



Co-funded by the
Erasmus+ Programme
of the European Union

QUESTION ANSWERING

Prof. Jesús Vilares

jesus.vilares@udc.es



UNIVERSIDADE DA CORUÑA



Introduction

- ▶ The goal of **Question Answering (QA)** is to return concrete answers to precise and arbitrary questions from the user:

“ Who is the French president ? ” Emmanuel Macron

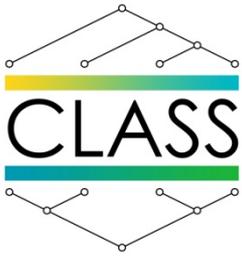
“ Where is the Louvre Museum ? ” In Paris, France

“ Which is the currency in China ? ” The yuan

- ▶ **Applications:** when the end user needs to know a specific piece of data without having to read the documents related to that search topic:

- Virtual assistants
- Online help systems
- Query interfaces for consulting technical manuals

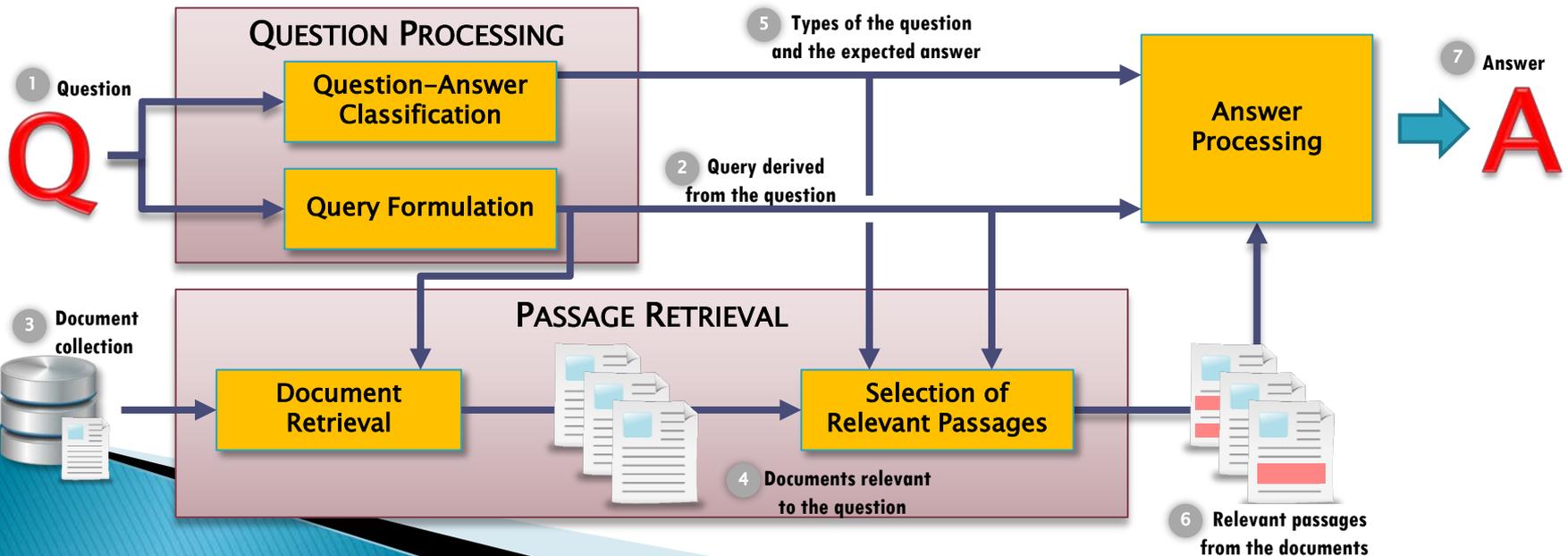




Architecture of a QA system

► Combines IR and IE techniques:

- IR systems: can locate documents related to the topic of the query but cannot extract the required information
- IE systems: can extract the required information but cannot process arbitrary queries (highly-specialized systems)





Architecture of a QA system

I. Question processing: processes the question to:

1. **Query formulation:** generates the query used to search the collection for documents relevant to the topic of the question (i.e. IR process)

Question: “ *Who is the French president ?* ”

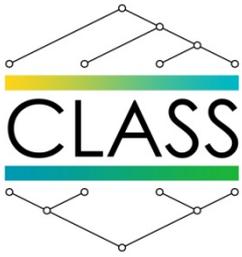
Query: president, french

2. **Question–Answer classification:** identifies the type of answer to be expected based on the type of the question

- E.g., in the case of *factoid questions* it refers to the **entity type** of the answer: person, location, date, etc.

Question: “ *Who is the French president ?* ”

Q-A type: person (PER)



Architecture of a QA system

II. Passage retrieval: locates those fragments (*passages*) of the documents of the collection which are relevant to the question topic. 2 phases:

- 1. Document retrieval:** by using the query previously obtained from the question, searches the collection for documents relevant to the question topic (i.e. IR process)
- 2. Selection of relevant passages:** processes the documents obtained in the previous phase to locate those fragments (*passages*) likely to contain the answer.

III. Answer Processing: extracts the answer from the candidate *passages* obtained. 2 approaches:

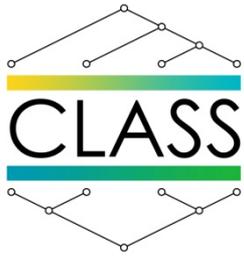
Question: “ *Who is the French president ?* ”

- 1. Returning the exact answer.**

Answer: Emmanuel Macron

- 2. Returning a text window from the document containing the answer (snippet).**

Answer: [the new French president, the centrist Emmanuel Macron]



Bibliography

- ▶ **[Jurafsky & Martin, 2009]** Jurafsky, D. & Martin, J.H. (2009). Chapter 23: Question Answering and Summarization. *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition (2nd ed.)*. Pearson–Prentice Hall.
- ▶ **[Pasça, 2003]** Pasça, M. (2003). *Open-domain question answering from large text collections*. CSLI Publications.