

# Multi-relational and Concept Analysis based Knowledge extraction in the Industry 4.0: A systematic mapping

Leutwyler Nicolás <sup>\*,\*\*,\*</sup> Lezoche Mario <sup>\*</sup> Torres Diego <sup>\*\*,\*</sup>  
Panetto Hervé <sup>\*</sup>

<sup>\*</sup> *University of Lorraine, CNRS, CRAN, F54000 Nancy, France,  
(e-mail: name.lastname@univ-lorraine.fr).*

<sup>\*\*</sup> *LIFIA, CICIPBA-Facultad de Informática, UNLP, La Plata 1900,  
Argentina, (e-mail: name.lastname@lifia.info.unlp.edu.ar)*

<sup>\*\*\*</sup> *Dto. CyT, UNQ, Bernal 1876, Argentina*

---

**Abstract:** Smart Enterprises, Smart Manufacturing, and Cyber-Physical Systems are gaining traction in many industry areas. On top of that, the amounts of available data grow rapidly, and organizations are eager to exploit their advantages. To accomplish that, it is mandatory to have a wide variety of methods and algorithms for knowledge extraction in order to fit the different needs and problems of the industry. In this study, we review and dissect the current state of the art in knowledge extraction applied to smart enterprises, smart manufacturing, and cyber-physical systems. More specifically, we provide a classification of the characteristics of the available methods in the literature according to their applications, and point out areas of improvement.

*Keywords:* Enterprise interoperability, AI-based enterprise systems, Systems interoperability, Cyber physical system, Smart factory

---

## 1. INTRODUCTION

Industry 4.0 is pushing organizations and enterprises alike to adopt practices such as the usage of data in order to optimize processes [Torres et al. \(2021\)](#); [Romero et al. \(2020\)](#). Additionally, the growth of Internet of Things (IoT) and Cyber-physical systems (CPS) over the last two decades leaves large amounts of heterogeneous data at their disposal. In this context, *Knowledge Extraction* (KE) is an essential step to consider not only in software development, but also in all the stages of industrial procedures. Moreover, the sharing of *knowledge* between different pieces of software, and enterprises (i.e., semantic interoperability defined by Sowa in [Sowa \(2000\)](#)) is a crucial step for certain processes. In addition, these areas leave room for multiple *heterogeneous* and *linked* sources, which consequently present other challenges to the area of data-mining. The name given to the process of doing KE from heterogeneous and linked sources is multi-relational data mining (MRDM), as described in [Džeroski \(2003\)](#).

In this context, KE has become a widely studied topic in the academy, leading to a great number of papers being written about it each year. Consequently, many of these academic works have been successfully implemented in the industry. In particular, there are a plethora of methods that are useful to extract different types of knowledge, e.g., Formal Concept Analysis (FCA) for extracting a Lattice from of a Formal Context, distance-based clustering methods such as k-means, ontology creation techniques, etc. Since these methods have their own specific theoretical advantages, and limitations, it is interesting to identify

where they are being used in order to recognize possible areas of improvement. More specifically, in this work we review the state of the art in the following three areas: Smart Enterprise (SE), Smart Manufacturing (SM), and CPS. Furthermore, understanding the characteristics and relationships of the KE methods between each other is necessary to make a better use of them.

To our knowledge, in terms of related and relevant works, there are three articles that are worth mentioning. Firstly, [Kalfoglou et al. \(2004\)](#) review FCA as a tool for *knowledge technologies*. They analyze how useful FCA is as a tool for several applications in different domains, by combining it with other knowledge technologies (i.e., a form of interoperability). In that review, considering that it has been published in 2004, the authors arrive to the conclusion that FCA still needs further techniques to be considered directly applied by knowledge engineers. Secondly, [Polovina et al. \(2016\)](#) review the combination of FCA and Conceptual Graphs (CGs) for highlighting the gaps in an Enterprise Information Model. And finally, [Poelmans et al. \(2013\)](#) review FCA as a general tool in the fields of software, web mining, text mining, knowledge discovery, and ontology engineering.

All these works are either specific to a single topic, general but not contextualized in the mentioned fields (i.e., SE, SM, and CPS), or do not provide a specific methodology for reproducing the review. Therefore, in this paper we aim to provide a sufficiently general review, providing a specific and systematic methodology (see Section 2) that

allows it to be reproducible in order to find the gaps in the aforementioned fields.

The article is structured in the following way: in section 2 we present the methodology used to conduct the review. In section 3 we discuss the obtained results. Later, the method used to map and analyze the different characteristics in each field is introduced, and the results are presented in section 4. Finally, section 5 unfolds a more general discussion, presents the work to be done in order to improve the review, and the future research directions.

## 2. METHOD

This work falls into the classification of Systematic Literature Mapping (SLM). An SLM aims to identify relevant studies, evaluate their quality, and qualitatively (or quantitatively) summarize the findings [Robinson and Lowe \(2015\)](#). Additionally, it is necessary to identify the tools and engines used to select the studies. The search carried out in this work is based on the guidelines presented by [Petersen et al. \(2015\)](#). In short, the methodology consists of three steps (1) define research questions and keywords, (2) apply search on databases, (3) remove duplicates, (4) apply criteria on abstract and full-text reading, (5) analyze the results. One of the reasons why it is interesting to do an SLM is that it leaves evidence on how to replicate the search using the criteria mentioned in it.

### 2.1 Research Questions

The goal of this mapping work is to identify how methods are utilized in the scientific literature of SE, SM, and CPS. This leads us to the following research questions:

- RQ1: When and where the articles have been published?
- RQ2: In which domains the articles are applied?
- RQ3: What are the problematics the papers aim to solve?
- RQ4: What are the methods utilized for MRDM or Concept Analysis in SE, SM, and CPS respectively?
- RQ5: How articles measure performance of the methods they propose?
- RQ6: What type of knowledge representation the methods extract or use?
- RQ7: How do the authors evaluate their solution?
- RQ8: What type of contribution the article provides?

This information, in combination with best practices defined in [Petersen et al. \(2015\)](#), is later used to propose a categorization and classification of the existing methods (see subsection 2.6).

Next, the search protocol, which has been followed by the first author and later reviewed by the rest of the authors, is presented.

### 2.2 Keywords and search strings definition

*Important concepts:* The following concepts will be used throughout the rest of the paper,

- *Data Mining* (DM) is the process of discovering knowledge or patterns from massive amounts of data [Han \(2009\)](#).

- *Knowledge Discovery in Databases* (KDD) is an automatic, exploratory analysis and modeling of large data repositories. KDD is the organized process of identifying valid, novel, useful, and understandable patterns from large and complex data sets [Maimon and Rokach \(2005\)](#).
- *Knowledge Extraction* is the creation of knowledge from structured (relational databases, XML) and unstructured (text, documents, images) sources [Unbehauen et al. \(2012\)](#).
- *Information Retrieval* (IR) is finding material (usually documents) of an unstructured nature (usually text) that satisfies an information need from within large collections (usually stored on computers) [Manning et al. \(2008\)](#). This definition differs to the Data Mining one in the sense that in IR the goal is to gather information from an unstructured source in order to *satisfy* some *known* need, e.g., browsing in google. However, it is important to notice that IR methods might use Data Mining algorithms in their internal process.
- We use *Concept Analysis* (CA) as the ensemble of the Formal Concept Analysis method, introduced by Wille [Wille \(1982\)](#), and its extensions, e.g., Relational Concept Analysis [Rouane-Hacene et al. \(2013\)](#); [Wajnberg et al. \(2021\)](#), Polyadic Concept Analysis [George Voutsadakis \(2002\)](#), etc.

Considering these definitions, it is our understanding that DM, KDD, and KE are three names for the same concept. On the other hand, IR, is a concept more related to the overall process that might or might not use DM algorithms in it, but also includes steps like data warehousing.

In order to identify keywords and define search strings, this work will rely on Population, Intervention, Comparison, and Outcomes (often referred by its acronym PICO) i) population: knowledge extraction or information retrieval studies, either using some form of concept analysis (e.g., FCA, RCA, Fuzzy FCA) or MRDM method, applied in the SE, SM, or CPS context, ii) intervention: algorithms, methods and tools, iii) comparison: different ways of dealing with the challenges imposed by the SE, SM, and CPS industries, iv) and outcomes: a mapping of the types methods in terms of their usages in the three already mentioned fields.

Thus, the identified keywords are divided in four sets, meaning that a paper is considered inside our search if it contains *at least* one keyword of each set.

- Set 1: Scoping the search for our domain with “Knowledge Extraction”, their synonyms, and “Information Retrieval”. In this set, the permutations of keywords, such as “Knowledge Mining” or “Data Retrieval”, are also considered by querying “Knowledge” OR “Extraction” OR “Data” Or “Mining”, and so on. This approach also includes papers only mentioning “Knowledge”, or paper including combinations we are not looking for such as “Mining Extraction”, but those are filtered later.
- Set 2: Related with the population, MRDM, “concept analysis” and their synonyms.
- Set 3: SE, SM, CPS, and their synonyms.

- Set 4: Terms related with the intervention, “algorithm”, “method”, and “tool”.
- Set 5: Search terms related with the classification, e.g., semantic, semantically, interoperability, interoperable.

### 2.3 Database selection

For the search, this work uses ACM, IEEE\_Xplore, Scopus, and Web of Science. The selection of databases is based on the recommendation in Petersen et al. (2015) i.e., it is sufficient to use IEEE and ACM plus two indexing databases. The query shown in Table 1 has been adapted and ran in each of the mentioned databases, obtaining the results depicted in Table 2. Zotero, a reference management tool, was used in order to delete duplicates and to manage the large amount of references. This study has been conducted during 2022, and more particularly, the first quarter of 2022 and before was considered during the search. Any article published after April 2022 has been ignored.

Search query
(method OR algorithm OR tool) AND (knowledge OR extraction OR data OR mining OR information OR retrieval) AND (rca OR "concept analysis" OR mrdm OR "multi relational") AND (semantic* OR interop*) AND (cps OR "cyber physical" OR sensor* OR sensing OR enterprise* OR industry OR smart* OR IoT OR "Internet of Things" OR cloud OR manufacturing OR logistics)

Table 1. Base query used in each search engine.

Database	Total	Unique	Duplicated
ACM	27	22	5
IEEE	26	23	3
Scopus	66	66	0
Web of Science	24	5	19
<b>Total:</b>	<b>143</b>	<b>116</b>	<b>27</b>

Table 2. Number of papers obtained on the queries execution on each database.

### 2.4 Study selection and quality assessment

From the results, firstly, we selected papers based on the titles and abstracts in a step named first screening. Secondly, we included or excluded papers based on a full-text reading, in a step named second screening. The criteria used for the selection can be seen in Table 3. It is important to mention that, when in doubt during the title and abstract reading, articles were taken to full-text reading instead of directly excluded. One of the threats to the reliability of the review is that the selection was conducted by only one author, and hence it could have missed some articles (more on that in section 5). For this step, we used the *Scolr* tool, which allows conducting both screenings.

The number of papers included/excluded in each screening step can be seen in Figure 1. After the full-text review, it became clear that some papers, though interesting, were out of the scope of this study, because they were either not

<i>C1</i>	
Inclusion	Exclusion
Papers written in English	Papers not written in English
Conference Papers	Literature reviews
Journal Papers	Non peer reviewed papers
Book Chapters/Sections	Proceedings
About MRDM or About Formal Concept Analysis	
<i>C2</i>	
Inclusion	Exclusion
Applied to IoT or CPS	Anything not satisfying the inclusion criteria
Applied to Enterprises	
Applied to Manufacturing	

Table 3. Criteria for selecting papers.

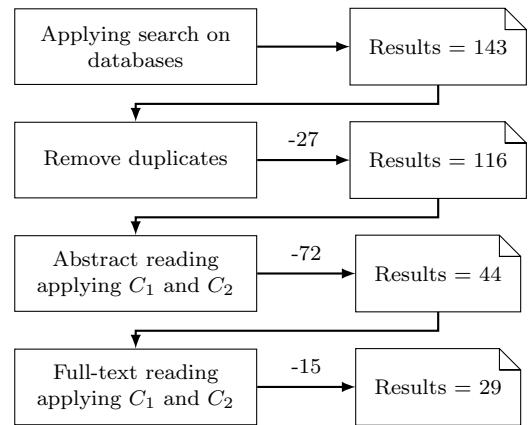


Fig. 1. Number of articles included in each step of the selection process

applied, or applied in a field other than the three that we are interested on.

The quality assessment consisted on comparing the final 29 primary studies with an independent set of papers we knew that ought to be in the final set Loia et al. (2010); De Maio et al. (2017); Morozov et al. (2018). Also, the following questions were answered to assess the quality of the selected articles:

- Is the application of the methodology clearly stated?
- Are the algorithms or architectures defined?
- Are experiments or discussions conducted to evaluate the results?

Thus, studies not presenting the methodology, algorithms, architectures or not doing experiments or discussions on the results were excluded. The final selection of articles is presented in Table 4.

### 2.5 Data extraction

The data extraction from the identified primary studies has been done considering the Table 5. Data extraction fields have a *data item* and a *value*.

Article ID	Article
1	Li and Li (2015)
2	Bordes et al. (2014)
3	Akmal and Batres (2013)
4	An and Li (2017)
5	Baxter et al. (2021a)
6	Buchmann et al. (2011)
7	Loia et al. (2010)
8	Cretan et al. (2021)
9	Fallon and Polovina (2016)
10	Cui et al. (2007)
11	Sithole and Marshal (2019)
12	Shin et al. (2013)
13	Hildebrandt et al. (2019)
14	Abuzaid et al. (2021)
15	Watmough (2014)
16	Polovina et al. (2019)
17	Moujabbar and Ramdani (2013)
18	Baxter et al. (2021b)
19	Mashhadi et al. (2015)
20	Galkin et al. (2016)
21	De Maio et al. (2017)
22	Yang et al. (2022)
23	Morozov et al. (2018)
24	Zhang et al. (2012)
25	Wajnberg et al. (2018)
26	Louati et al. (2016)
27	Siddavatam et al. (2017)
28	Gharbi and Mezni (2020)
29	Miguel-Rodríguez et al. (2016)

Table 4. Articles by ID (to be used in Hasse sub-lattices.)

## 2.6 Analysis and classification

The fields extracted from each selected primary article are tabulated and visually illustrated in Section 3. The papers were grouped and counted by each of the fields, and since some of them are multiple (e.g., methods) the total sum is greater than the total amount of papers reviewed.

For the field applications, while reading the selected articles, there were many types of applications to assess e.g., R&D data modelling, manufacturing process ontology creation, Smart Cities data processing, etc. For the purpose of this paper, we group them by their *most precise* domain, which is either SE, SM, or CPS. More precisely, articles are classified in SM if they are applied in an SM area such as manufacturing ontology processing, even if it can be argued that SM is inside SE. The same occurs with CPS. Furthermore, the analysis was conducted considering the groupings of the three categories SE, SM, and CPS,

Data item	Value	RQ
<i>General</i>		
Study ID	Integer	
Article Title	Name of the article	
Year of Publication	Calendar year	RQ1
Publication Venue	Name of publication venue	RQ1
<i>Process</i>		
Application	SE   SM   CPS	RQ2
Problematic	Lack of a method   Lack of a model   Lack of a specification   Lack of data   Many stand-alone methods   Poor data accessibility   Existing methods' performance	RQ3
Method	FCA related   Ontology related   Supervised AI   Traditional algorithm	RQ4
Performance	Big-o complexity discussion   Output informal analysis   Informal discussion   Table and graphs results   Does not provide any	RQ5
Output of the method	Ontology   Knowledge Graph   Lattice Related   Frequent Pattern	RQ6
Evaluation strategy	Accuracy   Effectiveness   Efficacy   Robustness   Performance Technical   Feasibility Operational   Feasibility Learning Capacity   Reliability   Validity   Scalability   Utility	RQ7
Type of contribution	Algorithms   Architecture   Complete Software   Informal implementation   In-depth implementation   Model Deployed   Math functions	RQ8

Table 5. Data extraction divided in two sub-categories: general and process.

allowing us to draw conclusions about their different characteristics.

The *problematics* field that we extracted for this study is related to the question “what did authors considered that they were addressing with the article’s solution?”. In particular, 7 different types have been used: lack of a method, lack of a model, lack of a specification, lack of data, lack of a general method (i.e., many stand-alone methods), poor data accessibility, and existing methods’ performance.

In terms of the methods, the classification used for the selected articles was conducted similarly to the applications field. Several methods were found e.g., FCA, Fuzzy FCA (FFCA), Relational Concept Analysis, Bayesian Belief Network, direct algorithms from relational databases to knowledge graphs, etc. Considering this, we counted the articles using the following general method names: *FCA related*, *Ontology related*, *Supervised AI*, and *Other*. *Other* represents the algorithms that directly extract patterns or knowledge from databases, without using any of the aforementioned methods. In particular, most of the *Other* occurrences were multi-relational data-mining algorithms using *tuple-id-propagation*.

This study also considers the way authors evaluated performance. The classification terms used are i) *Big-O complexity discussion*: when the evaluation of performance is based on the big-O asymptotic notation. ii) *Table and graphs results*: when there are experiments considering performance, and the authors clearly present the results in tables or graphs. iii) *Output informal analysis*: when there is an analysis of the output. Typically, this happens when the authors say they did experiments but only loosely present some results and speak about them without presenting a graphic or a table. iv) *Informal discussion*: when the authors only provide a discussion without much detail. v) *Does not provide any*: when there is no mention of performance.

To classify the most used knowledge representation types in the articles, we look into the input and output formats of the solutions proposed. On the one hand, the types of input of general data mining methods are very heterogeneous, ranging from formal contexts, and structured data, to data streams and unstructured data. On the other hand, outputs are usually less heterogeneous, and therefore we decided to use the categories related to the most common knowledge representation types i) Ontology, ii) Knowledge Graph, iii) Lattice related, iv) Frequent Pattern.

For the evaluation strategies of the articles' solutions, this work uses the taxonomy defined in [Prat et al. \(2015\)](#). Particularly, the evaluation types we used are *accuracy, effectiveness, efficacy, robustness, performance, technical feasibility, operational feasibility, learning capacity, reliability validity, scalability, and utility*.

Finally, the type of contribution was classified considering the type of contribution the paper presents. Does it provide a model? Does it present an entire suit of functions to implement and directly deploy it? Is it a high level discussion about a meta-solution? With these questions in mind, the labels we used were: i) Deployed: when the solution the article proposes is complete in the sense that they not only present and discuss the algorithms, but also state how it has been implemented, in which language, and how it was deployed. ii) Algorithms: when the solution presents the algorithms in pseudocode. iii) Architecture: when the proposed method is presented in the form of an architecture, not necessarily as a traditional algorithm. iv) Complete Software: articles whose solution includes not only the algorithms and architecture description, but also the discussion about the UI implementation and integration in the context of the particular study. v) Informal implementation: articles with an informal discussion about how to implement a particular function or idea without explicit algorithms. vi) In-depth implementation: articles with a more formal discussion about how to implement a particular function, but still without explicit algorithms. vii) Model: articles presenting models instead of methods. viii) Math functions: articles based on how a particular function would improve a certain method.

## 2.7 Validity evaluation

According to Petersen et al. [Petersen et al. \(2015\)](#), in software engineering, the following types of validity should be considered: descriptive, theoretical, generalizability, and interpretive validity.

*Descriptive validity* Descriptive validity is how accurately and objectively observations are described. Qualitative works have a greater threat to descriptive validity than quantitative ones. In order to deal with this threat, we created a data extraction form to help objectify the recording of data (Table 5).

*Theoretical validity* The ability of being able to capture what we intend to capture is called *theoretical validity*. For instance, biases can lead us to select articles we should not and to not select others that we should.

While searching, studies could have been missed. For example, two searches in the same field, could yield different sets of articles. To address this threat, backward snowball sampling of articles could be done after the full-text reading [Jalali and Wohlin \(2012\)](#), but addressing it is outside the scope of this paper.

Bias is also a threat in the phase of *data extraction and classification*. To address this problem, it is useful that one researcher performs the extraction while the other reviews it. However, as mentioned in [Petersen et al. \(2015\)](#), given the fact that the process involves human judgement, the threat is unavoidable.

*Generalizability* Petersen and Gencel [Petersen and Gencel \(2013\)](#) introduced a distinction between external and internal generalizability, i.e., between groups or organizations, and within a group, respectively. Internal generalizability does not represent a major threat because of the wide range of articles following the same strategies. In terms of the external one, it is not a major threat as well because the approach takes into account general metrics that can be applied to other fields of study (see Table 5).

*Interpretive validity* Interpretive validity refers to the idea of the conclusion being reasonable given the data. A threat to interpretive validity is, again, researcher bias. In our case, the first author's major field of expertise is algorithms and efficiency, and that could lead to a bias in the interpretation. However, the main field of work of the rest of the authors are knowledge representation and extraction, semantic web, and industrial engineering respectively, helping to reduce the threat.

*Repeatability* Repeatability refers to how repeatable the process is. It demands a thorough reporting of the research process. We reported the SLM process followed, and moreover, we explained the different possible threats and our actions to reduce them. Additionally, we provide the list of obtained articles in the dataset [Leutwyler et al. \(2022\)](#).

## 3. RESULTS OF THE REVIEW

In this section, we illustrate and present the findings of the mapping using the resulting data [Leutwyler et al. \(2022\)](#) after the data extraction.

### 3.1 Frequency of Publications (RQ1)

Papers have been counted by year and publication venue in Figure 2 and Figure 3 respectively. The years with more articles were 2016 and 2017 with 4 papers each, and most

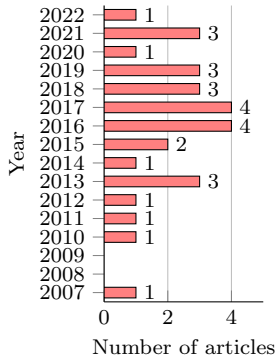


Fig. 2. Articles per year

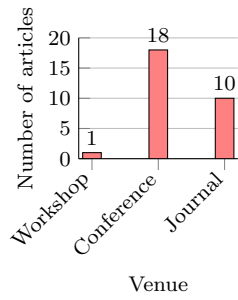


Fig. 3. Articles per publication venue

of the papers were published after 2014. Moreover, the years 2008 and 2009 did not have occurrences.

The growth in papers after 2014 showed that there is still an increasing interest in this topic in the scientific community. Additionally, one of the hypothesis we have about the reason this happens is the increase of application papers thanks to the theoretical foundations being more solid in the last years.

In terms of publication venue, most of the papers were published in conferences, about a third of them in journals, and only one in a workshop. Particularly, both the journals and conferences are heterogeneous, ranging from computer-science specific to engineering and manufacturing ones.

### 3.2 Applications (RQ2)

The applications of the articles were done by categorizing them in the three categories, SE, SM, and CPS. In particular, the number of papers in each of the categories is shown in Figure 4. There were 18 papers in the domain of SE, 2 in SM, and 9 in CPS. In addition, the subclassification of the highlighted CPS slice can be seen in Figure 5, where *General* are the articles in the CPS field without any particularity, SC are the articles about smart cities, and IoT are the articles working on the data transmission layer in CPS. In this category, there are 4 papers in General, 2 in smart cities, and 3 in IoT.

The most important suggestions of Figure 4 is the difference in the amount of papers from SE, doubling the ones in CPS and having more than six times the papers in SM. One of the possible reasons we did not get many SM articles could be related to the fact that SM could be seen as a particular case of SE. Conversely, it is also possible that SM papers are being published in other indexes.

More granularly, in Figure 5, we observe that there was a good balance between papers being applied to the CPS area in general, and papers applied to a particular one such as IoT or Smart Cities.

### 3.3 Addressed problematics (RQ3)

Addressed problematics were grouped in the categories: *Lack of a method*, *Lack of a model*, *Lack of a specification*, *Lack of data*, *Lack of a general method*, *Poor data accessibility*, and *Existing methods' performance*. In Figure 6,

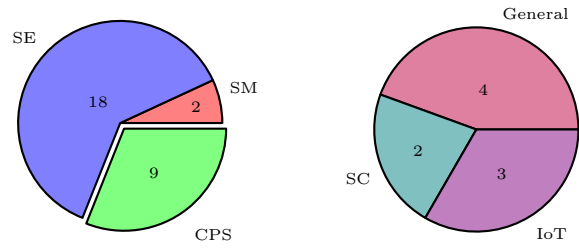


Fig. 4. Number of papers by application, where SM stands for Smart Manufacturing, SE for Smart Enterprise.

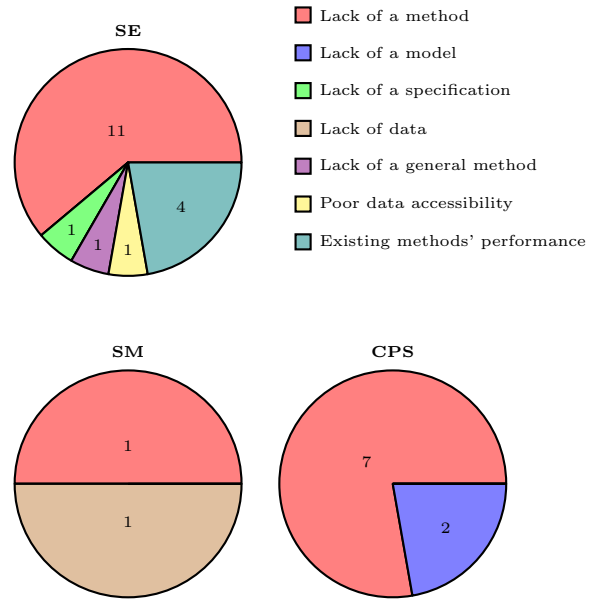


Fig. 6. Number of papers addressing each problematic by application.

it is possible to observe that *lack of a method* is the main problematic addressed by the selected articles, regardless of the application. Furthermore, SE is the only field in which the *existing methods' performance* was addressed. Moreover, the other two fields only deal with lack of methods, lack of models, or lack of data. Additionally, *lack of a method* covers more than 50% of occurrences in both SE (79%) and CPS (61%). Finally, more than 20% of papers in SE, address the problematic of the performance in previous methods.

This context helps to point out how more mature is the SE field with respect to the other two. SE not only sees a wider variety of problems to address, but also can be considered to be in a stage in which the existing methods are currently being improved. Furthermore, the CPS field has shown to be in an earlier stage, addressing mainly new solutions and models.

### 3.4 Methods distribution (RQ4)

For the methods, we considered four categories: FCA related, Ontology related, Supervised AI, and Other (i.e., these methods were algorithms related to manipulating databases directly using some *join* technique). In Figure 7,

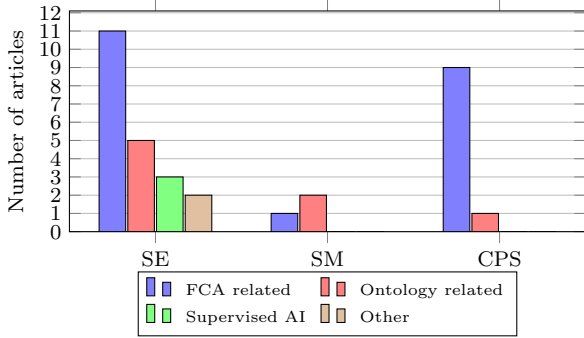


Fig. 7. Methods used in the articles grouped by application

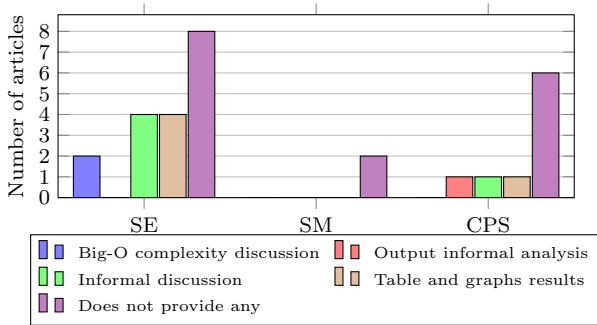


Fig. 8. Ways of measuring performance in the articles grouped by application

it can be seen that, FCA related, articles are the methods with more occurrences in total. In contrast, Supervised AI and Other only appear in SE, and have a total of 3 appearances among the 29 articles. This situation puts in evidence how relevant FCA related methods are in applied knowledge extraction. However, Ontology Related methods had relevance to some degree. In SE and CPS, it is the second most used method instead of Supervised AI, which was the one we expected, whereas in SM, Ontology Related had more prevalence than FCA.

### 3.5 Performance (RQ5)

The categories considered for the performance were Big-O complexity discussion, Output informal analysis, Informal discussion, Table and graphs results, and Does not provide any. Looking into Figure 8, we observe that all three application categories do not consider performance as an interesting metric to discuss in the majority of the papers, i.e., *Does not provide any* is the category with more occurrences. In addition, among the articles considering some performance metric, we found that most of the papers relied on either an informal discussion of the method itself, an informal discussion about the output (e.g., saying that comparing the proposed algorithms with the ones in the literature seemed to work faster, but not presenting the results), or a table/graph showing some particular results from the experiments (e.g., executed time, amount of generated association rules, etc.). Only two among all articles considered an asymptotic study to how the proposed solution scales.

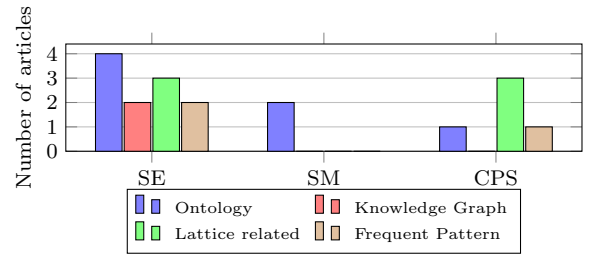


Fig. 9. Knowledge representation types in the output grouped by application

### 3.6 Knowledge representation types (RQ6)

For knowledge representation, we considered the categories: Ontology, Knowledge Graph, Lattice related, and Frequent Pattern. In Figure 9, we observe that *ontology* is the most used knowledge representation type in the output in total. However, *Lattice related* types are the most used in the CPS field, while *Frequent Patterns* and *Knowledge Graphs* total occurrences are in last place with 4 and 3 occurrences respectively. *Ontology* is the only knowledge representation type with occurrences in the three of the applications. This might be related with the generality of ontologies in relation with more specific structures like lattices.

### 3.7 Articles evaluation strategy (RQ7)

Articles evaluation strategies have been grouped in the following categories: Accuracy, Effectiveness, Efficacy, Robustness, Performance, Technical Feasibility, Operational Feasibility, Learning Capacity, Reliability, Validity, Scalability, and Utility. Figure 10 shows that *effectiveness* is the most considered evaluation strategy in both SE and CPS, while in SM it is tied with *efficacy* and *accuracy*. This is evident as by harnessing data and knowledge, companies can create an end-to-end visualisation of the production process. With knowledge and real-time information from advanced data analysis, previously unnoticed trends can be detected. This new knowledge and information can be contextualised and provided to all decision-making levels of the company for a faster and more accurate process. These connected systems are then linked to all functional systems of the enterprise, thus creating an efficient and productive intelligent environment. *Robustness* and *scalability* have only been considered in SE, whereas *utility* and *technical feasibility* occur only in CPS. *Performance* appeared both in SE and CPS, but in terms of percentage of total articles reviewed, it had a bigger impact in CPS (22% and 33% respectively).

Interestingly, as discussed in subsection 3.5, although a fair amount of articles studied the performance, only one of them considered scalability in its evaluation. This is related with Figure 8 since the fact that not many papers studied performance implies that not many of them studied scalability.

### 3.8 Type of Contribution (RQ8)

For the type of contribution, the categories considered were Algorithms, Architecture, Complete Software, Informal implementation, In-depth implementation, Model,

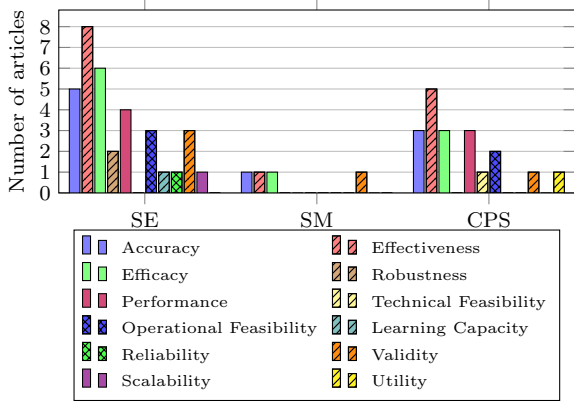


Fig. 10. Number of papers by application and evaluation strategy

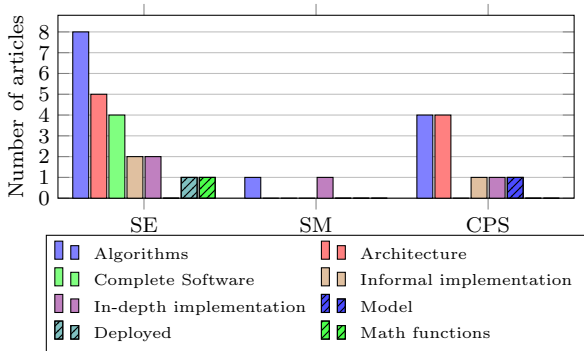


Fig. 11. Type of contributions in the articles grouped by application

Deployed, and Math functions. Figure 11 shows how most of the articles contributed *algorithms* and *architectures*. Additionally, the only field presenting a *complete software* and *deployed* solutions is SE. In terms of SM, the only two papers reviewed presented an in-depth implementation of the solution. Finally, CPS is the only field presenting a model as an outcome.

## 4. CLUSTERING TECHNIQUE

### 4.1 Formal context handling

For discovering knowledge and relationships between the reviewed articles in a more granular fashion, we used the mathematical framework FCA. In addition to the graphs and analysis provided in the last section, FCA yields a hierarchy of concepts as added and non-trivial knowledge from the articles reviewed. This hierarchy attains certain mathematical properties whose purpose is to represent the semantic understanding of “concepts”. In short, its input is a formal context  $K = (G, M, I)$  where  $G$  is a set of objects,  $M$  a set of attributes, and  $I$  an incidence matrix with  $I_{o,a}$  indicates whether the object  $o$  has the attribute  $a$ . From  $K$ , FCA extracts a set of formal concepts  $\mathcal{C}$ . A formal concept  $C = (X, Y) \in \mathcal{C}$  is a set where  $X \subseteq G$  is the extent,  $Y \subseteq M$  is the intent, the intersection of all attributes of objects in  $X$  is equal to  $Y$ , and the intersection of all objects held by attributes in  $Y$  is equal to  $X$ . Furthermore, if we consider the relation  $\leq_{\mathcal{C}}$  between concepts, where  $(X, Y) = C \leq_{\mathcal{C}} C' = (X', Y')$  means  $X \subseteq X'$ , we have that  $\mathcal{L} = (\mathcal{C}, \leq_{\mathcal{C}})$  forms a lattice (typically referred to

as *Concepts Lattice*). Two concepts worth mentioning are  $\top = (O, Y)$  and  $\perp = (X, M)$ , typically called *top* and *bot* respectively.

In this work, the objects  $G$  are the set of reviewed articles, identified by an *article ID* shown in Table 4. For the attributes  $M$ , we used the values of each of the extracted files, except for the year, and the publication venue. Furthermore, the multi-values have been put in the formal context as an attribute for each unique value, e.g., if the paper  $x$  contains the values *algorithms* and *deployed* in the column *type of contribution*, the context will include the attributes *algorithms* and *deployed*, and the incidence matrix will have  $I_{p,a}$  where  $p$  is the ID of the paper, and  $a$  the index of the mentioned attributes. To illustrate this, in the sets of concepts related to the formal context of the articles reviewed, there is a concept  $(\{22, 26\}, \{\text{boolean}\})$ , meaning that the only attribute these two articles have in common is *boolean*, and the attribute is *only* present in those papers.

### 4.2 Results

In order to generate the lattice, we used the tool [Conexp](#) with the formal context in [Leutwyler et al. \(2022\)](#).

The IDs and the articles considered in the lattice are shown in Table 4. The entire lattice is shown in Figure 12. Additionally, in Figure 13, Figure 14, and Figure 15, there are three sub-lattices depicted where the top node super-concept is the  $\top$  and the bottom concepts sub-concept’ is  $\perp$ . Above each node, we have the comma-separated attributes, and below them, we have the papers with an article ID  $i \in [1, 29]$ . The representation used is commonly known as compressed Hasse lattice representation, and the way to read it is the following: each circle represents a concept, concepts connected by lines inherit articles from bottom to top and attributes from top to bottom. Moreover, white-black circles only add objects, blue-white ones only add attributes, blue-black ones add both objects and attributes, and all-white ones do not add objects nor attributes. For instance, in Figure 15, the concept with the attribute *manufacturing* also contains the attributes *does not provide any* and *ontology* in its intent, and the articles 2 and 11 in its extent.

Firstly, Figure 13 highlights the sub-hierarchy of articles applied to SE (whose label in the lattice is *enterprise*). There, we notice how the immediate children of SE add the attributes *complete software*, *lack of a method*, *fca related*, *does not provide any*, *ontology*, *algorithm* and *table showing results*. From these sub-concepts, the only one that inherits entirely from *enterprise* is *complete software*, corresponding to the results presented in Figure 11. Furthermore, the first concept containing big-O complexity as a means of evaluating the papers is a direct child from *enterprise* and *algorithm*, meaning that in this context evaluating performance or scalability with big-O complexity implies SE and a contribution of the type *algorithm*.

Secondly, Figure 14 shows the sub-hierarchy of the articles applied to CPS. The first noticeable aspect of this sub-lattice is that the CPS concept has 3 direct sub-concepts adding the attributes *does not provide any*, *lack of a method*, and *fca related* respectively. Moreover, the graphic

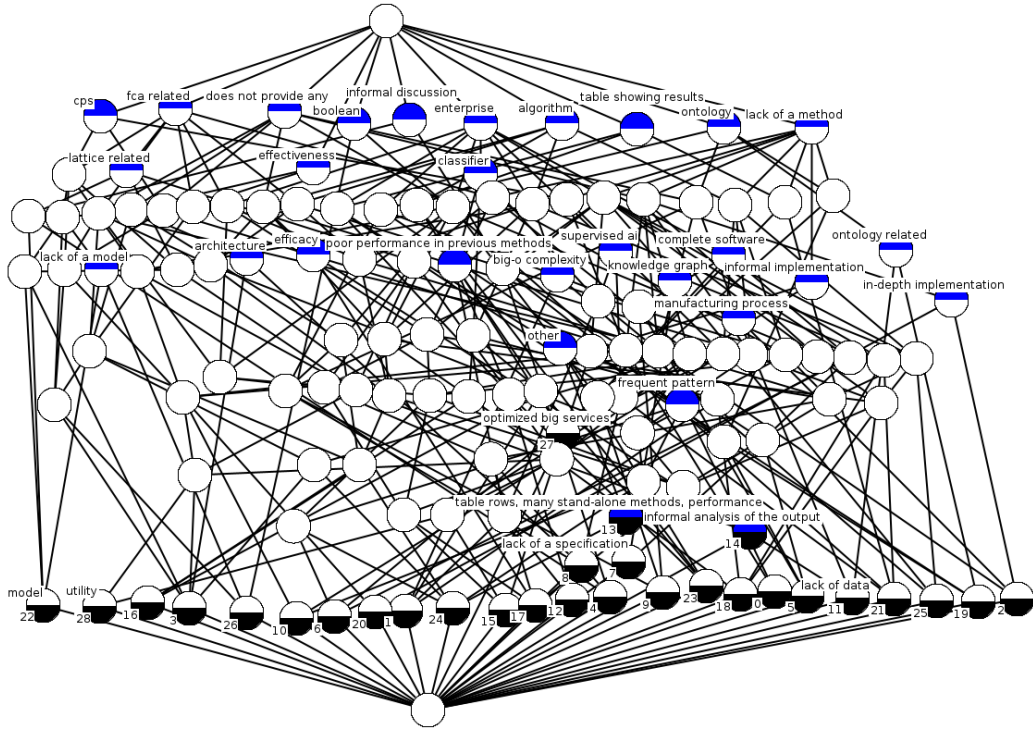


Fig. 12. Entire lattice calculated using Conexp.

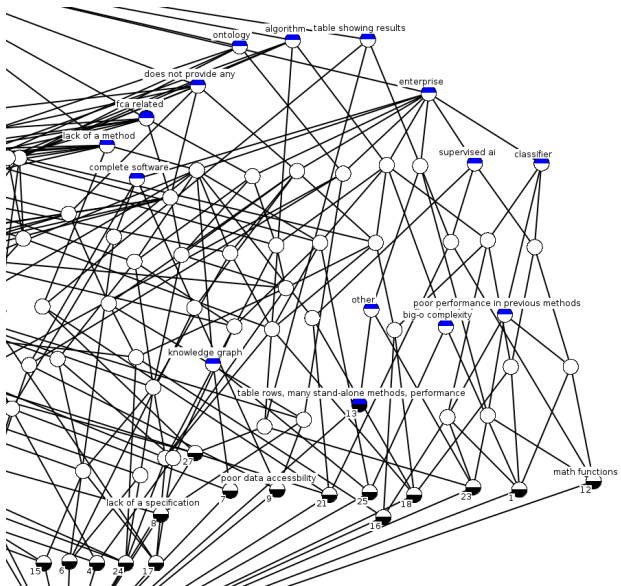


Fig. 13. SE sub-lattice calculated using Conexp.

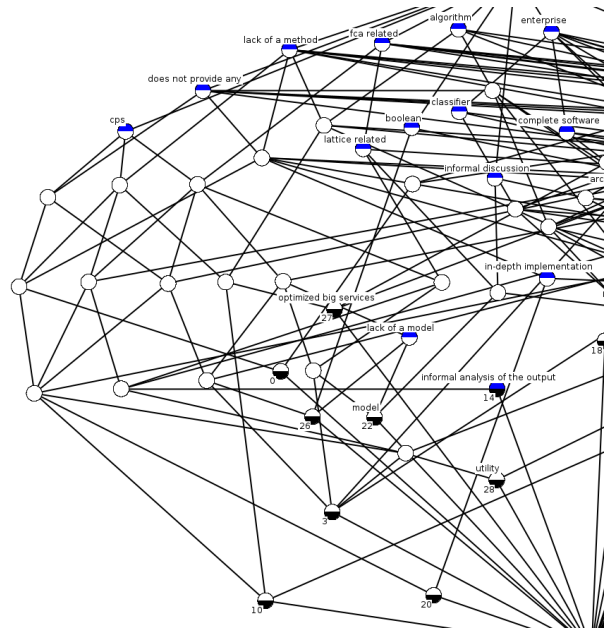


Fig. 14. CPS sub-lattice calculated using Conexp.

hints at the fact that although CPS attribute appears in a top concept (i.e., there are many articles sharing the attribute), the characteristics are shared with other top concepts in most of the cases. The only exceptions are *optimized big service* (that is a type of output), *lack of a model*, *model*, *utility*, and *informal analysis of the output*, being properties that are always accompanied by the CPS application.

Finally, Figure 15 depicts the hierarchy of articles in the SM application, which is not a top concept, and already has the attributes *ontology* and *does not provide any*. This means, in a conceptual level, that even though there are

only two papers in SM, both are part of the concept sharing the two properties of using an ontology as a type of knowledge representation in their developments, and to not provide any evaluation or mention of performance.

### 5. CONCLUSION AND FUTURE WORK

In this work, a systematic literature mapping has been conducted over a set of well-defined articles in the fields of SE, SM, and CPS. Many interesting results have been found directly from the data extraction, and were presented in section 3. Later, three concept hierarchies have

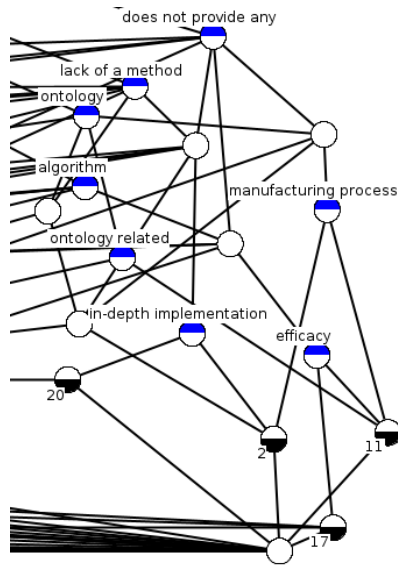


Fig. 15. SM sub-lattice calculated using Conexp.

been presented in section 4. The conclusions that we draw from the presented results are:

- (1) Due to the generality of SE in contrast to CPS and SM, the amount of publications found in the field are not only more numerous, but also provide a wider range of contributions. In addition, the methods used in this area are currently being improved in two ways: performance (e.g., time complexity, memory consumption, etc.), and data accessibility.
- (2) Methods for knowledge extraction in CPS are mainly new, i.e, the purpose they serve in their respective articles is to address the lack of a method or the lack of a model rather than improving existing ways of addressing the problem. This points out to the early stage in which the methods are in this field. Moreover, while conducting the full-text reading, we found that only 1 article addressed multi-relational data-mining, and it used Relational Concept Analysis as an auxiliary tool to the proposed solution.
- (3) There is a lack of publications in the area of SM, considering the indexes used in this work. Having said that, the articles found presented that Ontology related methods and knowledge representation types are more relevant than FCA in the field. Nevertheless, we consider that the amount of papers retrieved is not enough to draw a solid conclusion.

Another interesting point worth mentioning is that the FCA-related multi-relational data mining methods found during the review were RCA and Conceptual-Graph FCA. They were used either with feature engineering, or as an auxiliary tool in order to provide information about the frequent patterns as an input to other process. No Polyadic nor Triadic Concept Analysis have been used nor discussed in the articles, which was a surprise to us.

With these points in mind, we reach the conclusion that the **formal methods** in KE from **heterogeneous sources** still lacks research development, particularly in the area of **distributed architectures**, which plays a crucial role in the two fields of CPS and SM. Therefore, we will focus our efforts in addressing specially the research in

the area of lattice related KE from heterogeneous sources in distributed architectures.

The future work that can be applied in order to improve this article includes: 1) To broaden the search including more areas and to see whether similar conclusions apply there. 2) To conduct a cross-review of the search fields in order to ensure less subjectivity in the quality assessment. 3) To conduct backward and forward snowball steps to avoid missing relevant articles. 4) To analyze the association rules resulted from the given lattice.

## ACKNOWLEDGEMENTS

This work has been funded with the help of the French National Agency for Research and Technology (ANRT) and French National Syndicate of Ski Teachers (SNMSF).

## REFERENCES

- Abuzaid, F., Kraft, P., Suri, S., Gan, E., Xu, E., Shenoy, A., Ananthanarayan, A., Sheu, J., Meijer, E., Wu, X., Naughton, J., Bailis, P., and Zaharia, M. (2021). DIFF: a relational interface for large-scale data explanation. *The VLDB Journal*, 30(1), 45–70. doi:10.1007/s00778-020-00633-6. URL <https://doi.org/10.1007/s00778-020-00633-6>.
- Akmal, S. and Batres, R. (2013). A Methodology for Developing Manufacturing Process Ontologies. *Journal of Japan Industrial Management Association*, 64, 303–316. doi:10.11221/jima.64.303.
- An, J. and Li, G. (2017). A Novel Fuzzy Rough Context Preprocessing Method. In *2017 9th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC)*, volume 2, 141–144. doi:10.1109/IHMSC.2017.147.
- Baxter, M., Polovina, S., Laurier, W., and Rosing, M.v. (2021a). Active Semantic Relations in Layered Enterprise Architecture Development. In M. Cochez, M. Croitoru, P. Marquis, and S. Rudolph (eds.), *Graph Structures for Knowledge Representation and Reasoning*, Lecture Notes in Computer Science, 3–16. Springer International Publishing, Cham. doi:10.1007/978-3-030-72308-8\_1.
- Baxter, M., Polovina, S., Laurier, W., and von Rosing, M. (2021b). Generating Layered Enterprise Architectures with Conceptual Structures. In T. Braun, M. Gehrke, T. Hanika, and N. Hernandez (eds.), *Graph-Based Representation and Reasoning*, Lecture Notes in Computer Science, 34–47. Springer International Publishing, Cham. doi:10.1007/978-3-030-86982-3\_3.
- Bordes, A., Glorot, X., Weston, J., and Bengio, Y. (2014). A semantic matching energy function for learning with multi-relational data. *Machine Learning*, 94(2), 233–259. doi:10.1007/s10994-013-5363-6. URL <https://doi.org/10.1007/s10994-013-5363-6>.
- Buchmann, R.A., Meza, R., and Pulcher, D. (2011). An Approach to the Semantization of ERP Systems. In W. Abramowicz (ed.), *Business Information Systems*, Lecture Notes in Business Information Processing, 218–229. Springer, Berlin, Heidelberg. doi:10.1007/978-3-642-21863-7\_19.
- Cretan, A., Nica, C., Coutinho, C., Jardim-Goncalves, R., and Bratu, B. (2021). An Intelligent System to Ensure

- Interoperability for the Dairy Farm Business Model. *Future Internet*, 13, 153. doi:10.3390/fi13060153.
- Cui, Z., Damiani, E., and Leida, M. (2007). Benefits of Ontologies in Real Time Data Access. In *2007 Inaugural IEEE-IES Digital EcoSystems and Technologies Conference*, 392–397. doi:10.1109/DEST.2007.372004. ISSN: 2150-4946.
- De Maio, C., Fenza, G., Loia, V., and Orciuoli, F. (2017). Making sense of cloud-sensor data streams via Fuzzy Cognitive Maps and Temporal Fuzzy Concept Analysis. *Neurocomputing*, 256, 35–48. doi:10.1016/j.neucom.2016.06.090. URL <https://www.sciencedirect.com/science/article/pii/S0925231217304125>.
- Džeroski, S. (2003). Multi-relational data mining: an introduction. *ACM SIGKDD Explorations Newsletter*, 5(1), 1–16. doi:10.1145/959242.959245. URL <https://doi.org/10.1145/959242.959245>.
- Fallon, R. and Polovina, S. (2016). Automated REA (AREA): A Software Toolset for a Machine-readable Resource-Event-Agent (REA) Ontology Specification.
- Galkin, M., Auer, S., Kim, H., and Scerri, S. (2016). Integration Strategies for Enterprise Knowledge Graphs. In *2016 IEEE Tenth International Conference on Semantic Computing (ICSC)*, 242–245. doi:10.1109/ICSC.2016.24.
- George Voutsadakis (2002). Polyadic Concept Analysis. *Order*, 19(3), 295–304. doi:10.1023/A:1021252203599. URL <https://doi.org/10.1023/A:1021252203599>.
- Gharbi, M. and Mezni, H. (2020). Towards big services composition. *International Journal of Web and Grid Services*, 16(4), 393–421. doi:10.1504/IJWGS.2020.110946. URL <https://www.inderscienceonline.com/doi/abs/10.1504/IJWGS.2020.110946>. Publisher: Inderscience Publishers.
- Han, J. (2009). Data Mining. In L. LIU and M.T. ÖZSU (eds.), *Encyclopedia of Database Systems*, 595–598. Springer US, Boston, MA. doi:10.1007/978-0-387-39940-9\_104. URL [https://doi.org/10.1007/978-0-387-39940-9\\_104](https://doi.org/10.1007/978-0-387-39940-9_104).
- Hildebrandt, M., Sunder, S.S., Mogoreanu, S., Thon, I., Tresp, V., and Runkler, T. (2019). Configuration of Industrial Automation Solutions Using Multi-relational Recommender Systems. In U. Brefeld, E. Curry, E. Daly, B. MacNamee, A. Marascu, F. Pinelli, M. Berlingerio, and N. Hurley (eds.), *Machine Learning and Knowledge Discovery in Databases*, Lecture Notes in Computer Science, 271–287. Springer International Publishing, Cham. doi:10.1007/978-3-030-10997-4\_17.
- Jalali, S. and Wohlin, C. (2012). Systematic literature studies: database searches vs. backward snowballing. In *Proceedings of the ACM-IEEE international symposium on Empirical software engineering and measurement, ESEM '12*, 29–38. Association for Computing Machinery, New York, NY, USA. doi:10.1145/2372251.2372257. URL <https://doi.org/10.1145/2372251.2372257>.
- Kalfoglou, Y., Dasmahapatra, S., and Chen-Burger, Y.H. (2004). FCA in Knowledge Technologies: Experiences and Opportunities. In P. Eklund (ed.), *Concept Lattices*, Lecture Notes in Computer Science, 252–260. Springer, Berlin, Heidelberg. doi:10.1007/978-3-540-24651-0\_23.
- Leutwyler, N., Lezoche, M., and Torres, D. (2022). Systematic Literature Mapping - Selected Articles Data Extraction. doi:10.5281/zenodo.7307957. URL <https://doi.org/10.5281/zenodo.7307957>.
- Li, X. and Li, G. (2015). A Hybrid Context Inconsistency Resolution Method. In *2015 7th International Conference on Intelligent Human-Machine Systems and Cybernetics*, volume 2, 73–76. doi:10.1109/IHMSC.2015.121.
- Loia, V., De Maio, C., Fenza, G., Orciuoli, F., and Senatore, S. (2010). An enhanced approach to improve enterprise competency management. In *International Conference on Fuzzy Systems*, 1–8. doi:10.1109/FUZZY.2010.5584261. ISSN: 1098-7584.
- Louati, A., Aufaure, M.A., Cuvelier, E., and Pimentel, B. (2016). Soft and Adaptive Aggregation of Heterogeneous Graphs with Heterogeneous Attributes. In P. Molli, J.G. Breslin, and M.E. Vidal (eds.), *Semantic Web Collaborative Spaces*, Lecture Notes in Computer Science, 145–180. Springer International Publishing, Cham. doi:10.1007/978-3-319-32667-2\_7.
- Maimon, O. and Rokach, L. (2005). Introduction to Knowledge Discovery in Databases. In O. Maimon and L. Rokach (eds.), *Data Mining and Knowledge Discovery Handbook*, 1–17. Springer US, Boston, MA. doi:10.1007/0-387-25465-X\_1. URL [https://doi.org/10.1007/0-387-25465-X\\_1](https://doi.org/10.1007/0-387-25465-X_1).
- Manning, C.D., Raghavan, P., and Schütze, H. (2008). *Introduction to Information Retrieval*. Cambridge University Press. doi:10.1017/CBO9780511809071. URL <https://www.cambridge.org/highereducation/books/introduction-to-information-retrieval/669D108D20F556C5C30957D63B5AB65C>. ISBN: 9780511809071 Publisher: Cambridge University Press.
- Mashhadi, N.R., Jalali, M., and Jahan, M.V. (2015). Inference of mobile users' social relationships using Bayesian belief network. In *2015 International Congress on Technology, Communication and Knowledge (ICTCK)*, 232–240. doi:10.1109/ICTCK.2015.7582676.
- Miguel-Rodríguez, J.d., Galán-Páez, J., Aranda-Corral, G.A., and Borrego-Díaz, J. (2016). Urban Knowledge Extraction, Representation and Reasoning as a Bridge from Data City towards Smart City. In *2016 Intl IEEE Conferences on Ubiquitous Intelligence & Computing, Advanced and Trusted Computing, Scalable Computing and Communications, Cloud and Big Data Computing, Internet of People, and Smart World Congress (UIC/ATC/ScalCom/CBDCCom/IoP/SmartWorld)*, 968–974. doi:10.1109/UIC-ATC-ScalCom-CBDCCom-IoP-SmartWorld.2016.0152.
- Morozov, D., Lezoche, M., and Panetto, H. (2018). Multi-paradigm modelling of Cyber-Physical Systems. *IFAC-PapersOnLine*, 51(11), 1385–1390. doi:10.1016/j.ifacol.2018.08.334. URL <https://www.sciencedirect.com/science/article/pii/S2405896318314587>.
- Moujabbir, M. and Ramdani, M. (2013). Fuzzy Closed Concept for the Correction of Errors in the Dematerialization Process. In *2013 Eighth International Conference on P2P, Parallel, Grid, Cloud and Internet Computing*, 447–452. doi:10.1109/3PGCIC.2013.75.
- Petersen, K. and Gencel, C. (2013). Worldviews, Research Methods, and their Relationship to Validity in Empirical Software Engineering Research. In *2013 Joint Conference of the 23rd International Workshop on Software Measurement and the 8th International Conference on*

- Software Process and Product Measurement*, 81–89. doi:10.1109/IWISM-Mensura.2013.22.
- Petersen, K., Vakkalanka, S., and Kuzniarz, L. (2015). Guidelines for conducting systematic mapping studies in software engineering: An update. *Information and Software Technology*, 64, 1–18. doi:10.1016/j.infsof.2015.03.007. URL <https://www.sciencedirect.com/science/article/pii/S0950584915000646>.
- Poelmans, J., Ignatov, D., Kuznetsov, S., and Dedene, G. (2013). Review: Formal concept analysis in knowledge processing: A survey on applications. *Expert Systems with Applications: An International Journal*, 40, 6538–6560. doi:10.1016/j.eswa.2013.05.009.
- Polovina, S., Rosing, M.v., Scheruhn, H.J., and Weidner, S. (2016). Highlighting the Gaps in Enterprise Systems Models by Interoperating CGs and FCA. *International Conference on Conceptual Structures*, 22.
- Polovina, S., von Rosing, M., Laurier, W., and Etzel, G. (2019). Enhancing Layered Enterprise Architecture Development Through Conceptual Structures. In D. Endres, M. Alam, and D. Şotropa (eds.), *Graph-Based Representation and Reasoning*, Lecture Notes in Computer Science, 146–159. Springer International Publishing, Cham. doi:10.1007/978-3-030-23182-8\_11.
- Prat, N., Comyn-Wattiau, I., and Akoka, J. (2015). A Taxonomy of Evaluation Methods for Information Systems Artifacts. *Journal of Management Information Systems*, 32(3), 229–267. doi:10.1080/07421222.2015.1099390. URL <http://www.tandfonline.com/doi/full/10.1080/07421222.2015.1099390>.
- Robinson, P. and Lowe, J. (2015). Literature reviews vs systematic reviews. *Australian and New Zealand Journal of Public Health*, 39(2), 103–103. doi:10.1111/1753-6405.12393. URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/1753-6405.12393>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/1753-6405.12393>.
- Romero, M., Guédria, W., Panetto, H., and Barafort, B. (2020). Towards a Characterisation of Smart Systems: A Systematic Literature Review. *Computers in Industry*, 120, 103224. doi:10.1016/j.compind.2020.103224. URL <https://www.sciencedirect.com/science/article/pii/S0166361520300270>.
- Rouane-Hacene, M., Huchard, M., Napoli, A., and Valtchev, P. (2013). Relational Concept Analysis: Mining Concept Lattices From Multi-Relational Data. *Annals of Mathematics and Artificial Intelligence*, 67. doi:10.1007/s10472-012-9329-3.
- Shin, J., Kulvatunyou, B., Lee, Y., and Ivezic, N. (2013). Concept Analysis to Enrich Manufacturing Service Capability Models. *Procedia Computer Science*, 16, 648–657. doi:10.1016/j.procs.2013.01.068.
- Siddavatam, I.A., Parekh, S., Shah, T., and Kazi, F. (2017). Testing and Validation of Modbus/TCP Protocol for Secure SCADA Communication in CPS using Formal Methods. *Scalable Computing: Practice and Experience*, 18(4), 313–330. doi:10.12694/scpe.v18i4.1331. URL <https://www.scpe.org/index.php/scpe/article/view/1331>.
- Sithole, V. and Marshal, L. (2019). Building Concept Hierarchies for the Internet of Things Patterns Using Domain-specific Dependency Knowledge. In *2019 Open Innovations (OI)*, 1–7. doi:10.1109/OI.2019.8908182.
- Sowa, J. (2000). Knowledge Representation: Logical, Philosophical, and Computational Foundations. In *Knowledge Representation: Logical, Philosophical, and Computational Foundations*. Brooks/Cole Publishing.
- Torres, D., Lezoche, M., Fernandez, A., Antonelli, L., and Panetto, H. (2021). Knowledge discovering from multiple sources in agriculture value-chain. In *International Conference in Technology Innovation and Uncertainty Management in Agri-food Systems, RUC-APS 2021*. Online, United Kingdom. URL <https://hal.archives-ouvertes.fr/hal-03666559>.
- Unbehauen, J., Hellmann, S., Auer, S., and Stadler, C. (2012). Knowledge Extraction from Structured Sources. In S. Ceri and M. Brambilla (eds.), *Search Computing: Broadening Web Search*, Lecture Notes in Computer Science, 34–52. Springer, Berlin, Heidelberg. doi:10.1007/978-3-642-34213-4\_3. URL [https://doi.org/10.1007/978-3-642-34213-4\\_3](https://doi.org/10.1007/978-3-642-34213-4_3).
- Wajnberg, M., Lezoche, M., Blondin Masse, A., Valtchev, P., Panetto, H., and Tyvaert, L. (2018). Semantic interoperability of large systems through a formal method: Relational Concept Analysis. In *16th IFAC Symposium on Information Control Problems in Manufacturing, INCOM 2018*, volume 51 of *IFAC-Papersonline*, 1397–1402. Bergamo, Italy. doi:10.1016/j.ifacol.2018.08.330. URL <https://hal.archives-ouvertes.fr/hal-01813398>. Issue: 11.
- Wajnberg, M., Valtchev, P., Lezoche, M., Blondin-Massé, A., and Panetto, H. (2021). FCA went (multi-)relational, but does it make any difference? 27–38. Montréal, Canada.
- Watmough, M. (2014). Discovering the Hidden Semantics in Enterprise Resource Planning Data Through Formal Concept Analysis. 495, 291–314. doi:10.1007/978-3-642-35016-0-11.
- Wille, R. (1982). Restructuring Lattice Theory: An Approach Based on Hierarchies of Concepts. In I. Rival (ed.), *Ordered Sets*, NATO Advanced Study Institutes Series, 445–470. Springer Netherlands, Dordrecht. doi:10.1007/978-94-009-7798-3\_15.
- Yang, J., Yang, L.T., Wang, H., and Gao, Y. (2022). Multi-Relational Tensor Graph Attention Networks for Knowledge Fusion in Smart Enterprise Systems. *IEEE Transactions on Industrial Informatics*, 1–9. doi:10.1109/TII.2022.3190548. Conference Name: IEEE Transactions on Industrial Informatics.
- Zhang, X., Duan, N., and Dong, W. (2012). Relational rule learning in decoupled heterogeneous subspaces. In *Proceedings of 2012 IEEE International Conference on Service Operations and Logistics, and Informatics*, 66–71. doi:10.1109/SOLI.2012.6273506.