BIG DATA

Big data is a generic term for large or complex datasets that are difficult to store and analyse. Big Data can be described in terms of:

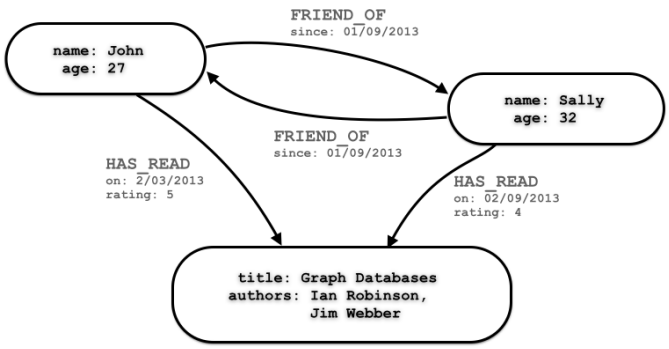
- volume – too big to fit into a single server
- velocity – streaming data, milliseconds to seconds to respond
- variety – data in many forms such as structured, unstructured, text, multimedia.

Modelling

Modelling: recreating a real-life situation on a computer

Graph schema: a method of defining a database in terms of nodes, edges and properties

- Node → entity
- Properties → information stored within each entity
- Edge → link and relationship between two nodes



Distributed processing

Distributed processing/computing: the principle of spreading large and complex tasks over a number of computers or servers

In imperative or object-orientated languages there are variables which are mutable (they change as the program runs) and so they have a state. This is said to have side effects (state of variable affects how the program runs) and so it is problematic when having distributed programming.

Instead functional programming, paradigm that uses functions to create programs, does not use mutable variables. Therefore functional programming is usually used to create distributed programs.