KDI : Knowledge and Data Integration

Representation

Phase: 1. Introduction & Representation Diversity

W1.L12.M1.T2

Contents

Contents

How to represent the world?

To represent the world, we use ...

- Words
- Sounds
- Pictures
- ...

How to represent the world?

To represent the world, we use ...

- Words
- Sounds
- Pictures
- ...

Definition

$$representation = language + knowledge + data$$

Example

Starry Sky is an art work drawn by Vincent Willem van Gogh.

Contents

What is language?

Language is a tool for representing what is the case in the world

What is language?

Language is a tool for representing what is the case in the world

Definition

We take language as a set of *terms*, each term being associated a *meaning*, standing for what is the case in the world.

$$\textit{language} = \textit{terms} + \textit{meanings}$$

What is language?

Language is a tool for representing what is the case in the world

Definition

We take language as a set of *terms*, each term being associated a *meaning*, standing for what is the case in the world.

$$language = terms + meanings$$

Example (Term 'chair' and its meanings)

Noun

- S: (n) **chair** (a seat for one person, with a support for the back) *"he put his coat over the back of the chair and sat down"*
- S: (n) professorship, **chair** (the position of professor) *"he was awarded an endowed chair in economics"*
- S: (n) president, chairman, chairwoman, **chair**, chairperson (the officer who presides at the meetings of an organization) *"address your remarks to the chairperson"*
- S: (n) electric chair, **chair**, death chair, hot seat (an instrument of execution by electrocution; resembles an ordinary seat for one person) *"the murderer was sentenced to die in the chair"*
- S: (n) **chair** (a particular seat in an orchestra) *"he is second chair violin"*

Verb

- S: (v) **chair**, chairman (act or preside as chair, as of an academic department in a university) *"She chaired the department for many years"*
- S: (v) moderate, **chair**, lead (preside over) *"John moderated the discussion"*

Contents

What is knowledge?

Knowledge is a tool for stating what is the case in the world

What is knowledge?

Knowledge is a tool for stating what is the case in the world

Definition

We take Knowledge as a set of **entity types (etypes)**, each associated with a set of **properties**

$$knowledge = entity\ types + properties$$

What is knowledge?

Knowledge is a tool for stating what is the case in the world

Definition

We take Knowledge as a set of **entity types (etypes)**, each associated with a set of **properties**

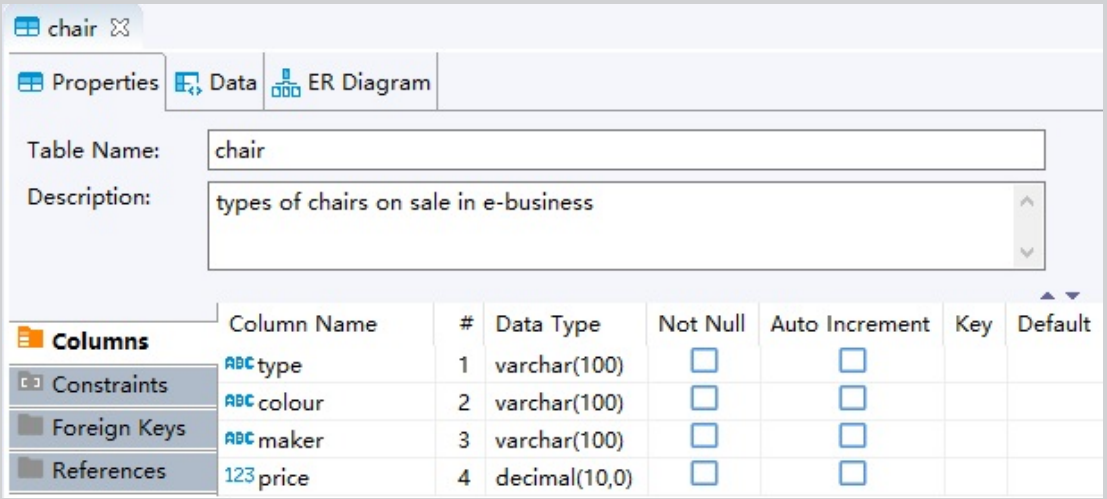
$$knowledge = entity\ types + properties$$

Think of etypes as universal statements about the world. They have two types of properties

- by Object property we mean a **relation** between entity types (e.g., all cities are near other cities).
- by Data property we mean an **attribute** describing some specific characteristic of the etype (e.g., all cities are inhabited by a certain number of people).

Example

Example (etype with data properties)



The entity type *chair* with the properties *type*, *colour*, *maker* and *price*.

Contents

What is data?

Data is knowledge about individuals.

What is data?

Data is knowledge about individuals.

Definition

We take data as a set of **entities**, each of a given etype, each associated with a set of **property values**

$$data = entities + property\ values$$

What is data?

Data is knowledge about individuals.

Definition

We take data as a set of **entities**, each of a given etype, each associated with a set of **property values**

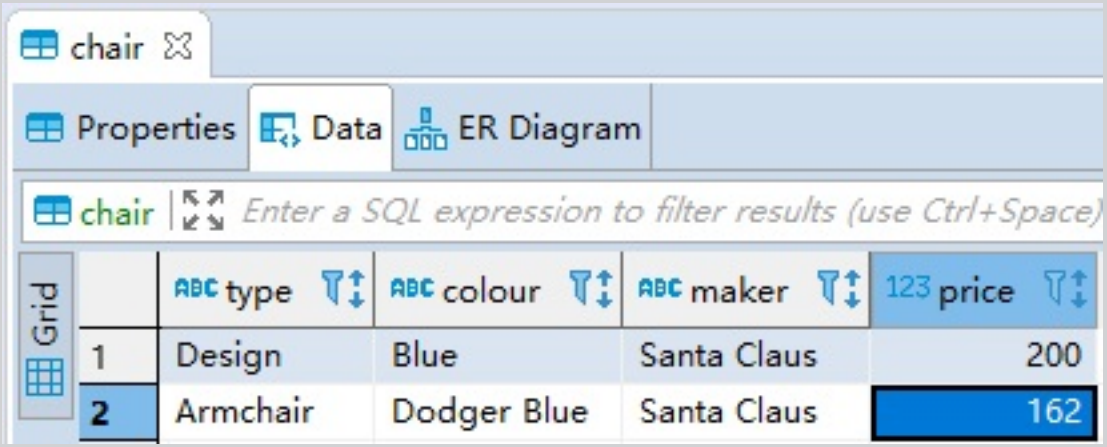
$$data = entities + property\ values$$

Think of entities as etypes where

- the etype consists of a single individual (e.g., the city of Trento)
- its object property values describe its relations with other entities (e.g., the city of Trento is near the city of Verona)
- its data property values, describe specific aspects of the entity (e.g., the population of the city of Trento is around 100,000 people).

Example

Example (Records in a legacy database)



The entity *#1* has property values: *Design* (of type), *Blue* (of colour), *Santa Claus* (of maker) and *200* (of price).

Types of Data

The direct example of a piece of data is a record in a table of a relational database. But there are various kinds of data in the world ...

- Unstructured Data
  Text without any metadata such as tags, relations, references ...
- Semi-structured Data
  Text with arbitrary defined tags such as html/xml files.
- Structured Data
  Data with structures such as records in (relational) databases and nodes and links in knowledge graphs.

Summary

A representation of the world is stated in a certain language, with respect to a certain universal knowledge, instantiated to specific data, where:

- language = terms + meanings
- knowledge = entity types + properties
- data = entities + property values

A Knowledge Base (KB) is a set of etypes.

A Database (DB) is a set of entities represented with respect to a certain KB.

In this course, we represent KBs and DBs as Knowledge Graphs (KGs).

KDI : Knowledge and Data Integration

W1.L12.M1.T2

Representation

Phase: 1. Introduction & Representation Diversity