



KNOWDIVE



KDI • **Knowledge and Data Integration**

Schema Alignment & ETG Generation

iTelos Formal Modeling Phase

Fausto Giunchiglia, Mayukh Bagchi

Contents

- 1 Introduction**
- 2 FT (*as is*)
- 3 ER (*as is*)
- 4 FT - ER (Schema) Alignment
- 5 ETG Generation
- 6 Summary

What We Have

- 1 The foundational teleology (FT) with the foundational primitives - Objects, Functions and Actions, and the foundational relations.
- 2 An ETG Model wherein the concepts are *uniquely* and *formally* expressed using GID and aligned to UKC KB via Language Alignment step

Where is the Gap? The ETG Model is till now ***formally separate*** from the Foundational Teleology (FT), ***without any formal alignment*** between concepts and relations as expressed in the ETG Model with their foundational counterparts.

Schema Alignment

- The schema alignment activity aims to exactly close the gap between the ETG model and the foundational teleology
- It formally grounds the ETG model in the foundational teleology, thus producing the final, fully formal ETG ready for the next phase - data integration
- Schema alignment involves aligning the objects, functions, actions and relations in the ETG model to their semantically corresponding foundational primitives via *intermediate concepts*
- The ETG created following the above process is (i) *shareable* as it is grounded in the foundational teleology, and (ii) *reusable* as a consequence of the domain compositionality principle of *iTelos*.

Contents

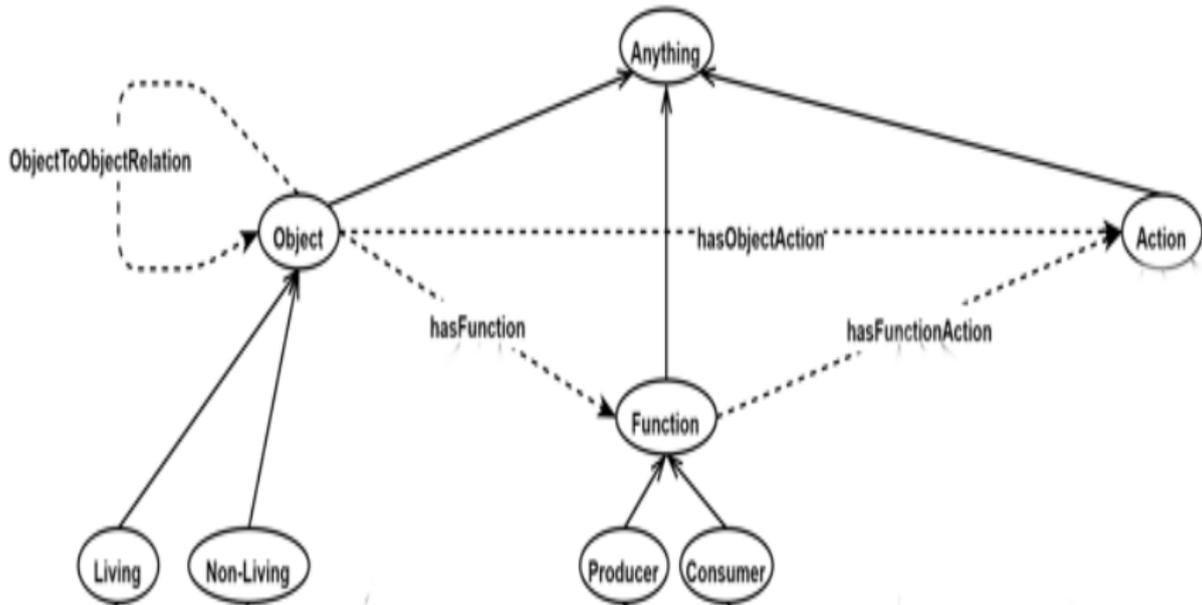
- 1 Introduction
- 2 FT (*as is*)**
- 3 ER (*as is*)
- 4 FT - ER (Schema) Alignment
- 5 ETG Generation
- 6 Summary

Foundational Teleology (FT) - Recap

- Our *foundational teleology* (FT) models *causality* via the following *foundational primitives*:
 - Objects (Living and Non-Living)
 - Functions (Producers and Consumers)
 - Actions
 - Space and Time (*a priori*)
- When we populate the foundational teleology (reflected in the UKC top-level) with *domain-specific concepts*, we get the *Entity Type Graph (ETG)* for a specific domain
- The *ETG* is the *design basis* on which *Entity Graphs (EGs)* are designed and founded, modelling domain-specific ground truth

Foundational Teleology - Diagram

The foundational teleology is visualized as -



Contents

- 1 Introduction
- 2 FT (*as is*)
- 3 ER (*as is*)**
- 4 FT - ER (Schema) Alignment
- 5 ETG Generation
- 6 Summary

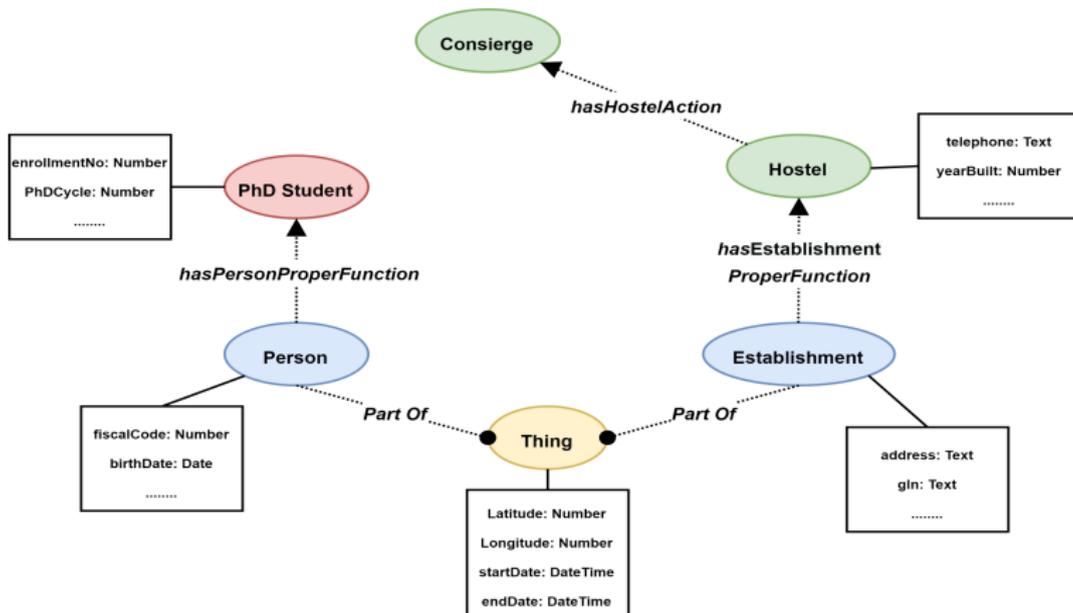
Modelling ER - Recap

The ER model is designed by performing the following (flexible) steps -

- 1 Specify the reference context, *viz. Thing*, in the context of our ETG
- 2 Instantiate the *Object Partonomy* with respect to *Thing*
- 3 Relate each *object* to its (proper) *function(s)*
- 4 Relate each *function* to its (admissible) *action(s)*
- 5 Relate each *object* to its (admissible) *action(s)*

NOTE: Specify properties and attributes (including data types) for each individual concept in (2), (3), (4) and (5).

ER Model - Informal Visualization



Legend



Contents

- 1 Introduction
- 2 FT (*as is*)
- 3 ER (*as is*)
- 4 FT - ER (Schema) Alignment**
- 5 ETG Generation
- 6 Summary

Methodology

The methodology for the schema alignment is as follows -

- 1 Align the objects in the ER model to the foundational distinction of *living* and *non-living* within the category *Object* in the FT
- 2 Align the functions in the ER model to the proper foundational distinctions of *producer* and *consumer* within the category of *Function* in the FT
- 3 Align the actions in the ER model to the foundational distinction of *Actions* in the FT
- 4 Align the relations in the ER model to the *foundational relations*

NOTE: The above methodology is completely flexible w.r.t project-specific modelling requirements. The alignment is dependent upon the key notion of linking via *intermediate concepts*

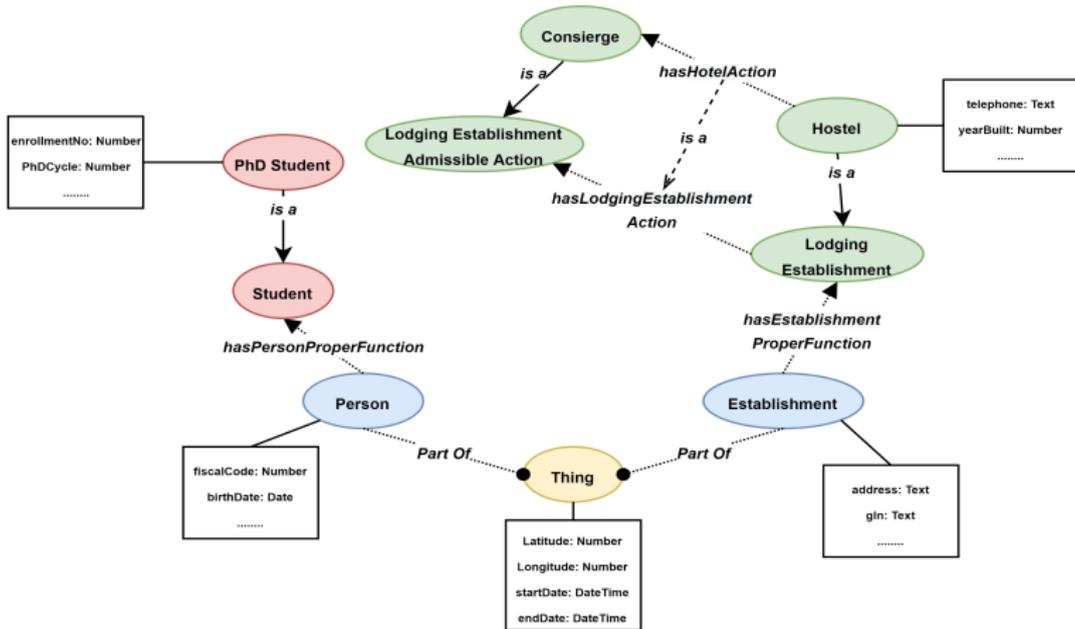
Modelling EER

- In formal rules of knowledge classification, the *principle of modulation* states that a hierarchy should comprise one concept “*of each and every order that lies between the orders of the first link and the last link*” of the hierarchy.
- These concepts are known as *intermediate concepts*. For example, in the hierarchy, World - Europe - Italy, *Europe* is the intermediate concept
- In other words, the principle asserts that a hierarchy shouldn't have any *missing link* relevant to the modelling purpose
- In fact, the *intermediate concepts* are *key and mandatory* for grounding the ETG Model in the foundational teleology, generating the Extended ER (EER) model

Modelling EER - Examples

- *Student* should be the primitive function of the object *Person*, not *PhD Student* as it is more specialized
- Similarly, *LodgingEstablishment* should be the primitive function of the object *Establishment*, not directly *Hostel* as it is more specialized
- Similarly, *LodgingEstablishmentAdmissibleAction* should be the intermediate concept linking the function *LodgingEstablishment* with its actions such as *Consierge*
- Each of these concepts are *intermediate concepts* which grounds the ETG Model in the foundational teleology (informal illustration in next slide)

Modelling EER - Visualization



Legend

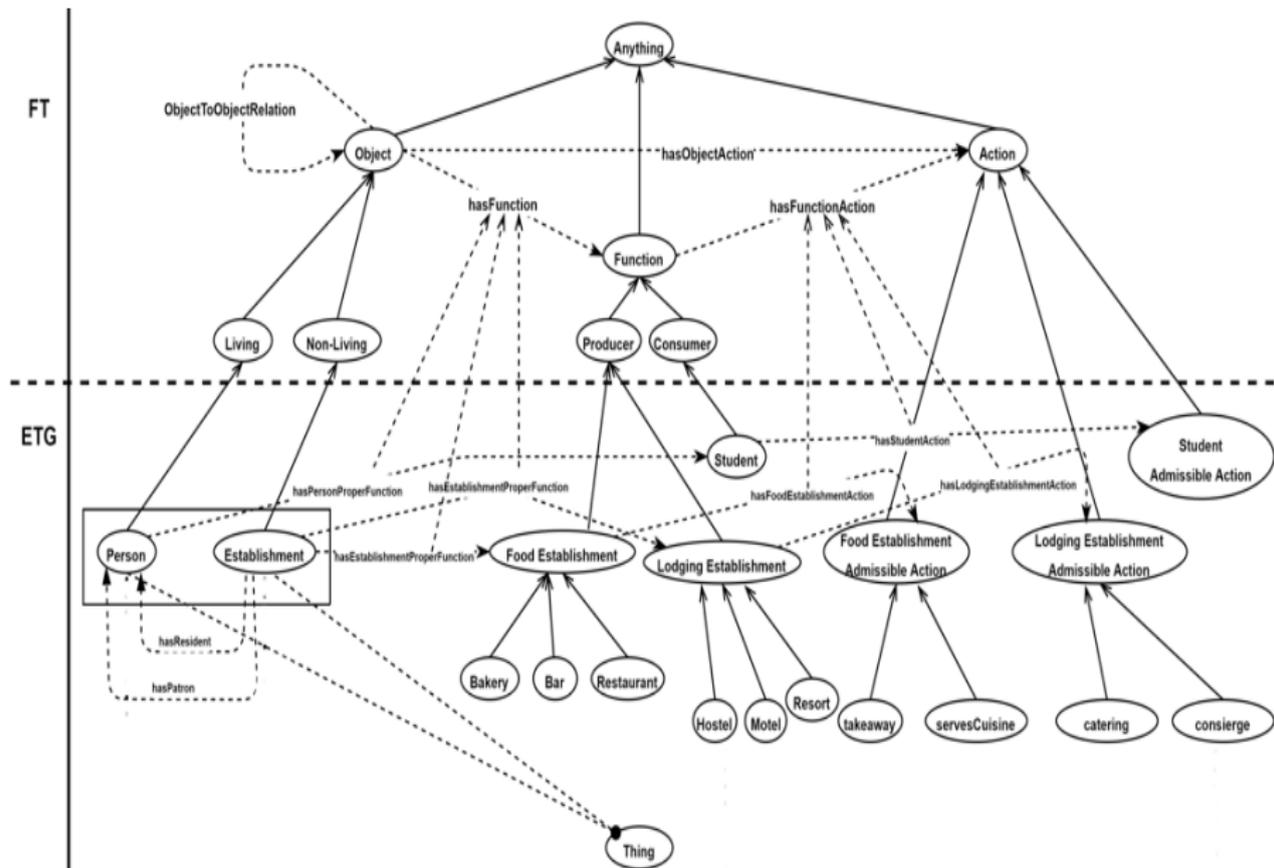
Reference Context

Common

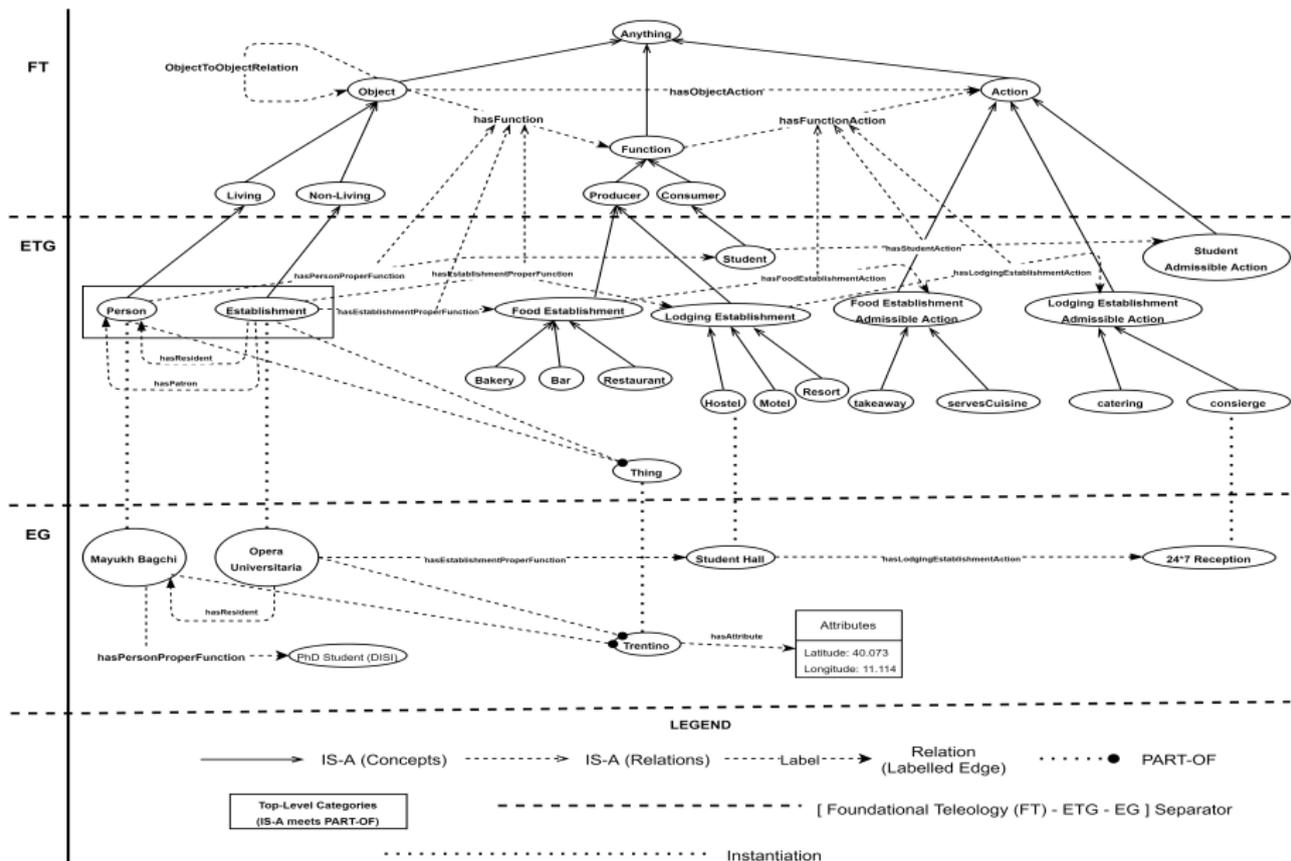
Core

Contextual

Visualization



Extended Visualization



Contents

- 1 Introduction
- 2 FT (*as is*)
- 3 ER (*as is*)
- 4 FT - ER (Schema) Alignment
- 5 ETG Generation**
- 6 Summary

ETG Generation

- The schema alignment activity is practically performed via defining and aligning the ETG Model to the foundational teleology (in the form of an OWL RDF/XML file) using the open source ontology editor [Protégé](#) (demo in formal modelling practice lecture)
- The outcome of the ETG Generation is a concrete OWL RDF/XML file where the formally defined ETG Model is grounded in the primitives of the Foundational Teleology, unified in a single artifact which we term as the *Entity Type Graph*
- The Teleology is finally imported into the KarmaLinker tool, which is used to populate the teleology with appropriately shaped data in the next phase of *iTelos* - data integration

Contents

- 1 Introduction
- 2 FT (*as is*)
- 3 ER (*as is*)
- 4 FT - ER (Schema) Alignment
- 5 ETG Generation
- 6 Summary**

Summary

- We learnt about the step-by-step general methodology for schema alignment
- We understood the need for extending the ER model into an Extended ER (EER) Model employing intermediate concepts for the alignment to the foundational teleology
- The formal ETG model becomes the basis on which heterogeneous datasets are integrated in a stratified manner
- THANK YOU !!!



KDI : Knowledge and Data Integration



Fausto Giunchiglia, Mayukh Bagchi



**Schema Alignment &
ETG Generation**

iTelos Formal Modeling Phase