



**Quickly find the best way to remove wine stains**

**Get your robot to understand pancake recipes on the Web**



## A lot of unstructured know-how

Have you ever wondered how to do something and found the answer on the Web?

Popular websites collecting know-how:



> 1.000.000 know-how articles



> 150.000 know-how articles

Recipes, step-by-step instructions and do-it-yourself guides are among the most common examples of know-how (or Procedural Knowledge).

Know-how in digital form is often unstructured (hard for a machine to understand) and it can come in many different formats. The most common are:

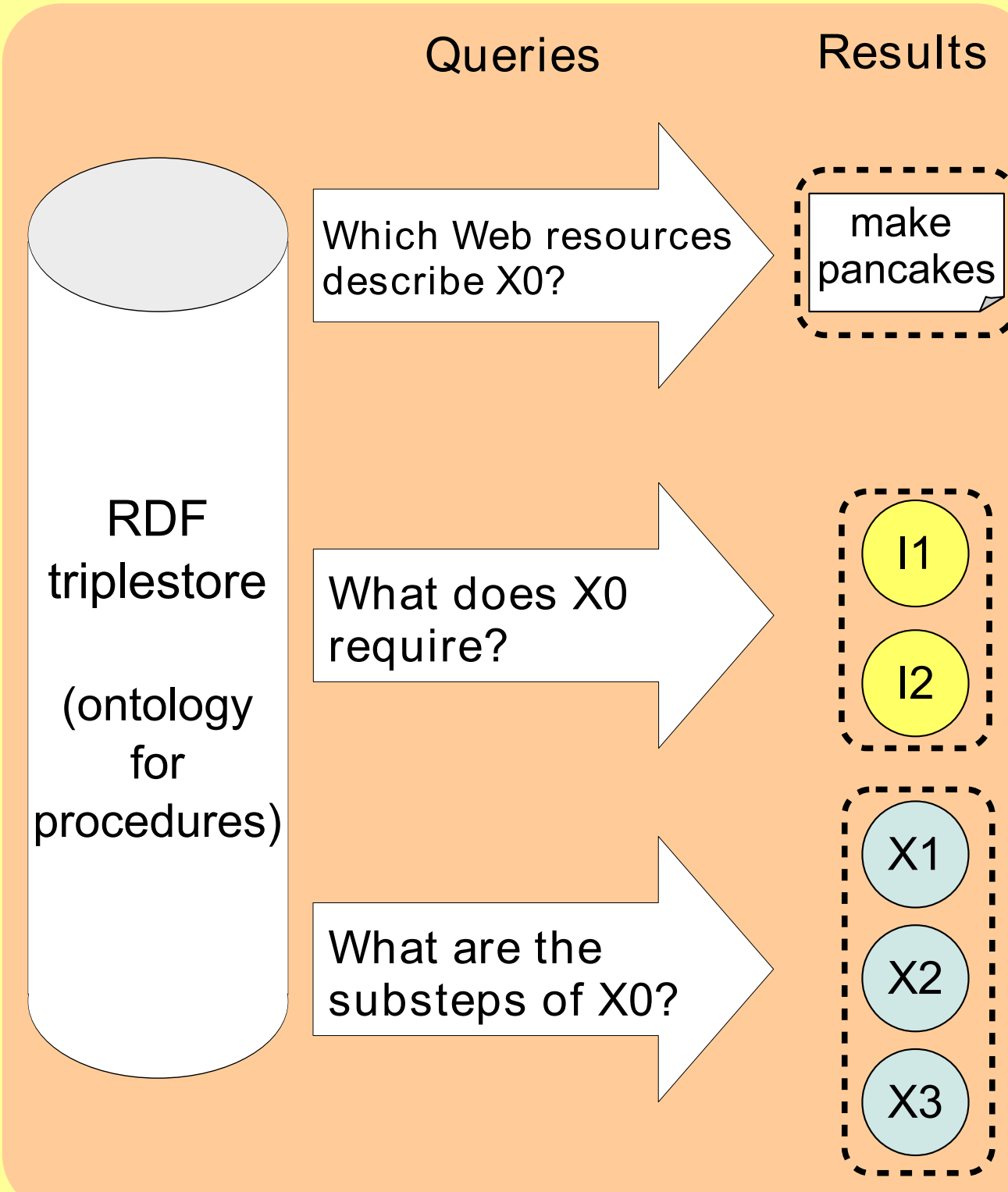
- text
- pictures
- video

Know-how can be:

- publicly available (e.g. articles on the Web)
- private (e.g. the standard operating procedures of an organisation)

## Querying the knowledge base: SPARQL

Software clients do not need external reasoners to interact with the procedural knowledge base, the SPARQL query language is enough



The SPARQL query language can be used both to **query** and to **update** the RDF triplestores containing procedural knowledge.

The diagram on the left shows some simple queries that can be formulated in SPARQL:

- Which Web resources describe X?
- What does X require?
- Which are the substeps of X?

SPARQL can also formulate more complex queries such as:

- Given the set of steps which I have already completed to achieve X, which is the step (or steps) that I can do next?
- Which are all the possible ways of achieving X?

## How can the Semantic Web help?

### Search

In the most simple scenario, a search for know-how can be improved by **limiting the results** to those that describe procedures. But a semantic representation of procedures can also be used to exclude processes that **require** a tool you don't have or that would take **too much time** according to your preferences.

### Compose

Discovering the **connections** between different procedures is not trivial. For example, how can you discover all the possible ways to get a European visa? On the Semantic Web it would be possible to create a **network of procedures** which can be used to **discover** related know-how and to **compose** it in **workflows**.

### Assist

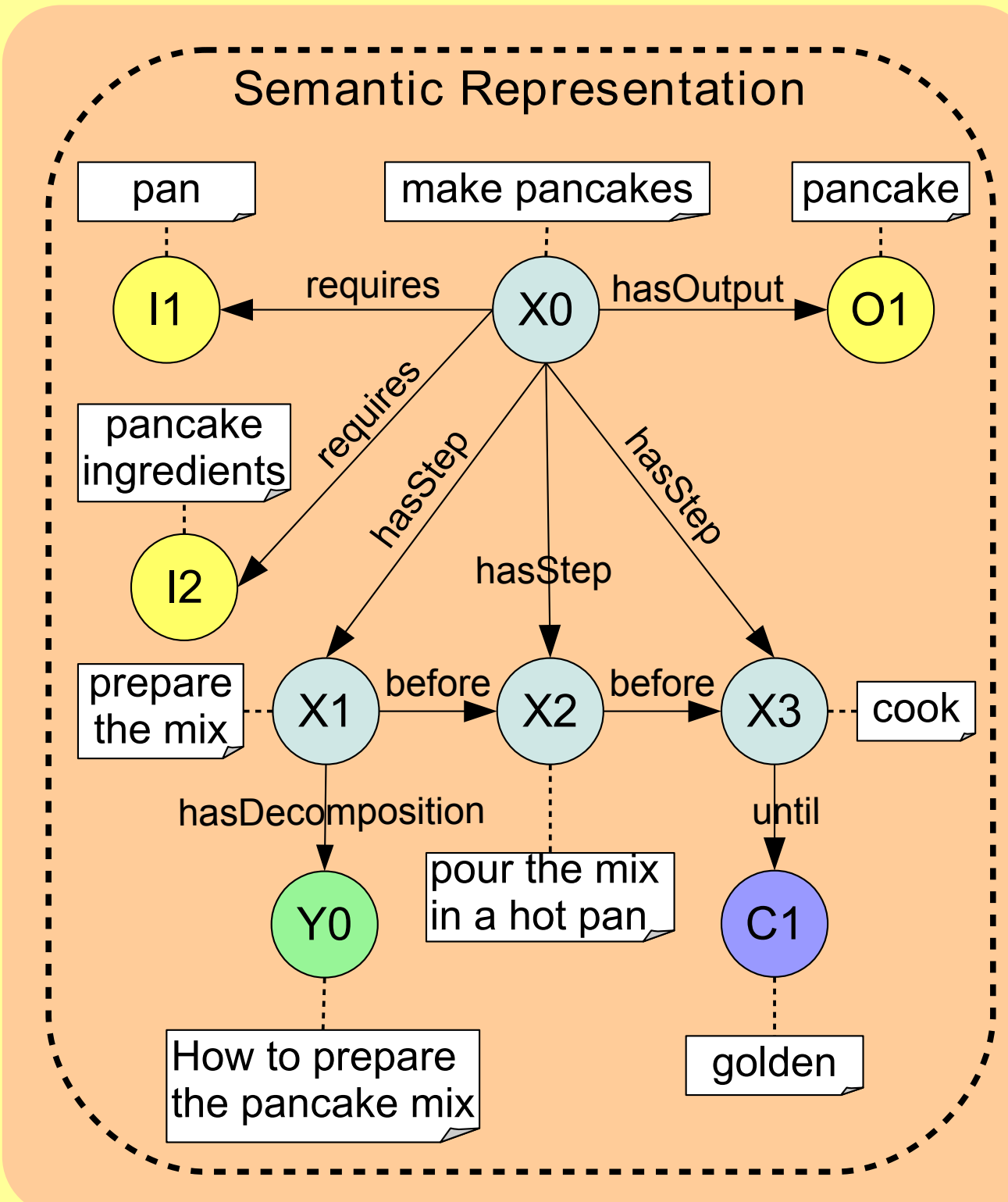
A system that understands know-how could recognize what activity you are trying to achieve (**Activity Recognition**) and assist you in completing the remaining steps. If you are trying to reach a certain destination, for example, it could give you directions on how to reach it.

### Automate

If a process is described with a sufficient level of detail, a system might be able to automate it (or automate some of its steps). For example it could be automated by a **Web Service** or maybe even by a **robotic agent**.

## An RDFS ontology for procedures

A simple vocabulary to represent the most common concepts and relations that occur in procedures



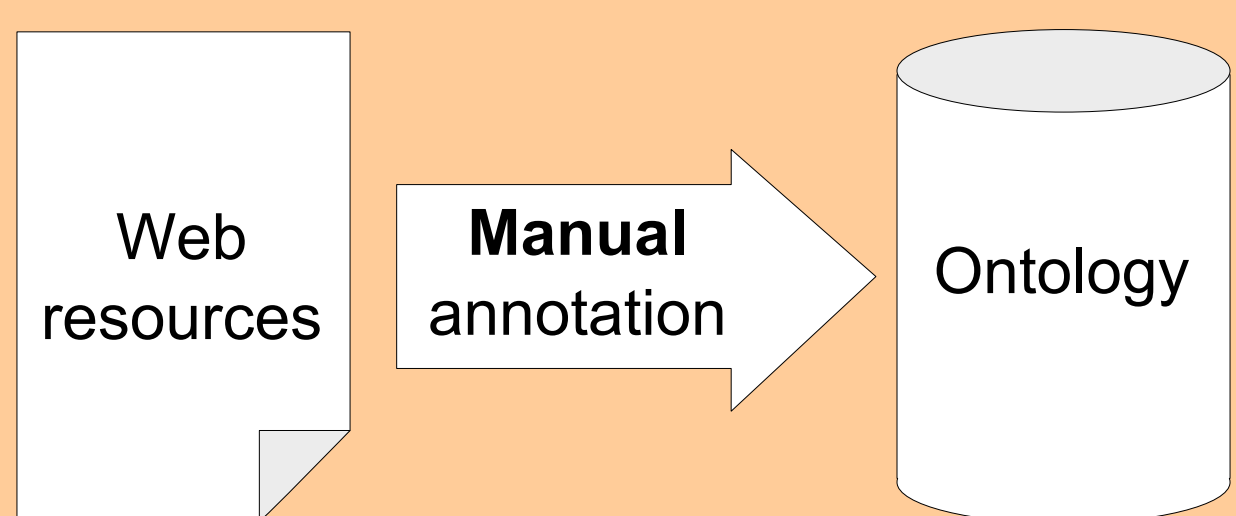
A procedure can be represented semantically at different levels of detail. The diagram on the left shows a more detailed representation of the procedure "How to make pancakes". This representation, for example, explicitly states that the procedure **requires** the "pan" object and that the step "pour the mix in a hot pan" can be done only **after** the mix has been prepared.

Some of the most common components of procedures are the following:

- Decomposition of a procedure into substeps
- Inputs and outputs
- Preconditions and effects
- Workflow operators such as the order in which the steps should be performed (e.g. the "before" relation) or loops (e.g. the "until" relation)

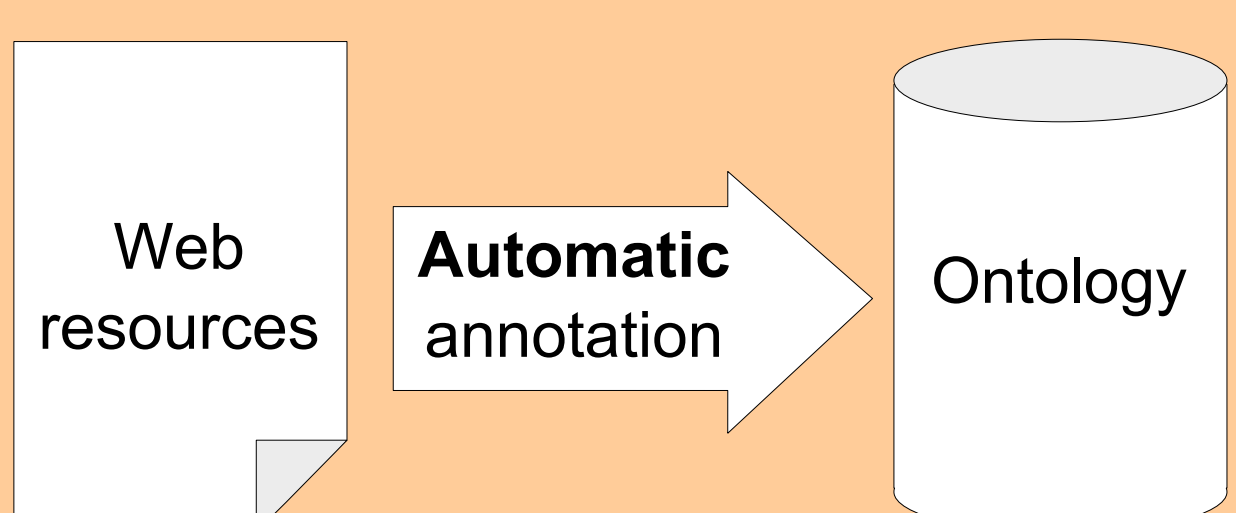
## Knowledge Extraction

Manual and Automatic annotations of unstructured know-how



### Manual annotation

- can be done using **annotation tools**
- can be made more scalable by adopting **Social Computation** strategies (e.g. engaging user communities)
- can help the development of automatic annotation strategies

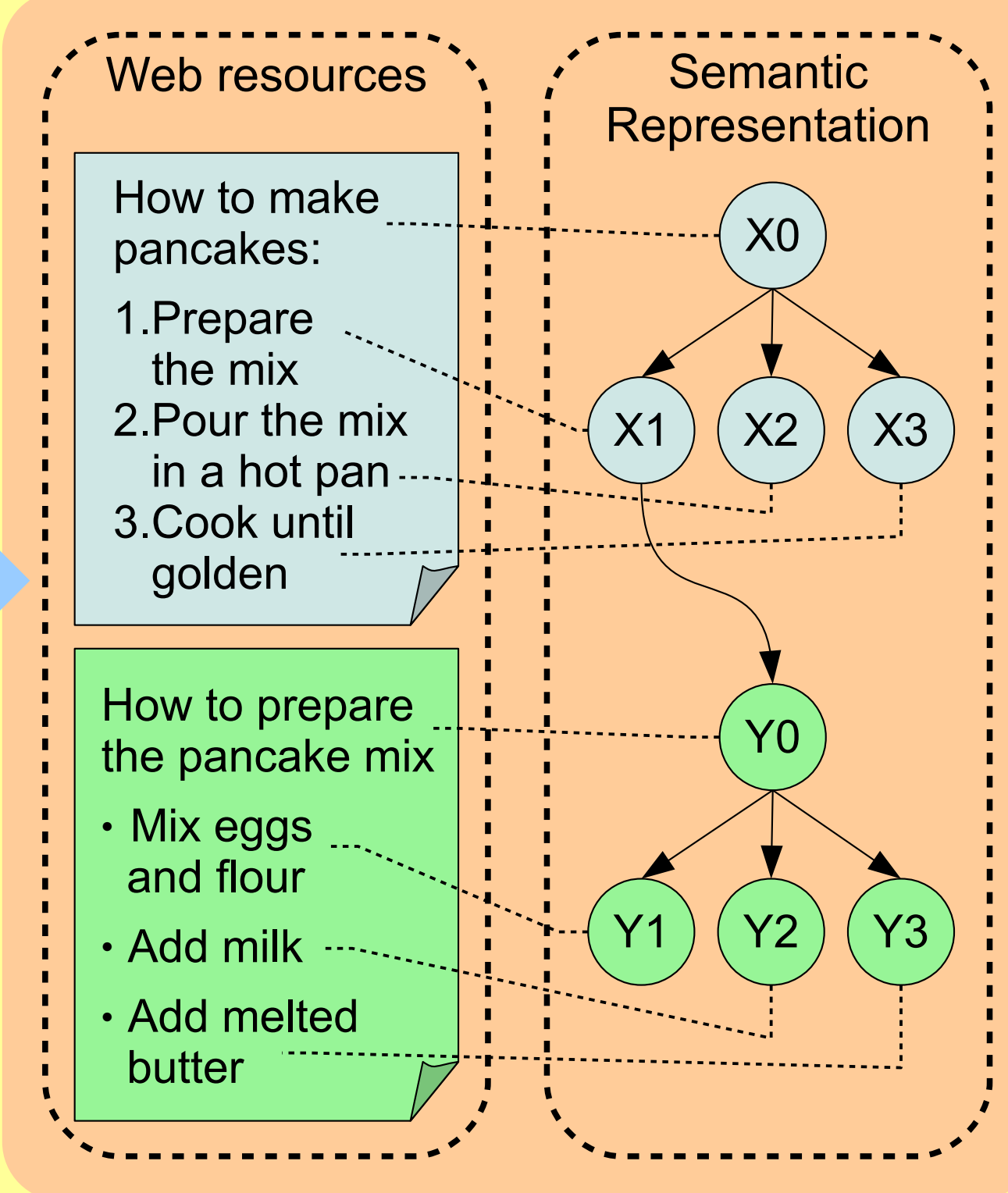


### Automatic annotation

- for textual resources, this can be done using **Natural Language Processing (NLP)**
- Several attempts have been made to extract procedural knowledge automatically. The results obtained are promising, but still limited in accuracy.

## Knowledge Representation in RDF

A representation of procedures based on triples



• The resources describing a process (e.g. the text where the instructions are written) can be connected to the semantic entities that represent them. This can be done using the Open Annotation Data Model.

In this example, the entity X0 is connected to the text fragment "How to make pancakes".

• Different parts of a process can be annotated as different entities.

In this example, the entities X1, X2 and X3 represent the three substeps of making a pancake described in the blue procedure.

• Entities can be related to each other

In this example, the blue procedure is connected to the green procedure by a relation between entities X1 and Y0. This relation could mean that process Y0 ("How to prepare the pancake mix") describes a possible way of achieving the substep X1 ("Prepare the mix").