Formalisation of models and knowledge extraction : Application to heterogeneous data sources in the context of the Factory of the Future

Habilitation à Diriger des Recherches de l'Université de Lorraine

12/01/2021

Mario Lezoche

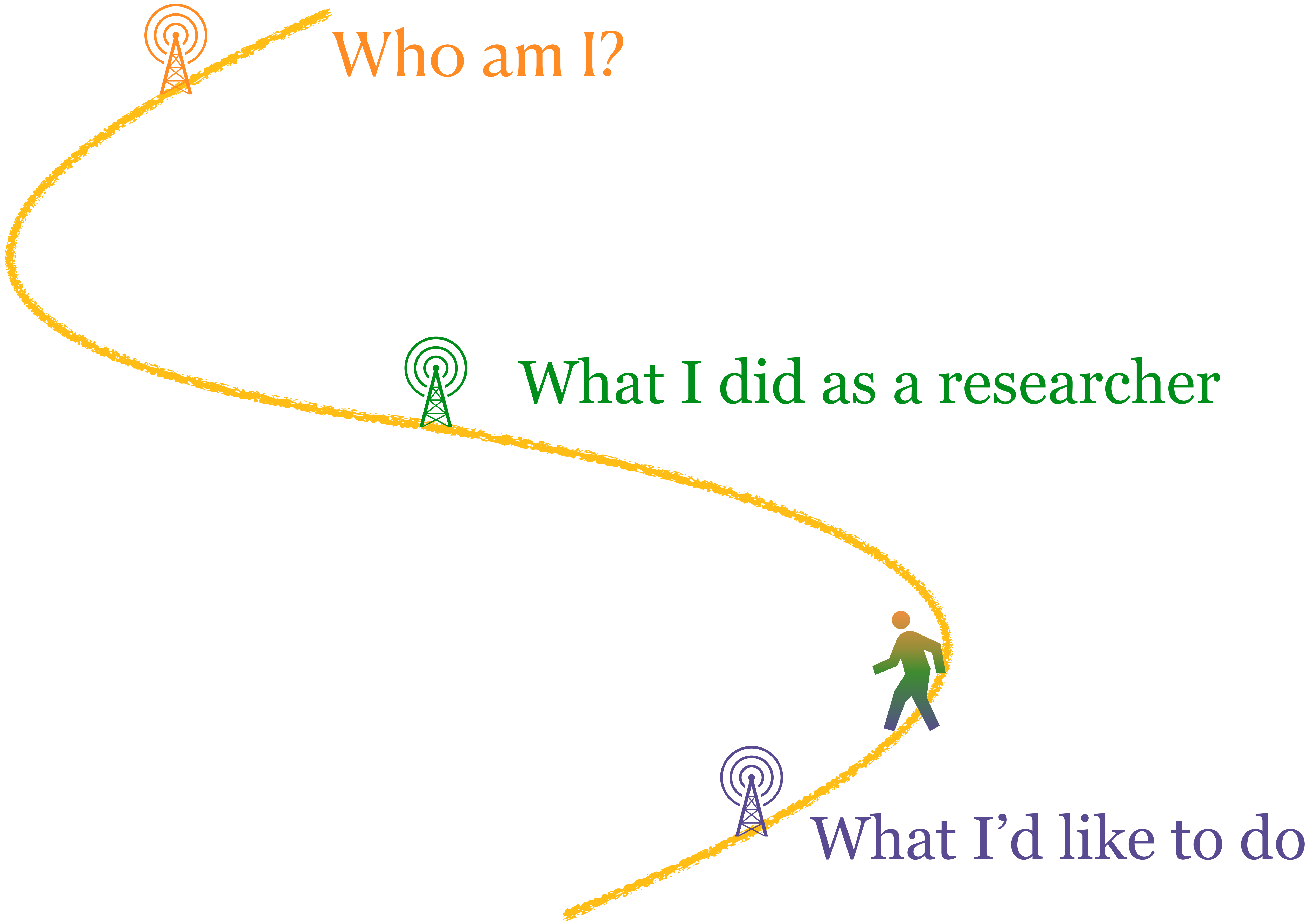
Maître de Conférences de l'Université de Lorraine
Docteur de l'Université ROMA TRE
Section CNU: 61

UMR 7039

Presentation plan

A path from the past to the future

- Overview of my career
  - Quantification in research and teaching
- Research domains I focused in
  - Scientific positioning
  - Scientific contributions
- The future
  - Research and teaching project
  - Conclusion



Presentation

Mario Lezoche
aka Raksati
Born 03/04/1974
Married

Research Master's Degree in Computer Science Engineering specialised in Data Bases and Ontologies, 2006, University of Roma TRE

Ph.D in Computer Science, 2009, CNR of Rome and Università ROMA TRE « Coherence problem between Business Rules and Business Processes »

Post-doc's Degree in Automation and Computer Science Engineering, 2011, Université Henri Poincaré – Nancy I, CRAN, UMR 7039, CNRS

Associate professor
1/9/2012
Teacher at IUT Hubert Curien Epinal, UL
Researcher at CRAN, UMR 7039, CNRS, UL

Co-Mentorship numbers
2 Post-Doc researchers
5 PhD students tutored
1 PhD student ongoing + 2 coming in 2021
2 Research Master degree students

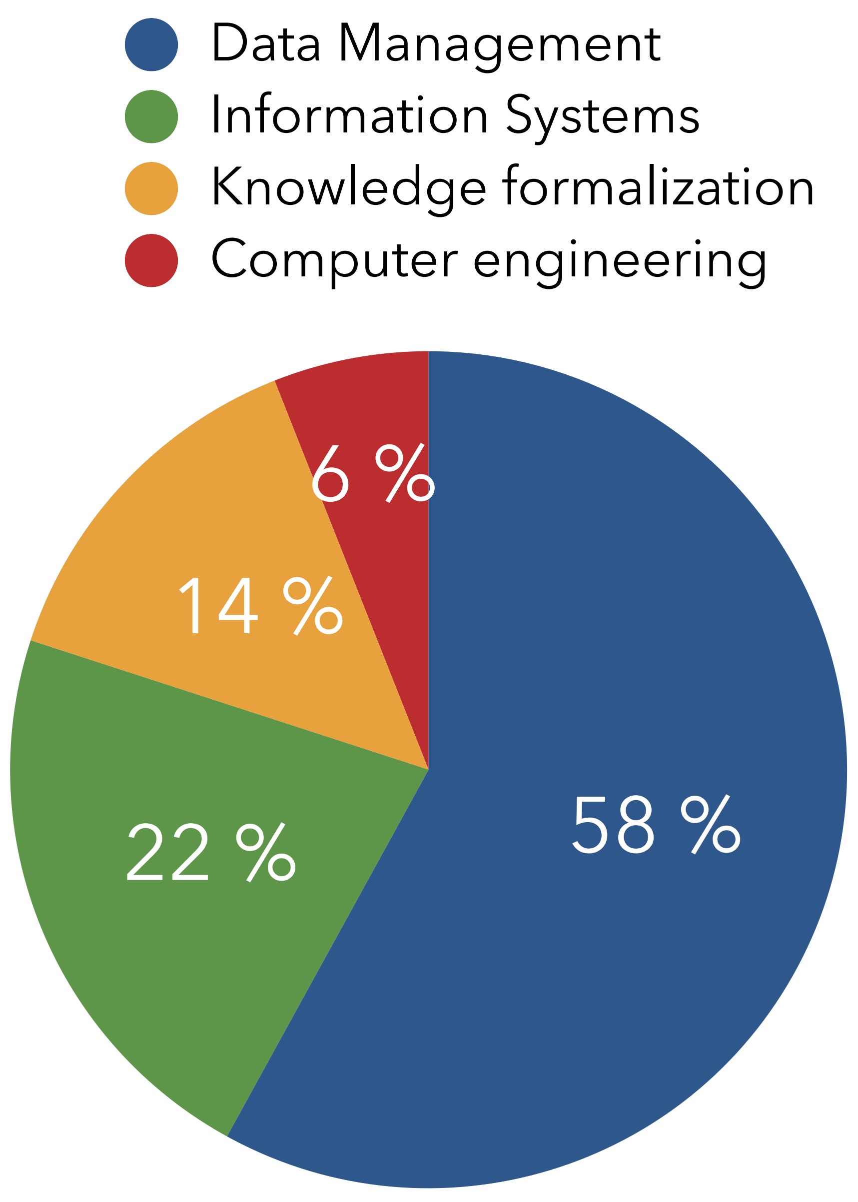
Project numbers
7 exchange researchers projects
2 H2020 + 1 ANR PRCE with responsibilities
1 STIC AmSud as coordinator
2 ANR PRCI registered

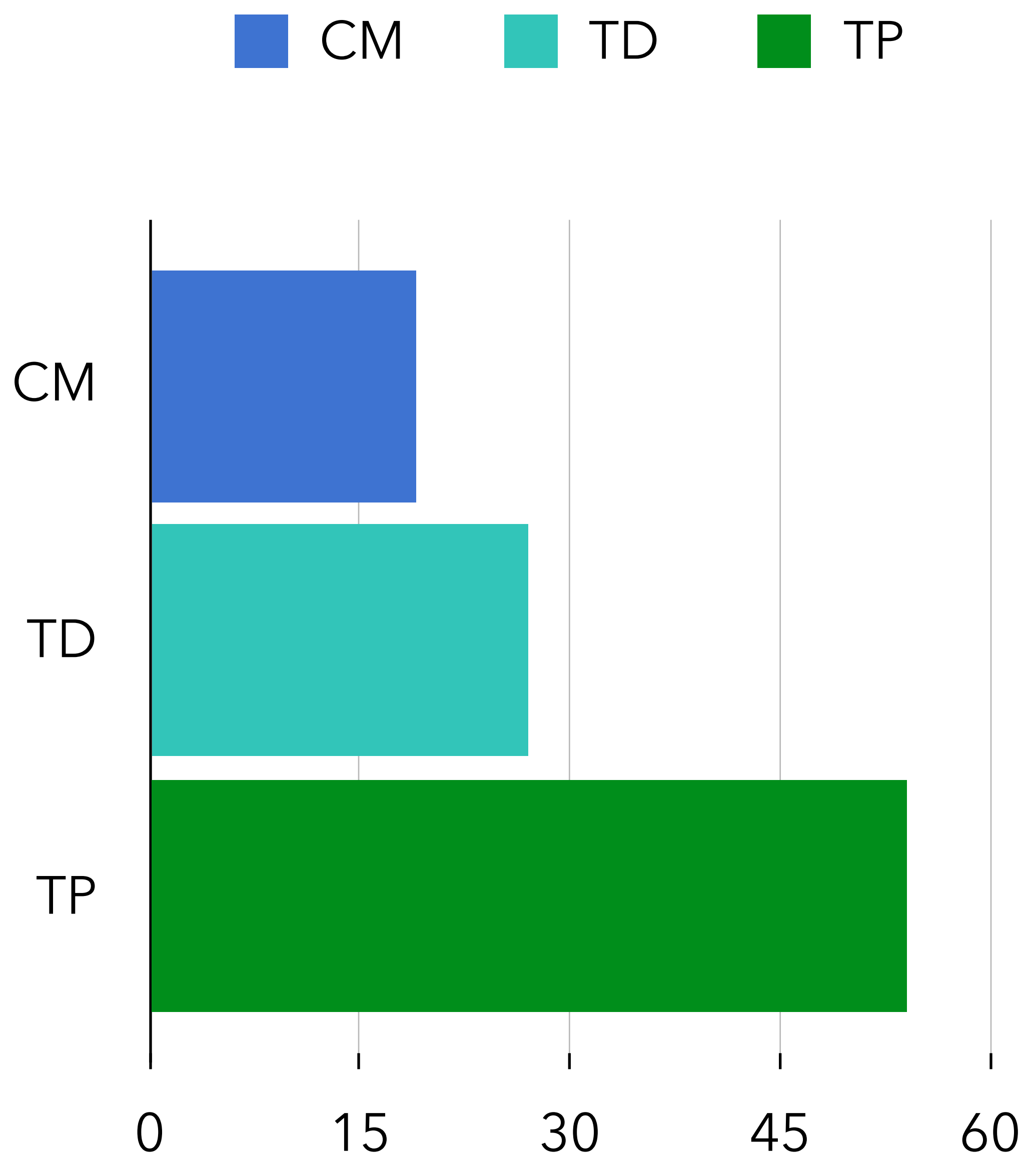
Events organisation numbers
2 Charles Hermite Federation Seminars
1 GDR MACS STP days
2 IPC in ICIST and PRO-VE
1 co-Chair in EI2N workshop

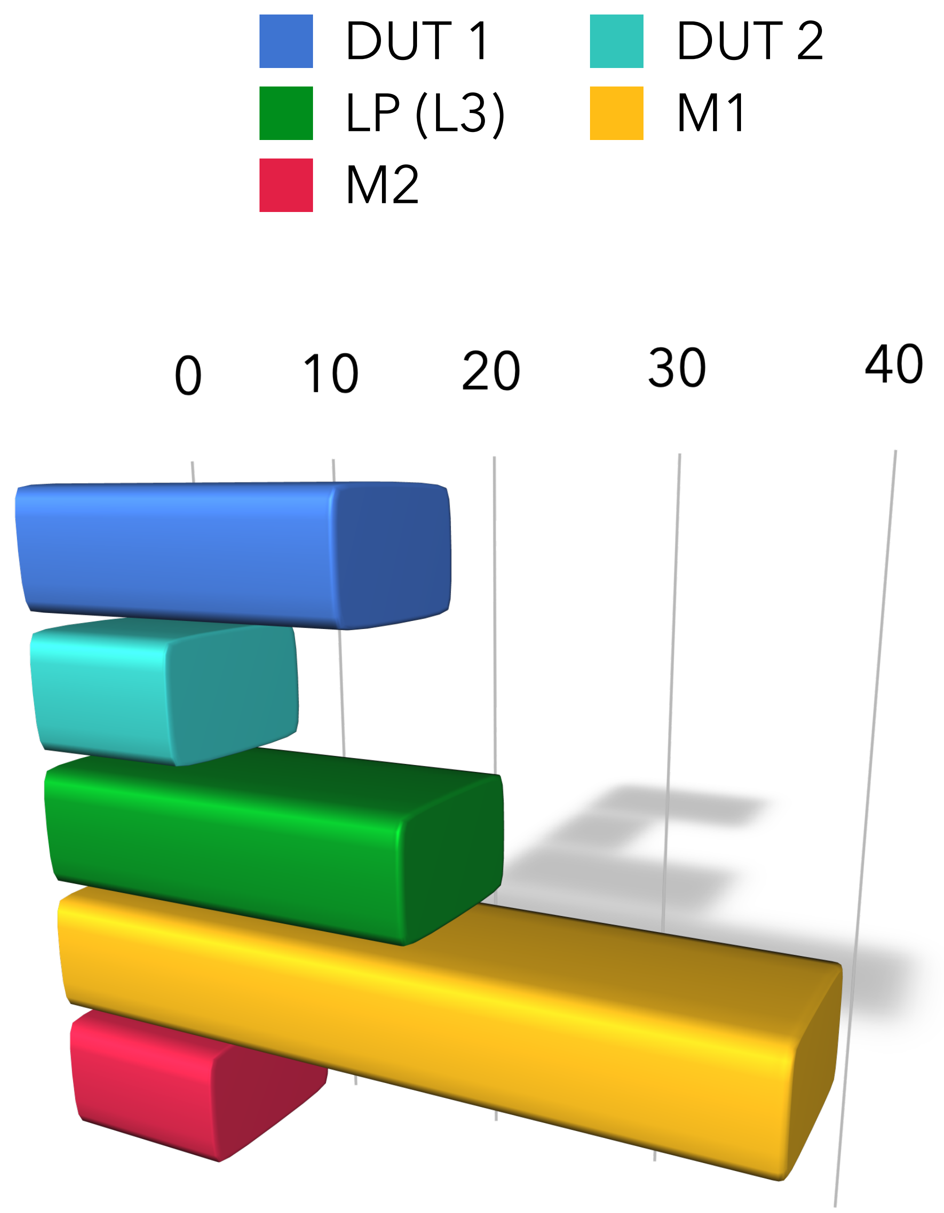
Quantitative research results
55 research publications in total
10 journal papers (9 JCR)
2 journal papers per PhD student
PEDR since 2017

Teaching responsibilities

Theme		Level	Total Eq.TD	Total by Theme	Total Eq TD by Theme
Data Management		DUT 1	280	1496	1628
		DUT 2	256		
		LP (L3)	352		
		M1	740		
Informations Systems		DUT 1	288	568	640
		M1	208		
		M2	144		
Knowledge formalisation		LP (L3)	256	368	392
		M2	136		
Computer Engineering		M1	176	160	176
Student Supervision	Projets	DUT 2	48	192	192
		LP (L3)	48		
		M1	96		
	Stage	DUT 2	48	360	360
		LP (L3)	168		
		M1	72		
		M2	72		
		Total	3388		







Administratif responsibilities

IUT responsibilities (since 2012)

Member of the directive institute council
Elected member of the institute council
Restricted committee of the institute council

Licences Responsible (since 2013 - Licence PIDDI and ECMN)

Admission commission
Development board
Development council (Transition from PIDDI to ECMN 2017)

DUT responsibilities (since 2012)

Member of the admission commission
Member of improvement council

Engineering school Responsible (2014 - 2017)

Specialization in Enterprise Information Systems (SIE)

Institutional positioning

Local Positioning

Researcher at CRAN, UMR 7039, Université de Lorraine, CNRS
Eco-Technic systems engineering (ISET) department
Research project team Intelligent System and Objects in Interaction (S&O-2I)

National and International Positioning

Since 2012: Member of the WG Easy-DIM (currently INE) of GDR MACS

Since 2013: Member of the Scientific Interest Group Interoperability-GR of the Greater Region.

Since 2014: Member of IFAC TC 5.3

Since 2020: Member of IFIP TC-12 (WG12.1 and WG12.6)

Research domain interests

Semantic interoperability in manufacturing enterprises and Factories of the Future (Industry 4.0)

Knowledge formalization in manufacturing enterprises and Factories of the Future (Industry 4.0)

Product Life Cycle in manufacturing with Knowledge management, Semantics

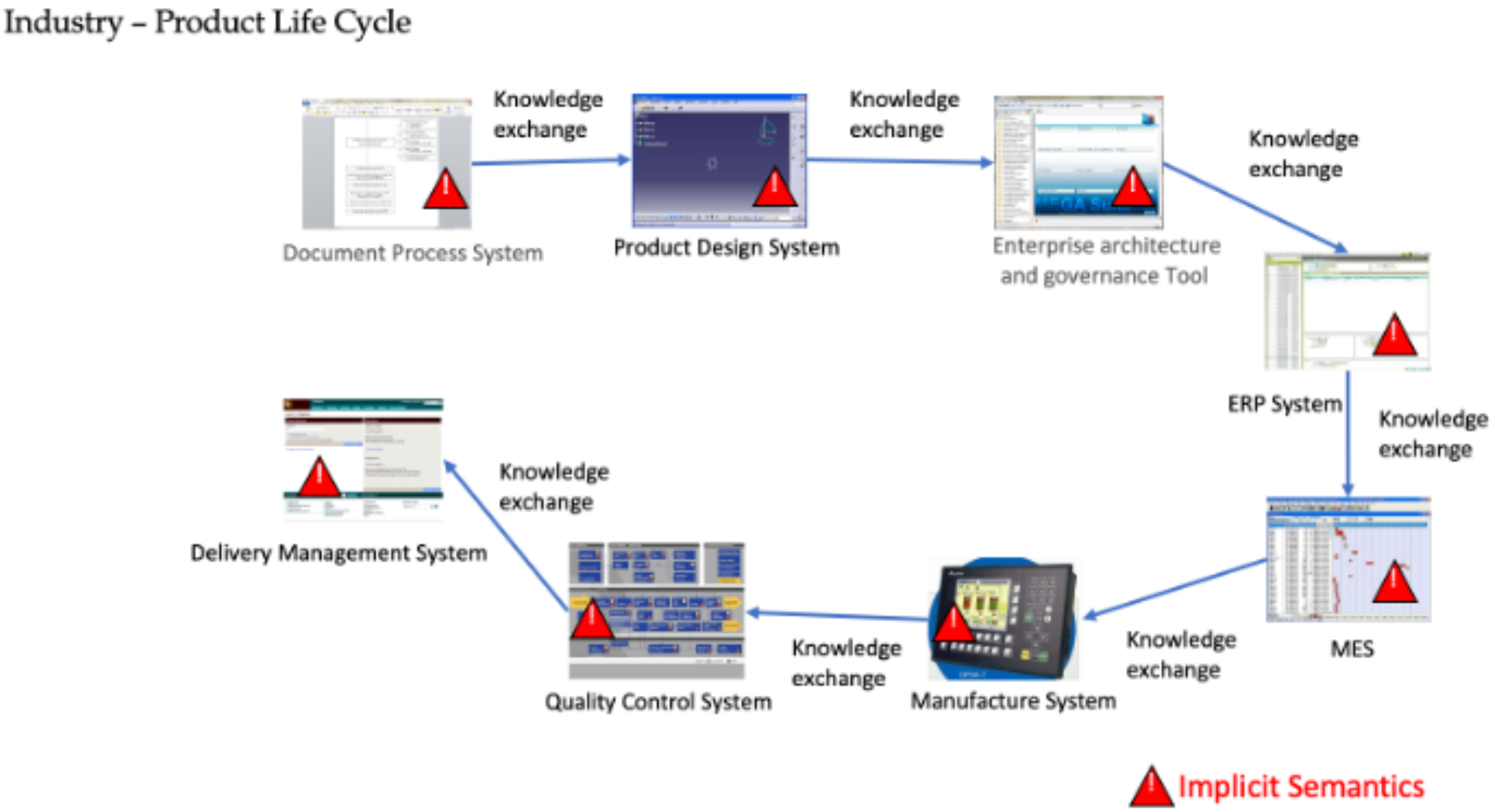
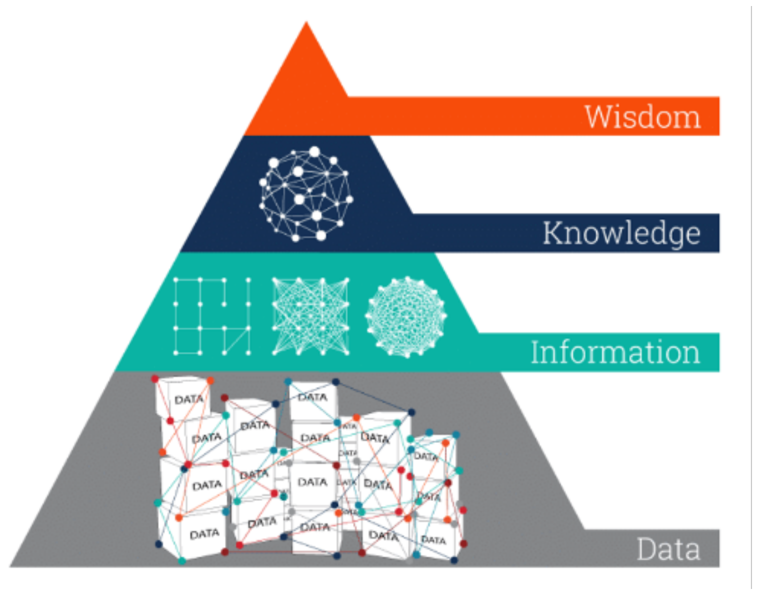


Image Source: Adaptation from [Liao, 2013]

From the past...



... to the Present

Industry (4.0), Knowledge management, data production…

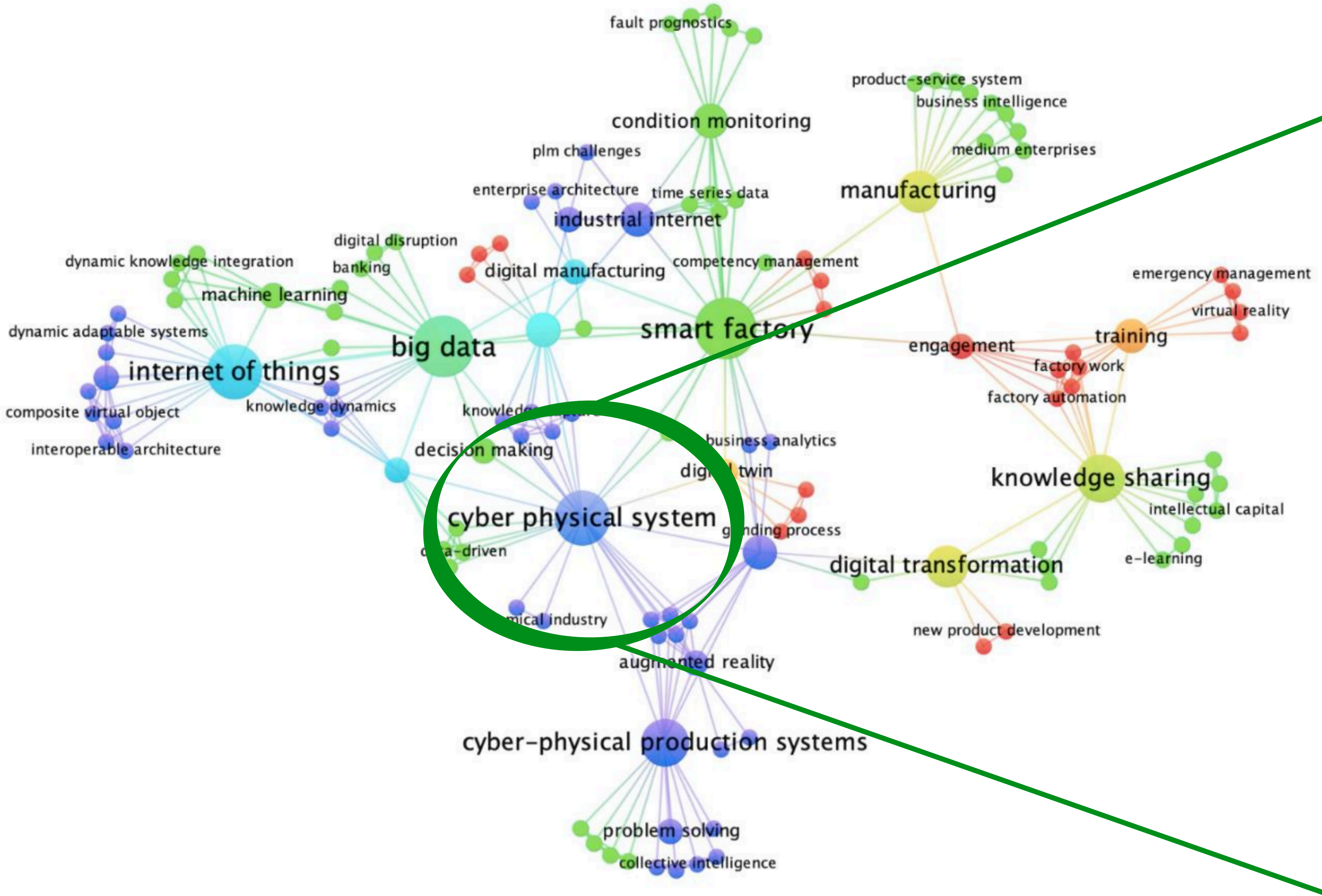


Image Source: M. Fakhar Manesh, et al. "Knowledge Management in the Fourth Industrial Revolution: Mapping the Literature and Scoping Future Avenues," in *IEEE Transactions on Engineering Management*, vol. 68, no. 1, pp. 289-300, Feb. 2021, doi: 10.1109/TEM.2019.2963489.

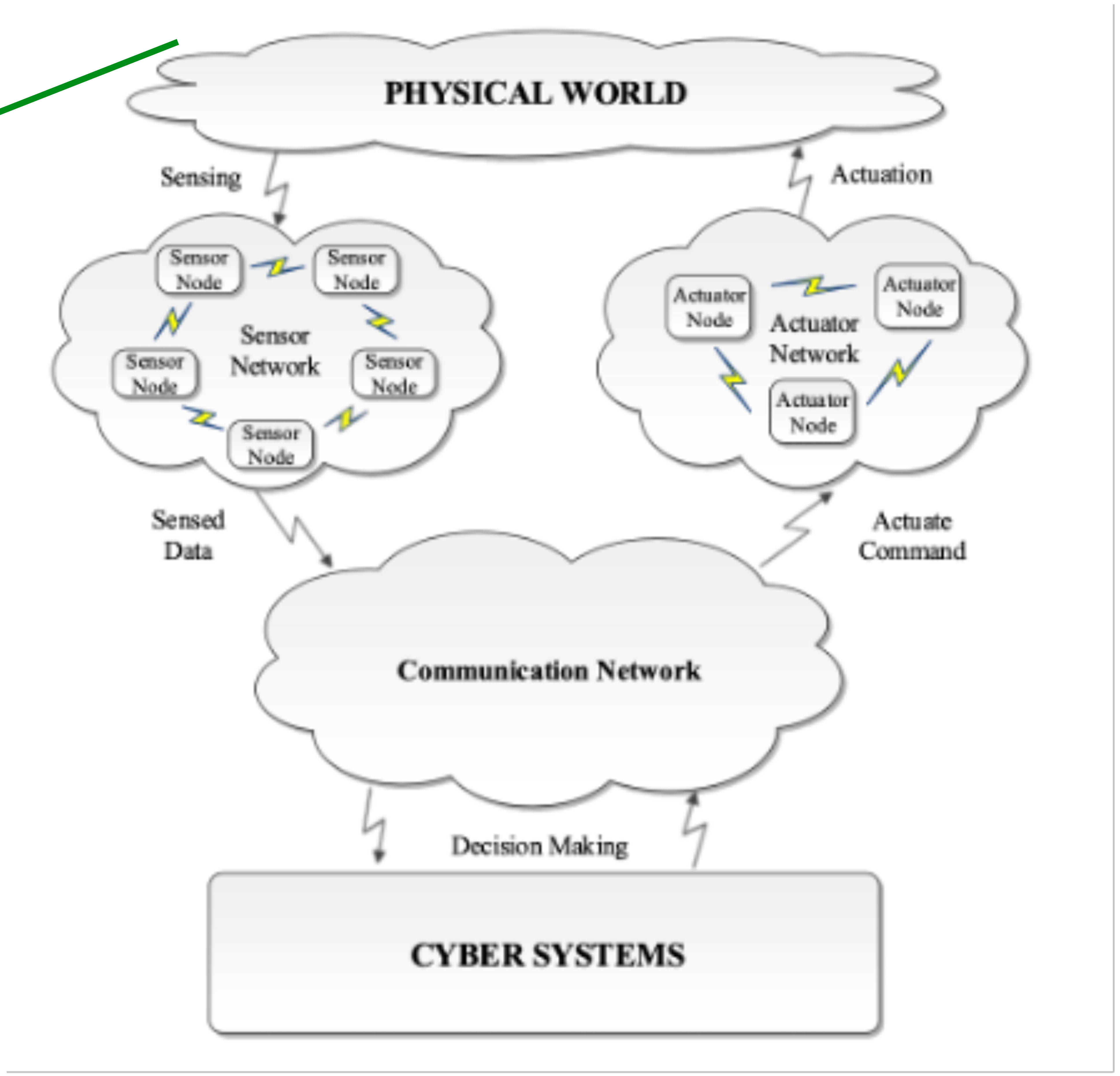


Image Source: CPS Holistic view [Gunes, 2014]

Industry 4.0 and CPS

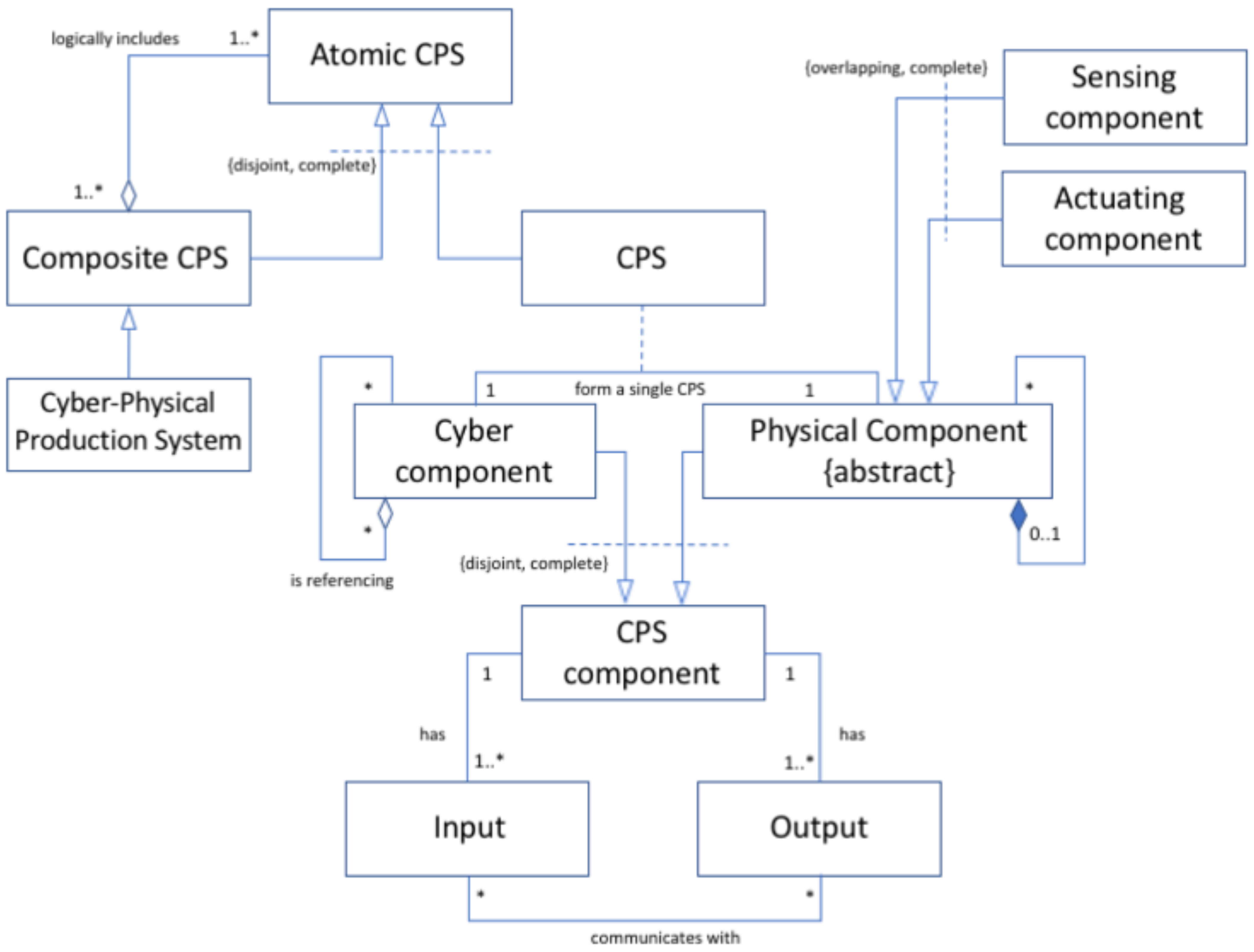


Image Source: [Lezoche, 2020]

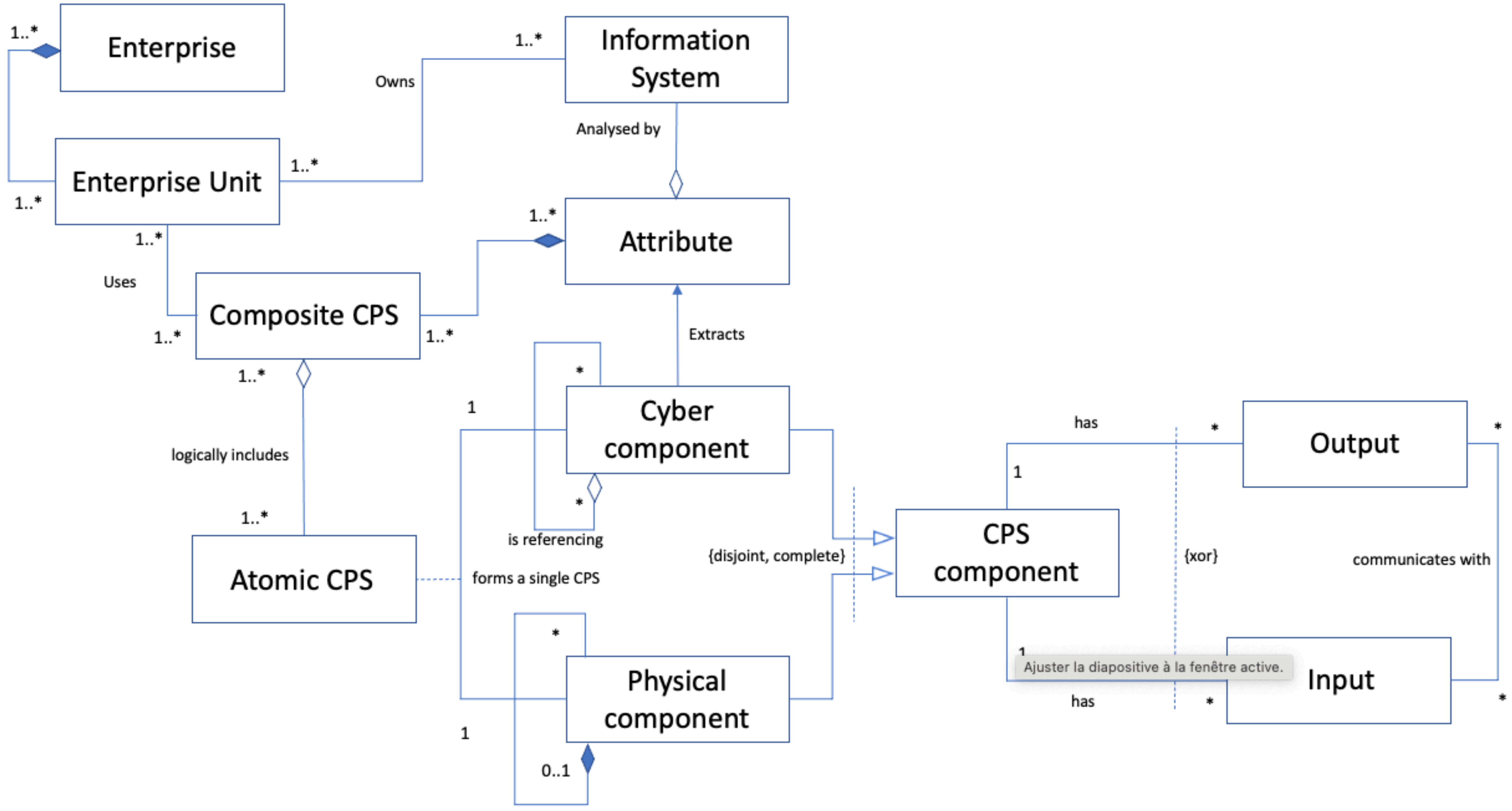
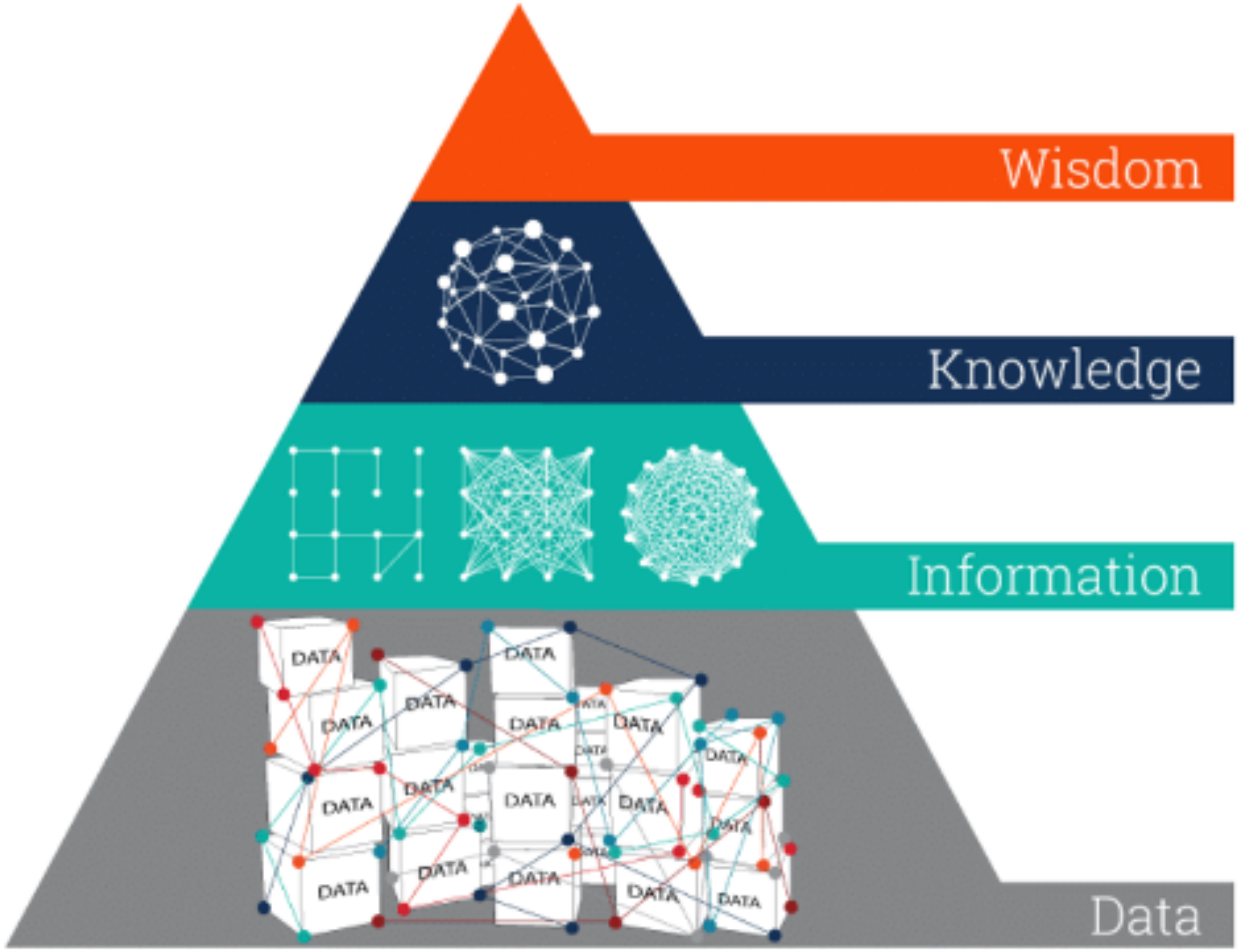


Image Source: Submitted paper Yasamin Eslami, Mario Lezoche, Philippe Kalitine, Sahand Ashouri, How the Cooperative Cyber Physical Enterprise Information Systems (CCPEIS) improve the Semantic Interoperability in the domain of Industry 4.0 through the Knowledge formalization, INCOM 2021

From rough data to Knowledge

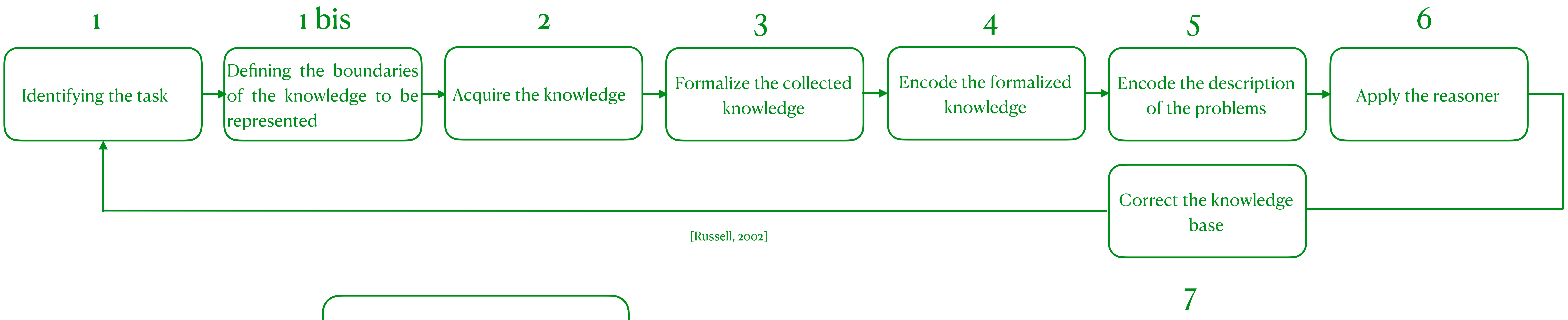


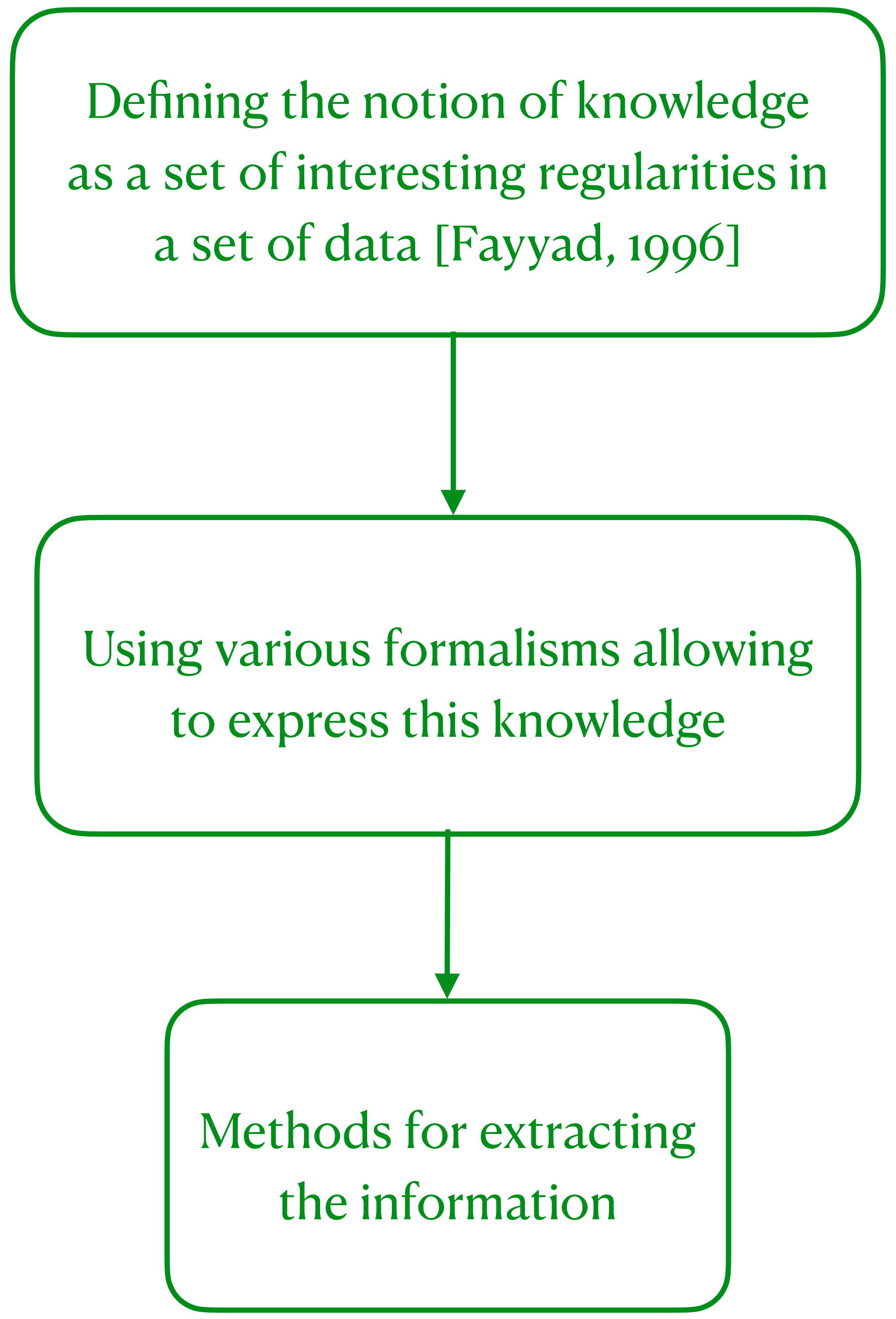
The progression from data to information, knowledge, and wisdom [Ackoff, 1989]

In [Fayyad, 1996] article, we encounter some interesting definitions:

- *"the notion of finding **useful patterns**"*
- *"we can consider a pattern to be knowledge if it exceeds some interestingness threshold"*

From rough data to Knowledge





Knowledge discovery

Knowledge representation

Knowledge extraction (Multi-Relational Data Mining (MRDM))
How to structure the knowledge

My research interests

Semantic interoperability

Formalisation of implicit knowledge in models

- Model-driven cooperative systems engineering, the cooperation concerning "actors"willing to interoperate.

- Computable formalisation of the models
- To make available and extend:
  - mathematical languages
  - modelling languages
  - tools

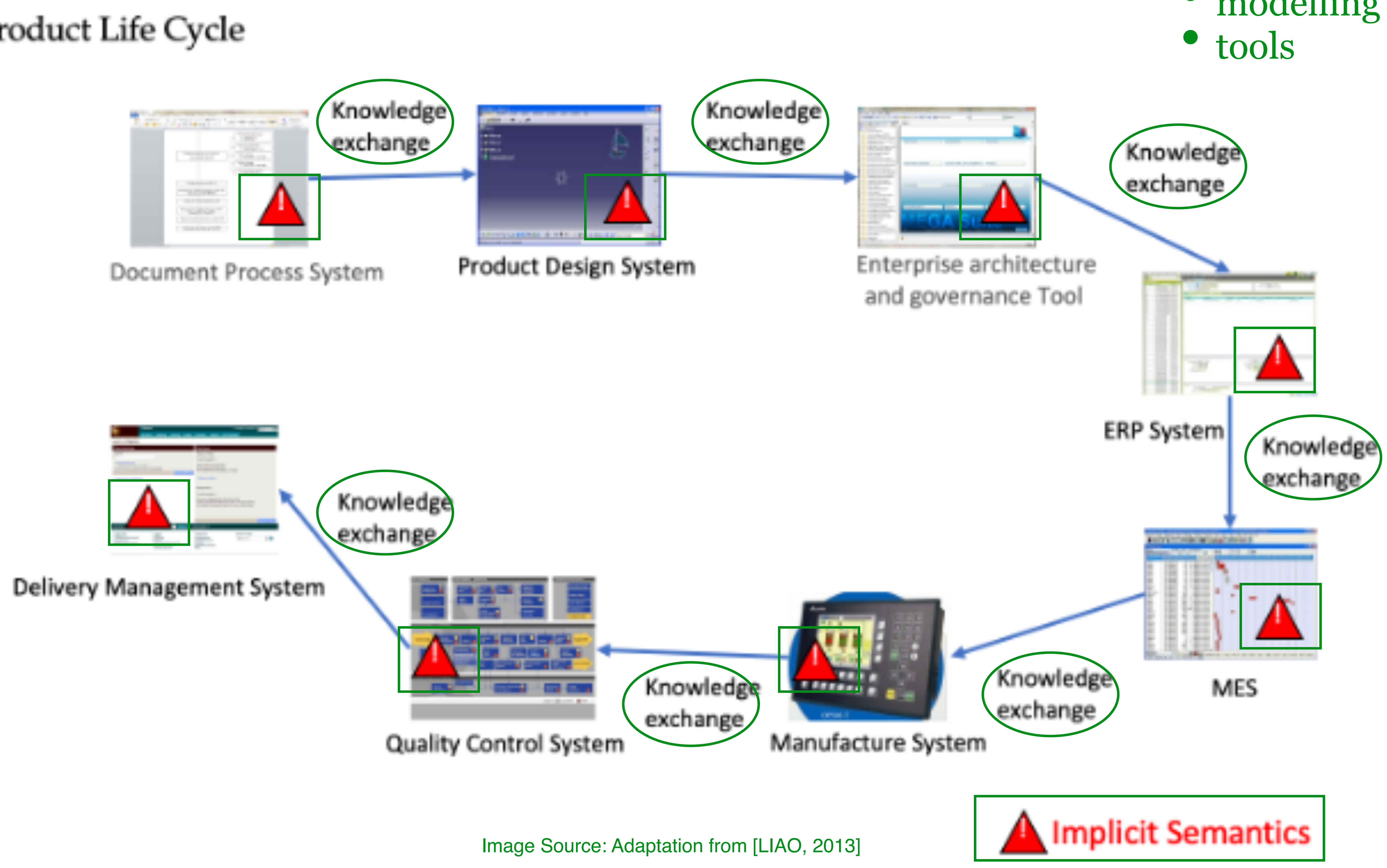
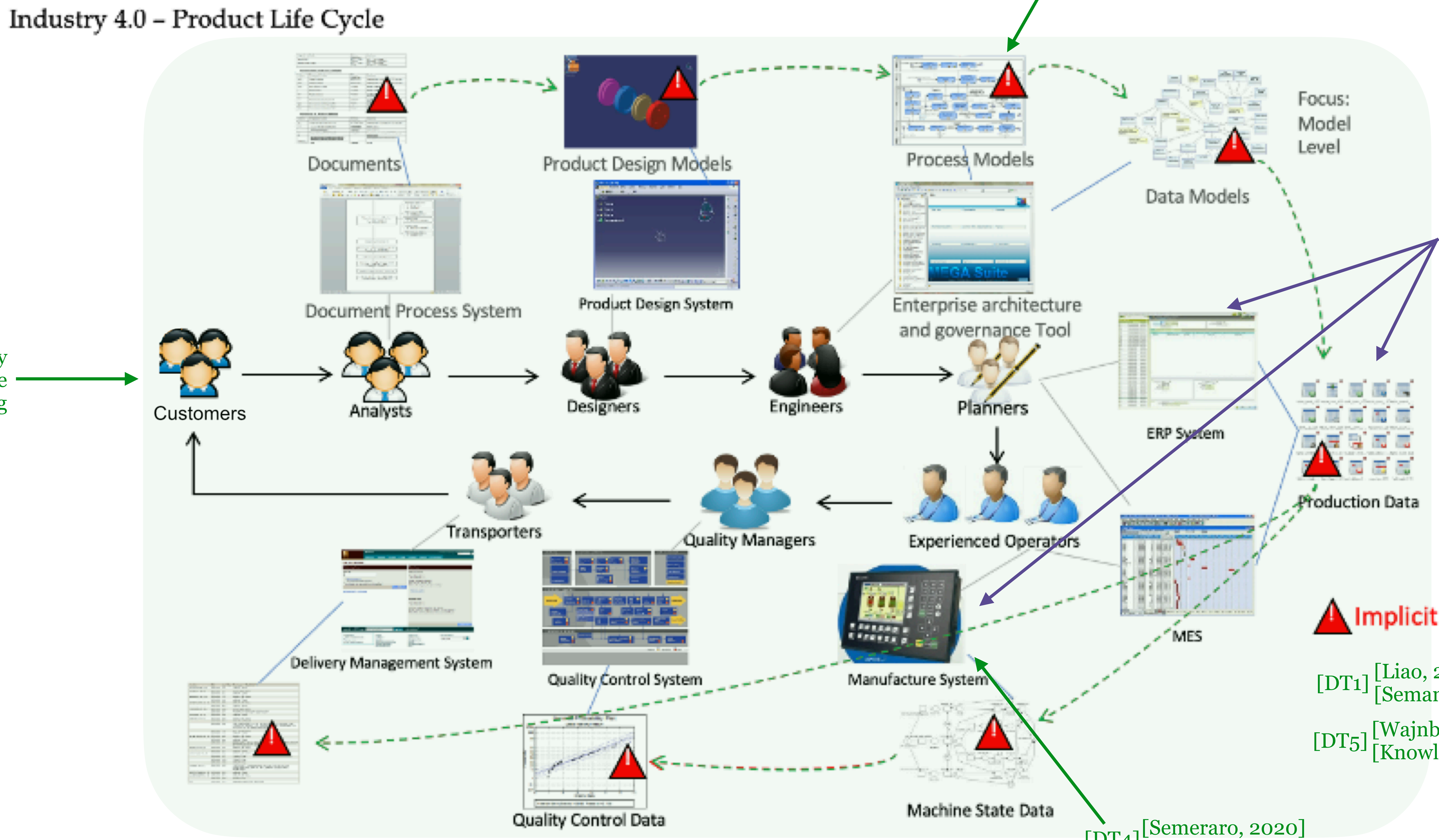


Image Source: Adaptation from [LIAO, 2013]

Scientific contributions from my PhD student's co-mentorships

[DT2] [Pereira Detro, 2017] [Semantic interoperability of processes]



[DT6] [NDIAYE] [Knowledge extraction, formalisation for optimising processing]

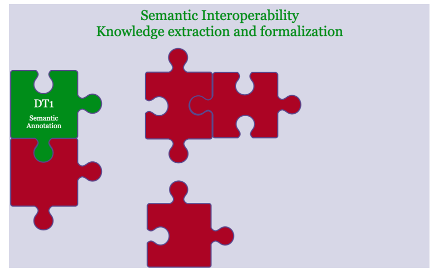
[DT3] [Eslami, 2019] [sustainability assessment in the manufacturing organisations]

[DT1] [Liao, 2013] [Semantic annotation of models]

[DT5] [Wajnberg, 2020] [Knowledge extraction through RCA]

[DT4] [Semeraro, 2020] [Semantic definition invariants of CPSs]

DT1: Yongxin LIAO PhD contributions [Liao, 2013]



Semantic annotations for system interoperability in a PLM context.

Semantic interoperability issues in a Product Lifecycle Management

Industry – Product Life Cycle

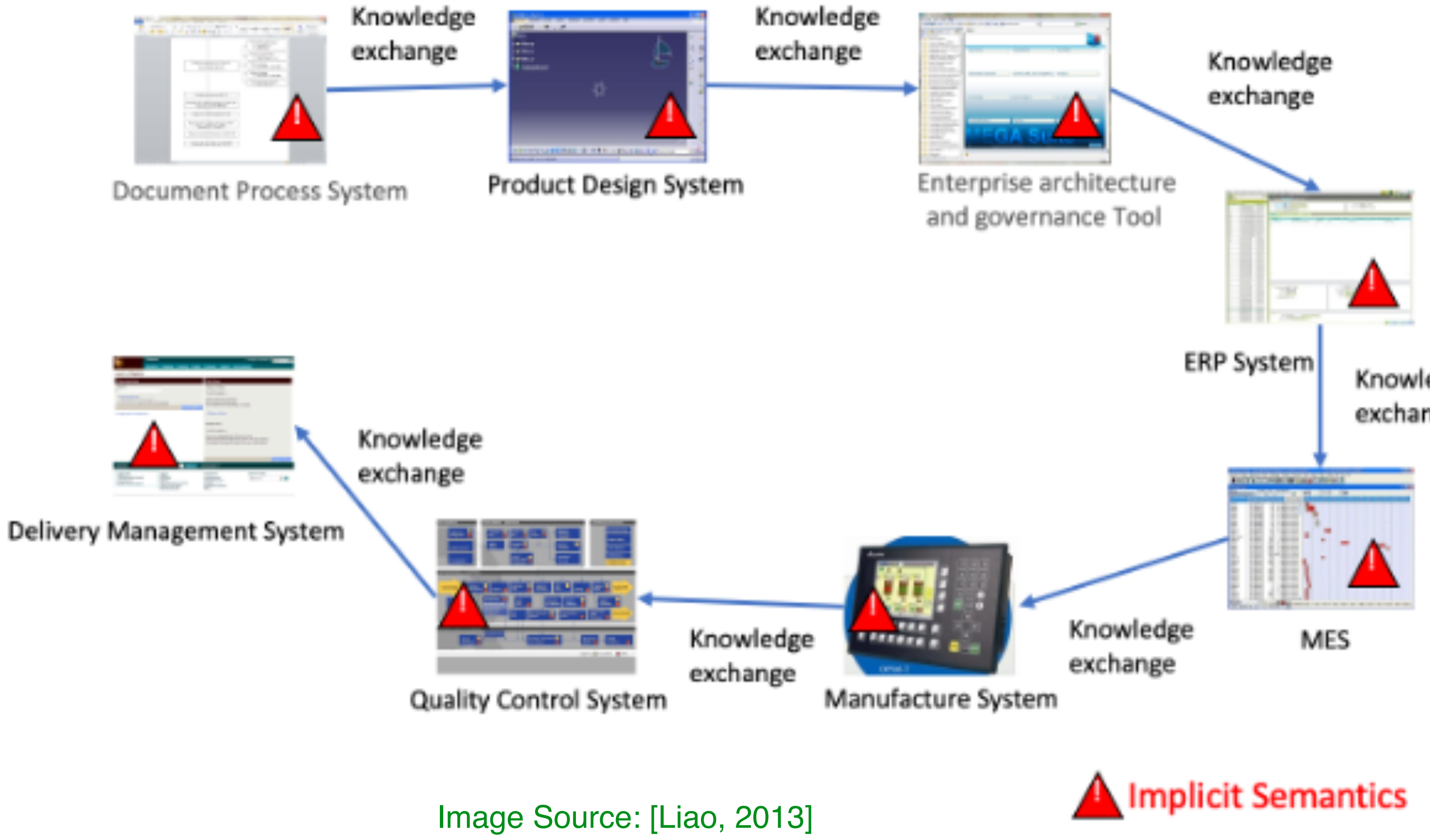


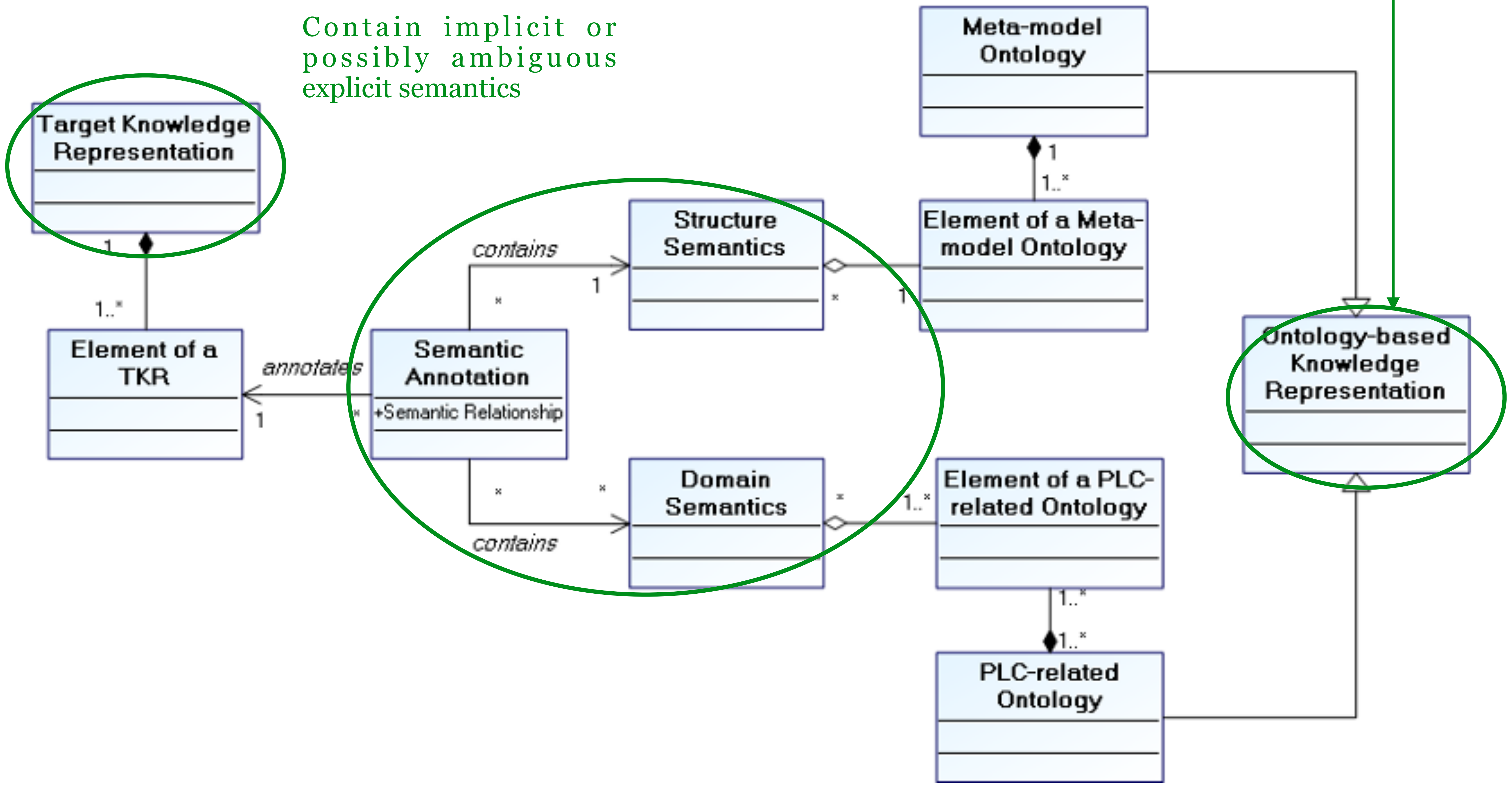
Image Source: [Liao, 2013]

Obstacles

- The implicit semantics that is necessary for understanding a knowledge representation.
- The lack of mechanisms to verify the correctness of explicit semantics in the exchanged knowledge representation.

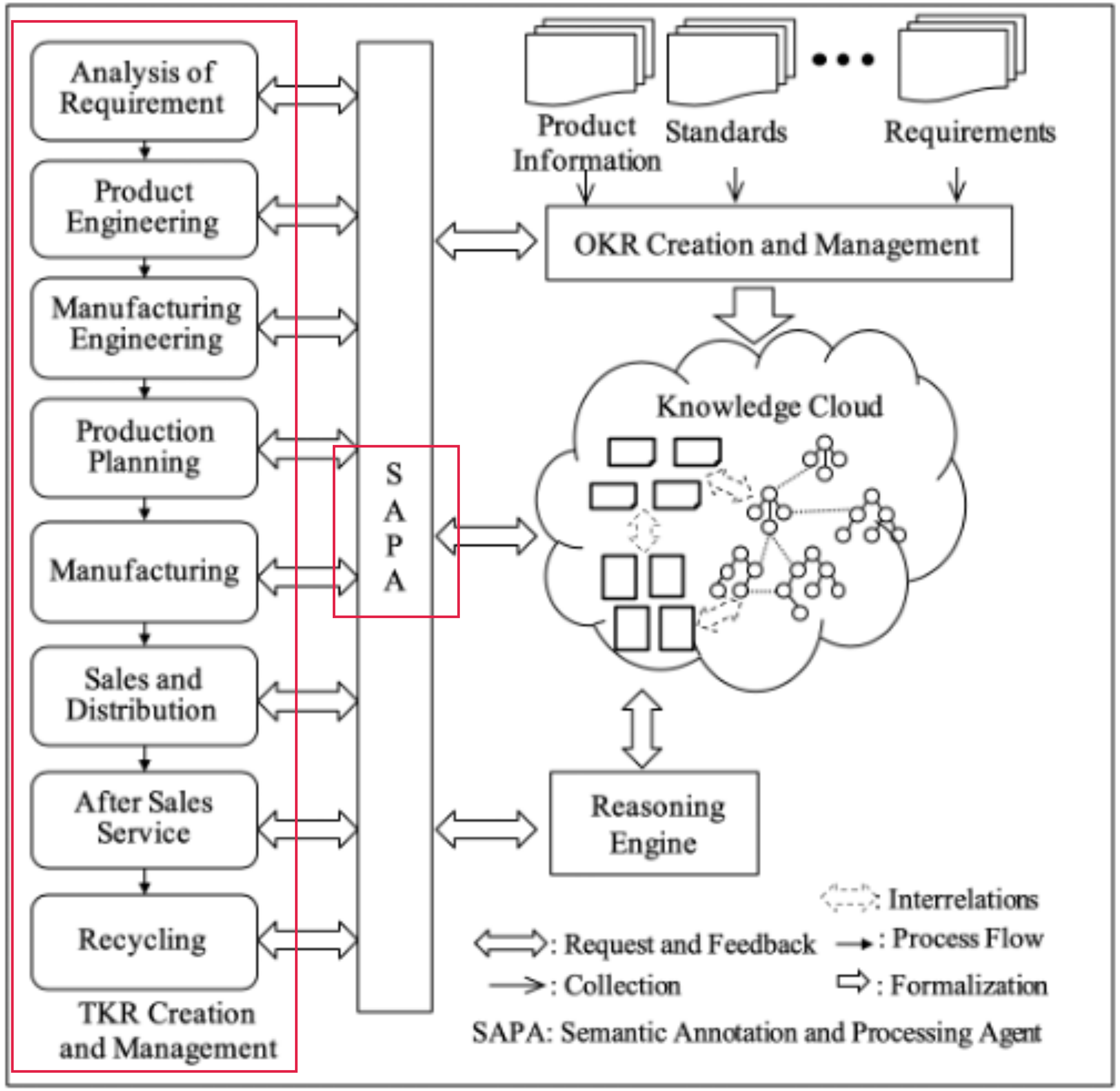
Ontologies that capture different aspects of knowledge and provide the common and shared conceptualisations

Contain implicit or possibly ambiguous explicit semantics



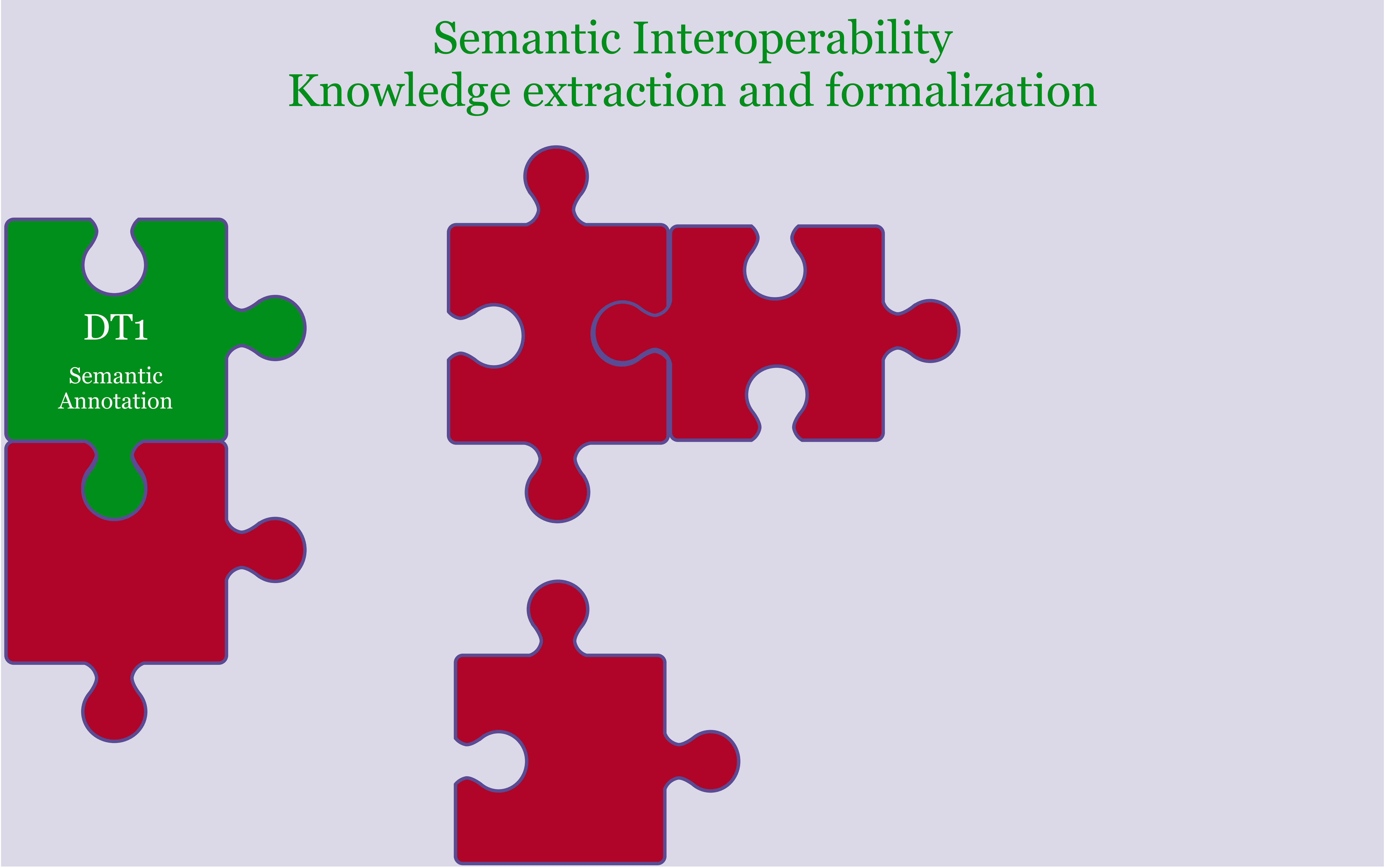
Semantic annotation meta-model -[Yongxin Liao et al., 2014]

DT1: Yongxin LIAO PhD contributions [Liao, 2013]



Semantic annotation framework architecture - Image source [Liao, 2013]

Main scientific challenges direction

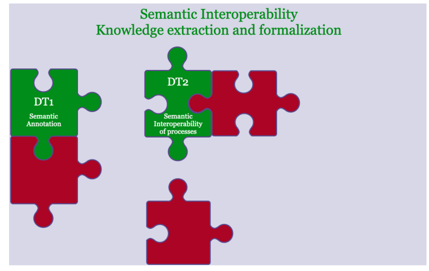


Results

DT1: [Liao, 2013]
Semantic Annotation metamodel
Semantic annotation framework

PhD contributions

DT2: Silvana PEREIRA DETRO contributions [Pereira Detro, 2017]



A framework for interoperability assessment in E-Health information systems using process semantics mining

Model customisation and process model selection related to various constraints

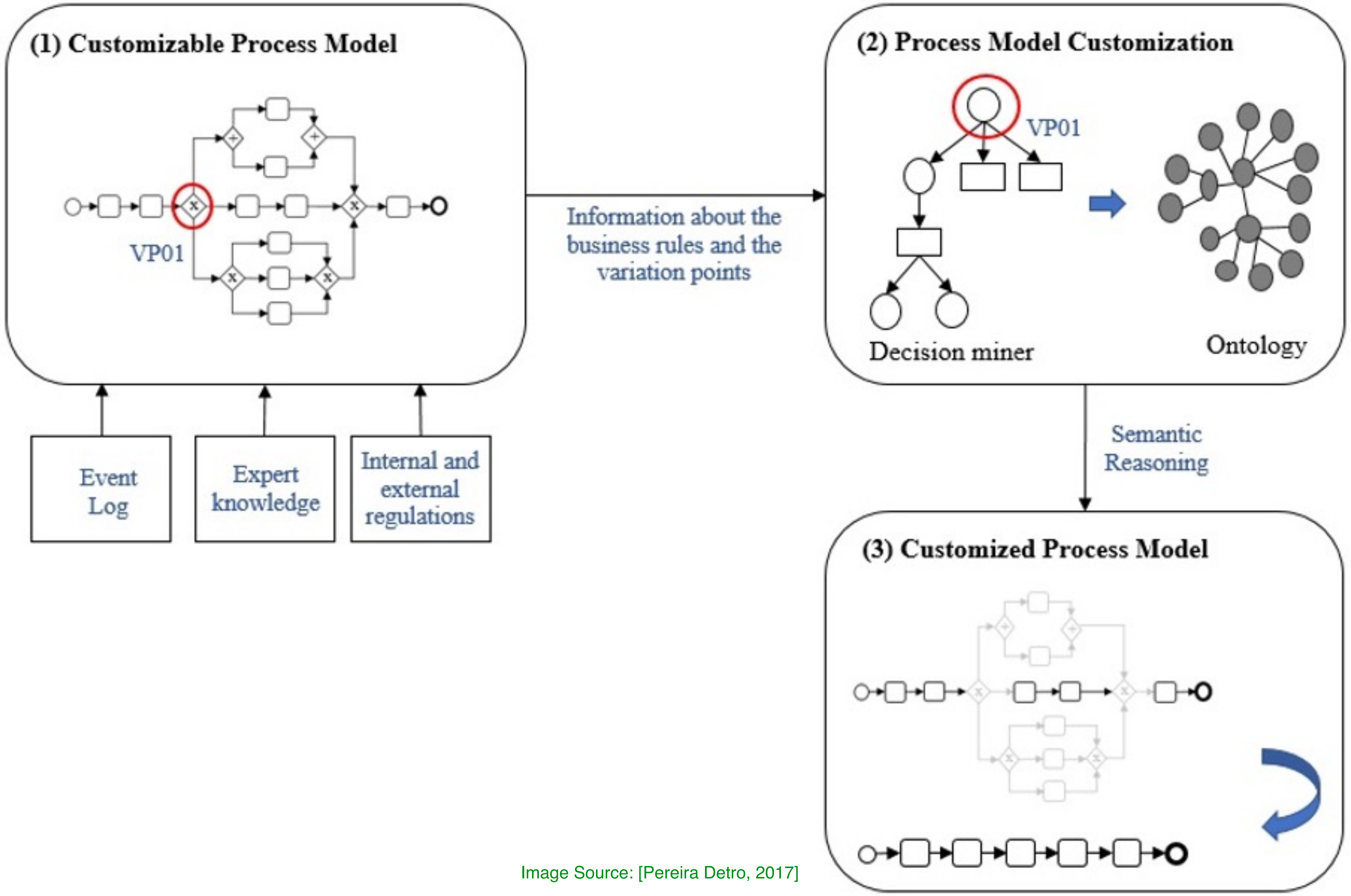


Image Source: [Pereira Detro, 2017]

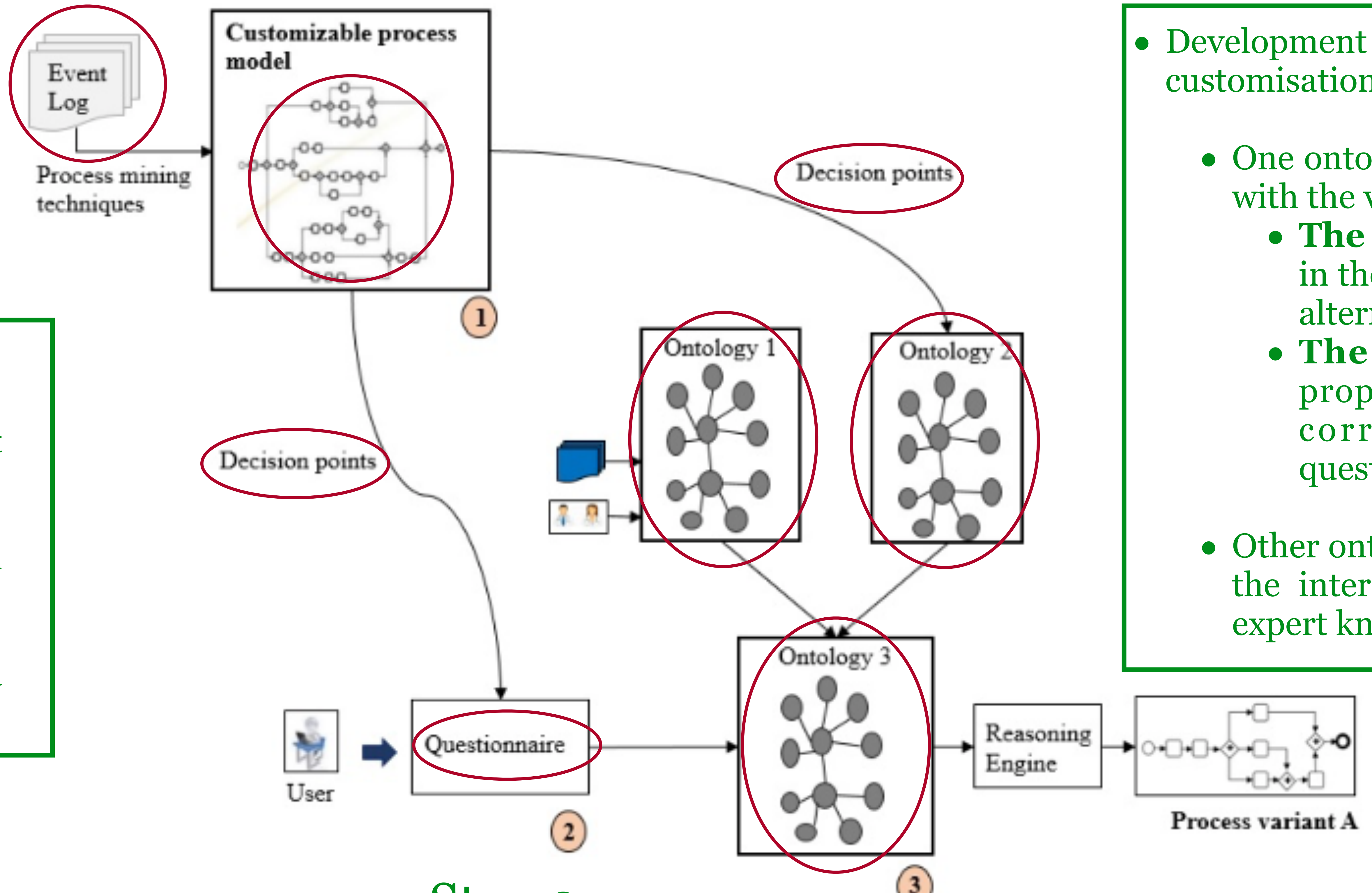
Research questions:

- How to customise a process model in order to obtain a process variant that correctly represent a business context?
- What are the theoretical and practical arguments motivating the application of process mining to discover customisable process models?
- What are the theoretical and practical arguments motivating the use of ontologies for process model customisation?

DT2: Silvana PEREIRA DETRO contributions [Pereira Detro, 2017]

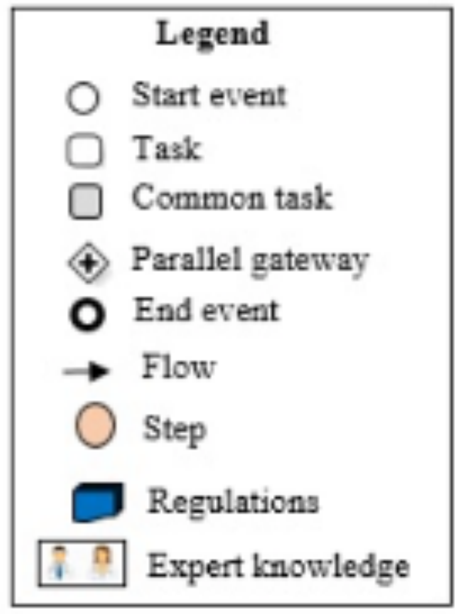
Step 1

- An event log is analysed
- A process model based on the event log, the expert knowledge and the internal and external regulations
- Identify the variation points not providing information for choosing the alternatives for each variation point
- The decision tree concept is used to carry out a decision point analysis



Step 3

- Development of the ontologies for process model customisation.
  - One ontology formalises the knowledge related with the variation points.
    - **The leaf nodes** are defined as concepts in the ontology, which correspond with the alternatives for the variation points.
    - **The branches** are defined as data properties in the ontology and they correspond with the facts in the questionnaire.
  - Other ontology formalises the knowledge about the internal and/or external regulations and expert knowledge.



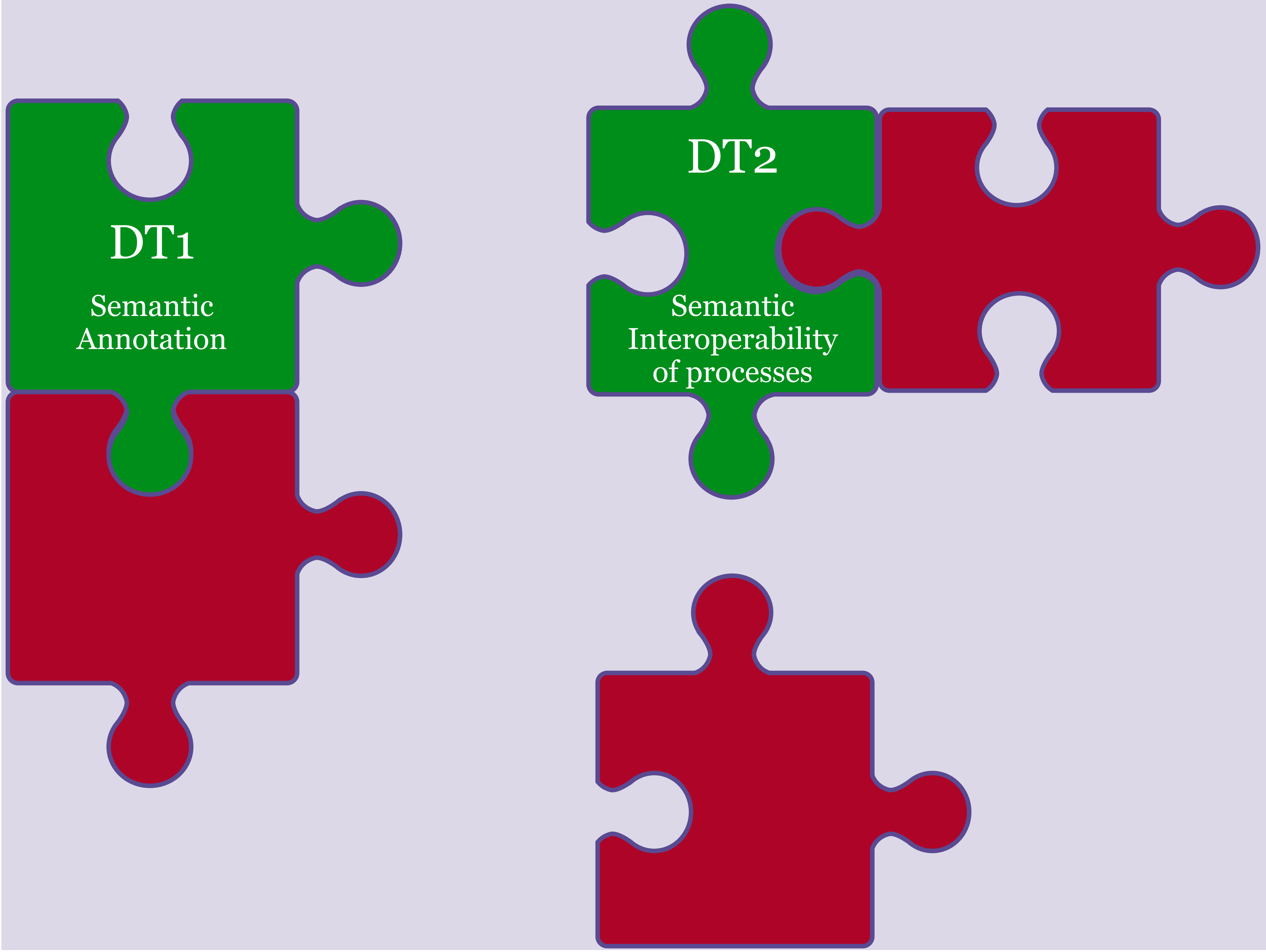
Step 2

- The questionnaire is applied to guide users in providing the information needed for process variant selection.
- Variation point definition enables interdependencies
- Definition of the order of dependence

Framework for customise process variants from - Image source [Pereira Detro, 2017]

Main scientific challenges direction

Semantic Interoperability
Knowledge extraction and formalization

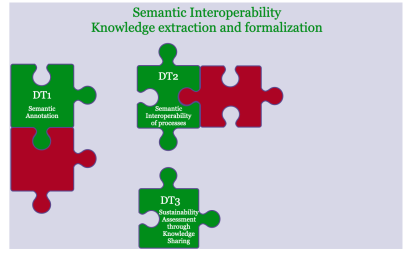


PhD contributions

Results

DT1: [Liao, 2013]
Semantic Annotation metamodel
Semantic annotation framework

DT2: [Pereira Detro, 2017]
Framework for customised process variants

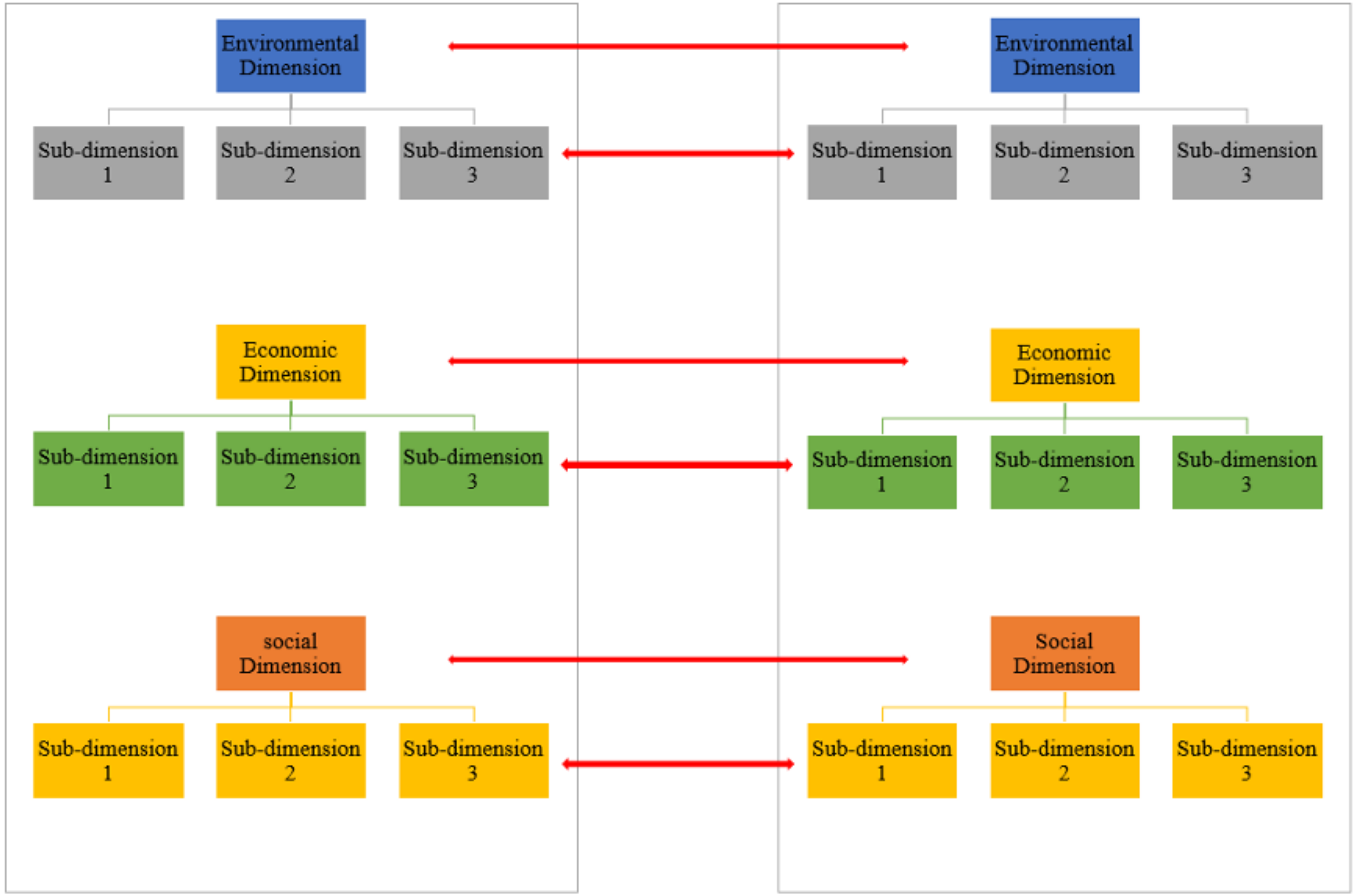


DT3: Yasamin ESLAMI contributions [Eslami, 2019]

A Modelling-Based Sustainability Assessment in Manufacturing Organisations

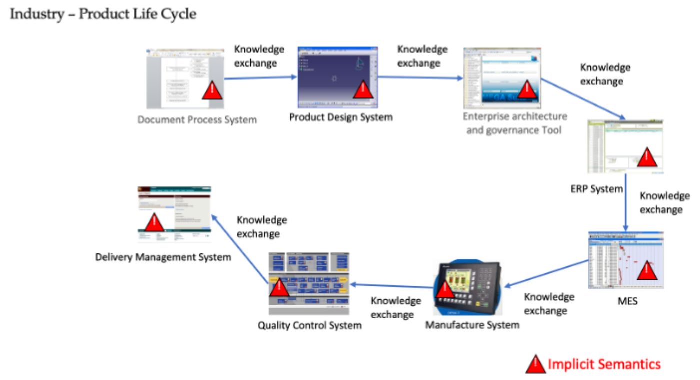
Solve the gap between the sustainability performance improvement of enterprise manufacturing needed and the efficiency and capability of the available assessment tools.

Manufacturing Domain in Practice **Scientific Domain in Literature**

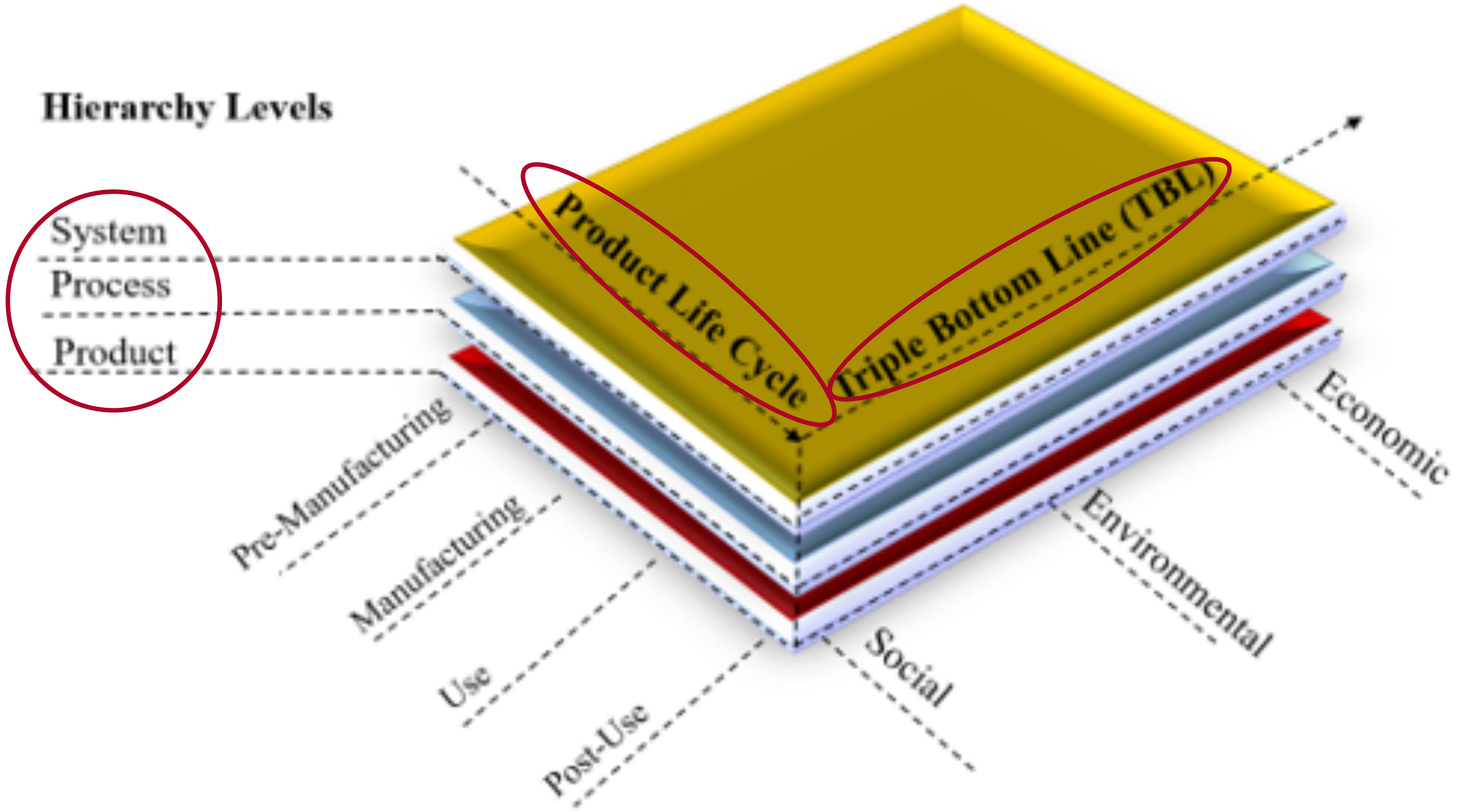


Sustainability dimensions comparison - Image source [Eslami, 2019]

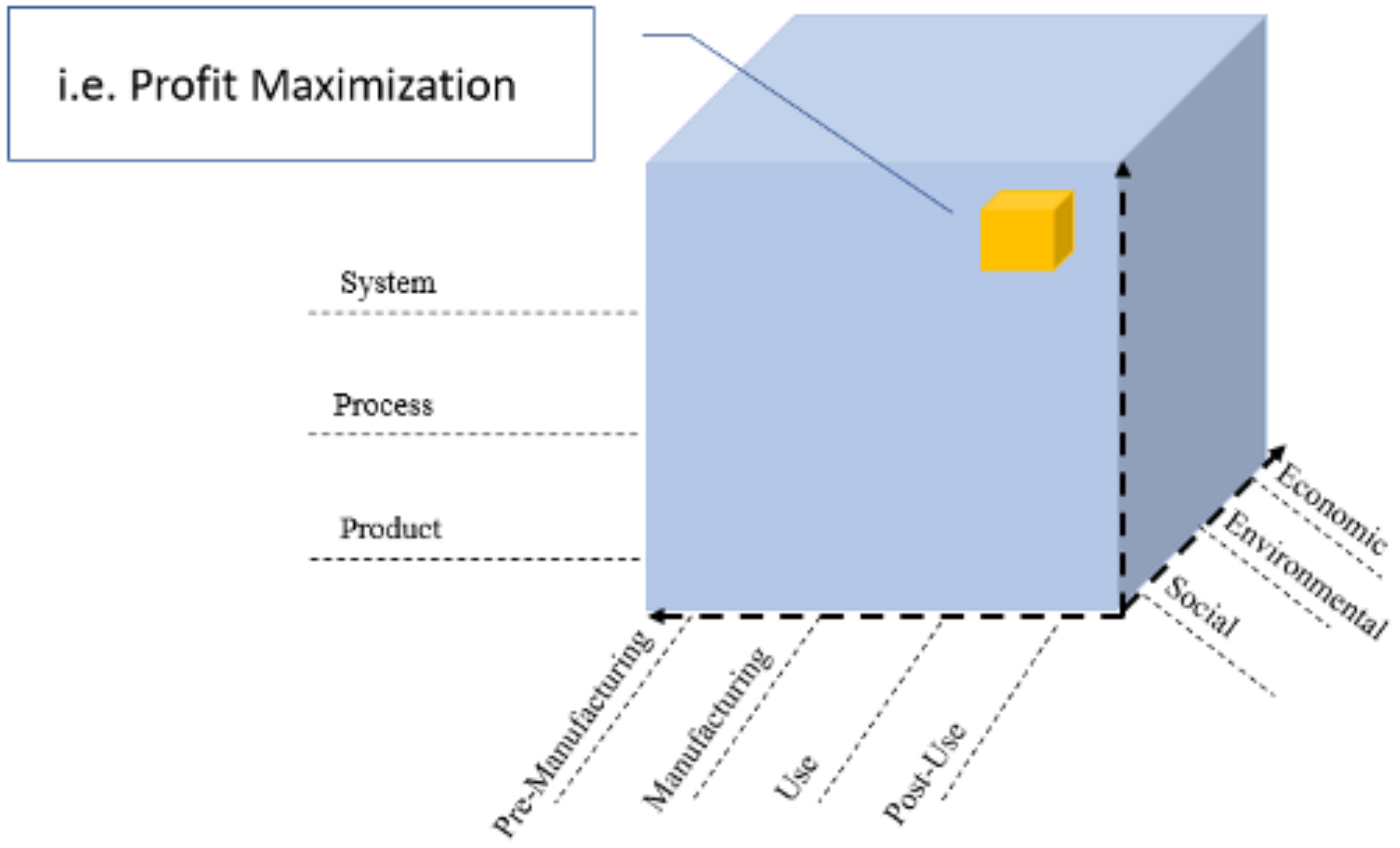
During all the Product Life Cycle



DT3: Yasamin ESLAMI contributions [Eslami, 2019]

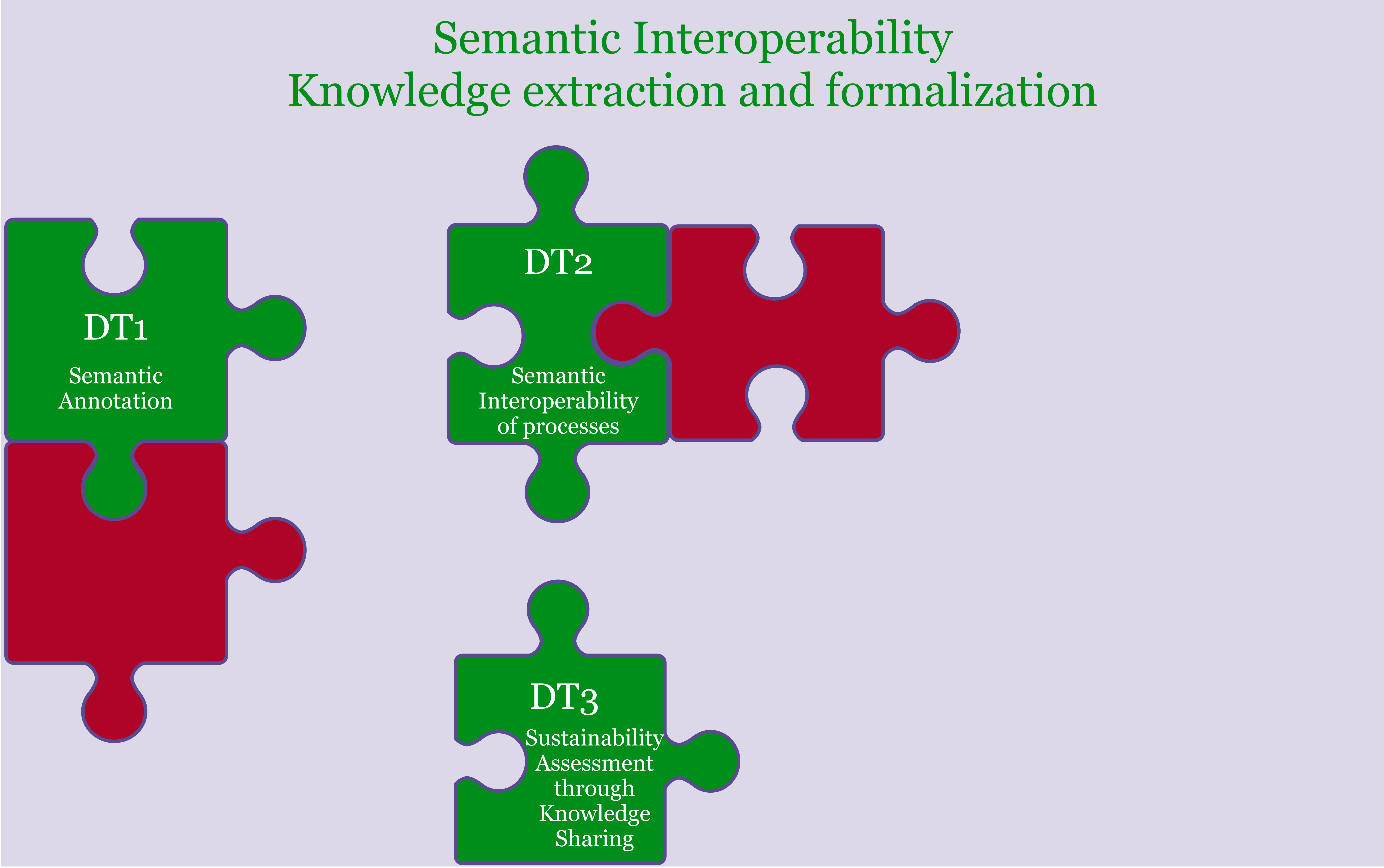


Three-Dimensional Model for Sustainability Assessment [Eslami, 2019]



An example of a sustainability cubical [Eslami, 2019]

Main scientific challenges direction



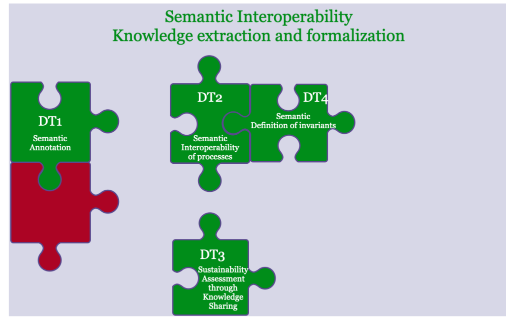
Results

DT1: [Liao, 2013]
Semantic Annotation metamodel
Semantic annotation framework

DT2: [Pereira Detro, 2017]
Framework for customised process variants

DT3: [Eslami, 2019]
Three-Dimensional Model for Sustainability Assessment
Cubical model from the Knowledge integration

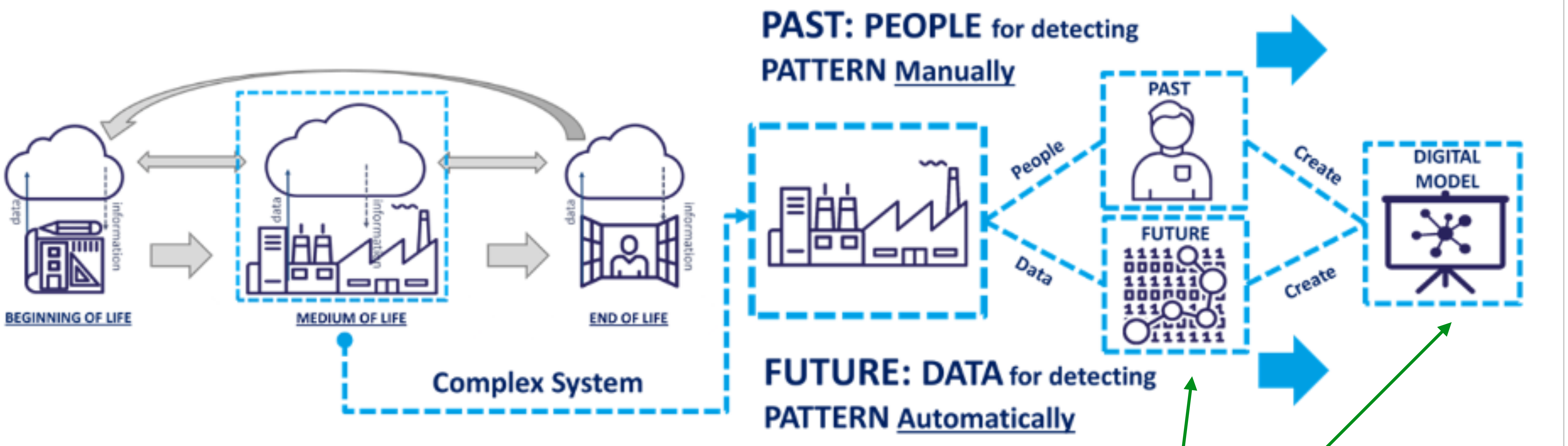
PhD contributions



DT4: Concetta SEMERARO contributions [Semeraro, 2020]

Contribution to the formalisation which is driven by the data of modelling invariants of cyber-physical systems

The thesis aims to identify an approach to formalize data-driven invariant modelling constructs for improving the smartness of manufacturing processes and products, involving networked components.



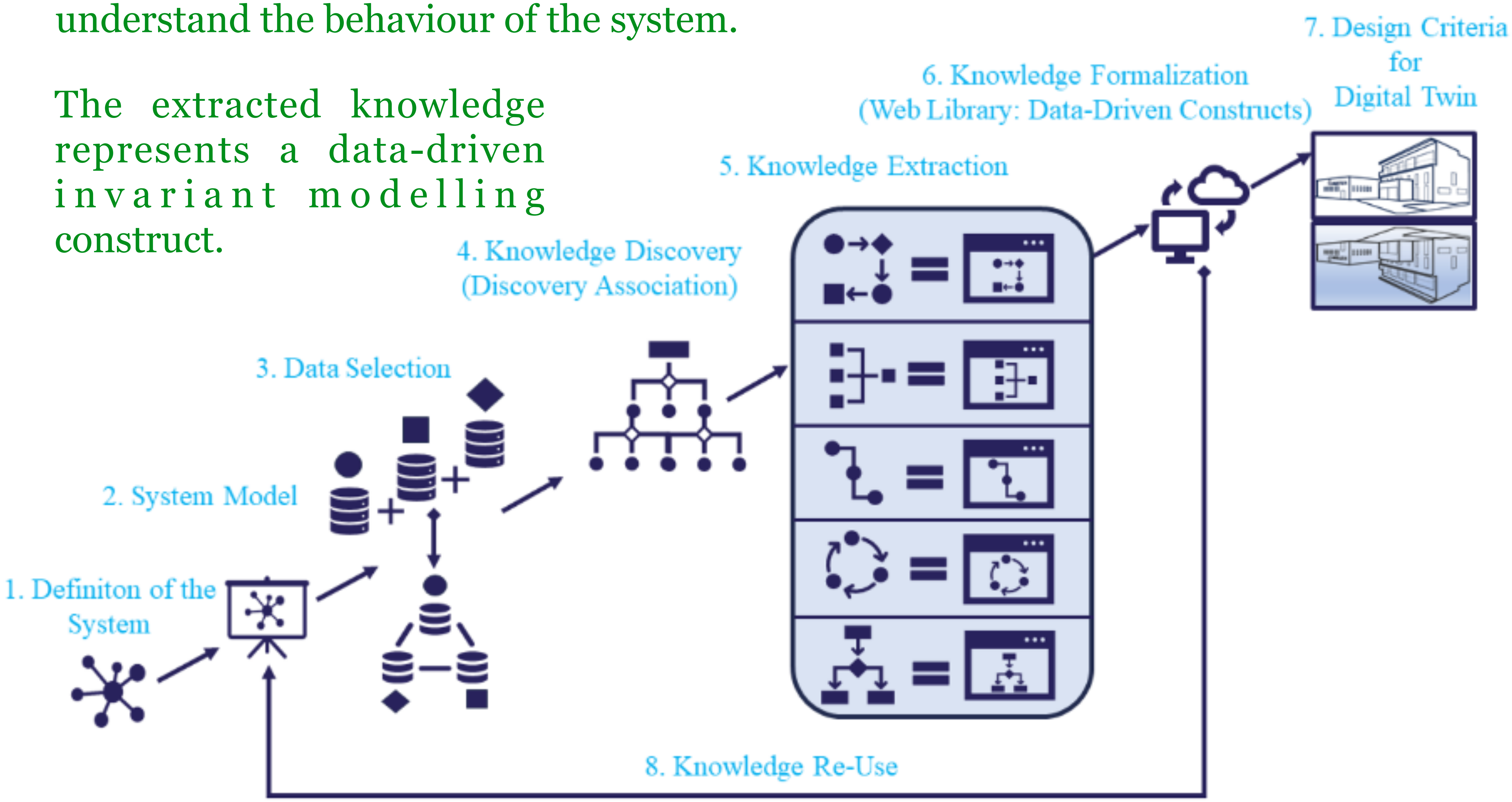
Research context - Image source [Semeraro, 2019]

To detect automatically from data invariant the modelling constructs.

DT4: Concetta SEMERARO contributions [Semeraro, 2020]

The associations can describe recurrent behaviours of the system and it can codify tacit knowledge that can be used to better understand the behaviour of the system.

The extracted knowledge represents a data-driven invariant modelling construct.



The Approach to Extract and to Formalize Data-driven Modelling Construct - Image source [Semeraro, 2020]

1) To identify the systems to analyse

2) To detail the representation of system

3) To enable the selection of data to analyse

4) To discover automatically associations and relationships among data.

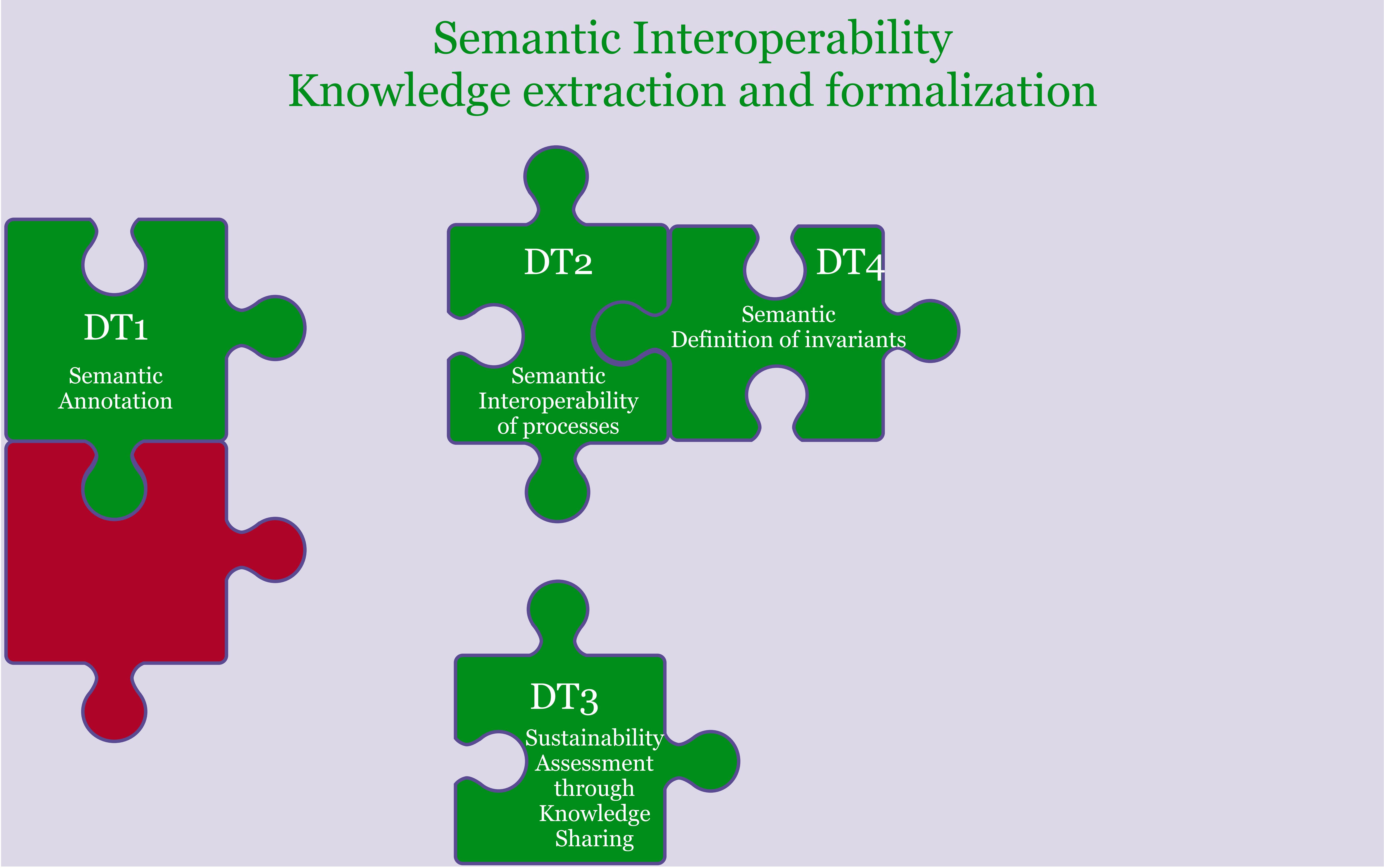
5) To extract knowledge from data and to define the physical meaning of the associations.

6) Formalization of Data-driven invariant modelling constructs

7) To design the virtual model of a system for realising its digital twin.

8) Reuse of the Data-driven invariant modelling constructs

Main scientific challenges direction



Results

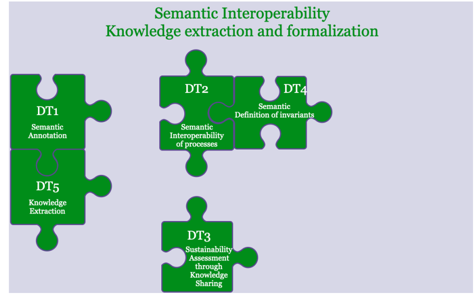
DT1: [Liao, 2013]
Semantic Annotation metamodel
Semantic annotation framework

DT2: [Pereira Detro, 2017]
Framework for customised process variants

DT3: [Eslami, 2019]
Three-Dimensional Model for Sustainability Assessment
Cubical model from the Knowledge integration

DT4: [Semeraro, 2020]
Formalized data-driven invariant modelling constructs to improve the smartness of manufacturing processes and products, involving CPS components.

PhD contributions



DT5: Mickael WAJNBERG contributions [Wajnberg, 2020]

Relational concept analysis: a versatile method for knowledge extraction
The thesis aims to develop the RCA approach, an extension of FCA, to extract association rules bypassing the presence of cycles in the object descriptions on any binarised multi-relational dataset.

Formal Concept Analysis (FCA)

A **Formal Context** is a triplet K = (O, A, I) where

Attributes

Object

K	sa	co	mc	te
de	x			x
le	x		x	
pe			x	x
tp		x		x

I is a binary relationship (O x A)

Derivation operation concept:

$$X' = \{a \in A \mid \forall o \in X, (o,a) \in I\} = \cap_{o \in X}\{a \in A \mid (o,a) \in I\}$$

The FCA aims to extract sets of objects with common attributes.

Formal Concepts

A pair C=(X,Y) belonging to P(O) x P(A) such that Y=X' and X=Y' is called a formal concept.

Concepts Lattices

Let K = (O, A, I) a formal context. Let us note:

- C_K the set of all the formal concepts of P(O) x P(A) and
- $\leq_K$ the relation of inclusion on the extensions of the concepts.

The partial ordered set (poset) $\mathcal{L}_K = (C_K, \leq_K)$ forms a complete finished lattice. It is called the **concept lattice** of context K [Ganter, 1999].

DT5: Mickael WAJNBERG contributions [Wajnberg, 2020]

Formal Concept Analysis exemple


Pepita (pe)


Demetra (de)


Talpi (tp)


Lea (le)

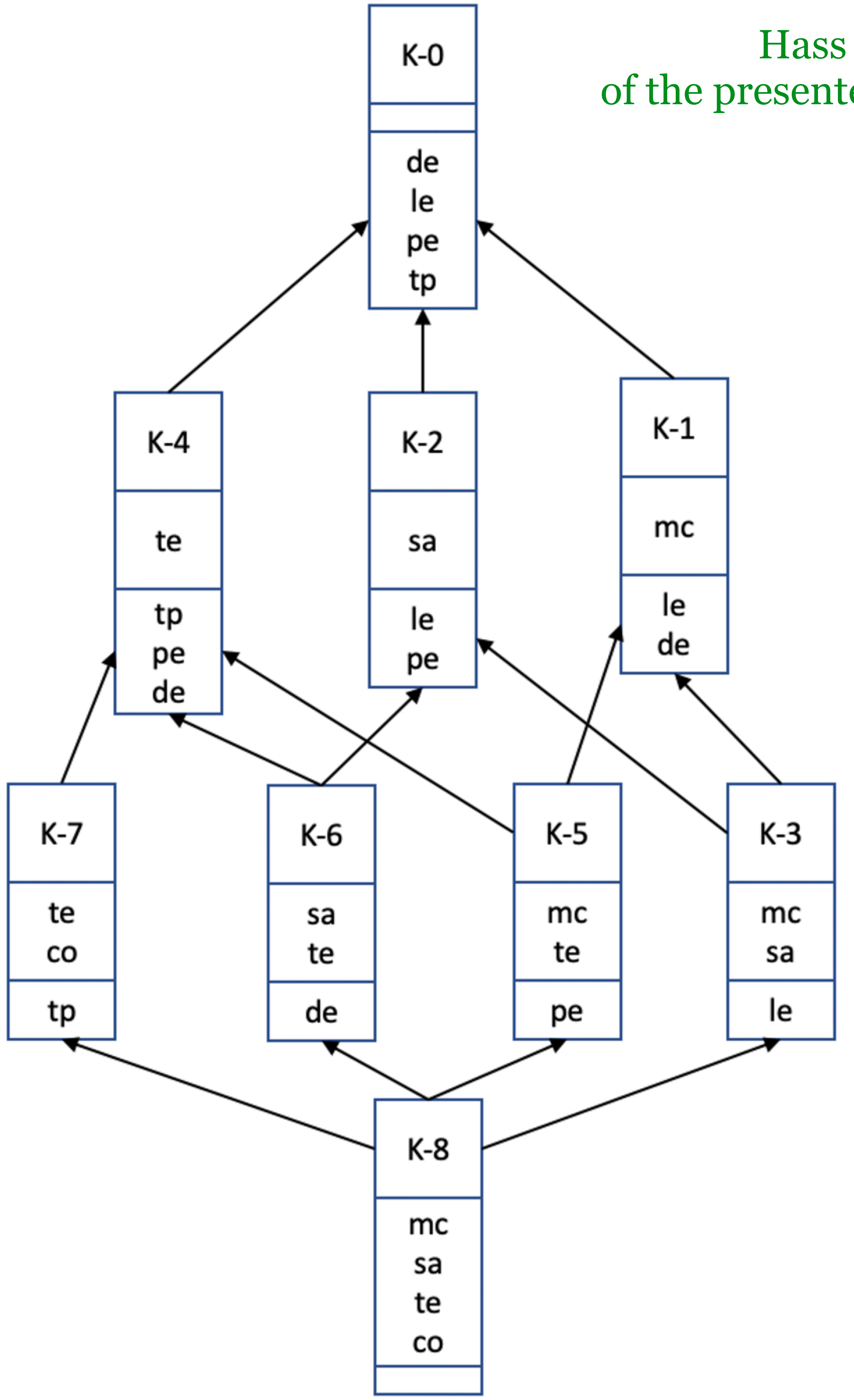
Properties:

- always being hungry (sa)
- always wanting cuddles (co)
- having a unique colour coat (mc)
- always protecting their territory (te)

The formal context K = (O, A, I) where
- O = { de, le, pe, tp }
- A = { sa, co, mc, te }
- I represented by the following unary table

K	sa	co	mc	te
de	x			x
le	x		x	
pe			x	x
tp		x		x

Derivation operation:
{le, de}’ = {sa}
{pe, de}’ = {te}
{te}’ = {tp, pe, de}
{mc, te}’ = {pe}



Hass diagram
of the presented formal context

DT5: Mickael WAJNBERG contributions [Wajnberg, 2020]

Relational Concept Analysis (RCA)

The RCA [Huchard, 2002] **extends the FCA** to relational data compatible with the **entity-association model** [Chen, 1976]. Such a model considers **binary relations between objects**.

Formal Contexts

$K_i = (O_i, A_i, I_i)$

Source

$K_j = (O_j, A_j, I_j)$

Target

Relation $\mathbf{R_{i,j,k}}$ is a binary relation subset of O_i x O_j

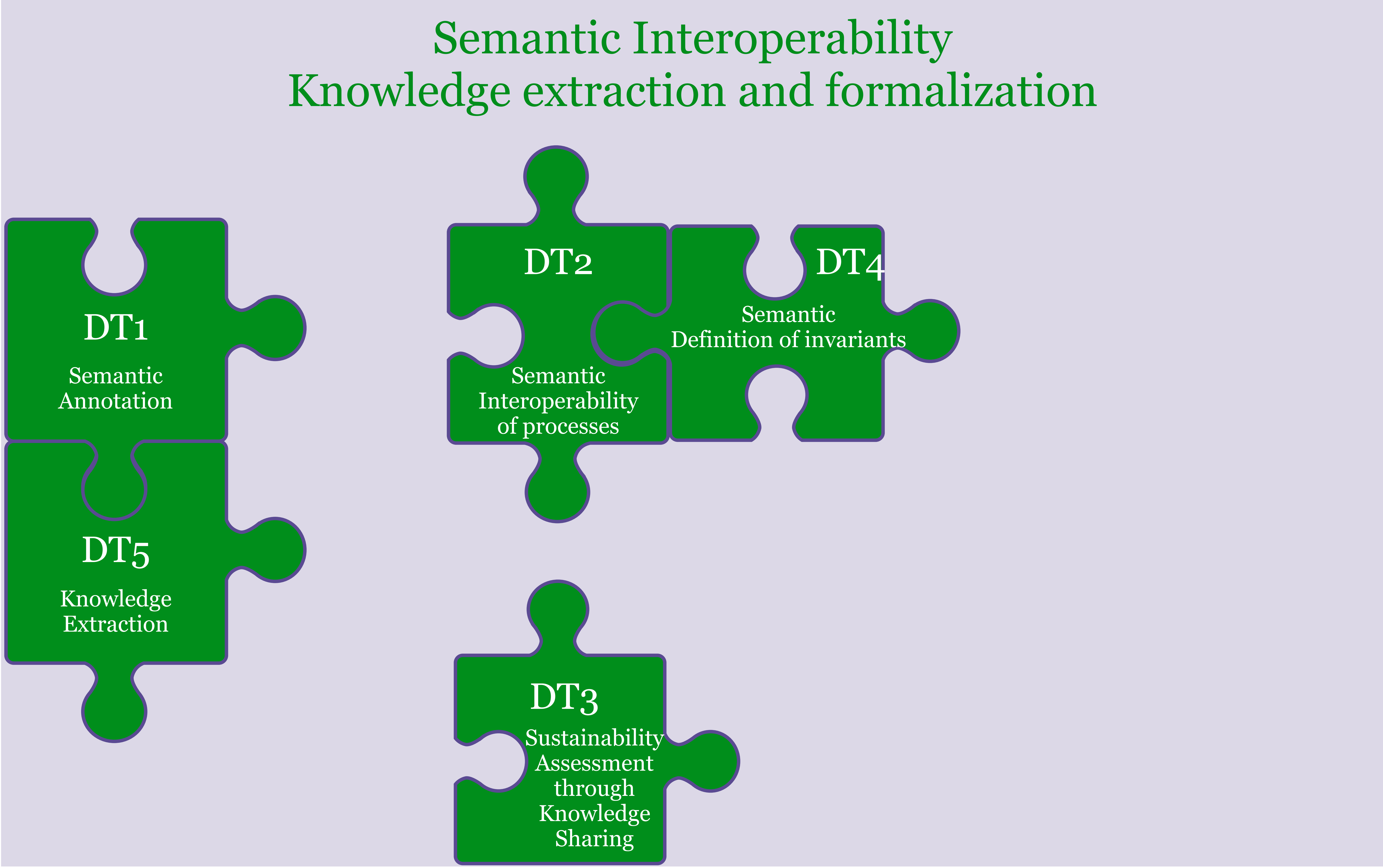
A relational family of contexts (RFC)

RFC is a pair (K,R) such that:
- K is a set of formal contexts $K_i = (O_i, A_i, I_i)$
- R is a set of relations $R_{i,j,k}$ belongs O_i x O_j for i, j belonging to {1, ..., |K|}

Results:

- redefinition of a formal concept within a framework of contexts extended in an iterative way
- demonstration that the rules of association could be extracted and defined without any attribute to be solved recursively.

Main scientific challenges direction



Results

DT1: [Liao, 2013]
Semantic Annotation metamodel
Semantic annotation framework

DT2: [Pereira Detro, 2017]
Framework for customised process variants

DT3: [Eslami, 2019]
Three-Dimensional Model for Sustainability Assessment
Cubical model from the Knowledge integration

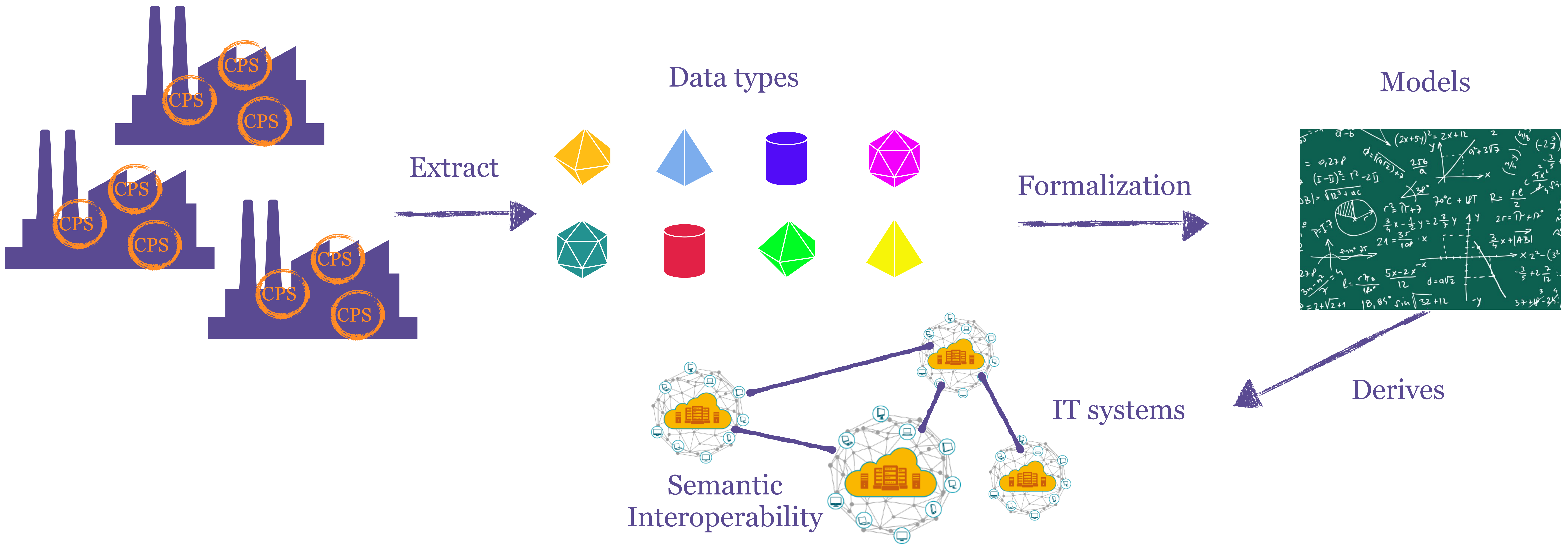
DT4: [Semeraro, 2020]
Formalized data-driven invariant modelling constructs to improve the smartness of manufacturing processes and products, involving CPS components.

DT5: [Wajnberg, 2020]
The thesis improves the RCA to allow the extraction of association rules to bypass the presence of cycles in the object descriptions on any binarized multi-relational dataset and at eliminating redundancy in the rules that could arise due to the inherent relationship between relational characteristics.

PhD contributions

My research project proposal - Context

Formal methods for extracting and reusing knowledge from heterogeneous sources for semantic interoperability of distributed architectures in a Factories of the future context.



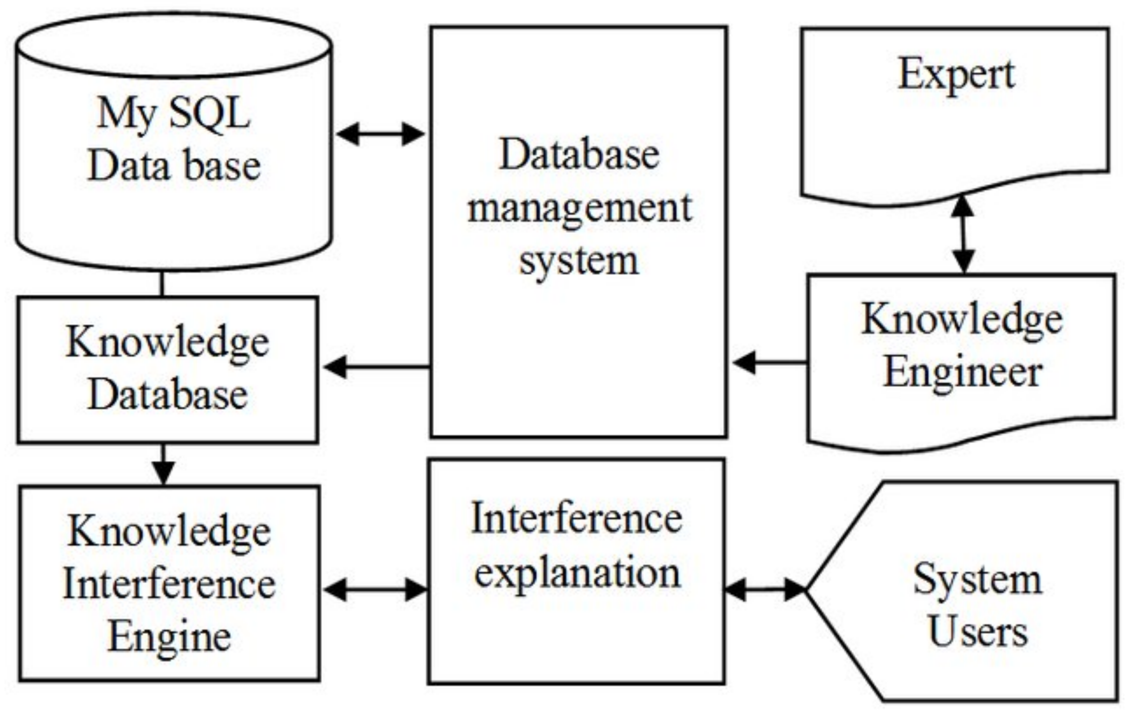
to ensure their common understanding and to facilitate their interoperability by minimizing semantic losses

My research project proposal - Positioning

Formal methods for extracting and reusing knowledge from heterogeneous sources for semantic interoperability of distributed architectures in a Factories of the future context.

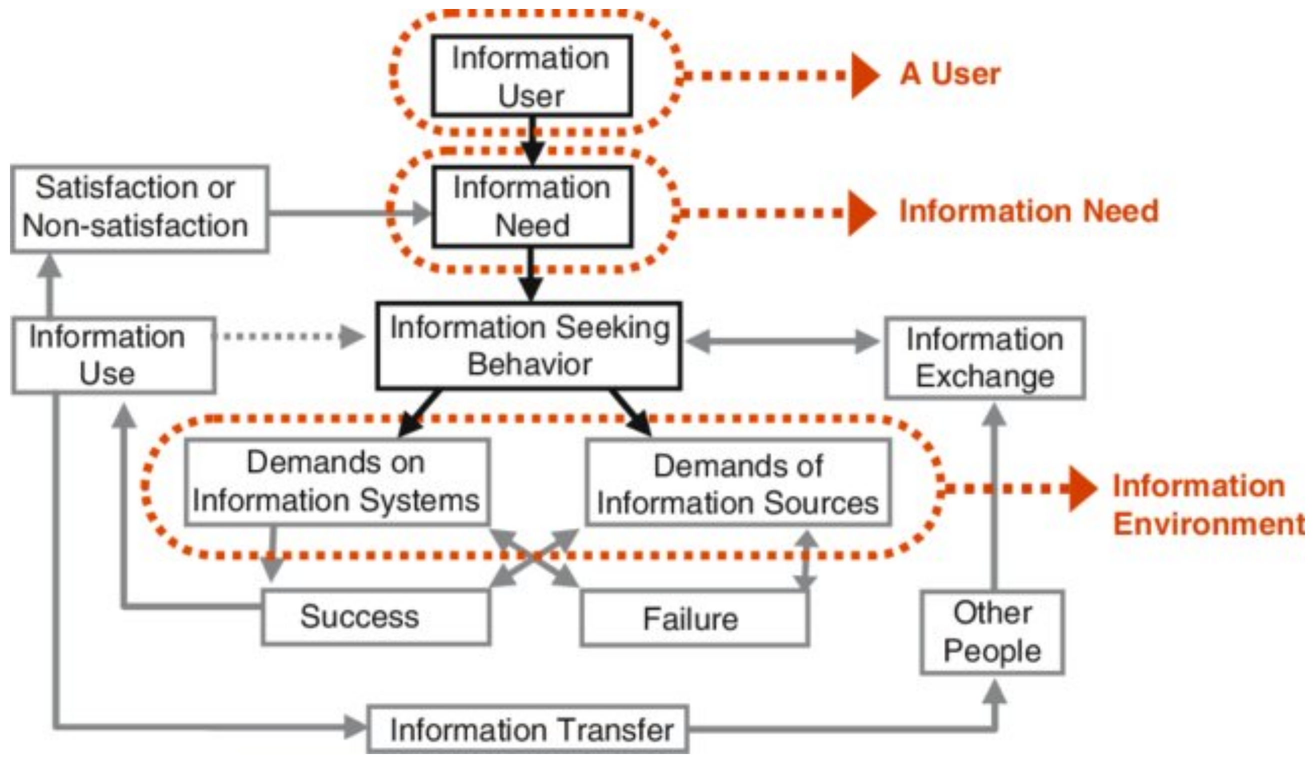
The engineering of interoperable systems [Ramos, 2011] and [Morel, 2003] will be used that consists of relying on different types and levels of abstraction or models.

Structure



[Thongmark, 2016]

Behaviour



[Wilson, 1981]

Constraints

Interoperability protocol(s)

Specific requirements of the system domain

[Zdravković, 2016]

Research project description - Objective

Objective
- to model data from heterogeneous sources and create knowledge from them
- to study the problems posed by model-driven engineering in **cooperative systems**

Complex systems [Camarinha-Matos, 2014]

Networks of CPSs

Specifying such a network involves an **interoperation paradigm** [Fisher, 2006].

The ability of these components to interoperate

To make available languages and modeling tools adapted to each project of systems

This challenge has two dimensions:
- on the one hand, the capacity of modeling to provide tools for business processes, which requires the definition and formalization of their invariants;
- the study of the conditions of use of the models in practice, which is always evolving and uncertain.

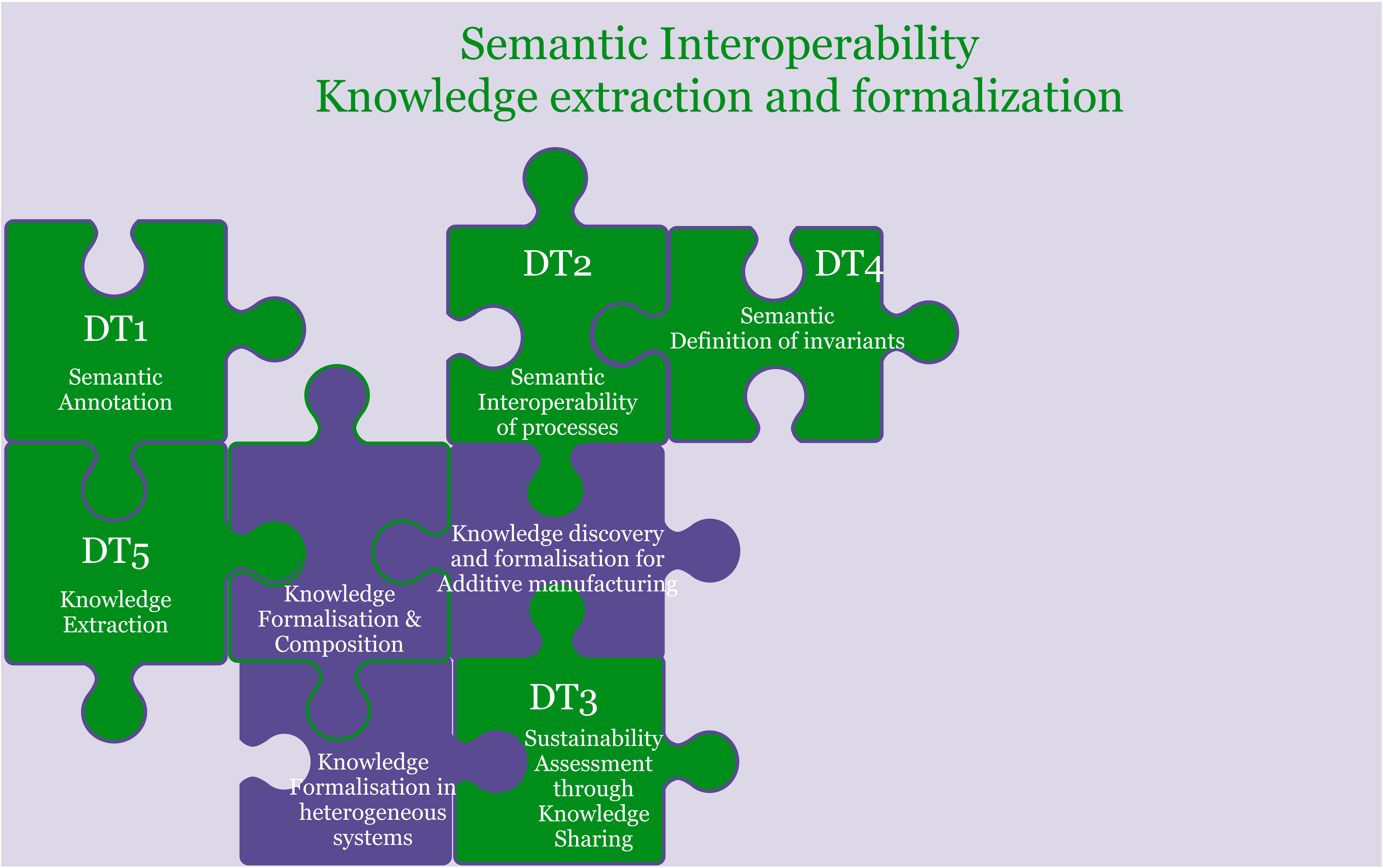
Research project description - Scientific locks

Which approach would be the most suitable, taken in account all the domain constraints, to model data from heterogeneous sources?

How to discover and formalize knowledge from the heterogeneous data?

How to solve semantic interoperability problems posed by model-driven engineering in heterogeneous and cooperative willing systems?

Main scientific challenges

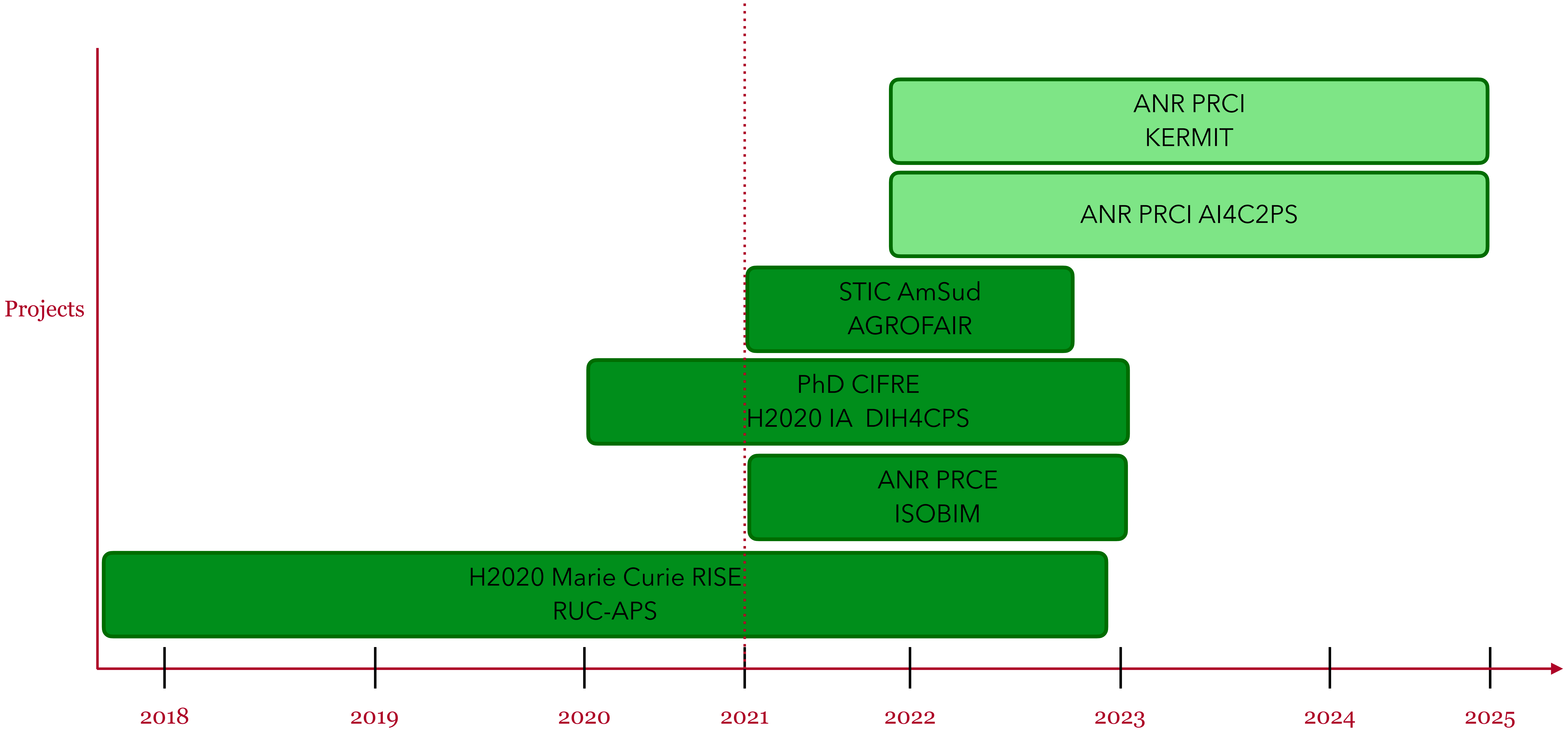


Methodologies for the clarification and formalisation of knowledge based on intelligent/connected objects in the factory of the future.

Automatic knowledge explicitation methodology

PhD contributions
My research Project

Research project positioning in relation to the research environment



Local context: CRAN - ISET - S&O2I

Regional context: The Grand Est region "Industry of the Future" plan Collaboration with the Epinal Chamber of Commerce and Industry and its incubator the Quai Alpha.

National context: The ANRT 2020 calls for proposals in the axe 5.2 - Artificial Intelligence (knowledge extraction, formalisation and management)

International context: The issue of Factories of the future and Industry 4.0 at European level Horizon Europe 2021-2027.

Teaching project

Industry 4.0 aspects:

The QLIO department asked and obtained 500,000 euros funding for the construction of a lean 4.0 atelier. This is allowing us, with the support of the whole pedagogical team, to structure an innovative didactic path.

As far as my contribution is concerned, I have structured, in the 4 modules in which I'm responsible a path that starts from the understanding of basic computer science with the generation of data to the structuring of information and finally to the formalisation of knowledge related to the enterprise systems.

Knowledge creation and formalisation aspects:

In the Telecom Nancy computer science engineering school I will deepen the issues related both to the formalisation of knowledge and Industry 4.0 specifically:

Enterprise 4.0: where I teach CPS programming and how to manage the data flow generated by smart sensors to create new knowledge formalised in taxonomies that can be used in the future. In addition, in the same module I will intensify the teaching of how to program a blockchain and a panel of possible use.

Artificial Intelligence: where, in collaboration with several teachers, we help students to understand the basic algorithms to understand the discipline itself and push them to search for new solutions related to the world of research.

Integrated Enterprise Management: in this module I introduce students to the knowledge of methods of model creation and model validation against specific software such as ERP.

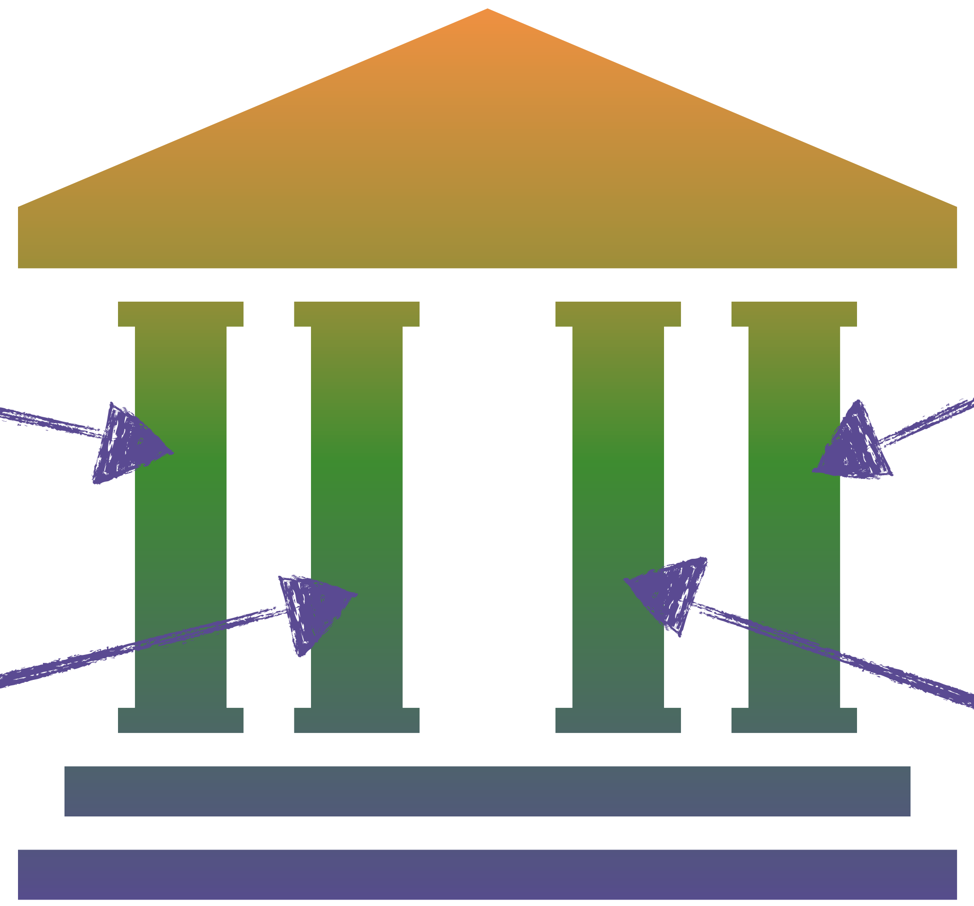
Conclusion

How I see the life of a Researcher/Mentor stands on various concepts :

Constant evolution

Scientist for research
- Publications
- Supervision of research students

Working in Team
- Pedagogical team
  - Team Teaching Units, Diploma Reflection
- Research team
  - local, national and international projects
- Research groups
  - Involvement in SAGIP, IFIP, IFAC...



Pedagogy to help people to improve
- Research training through research
- Project student supervision
- Teaching by projects

Collective to keep the structure alive
- Search for funding
- Participate in the bodies of research and teaching institutions

Basic Abilities
- Curiosity and Critical Mind
- Lifelong
- Openness and listening

Formalisation models and knowledge extraction : Application to heterogeneous data sources in the context of the Factory of the Future

THANKS FOR THE AUDIENCE!

NOW I'M HERE FOR YOUR INTERESTING QUESTIONS

UMR 7039

Hubert Curien