1. Describe your knowledge about knowledge graph.

A knowledge graph is a data structure that allows to contextualize entities, and organize correlations between entities or multiple type of entities. In this view, a knowledge graph is just a set of entities (Lionel Messi, Argentina, etc.), a set of relations between those entities (<plays_for>, <was_born_in>), and a set of facts. Facts are the combination of the previous two (<Messi, plays_for, Argentina>).

Entities are mostly persons, organizations, locations or products. These are usually called named entities because they refer to real world objects (physical or abstract) that bear a name. Barack Obama, Hawaii, Greece, Batman, or iPhone 7 are examples of named entities. Since named entities are ambiguous (multiple named entities share the same name), each named entity in the knowledge graph must be uniquely identified. The knowledge graph must be aware that they are different. For instance, consider the case of the US President Barack Obama, who has the same name as his father. The knowledge graph must then use a distinct and unique Id for each of them. For example, Barack_Obama_463 for Barack Obama and Barack_Obama_732 for Barack Obama.

Relations join entities together. One could think of relations as verbs or verbal phrases like <was_born_in>, <graduated_from>, <plays_for> or <acted_in> . Each relation must be unique, have a precise meaning, and a given scope, in the sense that they can only join specific classes of entities (<was_born_in> only involves persons and locations; <acted_in> only relates actors with movies, series or stage plays).

A fact is formed by joining entities through relations. For instance, <Obama, was_born_in, Hawaii> is a fact about entities Barack Obama and Hawaii, joined by the relation <was_born_in>, describing that the US president was born in the US state of Hawaii. Ultimately, the knowledge contained in a knowledge graph is represented by the set of facts it contains.

2. Why do we want to build such a knowledge graph?

knowledge graphs are the knowledge available to the machine, which in addition to some reasoning capabilities can be used by intelligent applications. In principle, a knowledge graph can be depicted as an oracle to whom a computer can ask anything. A nearby Restaurant, a list of US presidents, who is Angelina Jolie? Where was she born? The extent to which the application can answer these questions will depend on the quality of the knowledge graph and the type of information it contains.

The range of questions (usually referred as queries) that can be asked to a knowledge graph is broad. It can involve any combination of relations, entities, classes or facts. If the knowledge graph is relatively complete, it is guaranteed to provide high-quality answers in a minuscule amount of time.

3. What steps are required? Show your own workflow for this task.

To build a knowledge graph, you must first answer two questions required to build any graph:

1. What are the nodes? In a knowledge graph, they will be related to semantic concepts such as persons, entities, events etc.
2. What are the edges? They will be defined by relationships between nodes based on semantics.

Once you answered these two key questions, you can go to the next phase which is the data acquisition strategy. In addition to data acquisition, you will need knowledge processing where you come up with algorithms and heuristics to identify and extract the knowledge you are after from the data you acquire.

4. What are the challenges?

1) Noise in the source of publicly available information
2) Problems in dealing with information extraction methods
3)may contain many errors and inconsistencies in graph

5. Draw a knowledge graph for the given data.