



Knowledge Graphs

Lecture 5 – Ontological Engineering for Smarter Knowledge Graphs

5.5 Knowledge Graph Construction

Prof. Dr. Harald Sack & Mary Ann Tan

FIZ Karlsruhe – Leibniz Institute for Information Infrastructure

AIFB – Karlsruhe Institute of Technology

Autumn 2023



FIZ Karlsruhe

Leibniz-Institut für Informationsinfrastruktur

Knowledge Graphs

Lecture 5: Ontological Engineering for Smarter Knowledge Graphs

5.1 Beyond the Limits of OWL

Excursion 7: The Semantic Web Rule Language SWRL

5.2 How to design your own Ontology

5.3 How to design better Ontologies

5.4 Ontological Engineering

5.5 Knowledge Graph Construction

5.6 Ontologies & Knowledge Graphs – Best Practices

Knowledge Graph Construction

Knowledge Graph Construction and Enrichment

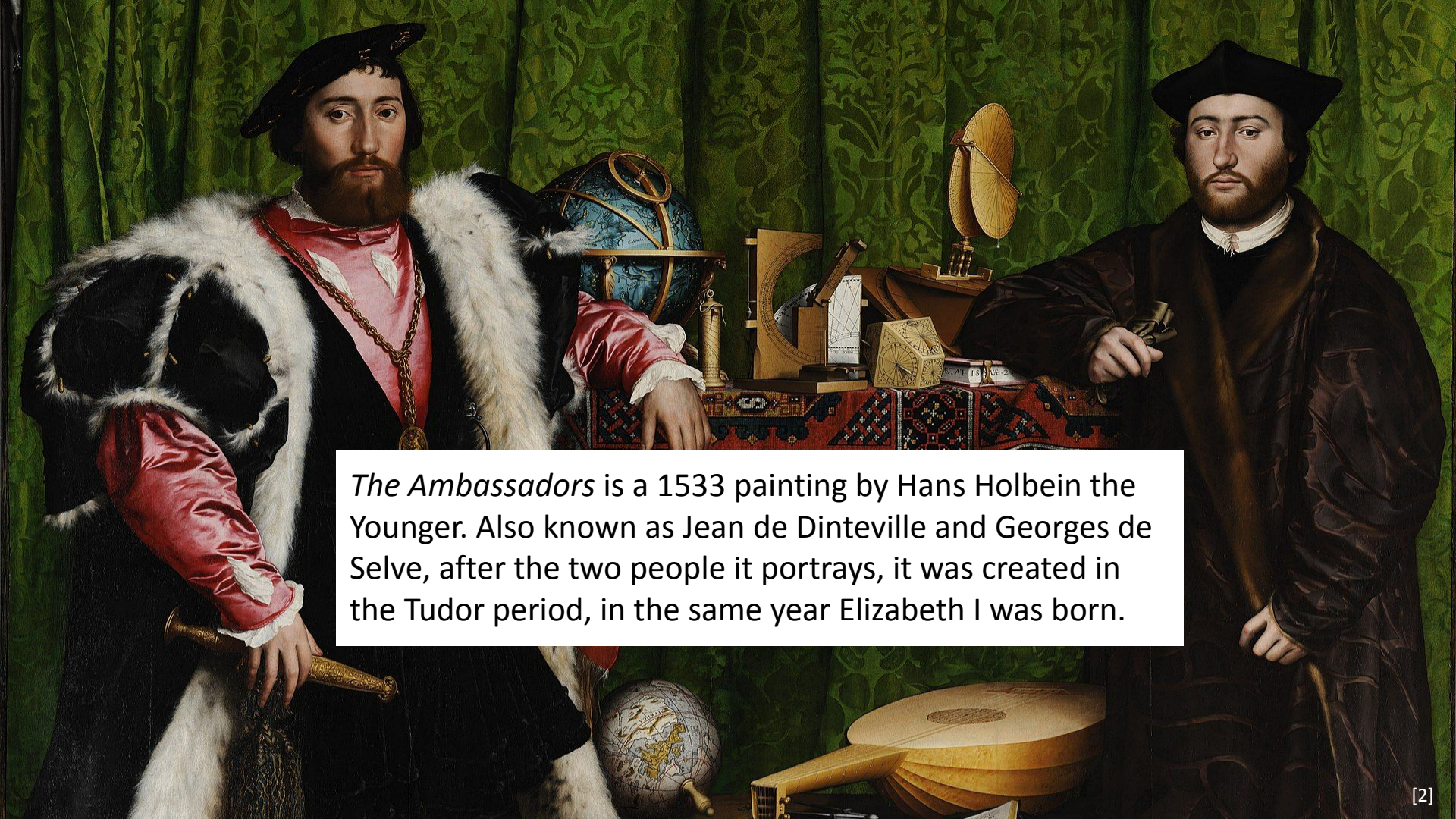
- Human Collaboration (manually)
- Automated Methods
 - from unstructured text sources
 - from structured data sources

Knowledge Graph Construction

from unstructured Text Sources

- The analysis of unstructured text requires techniques from **Natural Language Processing**
- NLP-based Knowledge Extraction:
 - Identify entities (**Entity Recognition**)
 - Map entities to a reference knowledge graph (**Entity Linking / Entity Resolution**)
 - Identify relations (**Relation Extraction**)
 - Map relations to a reference knowledge graph (**Property Linking**)
 - Further analysis to create meaningful triples:

POS Tagging, Constituency & Dependency Parsing, Co-Reference Resolution,
Semantic Role Labelling



The Ambassadors is a 1533 painting by Hans Holbein the Younger. Also known as Jean de Dinteville and Georges de Selve, after the two people it portrays, it was created in the Tudor period, in the same year Elizabeth I was born.

KG Construction from unstructured Text Sources

Part-of-Speech Tagging

— Text to annotate —

The Ambassadors is a 1533 painting by Hans Holbein the Younger. Also known as Jean de Dinteville and Georges de Selve, after the two people it portrays, it was created in the Tudor period, in the same year Elizabeth I was born.

— Annotations —

parts-of-speech x

— Language —

English

Submit

Part-of-Speech:

	DT	NNS	VBZ	DT	CD	NN	IN	NNP	NNP	DT	JJR																								
1	The	Ambassadors	is	a	1533	painting	by	Hans	Holbein	the	Younger	.																							
	RB	VBN	IN	NNP	NNP	NNP	CC	NNP	NNP	NNP	IN	DT	CD	NNS	PRP	VBZ	PRP	VBD	VBN	IN	DT	NNP	NN	IN	DT	JJ	NN	NNP	PRP	VBD	VBN	.			
2	Also	known	as	Jean	de	Dinteville	and	Georges	de	Selve	,	after	the	two	people	it	portrays	,	it	was	created	in	the	Tudor	period	,	in	the	same	year	Elizabeth	I	was	born	.

<http://corenlp.run/>

KG Construction from unstructured Text Sources

Named Entity Recognition

— Text to annotate —

The Ambassadors is a 1533 painting by Hans Holbein the Younger. Also known as Jean de Dinteville and Georges de Selve, after the two people it portrays, it was created in the Tudor period, in the same year Elizabeth I was born.

— Annotations —

named entities ✕

— Language —

English ▼

Submit

Named Entity Recognition:

DATE
1533

PERSON
Hans Holbein the Younger

1 The Ambassadors is a 1533 painting by Hans Holbein the Younger .

PERSON
Jean de Dinteville

PERSON
Georges de Selve

NUMBER
2.0

DATE
2023

PERSON
Elizabeth I

2 Also known as Jean de Dinteville and Georges de Selve , after the two people it portrays , it was created in the Tudor period , in the same year Elizabeth I was born .

<http://corenlp.run/>

KG Construction from unstructured Text Sources

Co-Reference Resolution

— Text to annotate —

The Ambassadors is a 1533 painting by Hans Holbein the Younger. Also known as Jean de Dinteville and Georges de Selve, after the two people it portrays, it was created in the Tudor period, in the same year Elizabeth I was born.

— Annotations —

coreference ✕

— Language —

English ▼

Submit

Speakers:

S(PERO)

1 The Ambassadors is a 1533 painting by Hans Holbein the Younger .

S(PERO)

2 Also known as Jean de Dinteville and Georges de Selve , after the two people it portrays , it was created in the Tudor period , in the same year Elizabeth I was born .

Coreference:

1 The Ambassadors is a 1533 painting by Hans Holbein the Younger .

2 Also known as Jean de Dinteville and Georges de Selve , after the two people CoreEntity11 it portrays , CoreEntity11 it was created in the Tudor period , in the same year Elizabeth I was born .

<http://corenlp.run/>

KG Construction from unstructured Text Sources

Semantic Role Labelling

— Text to annotate —

The Ambassadors is a 1533 painting by Hans Holbein the Younger. Also known as Jean de Dinteville and Georges de Selve, after the two people it portrays, it was created in the Tudor period, in the same year Elizabeth I was born.

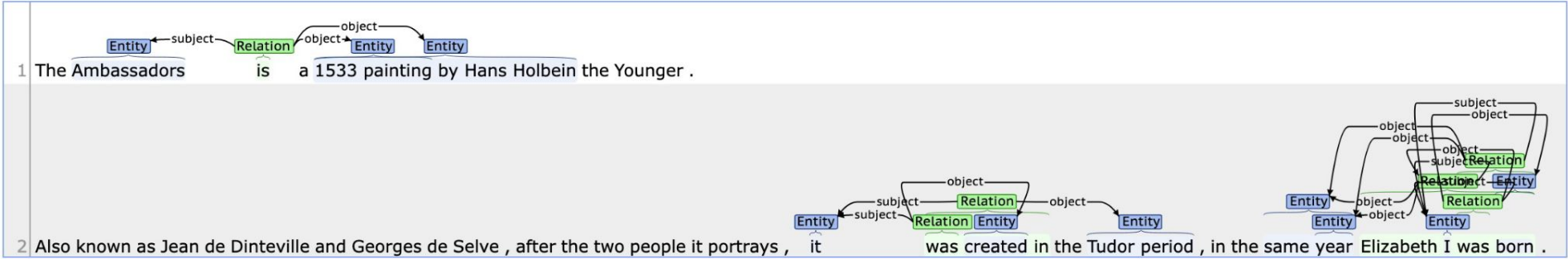
— Annotations —

openie ✕

— Language —
English ▼

Submit

Open IE:

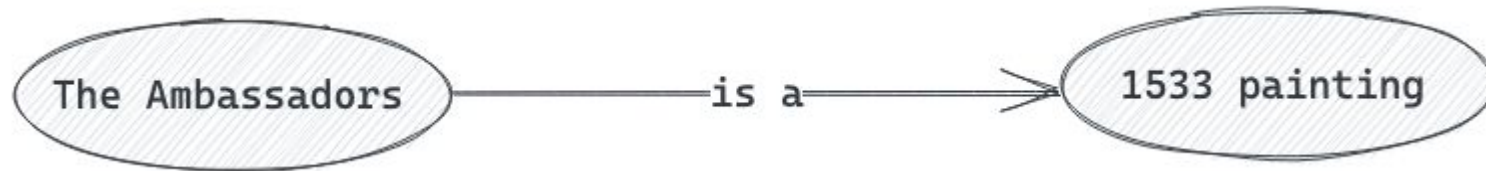


<http://corenlp.run/>

Knowledge Graph Construction

from unstructured Text

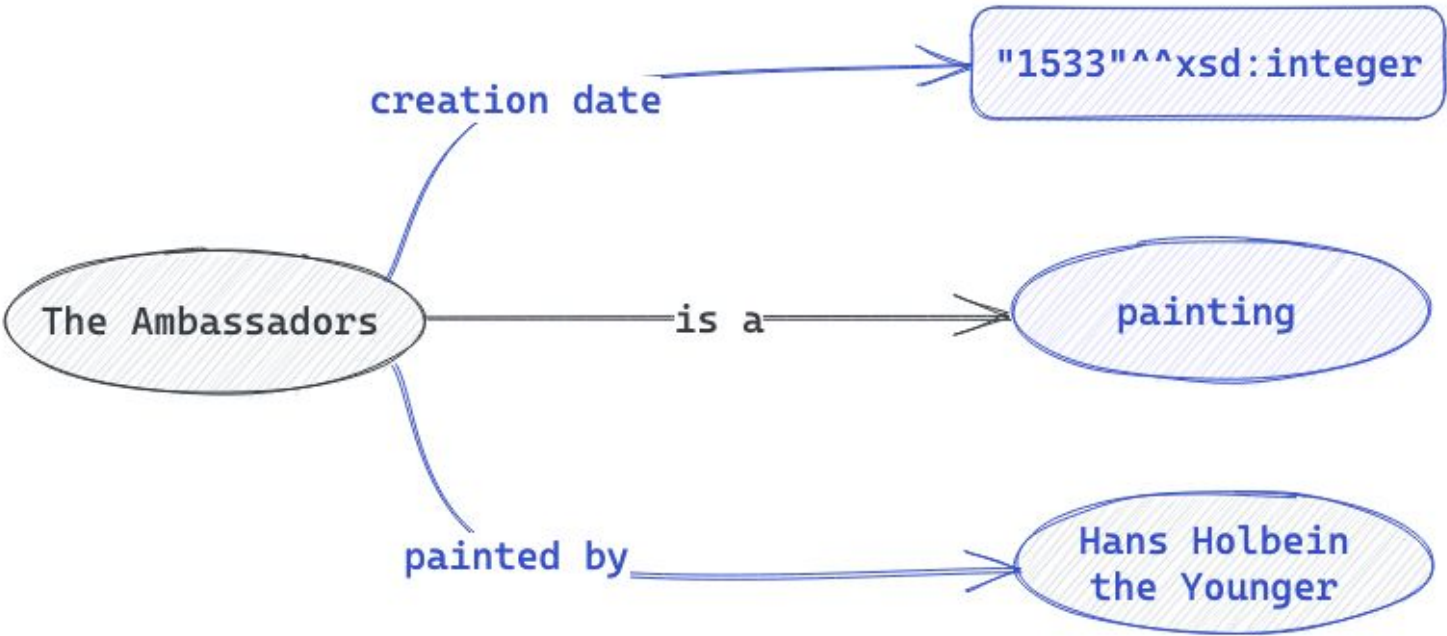
The Ambassadors is a *1533 painting* by *Hans Holbein the Younger*.



Knowledge Graph Construction

from unstructured Text

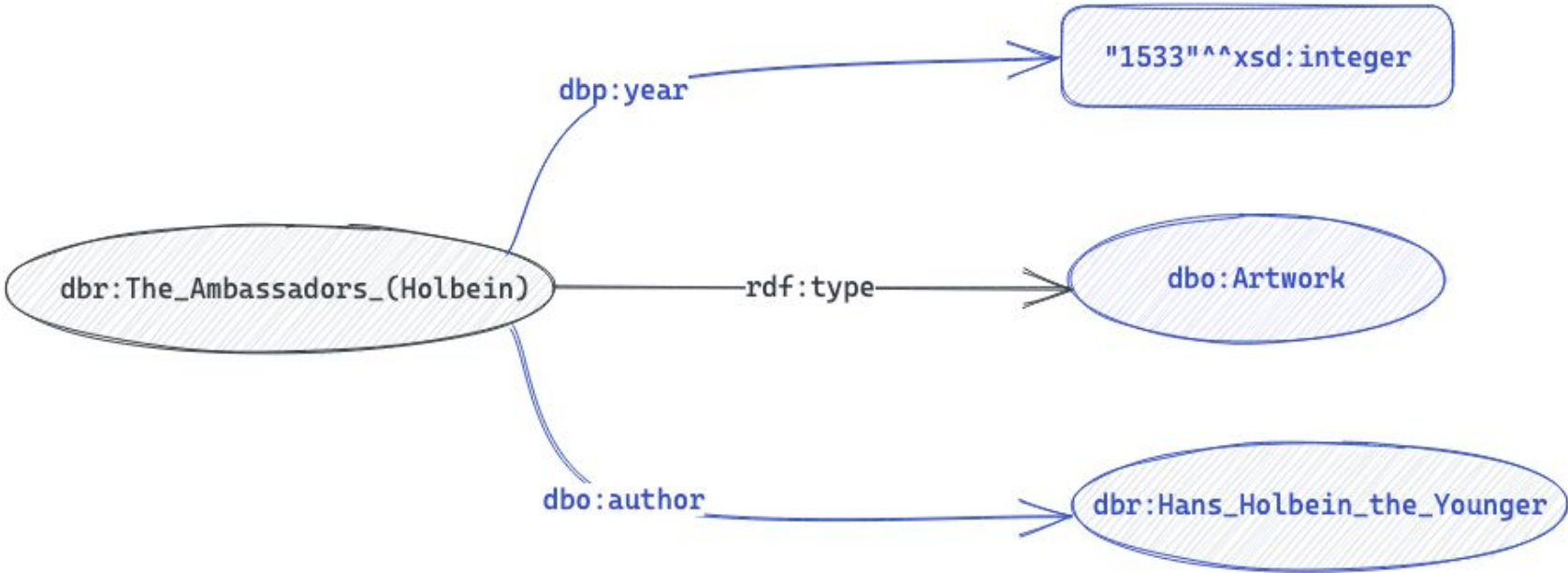
The Ambassadors is a **1533** painting by *Hans Holbein the Younger*.



Knowledge Graph Construction

from unstructured Text

The Ambassadors is a **1533** painting by *Hans Holbein the Younger*.



Knowledge Graph Construction



Create a knowledge graph from the following text:



The Ambassadors is a 1533 painting by Hans Holbein the Younger. Also known as Jean de Dinteville and Georges de Selve,[1] after the two people it portrays, it was created in the Tudor period, in the same year Elizabeth I was born.



The Ambassadors:



- Type: Painting
- Created by: Hans Holbein the Younger
- Created in: 1533
- Depicts: Jean de Dinteville and Georges de Selve
- Created during: Tudor period
- Year Elizabeth I was born: 1533

<https://chat.openai.com/chat>

Knowledge Graph Construction



<https://chat.openai.com/chat>

Create an RDF Turtle serialization of a knowledge graph from the following text:

The Ambassadors is a 1533 painting by Hans Holbein the Younger. Also known as Jean de Dinteville and Georges de Selve, after the two people it portrays, it was created in the Tudor period, in the same year Elizabeth I was born.



less

Copy code



```
@prefix ex: <http://example.org/> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .

ex:TheAmbassadors rdf:type ex:Painting ;
                  ex:wasCreatedBy ex:HansHolbeinTheYounger ;
                  ex:wasCreatedIn 1533 ;
                  ex:depicts ex:JeanDeDinteville , ex:GeorgesDeSelve ;
                  ex:wasCreatedDuring ex:TudorPeriod .

ex:ElizabethI rdf:type ex:Person ;
              ex:wasBornIn 1533 .
```


Knowledge Graph Construction



Could you please map the entities to wikidata identifiers?


<https://chat.openai.com/chat>



Sure! Here is the RDF Turtle serialization of the same knowledge graph with entities mapped to their corresponding Wikidata identifiers:



less

 Copy code

```
@prefix wd: <http://www.wikidata.org/entity/> .
@prefix wdt: <http://www.wikidata.org/prop/direct/> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .

wd:Q201812 rdfs:type wd:Q3305213 ;
wdt:P170 wd:Q159186 ;
wdt:P depicts wd:Q1060483 , wd:Q1060373 ;
wdt:P571 "1533-01-01T00:00:00Z"^^<http://www.w3.org/2001/XMLSchema#dateT
wdt:P1319 wd:Q11631 .
```

Knowledge Graph Construction

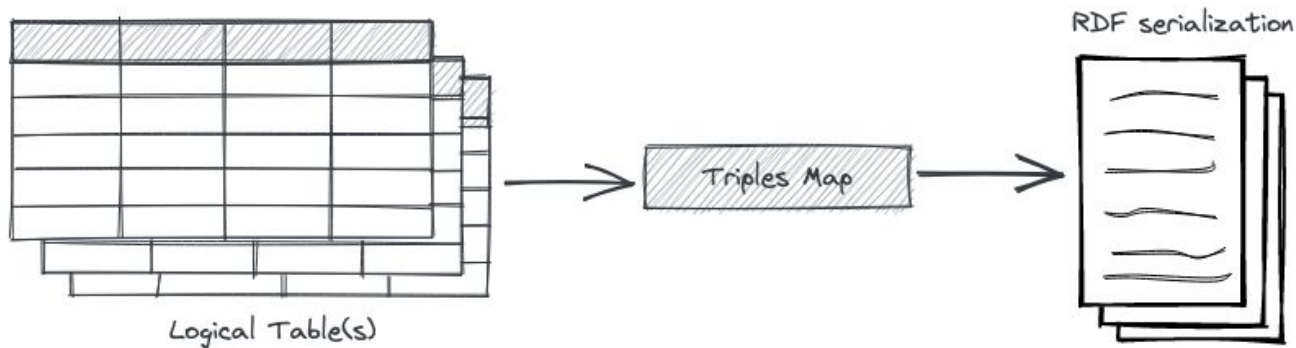
from structured Data Sources

- Structured data sources, e.g.,
 - Relational Databases, CSV files, etc.
 - Tree-based structured sources like XML, JSON, etc.
- Unlike **unstructured text**, where knowledge has to be **(imprecisely) extracted**, **structured data sources** often can be **(precisely) mapped** to a knowledge graph.
- 2-Step General Mapping:
 - (1) **Create a mapping** from the **source to a graph**.
 - (2) **Use the mapping** in order to **materialize the source data as a graph** or to virtualize the source (creating a graph view over the legacy data).

KG Construction from structured Data Sources

R2RML

- **R2RML** is a language for specifying mappings from relational data to RDF
- A **mapping** takes as input a **logical table**, i.e.,
 - a database table
 - a database view, or
 - an SQL query
- A logical table is mapped to a set of triples by a rule called **triples map**.



KG Construction from structured Data Sources

R2RML

- A **triples map** has two parts:
 - a **subject map**
 - several **predicate-object maps** (combining **predicate** and **object maps**)
- **Input** of a map:
 - a row of the logical table
- **Output** of a map: for each row,
 - a **subject resource** (IRI or blank node), often generated from primary key values
 - **several triples with the same subject**, but varying predicates and objects, generated from the attributes of the row

KG Construction from structured Data Sources

R2RML Simple Example

Paintings

ID <small>INTEGER Primary Key</small>	Title <small>VARCHAR(100)</small>	Creator <small>VARCHAR(100)</small>
123	The Ambassadors	Hans Holbein the Younger

```
@base <http://paintings.example.org/db/> .
@prefix rr: <http://www.w3.org/ns/r2rml#>.
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix dbo: <http://dbpedia.org/ontology/>.

<PaintingsMap> a rr:TriplesMap ;
  rr:logicalTable [ rr:tableName "Paintings" ];
  rr:subjectMap [
    rr:template "paintings.example.org/db/{ID}";
    rr:class dbo:Artwork;
  ];
  rr:predicateObjectMap [
    rr:predicate rdfs:label ;
    rr:objectMap [ rr:column "Title" ;
      rr:language "en" ];
  ];
  rr:predicateObjectMap [
    rr:predicate dbo:author ;
    rr:objectMap [ rr:column "Creator"
      rr:language "en" ];
  ];
].
```

```
@base <http://paintings.example.org/db/> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix dbo: <http://dbpedia.org/ontology/>.

:123 rdf:type dbo:ArtWork ;
  rdfs:label "The Ambassadors"@en ;
  dbo:author "Hans Holbein the Younger"@en .
```

```
<http://paintings.example.org/db/123>
<http://www.w3.org/1999/02/22-rdf-syntax-ns#type>
<http://dbpedia.org/ontology/ArtWork> .

<http://paintings.example.org/db/123>
<http://www.w3.org/2000/01/rdf-schema#label> "The
Ambassadors"@en ;

<http://paintings.example.org/db/123>
<http://dbpedia.org/ontology/author> "Hans Holbein the
Younger"@en .
```



Ontologies & Knowledge Graphs – Best Practices

Next Lecture...

Bibliographic References:

- Souripriya Das, Seema Sundara, Richard Cyganiak (eds.,2012), [R2RML: RDB to RDF Mapping Language](#), W3C Recommendation 27 September 2012.
- Aidan Hogan (2020), [The Web of Data](#), Springer.
Chap. 8.5.2 Relational Data to RDF, pp. 556–575.

Picture References:

- [1] “On this hyperrealistic scifi movie poster we see the scenery of Hans Holbein the Younger's famous painting "The Ambassadors" set into a retro futuristic environment on planet Mars showing countless small strange artifacts belonging to the ambassadors including a large distorted skull..”, created via ArtBot, Deliberate, 2023, [CC-BY-4.0], <https://tinybots.net/artbot>
- [2] “The Ambassadors”, Hans Holbein the Younger, 1533, [Public Domain], https://upload.wikimedia.org/wikipedia/commons/8/88/Hans_Holbein_the_Younger_-_The_Ambassadors_-_Google_Art_Project.jpg
- [3] “On this hyperrealistic scifi movie poster we see the scenery of Hans Holbein the Younger's famous painting "The Ambassadors" set into a postapocalyptic environment on planet Mars showing countless small strange and precious artifacts belonging to the ambassadors..”, created via ArtBot, Deliberate, 2023, [CC-BY-4.0], <https://tinybots.net/artbot>