KGE - Knowledge Graph Engineering

iTelos Methodology
General structure

Fausto Giunchiglia

Contents

iTelos - Structural features

- iTelos is a **phase-based** methodology.

- The execution of the different phases, **following the right order**, allows the implementation of a KGE process.

- Each phase is formed by one or more **specific activities** with the objective to solve the several KGE issues.

- The iTelos's phases are **standalone modules** which composed together form the methodology. Nevertheless, they can be **executed singularly or in composition** with one (or more) of the others.
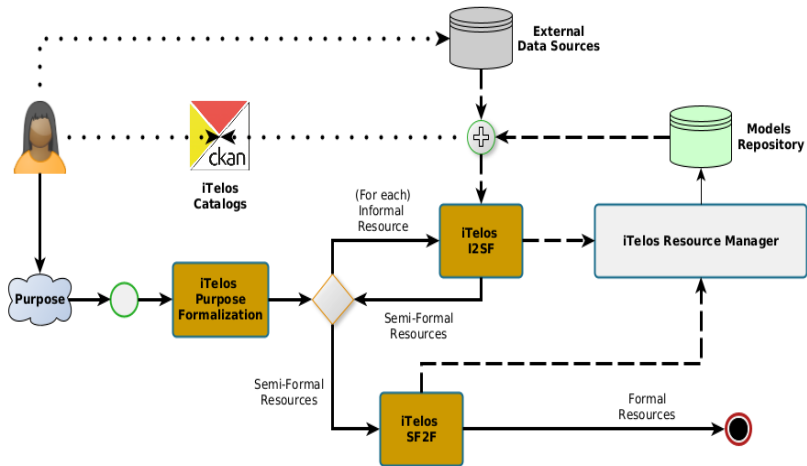
iTelos - Input & Output

- One more time ...

- **iTelos inputs**:
  - (Functional requirements) The Purpose.
  - (Non-Functional requirements) A list of data sources (optional).

- **iTelos outputs**:
  - (Explicit goal) The final (Purpose specific) KG.
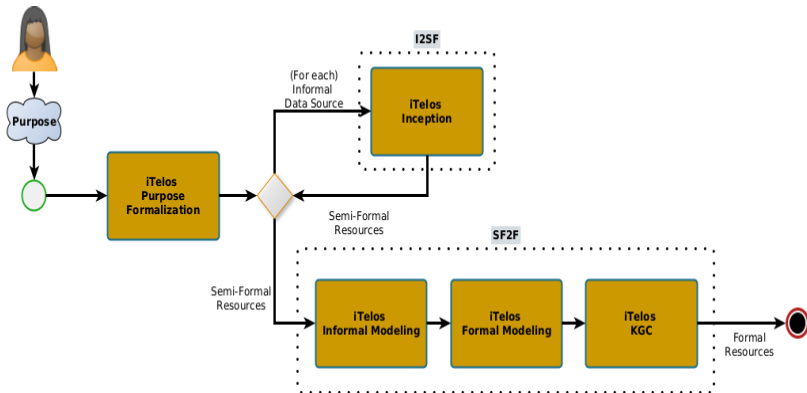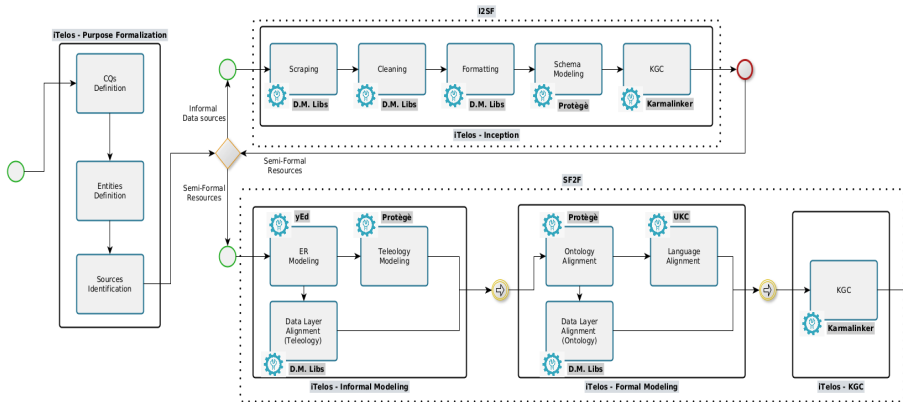 - (Implicit goal) A set of reusable resources.

Contents

iTelos top level view

iTelos top level view

iTelos top level view

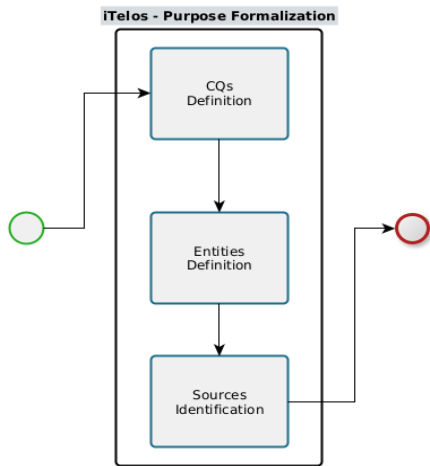**Note**: The Knowdive group developed a prototype tool able to lead the user along the different methodology's phases

Contents

Purpose Formalization

- **Input**: a natural language sentence representing the user's Purpose (plus optionally a list of already identified data sources).

- **Output**: a set of document in which Purpose's details are extracted and formalized (a first formalization step).

- **Objective**: to make explicit (in a more formal way) the functional requirements implicitly included in the input Purpose.

# Purpose Formalization - Note

- In the Purpose Formalization phase, iTelos aims to extract all the possible information from the Purpose expressed by the user.

- A key aspect to understand is, which are the **Common**, **Core** and **Contextual** elements considered by the Purpose.

- iTleos provides dedicated activities with the objective to classify the information into the popularity categories.

Contents

Inception
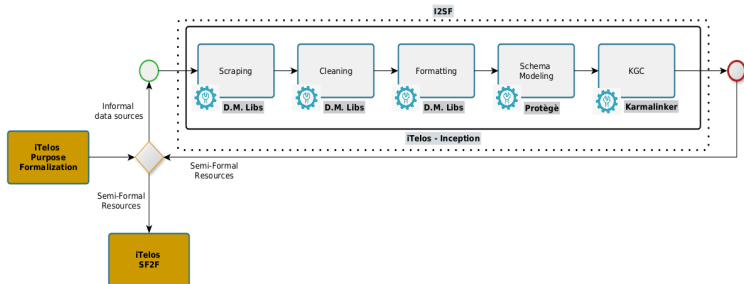
- **Input**: a set of data sources identified previously, plus the initial user's Purpose.

- **Output**: a set of semi-formal resources, created from the informal resources extracted by the data sources in input.

- **Objective**: to extract, clean, format and model the informal resources required to satisfy the Purpose, collected from the input data sources.

Inception - Note

- The Inception phase includes most of the effort in **data management**.

- The different activities, included in this phase, are executed following a specific order, with the objective **to create a semi-formal version of the resources** collected form the input data sources.

- Noticed how the **Purpose is fundamental** in order to properly built the semi-formal resources suitable for the project.

- The Inception phase is executed **iteratively** over each data source received in input. The internal activities will be in turn executed iteratively over each resource to be handled.

Contents

Informal Modeling

iTelos - Informal Modeling
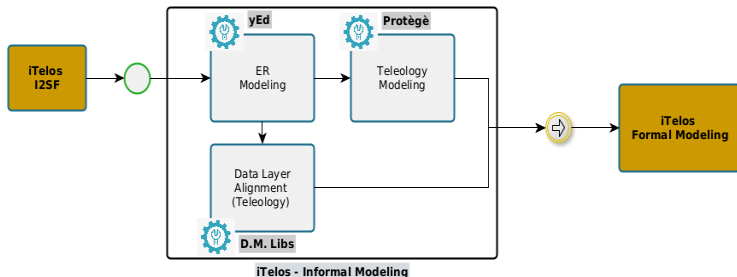
- **Input**: a set of semi-formal resources representing the information models that should be composed to build the final KG. Plus the Purpose.

- **Output**: an ER model plus the relative semi-formal purpose teleology. Additionally, the semi-formal datasets aligned with the teleology produced.

- **Objective**: to do a step forward the final KG's knowledge structure. In parallel to use the purpose model to align the dataset collected.

Informal Modeling - Note

- The Informal Modeling phase is the first macro activity in the SF2F iTelos macro phase.

- **Remember**: the SF2F macro phase, unlike the I2SF, aims at compose together the semi-formal resources (models) produced before. Therefore, the focus of the iTelos phases within SF2F is over the whole set of resources (not over each single ones).

- Noticed how the KG structure, divided in **Knowledge and data layer**, is more clearly reported starting from this phase, where specific activities are executed in parallel over the two different layers.

Contents

Formal Modeling

iTelos - Formal Modeling
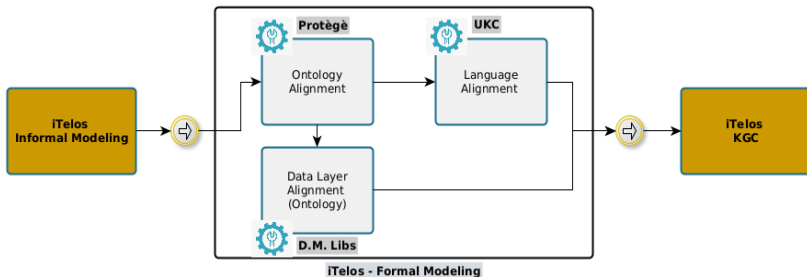
- **Input**: the teleology, representing the formalized Purpose, plus the semi-formal datasets aligned with.

- **Output**: the formal version of the dataset aligned to the purpose, plus the final KG's knowledge structure, the Entity Type Graph (ETG); plus the formal version of data resources.

- **Objective**: to fully formalize the semi-formal datasets, and produce a reusable knowledge structure for the final KG.
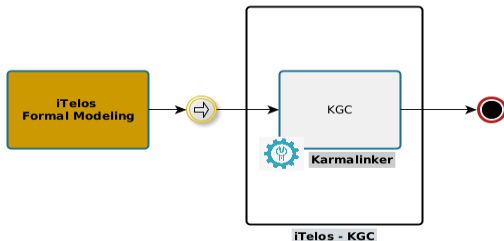
Formal Modeling - Note

- There are two key aspects in this phase:

    - The **reusability** implemented by grounding the teleology into existing reference ontologies.

    - The **(multilingual) interoperabiltiy** of the resources handled in this phase is enhanced by the language alignment.

- Noticed how the two key aspects above are fundamental in the SF2F conversion.

Contents

KGC

- **Input**: the final KG's ETG, plus the set of formal data resources.

- **Output**: the final KG.

- **Objective**: to merge the knowledge layer with data layer into a single exploitable resource.

KGC - Note

- The last iTelos phase aims to **compose all the information models** produced in the previous phases.

- In this phase, the KG knowledge and data layer constructed separately are **merged** together into a single resource.

- The final KG, as result of the KGE process, can be exploited as single resource, or it can be decomposed into the several models, thus exploiting them singularly.

General iTelos Notes

- Notice how each phase is a single module, solving a specific problem:

    - **Purpose Formalization**: it extracts the information requirements fro the purpose.

    - **Inception**: data collection and (semi) formalization

    - **Informal Modeling**: it defines a precise purpose's teleology (formal model of the purpose within a precise context).

    - **Formal Modeling**: it creates a reusable ETG.

    - **KGC**: it produces a single purpose-specific KG.

- Moreover, each phase produces a different kind of model relative to the same purpose, which needs to be properly stored for future reuse.

- To this end, each phase is responsible of producing **metadata** describing the relative information models.

Fausto Giunchiglia

iTelos Methodology
General structure