literals

Blank Nodes

DBpedia

Namespaces

Serialization

Programming

### Data Integration Example

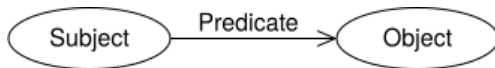
### Conclusions

Architecture

Examples

### Notes and Further Reading

<subject> <predicate> <object>



## Introduction

Motivation

Why Knowledge Graphs?

## History

Knowledge Representation

Semantic Web

Knowledge Graphs

## RDF

Introduction

**RDF Triples**

Literals

Blank Nodes

DBpedia

Namespaces

Serialization

Programming

## Data Integration Example

## Conclusions

Architecture

Examples

## Notes and Further Reading

## The Resource Description Framework (RDF)

W3C (World Wide Web Consortium) standard (“recommendation”)

- first public draft 1997
- RDF 1.0 in 1999; revised in 2004
- RDF 1.1 in 2014 (current version)

Family of standards: RDF, RDFS, RDFa, Turtle, N3, SPARQL, ...

## Format of triples

In RDF,

- Subject and predicate must be URIs (IRIs)
- Object can be IRI or **literal**

## Examples

```
<http://www.wikidata.org/entity/Q12418>  
  <http://purl.org/dc/terms/title>  
  "Mona Lisa" .
```

```
<http://www.wikidata.org/entity/Q12418>  
  <http://purl.org/dc/terms/creator>  
  <http://dbpedia.org/resource/Leonardo_da_Vinci> .
```

→ **Worksheet #1: Tasks 5 & 6**

### Introduction

Motivation  
Why Knowledge Graphs?

### History

Knowledge Representation  
Semantic Web  
Knowledge Graphs

### RDF

Introduction

**RDF Triples**

Literals

Blank Nodes

DBpedia

Namespaces

Serialization

Programming

### Data Integration

#### Example

### Conclusions

Architecture  
Examples

### Notes and Further Reading

## "Mona Lisa"

In this triple

```
<http://www.wikidata.org/entity/Q12418>  
<http://purl.org/dc/terms/title> "Mona Lisa" .
```

"Mona Lisa" is a **string literal**

## Things to know about literals

- Literals have a **datatype**, e.g., `string` or `int`
- Strings can have a **language tag**, e.g.,  
`"Leonardo da Vinci"@en`  
`"Léonard de Vinci"@fr`
- Strings are often used to provide human-readable **labels**  
*"Hey, how did you like the movie Q168154?"*
- For strings **only**, datatype can be omitted:  
`"Mona Lisa"` is equivalent to `"Mona Lisa"^^xsd:string`
- Again, literals can **only** appear in the **object** position of a triple `<s> <p> <o>`

### Introduction

Motivation

Why Knowledge Graphs?

### History

Knowledge Representation

Semantic Web

Knowledge Graphs

### RDF

Introduction

RDF Triples

Literals

Blank Nodes

DBpedia

Namespaces

Serialization

Programming

### Data Integration

#### Example

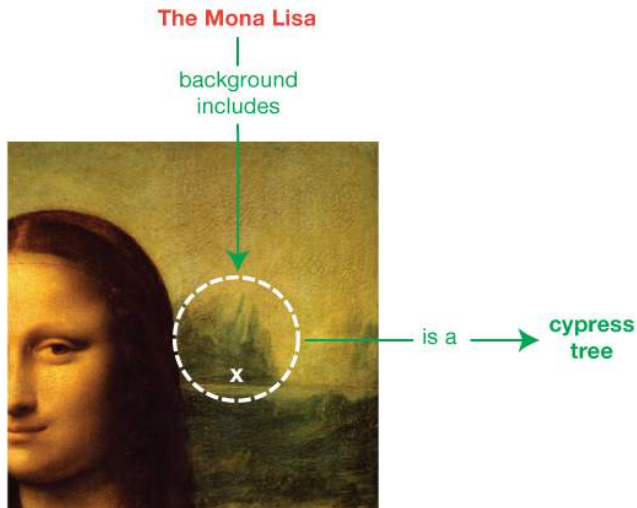
### Conclusions

Architecture

Examples

### Notes and Further

#### Reading



```
<http://dbpedia.org/resource/Mona_Lisa> <lio:shows> _:x .  
_:x a <http://dbpedia.org/resource/Cypress> .
```

## Introduction

Motivation  
Why Knowledge Graphs?

## History

Knowledge Representation  
Semantic Web  
Knowledge Graphs

## RDF

Introduction  
RDF Triples  
Literals

## Blank Nodes

DBpedia  
Namespaces  
Serialization  
Programming

## Data Integration Example

## Conclusions

Architecture  
Examples

## Notes and Further Reading

## About: [Leonardo da Vinci](#)

An Entity of Type : person, from Named Graph : <http://dbpedia.org>, within Data Space : [dbpedia.org](#)

Leonardo di ser Piero da Vinci (Italian: [leɔˈnardo di sɛr ˈpjɛːro da (v)ˈvintʃi] (); 14/15 April 1452 – 2 May 1519), known as Leonardo da Vinci (English: LEE-ə-NAR-doh də VIN-chee, LEE-oh-, LAY-oh-), was an Italian polymath of the Renaissance who is widely considered one of the greatest painters of all time (despite less than 25 of his paintings having survived). He is also known for his , in which he made drawings and notes on science and invention; these involve a variety of subjects including anatomy, cartography, and paleontology.

Property	Value
<a href="#">dbpedia:abstract</a>	<ul style="list-style-type: none"> <li>Leonardo di ser Piero da Vinci (Italian: [leɔˈnardo di sɛr ˈpjɛːro da (v)ˈvintʃi] (); 14/15 April 1452 – 2 May 1519), known as Leonardo da Vinci (English: LEE-ə-NAR-doh də VIN-chee, LEE-oh-, LAY-oh-), was an Italian polymath of the Renaissance who is widely considered one of the greatest painters of all time (despite less than 25 of his paintings having survived). He is also known for his , in which he made drawings and notes on science and invention; these involve a variety of subjects including anatomy, cartography, and paleontology. Born out of wedlock to a notary, Piero da Vinci, and a peasant woman, Caterina, in Vinci, in the region of Florence, Italy, Leonardo was educated in the studio of the renowned Italian painter Andrea del Verrocchio.</li> </ul>

## DBpedia URIs

Make sure you use the correct URI:

- <http://dbpedia.org/resource/>... is the canonical URI
- The DBpedia server returns either
  - <http://dbpedia.org/page/>... (HTML data, for a human)
  - <http://dbpedia.org/data/>... (RDF data, for an AI)

→ **Worksheet #1: Task 7**

### Introduction

Motivation

Why Knowledge Graphs?

### History

Knowledge Representation

Semantic Web

Knowledge Graphs

### RDF

Introduction

RDF Triples

Literals

Blank Nodes

### DBpedia

Namespaces

Serialization

Programming

### Data Integration

#### Example

### Conclusions

Architecture

Examples

### Notes and Further

#### Reading

## Shortening URIs

Instead of always writing full URIs (IRIs), we can split them into a **prefix** and **suffix**, e.g.: `<http://dbpedia.org/resource/Leonardo_da_Vinci>`

- We define a prefix **dbpedia**:

```
PREFIX dbpedia: <http://dbpedia.org/resource/>
```

- and now we can simple write:

```
dbpedia:Leonardo_da_Vinci
```

- Note: angle brackets `<>` only for full IRIs

→ reduces dataset sizes, easier to read

## Conventions

Commonly used URLs use the same namespace prefix

- E.g., FOAF (friend-of-a-friend):

```
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
```

- Lookup a prefix at <https://prefix.cc/>

→ **Worksheet #1: Tasks 8 & 9**

### Introduction

Motivation

Why Knowledge Graphs?

### History

Knowledge Representation

Semantic Web

Knowledge Graphs

### RDF

Introduction

RDF Triples

Literals

Blank Nodes

DBpedia

### Namespaces

Serialization

Programming

### Data Integration

#### Example

### Conclusions

Architecture

Examples

### Notes and Further

#### Reading

## Formats

There is no single format `.rdf` (like `.xml`), commonly used are:

**RDF/XML** for data exchange (somewhat deprecated)

**RDFa** for embedding RDF into web pages

**N-Triples (N3)** for streaming RDF data and bulk dataset up-/download

**Turtle** for human-readable files

**JSON-LD** for web applications

plus some variations/extensions.

## N-Triples

So far, we've mostly used the N-Triples format:

```
<http://www.wikidata.org/entity/Q12418> ↔  
  <http://purl.org/dc/terms/title> "Mona Lisa" .
```

each line in a file is one triple, full IRIs only (no namespace prefixes) and ended by a period `'.'`

### Introduction

- Motivation
- Why Knowledge Graphs?

### History

- Knowledge Representation
- Semantic Web
- Knowledge Graphs

### RDF

- Introduction
- RDF Triples
- Literals
- Blank Nodes
- DBpedia
- Namespaces

### Serialization

- Programming

### Data Integration

#### Example

### Conclusions

- Architecture
- Examples

### Notes and Further Reading



```

BASE <http://example.org/>
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema>
PREFIX schema: <http://schema.org/>
PREFIX dcterms: <http://purl.org/dc/terms/>
PREFIX wd: <http://www.wikidata.org/entity/>

```

```

<bob#me>
  a foaf:Person ;
  foaf:knows <alice#me> ;
  schema:birthDate "1990-07-04"^^xsd:date ;
  foaf:topic_interest wd:Q12418 .

```

```

wd:Q12418
  dcterms:title "Mona_Lisa" ;
  dcterms:creator <http://dbpedia.org/resource/Leonardo_da_Vinci> .

```

```

<http://data.europeana.eu/item/04802/243FA8618938F4117025F17A8B813C5F9AA4D619>
  dcterms:subject wd:Q12418 .

```

## Introduction

Motivation

Why Knowledge Graphs?

## History

Knowledge Representation

Semantic Web

Knowledge Graphs

## RDF

Introduction

RDF Triples

Literals

Blank Nodes

DBpedia

Namespaces

Serialization

Programming

Data Integration  
Example

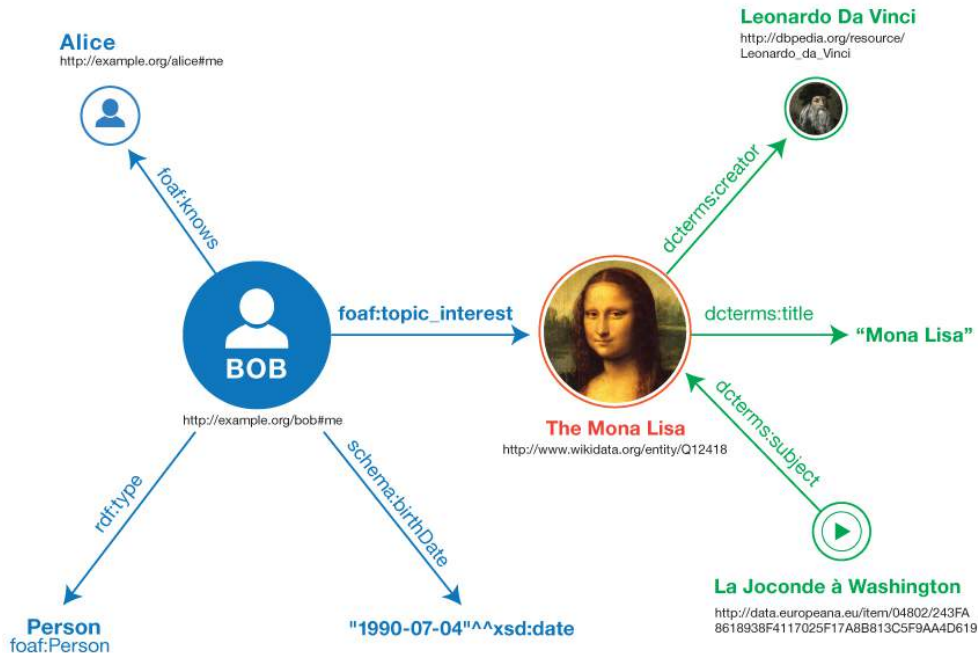
## Conclusions

Architecture

Examples

Notes and Further  
Reading

# Graph corresponding to the Turtle example



René Witte



## Introduction

- Motivation
- Why Knowledge Graphs?

## History

- Knowledge Representation
- Semantic Web
- Knowledge Graphs

## RDF

- Introduction
- RDF Triples
- Literals
- Blank Nodes
- DBpedia
- Namespaces
- Serialization
- Programming

## Data Integration Example

## Conclusions

- Architecture
- Examples

## Notes and Further Reading

# RDF in programming practice

---

- ▶ For example, using Python+RDFLib:
  - a “Graph” object is created
  - the RDF file is parsed and results stored in the Graph
  - the Graph offers methods to retrieve (or add):
    - triples
    - (property,object) pairs for a specific subject
    - (subject,property) pairs for specific object
    - etc.
  - the rest is conventional programming...
- ▶ Similar tools exist in Java, PHP, etc.

## Introduction

Motivation

Why Knowledge Graphs?

## History

Knowledge Representation

Semantic Web

Knowledge Graphs

## RDF

Introduction

RDF Triples

Literals

Blank Nodes

DBpedia

Namespaces

Serialization

Programming

## Data Integration

### Example

## Conclusions

Architecture

Examples

## Notes and Further

### Reading

# Python example using RDFLib

```
# create a graph from a file
graph = rdflib.Graph()
graph.parse("filename.rdf", format="rdxml")
# take subject with a known URI
subject = rdflib.URIRef("URI_of_Subject")
# process all properties and objects for this subject
for (s,p,o) in graph.triples((subject,None,None)) :
    do_something(p,o)
```

## Introduction

- Motivation
- Why Knowledge Graphs?

## History

- Knowledge Representation
- Semantic Web
- Knowledge Graphs

## RDF

- Introduction
- RDF Triples
- Literals
- Blank Nodes
- DBpedia
- Namespaces
- Serialization

## Programming

## Data Integration Example

## Conclusions

- Architecture
- Examples

## Notes and Further Reading

## 1 Introduction

### Introduction

- Motivation
- Why Knowledge Graphs?

## 2 History

### History

- Knowledge Representation
- Semantic Web
- Knowledge Graphs

## 3 The Resource Description Framework (RDF)

### RDF

- Introduction
- RDF Triples
- Literals
- Blank Nodes
- DBpedia
- Namespaces
- Serialization
- Programming

## 4 Example: Data Integration with Knowledge Graphs

### Data Integration Example

## 5 Conclusions

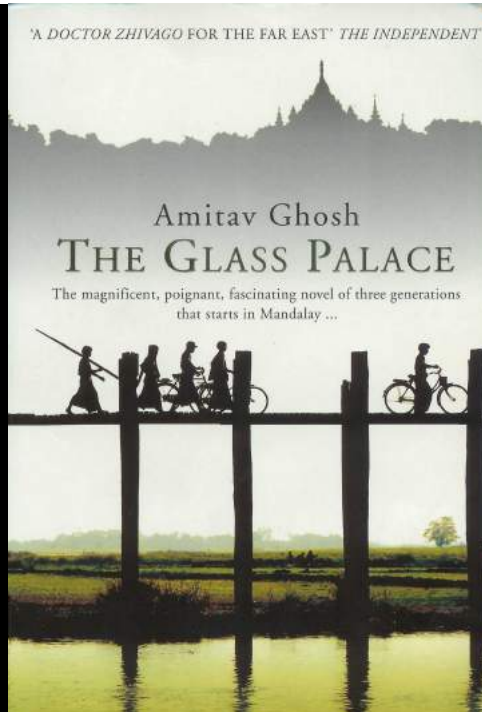
### Conclusions

- Architecture
- Examples

## 6 Notes and Further Reading

### Notes and Further Reading

## Let's start with a Book...



René Witte



### Introduction

- Motivation
- Why Knowledge Graphs?

### History

- Knowledge Representation
- Semantic Web
- Knowledge Graphs

### RDF

- Introduction
- RDF Triples
- Literals
- Blank Nodes
- DBpedia
- Namespaces
- Serialization
- Programming

### Data Integration Example

### Conclusions

- Architecture
- Examples

### Notes and Further Reading

# A simplified bookstore data (dataset “A”)

ISBN	Author	Title	Publisher	Year
0006511409X	id_xyz	The Glass Palace	id_qpr	2000

ID	Name	Homepage
id_xyz	Ghosh, Amitav	<a href="http://www.amitavghosh.com">http://www.amitavghosh.com</a>

ID	Publisher's name	City
id_qpr	Harper Collins	London

## Introduction

- Motivation
- Why Knowledge Graphs?

## History

- Knowledge Representation
- Semantic Web
- Knowledge Graphs

## RDF

- Introduction
- RDF Triples
- Literals
- Blank Nodes
- DBpedia
- Namespaces
- Serialization
- Programming

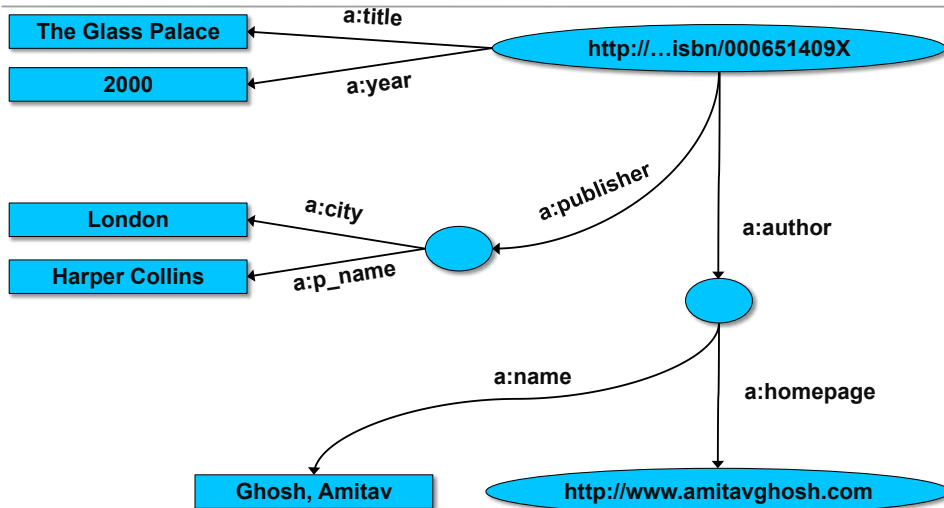
## Data Integration Example

## Conclusions

- Architecture
- Examples

## Notes and Further Reading

# 1<sup>st</sup>: export your data as a set of relations



## Introduction

Motivation  
Why Knowledge Graphs?

## History

Knowledge Representation  
Semantic Web  
Knowledge Graphs

## RDF

Introduction  
RDF Triples  
Literals  
Blank Nodes  
DBpedia  
Namespaces  
Serialization  
Programming

## Data Integration Example

## Conclusions

Architecture  
Examples

## Notes and Further Reading



# Some notes on the exporting the data

---

- ▶ Relations form a graph
  - the nodes refer to the “real” data or contain some literal
  - how the graph is represented in machine is immaterial for now

## Introduction

Motivation  
Why Knowledge Graphs?

## History

Knowledge Representation  
Semantic Web  
Knowledge Graphs

## RDF

Introduction  
RDF Triples  
Literals  
Blank Nodes  
DBpedia  
Namespaces  
Serialization  
Programming

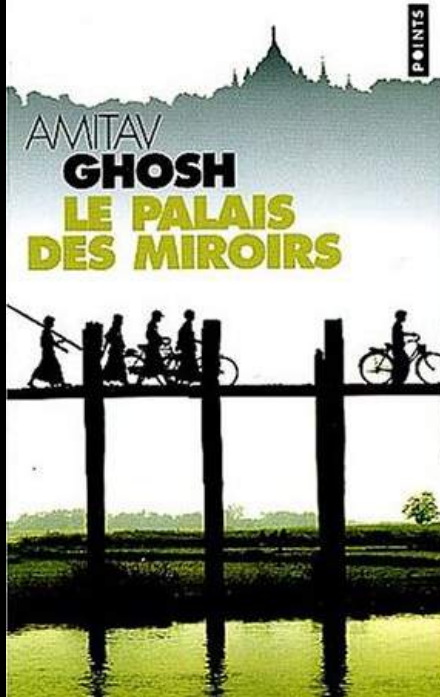
## Data Integration Example

## Conclusions

Architecture  
Examples

## Notes and Further Reading

Now the same book in French...



René Witte



## Introduction

Motivation

Why Knowledge Graphs?

## History

Knowledge Representation

Semantic Web

Knowledge Graphs

## RDF

Introduction

RDF Triples

Literals

Blank Nodes

DBpedia

Namespaces

Serialization

Programming

## Data Integration Example

## Conclusions

Architecture

Examples

## Notes and Further Reading

# Another bookstore data (dataset “F”)

René Witte



## Introduction

Motivation

Why Knowledge Graphs?

## History

Knowledge Representation

Semantic Web

Knowledge Graphs

## RDF

Introduction

RDF Triples

Literals

Blank Nodes

DBpedia

Namespaces

Serialization

Programming

## Data Integration Example

## Conclusions

Architecture

Examples

## Notes and Further Reading

A	B	C	D
1	<b>ID</b>	<b>Titre</b>	<b>Traducteur</b>
2	ISBN 2020286682	Le Palais des Miroirs	ISBN 0-00-6511409-X
3			
4			
5			
6	<b>ID</b>	<b>Auteur</b>	
7	ISBN 0-00-6511409-X	\$A11\$	
8			
9			
10	<b>Nom</b>		
11	Ghosh, Amitav		
12	Besse, Christianne		

## 2<sup>nd</sup>: export your second set of data

### Introduction

Motivation  
Why Knowledge Graphs?

### History

Knowledge Representation  
Semantic Web  
Knowledge Graphs

### RDF

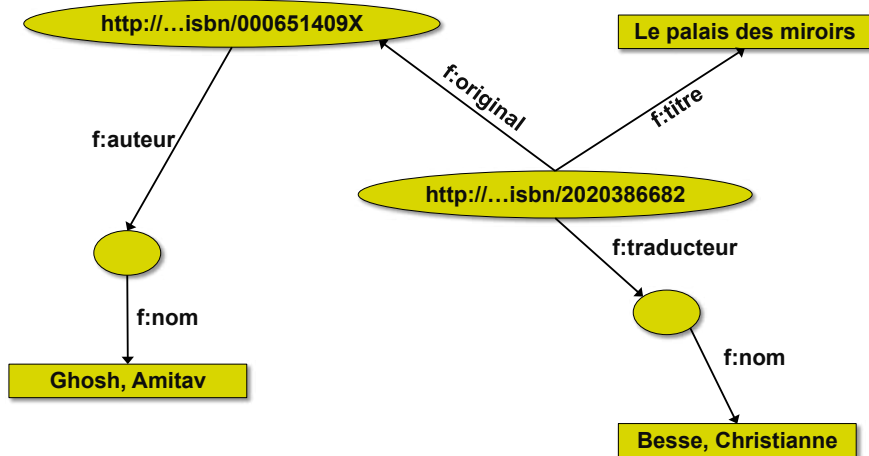
Introduction  
RDF Triples  
Literals  
Blank Nodes  
DBpedia  
Namespaces  
Serialization  
Programming

### Data Integration Example

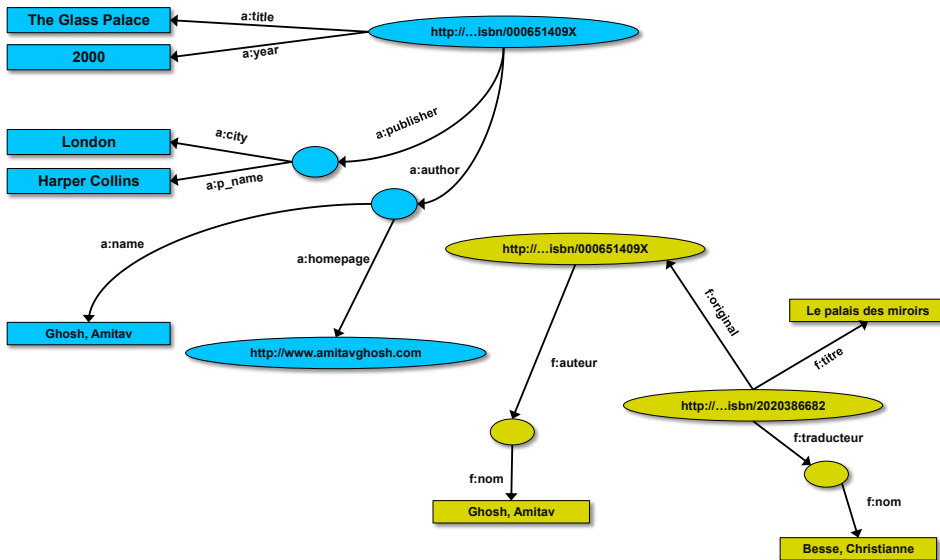
### Conclusions

Architecture  
Examples

### Notes and Further Reading



# 3<sup>rd</sup>: start merging your data



## Introduction

Motivation  
Why Knowledge Graphs?

## History

Knowledge Representation  
Semantic Web  
Knowledge Graphs

## RDF

Introduction  
RDF Triples  
Literals  
Blank Nodes  
DBpedia  
Namespaces  
Serialization  
Programming

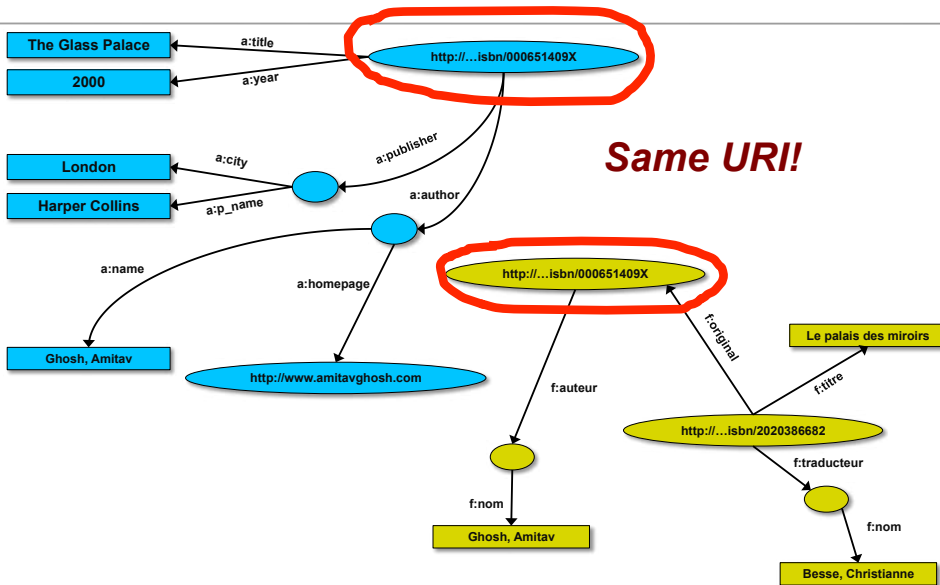
## Data Integration Example

## Conclusions

Architecture  
Examples

## Notes and Further Reading

# 3<sup>rd</sup>: start merging your data (cont)



## Introduction

Motivation  
Why Knowledge Graphs?

## History

Knowledge Representation  
Semantic Web  
Knowledge Graphs

## RDF

Introduction  
RDF Triples  
Literals  
Blank Nodes  
DBpedia  
Namespaces  
Serialization  
Programming

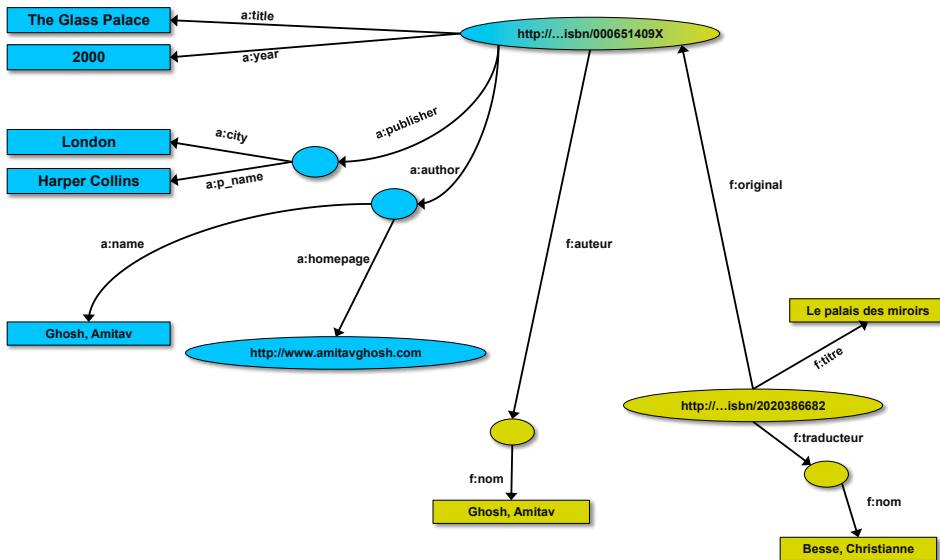
## Data Integration Example

## Conclusions

Architecture  
Examples

## Notes and Further Reading

# 3<sup>rd</sup>: start merging your data



## Introduction

Motivation  
Why Knowledge Graphs?

## History

Knowledge Representation  
Semantic Web  
Knowledge Graphs

## RDF

Introduction  
RDF Triples  
Literals  
Blank Nodes  
DBpedia  
Namespaces  
Serialization  
Programming

## Data Integration Example

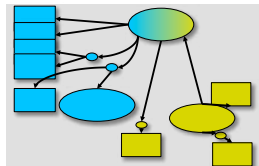
## Conclusions

Architecture  
Examples

## Notes and Further Reading

# Start making queries...

- ▶ User of data “F” can now ask queries like:
  - “give me the title of the original”
    - well, ... « donne-moi le titre de l’original »
- ▶ This information is not in the dataset “F”...
- ▶ ...but can be retrieved by merging with dataset “A”!



## Introduction

Motivation

Why Knowledge Graphs?

## History

Knowledge Representation

Semantic Web

Knowledge Graphs

## RDF

Introduction

RDF Triples

Literals

Blank Nodes

DBpedia

Namespaces

Serialization

Programming

## Data Integration Example

## Conclusions

Architecture

Examples

## Notes and Further Reading



# However, more can be achieved...

---

- ▶ We “feel” that a:author and f:auteur should be the same
- ▶ But an automatic merge does not know that!
- ▶ Let us add some extra information to the merged data:
  - a:author same as f:auteur
  - both identify a “Person”
  - a term that a community may have already defined:
    - a “Person” is uniquely identified by his/her name and, say, homepage
    - it can be used as a “category” for certain type of resources

## Introduction

Motivation

Why Knowledge Graphs?

## History

Knowledge Representation

Semantic Web

Knowledge Graphs

## RDF

Introduction

RDF Triples

Literals

Blank Nodes

DBpedia

Namespaces

Serialization

Programming

## Data Integration

### Example

## Conclusions

Architecture

Examples

## Notes and Further

### Reading

### 3<sup>rd</sup> revisited: use the extra knowledge

## Introduction

## Motivation

### Why Knowledge Graphs?

## History

- Knowledge Representation
- Semantic Web
- Knowledge Graphs

## RDF

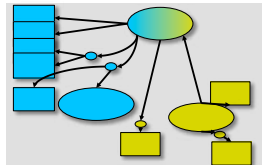
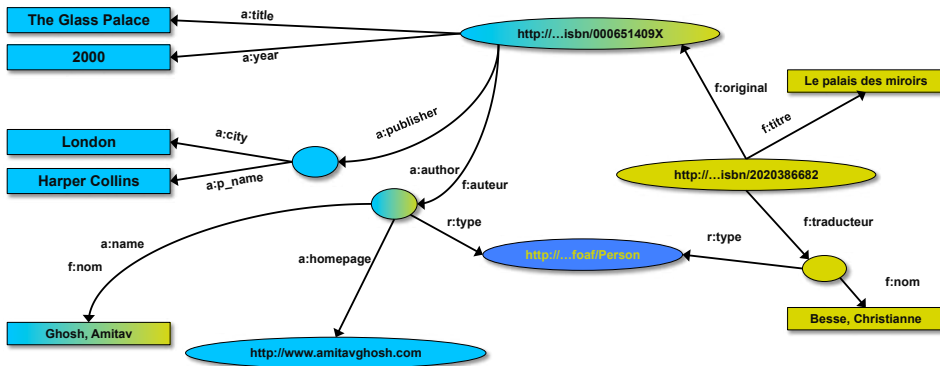
- Introduction
- RDF Triples
- Literals
- Blank Nodes
- DBpedia
- Namespaces
- Serialization
- Programming

## Data Integration Example

## Conclusions

Architecture  
Examples

## Notes and Further Reading



# Start making richer queries!

## Introduction

Motivation

Why Knowledge Graphs?

## History

Knowledge Representation

Semantic Web

Knowledge Graphs

## RDF

Introduction

RDF Triples

Literals

Blank Nodes

DBpedia

Namespaces

Serialization

Programming

## Data Integration Example

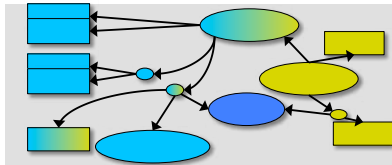
## Conclusions

Architecture

Examples

## Notes and Further Reading

- ▶ User of dataset “F” can now query:
  - “donnes-moi la page d’accueil de l’auteur de l’original”
    - well... “give me the home page of the original’s ‘auteur’”
- ▶ The information is not in datasets “F” or “A”...
- ▶ ...but was made available by:
  - merging datasets “A” and datasets “F”
  - adding three simple extra statements as an extra “glue”



# Combine with different datasets

---

- ▶ Using, e.g., the “Person”, the dataset can be combined with other sources
- ▶ For example, data in Wikipedia can be extracted using dedicated tools
  - e.g., the “[dbpedia](#)” project can extract the “infobox” information from Wikipedia already...

## Introduction

Motivation  
Why Knowledge Graphs?

## History

Knowledge Representation  
Semantic Web  
Knowledge Graphs

## RDF

Introduction  
RDF Triples  
Literals  
Blank Nodes  
DBpedia  
Namespaces  
Serialization  
Programming

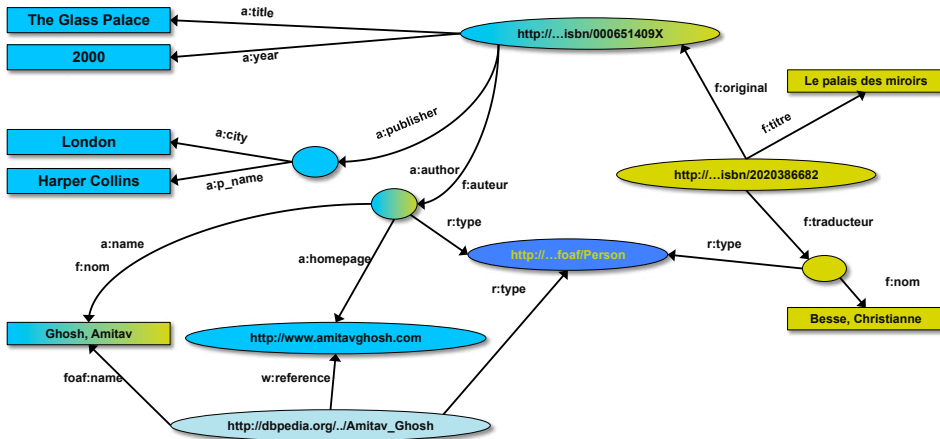
## Data Integration Example

## Conclusions

Architecture  
Examples

## Notes and Further Reading

# Merge with Wikipedia data



## Introduction

Motivation  
Why Knowledge Graphs?

## History

Knowledge Representation  
Semantic Web  
Knowledge Graphs

## RDF

Introduction  
RDF Triples  
Literals  
Blank Nodes  
DBpedia  
Namespaces  
Serialization  
Programming

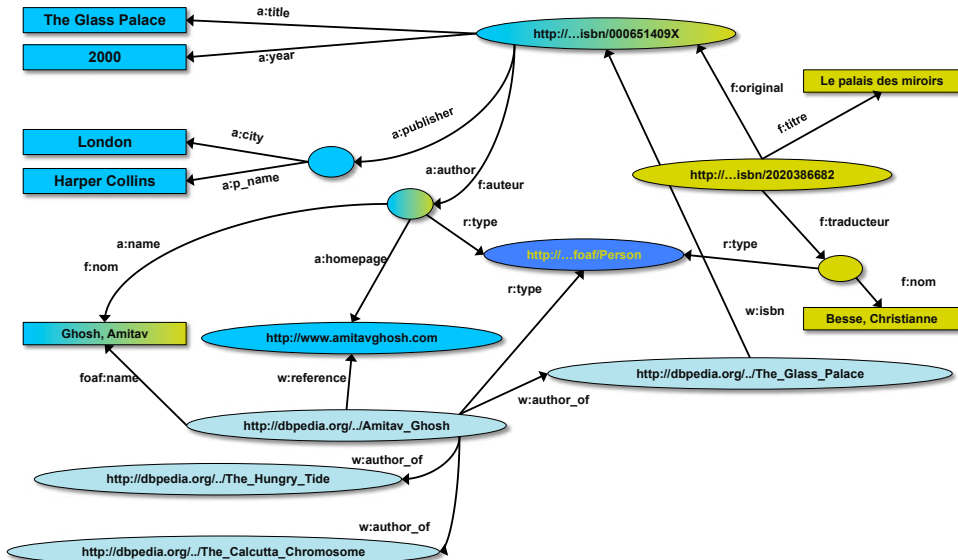
## Data Integration Example

## Conclusions

Architecture  
Examples

## Notes and Further Reading

# Merge with Wikipedia data



## Introduction

Motivation  
Why Knowledge Graphs?

## History

Knowledge Representation  
Semantic Web  
Knowledge Graphs

## RDF

Introduction  
RDF Triples  
Literals  
Blank Nodes  
DBpedia  
Namespaces  
Serialization  
Programming

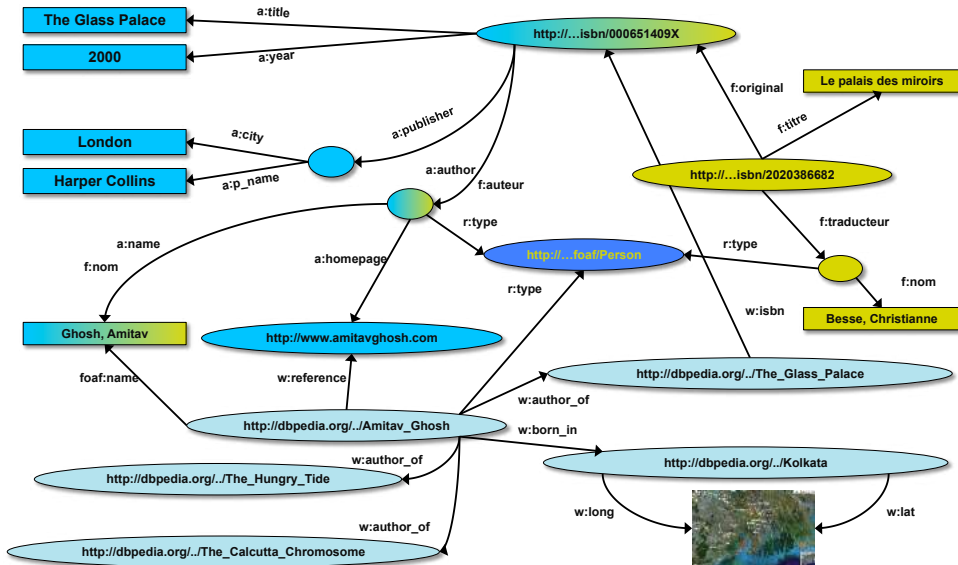
## Data Integration Example

## Conclusions

Architecture  
Examples

## Notes and Further Reading

# Merge with Wikipedia data



## Introduction

Motivation  
Why Knowledge Graphs?

## History

Knowledge Representation  
Semantic Web  
Knowledge Graphs

## RDF

Introduction  
RDF Triples  
Literals  
Blank Nodes  
DBpedia  
Namespaces  
Serialization  
Programming

## Data Integration Example

## Conclusions

Architecture  
Examples

## Notes and Further Reading

# Is that surprising?

---

- ▶ It may look like it but, in fact, it should not be...
- ▶ What happened via automatic means is done every day by Web users!
- ▶ The difference: a bit of extra rigour so that machines could do this, too

## Introduction

Motivation  
Why Knowledge Graphs?

## History

Knowledge Representation  
Semantic Web  
Knowledge Graphs

## RDF

Introduction  
RDF Triples  
Literals  
Blank Nodes  
DBpedia  
Namespaces  
Serialization  
Programming

## Data Integration Example

## Conclusions

Architecture  
Examples

## Notes and Further Reading

→ **Worksheet #1: Task 10**



## 1 Introduction

### Introduction

Motivation  
Why Knowledge Graphs?

## 2 History

### History

Knowledge Representation  
Semantic Web  
Knowledge Graphs

## 3 The Resource Description Framework (RDF)

### RDF

Introduction  
RDF Triples  
Literals  
Blank Nodes  
DBpedia  
Namespaces  
Serialization  
Programming

## 4 Example: Data Integration with Knowledge Graphs

### Data Integration Example

## 5 Conclusions

- Architecture
- Examples

### Conclusions

Architecture  
Examples

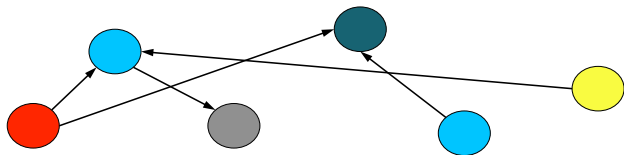
## 6 Notes and Further Reading

### Notes and Further Reading

# What did we do?



## Applications



## Data represented in abstract format

**Manipulate  
Query**  
...

**Map,  
Expose,**  
...



## Data in various formats

### Introduction

- Motivation
- Why Knowledge Graphs?

### History

- Knowledge Representation
- Semantic Web
- Knowledge Graphs

### RDF

- Introduction
- RDF Triples
- Literals
- Blank Nodes
- DBpedia
- Namespaces
- Serialization
- Programming

### Data Integration Example

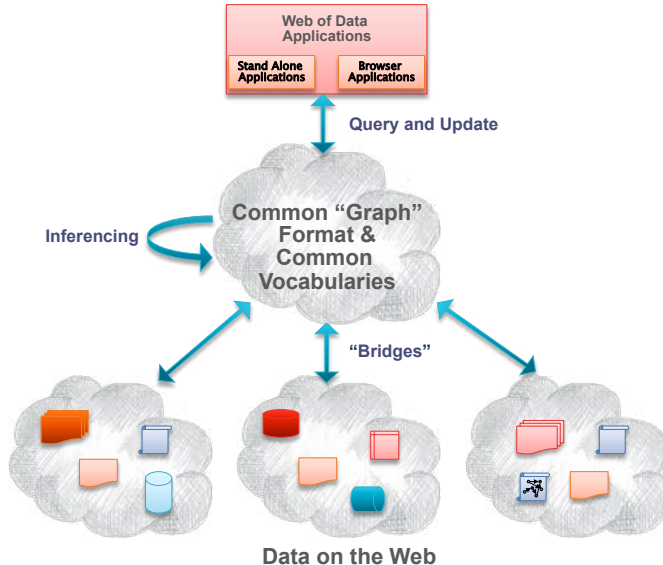
### Conclusions

- Architecture
- Examples

### Notes and Further Reading

# What did we do? (alternate view)

René Witte



## Introduction

Motivation

Why Knowledge Graphs?

## History

Knowledge Representation

Semantic Web

Knowledge Graphs

## RDF

Introduction

RDF Triples

Literals

Blank Nodes

DBpedia

Namespaces

Serialization

Programming

## Data Integration

### Example

## Conclusions

### Architecture

Examples

## Notes and Further

### Reading

# Success story: OpenIE (ReVerb)



Open Information Extraction

[openie.allenai.org](http://openie.allenai.org)



Argument 1: entity:The Beatles

Relation:

Argument 2:

All

Search

all location (21) film location (18) statistical region (16) name source (15) travel destination (14) misc.

more types

were bigger than **Jesus** (100)

came to America (95)

appeared on **The Ed Sullivan Show** (88)

broke up in 1970 (56)

Here Comes the Sun (46)

came to America (45)

is for the future (44)

are a great band (42)

perform on **The Ed Sullivan Show** (39)

were **Musical ensemble** (36)

are a great band »

Extracted Synonyms:

were  
is  
was

Extracted from these sentences:

are **The Beatles** are the best band , hands down but Oasis did make a great cover . (via ClueWeb12)  
**The Beatles** are a great band . (via ClueWeb12)  
**The Beatles** are the best band . (via ClueWeb12)  
**The Beatles** are the greatest band ... Started 1 month ago by georgedcc Yeah , Songs in the Key of Life is a bit much for 1 listen . (via ClueWeb12)  
**The Beatles** , arguably , are the greatest band , and may or may not have the greatest name . (via ClueWeb12)  
The point is , from my view , **The Beatles** are a good band , but way behind the greatest artists to ever grace rock . (via ClueWeb12)

## Introduction

Motivation

Why Knowledge Graphs?

## History

Knowledge Representation

Semantic Web

Knowledge Graphs

## RDF

Introduction

RDF Triples

Literals

Blank Nodes

DBpedia

Namespaces

Serialization

Programming

## Data Integration

### Example

## Conclusions

Architecture

## Examples

## Notes and Further Reading

# Success story: NELL

## Introduction

Motivation

Why Knowledge Graphs?

## History

Knowledge Representation

Semantic Web

Knowledge Graphs

## RDF

Introduction

RDF Triples

Literals

Blank Nodes

DBpedia

Namespaces

Serialization

Programming

## Data Integration Example

## Conclusions

Architecture

Examples

## Notes and Further Reading

### NELL Knowledge Base Browser

CMU Read the Web Project

log in | preferences | help/instructions | feedback

#### categories

#### relations











- everypromotedthing
- abstractthing
  - event
    - convention
    - musicfestival
    - protestevent
    - meetingeventtitle
    - conference
    - mlconference
  - weatherphenomenon
  - sportsevent
    - sportsgame
    - race
    - olympics
    - grandprix
  - crimeorcharge
  - earthquakeevent
  - election
  - bombingevent
  - militaryeventtype
    - militaryconflict
  - productionevent
  - roadaccidentevent
  - meetingeventtype
  - eventoutcome
  - malgorithm
  - physiologicalcondition
    - disease

### beatles (musicartist)

literal strings: [BEATLES](#), [Beatles](#), [beatles](#)

#### Help NELL Learn!

NELL wants to know if these beliefs are correct.  
If they are or ever were, click thumbs-up. Otherwise, click thumbs-down.

- [beatles](#) is a [musical artist](#)  
- [beatles](#) is a musician in the [genre classic pop](#) (musicgenre)  
- [beatles](#) is a musician in the [genre pop](#) (musicgenre)  
- [beatles](#) is a musician in the [genre rock](#) (musicgenre)  
- [beatles](#) is a musician in the [genre classic rock](#) (musicgenre)  

#### categories

- [musicartist](#)(100.0%)
  - M&L @198 (100.0%) on 07-feb-2011 [ Promotion of musicartist:beatles musicartistgenre:musicgenre:classic\_rock ]
  - CPL @1021 (80.9%) on 14-oct-2016 [ "numerous other artists including \_ " "traducidas de \_ " "incluidas en \_ " "had a guitar player" "early pioneers such as \_ " "controversial photo of \_ " "distressed image of \_ " "D-tracks of \_ " "Beatles Come Together \_ " "ohne die \_ " "opening band for \_ " "American acts like \_ " "classic acts like \_ " "performance footage of \_ " "were the perfect band" \_ " "record label" "record album by \_ " "les paroles de \_ " "never recorded the song \_ " "such renowned artists as \_ " "did a few songs" "top artists include \_ " "crazy lives of \_ " "UK artists such as \_ " "Lennon started \_ " "musical talent" \_ " "Birthplace" \_ " "harmonies" "Tour \_ " "starring \_ " "last days" \_ " "fourth album" \_ " "sixth studio album" \_ " "original recordings" "They were also pushing \_ " "She Said by \_ " "Other artists featured include \_ " "Post general comments related to \_ " "track also shows \_ " "such major artists as \_ " "time favorite band is \_ " "past masters such as \_ " "pop hooks of \_ " "popular musicians like \_ " "pop icons such as \_ " "music artists like \_ " "music bands like \_ " "pop stars such as \_ " "pop influenced by \_ " ]

## Introduction

## Why Knowledge Graphs?

## Knowledge Representation

## Knowledge Graphs

## Introduction

## Literals

Blank M

DBpedia

Namesp

## Serialization

## Programming

## Conclusions

## Architecture

## Notes and Further Reading

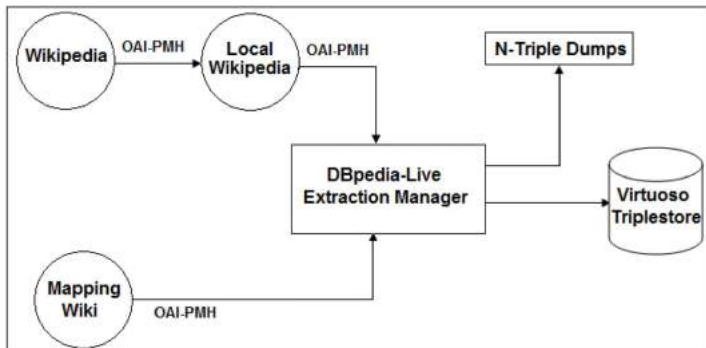
- 
- ```

graph TD
    MP[Max Planck] -- inGerman --> Stadt[Stadt]
    MP -- type --> city[city]
    MP -- type --> physicist[physicist]
    MP -- subclassOf --> scientist[scientist]
    MP -- subclassOf --> person[person]
    MP -- hasWon --> NP[Nobel Prize]
    MP --> 1919[1919]
    MP -- bornIn --> Kiel[Kiel]
    MP --> 45N07E[45°N07°E]
    MP --> 1858[1858]
    MP -- locatedIn --> Germany[Germany]
    MP -- initials --> Gitta[Gitta]
    city --> Stadt
    city -- type --> Kiel
    city -- type --> physicist
    city -- type --> scientist
    city -- type --> person
    city -- type --> NP
    city -- type --> 1919
    city -- type --> 45N07E
    city -- type --> 1858
    city -- type --> Germany
    city -- type --> Gitta
    physicist -- subclassOf --> scientist
    scientist -- subclassOf --> person
  
```

# Success story



- **DBpedia** is automatically extracted structured data from Wikipedia
  - 17M canonical entities
  - 88M type statements
  - 72M infobox statements



## Introduction

Motivation

Why Knowledge Graphs?

## History

Knowledge Representation

Semantic Web

Knowledge Graphs

## RDF

Introduction

RDF Triples

Literals

Blank Nodes

DBpedia

Namespaces

Serialization

Programming

## Data Integration Example

## Conclusions

Architecture

Examples

## Notes and Further Reading

# DeepDive



René Witte



## Introduction

Motivation

Why Knowledge Graphs?

## History

Knowledge Representation

Semantic Web

Knowledge Graphs

## RDF

Introduction

RDF Triples

Literals

Blank Nodes

DBpedia

Namespaces

Serialization

Programming

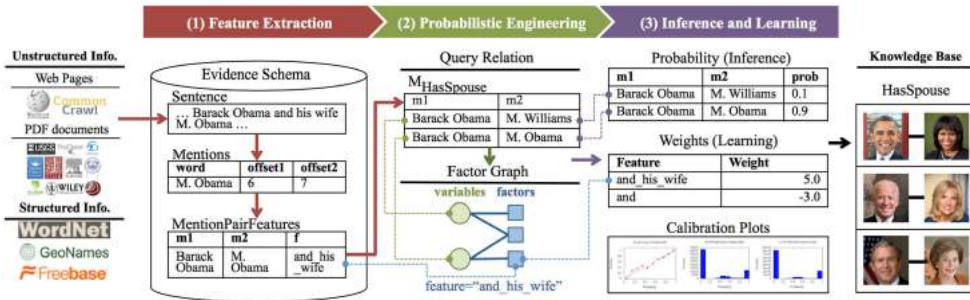
## Data Integration Example

## Conclusions

Architecture

Examples

## Notes and Further Reading

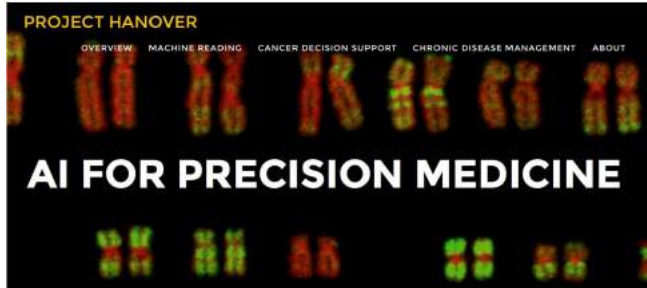


- Best Precision/recall/F1 in KBP-slot filling task 2014 evaluations (31 teams participated)



# Interesting application of Knowledge Graphs

---



Microsoft  
**Research**

## Chronic disease management:

develop AI technology for predictive and preventive personalized medicine to reduce the national healthcare expenditure on chronic diseases (90% of total cost)

### Introduction

Motivation

Why Knowledge Graphs?

### History

Knowledge Representation

Semantic Web

Knowledge Graphs

### RDF

Introduction

RDF Triples

Literals

Blank Nodes

DBpedia

Namespaces

Serialization

Programming

### Data Integration

### Example

### Conclusions

Architecture

Examples

### Notes and Further

### Reading

# Aristo Science QA challenge

- Science questions dataset

~5K 4-way multiple choice questions

Frogs lay eggs that develop into tadpoles and then into adult frogs. This sequence of changes is an example of how living things \_\_\_\_\_

- (A) go through a life cycle
- (B) form a food web
- (C) act as a source of food
- (D) affect other parts of the ecosystem

Science knowledge

frog's life cycle,  
metamorphosis



Common sense  
knowledge

frog is an animal,  
animals have life cycle

## Introduction

Motivation

Why Knowledge Graphs?

## History

Knowledge Representation

Semantic Web

Knowledge Graphs

## RDF

Introduction

RDF Triples

Literals

Blank Nodes

DBpedia

Namespaces

Serialization

Programming

## Data Integration

### Example

## Conclusions

Architecture

Examples

## Notes and Further

### Reading

# Knowledge Extraction

John was born in Liverpool, to Julia and Alfred Lennon.

Text

**NLP**

Lennon..  
John Lennon...

the Pool

Mrs. Lennon..  
.. his mother ..

his father  
he Alfred

Person

Location

Person

Person

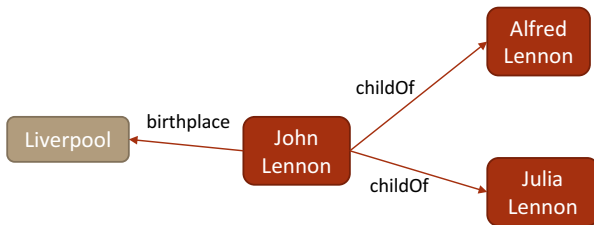
John was born in Liverpool, to Julia and Alfred Lennon.

NNP VBD VBD IN NNP TO NNP CC NNP NNP

Annotated text

**Information  
Extraction**

Extraction graph



[Introduction](#)

Motivation

Why Knowledge Graphs?

[History](#)

Knowledge Representation

Semantic Web

Knowledge Graphs

[RDF](#)

Introduction

RDF Triples

Literals

Blank Nodes

DBpedia

Namespaces

Serialization

Programming

[Data Integration](#)

[Example](#)

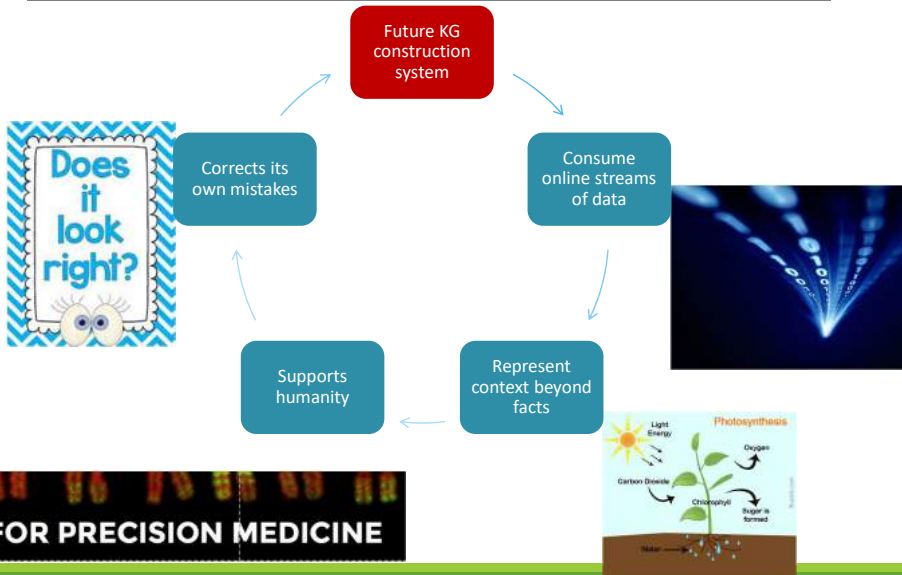
[Conclusions](#)

Architecture

Examples

[Notes and Further Reading](#)

# Future.....



## Introduction

Motivation  
Why Knowledge Graphs?

## History

Knowledge Representation  
Semantic Web  
Knowledge Graphs

## RDF

Introduction  
RDF Triples  
Literals  
Blank Nodes  
DBpedia  
Namespaces  
Serialization  
Programming

## Data Integration Example

## Conclusions

Architecture

Examples

## Notes and Further Reading

## 1 Introduction

### Introduction

- Motivation
- Why Knowledge Graphs?

## 2 History

### History

- Knowledge Representation
- Semantic Web
- Knowledge Graphs

## 3 The Resource Description Framework (RDF)

### RDF

- Introduction
- RDF Triples
- Literals
- Blank Nodes
- DBpedia
- Namespaces
- Serialization
- Programming

## 4 Example: Data Integration with Knowledge Graphs

### Data Integration Example

## 5 Conclusions

### Conclusions

- Architecture
- Examples

## 6 Notes and Further Reading

### Notes and Further Reading

## Introduction

Motivation

Why Knowledge Graphs?

## History

Knowledge Representation

Semantic Web

Knowledge Graphs

## RDF

Introduction

RDF Triples

Literals

Blank Nodes

DBpedia

Namespaces

Serialization

Programming

## Data Integration

### Example

## Conclusions

Architecture

Examples

## Notes and Further Reading

## Required

- [Yu14, Chapters 1, 2] (Introduction, RDF)

## Supplemental

- [Wor14] (RDF Primer)
- [RN10, Chapter 12] (Knowledge Representation)
- Graph databases: The best kept secret for effective AI,  
<https://www.youtube.com/watch?v=2ZzGMzitNgo>

- [Her] Ivan Herman.  
Tutorial on Semantic Web Technologies.  
<http://www.w3.org/People/Ivan/CorePresentations/RDFTutorial/>.
- [RN10] Stuart Russell and Peter Norvig.  
*Artificial Intelligence: A Modern Approach*.  
Prentice Hall, 3rd edition, 2010.  
[https://encore.concordia.ca/iii/encore/record/C\\_\\_Rb2591108?lang=eng](https://encore.concordia.ca/iii/encore/record/C__Rb2591108?lang=eng).
- [Wor14] World Wide Web Consortium (W3C).  
RDF 1.1 Primer.  
<http://www.w3.org/TR/rdf11-primer/>, 24 June 2014.
- [Yu14] Liyang Yu.  
*A Developer's Guide to the Semantic Web*.  
Springer-Verlag Berlin Heidelberg, 2nd edition, 2014.  
Available online at  
<https://concordiauniversity.on.worldcat.org/oclc/897466408>.