

# Lecture 2

## Knowledge Graphs

COMP 474/6741, Winter 2022

### 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
- 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

## 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

The screenshot shows a window titled "doctor" containing a text-based chat interface. The menu bar includes File, Edit, Options, Buffers, Tools, Text, and Help. The main area displays the following conversation:

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](#)



What is Concordia University?



**Concordia University** (French: Université **Concordia**; commonly referred to as **Concordia**) is a public comprehensive research **university** located in Montreal, Quebec, Canada. ... **Concordia** is a non-sectarian and coeducational institution, with more than 215,000 alumni worldwide.

Athletics brand: Concordia Stingers

Subsidiary or constituent schools: Loyola Campus, Faculty of Arts

Date founded: August 24, 1974

Geographic scope: Canada



## 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

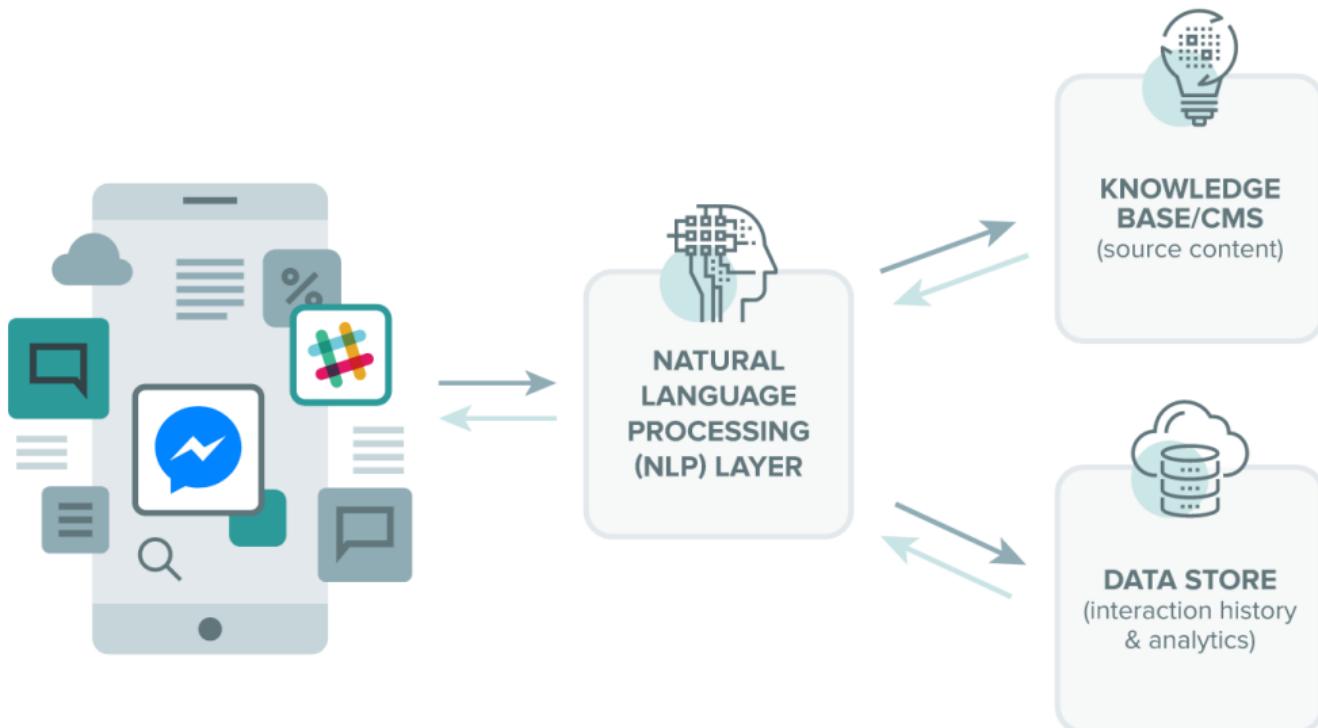
Architecture

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

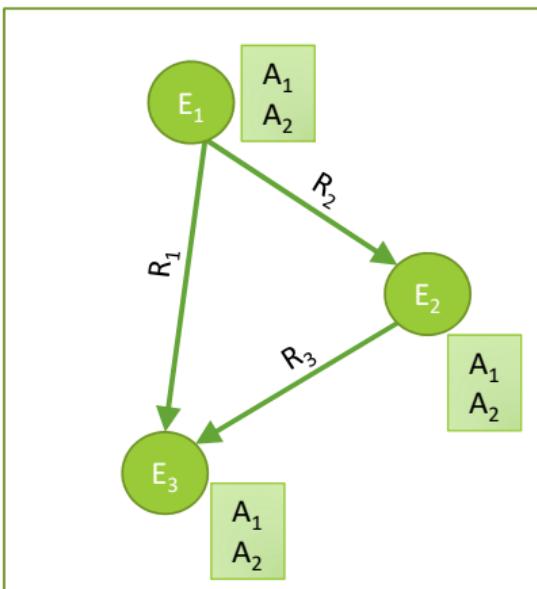
Examples

## Notes and Further Reading

[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?

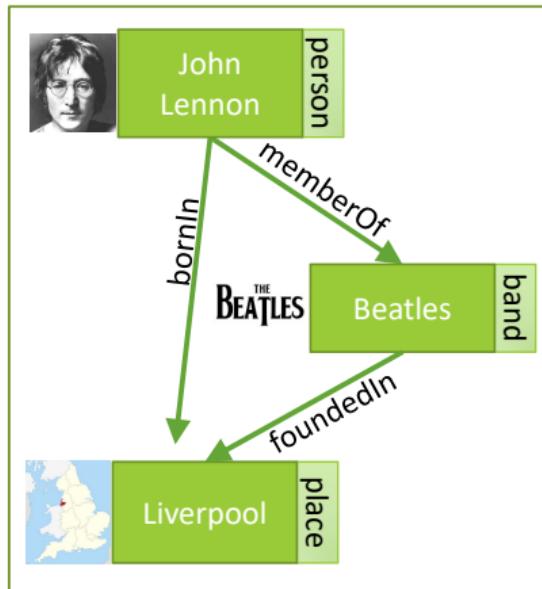
- Knowledge in graph form!
- Captures entities, attributes, and relationships
- Nodes are entities
- Nodes are labeled with attributes (e.g., types)
- Typed edges between two nodes capture a relationship between entities



[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](#)

# 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

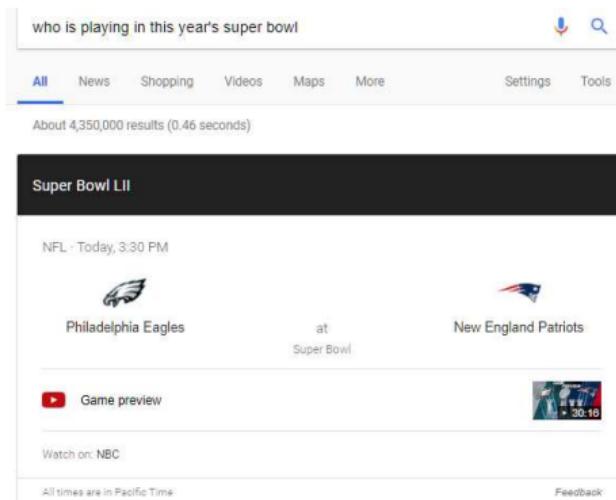


# Why knowledge graphs?

- Humans:
  - Combat information overload
  - Explore via intuitive structure
  - Tool for supporting knowledge-driven tasks
- Als:
  - 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



The screenshot shows a search results page for "who is playing in this year's super bowl". The top navigation bar includes "All", "News", "Shopping", "Videos", "Maps", "More", "Settings", and "Tools". The search bar shows the query. Below the search bar, it says "About 4,350,000 results (0.46 seconds)". The main content area is titled "Super Bowl LII" and shows the NFL game between the Philadelphia Eagles and the New England Patriots, scheduled for Today, 3:30 PM at the Super Bowl. It includes a "Game preview" section with a video thumbnail and a "Watch on: NBC" section. At the bottom, it notes "All times are in Pacific Time".

[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

IBM Watson Knowledge Studio

View Details Attribute View View Guidelines Completed 5 Close Alpha... 14pt 1

**2004-49-168A.txt**

A Merlin

1 V1, 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 V2, 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 front of V1 impacted the front of V2.

4 V1 rotated clockwise as V2 rotated counter-clockwise, and the left side of V1 impacted the right side of V2 in a sideslip configuration.

5 Both vehicles moved southwest to final rest.

6 Both vehicles were towed due to damage.

7 The unrestrained driver of V1 was hospitalized with foot and rib fractures as well as a liver laceration.

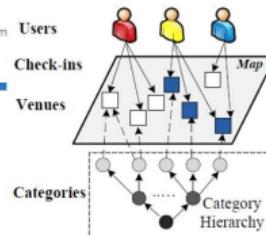
8 The restrained driver of V2 was treated and released with minor abrasion and contusion as well as a finger fracture.

9 The restrained male right passenger in V2 was pronounced brain dead two days later from brain injuries.

10 V1 was equipped with airbags and dual front airbags which deployed.

Entity Mention

Type	Subtype	Role
a	ACCIDENT_CAUSE	
b	ACCIDENT_OUTCOME	
c	CONDITION	
d	IMPACT	
f	MANUFACTURER	
m	MODEL	
y	MODEL_YEAR	
P	PART_OF_CAR	
p	PERSON	
s	STRUCTURE	
v	VEHICLE	



[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 3: Fueling Discovery

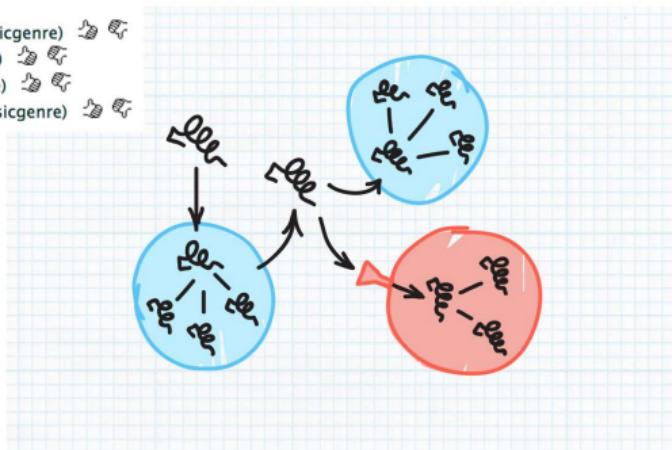
## beatles (musicartist)

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](#)

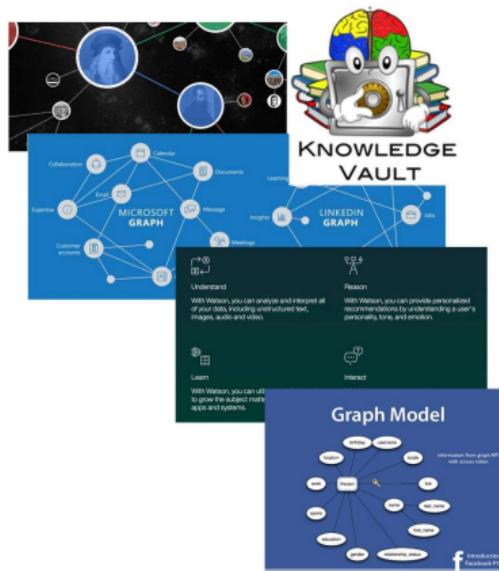
[Architecture](#)

[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						
	Mon 30th	00:18	07:06	12:36	19:32	
	Jan 2017	9.15m H	1.34m L	9.50m H	1.20m L	
	Tue 31st	00:55	07:43	13:14	20:10	
		9.18m H	1.36m L	9.49m H	1.25m L	
	Wed 1st	01:33	08:21	13:53	20:47	
	Fri 2017	9.10m H	1.51m L	9.37m H	1.42m L	
	Thu 2nd	02:14	08:59	14:36	21:27	
		8.91m H	1.76m L	9.15m H	1.70m L	
	Fri 3rd	09:00	09:42	15:31	22:12	
		8.86m H	2.00m L	8.84m H	2.04m L	
	Sat 4th	03:52	10:34	16:21	23:09	
		8.27m H	2.43m L	8.45m H	2.39m L	
	Sun 5th	04:59	11:42	17:34		
		7.95m H	2.71m L	8.13m H		
	Mon 6th	00:24	06:20	13:09	18:57	
		7.83m L	7.82m H	2.73m L	8.06m H	
	Tue 7th	01:49	07:39	14:31	20:13	
		2.56m L	8.03m H	2.42m L	8.29m H	
	Wed 8th	03:03	08:49	15:43	21:18	
		2.23m L	8.46m H	1.93m L	8.69m H	
	Thu 9th	04:08	09:47	16:45	22:14	
		1.82m L	8.94m H	1.41m L	9.07m H	
	Fri 10th	05:03	10:36	17:38	23:01	
		1.44m L	9.34m H	0.98m L	9.35m H	
	Sat 11th	05:51	11:21	18:24	23:44	
		1.17m L	9.81m H	0.75m L	9.47m H	
<small>The Beatles Total Album Sales Statistics</small>						
Labels	Total Number of Beatles albums sold					
	EMI - Polydor					
	Swan - Vertigo					
Associated acts	Total Albums Sold on Tunes					
Website	United Artists					
Past members	Total Singles Sold on Tunes					
	The Quarry - Cover Story					
	Preston - P - Discogs					
	thebeatles - United Kingdom					
	John Lennon - Germany					
	Paul McCartney - Australia					
	George Harrison - Japan					
	Ringo Starr - Argentina					
	Brazil					
	Sweden					
	Austria					
	Belgium					
<small>Beatles Billboard Chart Statistics</small>						
Total weeks on chart	1,279 weeks					
Total weeks at number one	16 weeks					
Total weeks at number one	178 weeks					
Album with longest time spent at number one (Please Please Me)	30 weeks					
<small>Beatles DATA SET</small>						

[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

## 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 Eltink, Producer [CONNECT](#)



(KFVS) - 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 photo) By ALLAN KOZINN DEC. 31, 2016

**dui** The Beatles  
The Harrison family is proud to announce the release of George Harrison - 1

**Ltd** The Beatles  
GEORGE HARRISON - THE JOHN COLLECTION

Released on 24th February, 2017, the vinyl box set includes all twelve of George's solo albums with exact replicas of the original track listing and artwork. Pre-order now and the box set can be ordered from [amazon.co.uk](#)

**WWD** George Harrison - The Vinyl Collection - Release February 2017

**YOUTUBE.COM** George Harrison - The Vinyl Collection, available to pre-order now with an exclusive & limited edition.

Like 0 Comment 0 Share

Like 0 love 0 comment 0 Share

Write a comment...

Jeffrey Smith What I would really be interested in is an "All Things Must Go" book on the Beatles. I would like to see what happened to the original record. I'd bet it would sound really good and I would buy it in a heartbeat. Like 0 reply 0 ago · January 17 at 10:10am · Edited

Dave Branning I can just see the grannies from the 60s and the grandmas in today's music stores holding their hands with glee over more while these various methods to make people buy their already bought and paid for record collections.

Like 0 reply 0 ago · January 17 at 10:10am · Edited

View more comments

The Beatles - George Harrison - 1

"Of very few individual songs can it be said, "This changed the course of popular music." "A Day in The Life" is one such song." - Richard Havers

The Beatles - A Day In The Life  
A Day in The Life Beatles Video Collection  
Or Now Get your copy here!



anager of the Beatles in 1960, he sent them on a stint in Germany

tagcraft. Press Association, via Associated Press

Introduction

Motivation

Why Knowledge Graphs?

History

Knowledge Representation

Semantic Web

Knowledge Graphs

RDF

Introduction

RDF Triples

Literals

Blank Nodes

DBpedia

Namespaces

Serialization

Programming

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



- 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

The Beatles - Topic

Home Videos Playlists Channels About

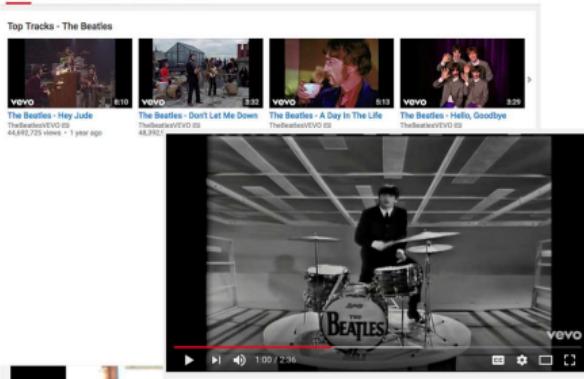
Top Tracks - The Beatles

The Beatles - Hey Jude The BeatlesVEVO 816 44,692,723 views • 1 year ago

The Beatles - Don't Let Me Down The BeatlesVEVO 832 48,751 views

The Beatles - A Day In The Life The BeatlesVEVO 839 329 views

The Beatles - Hello, Goodbye The BeatlesVEVO 839 329 views



The Beatles - I Want To Hold Your Hand - Performed Live On The Ed Sullivan Show 2/9/64



BED PEACE starring John Lennon & Yoko Ono

Yoko Ono 17,669 subscribers 852,022 views

[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](#)

# Outline

René Witte



## 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

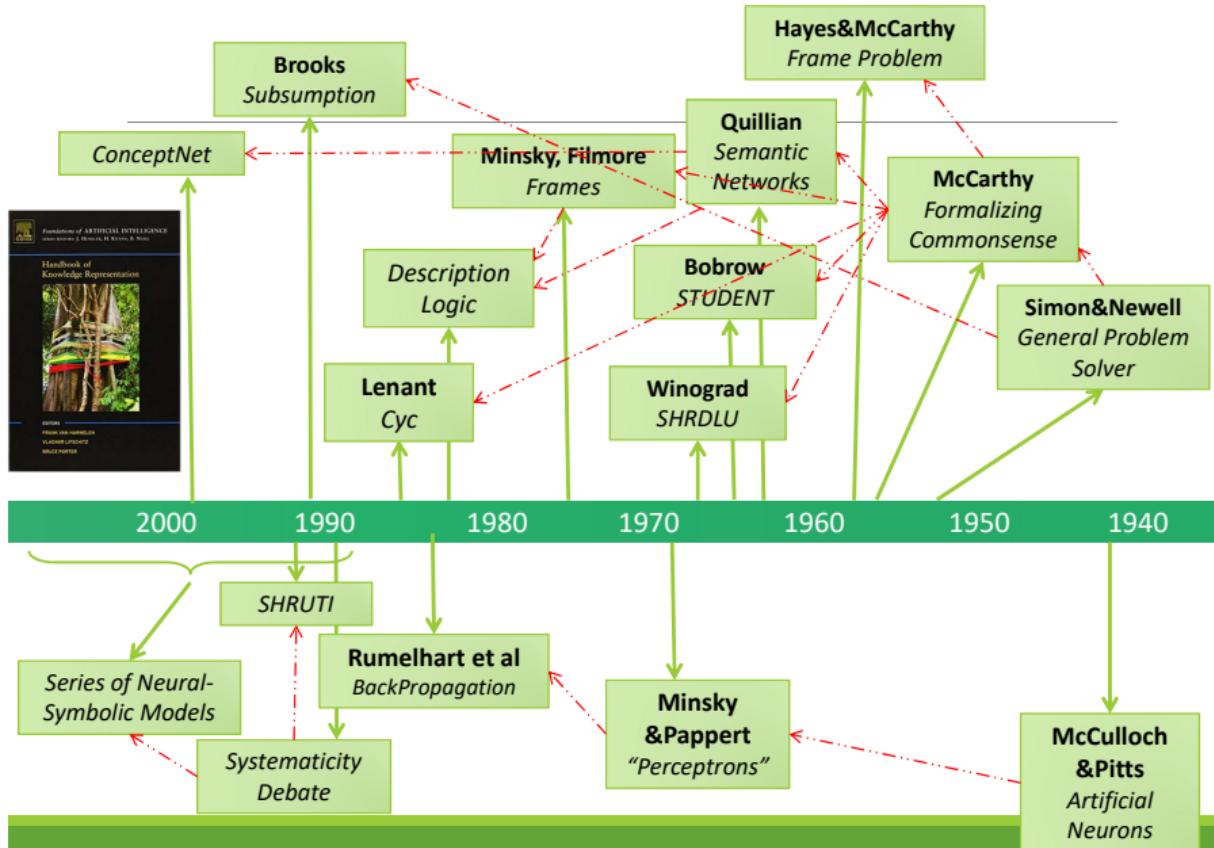
Examples

## 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

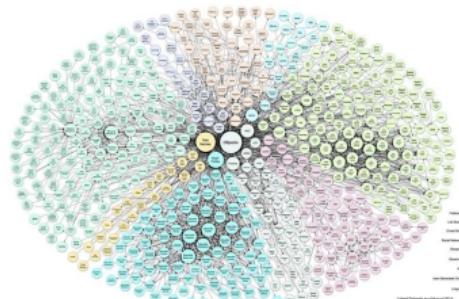
# Knowledge Representation

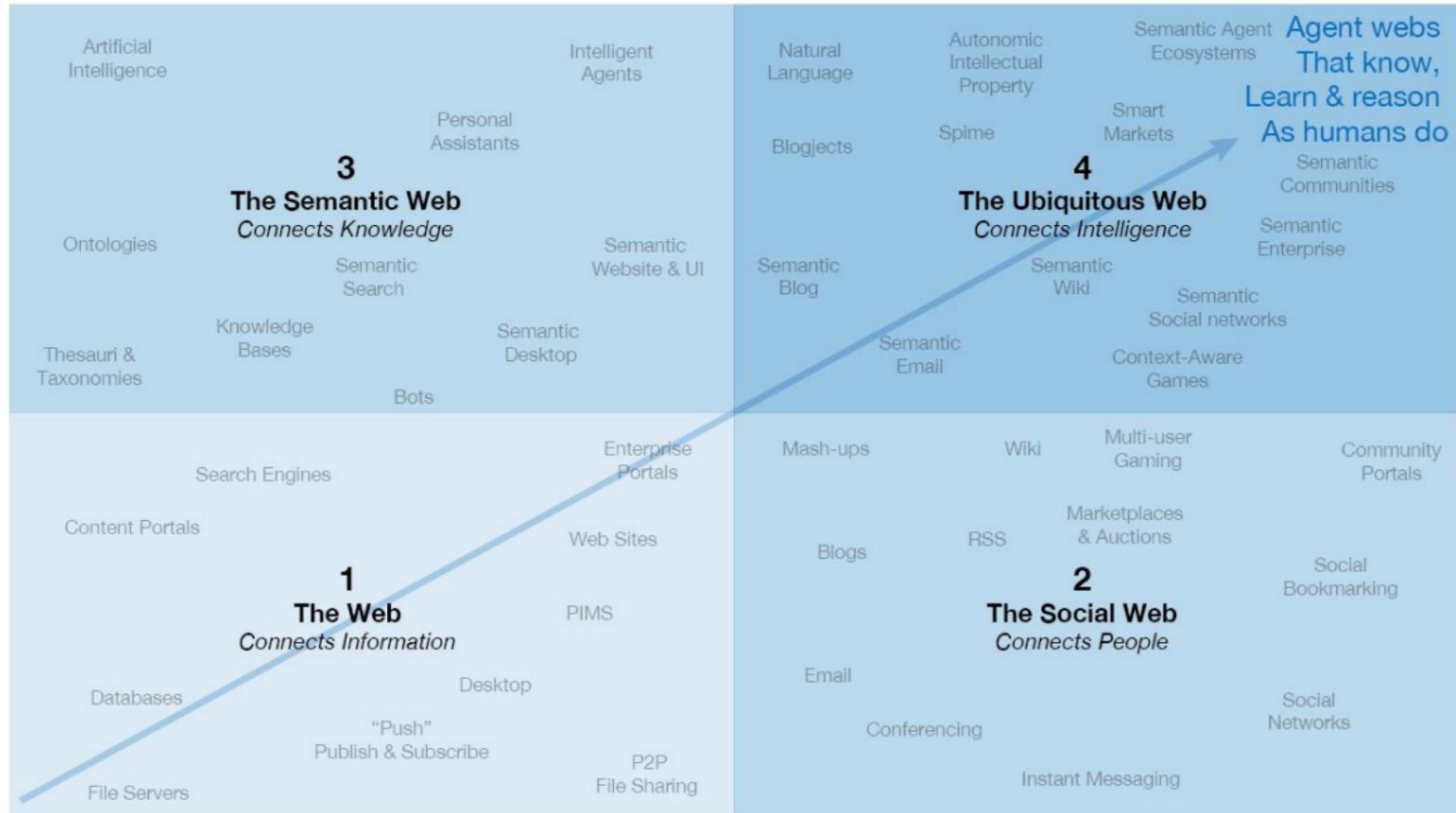
- Decades of research into knowledge representation
- Most knowledge graph implementations use RDF triples
  - <rdf:subject, rdf:predicate, rdf:object> : r(s,p,o)
  - Temporal scoping, reification, and skolemization...
- ABox (assertions) versus TBox (terminology)
- Common ontological primitives
  - 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](#)



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

René Witte



Apple Knowledge Navigator Video (1987)



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

## 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

## 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

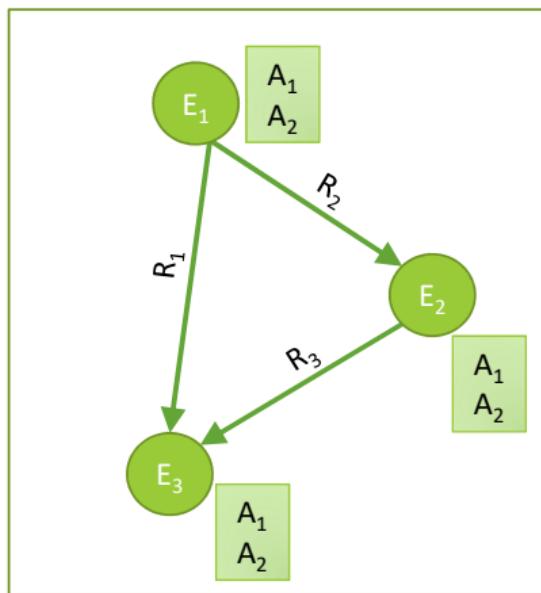
## Technology

- 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)

[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



[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](#)

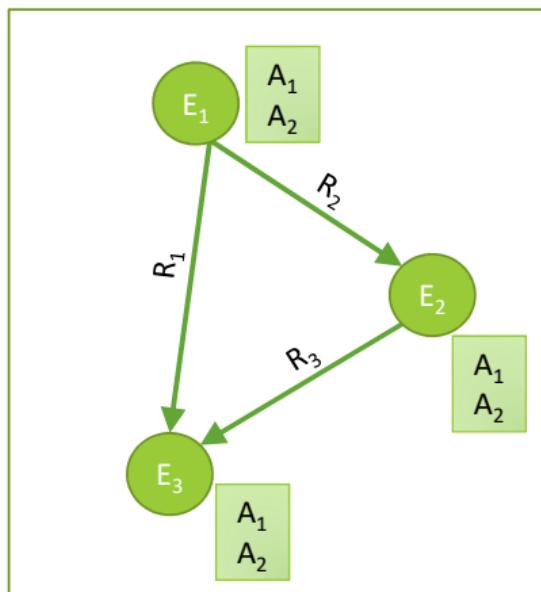
[Architecture](#)

[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

Introduction

Motivation

Why Knowledge Graphs?

## 2 History

History

Knowledge Representation

Semantic Web

Knowledge Graphs

## 3 The Resource Description Framework (RDF)

Introduction

RDF Triples

Literals

Blank Nodes

DBpedia

Namespaces

Serialization

Programming

RDF

Introduction

RDF Triples

Literals

Blank Nodes

DBpedia

Namespaces

Serialization

Programming

Data Integration  
Example

Conclusions

Architecture

Examples

Notes and Further  
Reading

## 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

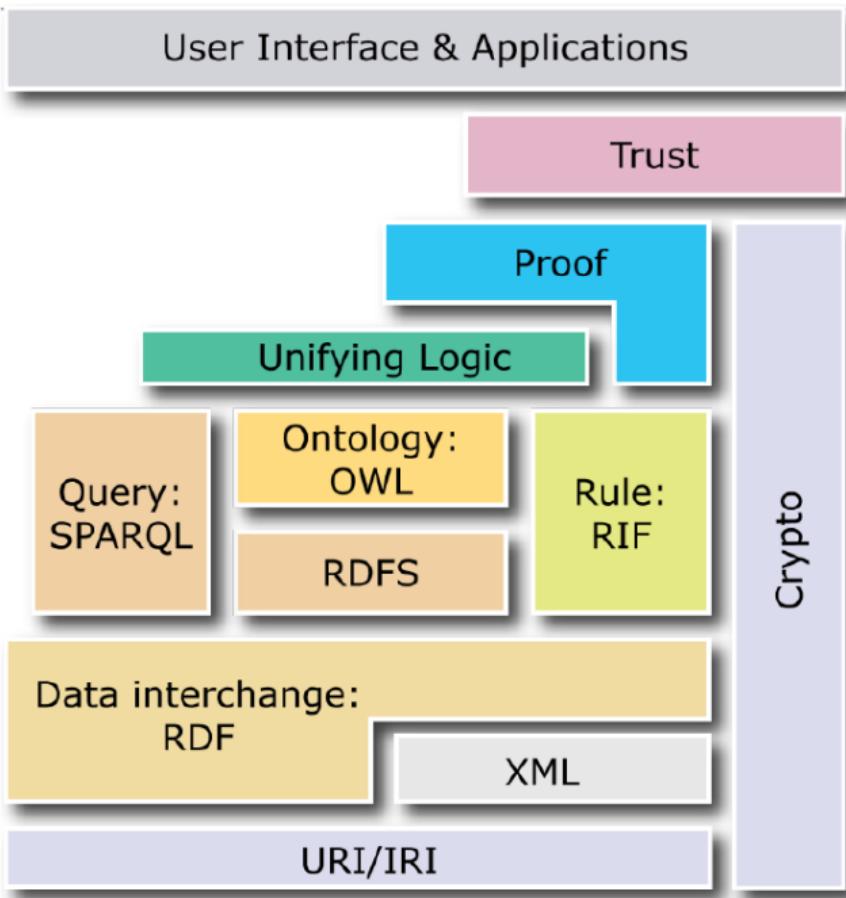
## Notes and Further

### Reading

# The Basis: RDF

# 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](#)

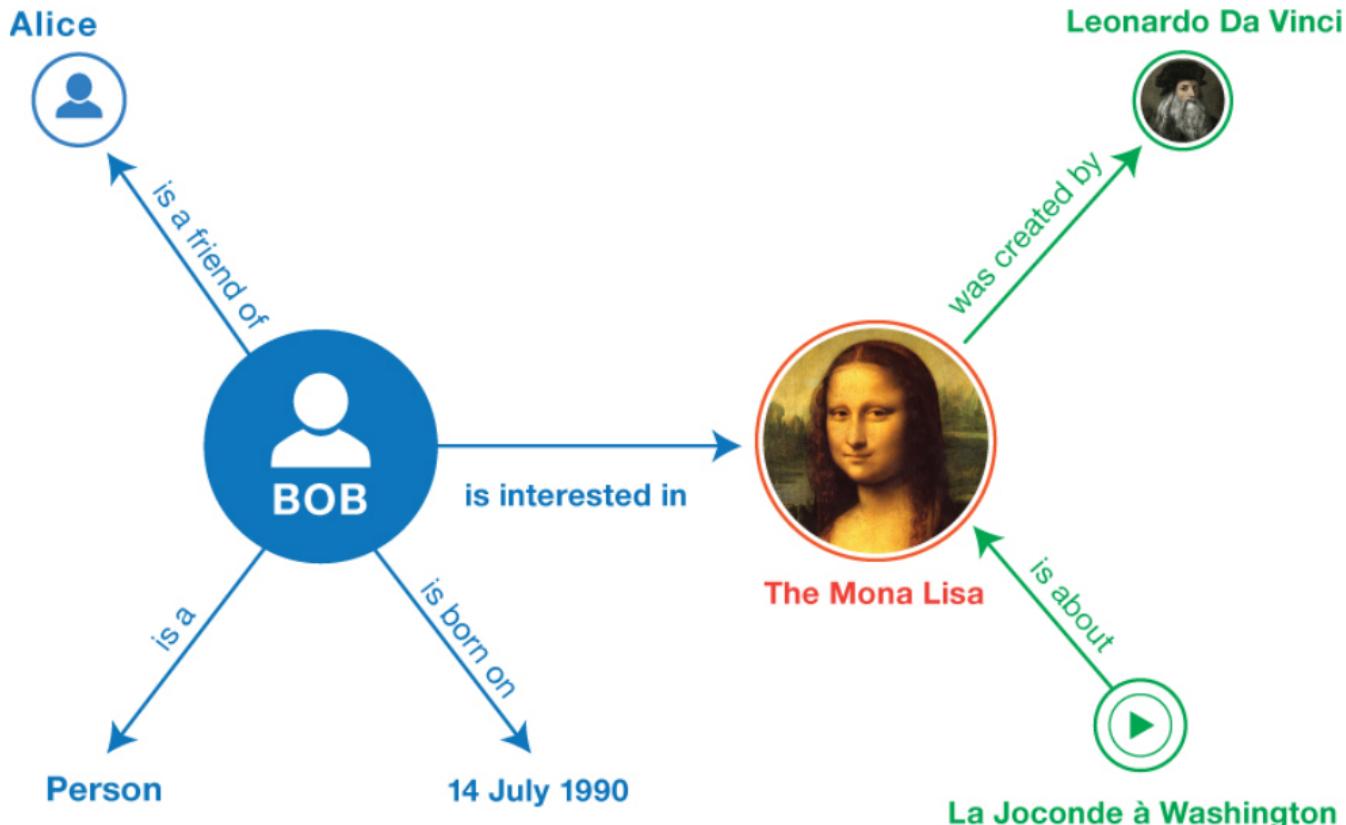
[Examples](#)

[Notes and Further](#)

[Reading](#)

# Knowledge as Graphs

René Witte



→ Worksheet #1: Tasks 1 & 2

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



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

## 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>.

[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](#)

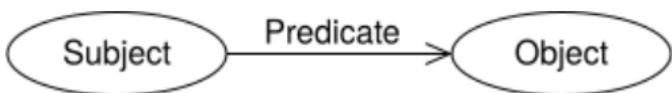
[Examples](#)

[Notes and Further](#)

[Reading](#)

→ **Worksheet #1: Tasks 3 & 4**

<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

[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

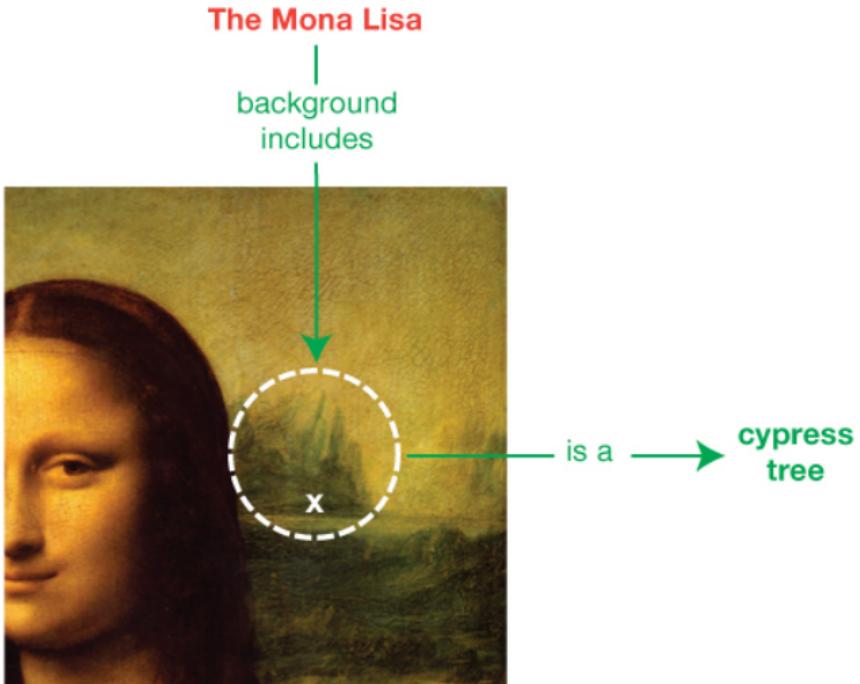
"Léonard de Vinci"@fr

- 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>



```
<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



Browse using ▾ Formats ▾

Faceted Browser Sparql Endpoint

## 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: [leo'nardo di 'ser 'pjero da (v)'vintʃi] (); 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="#">dbo:abstract</a>	■ Leonardo di ser Piero da Vinci (Italian: [leo'nardo di 'ser 'pjero da (v)'vintʃi] (); 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 Tuscany, Italy, Leonardo was educated in the studio of the renowned Italian painter Andrea del Verrocchio.

## DBpedia URIs

Make sure you use the correct URI:

- [http://dbpedia.org/resource/...](http://dbpedia.org/resource/Leonardo_da_Vinci) is the canonical URI
- The DBpedia server returns either
  - [http://dbpedia.org/page/...](http://dbpedia.org/page/Leonardo_da_Vinci) (HTML data, for a human)
  - [http://dbpedia.org/data/...](http://dbpedia.org/data/Leonardo_da_Vinci) (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](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/>
```

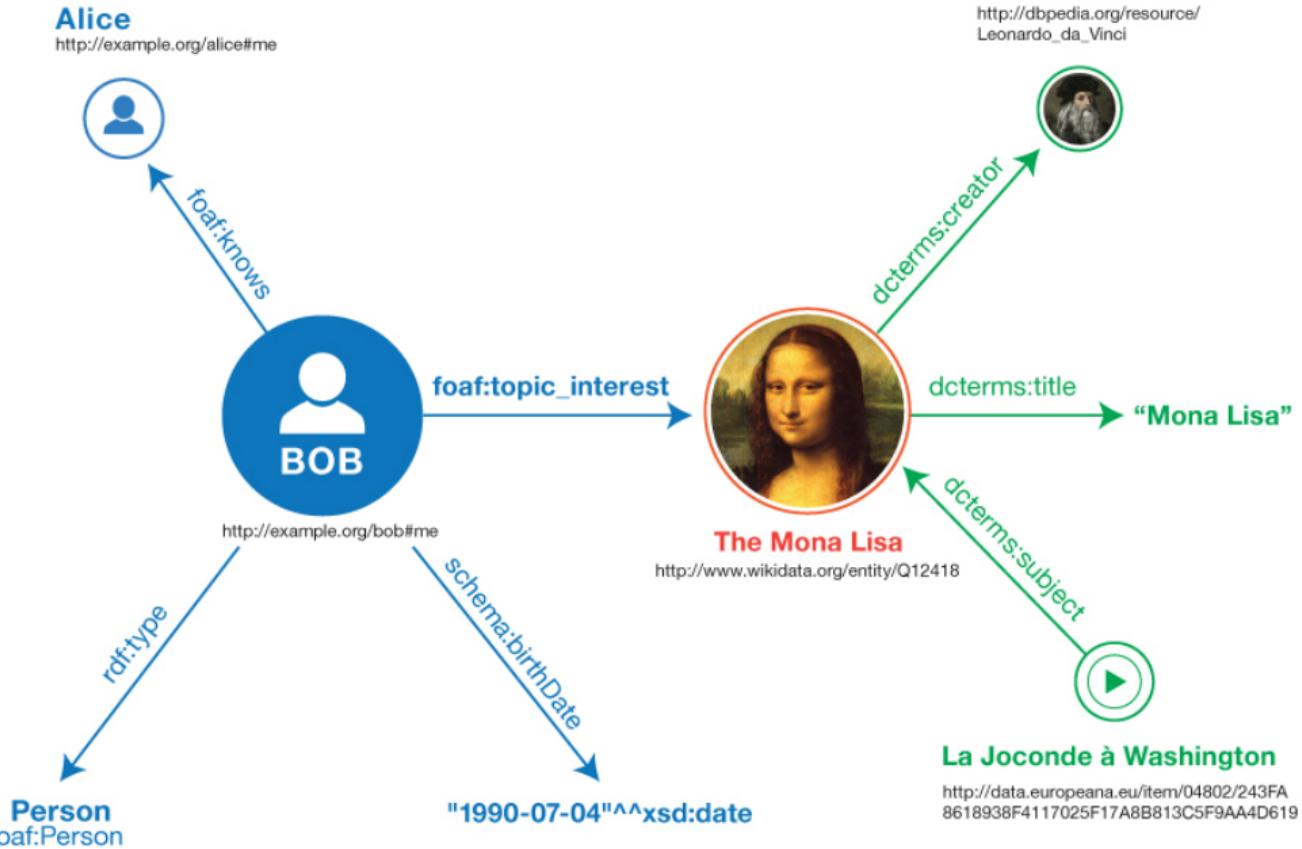
```
<bob#me>
  a foaf:Person ;
  foaf:knows <alice#me> ;
  schema:birthDate "1990-07-04"^^xsd:date ;
  foaf:topic_interest wd:Q12418 .
```

```
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



# 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="rdfformat")
# 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
- 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

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

René Witte



'A DOCTOR ZHIVAGO FOR THE FAR EAST' THE INDEPENDENT

# Amitav Ghosh THE GLASS PALACE

The magnificent, poignant, fascinating novel of three generations  
that starts in Mandalay ...



### 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”)

René Witte



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

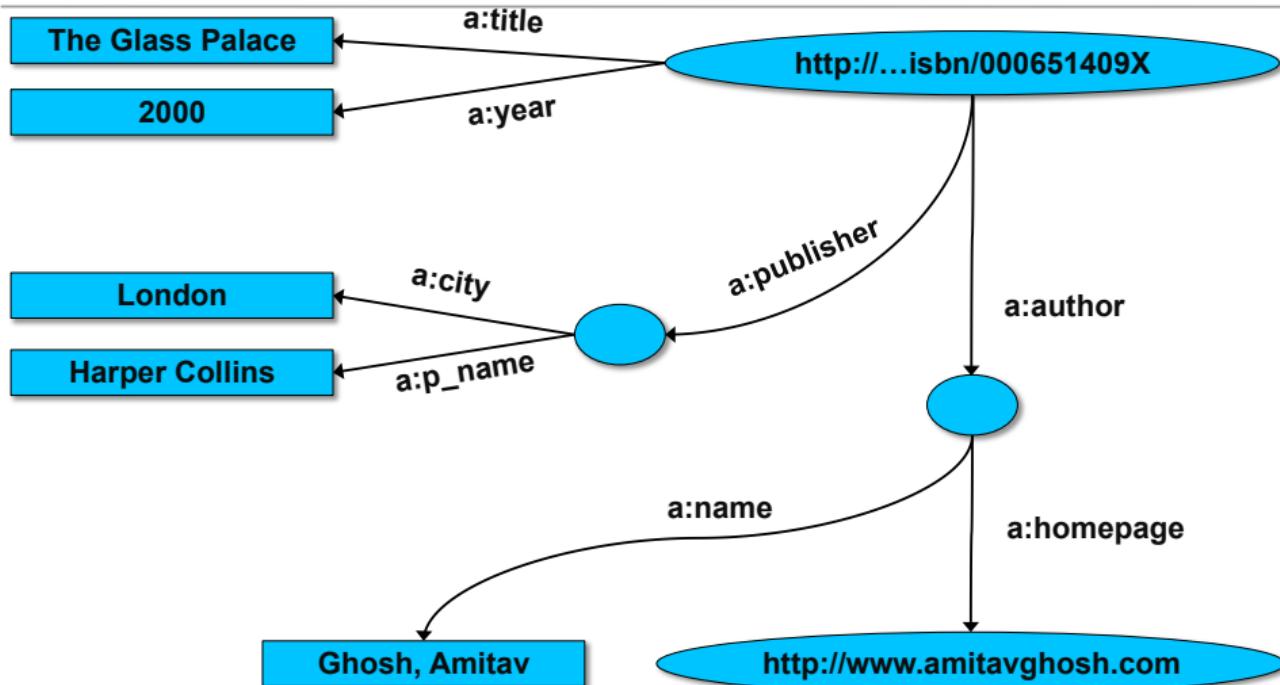
- Introduction
- RDF Triples
- Literals
- Blank Nodes
- DBpedia
- Namespaces
- Serialization
- Programming

## 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

---

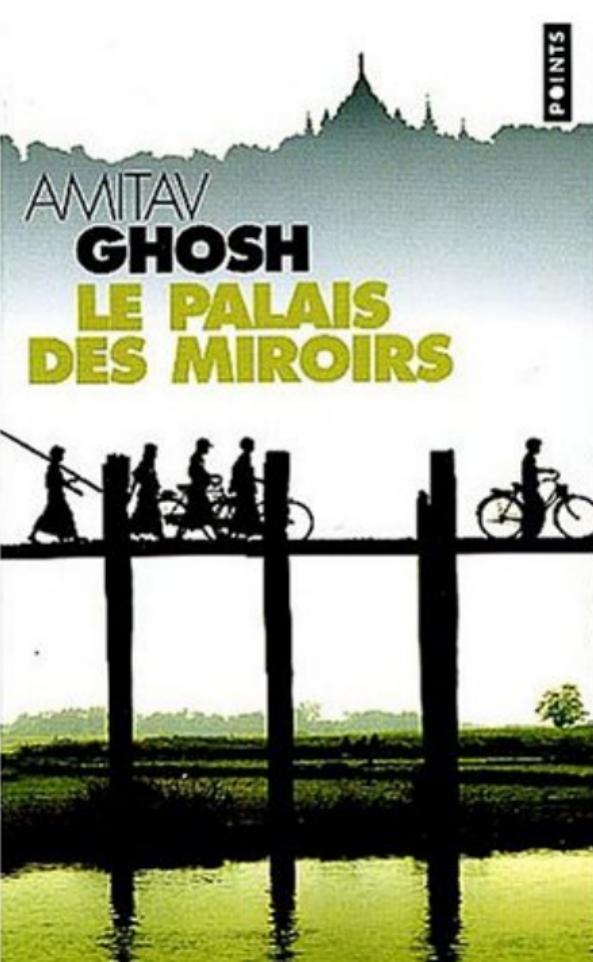
## ► Relations form a graph

- 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

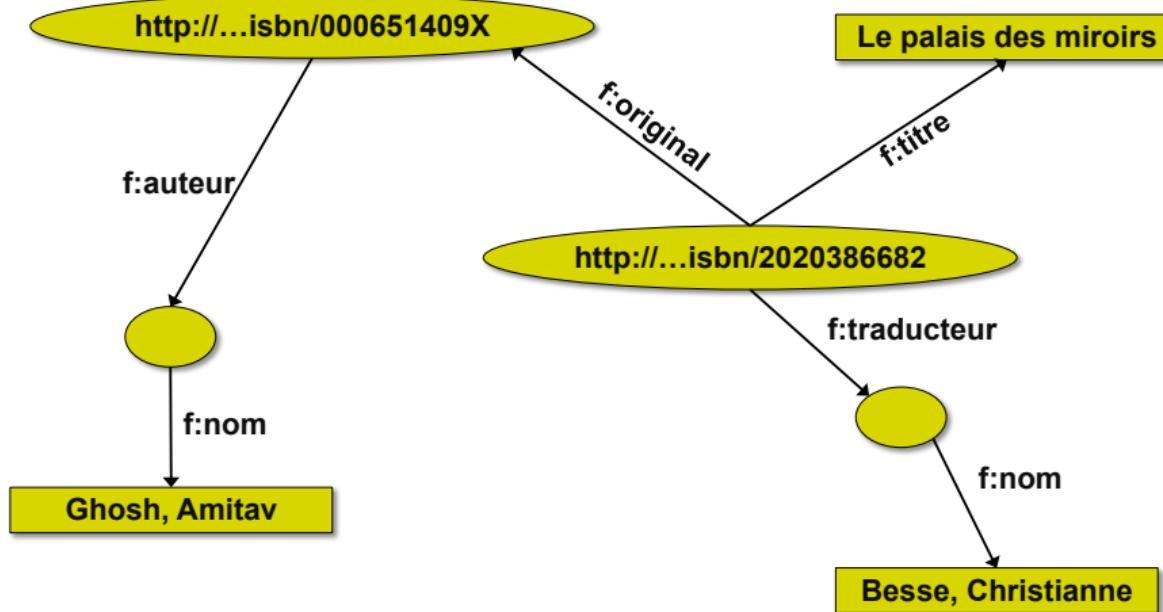
Architecture

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	\$A12\$
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

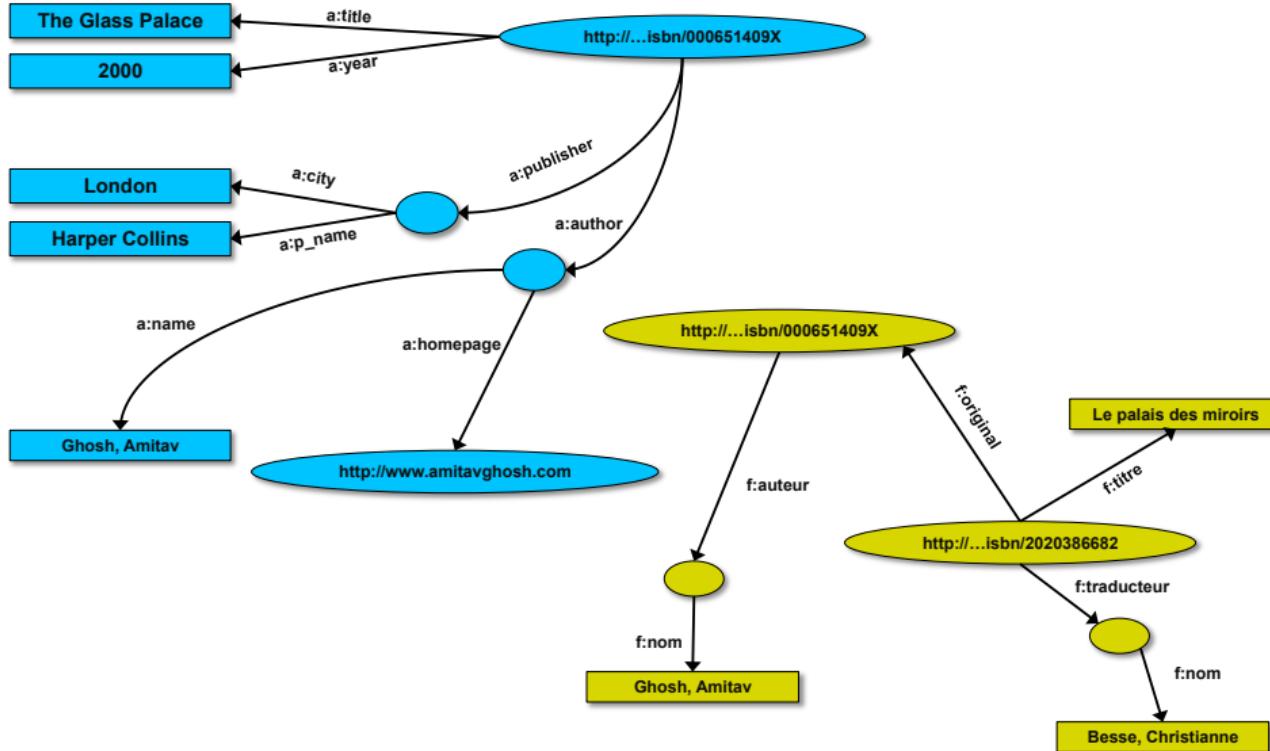
Architecture

Examples

Notes and Further  
Reading

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

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

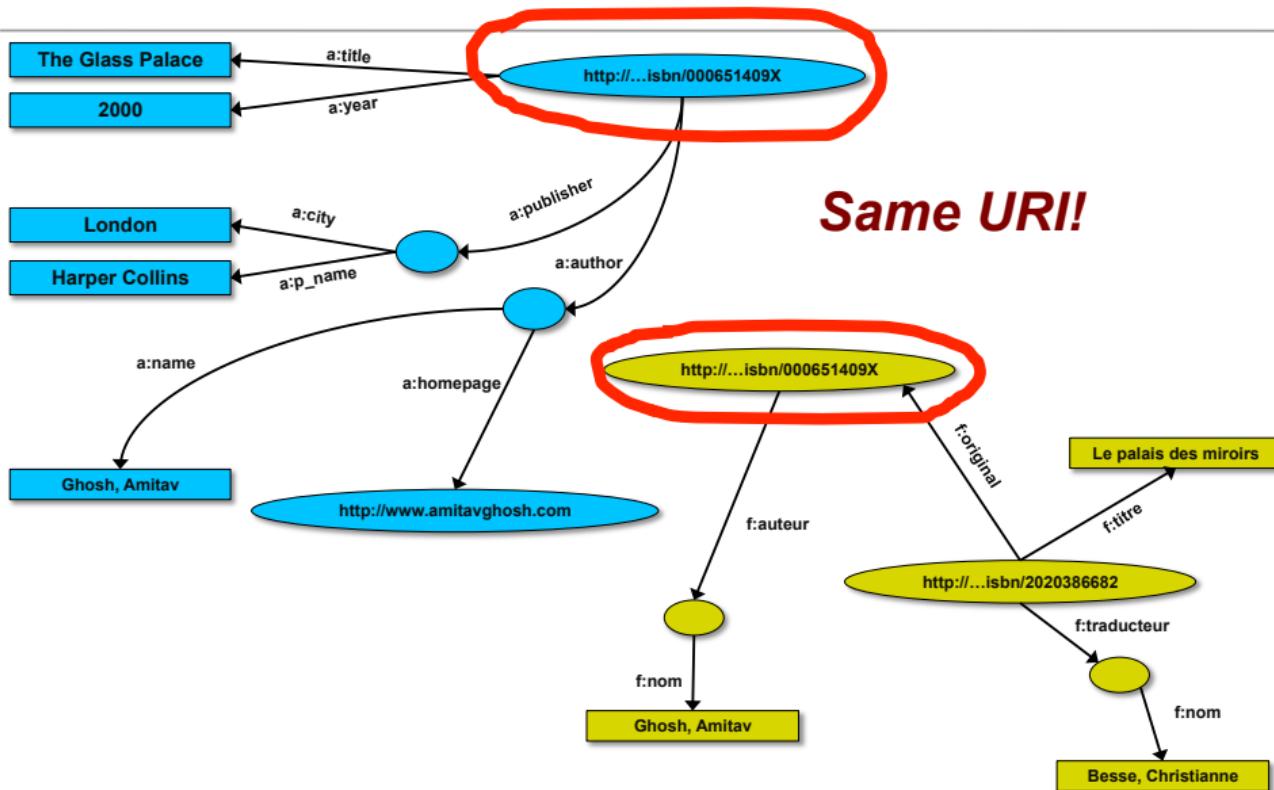
## Data Integration Example

- Conclusions
- Architecture
- Examples

## Notes and Further Reading

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

René Witte



## Introduction

- Motivation
- Why Knowledge Graphs?

## History

- Knowledge Representation
- Semantic Web
- Knowledge Graphs

## RDF

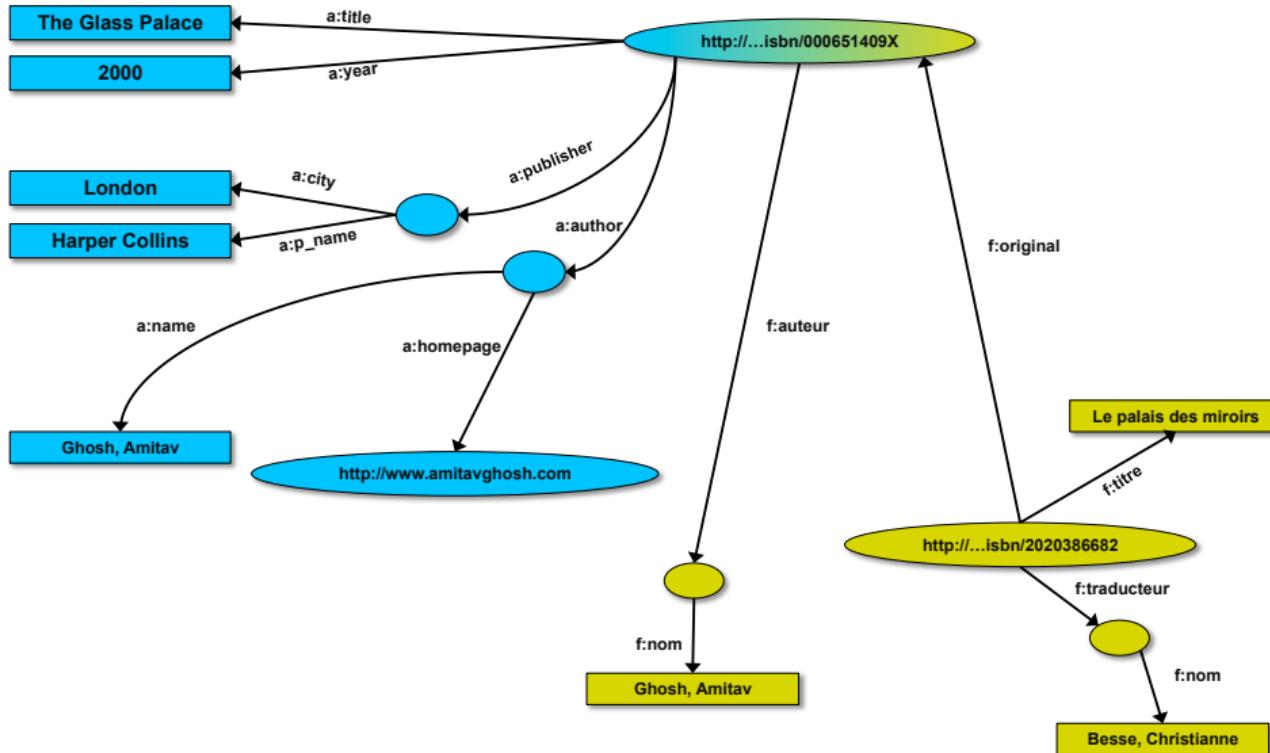
- Introduction
- RDF Triples
- Literals
- Blank Nodes
- DBpedia
- Namespaces
- Serialization
- Programming

## Data Integration Example

- Conclusions
- Architecture
- Examples

## Notes and Further Reading

# 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

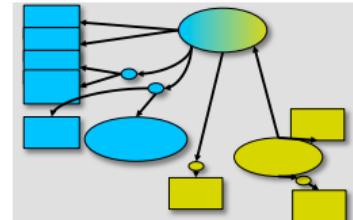
## Data Integration Example

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, ... « donnez-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](#)

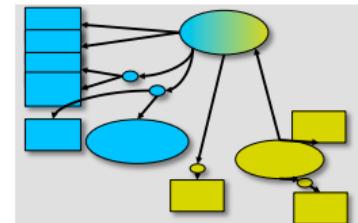
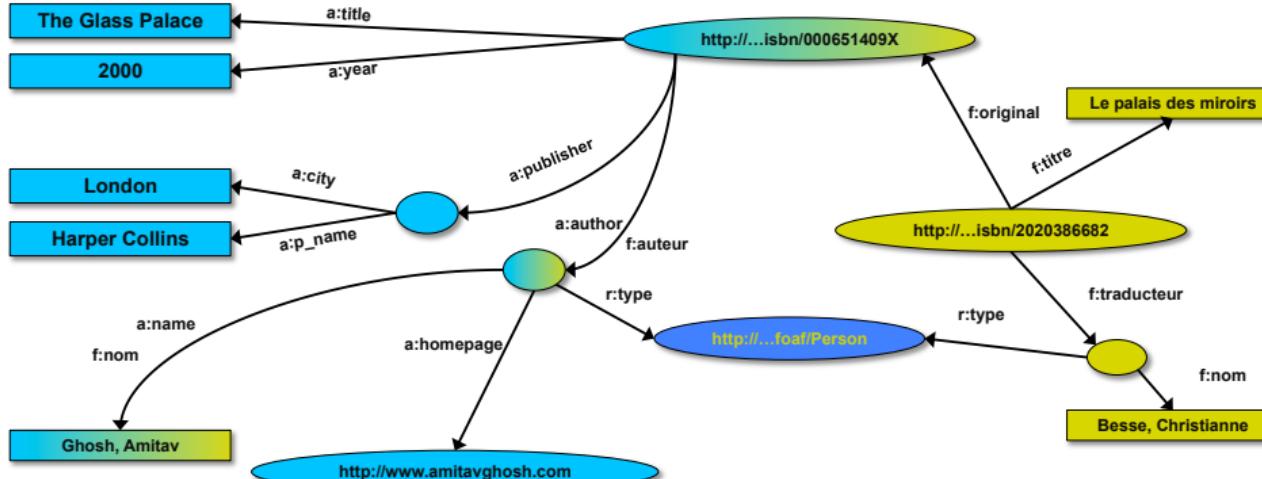
# However, more can be achieved...

---

- ▶ 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

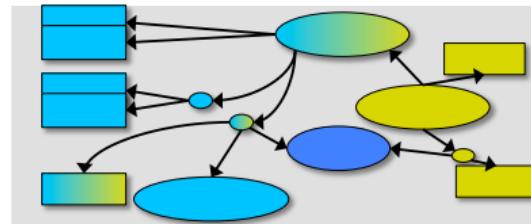
## Data Integration Example

Conclusions  
Architecture  
Examples

Notes and Further Reading

# Start making richer queries!

- ▶ 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”



## 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

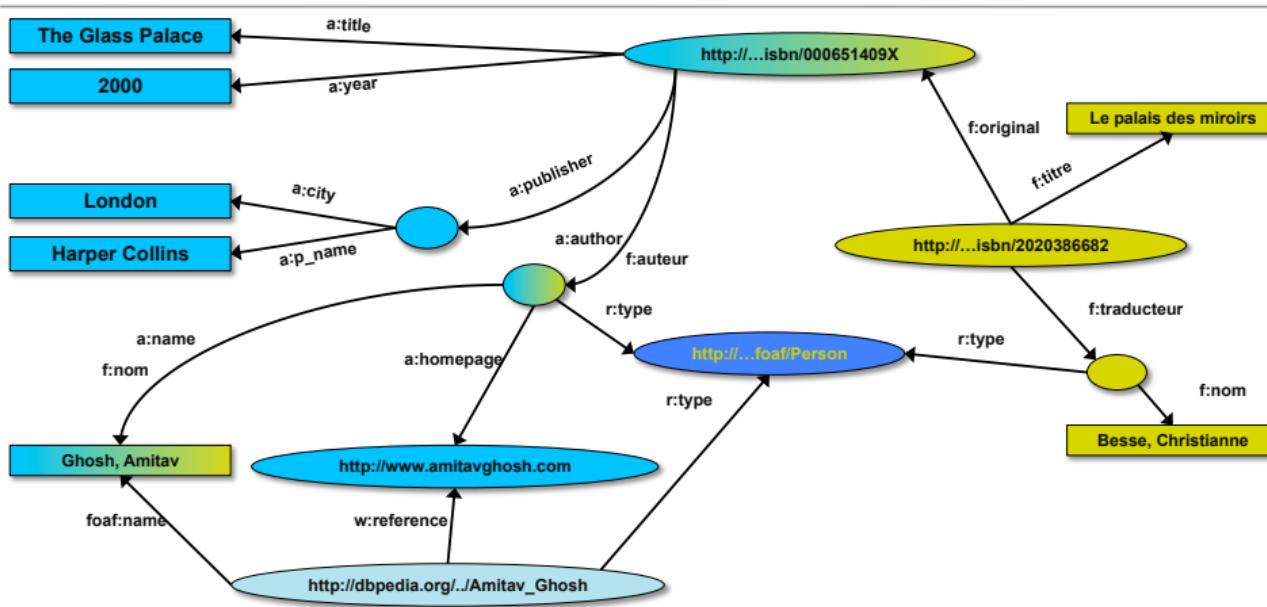
# Combine with different datasets

- ▶ 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

René Witte



## Introduction

- Motivation
- Why Knowledge Graphs?

## History

- Knowledge Representation
- Semantic Web
- Knowledge Graphs

## RDF

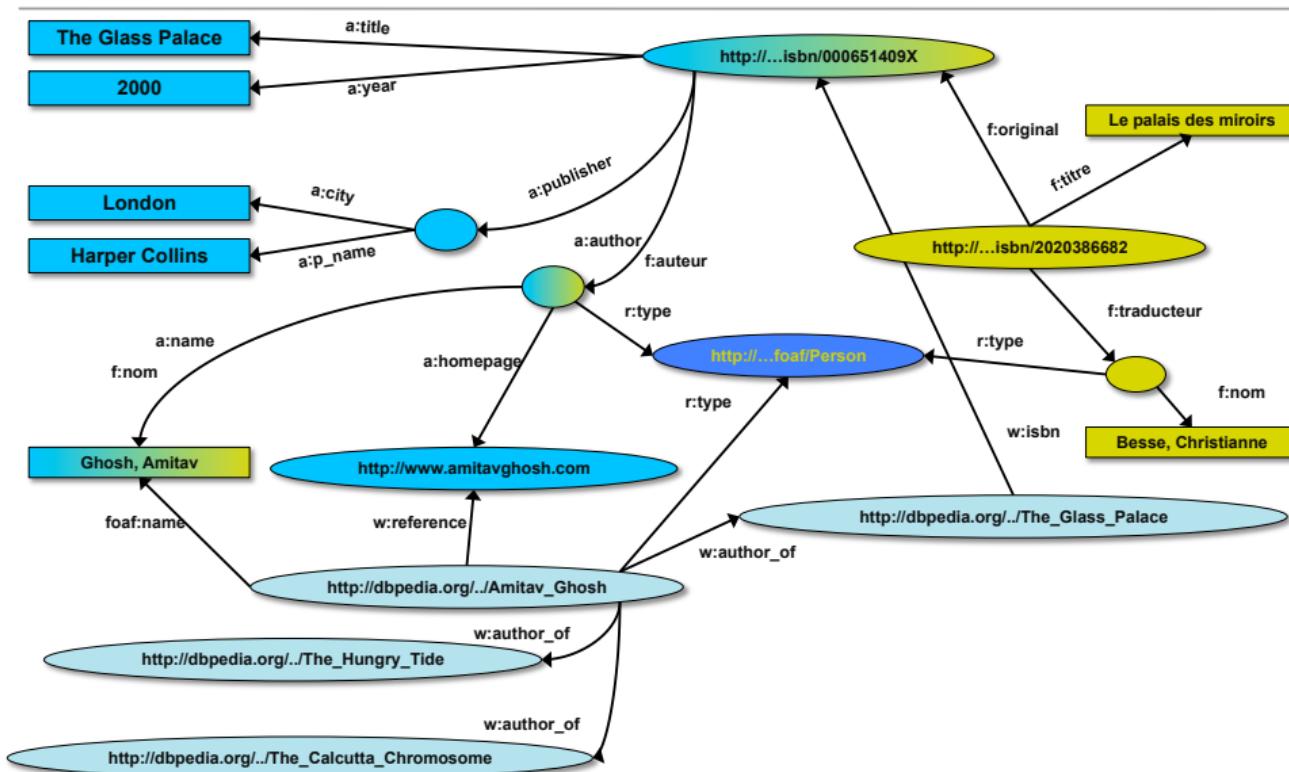
- Introduction
- RDF Triples
- Literals
- Blank Nodes
- DBpedia
- Namespaces
- Serialization
- Programming

## Data Integration Example

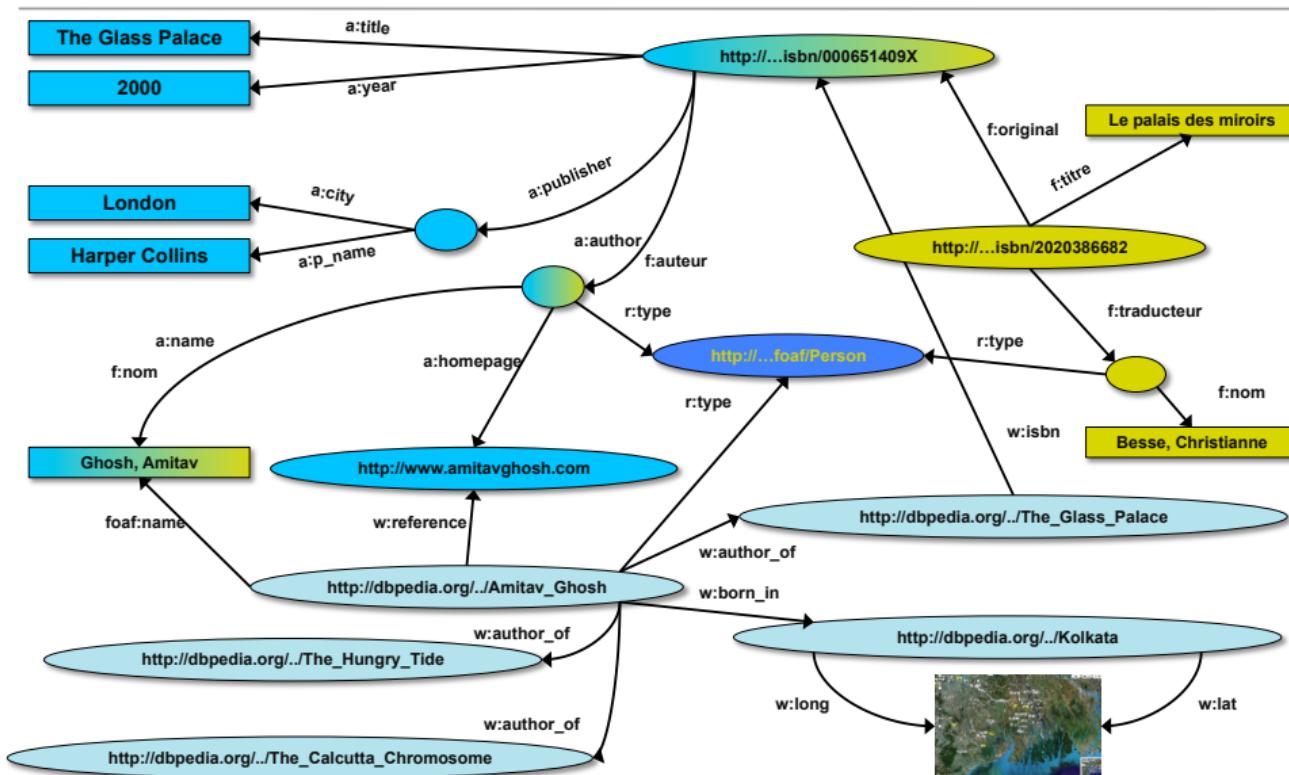
- Conclusions
- Architecture
- Examples

## Notes and Further Reading

## Merge with Wikipedia data



# 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

→ **Worksheet #1: Task 10**

[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](#)

# Outline

René Witte



## 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

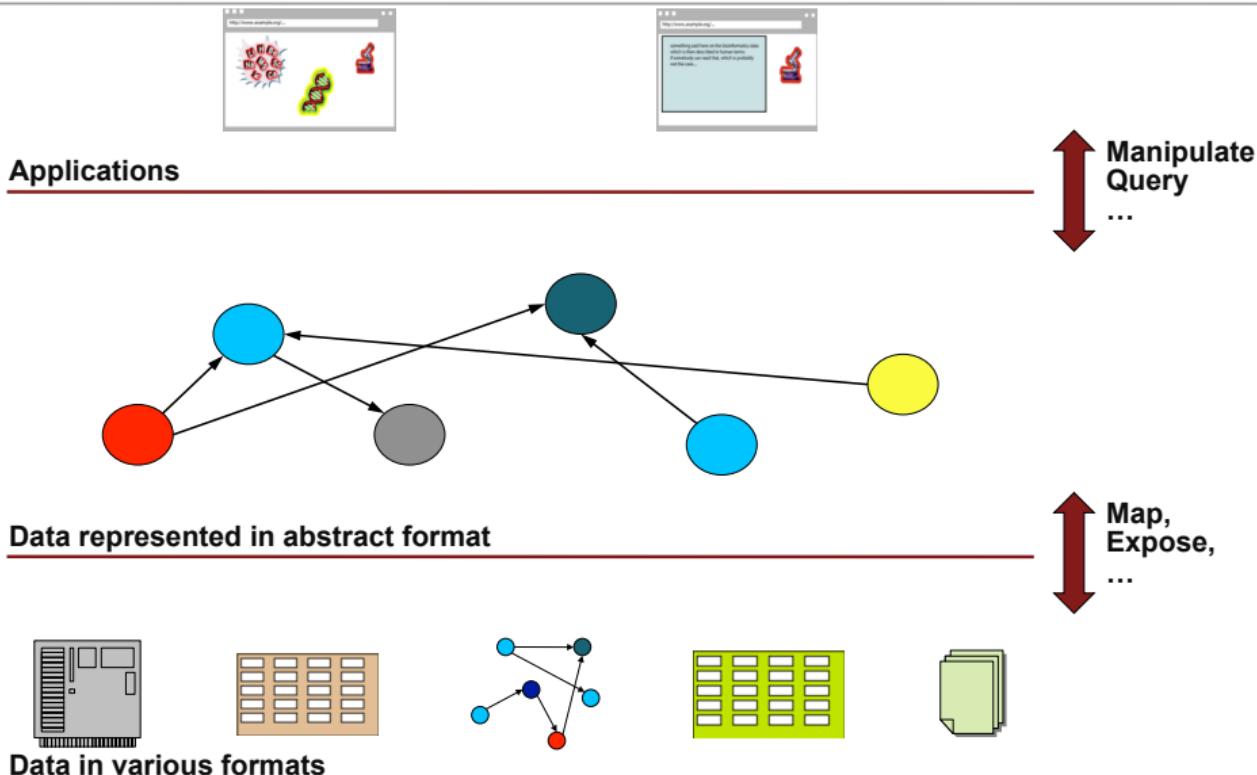
Examples

## 6 Notes and Further Reading

Notes and Further  
Reading

# What did we do?

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

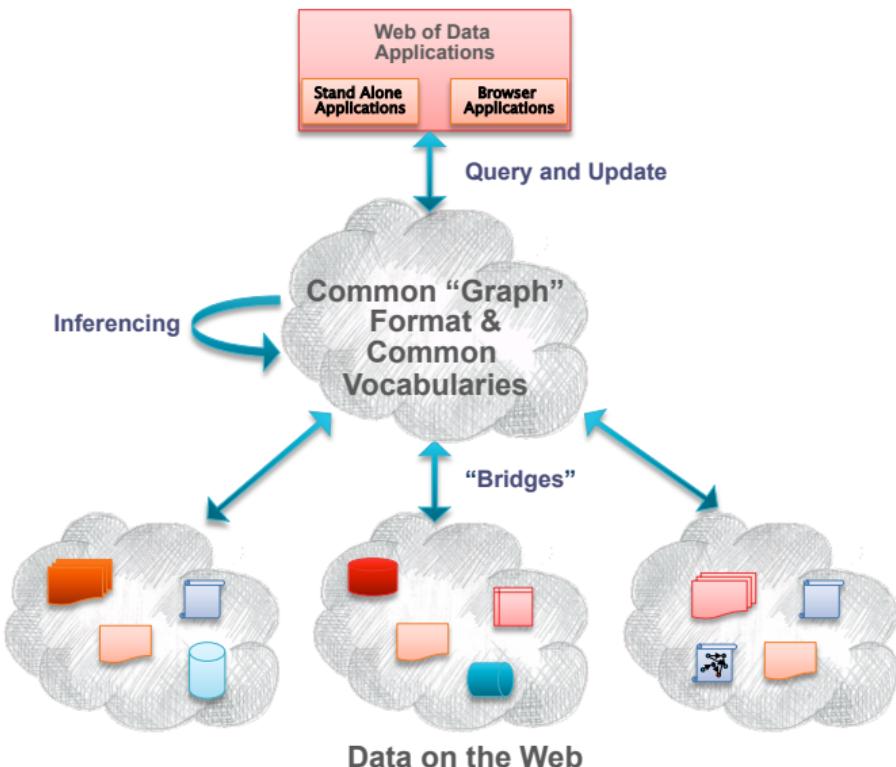
## Data Integration Example

- ### Conclusions
- Architecture
  - Examples

## Notes and Further Reading

# What 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)

Hosted by  Created at 

---

Argument 1:

Argument 2:

Relation:

All

[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

**NELL Knowledge Base Browser**

CMU Read the Web Project

Search

log in | preferences | help/instructions | feedback

**categories**    **relations**

---

**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%)
  - MBL @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 also were 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

## Motivation

## Why Knowledge Graphs?

## History

Knowledge Representation

Semantic Web

## Knowledge Graphs

RDF

## Introduction

RDF T

## Literals

Blank Note

DBpedia

## Namespaces

## Serialization

Programming

## Data Integration

## Exercises

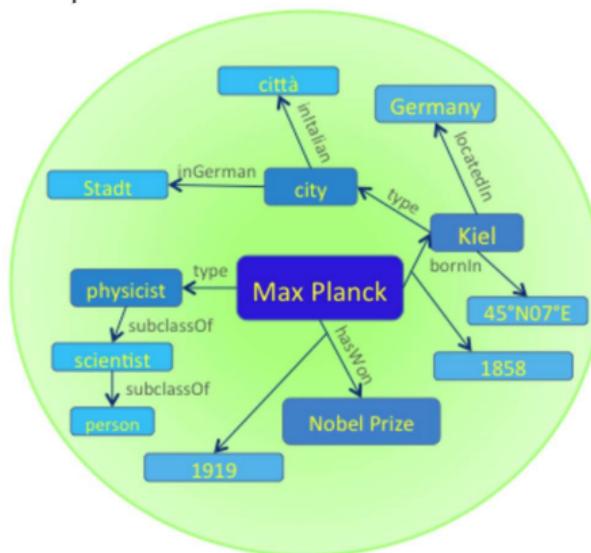
## Conclusion

Architecture

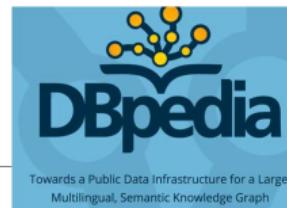
## Notes and Further

# Success story: YAGO

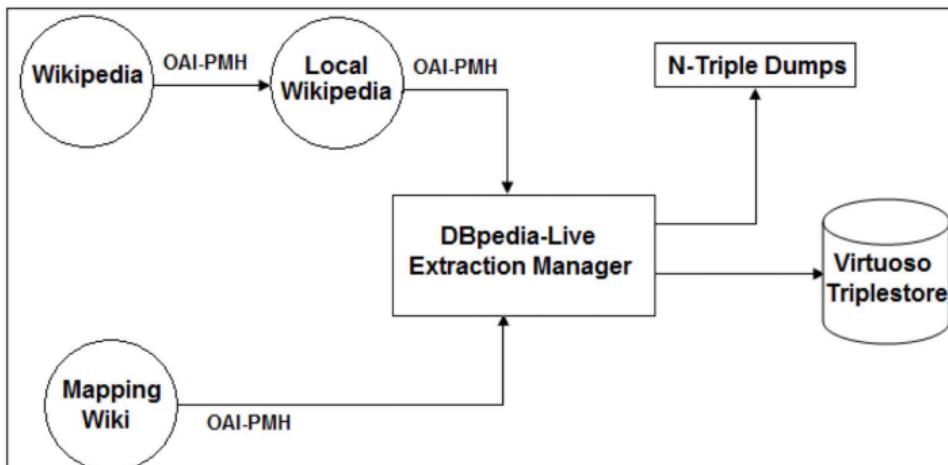
- **Input:** Wikipedia infoboxes, WordNet and GeoNames
- **Output:** KG with 350K entity types, 10M entities, 120M facts
- Temporal and spatial information

[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



- 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](#)

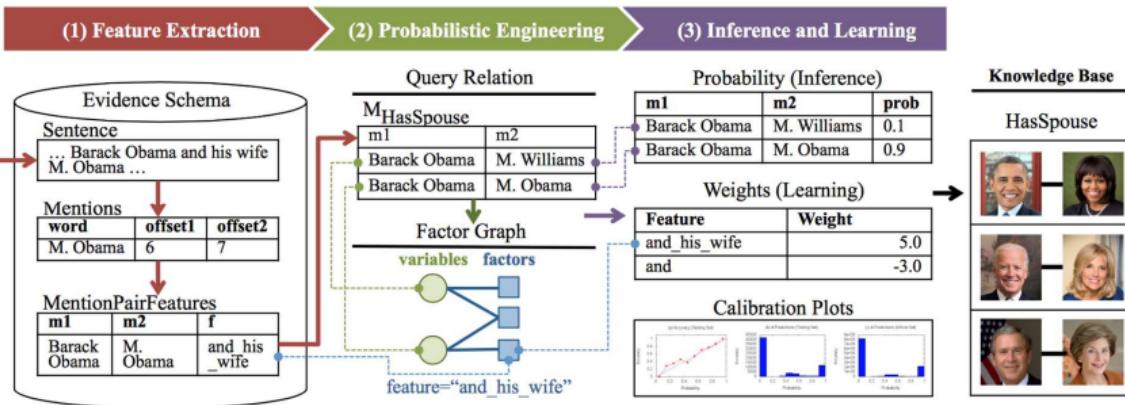
# DeepDive



Unstructured Info.
Web Pages
Common Crawl
PDF documents
USGS ProQuo
Wiley

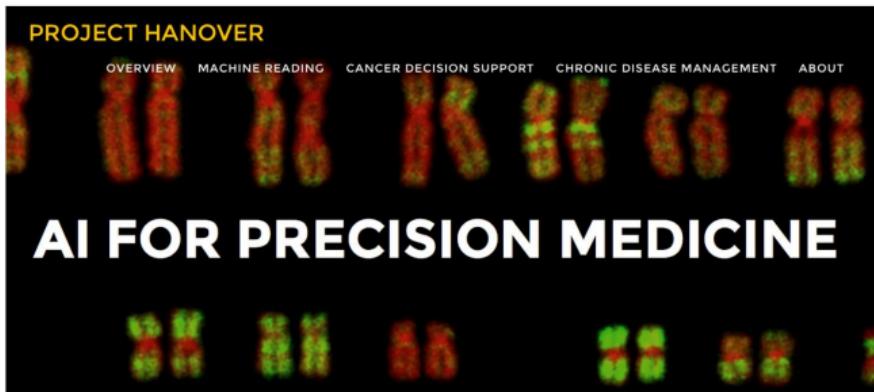
Structured Info.
WordNet
GeoNames
Freebase



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

[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](#)

# 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](#)

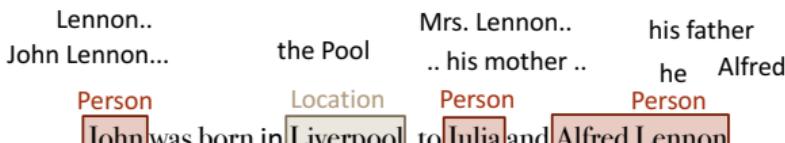
[Reading](#)

# Knowledge Extraction

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

Text

NLP

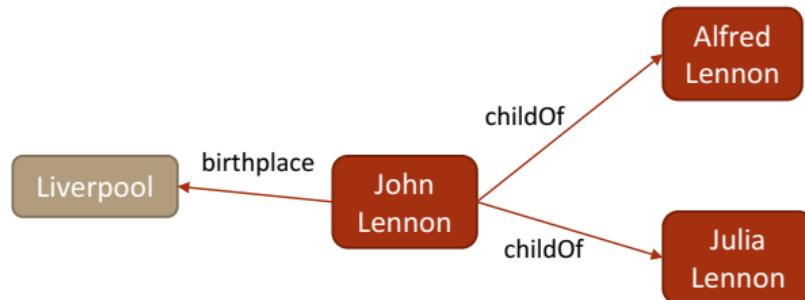


Annotated text

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

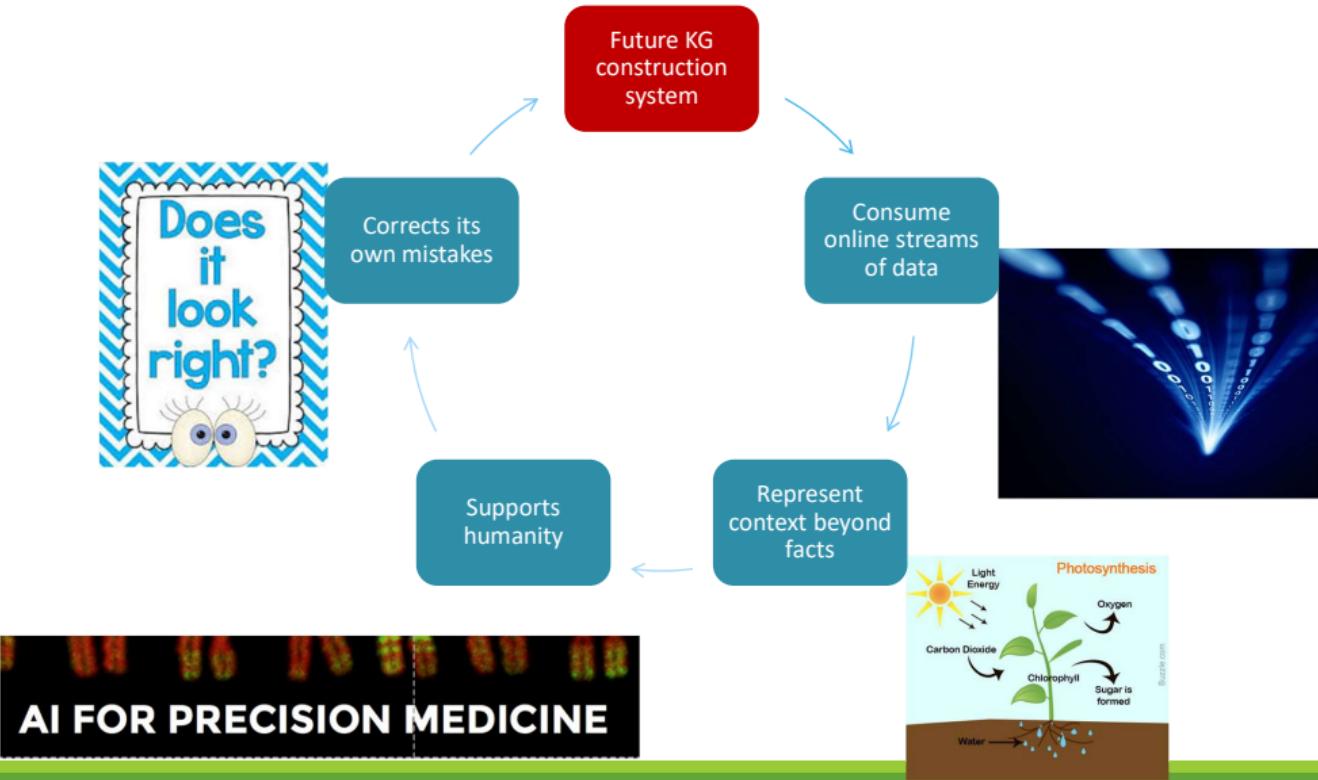
Conclusions

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
- 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

## 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>

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

# References

René Witte



- [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://concordiauniversity.on.worldcat.org/oclc/359890490>.
- [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.  
<https://concordiauniversity.on.worldcat.org/oclc/897466408>.

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