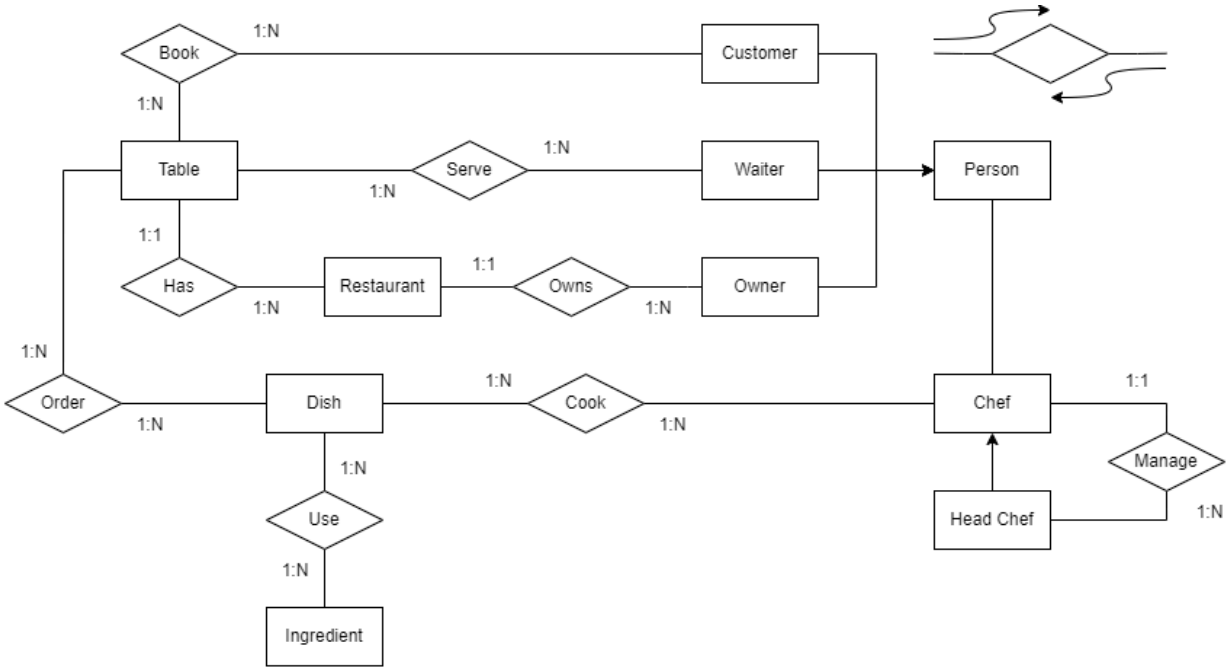
SYSTEMS AND METHODS FOR BIG AND UNSTRUCTURED DATA - Prof. Marco Brambilla - Sept 6, 2024

Last Name (Surname, Cognome) _____ First Name _____ Codice Persona _____

Consider the following ER Diagram



The following attributes describe the entities. The primary keys are underlined.

- **Person** - Personal_ID, Name, Surname, Birth Date, Phone Number
    - **Owner** - (...), Owner Since, E-mail
    - **Customer** - (...), E-mail
    - **Waiter** - (...), Hiring Date
    - **Chef** - (...), Hiring Date
        - **Head Chef** - (...), Head Chef Since
- **Restaurant** - Restaurant_ID, Name, Maximum Number of Guests, Address, Phone Number
- **Table** - Table_ID, Table Number, Max Seats
- **Dish** - Dish_ID, Name, Price, Recipe, Description
- **Ingredient** - Ingredient_ID, Name, Allergies (Yes/No)

The following attributes describe the relationships.

- **Book** - Number of People, Booking Date

N.B. There is only one Head Chef per Restaurant. Tables are served by more Waiters overall, but only one at a time. Dishes are cooked by more Chefs overall, but only one at a time.

Exercise 1 - Unstructured data models (1 PT)

Which DB technology would be the best to store the entities/rel. of the ER diagram? [you may not need all the rows of the table] (1 PT)

ENTITIES / REL.	DB TYPE	MOTIVATION

Exercise 2 - Neo4j (4.5 PT)

Consider the entities Customer, Table, Waiter, and Dish from the ER model and suppose you want to store the respective data instances in a graph database.
Sketch a graph model/example describing the nodes, main attributes, and edges.
Either show an example graph or a graph with types. (1 PT)

Write a Cypher query to collect the total number of different dishes that were served at each table, considering only tables whose max seats is greater than 10 and served by at least one waiter whose name is “Federico” (1.5 PT)

Write a Cypher query to collect the total number of customers served by each waiter, considering only customers who booked a table before 18/08/2024 and who ordered at least 5 dishes during one of their reservations. (2 PT)

Exercise 3 - MongoDB (6 PT)
Consider the Head Chef, Chef, Dish, and Ingredient. How many collections would you define? How would you implement the relations between the concepts? Provide a simple documental representation. **N.B.** A Head Chef is a Chef (1 PT)

Write a query to count the number of chefs managed by a chef named “Antonio” who became a head chef after 02/01/2023. Perform the query starting from the **Chef Collection or Sub-Collection**. (1 PT)

Write a query to collect the total amount of each ingredient that each chef named “Matteo” used to cook. Return only its Product_ID. Perform the query starting from the **Chef Collection or Sub-Collection**. (2 PT)

Write a query to count the number of chefs who cooked at least one dish named “Octopus with Potatoes”, which cost 25.00€. Perform the query starting from the **Chef Collection or Sub-Collection**. (2 PT)

Exercise 4 - Elasticsearch (4 PT)

Consider the Dish entity.

4.1. Provide the complete mapping of the index (i.e., field name, field type, the structure of the mapping, etc.) (1 PT)

PUT …

Write a query to return the list of all the dishes whose description and name include the word “salmon”, prioritising the one whose recipe includes the words “stir” or “boil”. (1.5 PT)

Write a query to return the list of all the dishes whose price exceeds 10.00€, whose name is “focaccia ligure”, and whose descriptions do not include the word “butter”. The condition on the description attribute must not affect the final score. (1.5 PT)

Exercise 5 - Cassandra (4 PT)
Consider the Chef table. Write a Cassandra script to perform the operations listed below.

Create the Chef table, with Personal_ID as the partition key and Phone Number as the clustering key, ordering the table based on the Phone Number. (1.5 PT)

Write a CQL query to collect the chef whose name is “Riccardo”. If any further operation is required, write its CQL code. (1 PT)

Create a Custom data type named “contact” that contains the name, surname, and phone number of the chef (1.5 PT)