

SYSTEMATIZED ANALYSIS USING DATA MINING'S METHODOLOGY ON THE TOPIC OF REGIONAL INDUSTRIAL SYMBIOSIS AND ITS NETWORKS

Erika Džajić Uršič

School of Advanced Social Studies in Nova Gorica, Slovenia

E-mail: erika.ursic@fuds.si

Abstract: *The aim of this paper is the analysis and a presentation of texts and terminological ontology of scientific and professional publications in the field of regional Industrial Symbiosis and Industrial Symbiotic Networks. With computer-aided text analysis it is possible to have a comprehensive overview of the publications of works in this field and a more detailed presentation of the most fruitful concepts of publications.*

Key words: Data mining, knowledge, industrial symbiosis, industrial symbiotic networks, ontology, OntoGen

1. Introduction

The development of European society is based on knowledge and the ability to implement it in all areas of human activity. Knowledge is, for example, also present in the recent EU grand strategy Europe 2020, aiming to achieve growth, which is also *smart* (Rončević 2019), although pursuit of this strategy is not always completely successful (Makarovič et al. 2014). In addition to its growth being smart, the European Union also strives for it to be inclusive and sustainable, hence the immense focus on circular economy (Mileva Boshkoska et al. 2018; Fric et al. 2020). In this article we are going to focus on the knowledge relating to the phenomena of industrial symbiosis and its networks.

We must know how to properly present our own or acquired knowledge and attract attention to individuals or/and organizations that can use this knowledge usefully in their work or field. It is the organizations themselves making sure the differences among them exist mainly due to ability to detect and absorb knowledge (Rončević and Valič 2019, 330).

Knowledge modeling provides a comprehensive overview of the issues under consideration or the domain being studied, so this paper focuses on the

field of knowledge modeling and its presentation in the form of terminological ontologies.

The following organization of the data from the visible journal on web into the text analysis of published titles of articles, will be supported by a computer programme. This programme will enable semi-automatic construction of ontologies from texts. By modelling the domain knowledge in the form of ontologies based on the titles of articles, is intended to examine the main research areas, which can be reflected from the publications. The result of the analysis was performed as a picture of areas in which the research activities of the topic on regional industrial symbiosis and industrial symbiotic networks was the most outlined and representative. It was taken into account that was selected only one Journal at this stage and a determined number of titles in terms of publications during the examined period of the last 25 years.

2. The concepts of Industrial Symbiosis and Industrial Symbiotic Networks

A resource sharing among co-located firms referenced in the Industrial Ecology (IE) literature as Industrial Symbiosis (IS) is concerned with the cyclical flow of resources through networks of businesses. Industrial symbiosis is mostly characterized with the cyclical flow of resources through networks of businesses as a means of cooperatively approaching ecologically sustainable industrial activity. As such it is motivated by economic considerations such as lowering costs for waste disposal, as well as by environmental ones (Beers et al. 2007). The main goal of industrial symbiosis is to aid in industrial organization by pushing firms to think beyond individual firm boundaries into a broader systems level (P7 Modelling industrial symbiosis to find the potentials and barriers in Aalborg, Denmark 2017). Firms belonging to an industrial ecology utilize industrial symbiosis as a collective approach to competitive advantage and simultaneously realize economic and environmental benefits, but also emerge social considerations, in our case Industrial Symbiotic Networks (ISNs).

Industrial symbiosis, in other words is a few dozen partners, means for trust-building include personal meetings, regular communication, sharing of information and knowledge, and stable rules of the game (Fric 2016).

In 1947, the term “industrial symbiosis” was first utilized by Renner in the economic geography literature to describe “organic relationship” between dissimilar industries, including the use of the waste products from one as input to another. The development of industrial symbiosis in an international

perspective gas grown substantially since the late 40's until now (Chertow 2000, 313–37; Džajić Uršič 2020).

2.1 Industrial symbiosis and the social sciences

When was tried to construct a proposed hierarchy in this paper, it has been “reduced” a perceived reality of industrial symbiosis systems so the research encompasses and gives importance to the social interactions as trust and willingness (Black Sea Industrial Symbiosis Platform 2017). For the success of industrial symbiosis (Schiller et al. 2014; GILG 2013) is firstly vital the exchange of material by-products and energy, secondly some forms of social exchange, in other words, inter-firm networking, trust and collaboration. Many attempts of regional projects to build industrial symbiosis failed because is it hard to find firms willing to co-locate and link their processes with other firms they do not yet know or trust. This is perceived as simply too great a risk to take. Trust is needed in exploring the possibilities for by-products exchange because firms need to share (possibly sensitive) information about their inputs and outputs (Gilg 2013; Fric 2016).

New economic sociology also pays a little attention to the influence of institutions; might be, that the institutional approaches play a secondary role in the new economic sociology in general, while networks and cultural explanations prevail. Even in studies that are informed by institutionalism, the concept of institution that is being used leans towards the new sociological institutionalism (CRESSI 2017; Kalundborg Symbiose 2017) (and focuses on questions of legitimacy and diffusion of institutional models). Although sociology has lagged behind other social sciences in appreciating new methodologies and a distinctive sociological contribution is evident here. Information about relative data of various titles can assist in the analysis of dozen of titles” texts and following hierarchical ontologies (Džajić Uršič 2020).

The development of industrial symbiotic networks is complex and can be, as a consequence of several factors influencing the evaluation process, uncertain and unclear. The complexity of the process has increased in the period. Even though the exchanges of waste were not called industrial symbiosis or synergies of networks, the early process began at that time. Even if the term of industrial symbiosis may appear to have fully grown onto the sustainability stage, resources exchange and processes of trading are basic aspects of the developing world (Chertow 2000, 313–37).

The importance of the social dimension of industrial symbiosis is very commonly acknowledged in the literature. Trust and willingness on some level is definitive to cooperate and it is seen as requirements for “industrial

symbiosis proper” (exchange of materials and energy) to take place (Ehrenfeld 1997; Gibbs and Deutz 2005; Gibbs and Deutz 2007; Tudor et al. 2007). Many of those industrial symbiosis projects failed because it is hard to find firms willing to co-locate and link their processes with other firms they do not yet know or trust. This is perceived as simply too great a risk to take. Trust is needed before interdependencies through by-products exchange can be set up, because certainty and continuity of supply are extremely important to industrial firms (Tudor et al. 2007).

3. From data to information and knowledge

Data are chaotic, unprocessed essentials that tell us nothing in themselves until they are properly processed, but only when we process the data and/or compare between, we find out their meanings. Data doesn't tell us anything about motivation, the quality or characteristics of the analyses, by which they were obtained, but are a prerequisite for information obtained as a result of researches. The problem arises when we have too much data and we do not know what decisions make to reach the desired goal (Bizjak 2014, 9–11).

The knowledge is a base in this case because helps us at work, in everyday life and has several meanings depending where it is applied. The knowledge is not just information like even information is not just data and knowledge comes from information in the same way as information comes from data.

Awad and Ghaziri (2007) argue that information helps to understand relationships between them and give meaning to the unorganized data, that can be later reorganized, statistically analyze, avoided from errors, or otherwise process them to obtain their meaning. Only when we collect enough information, we understand a phenomenon or problem and talk about knowledge (Girard and Girard 2015; Awad and Ghaziri 2007; Szilva et al. 2018).

Processes for discovering useful knowledge from data are usually based on data mining, which allows the use of special computer algorithms obtaining and displaying meaningful models from input data. Knowledge discovery processes can also be based on text collections and the procedure from an input text extracted structured information is called *text mining*. Similar to how databases can be analyzed by data mining in the case of text databases (example: bibliographic databases of scientific records, professional publications, etc.) uses text mining procedures (Sebastiani 2002; Feyyad 1996). Modern databases offer fast-growing volumes of text and other formats of data. Finding the proper data and getting information based on it, among such a crowd of data, is difficult, especially, if we want understanding

areas on which we are not professionally knowledgeable. Therefore, there is a growing need for methods and tools that would be able to extract useful knowledge from the mass of data (Fayyad et al. 1996).

Exploring data means looking for characteristics and display them for the group we are processing. This includes some important steps (Awad and Ghaziri 2007):

- definition of primary hypotheses and prediction of further measures,
- identification of outliers who are outside the expected crowd results,
- display of key features and
- selection of interesting data groups for further research.

For this case, we can use modern program tool that enable automatic or semi-automatic analysis displaying data. The tool called OntoGen in this research is used for textual analysis of dozen of titles of article where we can present results in the form of content *hierarchies* called *ontologies*.

3.1 Data mininig

Data mining and machine learning have a goal to extract knowledge (in a human-understandable structure) from large quantities of data (for example to illustrate relations between observed variables). The automatic or semi-automatic data mining process analyzes data and discovers new outlines to extract previously unknown remarkable outlines such as simplifications that can predict values of a certain variable in previously unseen examples from the known values of other variables (inquiring groups of data records, rare records and needs in the data).

The process of data mining can be seen as a procedure where signify terms from the collection of input texts used *Bag of words*, which show the analyzed texts by extracting terms from them, without taking into account grammatical rules and word order.

Firstly, what we need to do is to collect different groups of data and display from them statistically significant data to display characteristic for the group we are processing. At this stage are important some steps: (1) definition of initial hypotheses, (2) presentation of key attributes, (3) recognition of outliers outside of the crowd results (4) and selection of remarkable data for further studies.

3.2 What are the ontologies in this case?

An ontology is a presentation of a multitude of concepts on some field of human knowledge where those concepts can be categorized. The term "*ontology*" comes from philosophy and is often used to define philosophical disciplines for displaying some basis, reasons and typical characteristics of

reality. Ontologies can be used as a model of formal structure for unchanging picture and semantics of information (Lavbič and Krisper 2005). Depending on the degree of formality, ontologies are:

- descriptive (defined by terms that are often used to present knowledge and describe a specific thematic area),
- formal (consists only of formally accepted descriptive terminology ontologies, which denotes different aspects or types of a given area), or
- formalized (contain formal specifications of a field in the strictest meaning of the formalization of the presentation (Poli and Seibt 2010).

Formalized ontologies are glossaries of terms (words and their synonyms) that semantically structure a particular area and at the same time allow to show the hierarchical structure of relations between concepts. In this sense, ontologies have great potential especially in areas of knowledge management, information gathering, integration of intelligent systems and e-commerce (Lavbič and Krisper 2005).

The philosophy of the regional industrial symbiosis is concerned with identifying the kinds of things that exist in industrial symbiotic networks so, ontological (and epistemological) ideas guide not only to the selection of the research focus but also to the methodology assumed and the expected outcomes. Industrial symbiotic networks are in this research observed as social structures that are simultaneously constructed and reflected upon by the researcher and the stakeholders that participate into the system – this approach understands their identification first as a singular phenomenon, and then as a sustainable instrument, part of the process of co-creation of meaning and social interaction (Džajić Uršič 2020).

This ontological (and epistemological) position explains the focus on a small part of this research on the social processes. These processes are behind the articulation of industrial symbiotic networks and the emphasis given to structural and discursive dimensions in the analysis. When trying to conceptualize a proposed research and “reduce” a perceived reality of industrial symbiosis systems, we encompass and give importance to social interactions (Džajić Uršič 2020).

However, the new economic sociology pays just a little attention to the influence of the important *institutions* (from authors’ perspective); might be that the institutional approaches play a secondary role in the new economic sociology in general, while networks and cultural explanations prevail. Even in studies that are informed by institutionalism, the concept of an institution that is being used leans towards the new sociological institutionalism

(Howard-Grenville and Paquin 2008; Scott 2004) and focuses on questions of legitimacy and diffusion of institutional models. While sociology has padded behind other social sciences in valuing this new methodology, a distinctive sociological contribution is evident here (Macy and Willer 2002, 143–166). Theoretical interest focuses on dynamic social networks that shape and are shaped by agent interaction and to perform experiments to test macro-sociological theories by manipulating structural factors like network topology, spatial mobility and social stratification here (Macy and Willer 2002, 143–166). In the following chapter are presented and discussed ontologies in terms of informal ones, as semantic conceptual representations of domain knowledge. Sowa (2000) defines such ontologies as terminological, showing general concepts and relations of particular domain knowledge, as opposite to formal ontologies that precisely represent domain expressions and specific types of relations (such as is-part, is-case) among them. Terminological ontologies are characterised by the words and axioms in natural language, while formal ontologies are obtainable in a logical language, based on concept dictionaries and sentences that show the relationships between concepts. Ontologies representing domain knowledge can enable a common understanding of the domain (Lavbič and Krisper 2005), which simplifies communication among people (terminological ontologies) and computer systems (formal ontologies).

4. Methodology

The data are obtained primarily from website *Sciencedirect* (Sciencedirect 2020), where scientific and professional publications on the topic of “Industrial Symbiosis”, and topic “Industrial Symbiotic Networks” is built a model of knowledge. Those two topics or better keywords are analyzed. The data from the website are organized into a collection of posed titles, in plain text with a text editor. This was the basis for the text analysis performed with the computer tool *OntoGen* (Fortuna et al. 2001; 2006; 2007), which enables the semi-automatic construction of knowledge models in the form of ontologies (Ontogen 2020; *OntoGen* 2020).

Semi-Automatic means that *“a system is an interactive tool that aids the user during the ontology construction process. The system suggests relations, concepts, names, automatically assigns cases to concepts and provides a good overview of the ontology to the user trough concept browsing and visualization. At the same time, the user can fully adjust all the properties of the ontology by manually adding or deleting concepts, relations and reassigning cases”* (Fortuna et al. 2001, 1).

In this case, for this paper, somehow it is very difficult to manage a lot of web pages and textual documents and in this case ontologies play a very important role for them. Ontologies offer additional help to reduce overloaded information for a specific area.

OntoGen is a “data-driven and semi-automatic” system which is used for generating topic ontologies. *“The functionality of the system (machine learning algorithms, data handling and ontology visualization) is based on the Text Garden library [Grobelnik06] and is efficiently implemented in C++. The user interface is written in C# and requires Microsoft .NET framework 2.03 for running. The heavy-duty functionality is now compiled in native code and stored in DLL library. The library is accessed from the C# and used to provide the functionality to the user interfaces. Previous version used Managed C++ so the whole program was compiled into bytecode. By compiling the core functionality directly into native code the machine learning algorithms used in OntoGen run approximately twice as fast as in the previous version”* (Fortuna et al. 2001, 1).

“The old version of OntoGen is only used to edit the different types of ontology which was connected with different kinds of relations. It is difficult for user because user spend lot of time to edit ontology. But the new version of OntoGen is providing attractive interface for user which reduces time and complexity for user. The system is attractive and provides helps for user during the ontology construction process. It give suggests about concepts and relations between concepts and automatically assigns cases to concepts. For the construction process of ontology in OntoGen, data is provided by user. OntoGen is providing support for user to give automatic concepts and also describe the relationship between these concepts according to data” (Fortuna et al. 2001; Bajwa et al. 2019).

The next main steps with the semi-automatically managed OntoGen program are:

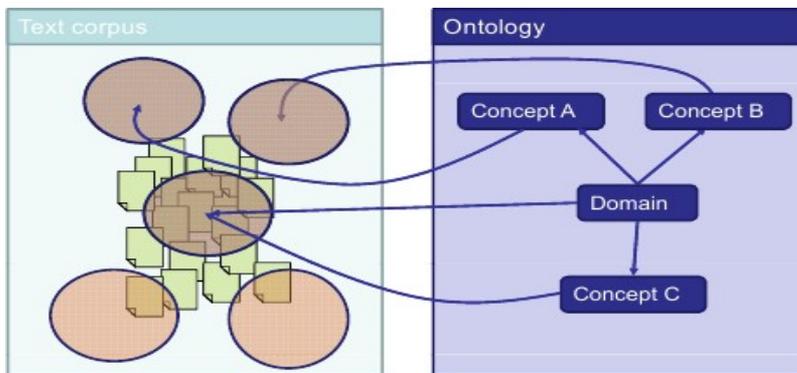
- presentation of knowledge modeling, from when the data are captured to information acquisition and knowledge demonstrations;
- definition of knowledge representation in the form of ontologies, classification of ontologies and the importance of terminological ontologies in knowledge modeling;
- presentation of a case study with a more detailed description of the preparation of data, for text analysis and the sequence of text mining with the computer tool OntoGen;
- presentation of the results of text analysis (Guarino 1995).

5. Case Study

The data included in the study refers to publications from 2000 to the end of 2019 and contains the titles of scientific publications on the topic of industrial symbiosis and its networks. With the tool for word processing and analysis, are processed then the titles. Due to the needs of textual address analysis publications, is chosen the format of the XML output format. On this way, we got 267 titles of units.

With OntoGen is build a “topic *ontology*” from a multitude of documents and this topic ontology is a set of topics related to relations and each topic has a set of documents (see Figure 1).

Figure 1: Topic ontology example



Source: Ontogen (2014)

The data can be processed with the program OntoGen after we prepared a file format as “*naked text*” in which each line contains data – title of the individual publications. In our case, each line of the input file contains the title address of individual publications of the topic on industrial symbiosis and industrial symbiotic networks. With this purpose are data obtained from Web source in XML format and edited with Microsoft Word text editing program. Thus, we obtained a collection of relevant data of industrial symbiosis and industrial symbiotic networks publications which contained titles of publications, among which appeared the titles of collections of scientific papers and electronic books. The consolidated text was then saved in file in plain text format and used then as an input file to be analyse with OntoGen.

The program itself suggests keywords and automatically assigns examples allows text to match the content of the suggested keywords. It allows examples of texts combined in concepts and provides a hierarchical overview

- *other*: restrictions, rules, axioms, events.

The relations into OntoGen have two relations: *is-a (subconcept-of)* and *is-similar* (Ontogen 2014).

Building ontologies the OntoGen program use a sorting technique in groups with the *k-means algorithm*, which classifies input elements into k-groups based on similarity of elements and looking for the most appropriate value of the *parameter k* for this case study. The k-means algorithm sorts the units from the input collection data by groups, so that the similarity of the cases is within individual groups greater compared to the similarity between specimens from different groups. For simplicity and efficiency, the k-means algorithm is one of the most commonly used classification algorithms in practice (Duda et al. 2012). There are three clustering methods implemented in OntoGen: already mentioned k-means, LSI and PH k-means. *“The first two were already part of the first version while the last method is newly added in the latest version. There is also a suggestion method called category which just groups the cases according to the labels included in the input data. The user can select which method the system should use for generating suggestions”* (Fortuna et al. 2001, 4).

In the case study, we founded 267 titles, that had been obtained from articles containing a total of 4546 words. To the list of blocked words, we added conjunctions, numerators, pronouns, etc., that have no special substantive weight. Into OntoGen we can also construct three types of ontologies: visualization – typically in a 2D grouping of documents, method of k-means grouping – mainly used and the third construction of ontologies can be possible as active learning; that means that a user can help a computer program to learn the concepts.

When building ontologies with OntoGen, we decide to have a higher parameter $k = 7$ (number of text groups) relative to the input size of the files, which in our case contained 267 article titles. With this value, we then looked for intermediate settings that the program shows us as a homogeneous as possible and at the same time sufficiently representative groups of titles, which the program names with concepts or keywords. Then we built an ontology with seven automatically proposed concepts (parameter value $k = 7$) at the first level of classification of article titles and added three subconcepts each, that were automatically suggested from OntoGen. In the first concept were sorted 35 titles where the program assigned the following keywords: *development, waste, development_industrial, development_industrial_symbiosis, resource, regional, case_study, industrial_park, park, application*. In the second concept were sorted 33 titles where the program assigned the following keywords: *approach, enterprise,*

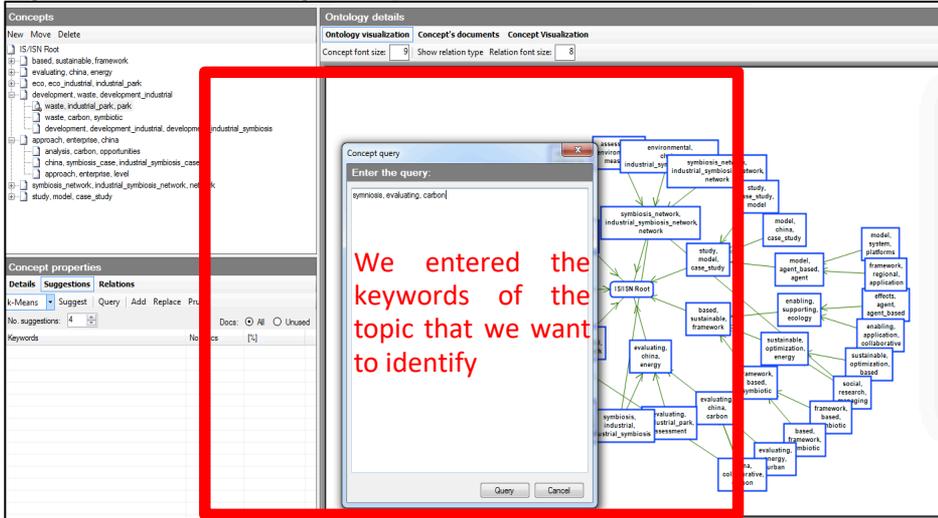
china, case_study, analysis, carbon, study, symbiosis_case, industrial_symbiosis_case, opportunities. In the third concept were selected 52 titles where OntoGen assigned keywords: *symbiosis_network, industrial_symbiosis_network, network, environmental, performance, measures, assessment, economic, indicators, environmental_performance.* In the fourth case, the OntoGen assigned keywords as *study, model, case_study, product, integrating, system, case, china, effects, building* using 35 titles. For the next concept, the fifth, the program sorted 41 titles and assigned *based, sustainable, framework, social, optimization, symbiosis_based, symbiotic, managing, review, application* as keywords. In the sixth case the program assigned keywords: *evaluating, china, energy, case, carbon, assessment, benefits, industrial_park, industrial_symbiosis, park* sorted and selected 27 titles. In the final case, the seventh, OntoGen used 44 titles and assigned *eco, eco_industrial, industrial_park, design, eco_industrial_park, park, system, process, product, analysis.*

The obtained results show that the largest group of publications use the general term of *symbiosis and networks and environmental performance and measures.* Right afterwards are used words as *eco/eco industrial parks and sustainable, framework.*

This feature in OntoGen is a controlled method for adding concepts, based on SVM (Support Vector Machine) active learning method. Enquiring and active learning is only applied to the cases from the selected concept. The idea of this function is that a user can tell/write the concept and some documents related to this concept and than OntoGen can learn the description of this concept and other related documents and discover if those documents are connected and related to the concrete concept.

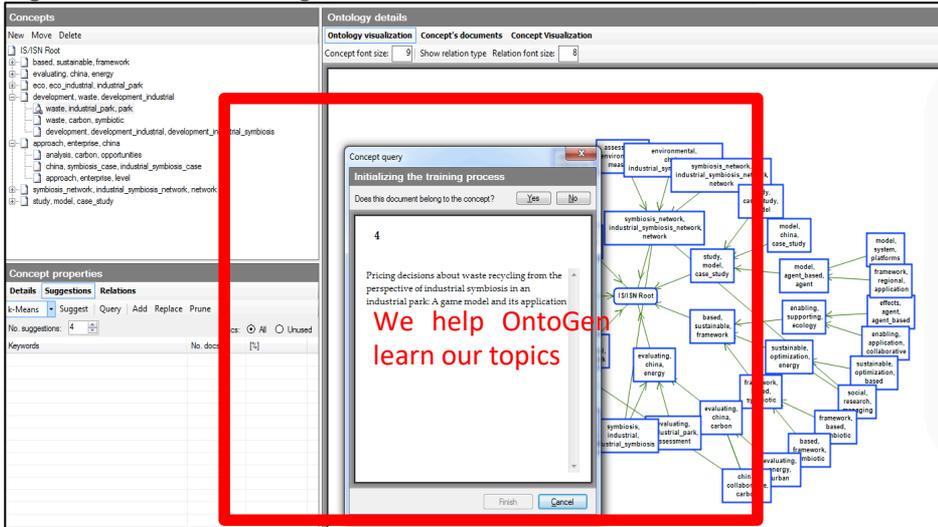
In the feature of *active learning*, a user can start a method by clicking "Query" button in Figure 3. So the system launches a dialogue box that takes the query from the user and after the user enters a query the active learning system starts asking questions and labelling the cases (in Figure 3). On each step, the system asks if a particular case belongs to the concept and the user can select "Yes or No" (Figure 4). *"Questions are selected so that the most information about the desired concept is retrieved from the user. After some initial labelled sample is collected from the user the system displays some additional information about the concept. It displays the current size (number of documents positively classified into the concept) and most important keywords for the concept (using SVM keyword extraction). The user can continue answering the questions or finish by clicking on the Finish button. The more questions that the user answers the more correct assignment of cases in the final concept are"* (Fortuna et al. 2001, 5).

Figure 3: Active learning 1. step



Source: Own data in OntoGen (2020)

Figure 4: Active learning 2



Source: Own data in OntoGen (2020)

With the function *Concept Visualization* (see Figure 5) in OntoGen we can observe four evident groups of keywords from the titles. The group appeared as four white areas.

Figure 5: The grouping of four concepts



Source: Own data in OntoGen (2020)

From Figure 5 we can see that some words are repeated several times in certain documents. Those words are for example *analysis*, *network*, *symbiosis*, *energy*, *case*, *China*, *industrial*, etc. That words can not be added to just one concept, also because the documents were very similar and in this case, we got a fairly diffuse distribution of concepts. However, with the help of the map, we can also identify the main content sub-areas, adding them to the ontology as a subconcepts. This method is also called a map of divided words.

Ontology in this research has seven concepts and content areas, each of which is divided into three subconcepts (see Figure 1). The first concept *eco*, *eco_industrial*, *industrial_park*, *design*, *eco_industrial_park*, *park*, *system*, *process*, *product*, *analysis* is divided into three subconcepts and these have assigned keywords: (1) *eco*, *eco_industrial*, *industrial_park*, *eco_industrial_park*, *park*, *system*, *analysis*, *social*, *urban*, *economy*; (2) *design*, *process*, *product*, *experiences*, *energy*, *objective*, *system*, *optimization*, *multi*, *design_industrial*; (3) *implementation*, *process*, *efficiency*, *methodology*, *role*, *economy*, *circular*, *circular_economy*, *institutional*, *building*.

The second concept *development*, *waste*, *development_industrial*, *development_industrial_symbiosis*, *resource*, *regional*, *case_study*, *industrial_park*, *park*, *application* has (1) *waste*, *industrial_park*, *park*, *case_study*, *application*, *resource*, *symbiosis_industrial*, *model*, *perspective*, *collaborative*; (2) *waste*, *carbon*, *symbiotic*, *managing*, *approach*, *data*, *technology*, *exploring*, *integrating*, *organize*; (3) *development*, *development_industrial*, *development_industrial_symbiosis*, *emergence*, *regional*, *analysis*, *resource*, *experiences*, *clusters*, *network*.

The third concept *approach*, *enterprise*, *china*, *case_study*, *analysis*, *carbon*, *study*, *symbiosis_case*, *industrial_symbiosis_case*, *opportunities*: (1) *analysis*, *carbon*, *opportunities*, *system*, *low_carbon*, *low*, *case_study*, *technical*, *approach*, *industrial_symbiosis_network*, (2) *china*, *symbiosis_case*, *industrial_symbiosis_case*, *study*, *case_study*, *enterprise*, *carbon*, *promote*, *process*, *emergence* and (3) *approach*, *enterprise*, *level*, *network*, *enabling*, *technology*, *innovation*, *study*, *input_output*, *output*.

The fourth concept: *symbiosis_network*, *industrial_symbiosis_network*, *network*, *environmental*, *performance*, *measures*, *assessment*, *economic*, *indicators*, *environmental_performance* is divided into (1) *assessment*, *environmental*, *measures*, *indicators*, *approach*, *effects*, *energy*, *emission*, *reduction*, *benefits*, (2) *environmental*, *chain*, *industrial_symbiosis_network*, *supply_chain*, *supply*, *symbiosis_network*, *economic*, *performance*, *improve*, *collaborative* and (3) *symbiosis_network*, *industrial_symbiosis_network*, *network*, *environmental*, *performance*, *comparable*, *case*, *performance_industrial*, *role*, *structural*. The fifth concept *study*, *model*, *case_study*, *product*, *integrating*, *system*, *case*, *china*, *effects*, *buildings* composed by subconcepts: (1) *study*, *case_study*, *model*, *case*, *building*, *product*, *integrating*, *development*, *study_industrial*, *business*, (2) *model*, *china*, *case_study*, *effects*, *insights*, *optimization*, *flows*, *product*, *integrating*, *resource* and (3) *model*, *agent_based*, *agent*, *agent_based_model*, *based_model*, *system*, *data*, *framework*, *platforms*, *product*.

The sixth concept is *based, sustainable, framework, social, optimization, symbiosis_based, symbiotic, managing, review, application* divided in (1) *enabling, supporting, ecology, platforms, application, review, collaborative, enabling_industrial_symbiosis, enabling_industrial, effects*, (2) *sustainable, optimization, energy, based, managing, symbiosis_based, industrial_park, social, park, research* and (3) *framework, based, symbiotic, social, optimization, analysis, role, business, symbiosis_based, symbiosis_system*, subconcepts. And the final – seventh concept *evaluating, china, energy, case, carbon, assessment, benefits, industrial_park, industrial_symbiosis, park* is constructed from three subconcepts: (1) *evaluating, china, carbon, energy, urban, benefits, product, reduction, collaborative, case*, (2) *evaluating, industrial_park, assessment, park, case, china, cycle, life_cycle, life, energy* and (3) *symbiosis, industrial, industrial_symbiosis, agent, agent_based, evaluating, simulation, based_simulation, indicators*.

6. Discussion

In the first concept of ontology, denoted by the keywords *eco, eco_industrial, industrial_park, design, eco_industrial_park, park, system, process, product, analysis* is 44 titles describing the structure of Eco-Industrial Parks (EIP) composed by several examples of industrial symbiosis, which allow energy/material exchanges among the different industrial enterprises within. Social, economic and environmental performances can be improved in these cases.

The second concept of the ontology includes 47 titles that include *development, waste, development_industrial, development_industrial_symbiosis, resource, regional, case_study, industrial_park, park, application*. In this group, we find the titles including insistent ecological challenges that danger social progress and economic prosperity in the regions, including unsustainable resource management and natural resource reduction, ecosystem degradation and bio-variety damage, contamination, waste and climate changes. This concept is also combining thematics as values of regional cooperation and of the means of implementation of the 2030 Agenda for Sustainable Development to join benefits and promote keys for better environment (Desa 2016).

The third concept of ontology: *approach, enterprise, china, case_study, analysis, carbon, study, symbiosis_case, industrial_symbiosis_case, opportunities*, includes 33 titles considering the specific situation where there are so many state-owned big firm groups as in China and uncovered opportunities of a low-carbon cities promotions with the industrial system.

In the fourth concept of ontology: *symbiosis_network, industrial_symbiosis_network, network, environmental, performance, measures, assessment, economic, indicators, environmental_performance* have 52 titles, among which are studies on ecological characteristics, impacts, whose activities are planned in the actions and socio-economic aspects; firms that measure, achieve and communicate their environmental performances are integrally well placed, by improving their processes, reduce costs, conform with regulatory requirements and stakeholder expectations and take advantage of new market opportunities (Oertwig 2017).

The fifth concept with keywords: *study, model, case_study, product, integrating, system, case, china, effects, building* include 35 titles where the main context focuses on the influence of micro-mezzo-macro economic factors and the social distances to the integration process in industrial symbiotic networks. Projects of macro-economic factors and social distances influence on the integration process of new business objects in some region (Schiller et al. 2014).

The next concept – the sixth is composed of keywords: *based, sustainable, framework, social, optimization, symbiosis_based, symbiotic, managing, review, application* and covered by 41 titles. Point groups or categories including industrial symbiosis points, circularity points, eco-innovation points, sustainability and resource efficiency points, and network strength studies points, were evaluated. Through reviews, a lot of strategies of implementation industrial synergies were stressed and argued for each of the above-mentioned categories (Department of Economic & Social Affairs – United Nations Publications 2008).

The seventh concept within 27 titles: *evaluating, china, energy, case, carbon, assessment, benefits, industrial_park, industrial_symbiosis, park* primarily deals with the issue of the plan stages of the removals and treatments in landfills. These stages generate an important economic and environmental cost, mainly caused by carbon dioxide emissions due to the transports. Therefore, several actions have been taken to encourage and disseminate eco-industrial development, such as the implementation of eco-industrial parks (EIPs) and their networks (Demirbas et al. 2004).

7. Conclusion

In the construction of terminological ontologies based on the titles on a field of industrial symbiosis and industrial symbiotic networks, we found that the precision of modeling and demonstrations of domain knowledge mostly depends on the settings of the text analysis parameters as in this case (paper)

is the value of the parameter “k”. K-means algorithm is used to sort input elements (article titles in our case) in “k groups” based on the substantive similarity of the elements. As the best choice in the case study was the division of input texts into seven groups (concepts) that was established at the first hierarchical level of the ontology and with three subgroups (subconcepts) at the second level. The size of the input file is also important and in present the number of titles covered in the primary reserch. Input data must be sufficient quantitatively, so the program tool would be able to analyze them and obtain meaningful analysis and representative results.

The fourth concept of ontology is numerically the most extensive, as it includes 52 titles. The articles that are mostly categorized in the fourth concept of ontology are: *environmental performances, measures, assessment, economic indicators*. Much of the articles describe the development of industrial symbiosis and their combination with other analytical methods. Among the results, we also expected more data about the networks that are spinning arround into the regional cases of industrial symbiosis.

OntoGen is recognised to be a useful tool for modeling domain knowledge. OntoGen is very logical in a detailed review of keywords that are automatically identified. Another plus of OntoGen is, that helps to recognize keywords displayed that could be manually improved or renamed (Bizjak 2014, 35–39). For an even more detailed implementation of domain knowledge it would also make sense to include summaries of articles from the period under consideration and thus obtain even more detailed demonstrations of publications about the field of industrial symbiosis and their networks.

References

Awad, Elias. M., Ghaziri, Hassan M. 2007. *Knowledge Management*, Dorling Kindersley (India) Pvt. Ltd., Delhi.

Bajwa, Imran Sarwar, Kamareddine, Fairouz, Costa, Anna. 2019. *Intelligent Technologies and Applications: First International Conference, INTAP 2018, Bahawalpur, Pakistan, October 23–25, Revised Selected Papers*. Vol. 932. Springer.

Van Beers, Dick, Corder, Glen David, Bossilkov, Alben, Van Berkel, Rene. 2007. “Regional synergies in the Australian minerals industry: Case-studies and enabling tools.” *Minerals Engineering* 20 (9): 830–841.

Black Sea Industrial Symbiosis Platform. 2017. European MSP Platform. Accessed 9 January 2017. <http://msp-platform.eu/practices/black-sea-industrial-symbiosis-platform>.

Chertow, Marian R. 2000. "Industrial Symbiosis: Literature and Taxonomy." *Annual Review of Energy and the Environment* 25 (1): 313–37. CRESSI Publications. Saïd Business School. 2017. <https://www.sbs.ox.ac.uk/faculty-research/research-projects/creating-economic-space-social-innovation-cressi/cressi-publications>. Accessed 26. 10. 2017.

Demirbas, M. Fatih, Bozbas, Kahraman, Balat, Mustafa. 2004. "Carbon Dioxide Emission Trends and Environmental Problems in Turkey." *Energy Exploration & Exploitation* 22 (5): 355–65.

Department of Economic & Social Affairs. 2008. *Annuaire Statistique*. United Nations Publications.

Desa, U. N. 2016. Transforming our world: The 2030 agenda for sustainable development.

Džajić Uršič, Erika. *Morphogenesis of Industrial Symbiotic Networks*. Peter Lang GmbH, Internationaler Verlag der Wissenschaften, 2020.

Duda, Richard O., Hart, Peter E., Stork, David G. 2012. *Pattern classification*. John Wiley & Sons.

Ehrenfeld, John, Gertler, Nicholas. 1997. "Industrial Ecology in Practice: The Evolution of Interdependence at Kalundborg." *Journal of Industrial Ecology* 1 (1): 67–79.

Fayyad, Usama. M. 1996 "From data mining to knowledge discovery in Databases: an Overview." *Advances in Knowledge Discovery and Data Mining*.

Fayyad, Usama M., Piatetsky-Shapiro, Gregory, Smyth, Padhraic. 1996. "From data mining to knowledge discovery in databases." *AI Magazine* 17 (3): 37–54.

Fortuna, Blaž, Grobelnik, Marko, Mladenič, Dunja. 2001. *Instructions for OntoGen 2.0*. <http://analytics.ijs.si/~blazf/ontogen/OntoGen2Help-2006-09-11.pdf>. Accessed 17. 8. 2020.

———. 2006. Semi-automatic data-driven ontology construction system. *In Proceedings of the 9th International multi-conference Information Society IS-2006, Ljubljana, Slovenia*.

———. 2007. OntoGen: semi-automatic ontology editor. *In Symposium on Human Interface and the Management of Information*, pp. 309–318. Springer, Berlin, Heidelberg.

Fric, Urška, Rončević, Borut, Džajić Uršič, Erika. 2020. “Role of Computer Software Tools in Industrial Symbiotic Networks and the Examination of Sociocultural Factors.” *Environmental progress & sustainable energy* 39 (2): 1–7, doi: 10.1002/ep.13364.

Fric, Urška. 2016. “Medorganizacijsko zaupanje in industrijska simbioza-predstavitev projekta/Interorganization confidence and industrial symbiosis-project presentation.” *Raziskave in Razprave* 9 (1–3): 233.

Gibbs, David, Deutz, Pauline. 2005. “Implementing Industrial Ecology? Planning for Eco-Industrial Parks in the USA.” *Geoforum* 36 (4): 452–64.

———. 2007. “Reflections on Implementing Industrial Ecology through Eco-Industrial Park Development”. *Journal of Cleaner Production*, 15 (17): 1683–95.

Girard, John P. Girard, John L. 2015. “Defining knowledge management: Toward an applied compendium.” *Online Journal of Applied Knowledge Management* 3 (1): 1–20.

Gladstone Industry Leadership Group (GILG). <https://gilg.com.au/>. Accessed 26. 10. 2017.

Guarino, Nicola, Giaretta, Pierdaniele. 1995. Ontologies and knowledge bases. *Towards very large knowledge bases*: 1–2.

Howard-Grenville, Jennifer, Paquin, Raymond. 2008. “Organizational Dynamics in Industrial Ecosystems: Insights from Organizational

Theory." *Changing Stocks, Flows and Behaviors in Industrial Ecosystems*, 122–39.

Kalundborg Symbiose – Verdens Første Industrisymbiose. <http://www.symbiosis.dk/>. Accessed 26. 10. 2017.

Lavbič, Dejan, Krisper, Marjan. 2005. "Semantika podatkov in ontologije." *Uporabna informatika* 3 (13): 121–135.

Macy, Michael W., Willer, Robert. 2002. "From Factors to Actors: Computational Sociology and Agent-Based Modeling." *Annual Review of Sociology* 28 (1): 143–66.

Makarovič, Matej, Šušteršič, Janez, Rončević, Borut. 2014. "Is Europe 2020 set to fail? The Cultural Political Economy of the EU Grand Strategies." *European Planning Studies* 22 (3): 610–626.

Mileva Boshkoska, Biljana, Rončević, Borut, Džajić Uršič, Erika. 2018. "Modeling and evaluation of the possibilities of forming a regional industrial symbiosis networks." *Social sciences* 7 (1), doi: 10.3390/socsci7010013.

Oertwig, Nicole, Galeitzke, Mila, Schmieg, Hans-Georg, Kohl, Holger, Jochem, Roland, Orth, Ronald, Knothe, Thomas. 2017. Integration of Sustainability into the Corporate Strategy. In *Sustainable Manufacturing: Challenges, Solutions and Implementation Perspectives*, edited by Rainer Stark, Günther Seliger, and Jérémy Bonvoisin, 175–200. Sustainable Production, Life Cycle Engineering and Management. Cham: Springer International Publishing.

OntoGen. <http://ontogen.ijs.si>. Accessed 22. 7. 2020.

Ontogen. 2014. Materiali. <http://ropot.ijs.si/UNG2014/materiali/ontogen.pdf>. Accessed 22. 7. 2020.

P7 Modelling industrial symbiosis to find the potentials and barriers in Aalborg, Denmark. 2011. <http://www.symbiosis.dk/sites/default/files/P7%20Industrial%20Symbiosis%20Aalborg%20University%20Christoffer.pdf/>. Accessed 2. 7. 2017.

Poli, Roberto, Seibt, Johanna. 2010. *Theory and Applications of Ontology: Philosophical Perspectives*. Springer Science & Business Media.

Rončević, Borut. 2019. "Cultural Political Economy of Europe 2020: Jean Monnet Chair CPE 2020 and its Impact." *Research in Social Change* 11 (2): 5–13.

Rončević, Borut, Besednjak Valič, Tamara. 2019. "How to Think about Regional Development Agencies as a Sociologist." *The Social Sciences* 14 (9): 326–34.

Schiller, Frank, Penn, Alexandra S., Basson, Lauren. 2014. Analyzing Networks in Industrial Ecology – a Review of Social-Material Network Analyses. *Journal of Cleaner Production* 76 (August): 1–11.

Sciadirect. <https://www.sciencedirect.com/>. Accessed 2. 1. 2020.

Scott, W. Richard. 2004. "Reflections on a Half-Century of Organizational Sociology." *Annual Review of Sociology* 30 (1): 1–21.

Sebastiani, Fabrizio. 2002. "Machine Learning in Automated Text Categorization." *ACM Computing Surveys (CSUR)* 34 (1): 1–47.

Sowa, John F. 2000. Ontology, Metadata, and Semiotics. In *Proceedings of the Linguistic on Conceptual Structures: Logical Linguistic, and Computational Issues*, pp. 55–81.

Szilva, Ivan, Caganova, Dagmar, Bawa, Manan, Pechanova, Lubica, Hornakova, Natalia. 2018. Knowledge Management Perception in Industrial Enterprises within the CEE Region. In *Cloud Infrastructures, Services, and IoT Systems for Smart Cities*, edited by Antonella Longo, Marco Zappatore, Massimo Villari, Omer Rana, Dario Bruneo, Rajiv Ranjan, Maria Fazio, and Philippe Massonet, 66–75. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering. Cham: Springer International Publishing.

Tudor, Terry, Adam, Emma, Bates, Margaret. 2007. "Drivers and Limitations for the Successful Development and Functioning of EIPs (Eco-Industrial Parks): A Literature Review." *Ecological Economics* 61 (2–3): 199–207.

Vogrinčič Bizjak, Jaka. 2014. Modeliranje znanja in prikaz v obliki ontologij s študijo primera raziskav v okolju: magistrsko delo. Univerza v Novi Gorici, Poslovno-tehniška fakulteta.