An Integrated Method for Information and Communication Technology (ICT) Supported Energy Efficiency Evaluation and Optimization in Manufacturing

Knowledge-based Approach and Energy Performance Indicators (EnPI) to Support Evaluation and Optimization of Energy Efficiency

Zur Erlangung des akademischen Grades **Doktor der Ingenieurwissenschaften** der Fakultät für Maschinenbau des Karlsruher Instituts für Technologie (KIT)

genehmigte

DISSERTATION

von

M. Sc. Hendro Wicaksono

Tag der mündlichen Prüfung: 30.08.2016

Hauptreferent: Prof. Dr. Dr.-Ing. Dr. h. c. Jivka Ovtcharova (KIT)

Korreferent: Prof. Dr.-Ing. Clemens Felsmann (TU Dresden)

Abstract

This thesis develops an approach which utilizes information and communication technologies and supports a holistic evaluation of energy efficiency in discrete manufacturing. The evaluation is performed both qualitatively and quantitatively to achieve accurate energy efficiency measures. The approach focuses on improving "Check" activities in the PDCA cycle of an ISO 50001 energy management system. The qualitative evaluation concentrates on the criteria that can be observed but not measured in figures. Thereby, the correlations between object states, such as energy-wasting situations or energy-efficient best practices are assessed. Those correlations are modeled in a knowledge base. The quantitative evaluation, by contrast, is performed by assessing energy efficiency using energy performance indicator (EnPI).

Compared to other related work, the innovation of this thesis consists in the integrated method applying an expressive and adaptive ontology knowledge base that is capable of learning as well as a sector-independent, straightforward, and directly usable energy performance indicator (EnPI) for evaluating and benchmarking different processes and organization levels within a company. Furthermore, this thesis utilizes the ontology knowledge base and EnPI in production planning and control, or more specifically, in production scheduling by means of a hyper-heuristics-based energy-optimized, flexible, and user-configurable scheduling approach.

The approach is validated using software prototypes and a system integration architecture that are applied in real manufacturing environments. These include a make-to-order manufacturing company, a serial manufacturing company, and a production system involving energy conversion, transport, ancillary, and production facilities.

Acknowledgements

This thesis is the result of my research work carried out from November 2012 to April 2016 at Institute of Information Management in Engineering (IMI), Karlsruhe Institute of Technology (KIT) and Research Center for Information Technologies (FZI).

Firstly, I would like to express my sincere gratitude to my advisor Prof. Dr. Dr.-Ing. Dr. h. c. Jivka Ovtcharova, head of IMI at KIT, for encouraging my research and for allowing me to grow as a research scientist. Her advice on both research as well as on my career has been invaluable.

I would also like to thank Prof. Dr.-Ing. Clemens Felsmann from Institut für Energietechnik, Technische Universität Dresden for his willingness to take on the role as co-supervisor. I also thank Prof. Dr. rer. nat. Frank Gauterin for being the chairman of the examination committee.

My sincere thanks also go to my colleagues at IMI and FZI, especially Dr. Sven Rogalski, for his support and willingness to discuss many things, Preslava Krahtova, Kiril Tonev, Kiril Aleksandrov, Viktor Schubert, and many more for the teamwork in many research projects, and for all the fun we have had. Also, I would like to thank my students Ernst Victor Prohl, Tim Belzner, Ardhi Irsyadi, Viet Bui, and many other students for their dedication and contributions to our projects.

Finally, I would like to thank my family for all their love and encouragement, especially to my parents who always support me in all my pursuits and make my studies possible. This work is dedicated to my mother who passed away last year.

Hendro Wicaksono August 2016

Table of Contents

List of Acronyms

AC	Alternating Current
ADACOR	Adaptive Holonic Control Architecture for Distributed Manufacturing Systems
AmI-MoSES	Ambient-Intelligent Interactive Monitoring System for Energy Use Optimization in Manufacturing SMEs
CAD	Computer Aided Design
CBR	Case-Based Reasoning
CRM	Customer Relationship Management
CSV	Comma Separated Values
DL	Description Logics
DTF	Deepest Task First
EBI	Energieeffizienz-Bedeutungs-Index
EDD	Earliest Due Date
EEI	Energy Efficiency Index
EEP	Energy Efficiency in Production
EII	Energy Efficiency Investment Index
EMDA	Energy Monitoring and Data Acquisition
EnMS	Energy Management System
EnPI	Energy Performance Indicator
ERP	Enterprise Resource Planning
ESI	Energy Saving Index
ESTOMAD	Energy Software Tools for Sustainable Machine Design
EWOTeK	Effizienzsteigerung von Werkzeugmaschinen durch Optimierung der Technologien zum Komponentenbetrieb: enhancing the efficiency of machine tools by optimizing the technologies for operating components
FSSP	Flow Shop Scheduling Problem
FU	Functional Unit
GB	Gigabyte
GDP	Gross Domestic Product
GEF	Graphical Editing Framework
GHz	Gigahertz

GSSP	Group Shop Scheduling Problem
HEMS	Home Energy Management System
HVAC	Heating, Ventilation, and Air Conditioning
ICT	Information and Communication Technology
IDE	Integrated Development Environment
IEA	International Energy Agency
IEC	International Electrotechnical Commission
ISES	Intelligent Services for Energy Efficient Design and Life Cycle Simulation
ISO	International Organization for Standardization
JDBC	Java Database Connectivity
JSSP	Job Shop Scheduling Problem
KDD	Knowledge Discovery in Database
KM	Knowledge Management
LFT	Longest Finishing Time
MASON	Manufacturing Upper Ontology
MES	Manufacturing Execution System
MSE	Manufacturing System Engineering
MSSP	Mixed Shop Scheduling Problem
OSGi	Open Service Gateway initiative
OSSP	Open Shop Scheduling Problem
OWL	Web Ontology Language
PC	Personal Computer
PDM	Product Data Management
PLC	Product Lifecycle
PPC	Production Planning and Control
QI	Quality Improvement
RBR	Rule-Based Reasoning
RCP	Rich Client Platform
RDF	Resource Description Framework
RDFS	Resource Description Framework Schema
regex	Regular Expression
REST	Representational State Transfer

SERENE	Smart and Integrated Energy Efficiency Evaluation
SERUM-IB	Smart Energy and Resource Usage Management in Building
SME	Small and Medium Enterprise
SOAP	Simple Object Access Protocol
SPC	Statistical Process Control
SWRL	Semantic Web Rule Language
SWT	Standard Widget Toolkit
TRL	Technology Readiness Levels
VDI	Verein Deutscher Ingenieure
WEKA	Waikato Environment for Knowledge Analysis
WPML	Work Process Modeling Language

List of Figures

List of Tables

1 Introduction

"Contrary to popular belief, we do not face a choice between economy and ecology unless we protect resources and the earth's natural capital."
~ Kofi Annan, UN Secretary-General (1997-2006)

1.1 Background Situation and Motivation

Energy and resource efficiency has been developing to one of the most crucial issues of the 21st century. The environmental goals, finite nature of fossil energy sources, and a constant population increase will lead this issue being considered to an increasing extent when defining environmental policies in the future. Continuous industrialization, especially in emerging countries, has led to a significant increase in energy consumption in the past decade. The International Energy Agency (IEA) reported alarming figures of carbon dioxide emissions in 2011, which reached a record value of 31.6 gigatonnes [IEA-12]. These alarming figures are mainly due to combustion of fossil fuel, such as coal, oil, and gas. They may prevent the goal of the 2-degree limit on global warming adopted by the G8 summit from being reached. The 2-degree limit means that by 2050 the carbon dioxide emissions will have to be half of the value of 1990 [Mer-09]. If energy continues to be generated by conventional means such as by the combustion of fossil resources, an increase in the carbon dioxide emissions cannot be avoided.

The world has taken several steps to overcome the problems. Political discussions and agreements, such as the Kyoto Protocol in 1997, are the reactions of the world to counteract these emerging problems. Industrial countries were asked to keep their emission on the same level or to reduce them by 6 - 8% until 2012. Climate policy of the German government has the defined objective to reduce carbon dioxide emission to 30-40% of the value of 1990 by 2020 [BMU-11]. The 2007 G8 summit pointed out that energy efficiency is one of the critical points to reduce gas emission [Mand-07]. The Federal Government of Germany later published an article on the official G8 summit website entitled *"G8: Verantwortung für Klima und Umwelt"* (G8: Responsibility for Climate and the Environment). This article outlined a statement that globally improved energy efficiency is the most sustainable and economically efficient option to reduce greenhouse gasses and to simultaneously improve the reliability of energy supply [BUN-07]. The term energy efficiency introduced here is defined as a reduction in the energy used for a given service such as heating and lighting) or level of activity [WEC-11].

Manufacturing that has a share of 22% in the GDP is one of the key industries in Europe. Around 70% of all jobs in the European Union directly and indirectly depend on manufacturing [Manu-04] [ECIS-09]. The discrete manufacturing sector that is the focus of this thesis employs 34 million people in Europe and has total revenue of around 1,500 billion Euros. This means that 30% of all workplaces in Europe are offered by discrete manufacturing industry [Frau-08]. The European Committee of the Machine Tool Industries (CECIMO) reported that machine tool industries that belong to the discrete manufacturing sector employ around 150,000 people at 1,500 companies in 15 European countries [Geer-13].

The significant impacts of manufacturing on public welfare make its energy efficiency efforts a critical contribution to achieving European and G8 policy objectives. Meanwhile, a study performed by the US the Energy Information Administration has revealed that this industry sector has the highest share of the total energy consumed by the end user (see Figure 1.1). Most of the consumption cover electric power and gas [UEIA-13]. Energy efficiency in manufacturing contributes not only to reduce CO2 emission but also to reducing Europe's dependence on energy imports. Thus, it helps to limit the inflation [ECIS-09]. In 2005, energy consumption in the industry was around 17-18% of total energy consumption in European Union countries, and more than 90% of this fraction were consumed by manufacturing sector [Mann-07].

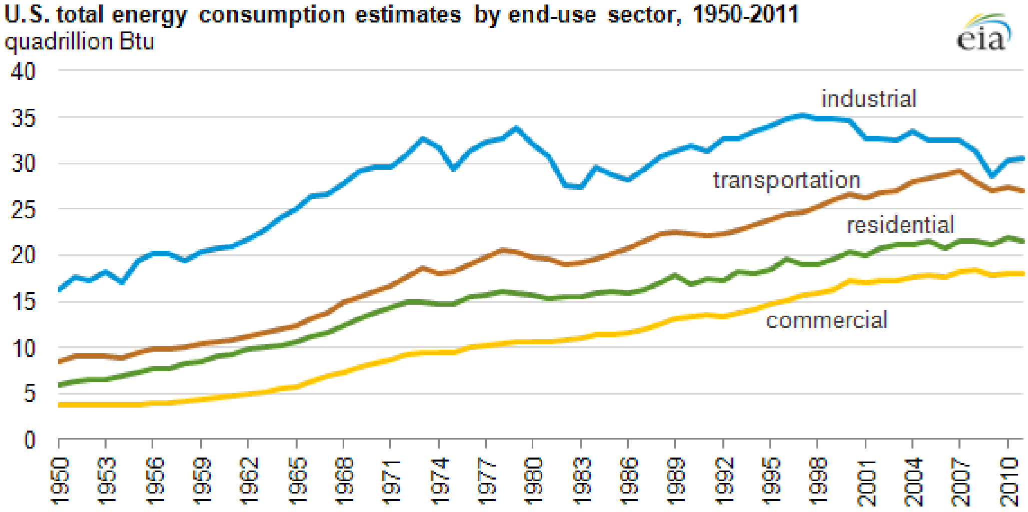


Figure 1.1 Energy consumption categorized by sector in the US [UEIA-13]

The European Energy Agency reported that the final energy consumption in the European Union increased by 7.1% between 1990 and 2010. However, it decreased by 3.2% between 2005 and 2012. This fact is depicted in Figure 1.2. As seen in the figure, industries have the highest final energy consumption of electricity up to 36.5% and second-highest final energy consumption of natural gas up to 32.5%. Although the final energy consumption of industry sector fell at an annual average of 1.1% between 1990 and 2010 because of a shift towards less energy-intensive manufacturing industries, there are still many improvement potentials through optimization of process management, notably through the application of Information and Communication Technologies (ICT) [ECIS-09] [EEA-13].

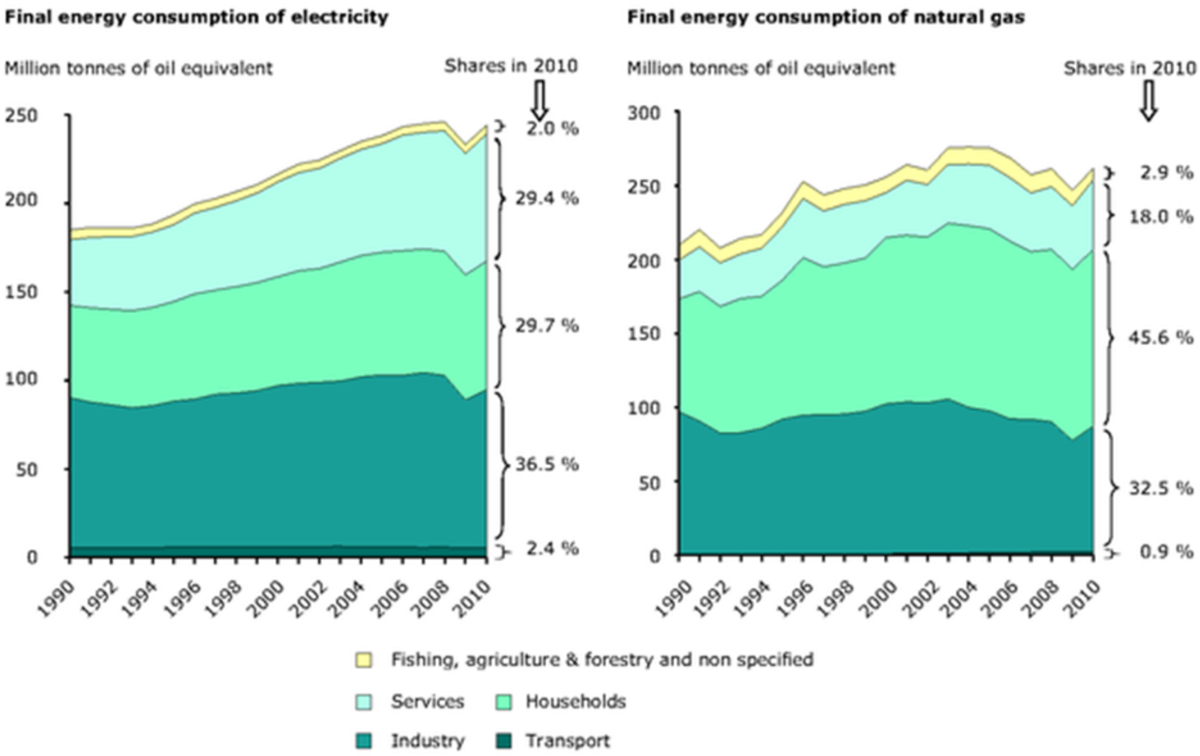


Figure 1.2 Total final energy consumption by sector in EU 1990-2010 [EEA-13]

Besides ecological, political, and social motivations, economic reasons play a significant role for manufacturing companies to consume energy in a more efficient way. Under today's market conditions, customers dominate manufacturers. The manufacturers are to fulfill the varied and complex customer requirements. Those manufacturers who fail to satisfy the customers' requirements will not be competitive on the market. They will lose profit because of decreasing sales in their respective market. According to a survey conducted in Germany, variation and diversity of metal and electrical industry products have increased. This increase is reflected by Figure 1.3 [Kink-05].

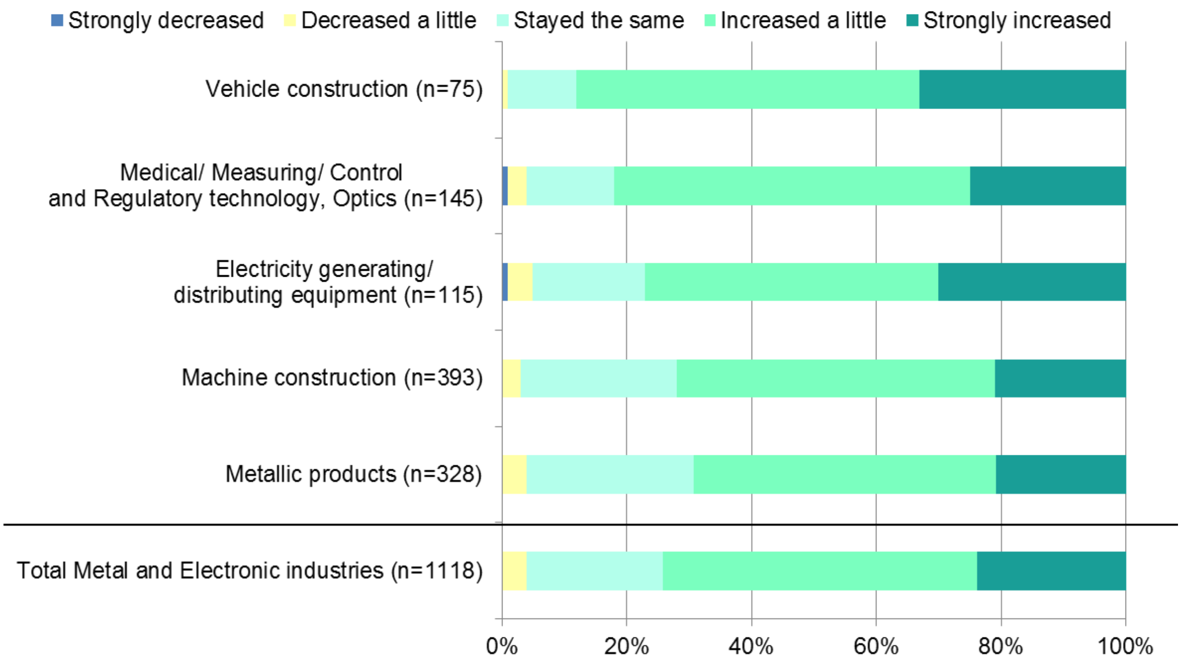


Figure 1.3 Increasing variant diversity in German metal and electrical industries [Kink-05]

Product variations have made manufacturing processes extremely complex. Manufacturers need to be flexible to satisfy the different customers' demands. Customers often demand complex and highly precise products. This makes the situation even more difficult for the manufacturers [Kink-05, p. 8]. The resulting manufacturing processes often are very energy-intensive and, hence, expensive. The continuing rise in energy prices is to be considered by manufacturing companies to an increasing extent since it may significantly affect their margin. Thus, energy efficiency, combined with the economization of production costs, is an essential factor of competitiveness, especially for energy-intensive industries. Energy efficiency and conservation, thus, start to play more decisive roles in business strategies of the manufacturing sector. Manufacturing companies need to find better, cheaper, and more efficient solutions. Large, as well as small and medium-sized enterprises, are required to consider energy efficiency as an important factor in their business strategy. Moreover, manufacturers need to produce high quality and individual products at lowest possible costs. In addition, they have to deliver the products as fast as possible that makes the product lifecycle shorter. These requirements for modern manufacturing environment are depicted in Figure 1.4.

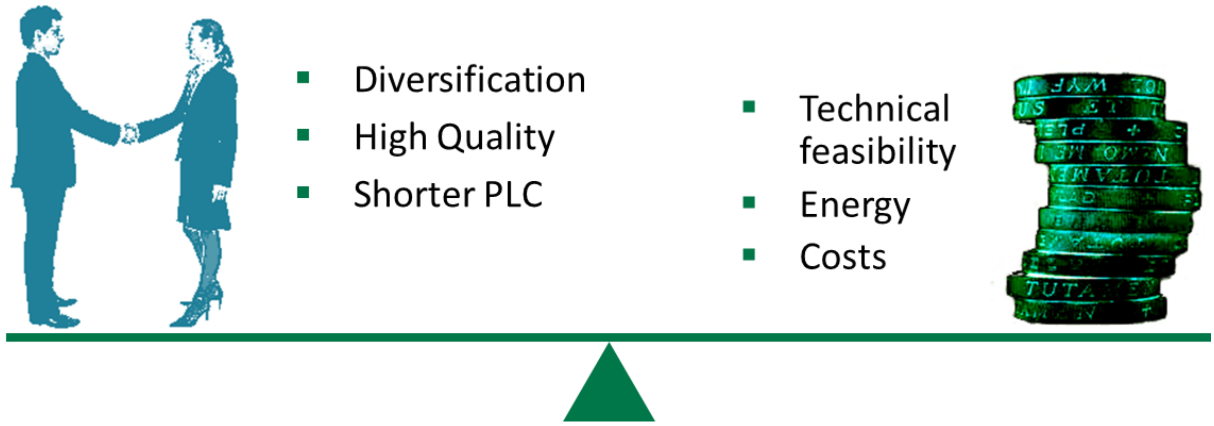


Figure 1.4 Requirements for modern manufacturing environment

A common framework to address these challenges is an energy management scheme which is to ensure minimal energy consumption of all processes and measures to cover a given demand. An energy management system (EnMS) is a systematic way to define the energy flows and a basis for the decision to improve energy efficiency. An EnMS includes the implementation of organization and information structures necessary for energy management, including the tools required for this purpose [KaEt-10]. ISO 50001 is used as the standard of EnMS in industrial companies. It describes the requirements for the implementation of the EnMS. To implement the EnMS, a company has to generate energy reports. Currently, no common reporting infrastructure makes the evaluation of energy performance more accurate.

Most of the manufacturing companies face problems in implementing the energy management standards. There are often no standardizations in their portfolio of operations such as plants, and sites [Meye-11]. The energy-consuming production processes, building infrastructures, and power plants are documented and managed separately and in an unstructured manner. The management is hardly aware of the consumption of energy in operation due to a knowledge gap between the staff members, such as managers and operators. Managers do not have the tools to manage energy use covering different levels of the organization in vertical and horizontal direction. The use of energy supplied by utility companies or power plants is not organized well in operation, thus leading to energy waste. Operators are often not aware of whether their activities and decisions create excessive energy use or demand. This problem is also due to the knowledge gap between them and a lack of a shared vocabulary. Best practices to avoid energy wasting are known by some employees. Other staffs do not have access to this best practice knowledge. Additionally, there are also interoperability problems among current technologies applied in manufacturing and energy system environments. These problems are caused by different vocabularies, application schemas, and data formats. The company thus is forced to correlate various types of data to support analysis and decision making [PaMo-08] [RuHV-13] [WeEt-15].

Energy consumption data are commonly stored in energy monitoring systems or simply in paper-based reports. Relations between the measured energy consumption and operational events, business processes, and resources in the company are lacking. The resulting allocation of energy costs to produced goods and services is neither accurate nor useful for understanding the actual cost of the product. Furthermore, the corporate reporting procedure using existing Key Performance Indicators (KPIs) related to energy is time-consuming and often not accurate [Meye-11]. No mechanism indicates the energy performance of different processes, resources, and organization parts within the company.

An important step in energy management is to measure energy consumption at various levels of granularity. This can be done using data acquisition systems. A data acquisition system consists of signals, sensors, signal conditioning, hardware, and a computer with software. Due to the various technologies of data acquisition systems, connectivity and standardization are important issues in establishing an automated energy management system. The meters have to be read manually. Thus, accuracy problems arise. Generally, information and communication technology (ICT) plays a major role in supporting energy saving in industrial companies, especially in industrial process and logistics optimization. Table 1.1 shows the energy saving potentials resulting from different techniques.

Table 1.1 Energy saving potential through various technologies [ECIS-09]

Process optimization	25-30%
Optimized logistics	16%
Integrated process chains	30%
Development of new products	10-40%
Intelligent motor drives	20-40%
Alignment with best performers	15%

According to ISO 50001, the baseline and objective criteria for an energy management system, such as the energy performance indicator (EnPI), are defined in the “Plan” phase. The objectives are implemented in the “Do” phase and evaluated during the “Check” phase. When implementing the energy management system, ICT can accelerate the evaluation process based on the objectives defined due to more automatic and intelligent measurements, data collection, data processing and analysis. The evaluation can be performed in two ways: Qualitatively and quantitatively. Qualitative evaluation focuses on the criteria that can be observed, but not measured in figures. Quantitative evaluation means an evaluation of energy efficiency using measurement figures. By using the figures, it is possible to determine the energy efficiency degree in different organizational parts or processes of a company.

1.2 Research Objectives

Based on the described problem, this thesis aims at developing an integrated approach utilizing information and communication technology to support a holistic evaluation of energy efficiency. The approach considers both qualitative and quantitative evaluation. This thesis focuses on the evaluation of energy efficiency in discrete manufacturing where the processes and products are considered as distinct entities.

The evaluation is performed by employing a knowledge base which is accessible by all staff members and systems in the company. The best practices and knowledge about energy-efficient and energy-wasting activities are standardized, formalized, and stored in the knowledge base. The knowledge base is used as the reference system to perform the evaluation. Based on it, this thesis has the following first objective:

Objective 1. *Development of a concept to support the evaluation using a manufacturing energy management knowledge base.*

In order to develop the knowledge base to support the evaluation, this thesis should answer the following questions:

- How is the knowledge represented in the knowledge base so as to allow for a shared understanding of different human and machine stakeholders involved in the energy management system?
- What are the methods to efficiently generate the knowledge base? How can the knowledge be obtained automatically using learning techniques?
- How can queries be made to the knowledge base in order to evaluate the current energy efficiency states?

An EnPI (energy performance indicator) quantifies energy performance and helps in the quantitative evaluation of energy efficiency in production. The EnPI is used to measure the energy efficiency degree and to benchmark different processes or levels of production organization within a company. Evaluation is to provide criteria that allow for the selection of relevant variables to simulate their effects on energy efficiency. Hence, the second objective of the thesis is as follows:

Objective 2. *Development of an energy performance indicator (EnPI) to measure the energy efficiency degree of different manufacturing organization*

In order to develop the EnPI, this thesis has to answer the following research questions:

- Which organization parts are considered to measure energy efficiency?
- What are the input parameters required to calculate the EnPI?
- How can the EnPI be represented to make the quantitative evaluation easier and straightforward?

- What does the calculation model of the EnPI look like?

One of the important aspects of manufacturing that significantly affect the energy performance of the company is production planning and control. This thesis focuses on production scheduling of discrete manufacturing as a specific field of production planning and control since many possible energy performance optimizations and improvements can be made, for example, energy savings and power load optimization. Thus, the third objective of this thesis is formulated as follows:

Objective 3. *Development of an approach for energy optimized production scheduling by using the knowledge base and EnPI*

- How can the scheduling be described and modeled in a flexible way but still represents real problems?
- How can the knowledge base be used for model development?
- Which optimization methods or algorithms are used to achieve energy-optimized production scheduling?

1.3 Thesis Organization

The research questions formulated in Section 1.2 leads to an approach to reaching the objectives defined. Figure 1.4 presents the approach to reach the objectives and how the approach is organized in the thesis.

Chapter 1 gives an introduction for the readers. First, it outlines the background that motivates this thesis. It deals with the importance of energy efficiency from both the ecological and the economic point of view. The challenges faced by manufacturing companies in implementing the energy management system in order to improve energy efficiency are also highlighted. Based on these problems and challenges, the main objectives of the thesis are defined. The objectives are pursued by means of research questions as guides to develop the thesis approach.

Chapter 2 discusses the fundamental theories and concepts of manufacturing, energy management, knowledge management and engineering, energy performance indicators, and optimization in manufacturing. Then, the state of the art of existing approaches relating to these concepts is evaluated and presented. The criteria to evaluate the different methods are also outlined to identify the strength and weakness of each approach as well as the need for further research. The analysis reveals how relevant this research is to the real problems, where this research is to be positioned compared to other research activities, and also whether the problems have already addressed by other researchers.

Based on the objectives defined in chapter 1 and the state of the art analysis in chapter 2, the requirements to be met by the concepts developed are formulated in **Chapter 3**. The requirements determine the development of the concepts and define the points that have to be fulfilled.

In **Chapter 4**, the concepts that are developed in this thesis are discussed in detail. First, an overview of the whole approach is given. Then, the steps needed to construct, represent, and use the knowledge base as the tool to support qualitative evaluation are described. This also includes knowledge generation through the learning technique. Energy system modeling, influencing parameters, and the calculation model to determine the energy performance indicators are presented afterward. This chapter ends with concepts for both energy- and cost-optimized production scheduling.

The concept is implemented as software prototypes which are described in **Chapter 5**. Verification of the results against the defined requirements is also discussed in this chapter. The chapter additionally presents the validation of the concepts in three different manufacturing environments. The last **Chapter 6** presents the summary of the scientific, technical, and industrial contributions of this thesis. It also gives an outlook on further research and development and gives recommendations for the application of the developed methods in practice.

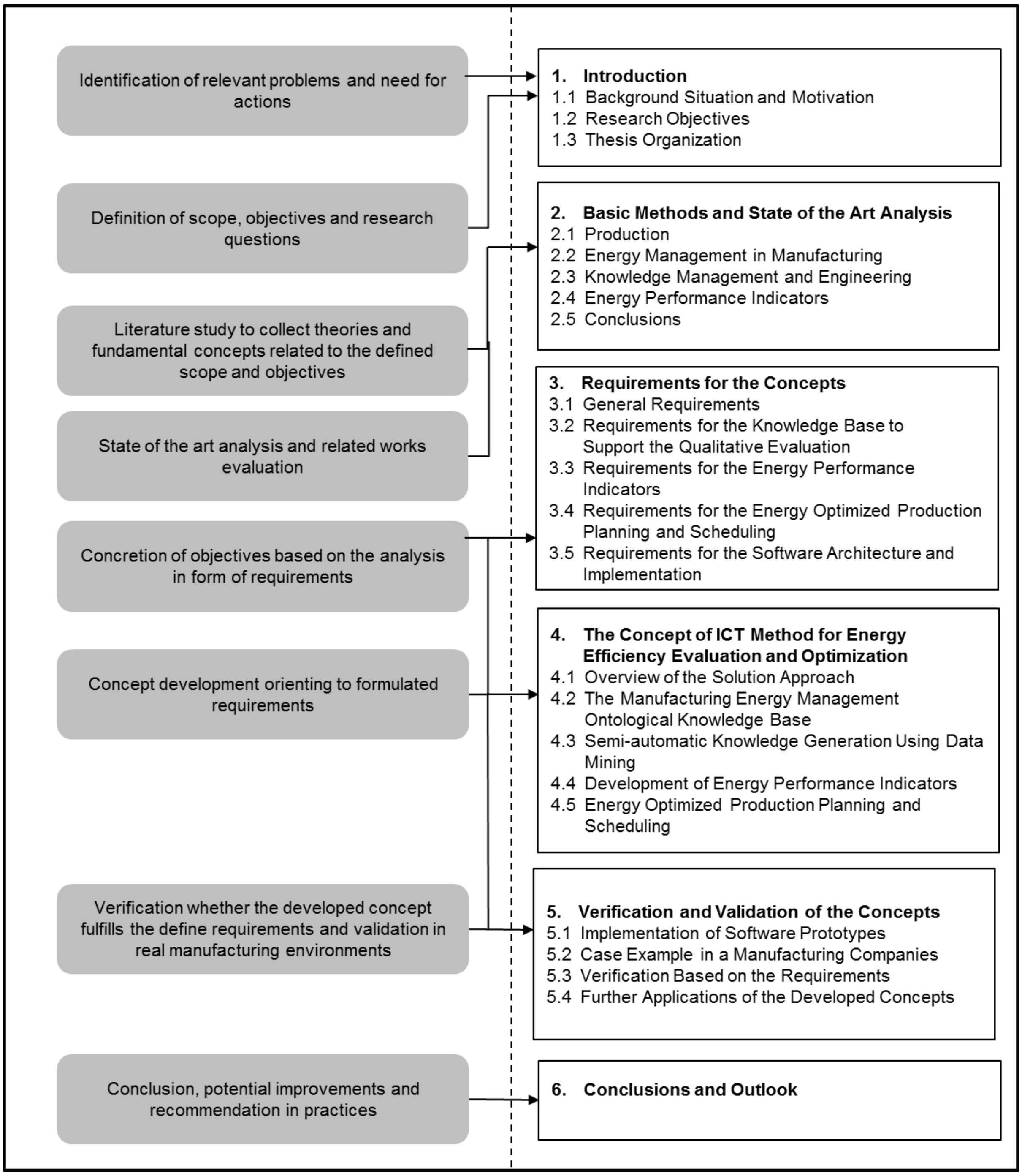


Figure 1.5 Thesis organization

2 Basic Methods and State of the Art Analysis

"Research is to see what everybody else has seen, and to think what nobody else has thought."
~Albert Szent-Gyorgyi (1893-1986)

This chapter aims to help the reader to get familiar with the problems and topics that are related to this work. It begins with the section that introduces production system as well as production planning and control which are considered as the application fields of the concepts developed in this thesis. It continues with a section containing the definitions of energy efficiency, energy management, applications of ICT and their focus on manufacturing. The basic concepts and state of the art of knowledge management and engineering approaches are outlined in the next section. The concepts are essential because this work develops a knowledge base system to support the energy efficiency evaluation as one of the important phases of energy management. Afterwards, a section that describes basic concepts of energy efficiency indicators and some related work is provided. Finally, this chapter ends with a conclusion outlining the main points that strongly pertains to the concept developments.

2.1 Production

This section starts with 2.1.1 that tries to make the definitions of production and manufacturing clearer. Next, Section 2.1.2 summarizes definitions of production system from literature. The different organization models of a production system that are relevant to this thesis are discussed in Section 2.1.3. At last, the fundamental concept of production planning and control (PPC), which is the application focus of this work, is described in Section 2.1.4. It includes a deeper explanation regarding the production scheduling as one of the tasks in PPC, and also the different approaches for solving production scheduling problem.

2.1.1 Production vs. Manufacturing

The terms manufacturing and production are often used interchangeably since both refer to the method and process to create products. According to the Oxford Dictionary the term manufacturing is defined as [Oxfo-12]:

"making (something) on a large scale using machinery" and *"the action of making or manufacturing from components or raw materials, or the process of being so manufactured."*

whereas production is defined in Business Dictionary in more general as [Busi-12]:

"the combination and transformation of input factors to products."

From these fundamental definitions, it can be implied that the term production is more general than manufacturing. Scientists have been developing the various definitions of production and all of them can be reduced to a *"transformation process from input to output."* The input and the output are not specified. It could be tangible goods or also intangible goods such as information and services [Schm-96] [Dang-01]. Manufacturing is referred in the scientific world as *"workpieces formed out of given materials following given geometrical parameters, adapted to its characteristics and composing functional products"* [Lanz-11]. Consequently, every manufacturing can also be called as production but not every production can be described as manufacturing. In the context of this work, manufacturing and production will be used synonymical and interchangeably.

2.1.2 Production System

A production system can be described by using separate definitions of a production and a system. The definition of production is already presented in Section 2.1.1. A system refers to a set of elements which are characterized by the abilities to exchange information, material, and energy with surrounding and other connected systems. Each element in a system can act as a system itself, whereas the total system could also be an element in a superordinate system [Ropo-99] [Trös-05] [Roga-09].

A production system can also be regarded as a bundle of resources with the intent to maximize profit. Production system resources can be classified into three main groups: personnel, material, and production equipment. Resources classified as production equipment are used to accomplish operational duties. In particular, this includes machines, tools, transport devices and office equipment [Zäpf-07] [Roga-09]. Every paid employee is referred as a personnel resource. They fulfill the demand for human work in a production system. Similarly, the equipment personnel can be classified as the personnel involved in the production process, for instance, machine operator and retooling personnel, and the personnel, who affect the production process indirectly such as accounting, IT, and warehouse staff [Aggt-90]. Materials are understood as raw materials or goods that can be transformed into final products within a production process if they are used in combination with production equipment and personnel resources. There are essentially three different types of material: direct material, auxiliary material, and supply. The characteristic of the direct material is the essential part of the product. Auxiliary material is the necessary good for the operation of production processes and is used as part of the product as well, for example, screw, glue, and paint. Unlike direct and auxiliary material, supply is not taken as part of the product. However, it is essential for the operational readiness of the production equipment. Examples of supply are cleaning materials, coolant, and energy sources such as electricity and gas [Kern-80] [Dang-03] [Roga-09].

2.1.3 Production System Organization

A factory as a production system is typically organized hierarchically into different layers or subsystems. Figure 2.1 depicts the organization layers in production system [Neuh-01] [MüEt-

09]. The lowest layer in the production system organization is the workstation. In a workstation, the process to merge resources (production equipment, personnel, and material) into the physical product is taken place. Each workstation contains a set of processes and the necessary execution time to manufacture a product which is dependent on the processing order. It is also the lowest consideration level in factory planning and operation [Neuh-01] [MüEt-09]. The production line layer consists of several workstations or production facilities. It is responsible for the linking and connection of working stations by using transporting systems and tools. Important tasks of the line layer are, for instance, minimizing idle time of working places (line adjustment) or managing the line-intern logistics [Neum-96] [Neuh-01].

In literature, it can be found different distinctions of the production system upper layers. The top layer is the integration of the factory to a (global) value network. This reflects the global division of production tasks. The following level embodies the factory as a whole, including production facilities, all buildings, and equipment. The layer combines all production-technical and production-organizational elements into a locally determined environment where commercial products are manufactured. A factory consists of different manufacturing areas or segments. The manufacturing segment could correspond to product type, process, or production volume and consists of several production lines [Roga-08] [MüEt-09].

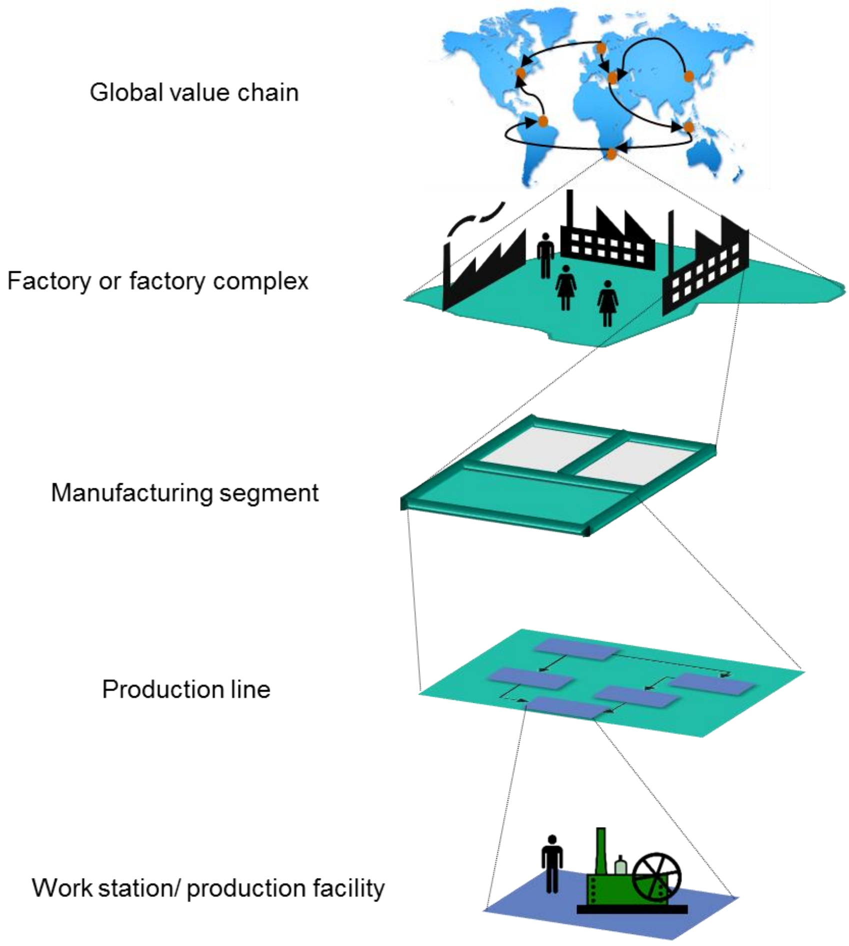


Figure 2.1 Production system organization based on [MüEt-09]

2.1.4 Production Planning and Control

Due to the intensive market requirements and internal complexity, planning and control become more and more important in business processes. The term production planning and control (PPC) refers to a part of operative production management and describes how production processes can be configured in terms of quantity, space and time, in order to produce output. The combination of different fields of tasks or subtasks results in the whole function of PPC. The differentiation of these areas of tasks depends on the PPC concept and differs in literature. Besides that, delimitation of planning and control is often not easy to determine. Following the ‘task view’ of the known Aachener PPC model, Figure 2.2 illustrates the different field of PPC tasks in a very general way [ScEt-06].

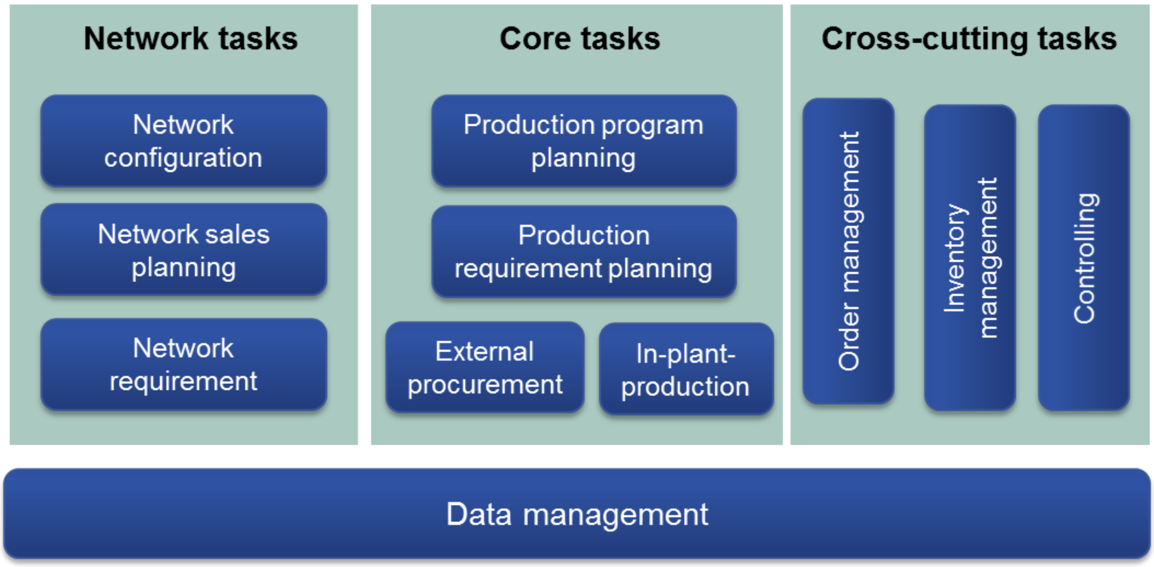


Figure 2.2 Aachener PPC model [ScEt-06]

Scheduling is one of the activities in the in-plant production. If the orders are received from customers, the production planner organizes the operations and resources through a PPC system. The scheduler determines the order of the work list for each resource. Setting the order is known in the business terminology as sequencing [Teic-98]. After determining the order, a precise temporal distribution of operations to specific machines is carried out. This task is referred as scheduling. The complexity of scheduling problems mainly depends on the form or arrangement of production [DoSV-93] [Kram-94].

Because of the different production environments, there are numerous variations of the scheduling problems. However, the solutions methodology developed in research and development are only based on the classical scheduling problems. Usually, the real manufacturing complex structures are not fully considered. Thus, the development of more flexible solution methodology that can be modified and used for several different cases is necessary for production management in practice. Production scheduling problems can be classified according to several criteria as depicted in Figure 2.3 [ZaTI-08]. The problems are classified based on the arrival time of the jobs, the inventory policy, whether the job processing times and machine availability is a priori known, in how many stages the jobs are broken down into the operations, and the job flow pattern towards resources.

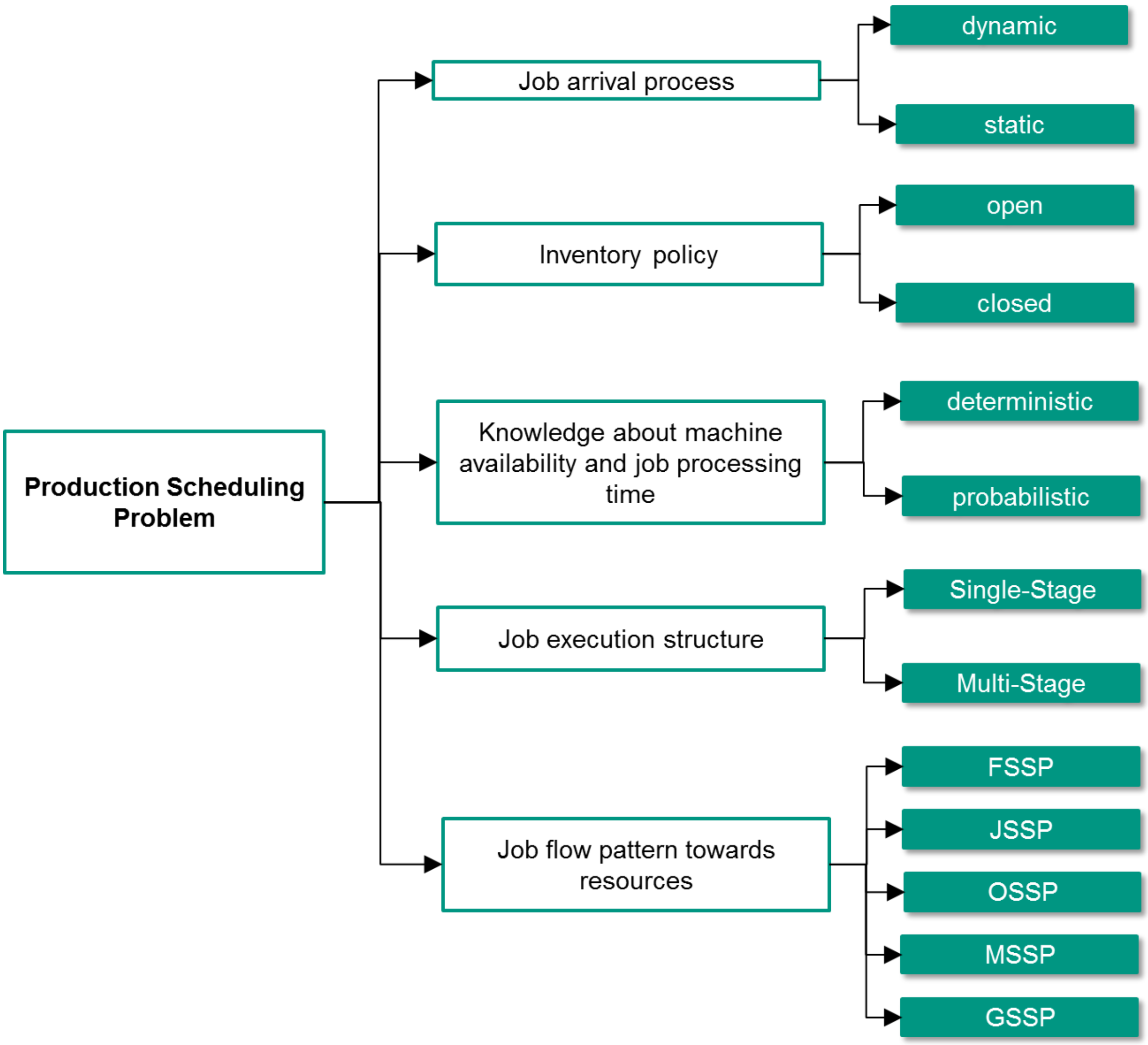


Figure 2.3 Classification of scheduling problem according to [ZaTI-08]

Production scheduling problems are examples of combinatorial optimization problems where the solutions are represented as discrete variables. A combinatorial optimization problem can be solved using exact or approximation methods. With the exact method, however, an optimal solution can be found within a limited time. The scheduling problems are mostly classified as NP-hard or NP-complete problems. The exact algorithms can be used for scheduling problems with a small number of instances but not for more complex and larger problems. Nevertheless, adding more machines, a job shop environment, preemption, precedence relations, release times, and multiple objectives will only complicate the problem. Thus, it is not possible to solve it in a reasonable time. Therefore, solving the scheduling problem using exact method should be avoided. A common technique categorized in the exact method is, for example, branch and bound [BlEt-07].

The good and valid solution can still be found using an approximation method. Figure 2.4 shows the classification of methods to solve combinatorial optimization problems. Whenever optimization fails or is too time-consuming, it is common practice to lose the optimality and decide for an approximate solution using a heuristic method. The calculated will be much faster by a heuristic. The heuristic method can be divided into two groups, i.e. tailored algorithms and meta-heuristics. Examples of tailored algorithms commonly used in production process scheduling are priority dispatching rules and shifting bottleneck procedures [Gere-66] [Peme-00].

In the field of metaheuristics, the problem model serves more as an interface between solution method and the problem itself. That means the problem is not so firmly incorporated into the solution method anymore. This contributes to more generic solution principles, yet it still needs to be assured that the model does not unintentionally relinquish generality. A local search method that seeks a solution in a provided neighborhood structure is considered meta-heuristic, for example, hill climbing, tabu search and simulated annealing [KiGV-83][Glov-89].

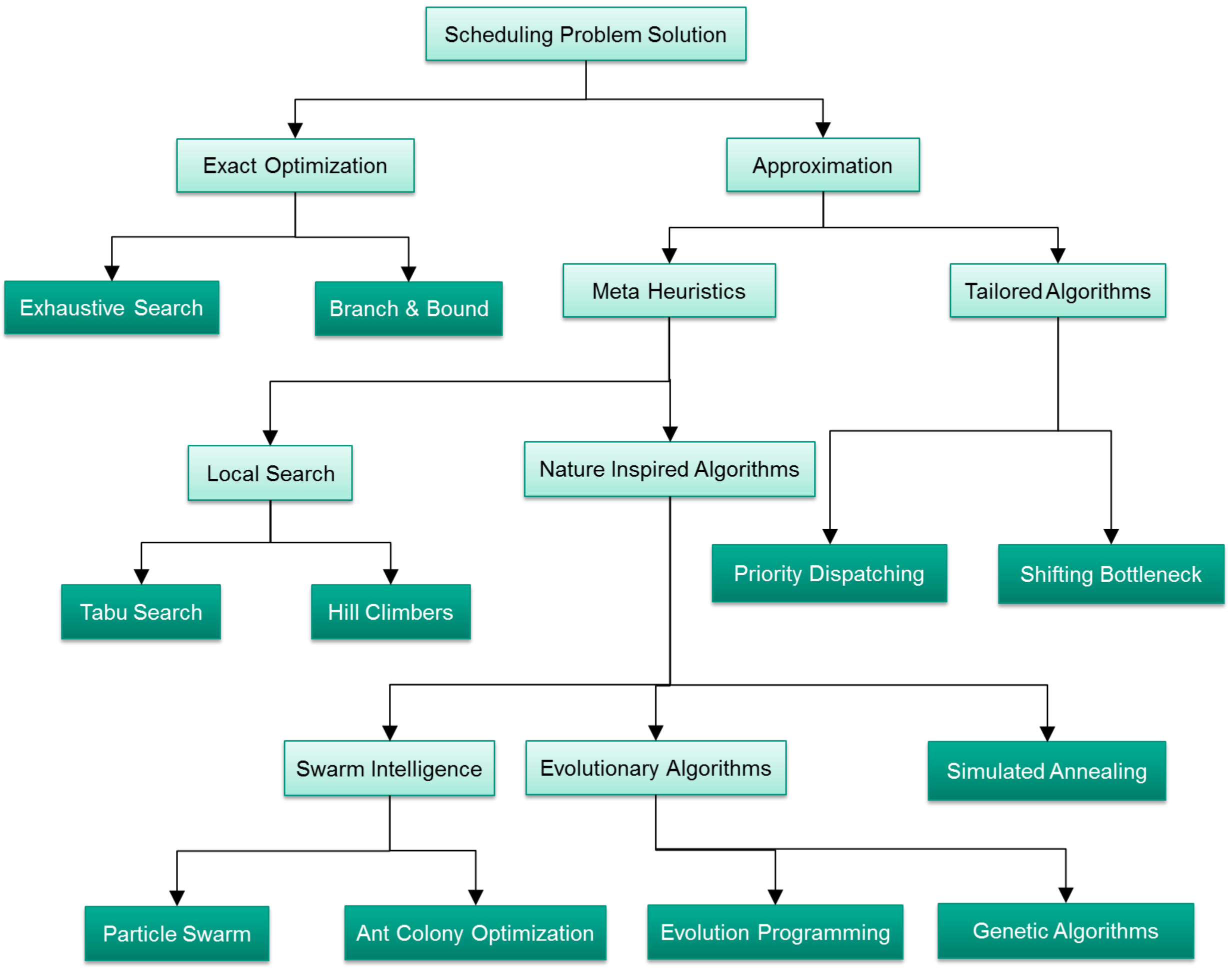


Figure 2.4 Classification of methods to solve combinatorial optimization problem

Metaheuristic focuses on defining the search space that guarantees feasibility of contained solution. However, the model used in metaheuristic is commonly too unrestrictive. Therefore, it is difficult to represent a very specific model in a generic model of metaheuristic. Hyper-heuristic is the method to transform the search space and is developed as an alternative due to incapability in finding a general model. The method chooses the suitable combination of heuristics on heuristics space that leads to the optimal solution in the solution space. The initial hyper-heuristic concept is to combine the priority dispatch and local search heuristic. The concept transforms the search spaces of a real world scheduling problem from solution space to heuristic space [StWV-92] [ChGT-96]. Burke et al. created a classification of hyper-heuristic based on two dimensions, i.e. the nature of the heuristic-based search space and the implemented learning strategy. It can be seen in Figure 2.5. Heuristic selection selects and combines a set of prefabricated heuristics to find a solution, whereas heuristic generation builds the heuristics from scratch. Moreover, on the low level not only construction heuristics can be used but also heuristics which iteratively perturb an existing solution. In that case, the choice of a perturbation strategy would be controlled by the hyper-heuristic. Offline learning requires training data from previous instances. On the contrary, online learning is done on the fly while the search is being performed [BuEt-03].

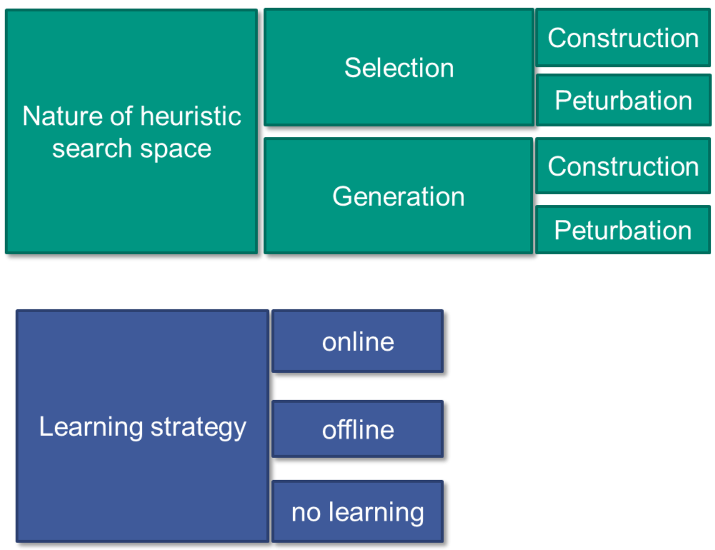


Figure 2.5 Classification of hyper-heuristic [BuEt-03]

2.2 Energy Efficiency and Energy Management in Manufacturing

This section begins with fundamental definitions of energy and energy efficiency. It includes the standard definitions from VDI. The section continues with the explanations of the terms energy management tasks and their application in manufacturing. Then, the existing energy management standards which are suitable to be implemented in manufacturing companies, such as ISO 50001, and EN 16001, are presented. The next part summarizes current techniques to improve energy efficiency from existing literature. Afterward, the role of ICT as one of the current techniques for energy management is elaborated. This section is finalized by a comparison of research and industrial works which deal with ICT supported energy management. In this part, the different works are evaluated and compared.

2.2.1 Basic Definitions of Energy and Energy Efficiency

Energy management is a term that has in recent years become more and more important. Reasons for this include scarce resources, rising prices, and increased sensitivity to environmental issues. The concept of energy management, both in terms of energy supply as well as in the energy and industrial plant, often used without further definition [Shie-06]. In academia and practice to date, no uniform views prevailed. This is due to the fact that the meaning of the concept of energy and management in different contexts vary widely [Kals-10].

In physics, energy is defined as power or the ability of a system to do work. In psychology, it is referred as the vitality or drive of a human. Moreover, in business, this term is used in relation to personal management and motivational theories. Scientific communities use the term energy in the context of economic production and logistics. *Verein Deutscher Ingineure (VDI)* defines energy as physical entity represented in the unit Joule (J) or Watt-hour (Wh). Energy is usually presented as a difference form from a standardized reference state of the system. According to the first law of thermodynamics, energy can neither be consumed nor generated but only transformed. Based on the second law of thermodynamics, the transformation requires restrictions as functions of the respective thermodynamic environments [VDI-03].

In the background situation as described in chapter 1, it is overall quite urgent to reduce the energy demands. This should however not put the prosperity and development opportunities at risk. Therefore, a significant increase in energy efficiency is demanded. In general, energy efficiency (η) is defined as ratios of supplied and used energy flows in a process or a plant in a stationary or quasi-stationary state [VDI-03]. The energy efficiency can be distinguished based on the target energy flow rates – such as electrical efficiency, thermal efficiency and so on. In manufacturing, energy efficiency is the desired benefit (products or services) with minimal use of energy to produce from a specific energy input to best possible use [Müet-09]. Improving energy efficiency means to consume less energy for producing the same quantity of products or services. Equation (2.1) calculates the energy efficiency in general.

$$\eta = \frac{used\ energy}{supplied\ energy} \tag{2.1}$$

2.2.2 Energy Management in Manufacturing

Energy Management is basically a management task. It is performed to reduce energy consumption and to estimate energy costs for long-term business objectives. It includes all processes and measures which are developed and implemented to achieve minimal energy consumption [DENA-11]. It also comprises the activities within an organization which coordinates the tracking of energy consumption [EO-11]. Energy management systems (EnMS) are used to fulfill these tasks.

Energy Management System (EnMS) encompasses organizational and informational infrastructures including sets of hardware and software tools which provide the systematic monitoring and data collection of consumed energy [VDI-07]. These infrastructures establish the basis for investment decisions concerning better energy efficiency [BMU-10]. An EnMS can help a company to improve its energy profile through systematic and continuous improvement cycles. The collected data give the starting point for saving energy. Through continuous monitoring, an EnMS can help identifying nontransparent and intensive energy consumptions.

The application objective of energy management systems is to identify energy saving potentials and to improve energy efficiency in companies and organizations. It results in cost savings. Thus, price competitiveness is improved. It saves resources of material and immaterial nature in order to fulfill economic, political, and ecological goals. A reduction of costs is becoming more and more important since energy prices are rising with growing resource shortage. Low energy consumption can also increase profitability despite the rising of energy prices that can lead to significant financial losses in energy-intensive sectors. The implementation of EnMS can also reduce the dependency on fossil fuels which become increasingly scarce. Companies aim to achieve an optimum energy purchase and maintain an efficient utilization throughout the organization. Energy costs are to be minimized and environmental effects reduced without affecting production and/or quality. Next to economic and ecological aspects, an EnMS can also assist in legal tasks such as complying with emission laws. Public buildings, organizations, and companies have been targeted by policy makers to apply EnMSs. For example, applying the DIN EN 16001/ ISO 50001 is a necessary precondition starting 2013 for energy and electricity tax reduction in Germany [Bunk-11]. Therefore, the application of energy management in manufacturing is becoming increasingly popular.

A systematic EnMS is a tool for depicting energy saving potentials and is used by companies worldwide [BMU-12]. Companies involved in the manufacturing industry are expected to have energy saving potentials up to 15% using EnMS tool. There are also technologies that play important roles in regulating instruments. These technologies are for example cogeneration systems, hydrogen technology, fuel cells, etc. Companies, who also consider lifecycle costs in calculating their investment costs, commonly implement the technologies [ShWB-09].

2.2.3 Energy Management Standards

The standards ISO 50001 and EN 16001 describe the requirements for energy management systems. It enables companies to improve their energy performance continuously through systematic approaches by taking into account legal requirements for the organizations.

A first national standard for energy management system in Europe was founded in 2000 in Denmark by the Association of Danish Industries, the Danish SME Association, the Danish Energy Agency, and several academic institutions. This was then followed by several other European countries, including Sweden in 2003 and Spain 2007. Industries have a vital role in developing such of standards that may exist on the market because standards can be implemented if they are accepted by those concerned. Due to the success of these standards at the national level, the establishment of an international standard was initiated in 2006 by European Committee for Standardization (CEN). The European Commission had also encouraged the introduction of energy management standards to improve energy efficiency in European industry. This results to the DIN EN 16001 in 2009. Finally, the ISO 50001 was published in June 2011 as international standards, as a mutual consent of Europe, USA, and Brazil [KaEt-10] [KaEt-12]. Figure 2.6 illustrates the history of energy management standards resulting the introduction of ISO 2011.

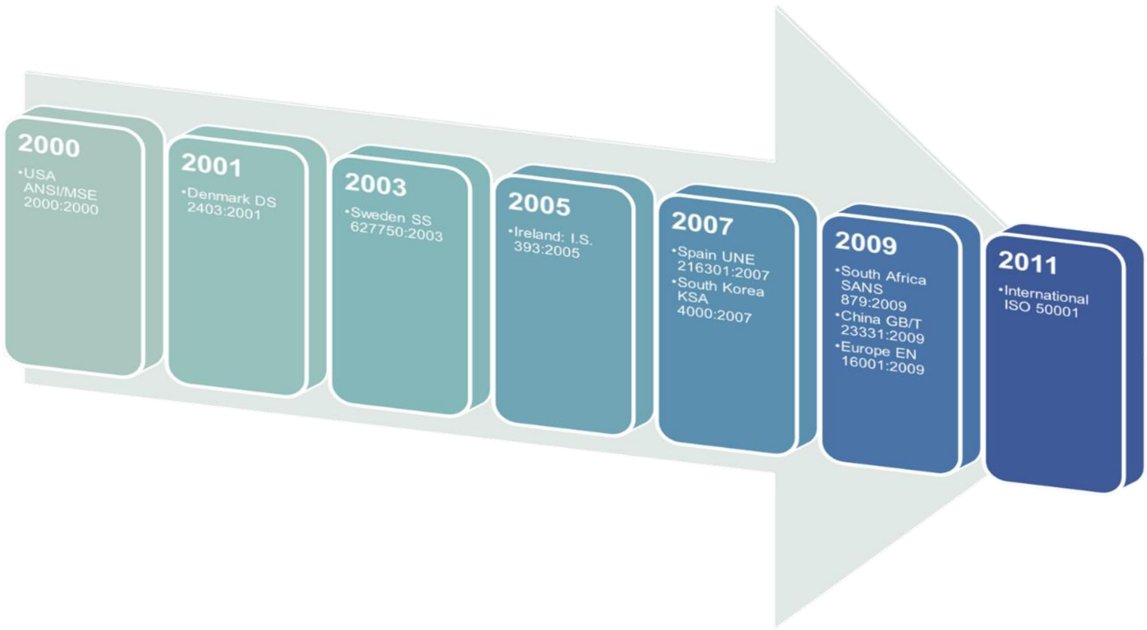


Figure 2.6 History of energy management standards [KaEt-12]

This standard defines the requirements for an energy management system so that companies can independently assess their energy consumption systematically. It also allows them to identify potential energy savings. The use of such potentials not only lowers operating costs but also increases the competitiveness of the company. Similar to classic management system standards, the standard is designed to be applicable to organizations of all types and sizes [Kots-08]. It depends neither on the energy usage intensity of the company, the industry branch, nor the company size. However, the most interesting is the introduction of an energy management system in energy-intensive companies where the significant saving can be achieved. The standard is based on the PDCA process (Plan-Do-Check-Act). It is an iterative four-step problem-solving process by which continuous improvement is possible. In each step, it involves both technical and managerial tasks [Stor-10, p.7]. Figure 2.7 depicts the energy management system model based on PDCA cycle according to [ISO-11].

The first step in PDCA process is planning. Here energy saving objective and strategies are defined. It includes the specification of energy efficiency measures and the determination of responsible stakeholders. An action plan and energy policies are established, and the necessary funds are provided. The optimizing measures are then carried out in the “Do” phase of the cycle. Besides obeying energy saving rules, these measures can be in the form of new efficient technologies or a process change. The management framework has to be customized to keep the process continuing [KaEt-12].

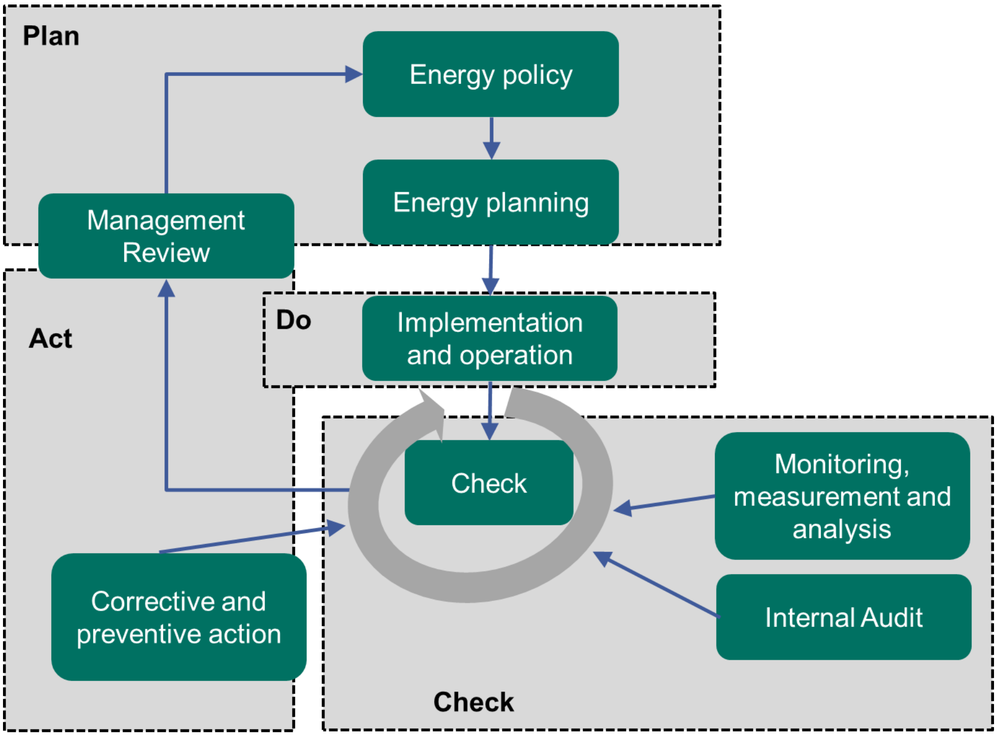


Figure 2.7 Energy management system model [ISO-11]

Next in the Check phase, the degree of achievements and effectiveness of energy management system implementation are assessed. This is performed through energy consumption monitoring, measurements, and analysis. Energy audits can help to brainstorm new ideas and can be supported by external specialists. The outcome of an audit is evaluated, and new insights are gained. These insights, current energy data, benchmark values, and the derivation of new targets are used for strategic optimization.

The last step “Act” can take place parallel to production activities. In this phase, the actual energy data, audit results, and energy efficiency related new findings are utilized to create optimization strategies. The new energy saving objectives are derived from the resulting strategies. Processes are executed continuously so that lower energy-related costs can be reduced significantly [BMU-12].

2.2.4 Current Techniques for Energy Management

The European Commission has published best available techniques for energy management in companies, especially manufacturing in a reference document. The techniques are classified into two categories [Euro-09, p.269]:

1. Energy efficiency at the installation level and techniques. It is independent of the system type, for instance, cogeneration, combustion, etc.

2. Energy efficiency in energy-consuming systems, processes, activities, or equipment. It depends on system specific techniques like reducing the head loss in combustion system through better insulation.

As mentioned in the Section 2.2.3 the energy management system is based on the PDCA cycle approach. According to the document published by European Commission, the energy efficiency management system must constantly monitor and measure energy efficiency indicators. The energy efficiency indicators are compared over time to determine the effectiveness of implemented energy management measures. To evaluate the energy efficiency, it is also important to consider the production parameters that influences the indicators, such as volume or number of produced units. To ensure the accuracy of statements represented by defined indicators, it is important to specify the boundaries of the evaluated system or system part clearly. The indicators should also be used to benchmark the energy usage of the current company to other companies from the same sector, region, or country. Information obtained from continuous measurement and benchmarking leads to the development of new measures to improve energy efficiency [Euro-09, p.276-281].

There are various techniques to improve energy efficiency at different steps in the installation phase. An energy efficient design should be performed before planning a new equipment installation or reconfiguring the existing one. Another technique is to include sensors in the installation to measure energy consumption and other parameters for calculating energy efficiency indicators. The process integration within the installation should also be optimized. It is also essential to perform maintenance in the installation phase. The organizations or personnel, who are responsible for planning and executing the maintenance tasks, has to be clearly defined. Through those tasks in the maintenance phase, defects affecting energy performance of the equipment, such as leaks or errors, can be detected in time. The execution of maintenance tasks has to be effectively monitored and controlled. It is done by having knowledge of parameters influencing energy efficiency of an installation, implementing measures to improve the energy efficiency and documenting them [Euro-09, pp. 282-296].

It is necessary to maintain expertise and impetus of energy efficiency programs within the organization to improve energy efficiency continuously over time. Maintaining impetus can mean implementing an energy management system standards, such as ISO 50001, gaining the certification of implemented standards, continuous benchmarking, creating financial profit centers for energy efficiency, or improving existing management system by using operational excellence. The company can recruit skilled or trained staffs, share in-house knowledge resources, or offer training by external specialists, to maintain the expertise in energy efficiency [Euro-09, p. 279-280].

2.2.5 The Role of ICT in Energy Management

According to the Directorates General (DG) for Information Society and Media of the European Commission, ICT will play a significant role to achieve the European Union environmental objectives and to help in accelerating economic growth as well as to improve energy efficiency [DG-8]. A comprehensive approach has been developed to facilitate the cooperation and communication between policy makers, business, and other stakeholders such as citizen. If it is implemented correctly, ICT can achieve various energy efficiency improvements [DG-08] [ECIS-09]:

1. ICT enables saving through intelligent observation and control. In machine level, it allows reduction of energy consumption without affecting the function of the machine tool.

2. Today's ICT architectures allow a wide range optimization through data exchange between automation, manufacturing execution system and enterprise resource planning system.

3. ICT allows advanced monitoring systems that analyze different data beside energy consumption, such as manufacturing processes, plant conditions, and life cycle information. ICT helps to structure and visualize the collected data, analyze them and report back for improvement. Furthermore, it gives first hints on where and when they are, and what has caused them. Monitoring is also needed to generate consumption profiles which in combination with knowledge management can aid in forecasting energy demands.

4. Advanced scheduling algorithms allow the increase of production effectiveness and through optimization of process arrangement. At production system level, it allows for a scheduling of energy intensive tasks when the lowest economic and ecological effects are to be expected. It could avoid peak loads resulting in the elimination of undesirable, high impact marginal peak load energy production, reduce the environmental footprint of a production facility as well as its operational costs [HuGa-08].

5. Modern ICT consume less energy consumption because of its advanced architecture. For example, multi-core processors make computing more energy efficient. It does not need several separate computers to solve a task. The product lifecycle footprint of this kind of equipment from the production to the disposal can be designed more efficiently through waste and resource reduction.

6. ICT solutions often help to reduce energy and resource consumption during manufacturing process sequences.

7. ICT digitalizes activities and processes (simulations, teleworking, e-government, smart grids, etc.). Therefore, it requires less material and energy.

2.2.6 Related Research and Industrial Works in ICT Supported Energy Management

By considering the essential roles of ICT in energy management as described in Section 2.2.5, different current industrial and research works are analyzed. Criteria for the consideration to carry out the analysis are developed. The criteria are deduced from research questions formulated in Section 1.2. The developed criteria are listed as follows:

- Interoperability and knowledge sharing capability. The tasks in energy management systems involve different organizational and informational infrastructures. Therefore, interoperability among different IT systems, between IT system and human stakeholders, as well as knowledge sharing among stakeholders is an important point for a successful energy management system implementation.

- Knowledge-based intelligent analysis. Information and knowledge coming from different sources have to be structured and formalized in a knowledge base. The knowledge base is then used as the reference to perform evaluation and analysis on current energy performance as part of check phase in energy management system (see Figure 2.7). The evaluation can be supported by such a decision support or expert system that works on the knowledge base.

- Adaptive and learning capabilities. The generation of the knowledge base is an extensive task. In order to accelerate knowledge generation process, a learning capability is a key feature to build a knowledge base system.

- Intelligent performance monitoring and benchmarking. The performance monitoring and benchmarking is an essential task to evaluate the energy performance using quantitative information. The benchmark should compare different organizations and resources in production.

- Energy optimized production planning and control. As mentioned in Section 2.2.5 point 4, ICT can help to improve energy efficiency and manufacturing productivity through optimized production planning and control activities, especially through energy optimized scheduling.

The first analyzed related work is the research project **EnHiPro** (energy and process materials optimized production). The project goal is to analyze and to evaluate measures to increase energy and auxiliary material efficiency in production systems, especially SMEs. Within the project, methods and tools that are applicable across different SMEs from different sectors are developed. The methods and tools enable continuous organizational and technical measures to increase efficiency and to assess their impact [enhi-12]. The project **e-SimPro** focuses on the reduction of production costs of machine tools. The developed software will help the assessment of the overall efficiency of machine tools by energy demand simulation. This research project is a component-based quantification of the energy consumption. In the first step, a study will be performed in order to identify the information relevant to the simulation components, then models and methods based on analysis results are developed. The next step is software module development and application in a pilot. Finally, the overall economic performance of the solution will be assessed [esim-12].

AmI MoSES is a project funded by the European Union that aims to improve energy efficiency through ambient intelligent approach and application of knowledge management technologies to implement a decision support system. Ambient intelligence is a combination of context awareness, ubiquitous computing, and communication. Through the combination energy consumption monitoring, knowledge management and ambient intelligence, feedback can be integrated into the process. Comprehensive information is provided and support is given for optimization [ami-08] [HeEt-10] [HeEt-11].

The goal of the project **EWOTeK** is to improve energy efficiency through a time-shifting of order processing and standby strategies. To achieve the goal, it requires an energy data acquisition system. Through a continuous data collection on individual machine component, it should be possible to evaluate the energy demand of each machine component based on currently executed production operations. It enables identification of energy waste both in machine and factories. Thus, effective solutions can be developed [RIG-11]. The project **Lifesaver** funded by EU in the frame program FP7 aims to provide a comprehensive system containing a collection of modules and service to support decision making for improving the energy performance of industrial companies. The system should support different time horizons of decision making. Lifesaver employs a knowledge base as the core component of the decision support system. The system provides intelligent monitoring to optimize energy saving in daily used of production equipment [life-12].

Another related project funded by EU in the program FP7 is the project **ESToMaD** (Energy Software Tools for Sustainable Machine Design). The project aims to develop methodology and tools to simulate and analyze energy flows for the design of new machines and analyzing energy performance of existing machines. The project extends existing software called AMESim to have the capabilities in computing the energy flow and loss between different machine components [esto-12]. The ESToMaD methodologies are integrated into existing ICT tools for design, modeling, and simulation. The project concentrates on machine construction industries. ESToMaD allows energy performance benchmark at the machine level and interaction between machine manufacturers and supplier of components and subsystems in designing machines [SyEt-13].

An ICT-based integral and intelligent energy management system is being developed in the project **Euroenergest** funded by EU in FP7 program. The project focuses on the energy management system in automotive industries to improve their competitiveness. In the project, an artificial intelligence methodology is being developed to perform energy prediction and optimization on the energy hub. A framework for load forecasting in the energy consuming side is also being developed using an adaptive neural network and multi-objective genetic algorithm. The framework is developed aiming at supporting the decision making in operational activities, as well as improvement of key performance indicators [euro-12] [CaEt-12].

The EU project **e-SAVE** started in 2012 aims to develop the information infrastructure, applications, and tools to support decision making in design and operations phases of supply chain management by taking into account the environmental KPIs and the dynamic energy profile of products and processes. The e-Save solution should support efficient information sharing among supply chain stakeholders throughout product lifecycle [esav-13]. The project **ISES** (Intelligent Services for Energy Efficient Design and Life Cycle Simulation) funded in the frame program FP7 develops an ICT solution to integrate and complement existing tools for energy efficiency evaluation, simulation, and optimization of products and facilities, such as buildings. The solution provides interoperability between energy analysis, product and facility design tools. An ontology-based method is developed to allow an intelligent and flexible interoperability [ises-13].

The FP7 EU project **DEMI** aims to develop a decision support system that assists engineers to design energy optimized discrete manufacturing processes, to monitor those processes in a more intelligent way, and to ensure the energy efficiency of the whole process life cycle. The project introduces knowledge-based approaches i.e. Case-Based Reasoning (CBR) and Rule-Based Reasoning (RBR), to provide the intelligence of the decision support system. Nevertheless, the knowledge acquisition is performed manually by domain specialists and developers using a wiki. The energy use is modeled and simulated, but KPIs to measure the energy performance are not developed. The project focuses on the process monitoring, not the production planning and control [Demi-10] [TeEt-11a] [TeEt-11b].

A project funded by German Ministry of Education and Research (BMBF) called **reBOP** develops integrated concepts to evaluate and optimize manufacturing process chains from resource efficiency point of views, including energy and material efficiencies. The concepts identify the correlations of different process states (e.g. idle, standby, retooling, start-up), disturbances and production values. The concepts are implemented as a software tool that enables the quantification of the energy and resource efficiency along process chains. This improves the transparency of weakness and optimization potentials in a manufacturing system [reBO-12].

Analysis of selected projects above results a number of ICT supported approaches and tools to improve energy efficiency in production. The comparative analysis result shows that there is no project that fulfills all criteria. Many projects address only several criteria or cover a specific criterion only partly. In contrast to the above-mentioned projects, this thesis aims to meet all of the criteria derived from the research objectives described in Section 1.2. The comparison results are summarized in Table 2.1.

Table 2.1 Comparison of projects in ICT-supported energy management

Criteria / Project	Interoperability and knowledge sharing	Knowledge-based intelligent analysis	Adaptive and learning capability	Performance benchmarking and monitoring of different production resources and organization	Energy optimized production planning and control
EnHiPro	+	-	-	+	-
AmI Moses	+	+	-	·	-
EWOTeK	·	-	-	·	+
e-SimPro	-	-	-	+	·
ESToMad	·	-	-	+	-
Lifesaver	+	+	-	-	·
Euroenergest	-	-	+	-	+
e-Save	+	-	-	+	·
ISES	+	+	+	-	-
DEMI	·	+	-	·	·
reBOP	·	-	-	+	·
This thesis	+	+	+	+	+

2.3 Knowledge Management and Engineering

This section gives an overview of the definitions, concepts, and existing work related to the first objective of the thesis. It begins with the fundamental definitions of data, information, and knowledge. It continues with the comparison and relations between knowledge management, knowledge engineering, and knowledge-based system, which are the basic concepts strongly related to the development of a knowledge base. This section describes then the different approaches for representing knowledge. Ontology as knowledge representation is taken into focus in this thesis due to its capabilities to solve interoperability problems and to allow reasoning. Different ontology languages are analyzed and compared. The related work in manufacturing and energy management that employ ontologies are also examined and presented. Different KDD techniques to extract knowledge from data are explored analyzed, and the results are demonstrated in this section. Finally, it presents the related work utilizing KDD in energy management and manufacturing.

2.3.1 Fundamental Definitions: Data, Information, Knowledge

Almost all of ICT-based systems deal with data. ICT helps human to retrieve, to process, to interpret and to store data in order to support their business or activities. The words data, information, and knowledge are often used to represent the same thing. The following definitions try to solve the disambiguation of the words.

Webster's dictionary defines data as follows:

"Data are factual information (as measurements or statistics) in forms of sets of symbols which are used as the basis for reasoning, discussion or calculation."

Bauer et al. defined information as follows [BaFG-91]:

„Information is an interpretation of data in a context."

Data become information if meanings are given to them in forms of relations, and they are interesting, meaningful, and understandable for a human. Information is gained from data through interpretation of data based on a reference system [Leut-02] [FeSi-98].

Knowledge represents the structure, meaning, application, interpretation and other functional characteristics of data [Lu-91][Rude-98]. In this work the following definition is used:

"Knowledge is a quantity of all interpreted and evaluated information which is integrated as a whole inconsistent and unambiguous manner. The structure of knowledge allows formulating the future events based on previous events that have already occurred and has been saved. In general, knowledge symbolizes the know-how coming from previous experiences."

The nature of knowledge enables to plan and design future events based on the experience gained from past ones [Ovtc-11]. Knowledge is the application of data and information answering the how question. Further knowledge can be inferred from the collected knowledge. [BeCM-04]. In literature, it also can be found different other definitions based on the different interpretation and application disciplines. Nevertheless, the relation between definitions of data, information, and knowledge can be summarized as shown in Figure 2.8.

Knowledge can be classified as tacit and explicit knowledge. Tacit knowledge is the knowledge that can be transferred through training or gaining experience but not codified. Tacit knowledge is normally transmitted by writing it down or verbalizing it. Tacit knowledge can be understood and transferred to the cultural context. Thus, it is difficult to transfer it to the people who do not belong to the culture. In order to transfer of tacit knowledge effectively, extensive personal contact, regular interaction, and trust are needed [Pola-66] [Coll-01] [Gour-02] [GoKo-11].

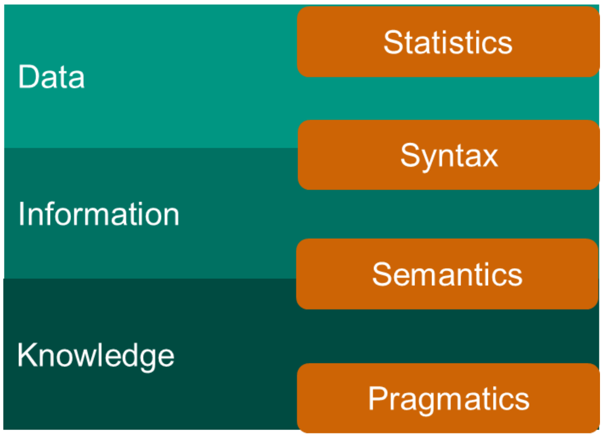


Figure 2.8 Relation between definitions of data, information, and knowledge [GoKo-11]

Explicit knowledge is the knowledge that has been articulated and codified so that it can be easily transferred. Explicit knowledge can be collected through practical experience or logical deduction. It can be formalized and stored in a single location. Tacit knowledge can be transformed into explicit knowledge through codification, articulation and specification techniques [Lam-00].

2.3.2 Knowledge Management, Knowledge Engineering, Knowledge Based System

Knowledge Management (KM) consists of strategies and practices to identify, collect, structure, represent, distribute, leverage, and share intellectual assets to enhance its performance and competitiveness. It consists of two main activities: (1) acquisition and documentation of explicit and tacit knowledge and (2) dissemination within the organization. KM is part of business strategy, provides the technical infrastructure for information management and is an important decision support in companies, especially the larger ones [AdMF-06]. In computer science, Studer and Maedche defined knowledge management as purposed use of knowledge resources and their targeted applications [StMa-02]. The knowledge management from computer science view has the following tasks [PrEt-00]:

1. To capture knowledge by means of knowledge acquisition process
2. To store knowledge by codifying and structuring the knowledge in knowledge bases using knowledge representation language
3. To deploy the knowledge in a knowledge server and to enable the access

According to Feigenbaum and McCorduck, Knowledge Engineering (KE) is defined as [FeMc-83]:

"KE is an engineering discipline that involves integrating knowledge into computer systems in order to solve complex problems normally requiring a high level of human expertise."

The paradigm of knowledge engineering has been shifting in recent years. The first paradigm viewed knowledge engineering as a transfer process that transforms the knowledge resources from human experts to a knowledge-based used by expert system programs. The paradigm changed to the modeling process view, due to the inconsideration of tacit knowledge to solve the problems. Many approaches in KE are nowadays influenced by problem-solving methods (PSMs) which allow inference to solve given tasks, reasoning to validate the knowledge base, and reuse of existing knowledge bases [HaWL-83] [BiKl-93] [StBF-98]. Knowledge engineering includes techniques for knowledge elicitation, modeling, representation, validation, automated reasoning, and justification aspects [Mott-99] [ScEt-00]. Figure 2.9 illustrates the aspects of knowledge engineering process [TuAL-05].

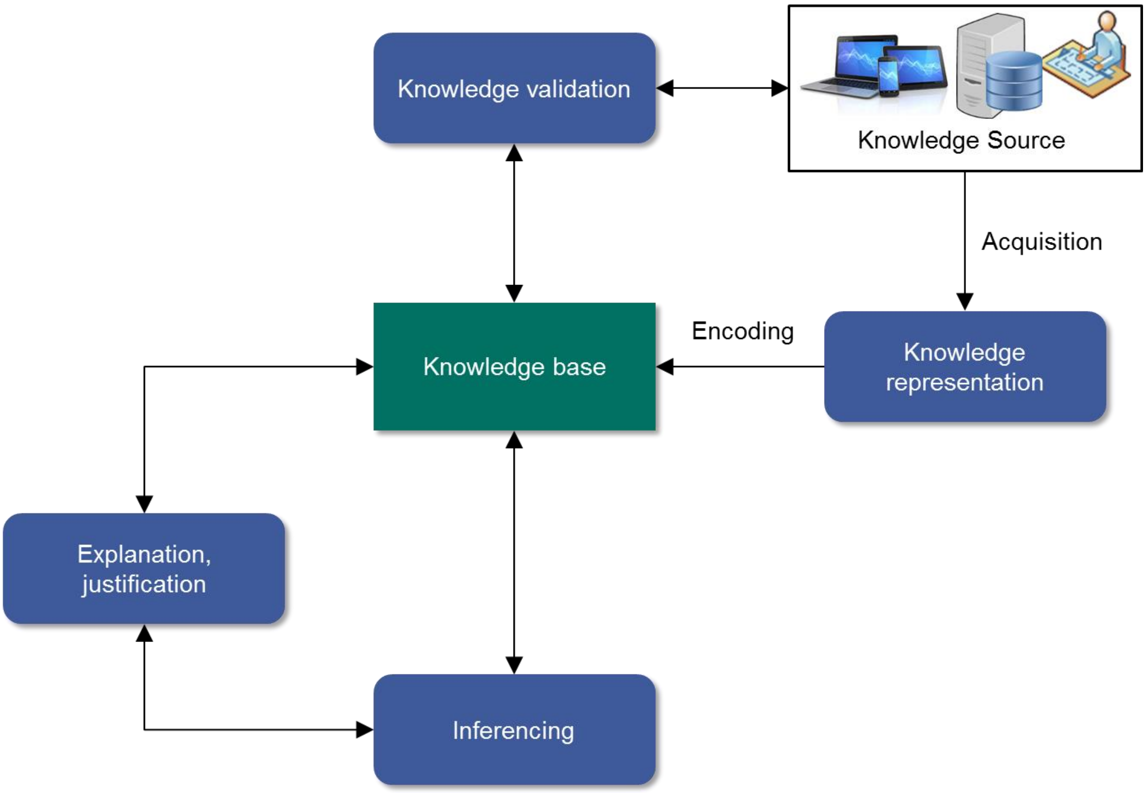


Figure 2.9 Knowledge engineering process [TuAL-05]

Knowledge management tasks concern with the knowledge needs of an organization, whereas knowledge engineering deals with technique and technologies needed to meet the organization’s knowledge management tasks. According to Preece et al., knowledge engineering techniques could give benefits in implementing knowledge management as follows [PrEt-00]:

1. Knowledge acquisition techniques lead to effective capture of valuable knowledge
2. Knowledge modeling techniques provide knowledge map of multiple knowledge sources within an organization

3. Structured and formal representation techniques allow easy access, knowledge verification, and integrity checking

The knowledge engineering can be regarded as tasks to build, maintain and develop knowledge-based systems. It is commonly used in the development of expert systems and decision support systems [Negn-05] [KeCr-07]. The knowledge-based system itself is defined as software systems containing a significant amount of knowledge in explicit and declarative form. The software is developed to support human decision making. It commonly contains an inference machine that performs reasoning and logical deduction based on the contained knowledge and specific problem [Kamp-99] [SpEt-01]. Section 2.3.3 describes different knowledge representations that are commonly used by knowledge-based systems to express the knowledge explicitly.

2.3.3 Knowledge Representation

The different kinds of knowledge types are stored formally through a form of knowledge representation. The knowledge representation allows access, usage and reasoning of the knowledge. Following is the definition of knowledge representation [Pars-12]:

“A knowledge representation is a representation of the world and our knowledge of it that is accessible to programs and usable. “

Knowledge representation has five following roles [DaSS-93][Pars-12]:

1. Surrogate, it represents abstractions that capture some part of the real world and is useful for a human. It is context sensitive and interest relative depending on the applications.

2. Expression of ontological commitment, expressing a way of thinking about the world. It consists of a set of concepts and connections between them.

3. Theory of intelligent reasoning. The representation is a model that allows human-like reasoning and acquiring information without manipulations.

4. The medium of efficient computation, so that computer programs able to work with it.

5. The medium of human expression, allowing interaction with a human through knowledge generation, building, and usage.

From literature, it can be found different approaches for knowledge representation (see Figure 2.10). The following sections describe the most relevant approaches related to this thesis:

Logic

Logic derives new knowledge from known facts using logical inferences. The logic based knowledge processing allows the introduction of variables for objects. The variables are used to formulate general statements [Lunz-94]. The objects or states and their complex relationships representing the real world are described as calculus statements. By using them, other statements can be deduced with the help of rules. In a logic-based representation, the calculus statements mean axioms, facts, or assumptions. The deduction process is a matter of logic [Pocs-00].

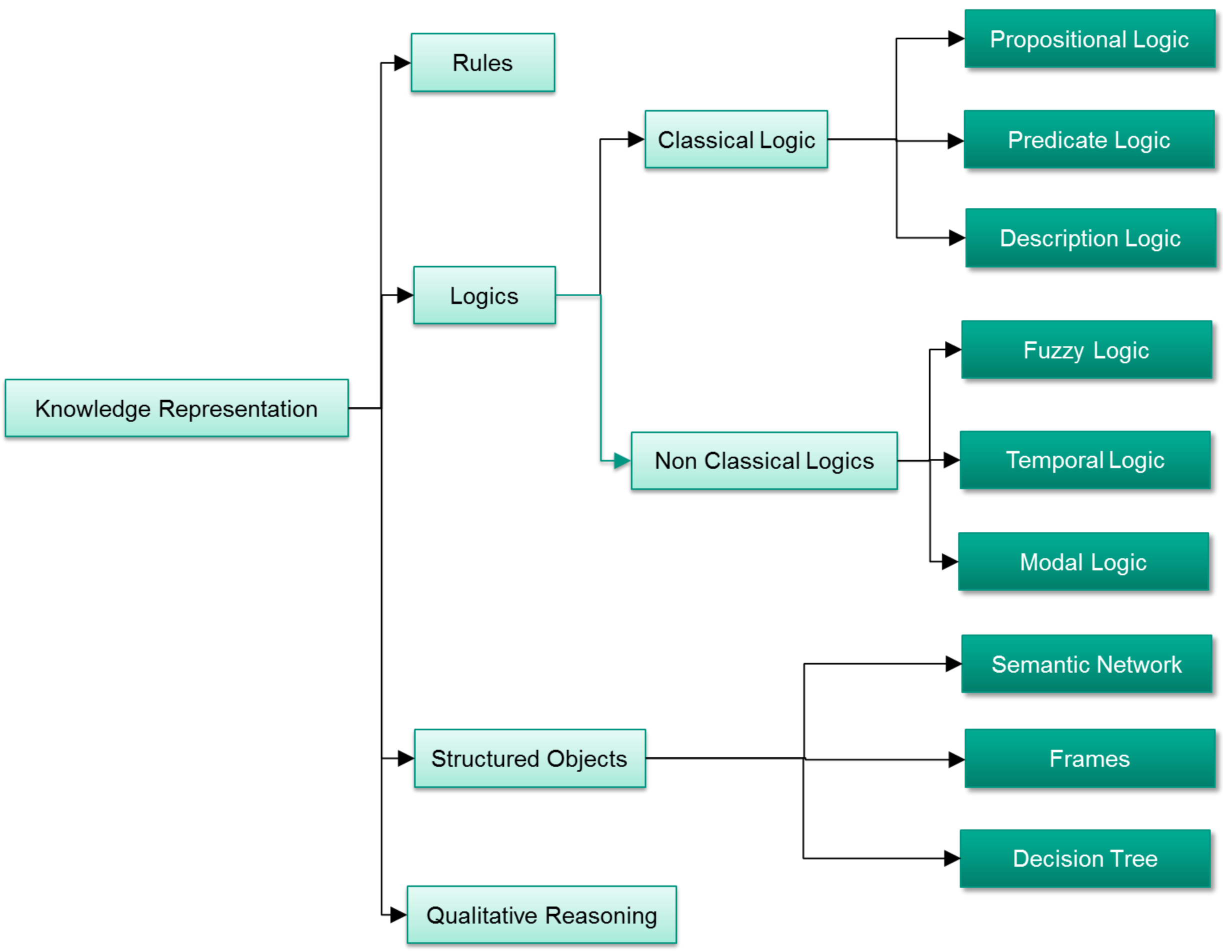


Figure 2.10 Overview of knowledge representation approaches based on [Jenn-01]

Semantic network

A semantic network is a graph that consists of nodes representing the terms (concepts) and the edges representing relations between concepts. In this manner, the "semantic memory" in the form of a network can be modeled [Klim-03]. A semantic network supports inheritance. Instead of storing each object all its properties, they are structured in a hierarchy. Only individual values are stored in an object. General characteristics of the predecessors of the object assigned to the hierarchy and all successors "inherited." In processing the queries, the value of properties of an object is first searched within the object. If the value is not found, it looks at his predecessor in the hierarchy. Semantic networks allow representation of objects having complex relations structures. Unfortunately, the complex structure of a semantic network makes the reasoning difficult because it does not support sub-network processing that could reduce the complexity [Rzeh-98] [Geba-01] [Leut-02].

Description Logics

Description logics (DL) represent the knowledge of an application domain in a structured and formally well-understood way. Description logic consists of a set of constructors that allows constructing complex concepts and roles from atomic one. Concepts refer to classes sets of objects, and roles are the binary relations on objects. DL contains formal and logic-based semantics. These elements do not exist in semantic networks and frames. DL has two main parts in expressing knowledge, namely TBox and ABox. TBox is the terminological part. It contains concept definitions that describe macros/names for concepts and axioms that restrict the knowledge model. In a database system, TBox corresponds to the schema. In object oriented concept, it has a role similar to classes [BaHS-08].

Concept can be defined as [HoSa-12]:

$$A \equiv C \tag{2.2}$$

where *A* is an atomic concept and *C* is a complex concept

Axiom is defined as [HoSa-12]:

$$C_1 \sqsubseteq C_2 \tag{2.3}$$

where C_1 and C_2 are complex concepts.

ABox is the assertion part that is built from concept assertions and role assertions. It is used to describe instances of concepts. It allows a concrete description of resources in a domain. ABox corresponds to objects in object oriented concept and data in database system [BaHS-08]. Concept assertion is formulated as [HoSa-12]:

$$a : C \tag{2.4}$$

where *a* is an individual name and *C* is a complex concept.

Role assertion is expressed as [HoSa-12]:

$$\langle a_1, a_2 \rangle : R \tag{2.5}$$

where a_i is individual names, and *R* is a role.

Modern DL systems provide capabilities to deduce implicit knowledge from explicitly represented knowledge automatically in a finite time. This is called reasoning service. Typically TBox is created first and then a reasoning process is carried out to ensure that all concepts are consistent. Researchers have been developed DL systems that are able to execute complete inference algorithms in large knowledge bases but still have surprisingly well performance. They have highly optimized tableau-based algorithms [BaHS-08] [HoSa-12].

Rules

In the approach, knowledge is represented with rules that formalized in the following form [Klim-03]:

$$if\ A\ then\ B \tag{2.6}$$

where A is the antecedent or premise, and B corresponds to conclusion or action. The conclusion can be retrieved, and action may be executed if the antecedent or premise is true. Due to their simple form, knowledge representation using rules is easy to understand and is also extendable. However, it is difficult to check the consistency of a set containing a large number of rules [Rzeh-98] [Leut-01] [Mahl-09].

2.3.4 Knowledge Representation Using Ontology

This section begins with the definition of ontology and the role of ontology in knowledge management. Then, it explains several techniques how to develop an ontology in a systematic way. The different technologies and languages for representing ontologies are then introduced. This section ends with the evaluation results of several related works employing ontologies in manufacturing and energy management.

Definitions

The term ontology comes from the Greek and is made up of two words, "on" that means the genitive of "be" and logos which refers to "study of" together. It is originated from metaphysics, a branch of philosophy that deals with the question of why something exists. It was used by Aristotle to classify things in the world. For many philosophers ontology (and metaphysics) is the core area or the rudiments of philosophy, other philosophers deny the possibility of ontologies. In information technology, the term has been taken up in the eighties of the last century by researchers in the field of artificial intelligence for modeling knowledge. In the literature, there are many definitions of ontology. The most common and widespread ontology definition in computer science is the one defined by Gruber [Grub-93]:

“An ontology is an explicit specification of conceptualization.”

Ontologies provide a shared and common understanding of a domain that can be communicated between people and application systems [Fens-00]. Studer, Benjamin, and Fensel refined this definition a few years later with the following definition which is also suitable for this work [StBF-98]:

“An ontology is a formal, explicit specification of a shared conceptualization.”

Formal means that an ontology should be readable by machines, which can be presented by means of a set of classes or concepts, relationships, individuals and logical statements. Explicit indicates that the type of classes and rules are determined exactly for their use. Specification is a declarative description. Shared means that the description should be based on the consensual understanding of the group of potential ontology users and not based on the perspective of some individuals. Conceptualization reflects an abstract presentation of some phenomena in the real world by identifying the relevant concepts of that phenomena.

Studer and Maedche defined ontology in a more formal way [StMa-02].

“An ontology is a 4-tuple W:= (K,is_a, R, ψ), where K is a set of concepts, ‘is_a’ is generalization relation on K, R is a set of relations, and ψ: $R \rightarrow \wp(K \times K)$ is a function.”

An ontology represents a network of hierarchies where information is linked with logical relations with each other. Ontologies consist of classes (also called concepts), properties (also called roles or slots), restrictions on properties (also called role restrictions or facets), and instances. Classes are the concepts of the domain in a hierarchical order. Relations are based on properties that its information must be specific assigned. Properties help to characterize the classes and their relations. Restriction can be added on properties. The class hierarchy, defined properties, and restrictions build the conceptualization components called TBox. In a real world scenario, the classes are instantiated resulting individuals or instances. The instance or assertion components are defined as ABox. TBox and ABox components make up a knowledge base [NoMc-01] [DaOS-03].

Example:

Drilling Machine is defined as an ontology concept or class. Drilling Machine is a type of Machine which is also a class. Compactor is also a type of Machine. Therefore, both Drilling Machine and Compactor are subclasses of Machine. This builds a class hierarchy. One might define that every compactor can process a maximum of five products. Process is an example of property with a restriction. Compactor100 is an instance of compactor and represents a real world object. It can process two products, e.g. P1 and P2. This example is illustrated in Figure 2.11.

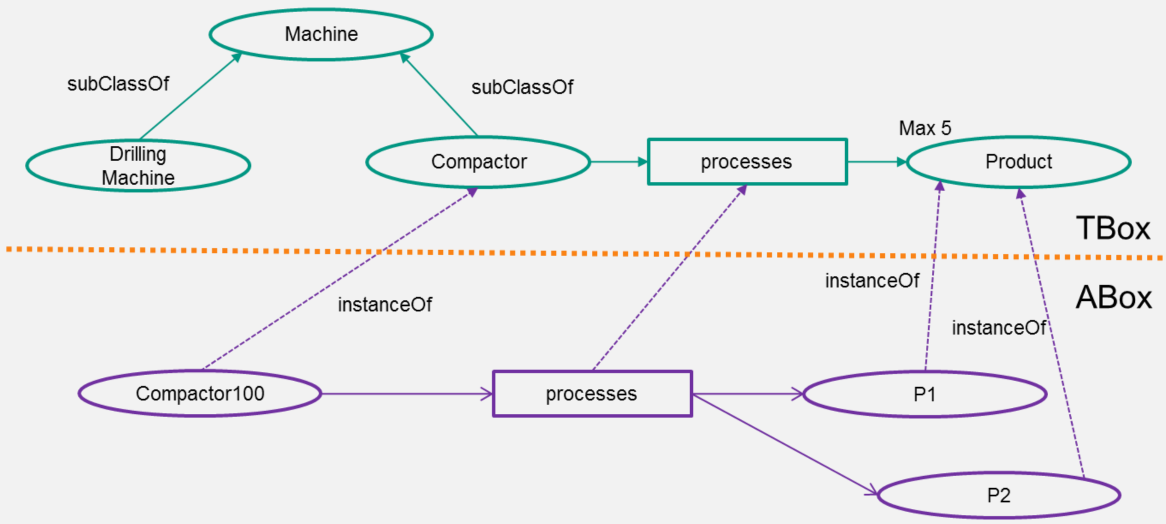


Figure 2.11 Example of an ontology

Ontology and Knowledge Management

One of the tasks in knowledge management is to store and to structure the knowledge using a knowledge representation language. Ontology is an approach for representing the knowledge. The goal of developing an ontology is to share a common understanding of the structure of information among people or software agents. It is done by specifying concepts and relations among those concepts for a specific domain. Thus, different applications using the same underlying ontology are able to share and to publish the same used terms. By using the ontology, the assumptions in the domain are expressed explicitly. If the knowledge about a domain changes, the changes can be propagated easily into the ontology. This is an advantage over hard-coding assumptions in the source code of the software. Hard-coded assumptions are not only hard to find but also difficult to change, in particular for people without programming expertise. It is not necessary to change the programming codes or application functionalities. An explicit representation allows new users easily learn the meaning of the terms in the domain.

The explicit and formal representation with an ontology makes the knowledge in the domain reusable. To create a new ontology, it is not always necessary to develop it from scratch. The existing ontologies from the relevant domains or general ontologies can be reused [NoMc-01].

Ontologies allow separation of the domain knowledge from operational knowledge. A domain expert models the knowledge domain, and the operational stakeholders create instances of the knowledge domain. For example, an energy management domain expert could define the general processes and entities involved in energy management. The responsible personnel in a company describe the company-specific processes and instances by following the definitions and structures created by the expert. A formal analysis of domain knowledge can be performed once a declarative specification is available. It is useful to evaluate the reusability of the knowledge. Ontologies consist of classes (also called concepts), properties (also called roles or slots), restrictions on properties (also called role restrictions or facets), and instances. Classes are the concepts of the domain in a hierarchical order. Properties help to characterize the classes and their relations. This is, for example, the number of doors a car has or its color. Adding restrictions on properties, one might define that the smallest possible number of doors of a car is two, and the maximum is four. In a real world scenario, the classes are instantiated to create a knowledge base. Each instance represents a real world object [McEt-00] [NoMc-01].

Methods of Ontology Development

In order to create ontologies, researchers have been developing different ontology development methodologies. A development methodology helps to formalize the ontology development process and ensures that the resulting ontologies are extendable, reusable, consistent, and complete with respect to the desired scope. There are different methods of ontology development that are discussed in the literature. In the following sections, the most relevant methods are briefly described.

MENTHONTOLOGY

The Laboratory of Artificial Intelligence at the Polytechnic University of Madrid developed the METHONTOLOGY framework that focuses on the ontology construction at the knowledge level. It includes the clear definition lifecycle in the development process which starts with a prototype. The prototype is evolving resulting more and more refined version. METHONTOLOGY defines particular techniques to perform each activity [FeGJ-97] [Góme-98] [Fern-99]. The ontology development process comprises clearly defined activities. The activities include project management activities (e.g. planning, control, quality insurance), development activities (e.g. specification, conceptualization, formalization, and implementation), and support activities (e.g. knowledge acquisition, evaluation, integration, documentation, and configuration management) It is not necessary to have knowledge about other description formats than ontology language itself to implement METHONTOLOGY [Fern-99]. METHONTOLOGY offers a flexible formalization of knowledge. However, an overview about existing domain ontologies to check for integration with the ontology is still required. Thus, the level of expertise required is medium. The clearly defined techniques for each activity can save the time needed to develop an ontology. METHONTOLOGY framework has been applied to develop ontologies from different disciplines such as chemicals, environmental pollution, referencing, knowledge acquisition, law, and education [Fern-99] [CoEt-05] [SiEt-11]. METHONTOLOGY methodology also contains an evaluation activity. It uses methods for evaluating knowledge sharing technologies as described by Gómez-Pérez. [JoBV-98] Those methods are extensive, but not exclusive to the METHONTOLOGY methodology.

TOVE

Toronto Virtual Enterprise (TOVE) is a methodology to integrate a set of ontologies to support enterprise modeling. The ontologies model the entities, structures, activities, processes, etc. of a company. TOVE is not just a theoretical methodology. It has been validated in IBM Canada, by modeling the business processes using ontologies [GrAF-00]. The methodology has also been applied for enterprise workbench design, where the user can explore various enterprise design and integrated supply chain management that organizes supply chain as a network of cooperating, intelligent agent. The various practical applications, even in large enterprises, indicates that the applicability and extensibility of the TOVE methodology are very good. The methodology does not focus on the lifecycle development process. It describes the start point to extend the ontology. However, it does not state the possibilities to modify existing definitions [Fern-99]. It can be concluded that TOVE methodology only contains a fundamental life-cycle process. In order to evaluate the completeness and the correctness of developed ontology, TOVE methodology uses defined motivating scenarios, formally defined competency questions and axioms. The axiom definition requires proficiency in formulating first-order logic problems. Therefore, the ontology developer is required to have more expertise in computer science in comparison to other methodologies. Formulating formal questions and axioms is a relatively time consuming iterative task. The methodology has been applied only in a same particular domain [Fern-99]. The methodology is not widely spread across different domains. Therefore, it is good but not perfect. The completeness and correctness can be simply, automatically and accurately evaluated. This is due to the formal competency questions and axioms that are formulated in the previous steps.

Enterprise Modeling Approach

The Enterprise Modeling Approach is developed by the by the Artificial Intelligence Applications Institute at the University of Edinburgh with its partners: IBM, Lloyd's Register, Logica UK Limited, and Unilever in the frame of Enterprise project. The approach results in the creation of Enterprise Ontology. Enterprise Ontology is the underlying ontology of Enterprise Toolsets that aims to model the knowledge for supporting enterprise change management [UsKi-95] [Fern-99]. However, Enterprise Toolsets seems to be the only application of Enterprise Model Approach ontology development. Therefore, the methodology is considered good but has lower applicability and extensibility rating. The methodology does not focus on a life-cycle oriented ontology development [Fern-99]. It describes basic steps for ontology development, i.e. objective definition, ontology building (ontology capture, coding, integration of existing ontologies), and evaluation. For this reason, the developers do not have to possess advanced expertise. However, due to the imprecise definition of techniques each step, the developers need to take more time to learn how to perform the steps. Complex projects use the Enterprise Modelling Approach [Fern-99]. Hence, it considered having a high degree of methodology spread. The evaluation is difficult to perform since the methodology defines evaluation activities only roughly. There is a lack of formal definitions that can be used as the references for the evaluation.

Ontology Development Approach Noy and McGuinness

The methodology was developed based on the experience using different ontology development tools, such as Protégé, Ontolingua, and Chimaera. Some part of the methodology originated from object-oriented software design. The methodology differs from object-oriented software design because the decision making is performed based on structural properties of a class, whereas object-oriented software design focuses on the operational properties of a class. The methodology places emphasis on accuracy, intuitiveness, extensibility, and maintainability of the developed ontology. It is a continuous iterative process along the entire ontology lifecycle. It comprises the following steps [NoMc-01]:

1. Determine the domain of the ontology. The domain of an ontology is determined by formulating competency questions that an ontology knowledge base should be able to answer [GrFo-95].

2. Reuse the existing ontologies. Reusing existing ontologies allows the interaction of other applications that already have committed to particular ontologies or vocabularies. It also could save the time in ontology development since it is not necessary to create the ontology from scratch.

3. Define important terms in the ontology. The first thing to do is to enumerate a list of terms and properties that will most likely be explained to the ontology users. Initially, it is important to get a comprehensive list of terms without worrying about the overlap between concepts they represent, relations among the terms, or any properties that the concepts may have, or whether the concepts are classes or slots.

4. Define classes and class hierarchy. From the terms listed in the 3rd step, the ones representing concepts or entities have to be identified. It results in a list of ontology classes. There are three possible methods to define the class hierarchy: top-down, bottom-up, and combined approach [UsGr-96].

5. Define classes' properties. The remaining terms listed in the third step that is not identified as classes are most likely to be properties of these classes. For each property, the associated classes have to be determined. These properties become slots attached to the classes. There are two types of properties: direct or intrinsic properties and indirect or extrinsic properties. Properties of a class are inherited by all its subclasses.

6. They are derived from the properties of parts of the object, or relationships to individuals that are members of other classes.

7. Define properties' facets that describe value types, cardinalities, domain, and range.

8. Create instances by considering class memberships and slot values.

The methodology has been implemented using widely spread ontology development tools, for instance, Protégé. Most of the ontology developers use Protégé as ontology editing tool. This leads to the wide-spread use of the methodology as well. The ontology developer is not required to have deep knowledge about ontological syntax. The tool offers a user-friendly Graphical User Interface (GUI) where the developer can model the domain knowledge in ontology graphically. Even complex axioms can be modeled using the tool by a simple mean. The competency questions can also be formalized using SPARQL. Therefore, the ontology evaluation can be performed automatically.

Ontology Representation

Ontologies are considered one of the pillars of the Semantic Web, and a Semantic Web vocabulary can be considered as a special form of ontology, or sometimes also merely as a collection of URIs with a described meaning [Sema-12]. The contents of Semantic Web are presented in explicit knowledge models which include relationships between concepts in the form of categories, hierarchies, rules, concept properties, patterns, relationships, and logic. The Semantic Web consists of basic technologies such as XML, RDF, OWL, which are depicted in Figure 2.12. The figure shows the hierarchy of those technologies and their functionalities and roles in layers.

Resource Description Framework (RDF) is able to represent a data model in a triple of resources, i.e. subject, predicate, and object. The triple is called statement. A resource has a URI as an identifier. It is usually used to describe resources in the web. RDF does not support the description of properties of an object [KlCM-14]. W3C has published RDF Schema (RDFS) to extend the functionality of RDF. RDFS defines the vocabulary of RDF. It allows defining a custom terminology, i.e. type or classes of an RDF resource. It also offers possibilities to define subclasses, properties with domain and range, and class hierarchy [BrGu-14]. Therefore, RDFS represents the TBox of description logic, whereas RDF corresponds to the ABox.

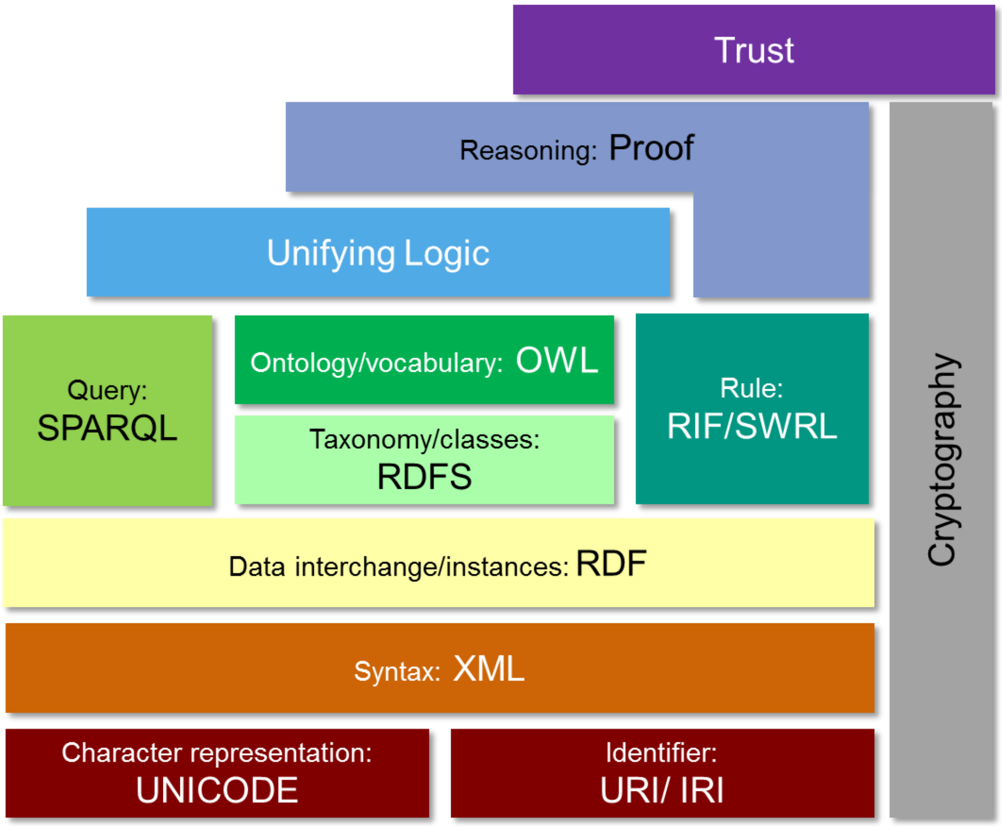


Figure 2.12 Semantic web layer cake [Sema-12]

OWL (Web Ontology Language) is an XML-based language to describe ontology in the semantic web that is developed by W3C. This language is the successor DAMN+OIL which based on XML and RDF [McVa-04]. OWL is a combination of RDFS and description logic. OWL allows advanced logic expression, such as complex class relations (conjunction, disjunction, and equality), complex instance relations, characteristics (transitive, symmetric, etc.), restriction, and cardinality of properties. According to the complexity of the language, there are three official sublanguages of OWL i.e. OWL-Full, OWL-DL, and OWL-Lite [HiKR-10]. The comparison between those languages is summarized in Figure 2.13. Rules can also be integrated into OWL represented ontologies using a language named Semantic Web Rule Language (SWRL). It is able to express logics bay combining sublanguages of OWL DL, OWL Lite, and RuleML [HoEt-04]. The different ontology languages and their expressiveness are presented in Appendix A. SPARQL Protocol and RDF Query Language (SPARQL) is a W3C standard for querying information represented with RDF. SPARQL has capabilities for querying required and optional graph patterns along with their conjunctions and disjunctions. It also supports extensible value testing and constraining queries by source RDF graph. The SPARQL queries results can be tabular results sets or RDF graphs [PrSe-08].

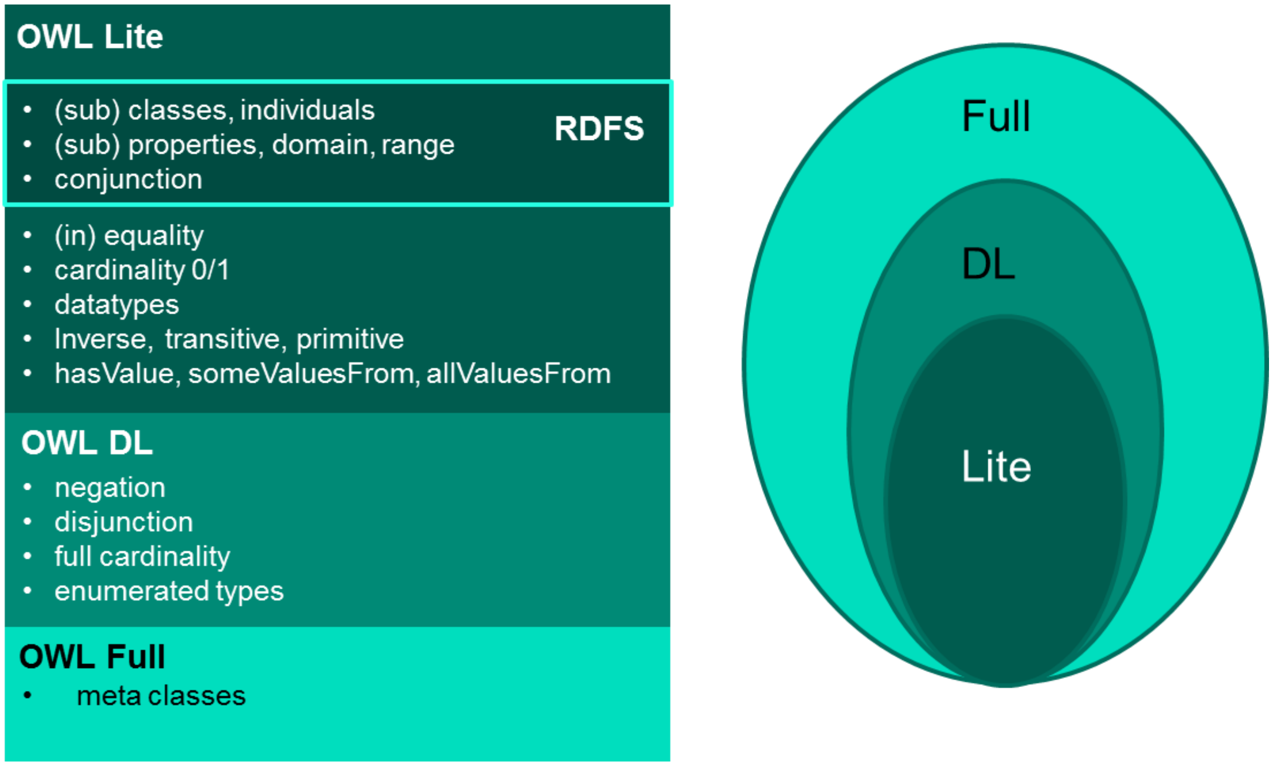


Figure 2.13 The comparison between ontology languages [HiKR-10]

Ontology in Manufacturing

Nowadays modern production systems employ heterogenic ICT systems to support planning and control activities. In order to accomplish the tasks, the different systems require interconnection. The interoperability is the crucial issue to be dealt with. To overcome the interoperability problems, researchers develop ontologies that offer formal definitions of entities and relationship in a particular domain [ChEt-07].

Schilberg and Meisen developed an ontology-based approach for solving interoperability problem in order to reduce the complexity of distributed numerical simulation for production. They introduced ontology to represent the simulations parameter, value range, and its meaning. It enables the user to observe the manufacturing process on a microstructure scale from the used material to the linked manufacturing process steps. All input and output data of the whole simulated process chain are correlated so that the loss of semantic information will be prevented. The approach allows a better understanding of semantic interconnections of distributed numerical simulation systems. It is then used as the basis for virtual production system implementation [ScMe-09]. Ontologies are also used to represent a common knowledge based model to describe processes workflows and events in a discrete production environment. They allow a semantic interoperability among manufacturing IT systems including schedulers, planners, legacy data transformation services and end-user client interfaces. For this, three ontologies are developed to model the core elements of a manufacturing system, the production flow, and events or disturbances [HaVr-15].

Ontology is also used to describe the processes and resources of assembly lines to address the interoperability problems in collaborative engineering processes. A set of rules is integrated into the ontology for the mapping between product requirements, product components, and production processes. It results in an interconnection between product attributes, manufacturing processes, and resources [FeEt-15]. Hai et al. developed ontology based a generic and extensible modeling language that models different types of work processes in chemical industries. Domain-specific knowledge is captured in the meta-model formally at different levels of detail. They introduced a modeling tool named WOMS-plus which is developed to apply the modeling language in industrial practice [HaTM-11]. An ontology-based approach is also proposed by Alsafi and Vyatkin to achieve fast reconfiguration of modular production systems. They developed an agent that uses ontology as knowledge base representing manufacturing environment to reconfigure the systems without any human interventions. It intelligently decides the new manufacturing configuration using the inferred facts from the ontological knowledge base [AlVy-10]. Ontology model is also developed to describe distributed manufacturing services and to facilitate efficient, accurate, and automatic retrieval of the required manufacturing services. It establishes a seamless interoperability in collaborative manufacturing environment [CaEt-11].

Panetto et al. developed an approach called ONTO-PDM representing product ontology that implicitly embeds the information about itself. The approach aims to denigrate the semantic uncertainty. The ontology provides an accurate and expressive information structure so that the search of product technical data can be easier. The implementation of the ontology has been done through (1) its translation into OWL language for its use with the Protégé ontology development and instantiation environment and (2) its translation into an Entity/Relationship model for its implementation into a database management system [PaDT-12].

Table 2.2 shows the comparison of related work in the application of ontologies in manufacturing. The comparison is based on the following criteria:

1. Knowledge acquisition and population method, how the knowledge is captured and formalized and how the instances are created.

2. The modeled entities and knowledge resources

3. Language to represent the ontology

4. Tool used to create ontology

5. Standards complying with the developed ontologies.

Table 2.2 Application of ontology in manufacturing

Approach	Knowledge acquisition and population method	Modelled knowledge	Representation	Tools	Compliances
Schilberg and Meisen [ScMe-09]	Data conversion, enrichment (parsing, reasoning)	Manufacturing processes, process linking, process properties (value range, unit, context)	n.a.	n.a.	n.a.
Alsafi and Vyatkin [AlVy-10].	Manual from experts	Resources (machines, tools), Operations (logistic, manufacturing), controller	OWL-DL	Protégé, Jena	IEC 61499 – distributed control and automation, MASON
Hai et al. [HaTM-11]	Manual from expert, WPML conversion	Operations: taxonomy/hierarchy, composition, function, steps, actor/controller, resources	OWL	Protégé, Jena, WOMS-plus (self developed)	WPML
Cai et al [CaEt-11]	Manual	Manufacturing equipment	OWL, SWRL	n.a.	n.a.
Panetto et al. [PaDT-12]	Manual	Product (definition), Process (segment, schedule, performance), Resources (material, equipment, personnel)	OWL, entity-relationship model	Protégé	IEC 62264
Ferrer et al. [FeEt-15]	Manual	Product, process, station, system	OWL, SWRL	Olingvo (Univ Tempere, Finland)	n.a.
Harcuba and Vrba [HaVr-15]	Manual	Product type, process, material, process goal	OWL	n.a.	n.a.

Ontology for Energy Management

The energy consumption in the building sector takes a significant proportion of the world energy consumption. It results in more emission of greenhouse gas. Energy efficiency improvement in buildings is necessary to reduce energy consumption and resulting greenhouse emission. Researchers have been developing ontologies as an underlying information model for energy management systems.

Van Dam and Luzko utilized ontology to model knowledge to improve the efficiency and reliability of energy and transport infrastructures. They developed a multi-agent based approach for controlling the infrastructures. They built a generic ontology of the infrastructure domain which enabled the extension and specialization to domain specific concepts. The proposed approach was applied in the electricity, process industry, and a case study in the transport sector domain. They developed a case study by cooperating with the Queensland University of Technology to evaluate different locations for a new freight hub. The result can be used to extend the database with ontologies and to gain knowledge about different control strategies. This modeling with ontologies can help them to improve the performance of existing infrastructures and a better understanding of decision making in the system. They developed local optimization algorithm to improve the system performance [VaLu-06].

Daouadji et al. presented an ontology-based virtualization to create an efficient resource description and discovery framework in a smart grid environment. It is developed particularly for ICT management purposes. The ontology modeling focuses on the energy-related semantic of resources and their properties, such as energy consumption. The developed resource discovery engine integrates a Bayesian network into the knowledge base in order to improve searching performance [DaEt-10]. Furthermore, Dodgu et al. introduced an ontology as an underlying information model of a decision support system. The system allows all stakeholders of the sector in taking the right decisions for more efficient energy generation, distribution and consumption [DoEt-14].

Han et al. developed ontology based building energy management system based on ontology, inference rule, and simulation. The ontology that offers information reusability, extensibility, and interoperability enables easy deployment of different Building Energy Management System (BEMS) technologies. Inference rules are generated based on both building energy characteristic and the domain knowledge. The number of generated rules contribute proportionally to the increase in energy efficiency. This model is expected to have a significant contribution to the energy saving and reduction of greenhouse gas emissions in building sector [HaJL-11]. Furthermore, Han et al. developed a rule-based ontology that provides semantic context awareness of building energy management and allows reasoning to detect energy waste and to recommend energy saving measures to eliminate the energy waste [HaJL-15].

Home Energy Management System (HEMS) allows the user to control devices in the home automation network through an interface and apply energy management strategies to reduce and optimize their consumption. However, there exist a number of devices and appliances in homes manufactured by different producers. They usually have different application models and communication technologies. Rosello-Busquet et al. introduce a semantic technology equipped home gateway that is able to integrate different devices within HEMS. It provides a common control interface and functionalities to implement energy management strategies. The home gateway is developed using OSGI-Equinox framework and Protégé-OWL API [RoEt-11]. Grassi et al. proposed a holistic approach for smart home environments that includes an ontology framework. The ontology acts as the information model representation of the smart home environment. They developed two ontologies which provide a flexible and extensible structure of information model. The ontologies make task management for energy consumption more intelligent. Thus, it leads to optimized energy management. All relevant information about the whole system is encoded into a global ontology knowledge base. The knowledge base can be exploited for providing efficient control logics and intelligent decision making related to energy consumption [GrNP-11].

An ontology-based holistic information modeling approach for building energy management is also developed by Wicaksono et al. The ontology is the core component of an intelligent building energy management system which incorporates existing BEMS. The developed ontology offers an expressive representation and reasoning capability of a building information model. The TBox components of the ontology are created by some experts of energy management system. The TBox model is aligned with existing Building Information Modelling (BIM) standard, i.e. IFC. The ontology populated using different means. The building topological instances are created semi-automatically by using OntoCAD tool that interprets the semantics of geometry data and maps the interpretation results to the respective ontology classes. The ontology contains occupant behavior model to depict user activities in the building. Furthermore, a data mining technique is incorporated to enrich the ontology with rules extracted from historical data. The resulting ontologies are exploited to detect the anomalies and energy wasting points, and also to gives recommendations of actions to the user to achieve desired energy efficiency targets [WiRo-10a] [WiAR-12] [WiEt-15].

Table 2.3 summarizes the related work in the application of ontologies in energy management. These work are compared based on the following criteria:

1. The scope of the ontology application, for example, home, public building, grid, etc.
2. Knowledge acquisition methods
3. The modeled entities and knowledge resources
4. Representation language
5. Tools used to create the ontology

Table 2.3 Application of ontologies in energy management

Approach	**Application scope**	**Knowledge acquisition and population method**	**Modelled knowledge**	**Representation**	**Tools**
Han et al. [HaJL-11]	Business building	Manual	Building architecture: area, zone, context: building operation , cause: building state, usage context, control: solutions	Hierarchy, rules	n.a.
Daouadji et al. [DaEt-10]	Energy grid (utility company)	Manual	Energy consumption, ICT devices, energy sources	RDF	n.a.
Rosselló-Busquet et al., [RoEt-11].	Household	Manual	Home appliance, commands, building environment	OWL, SWRL	Protégé, Jess rule engine, OSGI
Grassi et al. [GrNP-11]	Household	Manual	Service, user and usage context, device, energy	OWL	n.a.
Dogdu et al. [DoEt-14]	Energy grid	Manual	Building, electricity network, energy generation, intelligent device	OWL	n.a
[HaJL-15]	Building	Manual	Building, time, weather, space, temperature, etc.	OWL, SWRL	n.a.
Wicaksono et al. [WiRo-10a] [WiAR-12] [WiEt-15]	Household, building	Manual, data mining, 2D-drawing semi-automatic interpretation	Appliance, building environment, BAS, user activities, surrounding parameters	OWL, SWRL	Protégé

2.3.5 Knowledge Acquisition through KDD

It has been said that we are drowning in data, but starving for knowledge. Nowadays, modern computer systems are storing data routinely in databases. Therefore, the amount of data will always be larger. It has been estimated that the amount of collected information in the world doubles every 20 months. This information could be from a very wide variety of sources, such as the internet, industrial and commercial information. Data mining analyzes the data in databases and discovers patterns that are universal accepted, understandable, and necessary. Data mining helps to filter the available information, to find the desired information and to result in specific analysis from the information that already presents. Data mining is defined as the process of discovering patterns in data. The process must be automatic or (more usually) semiautomatic. The discovered patterns must be meaningful in that they lead to some advantage, usually an economic advantage [WiFH-05].

Definitions

The overall process of "Knowledge Discovery in Databases" or in short form KDD means all steps that are necessary to generate knowledge from given data. Besides the essential data-mining step, KDD involves another process on the data, such as selection, pre-processing, and transformation. Figure 2.14 gives an overview of the necessary steps for a successful knowledge creation by means of KDD.

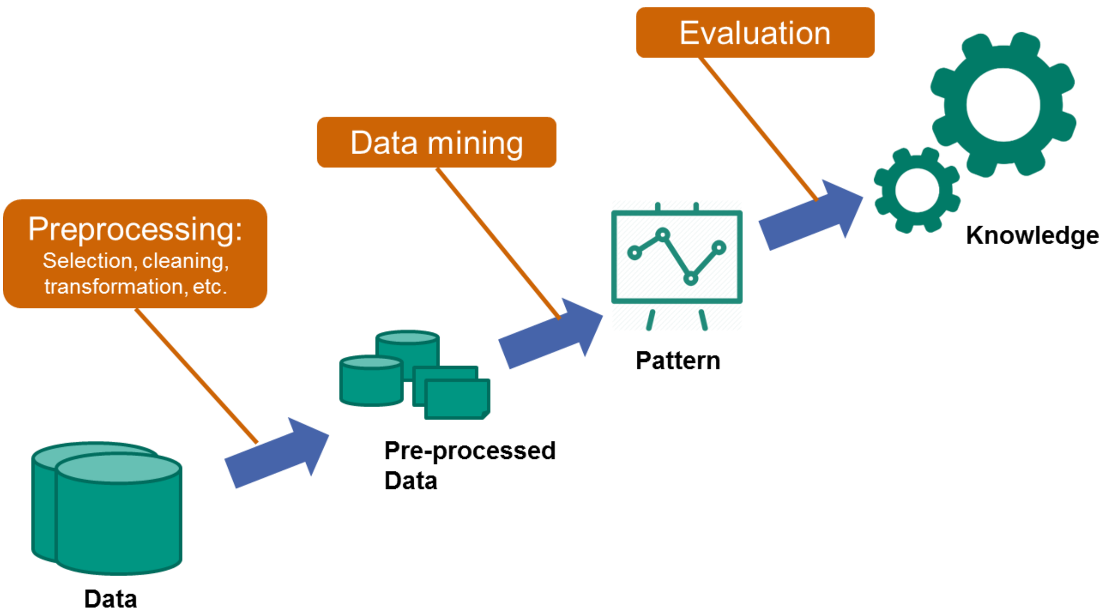


Figure 2.14 KDD process [Bram-13]

Data mining is the analysis step of Knowledge Discovery in Database (KDD) process as depicted in Figure 2.14. The data are collected from different sources. These data are then stored in a data storage such as a Database Management System (DBMS). If necessary, some of the data are pre-processed, for example, transformed into a standard format [Fark-11]. The standardized data format enables the execution of data mining algorithms. The data mining algorithm extracts the targeted patterns or rules from those data, according to the objectives of KDD process. Evaluation or interpretation of these results is then carried out to find potential knowledge that conforms to the KDD defined objectives. The disorganized, incomplete, and irrelevant raw data will be processed by different methods of data analysis such as association and classification. Before the data mining are further described in this section, it is important to understand the following basic terms [WiFH-05] [Bram-13].

- Dataset, a collection of data that normally presented in tabular form.
- Attribute, a property inherent in a database entity or associated with that entity for database purposes.
- Item, it represents a physical object like for example a product

The concept of data mining is in principle the process step of the analysis of data in terms of finding useful patterns and rules [FAPS-96]. Data mining, in this case, is a mixture of concepts and algorithms from the fields of machine learning, artificial intelligence, statistics, and data management and the application of these algorithms in order to find specific patterns in the data. In KDD except the pattern extraction and modeling, the interpretation of the data is also essential. It requires application-specific knowledge of the particular application area to make a good and purposeful result selection. The difficulty and the success of the knowledge discovery depend so strongly on the identification and distinction of a valid, new, potentially usable, and well-understandable knowledge [HaEt-06] [FAPS-96].

Data mining techniques are mostly inspired by the principle of machine learning. The difference is that, in data mining, the human plays a role in defining the goals and carrying out the process. There are two categories of machine learning techniques, i.e. supervised and unsupervised learning. Supervised learning gives statements or labels to the training examples. Classification of iris or techniques to determine an image is a human face or not, are examples of supervised learning. On the contrary, unsupervised learning no labels or statements are provided in the training data, for example clustering the different images based on their similarity without defining the name of each cluster. The following section elaborates some common techniques for data mining.

Data Mining Techniques

This focuses on the most widely applied data mining techniques which are potentially utilized to help the tasks of energy management in manufacturing. They are association rules, classification based on a decision tree, and neural network. The following chapter presents the classification using decision tree analysis. The explanation of association and neural network are presented in Appendix B since they will not be applied in this thesis.

Decision Tree Analysis

A decision tree is a visual and explicit representation of decisions in a decision-making process. In data mining, a decision tree describes data as inputs for decision making. Decision tree represents a hierarchical arrangement of decisions nodes and can be derived from the simple divide and conquer algorithm. Branches connect the nodes extending downward from the root node until terminating leaf node. The hierarchy is created by grouping on the basis of attribute's value from data. A decision tree is modeled through the following three steps [Bram-13]:

1. Attribute selection for splitting a node

   The attribute selection is a continuous step-by-step process. The attribute is selected based on its information content that is measured using entropy. Entropy is an information-theoretic quantity representing the 'uncertainty' contained in a training set, due to the presence of more than one possible classification. The entropy is denoted by *E*. It is measured in 'bits' of information and is calculated using equation (2.7). p_i is a value from 0 to 1 that represents the number of occurrences of class *i* divided by the total number of instances in the dataset. *K* is the number of possible class [Bram-13].

$$E = -\sum_{1}^{k} p_i \log_2 p_i \qquad (2.7)$$

2. Splitting the node

   The attribute will be divided into two or more subsets based on the selected node.

3. The termination criterion to determine an end node

   In order to determine end node, the expected information gain from splitting is measured. If the information gain is very small in statistical context, then the node become an end node. The real datasets that contain identical objects can lead to some errors. The tree pruning is carried out to minimize these errors.

4. Post-Pruning

   Classification performed on larger datasets may result in very complex trees that actually lead to overspecialization. There are numerous end nodes that contain only very few objects. This results in decreased generalization ability and a bad grouping of the objects of the test set. A separate post pruning process is, therefore, essential to generate a reasonable decision tree.

Classification and Classification Rules

Classification is a method to assign different objects into one of a number of mutually exclusive pre-defined categories called classes. Each object has to be assigned to precisely one class [Bram-07]. The popular algorithm in classification is Naïve Bayes Algorithm. The Naïve Bayes Algorithm is a probability-based classification method that based on the Bayes' theorem [WiFH-05]. The algorithm assigns each object to the class, where it has the highest probability to belong to the class [Neap-04].

A decision tree is a representation, from where a set of classification rules can be derived. The rules are often called decision rules. Other than association rules, classification rules cannot predict other than the class attributes. The conclusions of classification rules can have only one attribute i.e. class attribute. Meanwhile, the association ones can have several attributes. Classification assigns data in different pre-defined mutually exclusive classes or categories. Classification using decision trees is a widely used method for constructing a model from a dataset in the form of a set of classification rules.

A classification rule corresponds to a path in a decision tree from the root to the leaf. The leave represents the class attribute, whereas other nodes including the root and internal nodes correspond to the splitting attributes. Figure 2.15 shows an example of a decision tree having three splitting attributes and four leaves representing three different classes.

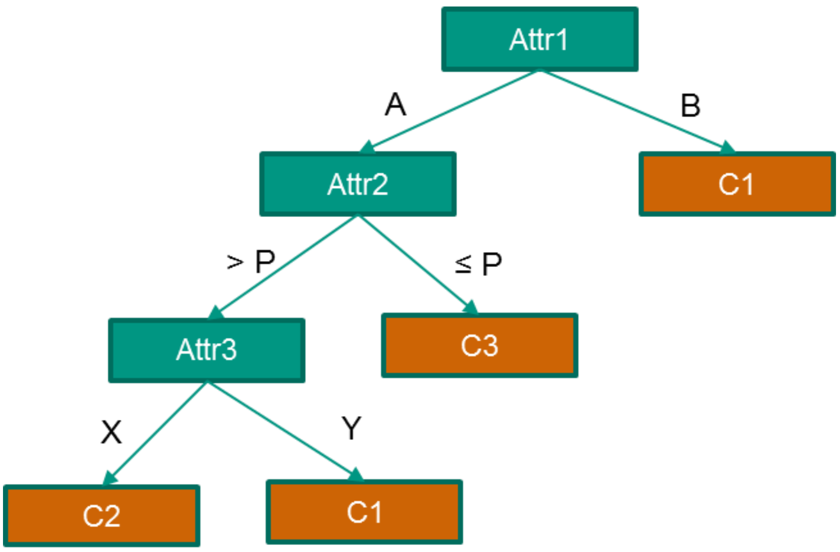


Figure 2.15 Example of a decision tree corresponding to a set of classification rules

From the decision tree shown in Figure 2.15, the following four classification rules can be derived. The number of rules corresponds to the number of leaves in the decision tree.

```
IF Attr1 = A AND Attr2 > P AND Attr3 = X THEN Class = C2
IF Attr1 = A AND Attr2 > P AND Attr3 = Y THEN Class = C1
IF Attr1 = A AND Attr2 ≤ P THEN Class = C3
IF Attr1 = B THEN Class = C1
```

(2.8)

The antecedent consists of a set of terms joined by the AND operator. Each term expresses the check on the value of an attribute. The decision rules are disjunctive among each other. The term representing the class attribute value constitutes the consequent part. This representation is often claimed that it has the advantage compared with other approaches due to being meaningful and easy to interpret [Bram-13].

Data Mining in Manufacturing

Application of Data Mining in Production Planning and Control

Facility layout planning could influence productivity and efficiency of manufacturing facilities significantly. It contributes directly to material handling costs. Altuntas and Selim proposed a data mining techniques with weighted association rule to solve facility layout problem. They introduced five different weights on each association rule. This is unlike the classical association rule approaches that assume an equal level of significance on each item. The following factors are considered to determine the weights: demand, part handling factor, and part handling factor and efficiency of material handling equipment. They develop a simulation to compare different performance measures, namely machine utilization, the total amount of products produced, cycle time, transfer time, and waiting time in queue [AlSe-12].

Shahzad and Mebarki developed a data mining technique to find unknown priority dispatching rules for job shop scheduling problem. The data mining discovers knowledge represented as a set of rules which determines the efficient solution for a job shop scheduling problem (JSSP). This set of rules is integrated with a tabu search based optimization module. The generated knowledge helps the tabu search algorithm to carry out the move more intelligently. The knowledge is also used to help the tabu search algorithm to reorient in large problems which usually occurs in real manufacturing. To achieve the objective Shahzad and Mebarki developed optimization module to find the most relevant solution to the scheduling problem, simulation module to measure the performance and functionality of the system, and learning module to generate the knowledge. The learning module uses decision tree based algorithm namely C4.5 because the representation of decision tree allows a better understanding and transparency of scheduling decision classification process [ShMe-12]. Bergmann et al. introduced data mining approach to approximate dispatching rules for production scheduling. The existing production scheduling models are normally too general and do not correspond to real-world problems. The proposed data mining approach generates the production scheduling simulation model that reproduces the dynamic behavior of real systems [BeFS-15].

A data mining based approach was also proposed by Windt and Hütt to identify the causes of lateness in multistage production systems demands by identifying intrinsic interdependencies between order and process parameters. They used cluster analysis combined with enrichment analysis adapted from gene expression analysis to extract the relationships between order, resource parameters, and lateness [WiHü-11].

Application of Data Mining in Production Optimization and Quality Control

According to Vazan et al. the tasks of production management are to achieve specified certain production related goals, such as utilization, quality, energy efficiency, etc. at a given time frame. Since many of these goals relate very complex and often also conflictive, a comprehensive optimization of all factors of production is very expensive. Many dependencies of different parameters are unknown, for example, the dependency between capacity utilization, cycle time, selected batch size, and variable costs. It is difficult for the company to determine the generated added value and the impact on the specified production targets. These problems and their dependencies can be solved using data mining. The data mining model and techniques are developed specifically for each method problem in manufacturing management. The discovered knowledge is tested using the developed model. The knowledge is then utilized to identify the impact of manufacturing parameters on the production process and to give recommendations for optimization measures. For example, the failures, failed parts, emergency situations or states that can negatively affect the production process can be predicted more accurately. Moreover, it acts as preventive control instrument of the production equipment, supports decision-making activities performed by manufacturing managers to forecast the cost of the production processes, and the customer's behaviors [VaTK-11]. Kamal developed data mining approach for improving the quality control process of manufacturing. He indicated that Statistical Process Control (SPC) played an essential role in understanding the variety of manufacturing process. SPC methods could help to improve different quality assurance processes. SPC is considered as data analysis method and a process management philosophy simultaneously. It employs data mining techniques for continuous improvement and enhancement [Kama-11].

To solve problems in manufacturing quality management, many existing Quality Improvement (QI), such as six sigma and kaizen involve data collection and analysis. Data mining has also been widely utilized. Data mining offers advanced techniques in data collection and analysis. Data mining supports the quality tasks, such as product/ process quality description, predicting quality, classification of quality, and parameters for optimization, are provided on data mining. These functionalities can be found in manufacturing industry applications. Köksal et al. presented the result of a comprehensive review of data mining applications for selected QI problems [KöBT-11]. Khan et al. proposed a data mining approach to extract valuable information from manufacturing data to optimize production processes. The approach generates summary statistics and predicts process performance based on categorical data that leads to the final cost reduction. It uses genetic algorithms to classify the categorical data which are generated during quality control process [KhEt-15].

The applications of data mining in manufacturing described above are summarized in Table 2.4. It elaborates the application field of data mining, types of processed data, data mining goals, and used techniques.

Table 2.4 Summary of data mining applications in manufacturing

Approach from	**Application**	**Processed data**	**Objective**	**Data Mining technique**
Vazan et al. [VaTK-11]	Production optimization	Capacity utilization, lot size, variable costs, supply, value chain	To maximize and capacity utilization and total number of finished part To minimize flow time and production costs	• Regression Analysis: neural network, CHAID, standard regression tree • Classification: neural network • Cluster analysis: K-Means, EM
Altuntas and Selim [AlSe-12]	Production planning	Demand, part handling factor and efficiency of material handling equipment	To maximize machine utilization, total amount of products produced, To minimize cycle time, transfer time, and waiting time in queue	Weighted association rules
Kamal [Kama-11]	Quality Management	Product and production process data	To improve manufacturing processes quality control	Decision tree and classification
Shahzad and Mebarki [ShMe-12]	Production scheduling	Scheduling data: jobs, tasks, products, workstations	To improve the optimization technique in solving Job Shop Scheduling Problem	Decision tree based classification: C4.5
Windt and Hütt [WiHü-11]	Production scheduling	• Order: type of product,	To identify and predict lateness	Cluster analysis

		dimensions, weight • Production: production depth • Process: sites change, rework intensity, quality, allocation change		
[BeFS-15]	Production scheduling	Production processes	To generates rules for dispatching decision	K-Nearest neighbors, naïve Bayes classifier, support vector machine, decision trees
[KhEt-15]	Quality management, production optimization	Production process, sensor data	To extract valuable information from manufacturing data for production process optimization	Genetic algorithm

Data Mining for Energy Management

In the recent years, each sector including building sector has been demanded to improve energy efficiency and to have green design features. Kim et al. state that the whole processes along entire building lifecycle have to be considered to improve the energy efficiency of buildings. It is not enough to apply only a certain technology, such as isolation technology. This consideration leads to the generation of a large amount of data during the building design and simulation. Kim et al. develop energy efficient building design process utilizing data mining to discover interrelationship and patterns of interest from a large dataset of energy consumption to improve the energy efficiency of building design during the design phase. The results could then be exploited as the guide for building designers or engineers to perform early design analysis based on energy simulation models [KiSK-11].

Yu et al. introduced data mining technique to extract the relations and associations among building operational data. The building operators and owners could obtain useful rule-based knowledge from these relations and correlations based on their domain expertise. The knowledge is used as to help the energy management activities in the building. By means the knowledge, HVAC (Heating, Ventilation, and Air Conditioning) system performance can be improved, and energy consumption can be reduced. The rules help the building operators and owner to identify energy waste, to detect faults in equipment, and to propose low-cost strategies for saving energy. The methodology employs a basic data mining technique association rule mining [YuEt-12].

A data mining based approach was also developed by Veleva and Davcev to analyze the collected energy consumption data from different appliances in a household in order to identify a cluster from each appliance or a cluster from a group of appliances plugged-in in the same time to a single power socket sensor. The accuracy of developed clustering algorithm enables the identification of the operating states of the household appliances. Additionally, an evaluation of energy consumption can be improved. The data mining algorithms are integrated into a Home Energy Management System (HEMS) that is able to control power socket, based on the exact identification of the plugged-in appliances [VeDa-12]. Zucker et al. also employed data mining algorithms to analyze the operation of HVAC systems. They developed X-Means clustering algorithm to detect the system states, for example, errors or failures, and to allocate duty cycle periods of the systems automatically with minimum pre-configuration [ZuEt-15].

The application of data mining in households was also a result of project KEHL funded by German Government. The developed data mining algorithms detect the energy consumption patterns and their relationships to user behaviors, building states, and surrounding parameters of household buildings. In their approach, they use different data mining techniques namely decision tree based classification and cluster analysis to generate knowledge that is exploited to support energy management activities in household buildings. The knowledge is then used to extract energy usage pattern, to recognize energy usage anomalies, to determine the classification of appliances based on their energy consumption, and to identify operation state of an appliance based on energy usage. This approach allows semi-automatic acquisition of knowledge components of an ontology represented knowledge base [WiRK-10] [WiRo-10b] [KEHL-11].

As the summary of this section, Table 2.5 compares different existing data mining approaches to improve energy efficiency based on the application area, the data mining objective, the types of input data, and the used techniques to extract the knowledge.

Table 2.5 Summary of data mining applications in energy management

Approach from	Application area of Energy Management	Objective	Processed data	Data mining technique
Kim et al. [KiSK-11]	Design phase of general building	• To identify correlations reflecting building actual performance • To find the most cost effective alternative in system design	Baseline building model parameters (users, area, floors), roof construction (material, insulation, air space, etc.) wall construction (material, insulation, construction methods), wall and roof insulation, etc.)	Decision tree based classification: C4.5
Yu et al. [YuEt-12]	Public building	• To detect energy wastage and faults in appliances • To achieve low-cost strategies for saving energy	Climatic data, building operation data, building physical parameters	Association rules mining
Veleva and Davcev [VeDa-12]	Household	• To identify a cluster from each appliance, or a cluster from a group of appliances plugged-in in the same time to a single power socket sensor • To identify operating states of appliances	Appliances’ energy consumption, appliances’ control states	Cluster analysis
Zucker et al. [ZuEt-14]	Building	To detect the states of HVAC systems in order	HVAC data	X-Mean clustering

		to optimize their operations		
Wicaksono et al. [WiRK-10] [WiRo-10b] [KEHL-11].	Household	• To identify energy usage pattern and usage anomalies • To identify appliances' operation states • To classify appliances based on their energy consumption	Energy consumption of appliances, appliances' states, building layout, user activities	Decision tree based classification, cluster analysis

As shown in Table 2.4 and Table 2.5, the decision tree approach has been commonly developed to support prediction and optimization tasks in manufacturing and energy management. This is due to the simple and easy-to-understand model offered by the decision tree. Many manufacturing activities deal with discrete or categorized data, for example, material types, machine categories, etc. The decision tree approach has been proven to be able to deal with those kinds of data. However, the common decision tree algorithms are only able to classify the instances into discrete classes. The application areas or decision tree are mostly predictions of certain values and optimizations that consider multiple parameters. There is still a lack of applications of decision trees that incorporates a classification method to predict numerical values in manufacturing energy management domain. This thesis tries to address this gap.

2.4 Energy Performance Indicators

This section deals with the fundamental concepts and state of the art related to the energy performance indicators which are developed in this thesis. The first part elaborates different energy-related figures. It is then continued with the discussion about existing energy performance indicators which have been developed by academia and industries. Finally, those energy performance indicators are compared by considering defined evaluation criteria.

2.4.1 Energy Related Figures

This section gives an overview of simple and widely used energy-related figures in industries and academia. It starts with the taxonomy of the figures. It continues with the introduction of several physical and economic figures which are interesting to achieve the objectives of this thesis.

Energy Figures Typology

Depending on the complexity of a system, there are different types of figures that give statements about the energy efficiency of the system. The classification of different types of these figures is shown in Figure 2.16 [Kals-10].

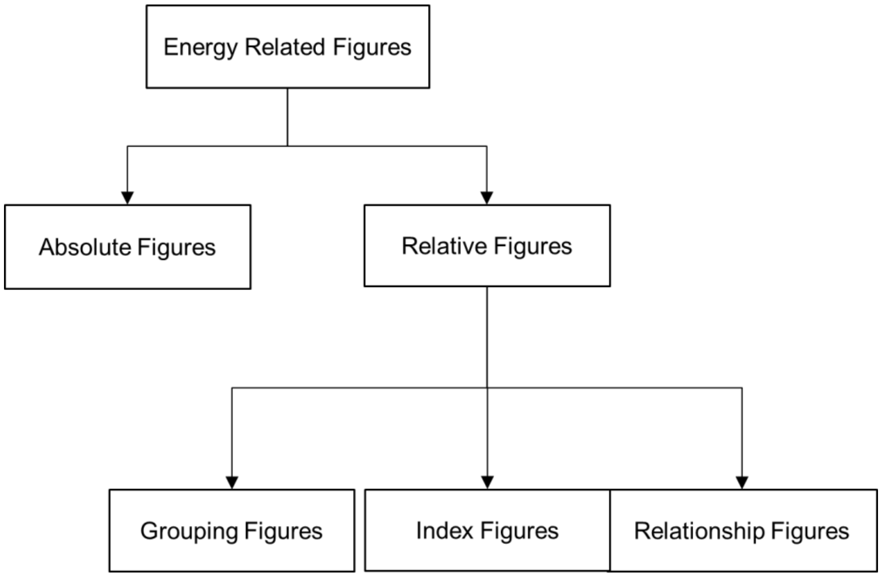


Figure 2.16 Classification of energy-related figures [Kals-10]

Absolute figures are for example electricity consumption of a factory hall per week, the gas consumption of an oven per day. Even if a time reference is given, such numbers are considered absolute rather than relative indicators. Since energy efficiency is always associated with the relation of efforts and benefits, absolute figures cannot be considered to represent efficiency. The figures can only give an indication which energy consumers should be paid more attention in the context of a corporate energy management. At the points of highest consumption can also have the highest savings are expected by increasing efficiency.

Grouping figures indicate the proportional relation of a parameter to a total amount, for example, the gas consumption of a hall in relation to the total gas consumption of the site. Grouping figures can be visualized well in pie charts. Like absolute figures, the grouping figures cannot indicate the efficiency. The grouping figures have still an advantage comparing to absolute figures. They can give more detailed statements which part of the factories consumes more energy, so the highest energy saving potential can be determined.

Relationship figures reflect the correlations between unequal numbers of different populations. They describe cause-effect relation, such as energy consumption per unit product, or energy consumption per unit of a machine. The effect indicator (energy consumption) is in the numerator of the fraction, the causative size (product or unit of work) in the denominator. Relationship figures are the appropriate metrics to evaluate the performance. The possible benchmarks will be discussed immediately.

Index figures describe changes that are dependent on time. Examples are indexes for energy prices, energy costs, or energy efficiency, for example, an index of changes in the output quantity in relation to the use of energy. They can be useful to check and to assess the success of the implementation of energy efficiency improvement measures.

Physical Figures

There are a number of general energy performance measures which are described in the VDI guideline 4661 [VDI-03]. The ones relating to energy efficiency are presented below.

"The energy efficiency η is defined as ratios obtained from target energy flows supplied and energy flows used in a process or in a plant in the stationary or quasi-stationary state."

Based on the target energy flow rates, a categorization of the electrical efficiency, thermal efficiency, etc. is made. The definition of energy efficiency also considers the thermodynamic aspect which takes into account the energy conversion. The quality of energetic conversion is represented as exergetic efficiency ξ or quality factor [VDI-03].

Utilization ratio $\boldsymbol{\eta}$ is the quotient of the usable target energy output in a particular time period and the total energy input.

In order to calculate the utilization ratio, a measurement period comprising operating, all pauses, standstills, idling, starting and stopping times is considered.

The specific energy consumption/ energy demand *w* refers to the use of energy in joules (J) or watt-hours (Wh) of a system related to the functional unit (FU) of a product or service. In determining the specific energy consumption, all supplied energy has to be taken into account. It means, in addition to the main energy sources such as fuel and the electric energy, the energy supplied from compressed air, heat, hot water, and hot oil, have to be considered. Table 2.6 describes the different specific energy consumption according to VDI standard 4661 [VDI-03].

Table 2.6 Specific energy consumption variables and units [VDI-03][Kals-10][BuEt-11]

Sector	**Reference variable**	**Energy consumption unit**	**Classification based on [Kals-10]**
Industry	Amount of production (PM)	MJ/PM	Relationship figure
	Mass	MJ/kg	Relationship figure
	Energy cost per production unit	GJ/prod unit, kWh/prod unit	Relationship figure
	Energy consumption per department	GJ/department	Grouping figure
	Power consumption per personnel	kW/person	Relationship figure
Energy conversion	Secondary energy, final energy	MJ/MJ	Relationship figure
Transport	Transport tons of goods	MJ/(t · km)	Relationship figure
	Transport of persons P	MJ/(P · km)	Relationship figure
Households, public buildings, offices	Air conditioning: housing volume and area	MJ/(m^3 · a), MJ/(m^2 · a)	Grouping figure
	resp. number of person P in a period	MJ/(P · a)	Relationship figure
	Lighting: luminous flux, illumination, etc.	MJ/lm, MJ/lx	Grouping figure
	Equipment: working periods (fridge, monitor, etc.)	MJ/h, MJ/a	Relationship figure
	Equipment: working processes: (washing , printing, etc.)	MJ/kg, MJ/Seite	Relationship figure

Gutowski et al. developed a specific indicator to evaluate the electrical power requirements of manufacturing processes. The indicator considers the series of production steps and production supporting activities, such as retooling activities, work handling, and lubrication. It also takes the energy requirements of different equipment states, such as start-up and ready states. The indicator represents the specific electrical exergy per unit of processed material (B_{elect}) and is calculated using equation (2.9) [GuDT-06].

$$B_{elect} = \frac{P_0}{\dot{v}} + k \; [kJ/cm^3] \tag{2.9}$$

where

P_0	=	Idle power [kW]
$\dot{v}$	=	Material processing rate [cm^3/s]
k	=	A constant [kJ/cm^3]

Po represents the electrical power from supporting equipment and does not correlate to the production process. It comes from for example energy consumption of coolant pump, hydraulic, and information infrastructures. The constant k depends on strongly on the characteristics of the processed materials or products, such as the hardness and the material type.

A similar approach is also considered by Patterson, who calculates the energy efficiency based on the ratio of output in tonne-kilometres to the energy input. The output does not only represent the production output. Additionally, it reflects the services delivered to the end customer since the product values also depend on the distribution and logistics processes [Patt-96]. The indicators introduced by Gutowski et al. and Patterson make the range of possibilities to measure energy consumption bigger, but it is more difficult to perform benchmarking. The energy requirement related to the amount of product is very dependent on the type and quality of the product. For example, the energy consumption to manufacture a truck is always higher than to manufacture a car.

Economic Figures

A figure that can overcome the problem in benchmarking is introduced as energy efficiency that is calculated using the following equation [MüEt-09]:

$$Energy\ Efficiency\left[\frac{€}{MWh}\right] = \frac{Netto\ productopm\ value\ [€]}{Primary\ energy\ usage\ [MWh]} \quad (2.10)$$

This energy efficiency calculation also allows for a comparison of plants that produce completely different products and can be applied in addition to the machine level and on the factory floor, even if there different products at the same time in different quantities or qualities. The primary energy usage as a reference takes into account the conversion losses (e.g., in coal plant). If they are outside the defined system boundaries, should not be considered. In the macroeconomic level, energy-related metrics are presented in relation to parameters, such as the gross domestic product (GDP). The GDP corresponds to the monetary value of the produced goods and services [DiEt-99]. Energy intensity is an example of these metrics. It is defined as the ratio of energy consumption to a monetary value, for example, the GDP [PhBW-97] [IrTh-06] [BuEt-11].

2.4.2 Existing Energy Performance Indicators

Several energy performance indicators have been developed to enable benchmarking and evaluation of energy efficiency among companies or organizations. They are mostly unitless. It brings different factories or plants in a relation thus allowing an assessment of the optimization potential. The benchmark is always based on the "best in class" corresponding to the factory which has the best figures. The data collection is here the major obstacle, as many companies do not want to publish their figures. They do not want to provide the data for the competitors. The other reason is that the anonym databases are scarce. The following sections give an overview of those figures.

Energy Efficiency Index from EEP Stuttgart

The Institute for Energy Efficiency in Production (EEP) at the University of Stuttgart, Germany published Energy Efficiency Index in July 2014. The policy makers use the index to measures energy performance of manufacturing in Germany in order to develop more accurate policies and incentives. Besides to measure the current energy performance (in the last six months), the indexes give a sight in the future (in the next 12 months). The Energy Efficiency Index enables the awareness of improving energy efficiency and competitiveness in the business. It supports investment decisions related to energy efficiency by allowing direct comparison to competitors from the same sector. Until 2015, the energy efficiency index of 28 sectors are captured and evaluated individually [EEP-15]. The data required for calculating the Energy Efficiency Index (EEI) are based on the basis data containing energy consumption, value creation, and investment provided by the government. They are then combined with data gathered from manufacturing companies through surveys. The survey includes questions about the current and future significance of energy efficiency, implemented and planned investment in energy efficiency, and targeted and planned energy saving [EEP-13].

Figure 2.17 depicts the calculation methodology of Energy Efficiency Index developed by the EEP University Stuttgart. The Energy Efficiency Significance Index (German: *Energieeeffizienzbedeutungsindex* (EBI)), Energy Efficiency Investment Index (EII), and Energy Saving Index (ESI) are calculated by considering both sector specific data collected through survey and statistical all-sector data provided by the government.

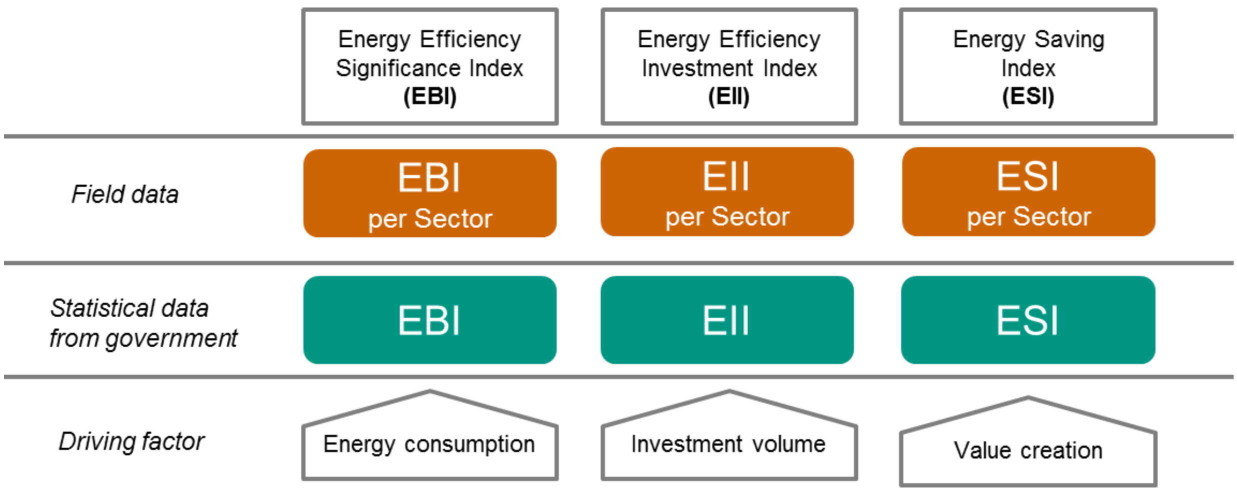


Figure 2.17 Calculation method of Energy Efficiency Index (EEI) developed by EEP University of Stuttgart [UnSt-13]

As seen in the Figure 2.17, the driving factors influencing the indexes are the measurable activities, such as the energy consumption, the investment volume related to the energy efficiency and the value creation. The Energy Efficiency Index (EEI) is computed using equation (2.11).

$$EEI = \sqrt{EII.ESI.EBI} \tag{2.11}$$

The EEI allows the benchmarking of energy efficiency over different companies in the same sector and over different sectors by taking into account the all-sector statistical data from the government. It also considers the investments to improve the energy efficiency. However, using the index, it is not possible to compare the energy performance different processes and organization units in a single manufacturing company.

EnergyStar

U.S. Environment Protection Agency (EPA) developed a benchmarking system within the program frame EnergyStar. The benchmarking system allows a comparison of different industry sectors in the United States. The EnergyStar Energy Performance Indicator (EnPI) measures only the energy efficiency in factory level of companies from certain sectors [Boyd-05][Boyd-06][Boyd-09][BoZh-11]. It is not possible to compare between facilities, process, and production organization within a single company by using EnergyStar EnPI. The EnergyStar EnPI is calculated by first collecting the energy-related data from member companies annually. The annual data collection makes a promptly energy efficiency assessment after changes in production, such as material flow changes or reconfiguration of the factory, impossible. In order to ensure comparability between individual factories, the benchmarking system was developed for various industrial sectors separately. Moreover, only certain manufacturing operations, activities, product components, and other factors that are similar across different companies in a sector are selected. They are considered to be able to represent the sector. The calculation also takes account the climatic condition by evaluating the daily degree days. In the automotive industry, for example, the information about wheelbase of the cars is gathered to determine the different energy requirements of different categories of vehicles [Boyd-05].

The result of the benchmarking calculation is the percentile (0%-100%) of energy efficiency of a factory in compare to the total number of registered factories. Therefore, as an example, the result value of 60% means that 60% of the registered factories in the same sector have worse energy efficiency. Factories having values between 75% and 100% are identified having sufficient energy efficiency. If the value lies under 75%, the company should develop measures to increase the energy efficiency. The companies having a score from 75% deserve to receive Energy Star Certificate.

By using Energy Star EnPI, the company receives a detailed assessment of various energy sources used. It is classified into electricity and non-electricity formed sources. It is a good tool to evaluate the energy efficiency of a factory compared to the most energy efficient in the same sector in the United States. At first glance, it makes clear whether there is potential for improvement and how significant the improvement is required. Nevertheless, the concept is unable to determine cause-effect relationships. It cannot evaluate the energy efficiency on the system level in shorter time interval either.

KURZ.Energiecheck

KURZ.Energiecheck is an internet based tool developed by energy agency the state of North-Rhein -Westphalia on behalf of the state government. It aims to support brief analysis of energy efficiency in companies from certain sectors, such as laundries, breweries, hospitals, and woodworking. The input data of the tool are the energy form, consumption and cost of each energy form, and yearly production and sales values. The tool gives different information as output, for instance, energy consumption per production units, the proportion of energy costs in sales, total energy consumption for heating, and total electrical power consumption. The tool also provides the comparison to other registered companies from the same sector. It offers a fast assessment of a company’s energy performance but the concrete measures are not provided. It has to be done through additional consulting service [ENRW-13].

Internet-Energie-Check

The internet energy check is developed by Bremer-Energie-Konsens GmbH and is launched by the climate protection agency of state Bremen. The aim of the Internet-Energy-Check is to identify the energy savings potential in small and medium enterprises and to suggest improvements. The user can perform a quick and detailed check using the Internet-Energy-Check website. The check is primarily targeted as the start point for handicrafts and service companies to deal with the energy efficiency [BEKG-13].

The check is based on the sector specific information. It requires input data, for instance, the number of employees, the amount of fuel consumption, the yearly electrical power consumption, and the energy costs. The internet portal calculates the energy costs and carbon dioxide emissions that could be reduced by an efficient use of energy and compares them to the average of the companies in the same sector. It also calculates the energy consumption per employee and compares the result with the average of the companies in the same sector. The detailed check considers the different energy forms and focuses on the monetary savings that can be achieved through technology and methodology improvement measures. The measures include pipe insulation improvements, optimizing the heating pumps operation time, cleaning the boiler, and exchanging the heat pumps. This proposed calculation method is very useful if the company wants to estimate the cost savings after procuring or installing new equipment. The calculation is also detailed and comprehensive. However, there is no relation to the values produced or provided by the corresponding energy effort. It is also not possible to assess the energy performance of different organization units.

ODEX

ODEX is energy efficiency indicator developed in Europe with the support from European Union Commission. It is used in the ODYSSEE-MURE project to measure the energy efficiency in three main sectors, i.e. industry, transport, households. It also measures the energy efficiency at end user or final customer level [Ener-10]. The ODYSSEE MURE is a project funded by the European Union in the frame of Intelligent Energy Europe Programme. Energy agencies from 27 EU member states plus Norway and Croatia participate in the project. The objective of the project is to monitor energy efficiency trends and policy measures in Europe [Ener-11]. ODEX is designed to track developments in the industry in terms of energy efficiency, not to compare individual companies or factories. Since the timing is important, the indicator figures are calculated based on a comparison between different years. The result is an index with a value range 0-100. A company having a value of 90, it means an improvement of energy efficiency by 10% has been reached. The values of the individual sectors are weighted according to their share of total energy consumption and offset to obtain an index for the whole economy. ODEX can be used to monitor the development of appropriate national economy compared to last year or a reference year [Ener-10] [GeBo-11].

MEEP

International Energy Agency (IEA) developed Measure(s) of Energy Efficiency Performance (MEEP) in 2008. MEEP comprises four types of figure which can be applied to different objectives as follows [Tana-08]:

1. Thermal energy efficiency of equipment – energy value available for production/operation divided by input energy value,
2. Energy consumption intensity – energy value divided by certain physical value,
3. Absolute amount of energy consumption – energy value, and

4. Diffusion rates of energy efficient facilities/types of equipment.

The first two types correspond to the figures representing energy efficiency and specific energy requirement as mentioned in Section 2.4.1. The third measure is the absolute energy consumption and thus is not a true indicator of energy efficiency, as the considered effort or input has no relation to the benefits or output. Interesting is the fourth type because by using it, it can be estimated how energy efficient a factory or a whole industry is. However, it cannot be applied at the machine and production line level. It requires additional measurement figures to determine the energy-efficient systems.

2.4.3 Comparison and Conclusion

This thesis analyzes and compares the existing energy-related figures and systematic EnPIs which are elaborated in Section 2.4.1 and Section 2.4.2 based on the following criteria:

1. Application possibilities at different levels of organizations and processes within a manufacturing company (see Section 2.1.3). These are for example production facility or workstation level, production line or segment level, and factory level.

2. Application for decision support in production planning. This thesis evaluates whether the figures or EnPIs can be used to benchmark the energy performance of different machines or production lines so that it allows optimal machine or process selections during production planning.

3. Benchmarking of any products, processes, and sectors. The figures and EnPIs are evaluated whether their measurements do not depend on products, processes, and sectors.

4. Consideration of energy-related costs/efforts. This thesis analyzes whether the existing figures and EnPIs consider all possible costs that represent the energy efforts of observed objects, for example, costs caused by peak loads and energy efficiency certifications.

5. Independent of energy forms. It is evaluated whether the figures and EnPIs are independent of energy form, such as electricity, gas, heat, etc.

6. Consideration of internal energy generation and conversion, for example, solar cells, transformators, etc.

Table 2.7 shows the result of the comparison of different figures that are introduced in section 2.4. As seen in the table, all physical and economics figures can be well used to support decision making in production planning. For example, using those figures, it can be determined, which machines have the better performance and should be included in the production line. However, the compared systems have to be a specific, for example, specific machines, HVAC, or lighting. It is difficult to apply the figures to the larger and complex systems such as a whole production line or factory since it involves multidimensional parameters with different measurement methods. Still, the economic figures are interesting due to the employment of a monetary unit because it allows comparison of different factories and units that produce different products. Those figures do not consider costs and investment to make the energy balance more positive, for example, costs of consumed different energy forms or investments for energy generation or conversion facilities. However, the consumption of energy having different forms can be aggregated resulting a single value representing the total energy consumption.

Most of the existing systematic energy performance indicators require general input data and information through surveys or online forms which result in general statements. The EEI from the University of Stuttgart, EnergyStar figures, and KURZ.Energiecheck figures are well applicable for the benchmarking of different manufacturing companies from the same sector. The ODEX is able to compare the energy performance of a single company over time, and Internet-Energy-Check focuses on the evaluation of energy efficiency related investment. Those indicators are not able to perform benchmarking of different organization units or processes within a manufacturing company. However, methods developed for ODEX could be applicable for evaluating any specific process or system over time within a manufacturing company.

All of the systematic energy performance indicators except the ones developed by Internet-Energie-Check and KURZ.Energiecheck, do not consider the costs of different energy forms and costs or investments to improve the energy efficiency. Internet-Energie-Check focuses on the investments, whereas KURZ.Energiecheck takes into account the different energy forms. The MEEP figures can be used for decision support in production planning. The MEEP type 3 can be used to compare different organization units within a company as long as the units give similar output products.

This thesis develops an energy efficiency indicator that takes into account the six criteria as shown in column headers of Table 2.7. The requirements for the concept are derived from the criteria and elaborated in section 3.3.

Table 2.7 Comparison of energy efficiency figures

	Application at different organization level	Application for decision support in production planning	Benchmarking of any products, processes, and sectors	Consideration of energy-related costs/ efforts	Independent of energy forms	Consideration of internal energy generation and conversion
Physical figures						
Energy efficiency η	·	+	−	−	−	−
Utilization ratio $\overline{\eta}$	·	+	−	−	−	−
Specific energy consumption (SEC)	·	+	·	−	·	−
Economic figures						
Energy efficiency [MüEt-09]	+	+	+	−	+	·
Energy intensity (EI)	+	+	+	·	+	−
Systematic performance measurement systems						
Energy Efficiency Index (EEI) from EEP Stuttgart	−	+	−	·	·	−
Energy Star	−	·	−	−	·	−
KURZ.Energiecheck	−	·	·	·	·	−
Internet-Energie-Check	−	·	·	·	−	−
ODEX	−	−	+	−	+	−
MEEP type 1)	−	+	−	−	+	·
MEEP type 2)	·	+	−	−	+	−
MEEP type 3)	+	+	−	−	+	−
MEEP type 4)	·	+	·	−	−	·

2.5 Conclusions

This chapter presents the results of literature studies of different fundamental concepts relating to this work. The basic concepts provide fundamental theories that help to achieve the defined objectives. The fundamental concepts related to production, energy efficiency and management, knowledge management, and energy efficiency indicators are presented briefly with references to the literature researched. In this way, the reader is given an overview of those concepts without going into details. It also helps those readers who are not familiar with the concepts.

After introducing the definitions of the terms manufacturing and production, which are interchangeable in this thesis, this chapter continues with a description of production systems and their common organizations. Production scheduling as one of the activities in production planning and control is also dealt in this chapter. There are several methods to solve production scheduling problems, and they are briefly introduced. According to the literature covering those methods, metaheuristics and hyper-heuristics offer a better flexibility in modeling the problem. However, metaheuristics require very general and unrestrictive models that are difficult to implement for very specific problems. Therefore, hyper-heuristics are more promising to be applied in this thesis, which considers non-standard parameters of production scheduling, such as energy efficiency. Furthermore, hyper-heuristics correspond to real problems better since they allow for customized strategies. Nevertheless, metaheuristics can be still utilized to find the solution in heuristic space [WiPO-13]. An additional interesting feature of hyper-heuristics is their mechanism of learning from feedbacks which permits a dynamic change of the preferred heuristics.

Energy efficiency improvement in manufacturing can be achieved by employing ICT-supported energy management. This chapter highlights the tasks of energy management in manufacturing and the importance of ICT to support those tasks. The work in ICT-supported energy management in production is evaluated, compared, and presented in this chapter. The evaluation is based on the criteria inferred from the research objectives, including the use of the knowledge base with interoperability, intelligent analysis and learning capabilities, energy performance measurement and benchmarking of different organization levels within a company, and energy optimization in production planning and control. The evaluation result is shown in Table 2.1. In some cases, the knowledge base is used to support energy management activities. In most cases, however, learning capabilities to acquire the knowledge are lacking, except for ISES. Previous work focusing on the knowledge-based approach mostly did not include energy performance benchmarking using quantitative indicators. No work using a knowledge base to optimize energy consumption in production planning and control has been carried out to date.

This thesis takes a deeper look at the knowledge management and engineering approaches which are essential in achieving the first objective of this thesis. According to the literature evaluation presented in this chapter, ontology is considered the suitable technology for knowledge representation in energy management, since it allows for a shared understanding of different human and machine stakeholders. Energy management systems involve different organizational and informational infrastructures in the company, including human actors, hardware, and software (see Section 2.2.2). There are different techniques to represent the ontology. According to the literature review, OWL is a language offering a higher degree of expressiveness. It covers both less expressive representations, i.e. RDF and RDFS. It can represent description logics which allow for decidable reasoning. OWL can be integrated with rules. Thus, it provides reasoning capability in the knowledge model. Several related works in manufacturing and energy management are evaluated. It is found that most of them use OWL for knowledge representation. Nevertheless, the application domain of ontology is energy management in either building or manufacturing. No applications of ontology in manufacturing energy management exist. The knowledge is still acquired manually by the experts. This fact is presented in Table 2.2 and Table 2.3. This thesis develops an OWL-represented knowledge base and methods to acquire the knowledge in a semi-automatic way.

The knowledge can be extracted from data by utilizing KDD techniques. This thesis evaluates different techniques potentially applied to acquire knowledge that is useful to support energy management activities in manufacturing. Due to model simplicity and the flexibility of handling different types of data, the decision tree is found to fit best to the purposes of this thesis. Literature review on the application of KDD shows that decision trees are commonly applied to solve problems in either energy management or manufacturing (see Table 2.4 and Table 2.5). Similar to ontology-based knowledge representation, applications in manufacturing energy management are still lacking.

As one of the tasks in EnMS, evaluation requires figures to measure the energy performance of the observed systems. The figures should be simple, straightforward, meaningful, and usable for benchmarking. Different existing figures are evaluated based on different criteria, such as the applicability to different organizational units, comparison of different products, processes, and sectors, and the consideration of energy flow and balance. As shown in Table 2.7, the existing figures do not fulfill those criteria. Therefore, new figures are required which will be developed in this thesis.

Figure 2.18 illustrates the conclusion of the state of the art analysis. As seen in the figure, there are three objectives and several aspects, e.g. knowledge acquisition and representation, considerations of the EnPI development, and PPC which are addressed in this thesis. These are represented as green blocks. In order to address those aspects, analysis of related works and state of the art has been performed (see the blue blocks) and resulted in some approaches and tools that will be applied or developed in this thesis (see the orange blocks). The conclusions are represented as orange trapezoids on the bottom of the figure. Those are the need of an OWL-based knowledge representation for manufacturing energy management, a decision tree based KDD approach for manufacturing energy management, novel intelligent energy performance indicators, and a hyper-heuristic based production scheduling approach.

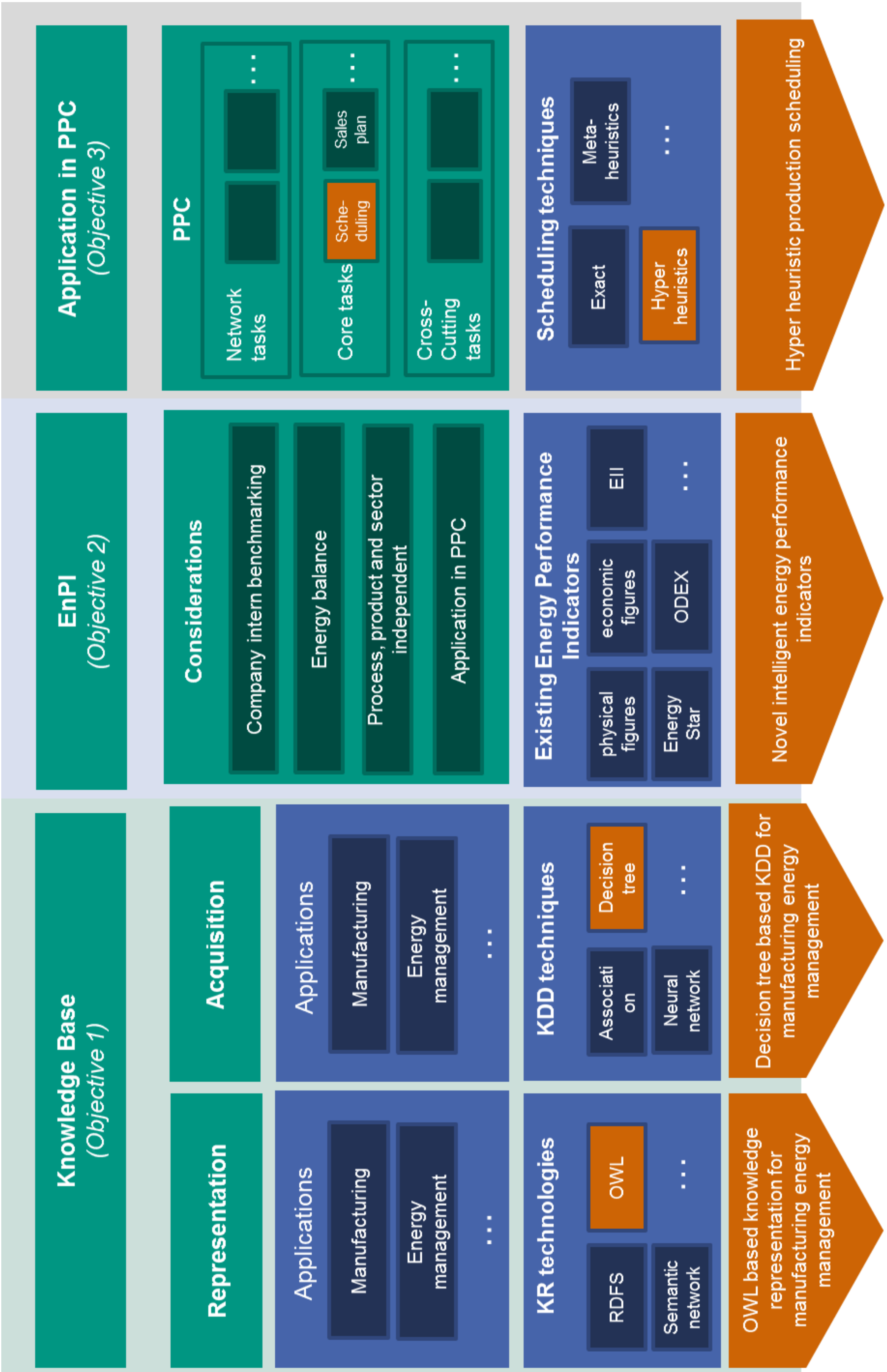


Figure 2.18 Conclusion of state of the art analysis

3 Requirements for the Concepts

"Let our advance worrying become advance thinking and planning."
~Winston Churchill (1874-1965)

As described in Chapter 1, this thesis aims at developing an integrated ICT-based approach to supporting the holistic evaluation and optimization of energy efficiency in production. There are three research objectives to achieve this goal. The first research objective is to develop a concept to support energy efficiency evaluation using a knowledge base. By using a knowledge base, qualitative evaluations, such as ordinal and verbal argumentative evaluations, are possible. The second research objective is to develop an indicator to support the quantitative evaluation of energy efficiency. Furthermore, the developed approach should be applied to production planning and control that is the third research objective.

This chapter outlines details of the requirements which have to be fulfilled by the concepts developed in this thesis. The requirements are derived from the research objectives, fundamental concepts, and state of the art analysis (see Section 2.5). Firstly, the requirements are to itemize the research objectives, but they also are to address the weaknesses and deficiencies of existing approaches, as shown in chapter 2. The requirements can be divided into four groups as seen in Figure 3.1. They are explained in detail in the following sections.

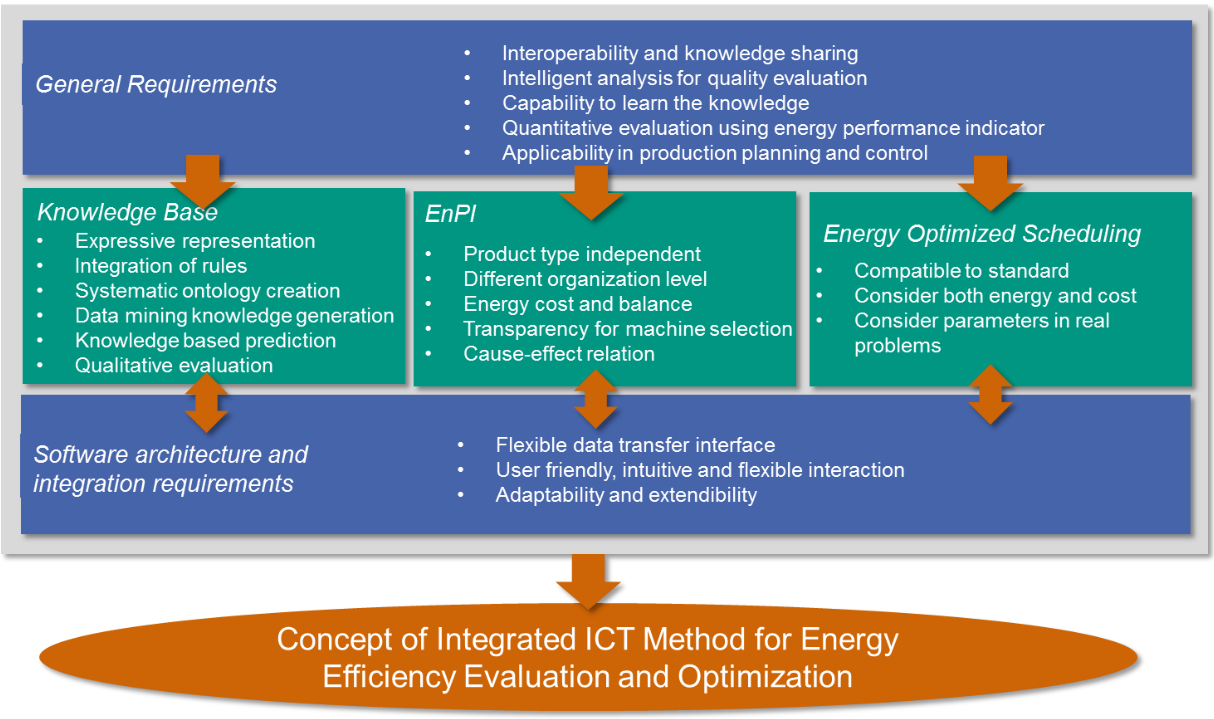


Figure 3.1 Requirements for the concepts

3.1 General Requirements of ICT Supported Energy Management

In Section 2.2.6, the different approaches related to ICT-supported energy management are described. The different criteria to evaluate the different approaches are also developed. They are summarized in Table 2.1. A set of requirements can be derived from the developed criteria and the evaluation results. This section presents the formulation of these requirements.

Requirement R1.1:

This thesis should enable interoperability and knowledge sharing among stakeholders involved in energy management systems in a company. As described in Section 2.2, the energy management system involves both organizational and informational infrastructures. Interoperability and knowledge sharing are essential to accelerate the achievement of the energy management systems' objectives. To allow interoperability and knowledge sharing between stakeholders, it requires a knowledge base that uses a knowledge representation which defines a consensus and makes the knowledge explicit. The requirements of this kind of knowledge base are further described in Section 3.2.

Requirement R1.2:

The knowledge base contains human knowledge and also best practice in energy management. It is used as a reference to evaluate the energy performance. The knowledge base should allow more intelligent and efficient analysis and evaluation of energy performance without an extensive examination of raw data. The knowledge should be expressed in such a way that a deduction and inference can be performed in order to draw a conclusion based on the stored knowledge. The requirement is considered as part of requirements for the knowledge base and expressed as Requirement R2.2

Requirement R1.3:

As presented in Section 2.2.6, the learning process can accelerate knowledge generation process. This is essential for building the knowledge base since it normally involves extensive tasks performed by an expert. Some knowledge that cannot be collected directly from experts or other human sources is not covered in the knowledge base. Some of the knowledge depend strongly on a specific company or a specific case. Hence, it is difficult to acquire this kind of knowledge. Thus, this thesis should develop an approach that able to learn the knowledge from the company or case specific data. The requirement is also considered as a requirement for the development of the knowledge base. After performing analysis of fundamental concept and state of the art, it can be derived to Requirement R2.4.

Requirement R1.4:

In performing evaluation during check phase of energy management system, quantitative information is useful to measure the current energy performance of a system or organization within the company. The quantitative information can also be used to benchmark different organizational or system entities in the company. This thesis should develop an approach that allows an intelligent energy performance measurement and benchmarking to compare different organization and resource in production. For this reason, this thesis develops an Energy Performance Indicator (EnPI). The requirements of EnPI are presented in Section 3.3.

Requirement R1.5:

As mentioned in Section 2.2.5 point 4, ICT can help to improve energy efficiency and manufacturing productivity through optimized production planning and control activities, especially through energy optimized scheduling. This thesis should develop a scheduling approach that considers energy consumption and costs optimization, for example, peak load avoidance. This requirement is broken down into detailed requirements described in Section 3.4.

3.2 Requirements for the Knowledge Base to Support the Qualitative Evaluation

As mentioned in the first objective of Section 1.2, the qualitative evaluation of energy efficiency in the manufacturing company can be done by using a knowledge base. The knowledge base makes the knowledge from different stakeholders explicit so that it is accessible as the reference to perform the evaluation. This section details the requirements for the knowledge base that is used to support the energy efficiency qualitative evaluation.

Requirement R2.1:

The knowledge base should use ontology as the representation. Literature analysis performed in this thesis has resulted that ontologies have been used to address the interoperability problems. The state of the art analysis has shown that there are different languages to express ontologies, such as RDF, RDFS, and OWL (see Section 2.3.3). Due to expressiveness provided by OWL, this thesis should use it to represent the ontology knowledge base.

Requirement R2.2:

As described in Section 2.3.3, reasoning capability is one of the essential properties of knowledge representation. The ontology represented in OWL allows reasoning. The reasoning is performed by a reasoning or rule engine. Therefore, the ontology should be equipped with a set of rules. In this thesis, SWRL is used as the language to represents the rules due to its compatibility with OWL ontologies.

Requirement R2.3:

The methodology to create the knowledge base is essential in order to get the high-quality content of knowledge and to have a knowledge base that is capable of solving the problems in desired way. There are different methods to create an ontology knowledge base in a systematic way. To choose the right method to create the ontology knowledge base, different criteria should be considered, for example applicability, extendibility, and the expertise requirements. The method should allow an efficient creation of the ontology.

Requirement R2.4:

Manual creation of knowledge base is an extensive task. Most of the knowledge come from the human knowledge that has to be modeled in the knowledge base. There is also knowledge that is hidden in different data and non-transparent for the human stakeholders. This thesis should allow an acquisition or an extraction of the non-transparent knowledge from different data in the company. An analysis of basic methods and state of the art described that this kind of knowledge can be acquired using knowledge discovery in database (KDD) (see Section 2.3.5). This thesis should use the data mining to generates the rules and relationships from different data related to energy efficiency semi-automatically. The result is then should be integrated into the ontology knowledge base in order to allow an incorporation of the new knowledge extracted from data and human knowledge that is already explicitly represented in the knowledge base.

Requirement R2.5

One of an essential task in energy management after performing evaluation is the prediction of energy consumption, in order to be able to perform accurate actions to improve the energy performance. Knowledge resulted from KDD on historical data can be utilized to predict the energy performance, i.e. energy consumption of products, processes, and resources. This thesis applies KDD to predict power consumption and duration of manufacturing processes based on the product properties. Thus, all necessary measures can be taken more accurately to improve the energy efficiency.

Requirement R2.6

The knowledge base should support the qualitative evaluation of energy efficiency degree in the manufacturing system. By consulting the knowledge base, it should be possible to identify the energy efficiency states of the objects, for instance, processes, machines, facilities, in the manufacturing. Additionally, the knowledge base should assist the detection of energy inconsistent performance or wasting. The evaluation should follow the criteria that correspond to the energy efficiency best practices.

3.3 Requirements for the Energy Performance Indicators

The concept of energy performance indicator (EnPI) should enable a holistic quantitative evaluation of energy efficiency in the manufacturing. As shown in Table 2.7, the existing energy efficiency indicators as the basis for the quantitative evaluation still have some weakness. Hence, this section formulates requirements to overcome them.

Requirement R3.1:

The EnPI should be applicable independent of the product types, variation or production process. The EnPI should also be able to compare different processes that produce the same products from different raw materials or intermediate products. It should not depend on the kind of units e.g. pieces, meters, kilograms, etc. outputted by the processes.

Requirement R3.2:

The developed EnPI should be able to measure the energy efficiency of different levels of production organization. The organization levels differ in their system boundary and thus to be considered in determining the influence factors on energy efficiency as well as in defining the energy flow model. The EnPI should measure energy efficiency at least on following levels:

- Machine, production facility, or workstation level. The EnPI should have the functionality to evaluate and to compare the energy efficiency of individual machines or production facilities.
- Production line level. The energy efficiency of the whole production line consisting of several machines or facilities should be evaluated by the developed EnPI. By comparing various machines, combinations and selections can be made.
- Factory level. At the factory level, the EnPI has to be able to support the measurement for selecting the most energy efficient factory location. In addition to the measurement of energy efficiency of individual factories, the various transport routes have to be considered.

Requirement R3.3:

The EnPI should not only consider the energy form. It should primarily take into account the energy-related costs. These include not only the costs of various energy sources themselves but also other energy costs, for instances energy peak load costs and costs for acquiring emission certificates. In this thesis, not only the energy costs coming from external sources or utility company are regarded. The energy costs and effort within the system boundary should be calculated and considered in the EnPI calculation. This is, for example, the internal energy generator and distribution network.

Requirement R3.4:

The EnPI should be able to support the production planning activities in evaluating the machinery and equipment in terms of energy efficiency. It should enable a systematic and transparent selection of the used machines. It should be used as a tool to assess and to simulate whether investment for new machines could improve the energy efficiencies in the company.

Requirement R3.5:

The EnPI should follow the principle of energy efficiency. It should describe the correlation between the output produced by the production system and the energy effort. Production system having higher EnPI should mean that it has better efficiency than the other. Therefore, other than relationship figure described in Section 2.4.1, the numerator should represent the work or output produced by the system, and the denominator should be the energy effort needed to produce the work or output.

3.4 Requirements for the Energy Optimized Production Scheduling

Requirement R4.1

There are many types of production scheduling problems, starting from simple one machine problems to complex environments involving different parameters. It is essential to describe a production scheduling problem in a simple and standardized representation. The approach developed in this thesis should follow a schema that is already well established, for instance, the approach developed by Blazewicz et al. [BlEt-07].

Requirements R4.2

The scheduling should give the solutions for real problems on the shop floor. Mass production companies, for instance, screw or small electronics component manufacturers, produce many items at once. A single operation could produce certain numbers of products called lot. Furthermore, in real shop floor, a machine could perform different operations. It needs to be retooled before it executes a different operation. A machine could be operated by personnel in different shifts, for example, two shifts a day and three shifts a day. The scheduling model should also consider that an operation could have different outputs. It should also take into account that a product can be produced by different kinds of operation from different machines.

Requirement R4.3

The optimization objectives should consider both cost and energy efficiency. The orders have to be finished before the given deadlines. The multiple optimization objectives should be weighted based on their importance.

3.5 Requirements for the Software Architecture and Implementation

The developed concepts the have to be feasible to be applied in the real world later on. Although this thesis develops only a software prototype to validate the concept, the software has to be designed by considering the typical system environment in manufacturing companies. Furthermore, the usability of the software system should be considered since the software should support the human stakeholders in implementing the energy management system in manufacturing. It is also essential to take into account the expandability of the software, due to the dynamic environment of IT systems in manufacturing companies. This section lists the requirements related to the software architecture and its integration with the IT environment.

Requirement R5.1

In order to make the developed concept applicable to the real manufacturing companies, the design of the software prototype should allow a simple and flexible integration to the typical IT environment of manufacturing companies, for example, PPC, ERP, MES and industrial automation systems. Therefore, a flexible interface serving the data transfer between the developed software and those systems has to be developed.

Requirement R5.2

The software will be used by different users since it supports the energy management activities in manufacturing involving different human stakeholders from manager to the operator level. The usability of the developed software should be ensured for the different types of user. The software should provide a user-friendly, intuitive, and flexible interaction interface with the user. It should not require any extensive user training in order to operate the software.

Requirement R5.3

Due to the complex problems faced in product development and manufacturing, the software technologies for manufacturing grow very rapidly. Several times in a year software technologies in the manufacturing could be newly released or updated. As a consequence, the software should be designed to address the problem. Much effort for adaptation, changes, and additional extensions in deploying and utilizing the software in the IT system environment of the companies should be avoided. A modular software architecture would be a possible approach for addressing the problem.

4 The Concepts of Integrated ICT Method for Energy Efficiency Evaluation and Optimization

"Research is creating new knowledge."
~Neil Armstrong (1930-2012)

The chapter describes the concepts developed to address the requirements formulated in Chapter 3. It begins with the overview of the approach which describes the main components of the solutions and how the components interact with each other. Each component is explained in more details in the next sections. The development of ontology knowledge base for manufacturing energy management is elaborated in the following chapter. It gives details on the development of knowledge model and rules, and how they are represented in an ontology. It continues with the description of the learning based approach to generate the knowledge semi-automatically. The integration of the results in ontology knowledge based is also introduced. The following chapter also features the development of energy performance indicators (EnPI) to achieve the second objective stated in Section 1.2. The EnPI development takes into account the requirements listed in Section 3.3. Finally, this chapter ends with the description of the approach for developing production process scheduling which considers the efficient energy consumption by utilizing the knowledge base. The approach should fulfill the requirements described in Section 3.4.

4.1 Overview of the Solution Approach

Based on the requirements stated in Chapter 3, four concepts are to be developed in this thesis. These concepts cover (i) the ontology knowledge base to address interoperability and knowledge sharing problems and to simultaneously ensure intelligent qualitative evaluation of energy efficiency; (ii) the KDD, including data mining, for the automatic generation of knowledge; (iii) the energy performance indicator for the quantitative evaluation of energy efficiency; (iv) the module for handling production process scheduling. The overview of the solution approach is depicted in Figure 4.1.

As seen in Figure 4.1, this thesis proposes an approach consisting of tools for both quantitative and qualitative energy efficiency evaluations to support the implementation of energy management in manufacturing companies. The company-specific data originate from the ICT systems deployed at the company (see Figure 4.1, no. 1). Data related to products, production processes, and resources are collected from different companies' software systems. For example, the data on resources, e.g. machines, human resources, etc. can be retrieved from the Enterprise Resource Planning (ERP) System, whereas the production processes including the planned operations and jobs that are executed at shop floors, can be collected from the Manufacturing Execution System (MES). The measurement data, for example, energy consumption, temperatures, light intensities, and other ambient factors are acquired by industrial automation systems including manufacturing automation systems (MAS) and building automation systems (BAS) (see Figure 4.1, no. 2). Those systems have different data communication interfaces, for example, CSV exports, web services, and SQL database exports. The data collection interface collects those data and then transforms, aggregates, and processes them to a certain representation. This thesis uses a relational (SQL) representation of data. The relational representation allows the knowledge generator (data mining) to process the data directly (see Figure 4.1, no. 3). The data collection interface is also responsible for the ontology population (see Figure 4.1, no. 5 and section 4.2.2). Ontology population means the process to create instances or individuals and to set their property values.

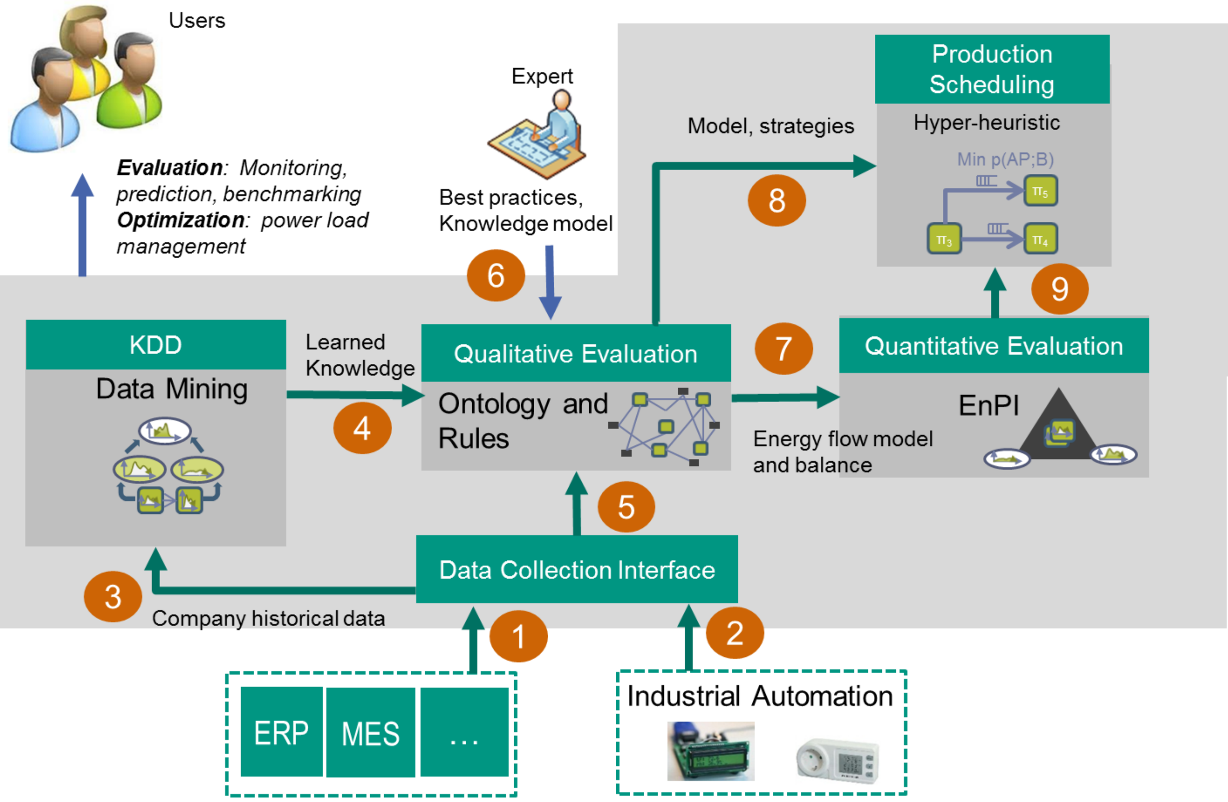


Figure 4.1 Overview of the solution approach

The knowledge generator extracts the meaningful relations between products, production processes, and resources, ambient parameters (e.g. temperatures), and energy consumptions. This is done semi-automatically by executing data mining algorithms. Rules representing the energy usage patterns are deduced from the data mining result. The relations may have the form of rules or equations. Rules allow for the identification of energy usage anomalies or inefficiencies. The rules are then transformed into Semantic Web Rule Language (SWRL), an OWL-compatible rule format. Afterward, they are integrated into the ontology knowledge base (Figure 4.1, no. 4). Section 4.2.3 gives the detailed description of the data mining development as a part of the Knowledge Discovery in Database (KDD) technique.

A further usage of the knowledge base is for energy efficiency qualitative evaluations. A qualitative evaluation focuses on the criteria that can be observed but not measured in figures, such as energy-wasting situations, or energy-efficient best practices. Qualitative evaluation is performed by employing a knowledge base which is accessible by all of the relevant staffs and systems in the company. The best practices and knowledge about energy efficient and energy wasting activities are standardized, formalized, and stored in the knowledge base. It is implemented as an ontology knowledge base combined with rules [WiEt-12].

The goal of the ontology development is to generate a holistic model of the energy management system that provides a shared vocabularies and knowledge structures, as well as avoids prevents ambiguities. This aims to address the interoperability and knowledge sharing problems among human and machine stakeholders. An effective qualitative evaluation of energy efficiency involves different systems and organizational units and requires a common understanding between them. The ontology should allow relating products, processes, resources as the main elements of a manufacturing system and the entities associated with energy, for instances, industrial automation for measuring and controlling energy consumption, building elements as energy consumers, and also human actors who affect the energy consumption in the manufacturing.

The knowledge model in the ontology is created by an expert in manufacturing energy management (see Figure 4.1 no 6). The references for performing qualitative evaluation are the states modeled in the ontology. The ontology allows for an automatic evaluation of the current states in manufacturing against the reference states. By using the ontology, it is possible to identify and to check the states in manufacturing, to which energy-efficient management measures can be applied. The entities involved in the applied measures are to be represented in the model. The energy-efficient states are expressed in Boolean representation which allows for judgment as to whether the value is true or false. A complex energy-efficient or energy-wasting state is modeled through rules that can be evaluated after the values of the properties involved in the rules have been compared with values measured by sensors of industrial automation systems or values from the IT systems. After identifying the energy-related states, the applicable techniques are described. In general, all techniques described in the European Reference Document cited in Section 2.2.4 belong to one of the following three types:

- Operational techniques which require process changes. For example, a technique to reduce peaks in energy consumption might be by not using two highly energy-intensive machines at the same time. Another possible operational technique is to perform maintenance checks more regularly.

- Design techniques that can be applied by adding, removing, or changing parts of the system. A typical example is to use more energy efficient machines if the lifetime cost benefit is positive.

- Configuration techniques that change the parameters of a certain machine or system, for example, changing the pressure of a compressed air system.

Ontology can be used to identify potentials for machine changes and the ideal type of machine to choose. The data relating to energy consumption have to be measured by measurement instruments that are part of the industrial automation system (see Figure 4.1). It is possible to introduce a rule to express the parameter optimization of a machine of a system, for example, the rule that expresses "the ideal RPM for a drilling machine processing a certain material type has to be a value between x and y." The optimization parameters might depend on multiple factors that can lead to a high complexity of the rule created. Ontology can also help to optimize the cycle for carrying out maintenance tasks. For example, it can formulate a rule that the best period to check an air compressor is every few months. A complex installation task involving process integration with different parameters and stakeholders can be modeled in ontology, although it is very extensive and requires special expertise. Energy efficiency management techniques on the processes or equipment level can normally be expressed by the ontology. The rules originating from data mining are used to detect the wasting of energy and energy usage anomalies. The system learns from historical data and identifies the usual energy consumption that depends on different parameters, such as product and process characteristics as well as ambient parameters. The method to develop the ontology is explained in more detail in Section 4.2.

A quantitative evaluation assesses energy efficiency using measurement figures called Energy Performance Indicators (EnPI). It is possible to determine the degree of energy efficiency in different organizational parts or processes of the company. In order to calculate the EnPI, a calculation model is developed that considers the energy flows and energy balances in the company. The values assigned to the calculation model are retrieved from the data collection interface. The entities and their relations considered in the calculation model are modeled in the ontology (see Figure 4.1 no 7). By using the developed EnPI, production planners or facility managers are able to simulate and compare different production and facility configurations in order to find the most energy-efficient ones. The details of EnPI development are presented in Section 4.4.

Production scheduling also uses the models representing the entities involved from the ontology (see Figure 4.1 no 8). In addition, it applies the EnPI to measure the energy efficiency of a schedule, consisting of a set of jobs that will be executed on the shop floor (see Figure 4.1 no 9).). It helps the production planners to perform production process planning that considers not only the costs but also energy efficiency. The approach uses a hyper-heuristic algorithm for finding the optimal production schedule that considers strategies tailored to the real energy efficiency problems in manufacturing. Section 4.5 outlines the development of the module.

4.2 The Manufacturing Energy Management Ontological Knowledge Base

As described in Section 2.3.4, OWL is an ontology representation language that offers a high degree of expressiveness. Therefore, it is widely used by different academia and industries to solve the problems in information and knowledge management in manufacturing and energy management. This thesis uses OWL to explicate the knowledge in a knowledge model for manufacturing energy management.

The knowledge model comprises classes, attributes, and relation definitions related to automation devices for instance sensors, energy meters, building environments, production facilities, products and resources; are created manually by experts. The ontology containing these hand-crafted elements builds the knowledge base corresponding to the manufacturing energy management domain knowledge. The ontology contains only the classes or TBox elements that describe the knowledge structurally and terminologically. It provides a common conceptual vocabulary in the manufacturing energy management domain. It does not contain any ontological individuals or ABox elements [WiRO-12].

The domain knowledge represents the model of a manufacturing energy management system. Therefore, it can be used and extended by any manufacturing companies. It does not contain any instance-specific information. A set of SWRL rules representing common practices of energy management in manufacturing is integrated into the domain knowledge. SWRL is a markup language that combines OWL and RuleML (Rule Markup Language) and enables the integration of rules in OWL ontology [HoEt-04].

The domain knowledge is then enriched with company-specific rules. Some rules are created by the knowledge engineers in the company. Other rules are created semi-automatically by executing data mining algorithms (see Section 4.2.3). The algorithms generate classification rules and equations corresponding to the knowledge that is extracted from the collection of energy, production, and energy-related infrastructure data. The rules are converted to SWRL format and integrated into the ontology [WiEt-14].

The Abox knowledge elements are created semi-automatically based on factory configuration and layout. For this purpose, a method for interpreting the semantic information from building and factory construction drawings has been developed [WiRo-10] [WiAR-12] [HäEt-13] (see Section 4.3.3). This will result in an instance of the manufacturing energy management domain ontology containing company-specific knowledge elements. Figure 4.2 depicts the overview of the approach for the knowledge base development.

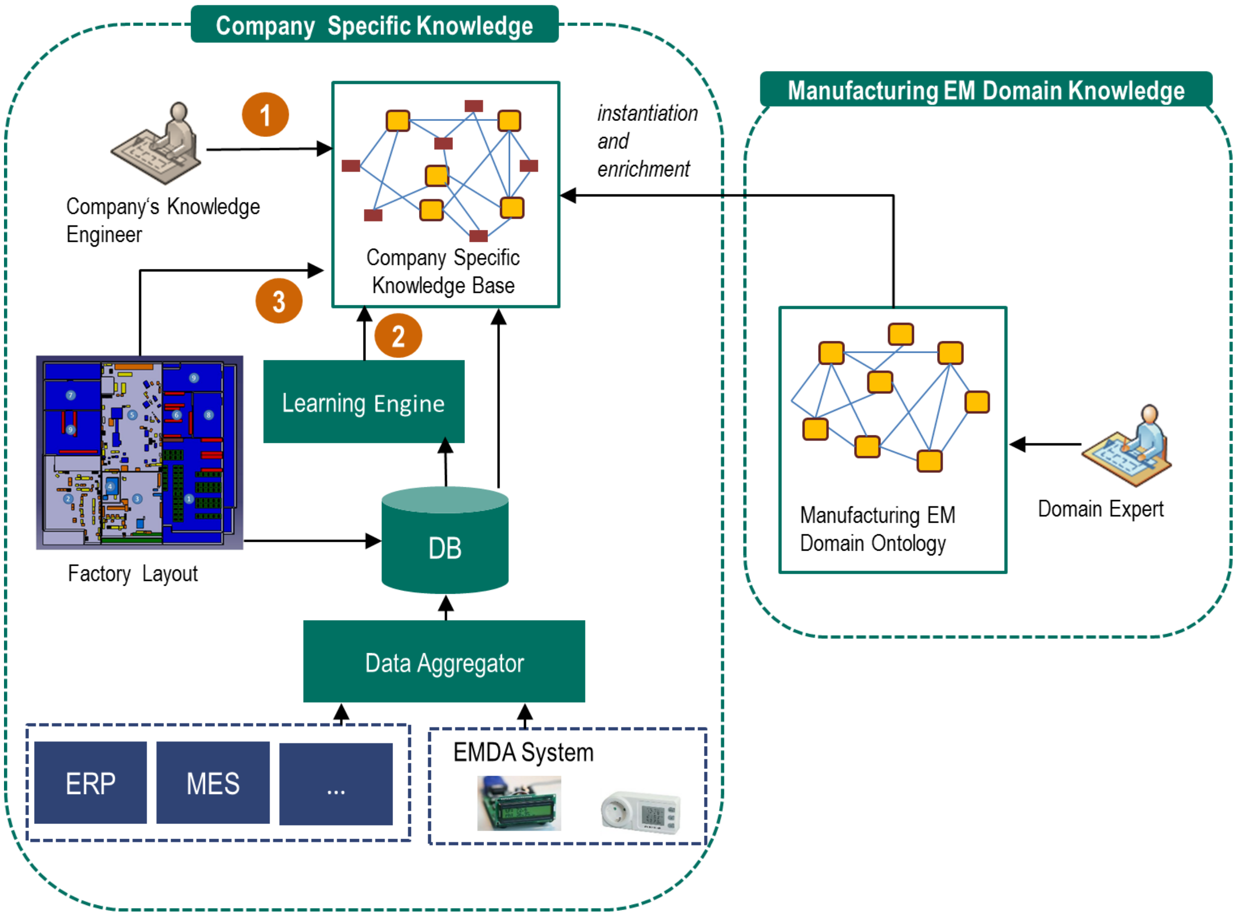


Figure 4.2 Methods to generate knowledge, published in [WiEt-14]

4.2.1 Manufacturing Energy Management Domain Ontology

Methodology Selection to Develop the Ontology

Before developing the manufacturing energy management domain ontology, existing methodologies of ontology development as presented in Section 2.3.4 are analyzed and evaluated. For this evaluation, several criteria are used which take into account the objective of ontology development. The following criteria are considered in evaluating the existing ontology development methods:

1. Applicability and extendibility of the resulted ontology
2. Expertise requirements
3. Time required for implementation
4. Manageability of changes
5. Verification possibility to evaluate correctness and completeness
6. Degree of the dissemination of the methodology

7. Support of ontology development tools

The development of the manufacturing energy management domain ontology is aimed at creating a generic ontology, containing TBox knowledge, i.e. concepts, relation definitions, and common rules that can be applied to any manufacturing company. The ontology is then extended and populated with company-specific information. For this purpose, the ontology has to be easy to extend to meet the company-specific requirements. The ontology should also be applicable to real scenarios in manufacturing energy management. To apply and to extend the ontology to a company, the company staff is needed to incorporate the ontology to meet the company's requirements and to make extensions, if necessary. An extension might be adding classes representing company-specific processes or adding new relations of new energy generation technologies to a specific machine type. For this, the ontology development methodology is applied. The chosen methodology is to provide a clear description and the minimum required skills to develop the ontology should be as small as possible. In order to measure these skills, the development process is described, and the necessary skills are deduced.

The development of the ontology should take a minimum period of time. The amount of time needed is estimated based on the number of required steps to implement the methodology. Furthermore, the strictness of definition of each step and the complexity of the definitions are important. By considering the dynamic environment of modern manufacturing, the chosen methodology should also be able to deal with changes. The chosen methodology should allow the user to perform the changes in the corresponding steps so that it is not necessary to restart the methodology from the beginning. Process description reveals how the chosen methodology manages changes. The dissemination degree of a methodology also is an important criterion to determine the quality of the methodology. If more projects and applications use ontologies created with a methodology, this means that the methodology is applicable to solving real world problems. For this criterion, the projects and applications using ontologies in the area of manufacturing and energy management are examined. The more different disciplines are covered by the methodology, the more universally applicable it is.

The next criterion is the possibility to verify correctness and completeness of the resulting ontology. After ontology creation is finished, it should be verified against the requirements defined before the development process. An evaluation step in the methodology should ensure this. The evaluation step should also be able to check the correctness and completeness after extensions or changes. The methodology step description is examined to evaluate the criterion.

The last criterion is how the methodology is supported by widely used ontology development tools such as Protégé[1], Neon[2], KAON[3], etc. Since the development of a complex ontology

[1] http://protege.stanford.edu/

[2] http://neon-toolkit.org/

[3] http://kaon2.semanticweb.org/

involves multiple actors with different backgrounds, the support of widely ontology development tool is very essential. This will allow for a collaborative development of an ontology by different actors. For example, the development of a manufacturing ontology requires collaboration of production engineers, automation engineers having knowledge about industrial automation, and actors who have knowledge on building management.

The methodologies presented in Section 2.3.4 are evaluated on the basis of the criteria mentioned above. The evaluation results are presented in Table 4.1. For each criterion, the methodology is rated positive (green), neutral (yellow), and negative (red). Positive means that the methodology fulfills the requirements of the corresponding criterion. Neutral implies that the requirements of the corresponding criterion are not fulfilled completely by the methodology, but that it is in line with them. Negative indicates that the methodology does not satisfy the requirements of the corresponding criterion, or it complies with the requirements to a small extent only.

As shown in Table 4.1, the methodology developed by Noy and McGuinness seems to fulfill most of the developed criteria. However, it is rated lower than the Enterprise Model Approach and equal to METHONTOLOGY when considering the expertise required in the ontology development. Some expertise is needed to analyze existing ontologies in the corresponding domain for their later integration into the developed ontology. Furthermore, TOVE has a better rating than Noy and McGuinness' approach regarding the verification possibilities. The formulated competency questions allow for a systematic verification, as to whether the ontology can be used to answer the questions. In this thesis, the ontology development is strongly based on Noy and McGuinness' methodology and on METHONTOLOGY, but it also uses a step included in the TOVE methodology.

Table 4.1 Comparison of ontology development methodologies

Criterion	TOVE	Enterprise Model Approach	METHONTOLOGY	Noy and McGuinness
applicability and extendibility	+	·	+	+
expertise requirement	-	+	·	·
required time for implementation	-	+	+	+
manageability of changes	·	-	+	+
dissemination degree	·	+	+	+
verification possibility	+	-	·	·
supports of ontology development tools	·	-	·	+

Specification of the Ontology

The objective of the specification step is to determine the purpose, the scope and the formality level of the ontology. The purpose gives the statements of the intended uses, the users and the usage scenarios. To classify the level of formality, they refer to the classification of Uschold et al. Finally, the scope of the ontology is determined, by defining the terms, their characteristics and granularity [UsGr-96] [FeGJ-97].

The purpose of the ontology development in this thesis is to create an ontology that is used as a knowledge base to support energy management tasks in manufacturing. The ontology should be applied in any manufacturing companies. It should also allow the identification and evaluation of current states of manufacturing in order to detect the energy saving potentials.

The common use case of using the ontology is depicted in Figure 4.3. There are two main roles of users. The first one is the expert. The expert formulates energy management best practices for manufacturing in the ontology. He can also add concepts and extends the ontology to fulfill specific requirements of energy management in the corresponding manufacturing company, for instances, adding new machine types, new production processes, etc. For this reason, the developed ontology has to be extendible. The second one is the manufacturing end users, for example, the production planners, the operational employee, etc., who queries the ontology to solve their problems. Some typical use case scenarios are for example querying the ontology to evaluate the current states of the manufacturing, to detect the energy saving potentials, and to detect anomalies based on the evaluated manufacturing states. The energy saving potentials and anomalies are detected based on the relations of different states of the manufacturing related facilities.

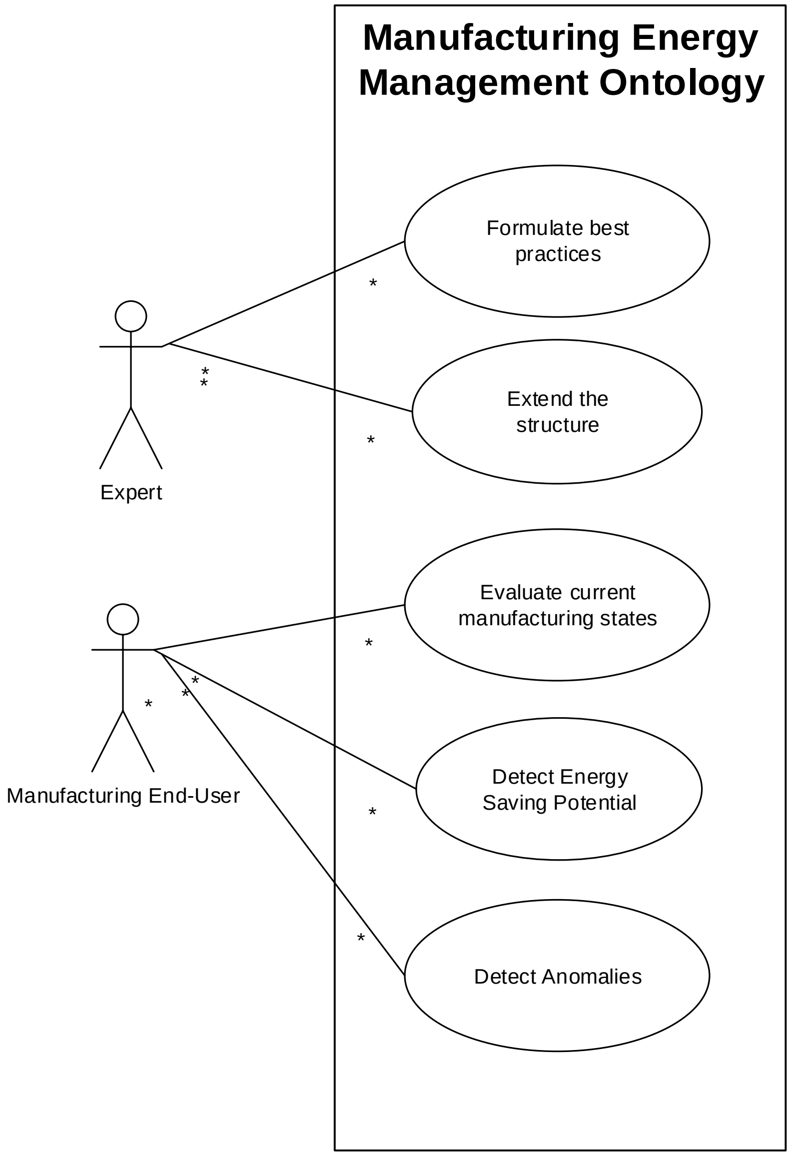


Figure 4.3 Use case diagram depicting users' interaction with the ontology [WiEt-14]

To address the intended use case scenarios, the ontology should be presented using a formal language in order to be understandable for both humans and machines or computers. By using a formal representation, it is more feasible to provide a formal proof to verify the correctness of the ontology. Ontology represented by OWL (see Section 2.3.4) follows the open world assumption. Therefore, it is not necessary for the ontology to model the domain completely. As a result, the ontology is designed in a semi-formal way.

The manufacturing energy management domain ontology has to be able to represent different entities and resources in manufacturing. Additionally, it has to contain classes representing supporting systems, such as energy conversion or energy supply facilities, transport facilities, HVAC, lighting, etc. These should be represented as upper classes of the ontology. The states of the manufacturing systems also have to be expressed in the ontology. The actual states of manufacturing processes, resources, and environment are collected using sensors that are connected to the installed automation system. Thus, the ontology needs to contain classes representing the industrial automation system. Evaluating the state of manufacturing systems give statement whether they are energy efficient or inefficient. Therefore, the ontology must define classes for describing the (in) efficiency and their reasons.

Reusing the Existing Ontologies

The existing ontologies in the manufacturing domain are examined before starting creating new ontologies. This aims to identify whether the structures, concepts, and relations can be reused. Furthermore, the examination of existing ontology helps to improve the familiarity to the domain and to understand the well-established ideas in the domain. In this thesis, there are three ontologies that are reviewed from the literature. They are ADACOR, MASON, and MSE. These ontologies are the most popular ones and often used or referenced to solve problems in manufacturing domain.

ADACOR (Adaptive Holonic Control Architecture for Distributed Manufacturing Systems) is an ontology for distributed manufacturing and is based on the Holonic Manufacturing Systems (HMS) paradigm which is inspired by living organism and social organizations. A holon means a part of a whole and a whole consisting of parts at the same time that forms a so-called holarchy of manufacturing. A holon is autonomous and can be physical or abstract. For example, a machine, a transporter, a manufacturing employee, or an order might be seen as a holon. All things that are considered as holon in manufacturing environment should be included in ADACOR ontology [BoLe-04] [Buss-98]. The ADACOR ontology is represented in object oriented frame-based manner recommended bay FIPA (Foundation for Intelligent Physical Agents). It introduces classes that represent products, processes, resources, and disturbances (see Figure 4.4). The class `product` represents the manufactured products. `ProcessPlan` and `Operation` describe the tasks in processing customer orders in the shop floor. The manufacturing `resources` include the transporter, production tools, and machines that execute the operations. `Disturbance` indicates the unexpected events on resources that degrade the planned processes [BoLe-04].

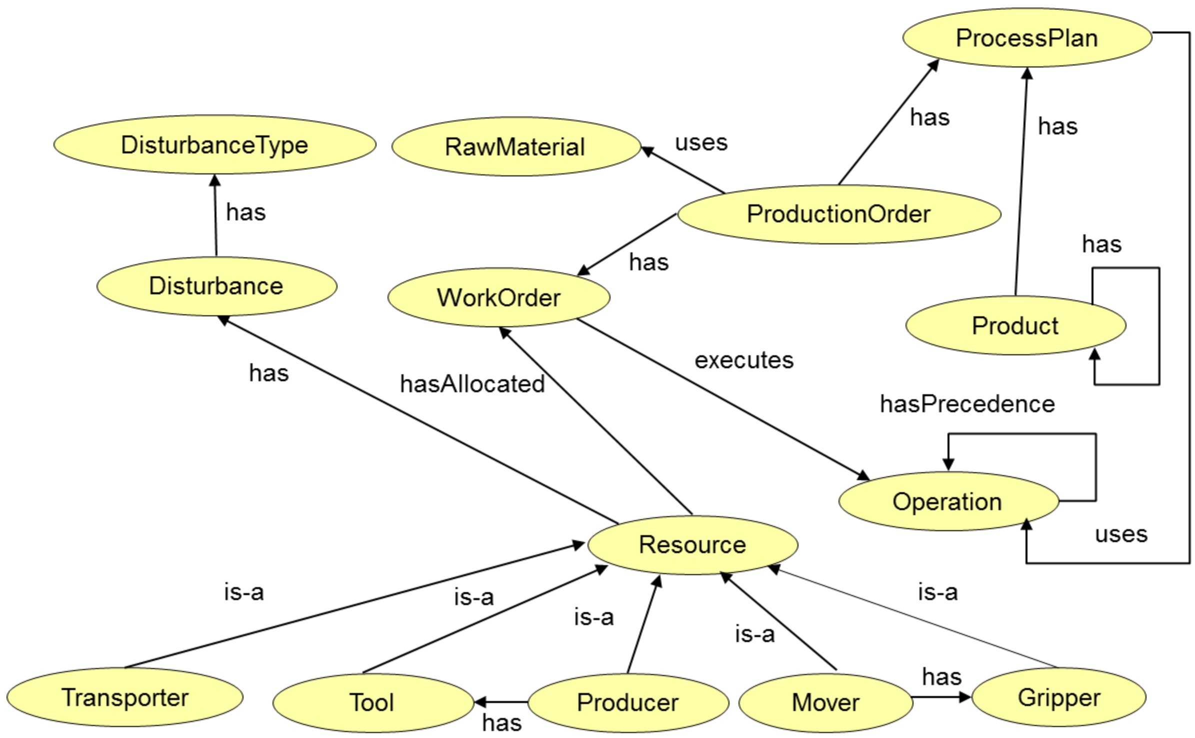


Figure 4.4 ADACOR ontology [BoLe-04]

Another relevant existing ontology is MASON (MAnufacturing Semantics ONtology). It is a manufacturing upper ontology that contains general concepts or upper classes in manufacturing domain. It can be extended to fulfill manufacturing specific needs. MASON contains three upmost classes, i.e. `Entity`, `Operation`, and `Resource` (see Figure 4.5). `Entity` is a common helper concept that specifies the products including properties and constraints. `Geometric`, `cost` and `material` are for example sub-concepts of `Entity`. `Operation` represents the manufacturing related activities, such as machining operation, logistic operation, human operations, and launching operations. `Resource` includes the possible manufacturing resources, for example, machines tools, personnel, as well as manufacturing locations, e.g. factory, workstation, etc. [SiEt-06].

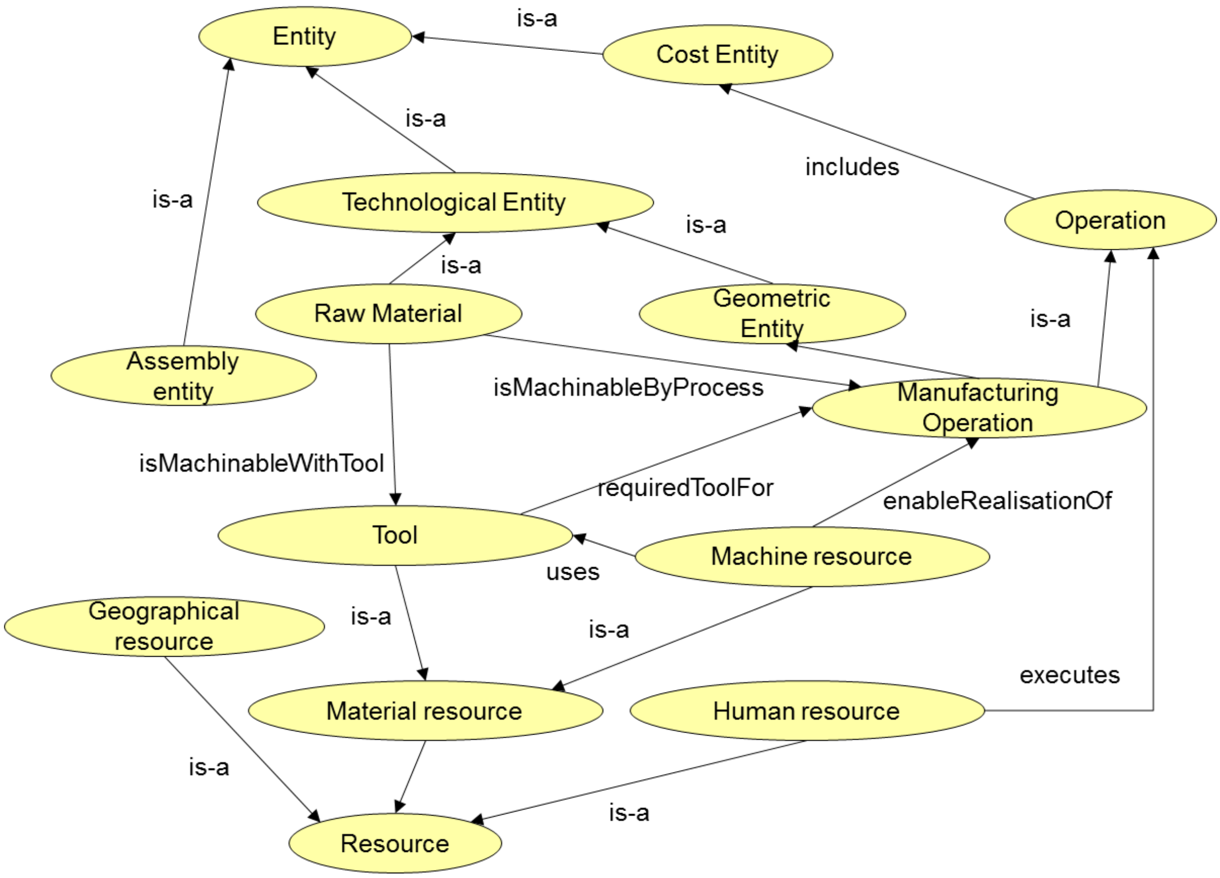


Figure 4.5 MASON ontology [SiEt-06]

In this thesis, the Manufacturing System Engineering (MSE) ontology is also considered as interesting ontology which offers the reusability. It aims to facilitate the communication of different software solutions across various organizations by providing a common structure of information. The structure has seven first level classes, i.e. `Resource`, `Process`, `Project`, `Flow`, `Enterprise`, `Extended_Enterprise`, and `Strategy`. This is depicted in Figure 4.6. Similar to MASON, the class `Resource` describes the human resource and manufacturing resource that can be derived to lower level abstraction, such as machine tools, and materials. `Process` also represents the business and manufacturing operations. `Project` describes products, documents, and programs that are produced by the enterprise. The `Project` and `Process` classes are related through the class `Flow`. `Strategy` defines the short-term operational and long-term business strategies. `Enterprise` represents the organization hierarchy of manufacturing, for instance, factory, manufacturing cell, workstations, etc. `Extended_Enterprise` aggregates the `Enterprise` and could contain one or more Enterprises. Besides adding the strategic component, the MSE ontology makes two additional extensions. It adds concepts for grouping `Process`, `Resource`, and `Strategy` concepts into different enterprises and aggregating those enterprises into groups [LiHa-07].

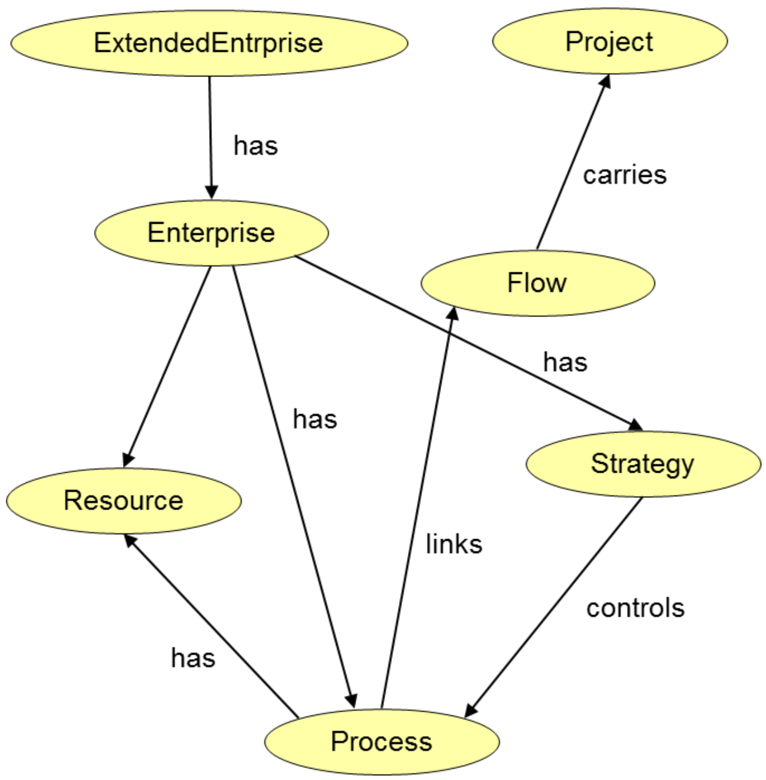


Figure 4.6 MSE ontology [LiHa-07]

In KnoHolEM[4] project, Wicaksono et al. developed an OWL ontology as the information for a holistic and intelligent energy management in building operational phase. The ontology is called Smart Energy and Resource Usage and Management in Building (SERUM-IB). The ontology consists of core classes representing entities involved in building energy management activities, i.e. `BuildingElement`, `BuildingControl`, `Actor`, `Behavior`, and `State`. The class `BuildingElement` models the building structures, for instances zones or rooms, HVAC elements, and electrical elements. Thus, the subclasses `Zone`, `ElectricalElement`, `HVACElement`, `StructuralElement`, etc. are created. Building elements are entities that are observed, examined and analyzed in energy management activities. The building elements are passive entities which have states but do not have capabilities to measure or to observe their own states. The class `BuildingElement` and its subclasses represent the fundamental of Building Information Model (BIM). It is aligned with the domain layer in IFC2x4 [WiEt-13].

The top class `BuildingControl` expresses the entities related to building automation system. It includes the sensors, actuators, controller, alarm, etc. which are elements of a building automation system. It has the capabilities to measure, to observe, and to control the state of `BuildingElement`. It aligns with entities in the IFC2x4 domain *IfcBuildingControlsDomain*. The top class `Actor` represents the human actors having behavior

[4] http://www.knoholem.eu

that can affect the states of building elements. The Actor can be organizations or persons, who have a name, postal address, telecom address, etc. It is aligned with the *IfcActorResource* in IFC2x4. The class `Behavior` represents the behavior performed by actors. The behaviors can affect the state of building elements. There are two methods to model the behavior in the building, i.e. bottom up and top down. The top class `State` indicates the state of building elements. It can be classified into `ComplexState` and `SimpleState`. The subclasses of `ComplexState` are for instances `ComfortState` and `EnergyEfficiencyState`, whereas subclasses of `SimpleState` are `WindowState`, `DoorState`, etc. [WiEt-15]. The whole class hierarchy is depicted in Figure 4.7.

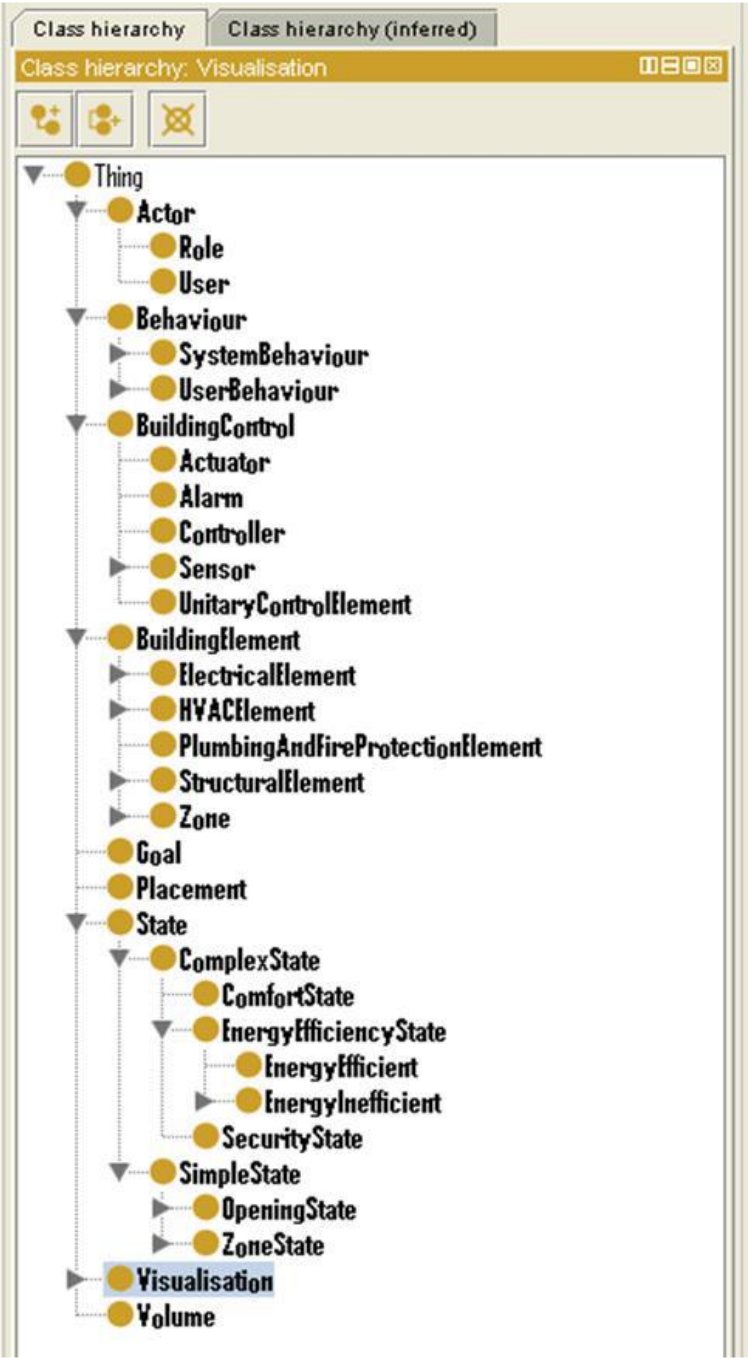


Figure 4.7 Class hierarchy of building energy management [WiEt-15]

The existing ontology development approaches for manufacturing domain always model the product, process, and manufacturing process as separated upper classes. However, these ontologies have different concept or class hierarchies. For example, MASON includes raw material in `Entity` hierarchy, whereas MSE assigns it as sub-concept of `Resource`. Besides containing the concepts representing product, process, and resource, both ontologies ADACOR and MSE introduce an additional concept representing the organization and strategic element in the manufacturing environment. MSE adds the class `Strategy`, whereas ADACOR adds `Supervisor` holons. In energy management domain, such as additional concepts might be useful to represent information how the operations are allocated on the machines to ensure the energy efficiency. Since the scope of this work focuses on supporting the evaluation phase of energy management system, the ontology should be able to describe the manufacturing topology and to give statements about the current states. It does not have to detail how such states are created.

The MSE ontology is also able to illustrate the production system organization. The manufacturing processes, resources, and strategies are organized in an enterprise and extended enterprise. This is very useful to represent the different organization levels of a manufacturing company and projects that involve more than one organizations. Additionally, MSE introduces a concept flow to connect the projects and processes indirectly. However, these extensions are not directly relevant to the energy management tasks. It can be concluded that the product, process, and resource hierarchies are crucial components of a manufacturing related ontology, and they provide a basis for the information modeling about manufacturing environment.

The KnoHolEM ontology contains concepts related to buildings and facilities that might be interesting for energy management in manufacturing. The concepts associated with states can represent energy efficiency states in a building. The concepts can easily be extended for manufacturing domain. The classes related to building controls can be adapted to model industrial automation systems that are used to capture the states in a manufacturing system, for instance, the power consumption of machines, the temperature of production halls, and machine states. The building element concept and its sub-concepts can be reused to constitute the building parts and ancillary facilities that assure production processes to function properly.

Knowledge Acquisition

The manufacturing energy management domain ontology should contain common knowledge of energy management that can be applied to any kinds of manufacturing companies. The domain ontology contains concepts that belong to part of the world, which in this case, is the manufacturing energy management. The ontology has to contain the best practices of energy management tasks that are applicable to different companies. The Reference Document on Best Available Techniques for Energy Efficiency that has been published by the European Commission and is briefly described in Section 2.2.4 provides a good start point to acquire knowledge that is conceptualized in the ontology. This section describes several examples of best practices that are modeled in the manufacturing energy management domain ontology. The best practices are exemplary listed in Table 4.2.

Table 4.2 Best practices of energy management with ontology as knowledge source

Best practice	Knowledge Source
Prefer the proportional pumps and avoid oversizing.	[Euro-09, p. 292]
Install high energy efficiency fans and cooling system.	[Euro-09, p. 294]
Ensure equipment to be operated below its rated voltage.	[Euro-09, p. 289]
Select power cables which has capacity corresponding to the power demand.	[Euro-09, p. 289]
Choose more efficient machines, usually the new ones.	Domain expert, common knowledge
Choose facilities that consume lower power than maximum allowed power.	Domain expert, common knowledge
Choose facilities that consume lower energy than specific energy consumption.	[ErWe-09]
The power factor of facilities has to be less than the high reactive power.	[Ries-12]
The peak power load to execute a job has to be less than the maximum allowed peak power load.	Domain expert, common knowledge
The utilization degree of a production facility has to be between minimum and maximum utilization.	[ErWe-09]
Jobs that consume high power should not be executed simultaneously if the accumulated consumed power leads to excess accumulative peak load.	Domain expert, common knowledge
Minimize the energy loss at start and stop of machines through reduction of start time and stop time.	[ErWe-09]
Minimize the energy consumption of standby state, but still considering the process security.	[ErWe-09]

The best practices are modeled as rules in the ontology. As shown in Table 4.2, some examples related to HVAC systems and electrical power supply are provided based on the best practice documents. Additionally, a discussion with a domain expert is conducted. The domain expert is a production planner in companies. Some of the best practices are also based on literature. By creating the rules additional required concepts or classes can be identified. If the required classes do not exist in the ontology yet, those classes are added in the corresponding position in the hierarchy.

The Knowledge Base Construction

As described in Section 2.3.4, OWL-DL offers an expressive representation comparing to RDF, RDFS, and OWL-Lite. It provides an implementation of the description logic. The knowledge base does not need to represent meta-model. Therefore, it is not necessary to use the OWL-Full. In the following the steps to construct the ontology knowledge base will be elaborated.

Important Terms in the Ontology

Manufacturing energy management considers not only production and energy aspects but also supporting entities, such as buildings, industrial automation, and organizations. Therefore, this thesis regards the following as essential terms in manufacturing energy management: Product, Process, Manufacturing Resource, Building Element, Automation, Energy Source, and State. Product represents purchased materials or products, the intermediate products that are manufactured in the factory, as well as the end products that will be sold to the customers.

Manufacturing processes are another important factors that influence energy consumption. Energy consumption of a process can be observed and measured thus the corresponding optimization techniques can be performed. The processes include production processes, logistic processes, and human processes. Manufacturing resources are required to execute the processes. Production facilities, such as machines and manufacturing equipment, transport facilities, and human resources, for example, production managers and operators, are considered as manufacturing resources.

Building elements of the factory, such the building structures (rooms, production halls, canteen, offices, etc.), lighting systems, and HVAC systems ensure the quality of manufacturing process execution. They also influence the energy consumption of the manufacturing system. The European Union Reference Document classifies those systems into three categories: energy consuming subsystems, energy transporting subsystems, and energy generating or reusing subsystems [Euro-09]. Furthermore, rated power and rated voltage of components are crucial information for energy efficiency rules in electric power supply systems. The energy generation sources that supply energy to the manufacturing system are also important aspects in manufacturing energy management. They can be from a utility company or internally generated, such as photovoltaic or thermal co-generation. It corresponds to the third category of subsystems that are introduced in the European Union Reference Document.

The ontology knowledge base has to be able to give the statements about the manufacturing system states. It should indicate the states of machines, processes, production lines, factory halls, etc. The states describe the energy efficiency that should be achieved and energy wasting or peak loads that have to be avoided within energy management activities. It is required to install industrial automation system to give the statements of states. It comprises sensors that measure the states of manufacturing resources, processes, and building elements, and actuators that can control the production and building facilities.

Definition of Classes and Class Hierarchy

The OWL class Product is modeled as a top class under the concept Thing, and the classes EndProduct, IntermediateProduct and RawMaterial as its subclasses.

A top level class Process is created to represent the manufacturing processes. The process is classified into ManufacturingOperation and ManufacturingJob. ManufacturingOperation represents the operations that can be executed by production facilities, for instance cutting a 50 cm metal, whereas ManufacturingJob describes concrete manufacturing operations that have already been launched into shop floor. The hierarchies of both ManufacturingOperation and ManufacturingJob are developed by adapting the process hierarchy of MASON consisting of ManufacturingOperation, HumanOperation, LogisticOperation and ManufacturingJob, HumanJob, LogisticJob respectively. All of those processes can affect the energy consumption in the company. As a fundamental of the manufacturing hierarchy, the following sub-classes are assigned to the ManufacturingOperation: Forming, Joining, Assembly, Machining, Molding, and Casting.

The class ManufacturingResource consists of the subclasses ProductionFacility, ProductionFacilityComponent, TransportFacility, and HumanResource. ProductionFacility includes machines where the manufacturing processes run. The classes Driller, Oven, Miller, etc. are created as subclasses of ProductionFacility. ProductionFacilityComponent describes the components of a complex ProductionFacility which can be controlled to improve the energy efficiency. HumanResource describes the personnel involved in the manufacturing system. It includes Operator, ManufacturingManager, etc.

The class BuildingElement models the building structures, such as rooms, production halls, canteen, offices, as well as the energy consuming facilities that indirectly affect the production processes, such as lighting and HVAC systems. Those facilities are included in the class AncillaryFacility. It is a direct subclass of BuildingElement By using the approach in interpreting the factory layout drawing, the ontological elements of the classes can be generated semi-automatically [WiRo-10a].

Corresponding to the first two categories of facilities described in European Union Reference Document [Euro-09], the classes EnergyConsumingAncillaryFacility and EnergyTransportingAncillaryFacility are created as subclasses of AncillaryFacility. The EnergyTransportingAncillaryFacility includes for example facilities related to electrical energy supply system commonly installed in manufacturing companies, cables, transformers, filters, capacitors, circuit breakers, supplies, and loads, etc. and facilities related to pumping, for instance, pump, fluid segment, pipe segment, etc. EnergyConsumingAncillaryFacility represents energy consuming subsystems includes lighting, HVAC, combustion, drying, separation and concentration, and electrical motor driven subsystems. The energy supply sources for the manufacturing system are represented by the OWL class `EnergySource`. The different methods and technologies to generate and to supply the energy are modeled as subclasses of `EnergySource`, for example, `ExternalPowerSource`, `Photovoltaic`, `ThermalCoGeneration`, etc.

The class `Enterprise` represents the different organization units of a manufacturing company where energy efficiency evaluation takes place. The subclasses of Enterprise are for example `Factory`, `ProductionLine`, `ProductionSegment`, and `Workplace` or `Workstation`. The class `Order` describes the customer orders that will be processed in the manufacturing. It has subclasses `CustomerOrder` and `ProductionOrder`.

The class `State` gives statements about the state of a manufacturing system. For manufacturing energy management purpose, the subclasses `Disturbance` and `EnergyEfficiencyState` are considered. The class `Disturbance` describes errors or disruptions that can affect the energy consumption of facilities or components. The class `EnergyEfficiencyState` refers to energy efficiency that should be achieved and energy wasting or peak loads that have to be avoided within energy management activities. Through this class, we can classify which practices or constellations could improve the energy efficiency, or which ones cause energy inefficiency. Power peak load is considered as a state that should be prevented since it can instantaneously cause high energy allocation thus causing higher energy costs.

The class `Automation` represents the integrated application model for different industrial and building automation technologies. It consists of two main subclasses i.e. `Actor` and `Sensor`. The actors for controlling the production or ancillary facilities are modeled as subclasses. The energy consumption meter, temperature sensor, occupancy sensor and other sensors are also included as subclasses of `Sensor`. Table 4.3 describes the existing ontologies that are aligned to the ontology classes.

Table 4.3 The reuse and alignment of existing ontologies

First level class	**Important subclasses**	**Reused/ aligned ontologies**
`Automation`	`Actor`, `Sensor`	[WiEt-15]
`BuildingElement`	`AncillaryFacility`, `EnergyConsumingAncillaryFacility`, `EnergyTransportingAncillaryFacility` `Zone`	SERUM-IB [WiAR-12][WiEt-15]
`EnergyEfficiencyState`	`EnergyEfficient`, `EnergyWasting`, `PeakLoad`	SERUM-IB [WiAR-12] [WiEt-15]
`EnergySource`	`ExternalEnergySource`, `InternalEnergySource`	[WiEt-15]
`Enterprise`	`Factory`, `ProductionLine`, `ProductionSegment`, `Workplace`	MSE [LiHa-07]
`ManufacturingResource`	`HumanResource`, `ProductionFacility`, `TransportFacility`	MASON [SieEt-06]
`Order`	`CustomerOrder, ProductionOrder`	ADACOR [BoLe-04]
`Process`	`Job: HumanJob, LogisticJob, ManufacturingJob` `Operation: HumanOperation, LogisticOperation, ManufacturingOperation`	MASON [SieEt-06]
`Product`	`EndProduct, IntermediateProduct, RawMaterial`	MASON [SieEt-06]

Definition of Properties and Their Facets

To characterize the products and to evaluate their influence on energy consumption, the following properties of `Product` are defined: `hasVolume`, `hasWeight`, `hasMaterialType`. Based on these properties, it can be estimated which machine the product should be manufactured with and how much energy it consumes.

The object property produces is added to ManufacturingOperation, to relate it to Product. The object property operableOn relates ManufacturingOperation to ProductionFacility. The relation is a many-to-many relation. The object property canOperate is an inverse of operableOn. The concept LogisticOperation that represents the intra-logistic operation or transport between production facilities has the properties hasOrigin and hasDestination with ProductionFacility being the range of both properties.

ManufacturingOperation has a data type property hasEnergyConsumption and hasPowerConsumption to express the energy and power consumption to execute the operation. The object property isEqualOperation is defined to express the equality whether two machines can execute the same operation but with different consumption of energy. This property is reflexive and symmetric. In contrast to that, a ManufacturingJob expresses a concrete job launched to the shop floor. A ManufacturingJob is related to a ManufacturingOperation via the class property performsOperation. Additionally, it has hasStartTime and hasEndTime data properties to describe its planned start and end. The object properties operates and operatedOn assign a ManufacturingResource to a ManufacturingJob. The number of operates properties is restricted to only one ManufacturingJob that has to be the currently executed or the next one to be executed. By adding this restriction, we are able to make an assertion about a machine being not in use during a job operated on another machine.

The class ProductionFacility has a relationship with ManufacturingOperation through the object property canOperate and a relationship with ManufacturingJob through the object property operates.

The object properties hasElectricalPowerInput and its inverse hasElectricalPowerOutput express the energy flow model using the EnergyTransportingAncillaryFacility in the manufacturing. They are sub-properties of hasEnergyInput and hasEnergyOutput that model the energy flow in general.

In addition, rated power and rated voltage of components are crucial information for energy efficiency rules in electric power supply systems. To capture this information, the data properties hasRatedPower and hasRatedVoltage are introduced. The HVAC system of a manufacturing company consists of ducts, fans, intakes, and air cleaners. They are considered as subclasses of the HVAC class. The structure of an HVAC system can be expressed similarly to the electric power supply system by using object properties hasAirInput and hasAirOutput to describe the topology of the HVAC system. Similar to the electrical power system, those properties are sub-properties of hasEnergyInput and hasEnergyOutput.

The EnergySource has the object properties hasEnergyInput to describe the energy source that is converted to another form of energy, and hasEnergyOutput to relate the EnergySource to the energy consuming facilities, i.e. ancillary, production, and transport facilities. The object property isAttachedOn is created to relate the classes Sensor and Actor to BuildingElement and ManufacturingResource. Table 4.4 lists the essential object properties, their domains and ranges.

Table 4.4 Essential object properties of manufacturing energy management ontology

Object property name	Domain	Range
hasAttachedSensor (inverse isAttachedTo)	BuildingElement, ManufacturingResource	Sensor
canOperate (inverse isOperableOn)	ProductionFacility	ManufacturingOperation
hasPowerInput	BuildingElement, ProductionFacility, TransportFacility	EnergySource
operates (inverse isOperatedOn)	ProductionFacility	ManufacturingJob
operationMoreEfficient	Operation	Operation
performsOperation	ManufacturingJob	ManufacturingOperation
produces (inverse isProducedBy)	Process	Product

Definition of Rules

The domain knowledge contains rules representing the best practice to improve energy efficiency that is defined in European Union Reference Document [Euro-09] and other literature [ErWe-09][Ries-12], and also based on common knowledge about energy efficient practices. These kinds of rules can be applied in different manufacturing companies. Thus, they are included in domain knowledge. The rules are expressed in SWRL.

The first rule representing the best practice as described in the document is related to electrical power supply system. The rule can be used to identify whether a building element or a manufacturing resource operates above its rated voltage (see equation (4.1)). For every BuildingElement or ManufacturingResource *?r*, it is checked whether it has a VoltageSensor *?s* attached to it. Then, it is examined whether the current voltage of the equipment is less than the rated voltage. If the current voltage exceeds the rating, *?r* is additionally assigned to the `VoltageExceedsRating` which is a subclass of `AncillaryFacilityEnergyWasting` [WiEt-14].

```
(BuildingElement or ManufacturingResource)(?r) ∧
VoltageSensor(?s)∧hasAttachedSensor(?r,?s)∧
hasRatedVoltage(?r,?rv)∧
hasValue(?s,?sv) ∧ swrlb:greaterThan(?sv,?rv)→
VoltageExceedsRating(?r)
```
(4.1)

The second rule also comes from the EU document. It evaluates if any cable in the electrical supply network is operated above its rated power (see equation (4.2)). For every `BuildingElement` or `ManufacturingResource` *?r* with an attached `PowerSensor` it is checked whether the current power demand exceeds the rated power in the `Cable` connected to *?r*.

```
(BuildingElement or ManufacturingResource)(?r)∧
PowerSensor(?s)∧ hasAttachedSensor(?r,?s) ∧ InternalNetwork(?n)
∧ Cable(?c)∧ uses (?n, ?c) ∧ hasRatedPower(?c, ?pr)∧
hasValue(?s,?ps)∧ swrlb:greaterThan(?ps,?pr)→
CablePowerExceedsRatedPower(?c)
```
(4.2)

Related to HVAC, a rule referring the inefficient and efficient fan is modeled in the ontology based on the EU document. The maximum efficiency of a fan is normally 60-85% [Euro-09, p. 240]. If a `Fan` has an efficiency below 60%, it can be considered that the fan is currently inefficient, and it should be replaced. The rule checks the given rated maximum efficiency value of a fan instance. In case the value is below 60%, the instance is assigned to the class `IneffficientFanUsed` if true (see equation (4.3)). `IneffficientFanUsed` is a subclass of `EnergyInefficient`.

```
Fan(?f)∧hasRatedMaximumEfficiency(?f,?e)∧
swrlb:lessThan(?e,0.6f)→ IneffficientFanUsed(?f)
```
(4.3)

SWRL rules can be applied to examine the energy consumption of operations and to identify which operation consumes less energy. The rule expressed in (4.4) checks whether the energy consumption of *?o2* is less than the energy consumption of *?o1* by comparing the `hasPowerConsumption` properties of both operations using the rule-inferred property `operationMoreEfficient` [WiEt-14]. The value of `hasPowerConsumption` represents the typical power consumption of the operation that is calculated based on historical data.

```
Operation(?o1)∧Operation(?o2)∧hasPowerConsumption(?o1,?c1)∧
hasPowerConsumption(?o2,?c2)∧swrlb:lessThan(?c2,?c1)→
operationMoreEfficient(?o2; ?o1)
```
(4.4)

The exceeding power load of an ancillary or production facility can be detected using SWRL rule as well. The rule in (4.5) indicates that if a production or ancillary facility *?f* has power input *?rp* greater than the threshold power peak *?rpp*, the rule engine will set the facility *?f* as an individual of the class `FacilityExceedsReferencePeakDemand` which is a subclass of `EnergyInefficient`. A similar rule can also be applied to detect exceeding energy consumption by comparing the current energy consumption value to an assigned threshold.

```
(EnergyConsumingAncillaryFacility or ProductionFacility)(?f)∧
hasPowerConsumption(?f,?rp)∧
hasReferenceMaxPowerInput(?f,?rpp)∧
swrlb:greaterThan(?rp,?rppi)→
FacilityExceedsReferencePeakDemand(?f)
```
(4.5)

The object property `hasPowerConsumption` represents the power consumption of a facility. The value of the property is determined by the power sensor that is attached to the facility. The following rule expresses this relationship.

```
(EnergyConsumingAncillaryFacility or ProductionFacility)(?f)∧
PowerSensor(?s) ∧ hasAttachedSensor(?f,?s) ∧ hasValue(?s,?v) →
hasPowerConsumption (?f,?v)
```
(4.6)

Rules and (4.6) show that ontologies equipped with SWRL allows a collaborative construction of the knowledge base. For example, a production planner or facility manager models the rule (4.5) that has a higher level of abstraction. An electronic or automation engineer creates the more technical rule (4.6) because he has more knowledge of sensors, their characteristics and how they are connected to the facilities.

For the facilities powered through Alternating Current (AC), the reactive power is an important factor that determines the energy consumption value. Reactive power is defined as the power for setting-up and down of the magnetic or electric field between the generator and consumer in the pulse of network frequency power [Ries-12, p.5]. The oscillation between electricity network and consumer reactive current is converted into heat in the electrical network. Thus, losses occur. From the perspective of the energy supplier, losses lead to an increase in investment and maintenance costs for the utility grid. The costs are charged to the energy consumer [Ries-12]. Therefore, the analysis of the reactive power is very important.

The energy loss because of high reactive power can be detected by measuring the minimum power factor and compare the measured value to the threshold value defined by the utility company. Usually, the cost of reactive energy is calculated when it exceeds 50% of the active load. This corresponds to a power factor cos φ = 0.9. There is no requirement that the value of 0.9 must never be exceeded. In some public utility districts, other performance factors are required, for example, 0.85 or 0.95 [Ries-12]. Equation (4.7) expresses the rule to detect energy inefficiency because of the high reactive power. A more technical rule supplemented with mathematical formulas can be created by an electrical engineer to get the minimum power factor of a facility.

```
(EnergyConsumingAncillaryFacility or ProductionFacility)(?f)∧
hasMinPowerFactor(?f,?v)∧swrlb:lessThan(?v,0.9f)⟶
FacilityHasHighReactivePower(?f)
```
(4.7)

Energy inefficiency of a single manufacturing process can be detected using SWRL rules as well. Rule (4.8) describes a simplified example of a rule that detects the power value *?v* of a manufacturing job *?j* exceeding the given threshold value *?rv*. The rule is a high level rule that is modeled typically by a production planner to avoid creating many peak power loads.

```
ManufacturingJob(?j)∧ hasPowerConsumption(?j,?v)∧
hasReferenceMaxPowerInput(?j,?rv)∧
swrlb:greaterThan(?v,?rv)→
JobExceedsReferencePeakDemand(?j)
```
(4.8)

The power consumption of a manufacturing job can be obtained through the rule (4.9) which is modeled by a more technical person, for example, production engineer. The power consumption of the facility is obtained by using the rule modeled by an electronic engineer as expressed in (4.6).

```
ProductionFacility (?p) ∧ ManufacturingJob(?j)∧ operates(?p,?j)
hasPowerConsumption(?p,?c)→ hasPowerConsumption(?j,?c)
```
(4.9)

The production facilities have to operate in minimum energy intensity in the corresponding production cycle. Figure 4.8 shows the point in the production cycle, where the optimal load i.e. the lowest energy intensity is achieved. It lays between the partial load and overload points.

The rules in (4.10) and (4.11) detect the production facilities that operate outside the range illustrated in Figure 4.8.

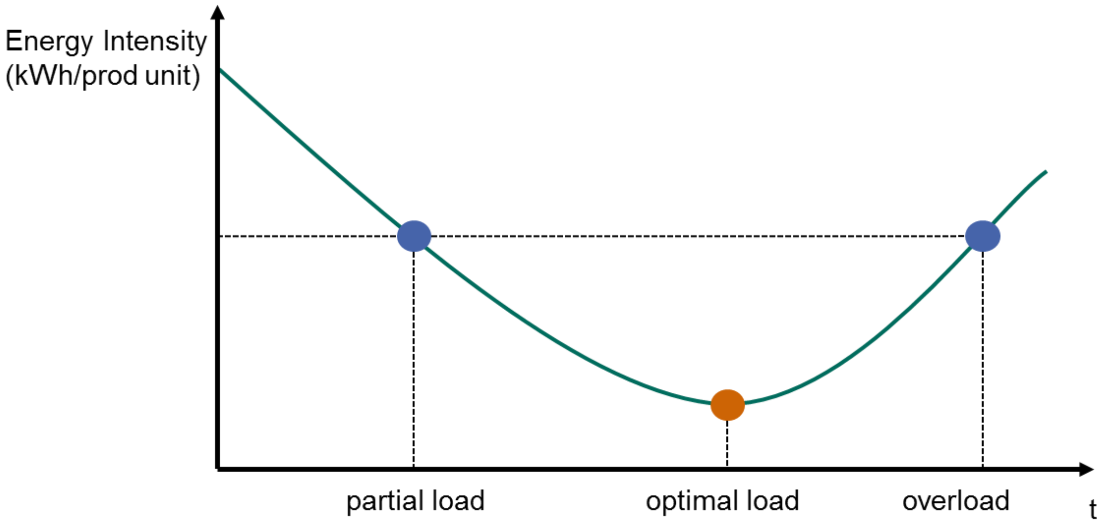


Figure 4.8 Optimal utilization of a production facility [ErWe-09]

```
ProductionFacility(?p)∧hasCapacityUtilization(?p,?cu)∧
hasReferenceMinCapacityUtilization(?p,?rmcu)∧
swrlb:lessThan(?cu,?rmcu)→ProductionFacilityHasPartload(?p)
```
(4.10)

```
ProductionFacility(?p)∧hasCapacityUtilization(?p,?cu)∧
hasReferenceMinCapacityUtilization(?p,?rmcu)∧
swrlb:greaterThan(?cu,?rmcu)→
ProductionFacilityHasOverload(?p)
```
(4.11)

The power input value of a job can be derived from the power input value of the production facility which currently executes the job. This relation is expressed in (4.12).

```
ManufacturingJob(?j)∧ProductionFacility(?f)∧
operates(?f,?j)∧ hasPowerConsumption(?f,?v)→
hasPowerConsumption(?j,?v)
```
(4.12)

SWRL rules can be used to detect exceeding power load caused by some jobs which are executed simultaneously. An example of this kind of rules is described in (4.13). If the power load value *?v* of the energy network *?n* exceeds the given threshold *?rv* then the network is considered in an energy-inefficient state. The corresponding instance of the class

`EnergyNetwork` is then assigned to the class `NetworkOverloadedByParallelJobs`.

```
EnergyNetwork(?n)∧hasTotalPowerDemand(?n,?v)∧
hasReferenceMaxPowerConsumption(?n,?rv)∧
swrlb:greaterThan(?v,?rv)→ NetworkOverloadedByParallelJobs(?n)
```

(4.13)

The value of `hasRatedOperationalTotalPowerDemand` can be determined using relation described in (4.14). It is assumed that the end time is always greater than the start time. Both production facilities *?f1* and *?f2* are connected to the same energy network. They execute manufacturing *?j1* and *?j2* simultaneously.

```
EnergyNetwork(?n)∧ManufacturingJob(?j1)∧
ProductionFacility(?f1)∧operates(?f1,?j1)∧
ManufacturingJob(?j2)∧ProductionFacility(?f2)∧
operates(?f2,?j2)∧isInNetwork(?f1,?n)∧isInNetwork(?f2,?n)∧
hasStartTime(?j1,?s1)∧hasEndTime(?j1,?e1)∧
hasStartTime(?j2,?s2)∧hasEndTime(?j2,?e2)∧
swrlb:greaterThanOrEqual(?s2,?s1)∧
swrlb:lessThanOrEqual(?s2,?s1)∧hasRatedPowerInput(?f1,?v1)∧
hasRatedPowerInput(?f2,?v2)∧swrlb:add(?v,?v1,?v2)→
hasRatedOperationalTotalPowerDemand(?n,?v)
```

(4.14)

A rule can be used to evaluate the states of a production facility, whether energy loss through long start or stop phase occurs. The duration of start and stop phase has to be less than maximum corresponding reference values. The equations (4.15) and (4.16) express this rule in SWRL.

```
ProductionFacility(?f)∧hasState(?f,?s)∧StartState(?s)∧
hasDuration(?s,?d)∧hasReferenceMaxStartDuration(?f,?md)∧
swrlb:greaterThanOrEqual(?d,?md)→
ProductionFacilityHasStatePowerLoss(?f)
```

(4.15)

```
ProductionFacility(?f)∧hasState(?f,?s)∧StopState(?s)∧
hasDuration(?s,?d)∧hasReferenceMaxStopDuration(?f,?md)∧
swrlb:greaterThanOrEqual(?d,?md)→
ProductionFacilityHasStatePowerLoss(?f)
```

(4.16)

In a real manufacturing environment, the production manager creates the rules (4.15) and (4.16). A production engineer, who has more knowledge of the data obtained from MES, expresses the `StartState` as rule (4.17). With a similar principle he can create a rule to express `StopState`.

```
ProductionFacility(?f)∧ ManufacturingJob(?j) ∧
operates(?f,?j)∧ isActive(?j, false) ∧ hasStartTime (?f,?sf)
∧ hasStartTime (?j,?sj)  ∧ swrlb:lessThan(?sf,?sj) →
hasState(?f,?s) ∧  StartState(?s)
```
(4.17)

SWRL rules can also be used to detect the energy loss due to the high energy consumption during the standby state or too low power consumption that can endanger the continuity of manufacturing processes. The equations (4.18) and (4.19) express the corresponding rules.

```
ProductionFacility(?f)∧hasState(?f,?s)∧StandByState(?s)∧
hasPowerConsumption(?s,?p)∧
hasReferenceMaxPowerConsumption(?s,?m)∧
swrlb:greaterThanOrEqual(?p,?m)→
ProductionFacilityHasStandbyPowerLoss(?f)
```
(4.18)

```
ProductionFacility(?f)∧hasState(?f,?s)∧StandByState(?s)∧
hasPowerConsumption(?s,?p)∧
hasReferenceMinPowerConsumption(?s,?m)∧
swrlb:lessThan(?p,?m)→
ProductionFacilityEndangersProcess(?f)
```
(4.19)

4.2.2 Population of Ontology with Company Specific Information

The individuals are created through different ways. Figure 4.9 summarizes the processes to create the individuals and to populate of the ontology using company specific information. The instances of the class `Automation` and its subclasses are mostly created by implementing software to map the application models of the automation technologies into the ontology (see Figure 4.9, step 1). The population would be more efficient if the company applies automation technologies which uses standardized information model such as SensorML[5].

The factory layouts are commonly created as two-dimensional drawings using tools such as AutoCAD[6]. The tools are commonly used to draw building spatial elements and the ancillary facilities in the factory, such as HVAC, access control, production machines, and energy

---

5 http://www.sensorml.com/

6 http://www.autodesk.com/products/autocad/overview

conversion or generation facilities such as solar cells [KrEt-09]. The complexity of objects drawn in a layout may vary depending on the way the engineers express the real world in geometrical representations. CAD is one of the easiest and oldest technologies used in the industry. However, it is ineffective for building information modeling due to the great amount of effort required to perform the activity. Recent researches show the evidence that the technologies do not ensure high quality, reliable, and coordinated information that the higher level of BIM produces [Dona-08] [VaNC-08].

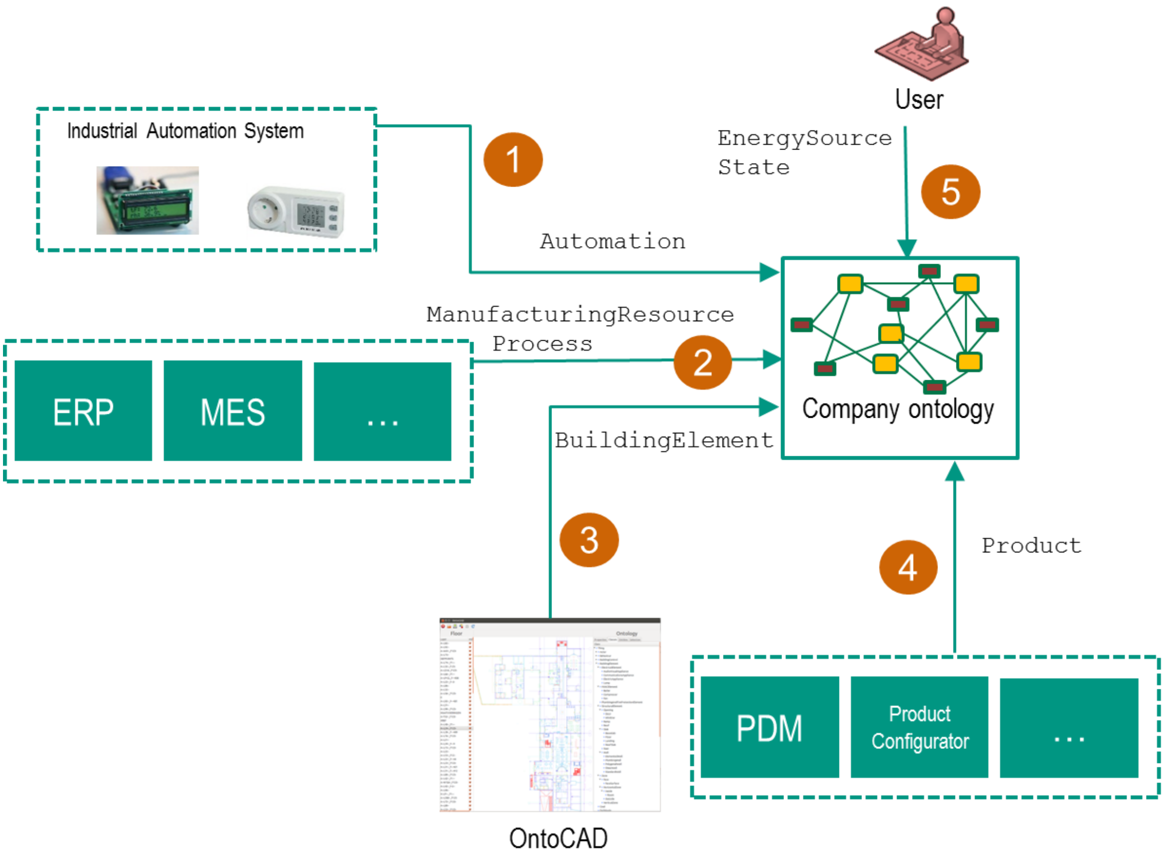


Figure 4.9 Ontology population process

OntoCAD Tool developed by the Institute for Information Management in Engineering (IMI), Karlsruhe Institute of Technology (KIT) extracts the semantic information from CAD drawing and populates the ontology using this extracted semantic information (see Figure 4.9, step 3). First, A CAD drawing is exported into exchange format DXF using CAD tool. Then the drawing primitives in the DXF file are imported into the OntoCAD tool. This vector based data representation is used as the basis for the viewer in OntoCAD GUI and the pattern matching algorithms. The different layers contained in the drawing are listed and visualized in different colors (see Figure 4.10, Layers). Using the OntoCAD GUI the user specifies the mapping of the ontology specific data and object properties with the OntoCAD functions, for instance, the computation of the object position. The pattern matching and classification algorithms recognize building elements based on user defined templates. The user selects an object from which an individual should be created (see Figure 4.10, Selection) and chooses the corresponding ontology class shown in the hierarchy (see Figure 4.10, Ontology). Using the GUI, the user can also search for similar objects and then populate all of them at once. The user has the possibility to validate the result directly and if necessary apply some corrections to the results. The results are continuously and automatically saved to the ontology [HäEt-13] [WiEt-15].

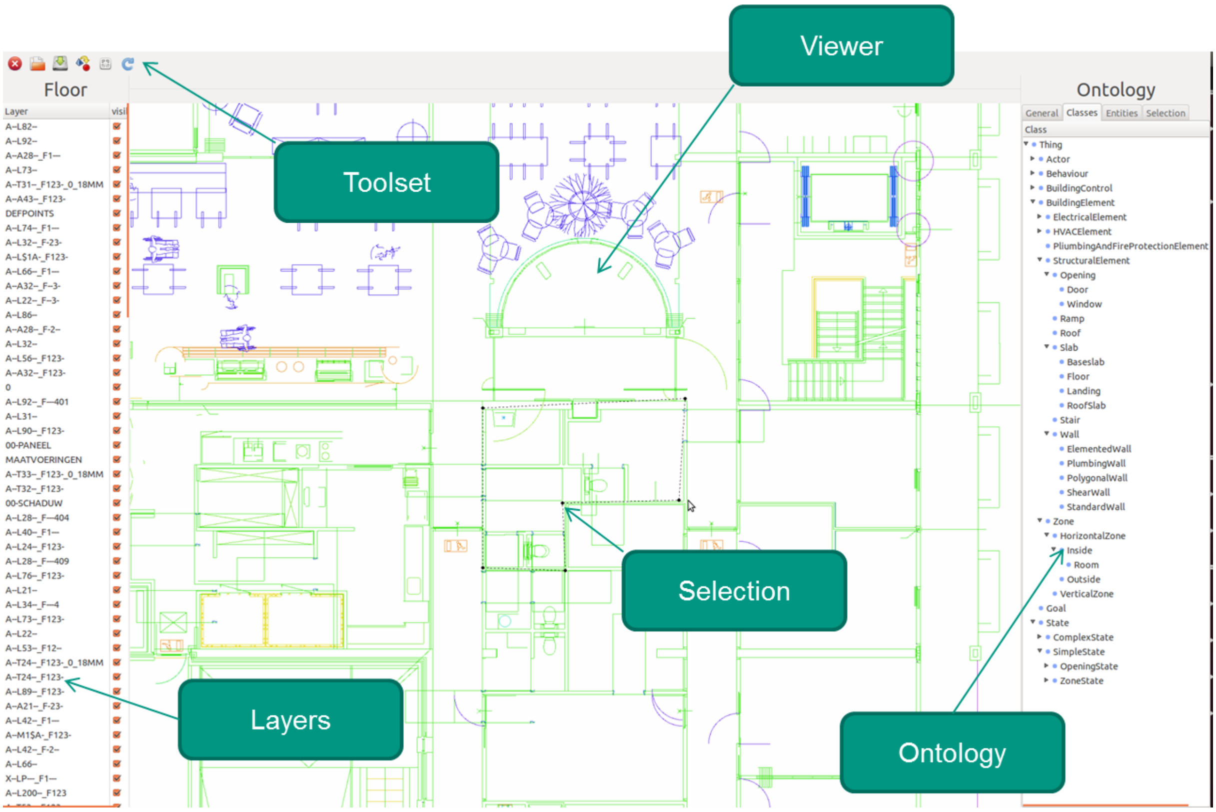


Figure 4.10 OntoCAD GUI

The individual of process and resource related classes are created from the information coming from ERP (see Figure 4.9, step 2). Most of the ERP systems store the information about the human resource and production facilities, such as machines. The list of available manufacturing operations that can be executed on those machines can also be found in the ERP. Individuals corresponding to the jobs executed on the shop floor are created based on information available in the ERP and MES. The product classes are populated by using the information stored in product configuration system or PDM (see Figure 4.9, step 4). Due to the availability of information in digital formats, the population of product, process, and resource related classes can be performed using software interface that maps the schema of PDM, ERP, or MES installed in the company.

The population of some other classes may have to be done manually by experts. This is, for example, the population of the part related to the energy source and manufacturing states (see Figure 4.9, step 5). Some of the individuals of classes related to the states are created automatically by the reasoning process on the rules and axioms that represents the relations leading to state changes.

4.2.3 Querying the Knowledge Base for Qualitative Evaluation

The main objective of the development of the ontology knowledge base is to improve the qualitative evaluation through automatic identification of energy efficiency states. It is achieved by performing reasoning on the knowledge base based on information fetched from ICT systems in the company and querying the knowledge base using SPARQL. Figure 4.11 illustrates the process to perform the energy efficiency qualitative evaluation through querying the knowledge base.

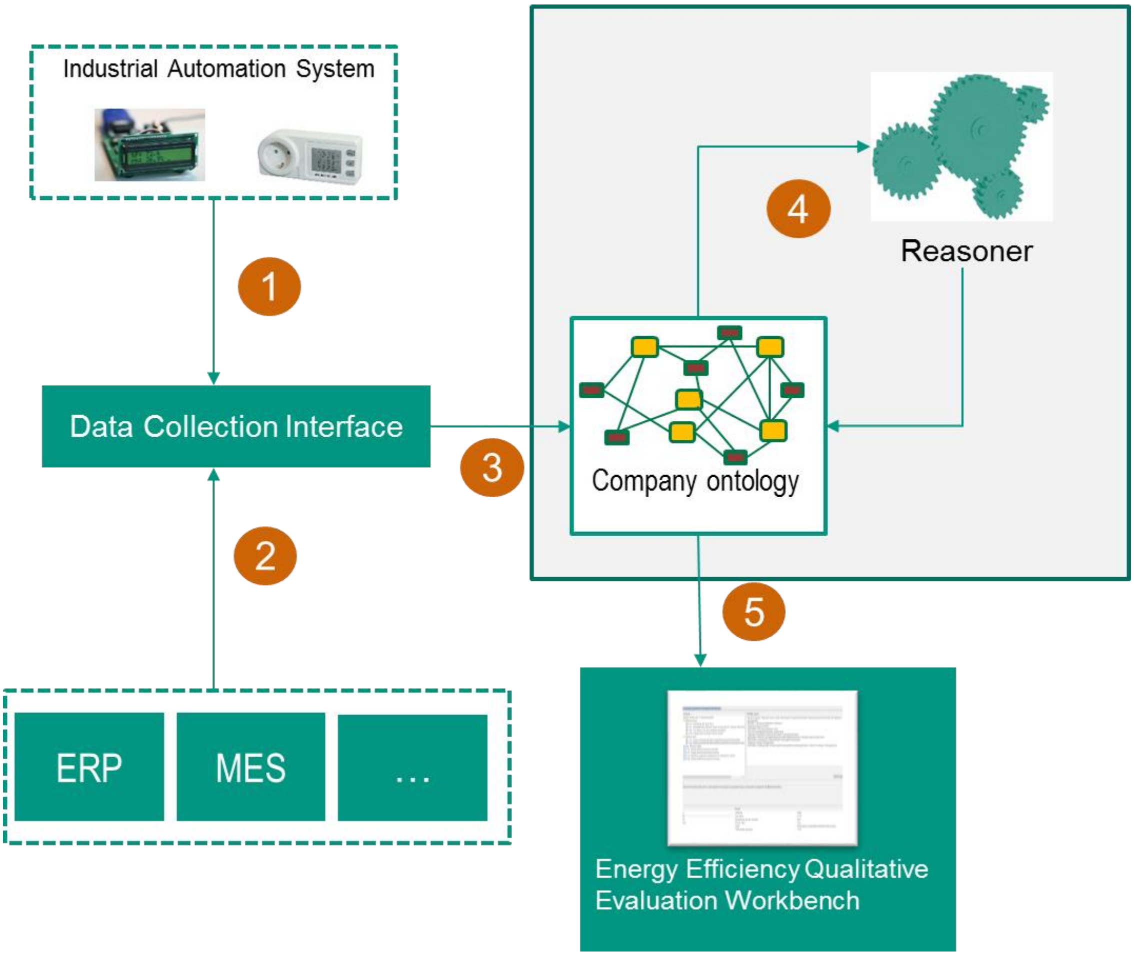


Figure 4.11 Energy efficiency qualitative evaluation by querying the knowledge base

As seen in Figure 4.11 the Data Collection Interface fetches the sensor data, for example, the power consumptions, temperatures, generated power, etc. periodically, for example, every 15 minutes (see Figure 4.11, step 1). Simultaneously, the interface retrieves the information about the states of the current manufacturing system, for instance, process states, orders, and processed products from manufacturing IT systems, such as ERP, MES, and PDM (see Figure 4.11, step 2). The data properties of individuals corresponding to industrial automation are updated periodically using the values coming from the industrial automation system. The properties are for example `hasPowerConsumption`, `hasMinPowerFactor`, etc. Meanwhile, the data properties of individuals corresponding to manufacturing entities are assigned with values coming from manufacturing IT systems. The properties `hasState`, `hasStartTime`, `hasEndTime`, etc. are examples of this category. The property values updating process is described as step 3 in Figure 4.11.

At the next second, the reasoning engine is executed to fire the SWRL rules in the knowledge base resulting new knowledge, for instance, new property values, new class-individual assignments and new relations (see Figure 4.11, step 4). These new knowledge represent the states of manufacturing objects at the current time. As seen in step 5, the states are queried by utilizing SPARQL. The Energy Efficiency Qualitative Evolution Workbench is a GUI for querying the knowledge base and evaluation the current states in the manufacturing. The following gives several examples of SPARQL to evaluate the energy efficiency states.

Find all production facilities which currently operate efficiently

```
SELECT ?facility
WHERE { ?facility rdf:type
      serene: ProductionFacilityEnergyEfficient }
```
(4.20)

Find all jobs and their corresponding production facilities which currently consume more power than the allowed maximum value

```
SELECT ?job ?facility
WHERE { ?facility serene:operates ?job .
?job rdf:type
      serene: JobExceedsReferencePowerDemand }
```
(4.21)

Find all facilities (production, transport and ancillary facilities) which currently consume more power than the allowed maximum value.

```
SELECT ?facility
WHERE {?facility rdf:type
      serene: FacilityExceedsReferencePeakDemand }
```
(4.22)

Example

Consider M1 and M2 are production facilities, i.e. instances of the class `ProductionFacility`. The maximum allowed power consumption values of M1 and M2, i.e. values of `hasReferenceMaxPowerInput` are 20 kW and 50 kW respectively. The values of `hasPowerConsumption` of M1 and M2 are updated with values coming from sensor readings. Table 4.5 lists the power consumption of M1 and M2 over time.

Table 4.5 Values of the property hasPowerConsumption over time

Time (T)	hasPowerConsumption	
	M1	M2
T1	18	45
T2	19	52
T3	22	52

At each T the reasoning engine is executed and fires rules. By firing the rule (4.5), the reasoning engine gives different results at T1, T2, T3. At T=T1, the query (4.22) give no results because the class `FacilityExceedsReferencePeakDemand` has no instance (see Figure 4.12). At T=T2, the power consumption of M2 exceeds the maximum allowed value. The reasoning engine sets M2 as an instance of `FacilityExceedsReferencePeakDemand` (see Figure 4.13). Finally, the power consumptions of both M1 and M2 exceed the maximum allowed values. The query (4.22) returns both M1 and M2 because the reasoning engine makes them as instances of `FacilityExceedsReferencePeakDemand` (see Figure 4.14).

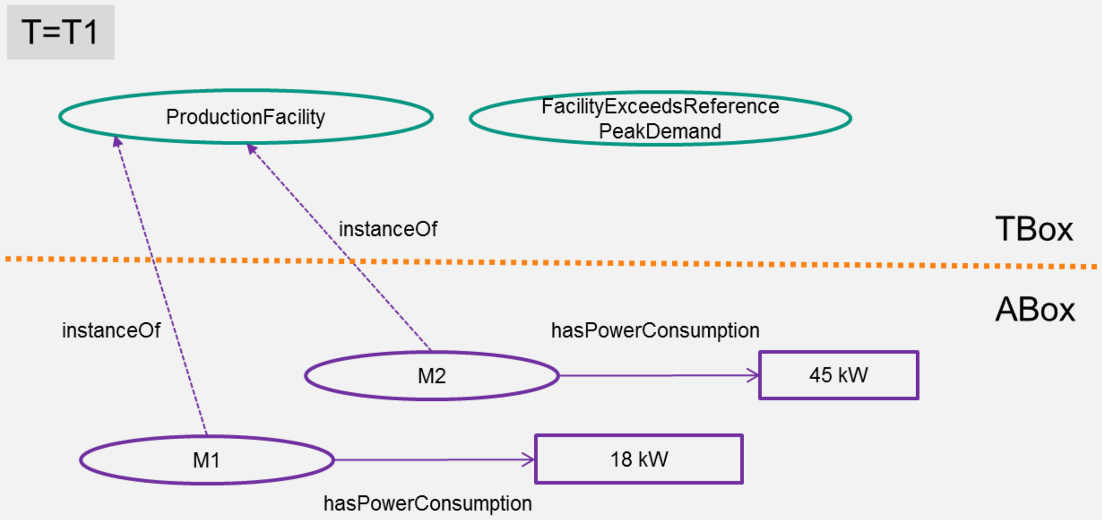


Figure 4.12 The ontology state at T=T1; M1 and M2 are instance of ProductionFacility

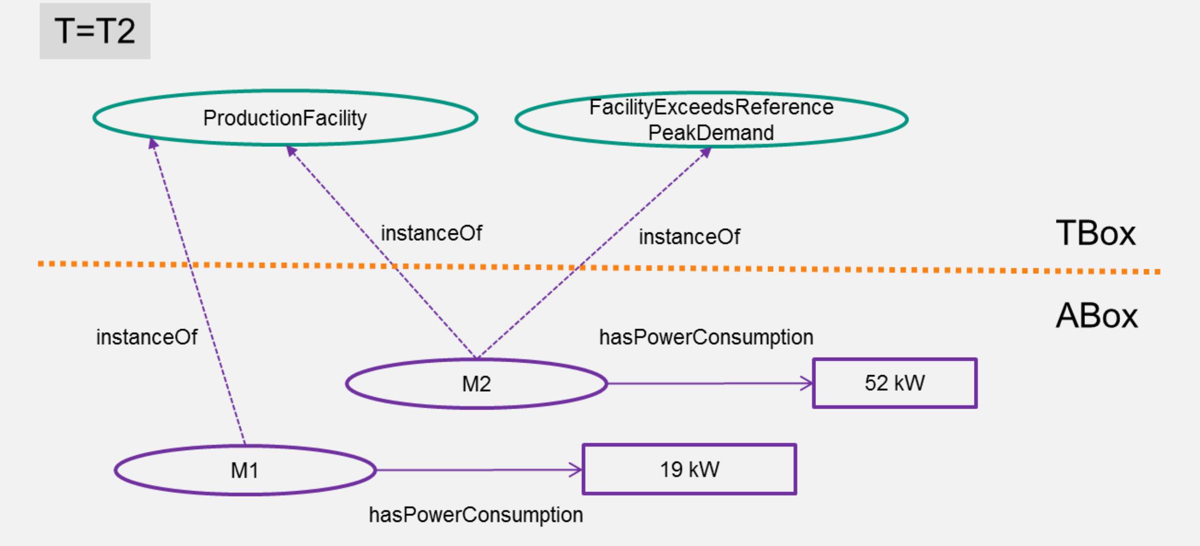


Figure 4.13 The ontology state at T=T2; M1 and M2 are instance of ProductionFacility; M2 is also instance of FacilityExceedsReferencePeakDemand

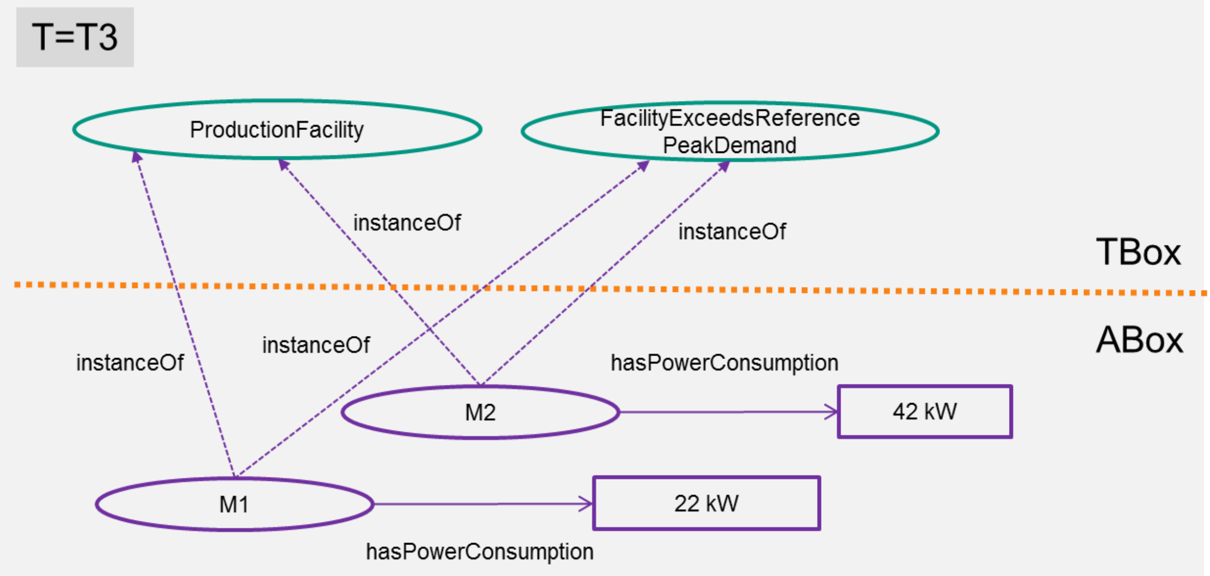


Figure 4.14 The ontology state at T=T3; M1 and M2 are instances of ProductionFacility and FacilityExceedsReferencePeakDemand

4.3 Semi-automatic Knowledge Generation Using Data Mining

A production process planning strongly depends on received customer orders. The properties of ordered products determine the sequence of production steps and the selection of the machines where the products should be manufactured. The machine selection is performed based on the knowledge of process planners or some experienced employees in the company. The knowledge concentrates mainly on the technical feasibility of the machine to accomplish the process and how process sequences generate the lowest costs. It often disregards the relations between the product properties and energy consumption of the machines where the product is manufactured.

The information about product profiles, process portfolios, and energy-related information are spread throughout data from different IT systems in the company. For instances, the profiles of ordered products are stored in the order management system or in the Customer Relationship Management (CRM) system. Production processes are recorded in Material Execution System (MES). Energy consumption data are collected by the Energy Monitoring and Data Acquisition (EMDA) or industrial automation system. The data from these different systems are not related automatically. Production or process planners have to analyze the data individually and relate them manually.

This section demonstrates the knowledge extraction from different data using data mining as a part of knowledge discovery in database (KDD) techniques. Data mining generates rules and relationships from different data related to energy efficiency semi-automatically. The rules and relationships are then incorporated into the ontology knowledge base. This procedure aims at fulfilling the Requirement R2.4 described in Section 3.2. In this thesis, the extracted knowledge is used as knowledge to decide on machine selection during the production process planning by considering the energy performance and the production costs of the machines. Furthermore, it is used to predict the energy consumption, order processing time, production and energy costs of an individualized configured product.

4.3.1 Definition of Company Data Sources

The properties or configuration of ordered products influence the machine selection during the process planning or shop scheduling, in the case when a production process step can be carried out on different machines. These properties also influence the process characteristics such as process time and energy consumption. These process characteristics are determined by machine selection. For example, a machine with higher processing capacity consumes more energy per processing time unit, but having a relatively short processing time. On the contrary, a machine with lower processing capacity consumes lower energy per processing time unit, but it needs longer processing time. These dependencies are depicted in Figure 4.15 [WiOv-12].

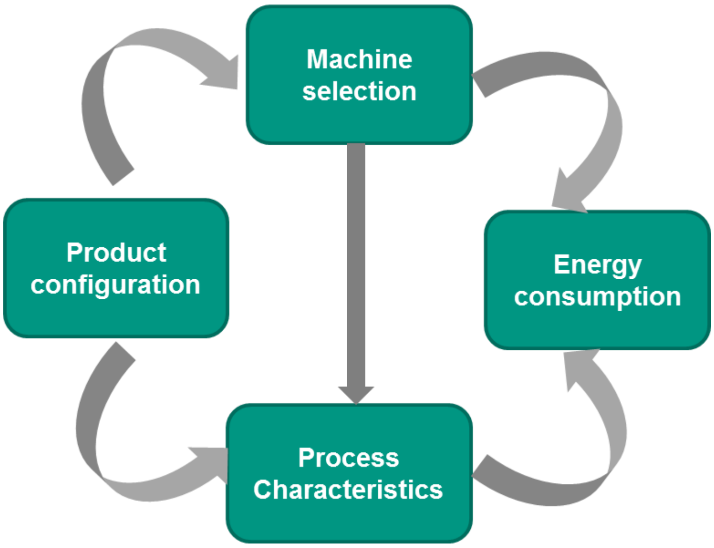


Figure 4.15 The influence of product configuration on energy consumption [WiOv-12]

For each process that contributes in the manufacturing of a product, the decisive influencing factors for machine selection, process characteristic, and energy consumption must be identified. Theoretically, it can be assumed that for each production process, one critical product characteristic exists, i.e. weight, size, surface or chip volume, etc. [ErWe-09]. The factors that influence energy consumption of production processes can be very different for each product and process. For example, in the separation processes such as turning or milling, the volume that has to be removed influence the energy consumption, whereas in the production processes that modifies material properties, the energy consumption depends strongly on the weight. Therefore, for each step in a manufacturing job sequence, the decisive product characteristics that influence energy consumption have to be identified individually. Thus, it is important to understand the production processes. Nevertheless, data mining can be used to identify these decisive characteristics.

Product profile or product configuration data are extracted from the order management systems, CRM, or product configuration systems. Through a filtering step, the most important product properties are selected. Some of those data have to be transformed, in order to match to the already defined data schema. The same steps are performed to process the resource data. The resource data come from the ERP system and contain information about production resources that are used by the manufacturing processes of the company. The most important resource data in this thesis are data of available machines and related costs. Production process log or machine activities data are collected using MES. The data describe processes characteristics, the processed order or product of each process, and the start and finish time of each process. Energy data are gathered using industrial automation systems. By using an energy meter on each machine, the energy consumptions of the corresponding machine are measured in a certain period, for instance, 15 minutes. The measurement data are then saved in a database. Figure 4.16 depicts a simplified Entity-Relationship diagram defining the data required for the input of KDD process. The entity `Product` represents the product with their properties that influences energy consumption of the production facilities which process the product, for example, the dimension or material type. `Order` is the customer order containing the product specifications. `ProductionProcess` models the processes that process the orders on machines. Processes cause machines to consume energy [WiOv-12].

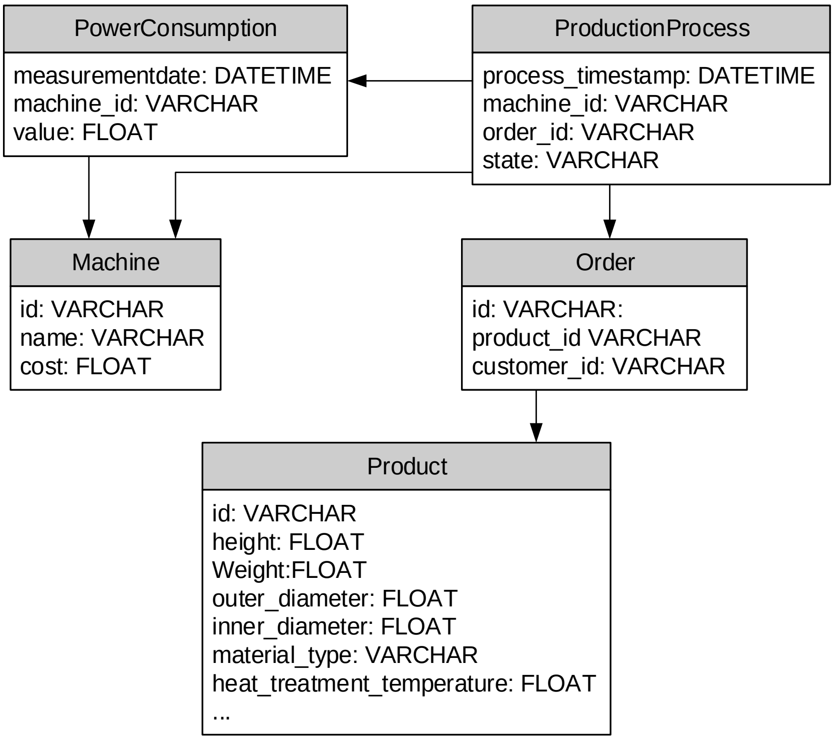


Figure 4.16 Entity-Relationship (ER) diagram describing the required data for data mining [WiOv-12]

4.3.2 Data Pre-processing

Data pre-processing is critically important and determines the quality of data mining applications. Therefore, it forms an important basis for this thesis. The goal of data pre-processing is to improve the quality of the processed data. Thus, it leads to satisfactory results. The data commonly originate from multiple and heterogeneous sources. It most likely contains noise, missing values, or inconsistent data, which have to be removed or manipulated. This section explains different preprocessing techniques that are implemented in this thesis in order to improve the quality of data mining results as well as the efficiency of the data mining process. The whole preprocessing steps are illustrated in Figure 4.17.

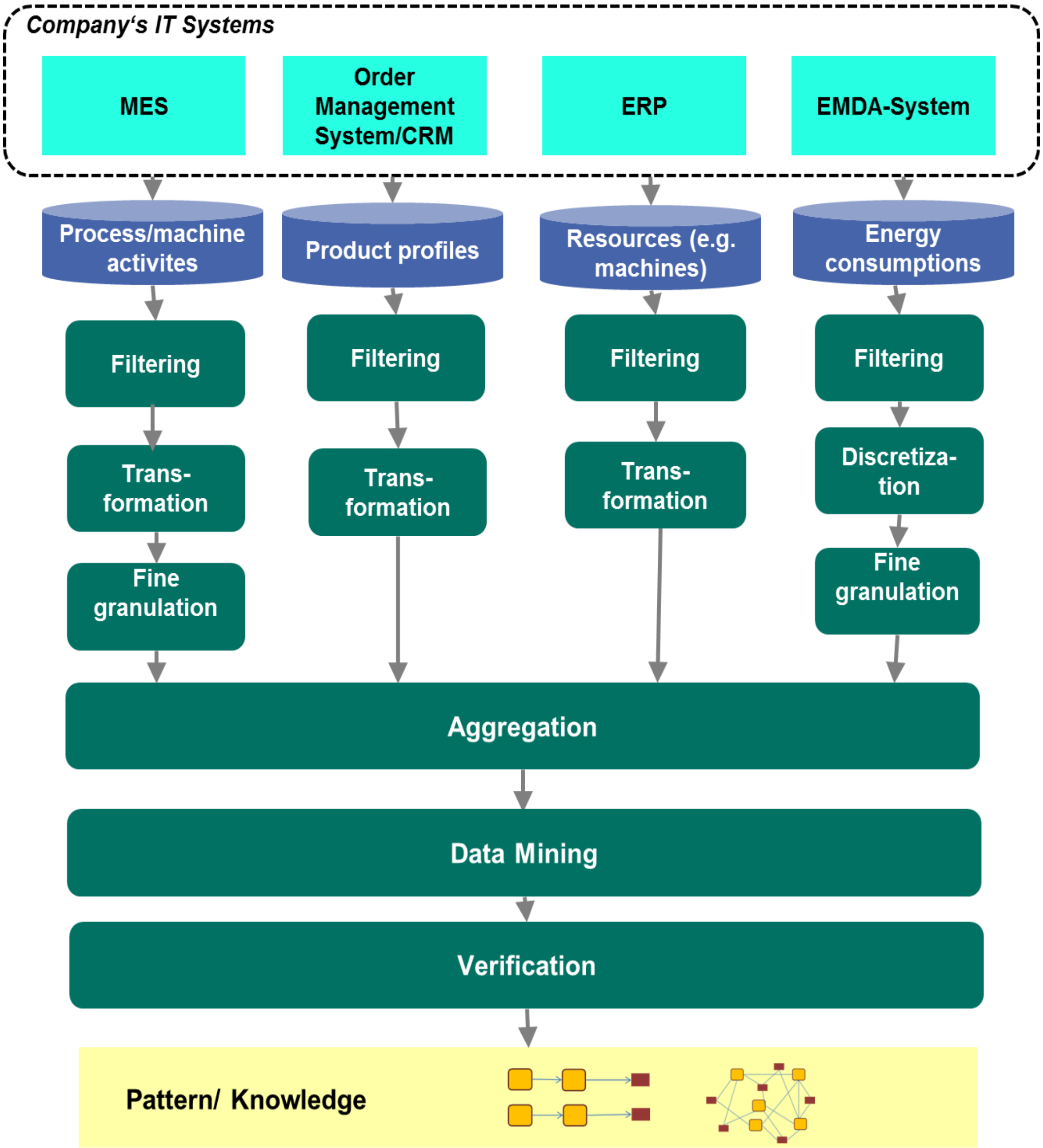


Figure 4.17 The steps to extract knowledge using KDD technique [WiOv-12]

Selection

Selection refers to the purposeful selection of data from one or more sources, e.g. data warehouse systems, databases, or data with experience-based knowledge, for further processing, analysis, and evaluation. During the selection phase, only important and relevant data are collected. In this step, meta information of the data have to be already figured out, for example, what is the goal of the data mining, what kind of knowledge have to be revealed from the data, where are the data located. This thesis selects the machine activity log data coming from MES, product profiles including dimensions, materials, quantities from the Order Management System, machines with their costs, operations, associated human resources, and metering data, such as energy consumptions and temperatures (see Figure 4.17).

Data Cleaning

The aim of data cleaning is to eliminate the noise and erroneous data as well as to correct the corrupt or inaccurate data. The validity of the data is checked. Therefore, it increases the accuracy of the next steps. In data cleaning step, incorrect or incomplete data sets are identified, corrected or specially labeled. For example, missing data can be assigned predefined values. The parameters of the records can also be reduced to the necessary parameters for use. In this thesis, the data cleaning step consists of a step to remove erroneous values and outliers. The erroneous values from sensor measurement, such as energy consumption and temperature, have to be removed and excluded from the execution of data mining algorithms. Since the data coming from measurement represent the consequences of the influencing parameters (e.g. features of the processed products, production process characteristic (duration, tool, etc.)), manipulations should be avoided. The measurement data may contain outliers resulting from measurement errors or exceptional conditions. In this thesis, the outliers are detected using interquartile ranges method that is already implemented in data mining tool WEKA. Equation (4.23) defines the outliers, where o_i is the outlier, Q_l is the lower quartile, Q_u is the upper quartile, and k is a given constant [WiFH-11] .

$$o_i < Q_l - k(Q_u - Q_l) \;\; OR \; o_i > Q_u + k(Q_u - Q_l) \tag{4.23}$$

Missing values could affect the quality of data mining results. If values of several attribute in a tuple are missing, it is difficult to extract the relation among attributes. The simplest method to overcome this is to ignore or remove the tuple containing missing values. However, the method is not effective, if the data set contains a lot of missing values since it leads to mass removal of tuples. Alternatively, a global constant such as “unknown” or “N.A.” could be assigned to the place of missing values. In the case of numeric attributes, an impossible value, such as “-999”, can fill the missing values. Other possibilities are to fill the missing values with the average value of the corresponding attribute or with the most probable value [HaKa-01]. In this thesis, the missing measurement values are removed since they only have a share lower than 1% of the whole dataset.

Transformation

The objective of transformation step in KDD is to prepare the data in a schema and format that can be processed by data mining algorithm. Commonly, the transformation consists of data mapping and data generation steps. The data mapping maps the source data to the desired schema or format. The data generation generates the data automatically based on the defined mapping. Depending on the type and source, data might have different file formats. For further analysis, these must be transformed to a specific format depending on the data mining application, for example, the ARFF format used in data mining tool WEKA which can easily be generated from CSV files. In this thesis, the data used as input for data mining algorithms are relational data that organized in tables and columns. The following transformations are performed in this thesis:

1. Splitting of order data into product and order data. An order entry contains both order related information such as due date, customer, product quantity, release date, and product characteristics. This should be separated in order to have an accurate relation between product characteristics and energy consumption of machines, on where the product is processed.

2. Transformation involving date attributes. Two date attributes are subtracted to acquire a third attribute representing the time difference. For example, the machine activity logs contain only information about start and end time of production operations. An attribute representing the duration of production operations can be created by subtracting the end time with the start time.

3. Text to attribute transformation. The source data contain attributes which have text as data types. The text holds multiple information that has to be broken down into several attributes in order to get low granulated data. Thus, the data mining results can be more meaningful and accurate. For example, the product dimensions are stored as a long description text. It should be decomposed into different attributes, for instances, inner diameter, outer diameter, height. A parsing technique is required to do this. In this thesis, the parsing is conducted by pattern matching using regular expression (regex, regexp). Regular expression allows flexible ways to match text, such as particular characters, words, or pattern of characters [NN-04].

4. Data type transformation, such as transformation from string to date, from string to number, and date-time formatting.

5. Data format transformation that transforms the text into column or field distinctive format, such as CSV and ARFF.

Discretization

Some classification and clustering algorithms are only able to deal with nominal attributes and cannot handle numerical attribute. Algorithms which are capable of processing numeric attributes work even faster and produce more satisfying results if discretization is performed on the attributes first. This is because continuous values are difficult to predict. The simplest way to carry out the discretization is to sort the instances by the attribute's value, to split the attribute into ranges, and then to assign the instances into the corresponding ranges. Each range has to be filled with a certain minimum number of instances. This technique is the common method to discretize continuous attributes before performing the learning in data mining phase.

Discretization methods can be classified into supervised and unsupervised discretization. Supervised discretization performs the discretization process by taking into account known classes. On the contrary, in unsupervised discretization, there is no knowledge about classes of the instances in the training set when discretization process is carried out. In supervised discretization, the attribute values are divided into a predefined number of equal intervals. In unsupervised discretization, the intervals are created during the discretization process. There are two methods to define the discretization intervals. The first one is equal-interval binning (grouping into equal intervals). In equal-interval-binning, instances are often distributed very unevenly. It may be possible that some intervals contain many instances, but others not. This may produce unsatisfactory results after executing the data mining algorithm. It is better to allow different sized intervals where each interval holds the same number of training examples. The second method is equal frequency binning i.e. grouping with the same frequency where the splitting is based on the distribution of training examples [WiFH-05].

In this thesis, the data mining is used to predict the energy consumption of machines based on the characteristics of processed product. It is also exploited to detect the energy consumption anomalies. In this thesis, an equal interval binning is implemented, owing to the energy consumption data that are convergent to a value or interval corresponding to the particular product characteristis. Each range is supposed to be filled with a similar number of instances. Values lying far outside the desired interval are considered as anomalous values and then to be removed. Algorithm 4.1 demonstrate the implemented discretization.

Algorithm 4.1 Discretization Algorithm

require: number of instance *n*; number of interval *m;*
max value of discretized attribute i_{max}; min value of discretized attribute i_{min}
while i < n **do**
m ← **min** (n,m)
bin_size ← $i_{max} – i_{min}$ / m
lower_bound ← max (i_{min}, i_{min} + **floor** ((i- i_{min} /bin_size) * bin_size))
upper_bound ← lower_bound + bin_size
end while

Interpretation

In the final evaluation, experts or users who are familiar with the application area have to analyze and to differentiate, which extracted knowledge is important. The aim is also to identify relationships between parameters, to verify the analysis, and to document the pattern found from the previously unused record.

4.3.3 Data Mining Algorithm for Knowledge Generation

In this thesis, a set of data mining algorithms is applied in order to extract the hidden pattern in the company’s historical data semi-automatically resulting in a set of rules and interrelationships. The results are incorporated in the ontology knowledge base and extend it with learned knowledge. The learned knowledge is used to perform predictions that lead to the energy efficiency measures. For example, it allows predicting the energy consumption and cost of an individually customized product. Therefore, the company will be able to estimate the energy effort of the customized product. Furthermore, the extracted knowledge can be used to help the production planners in selecting machines and material flow sequences to produce the ordered customized product, with the objective of achieving the minimum energy consumption.

In order to predict the energy consumption the relation between product configuration, process characteristics, machine selection, and energy consumption depicted in Figure 4.15 have to be considered. First, the properties of a product, which influence the energy consumption of the machines processing it, have to be identified. For example, to predict the energy consumption in producing a customizable stainless steel product having a cylinder shape, the diameter, the length, and the material are the important factors determining the energy consumption.

As seen in the ER diagram depicted in Figure 4.16, the entity `Product` containing relevant properties is related to the entity `Order`. It is implemented through the property `product_id` of the entity `Order`. The `ProductionProcess` has properties `order_id` referring to the entity `Order` and `m_id` referring to the machine id in the entity `Machine`. `PowerConsumption` contains data about the power consumption of each machine at the certain time points. These relationships with some simplified example representing the pre-processed data are illustrated in Figure 4.18. As seen in the figure, the tables `PowerConsumption` and `ProductionProcess` contain a column date. The columns represent the time points, where the energy consumptions are measured, and where the production processes are recorded. Using the columns, the relation between process and energy consumption can be built. For example, the production process that processes the order `10-3661-1` on machine `95512` consumes the electrical power of `56` kW at `2012-01-19 13:30:00`.

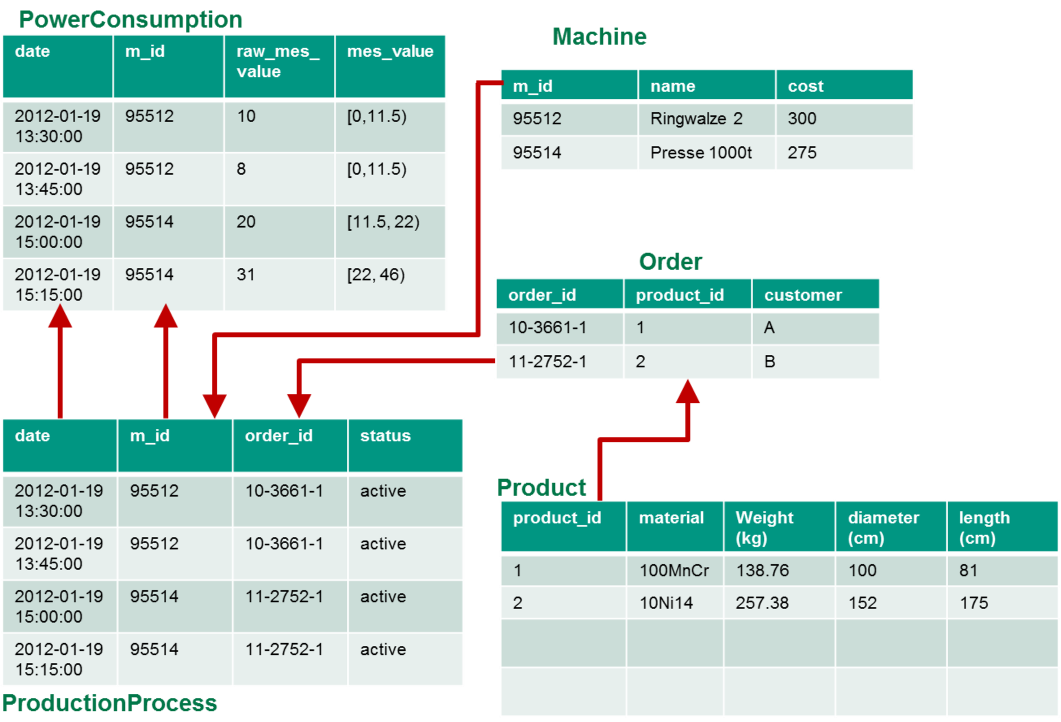


Figure 4.18 Relationships between entities containing pre-processed data

The data mining algorithms are executed on the preprocessed data. As described in Section 2.3.5, the decision tree is a method to generate a set of classification rules. Figure 4.19 illustrates a decision tree representing a set of rules for predicting power consumption patterns by means of product properties, if the product is processed in a certain machine, for example, a rolling machine. As seen in the tree, a node tests the values of a certain attribute. If the attribute type is nominal, for example, material, the number of branches is equal to the number of possible values of the attribute. If the attribute type is numeric, i.e. weight and diameter, the split value is calculated, and it results in two branches. The split attribute is determined by entropy representing the information gain (see Section 2.3.5). The class attribute is discretized `mes_value` which indicates power consumption range. One of the common algorithms that implement decision tree to find classification rules is C4.5 [Quin-93].

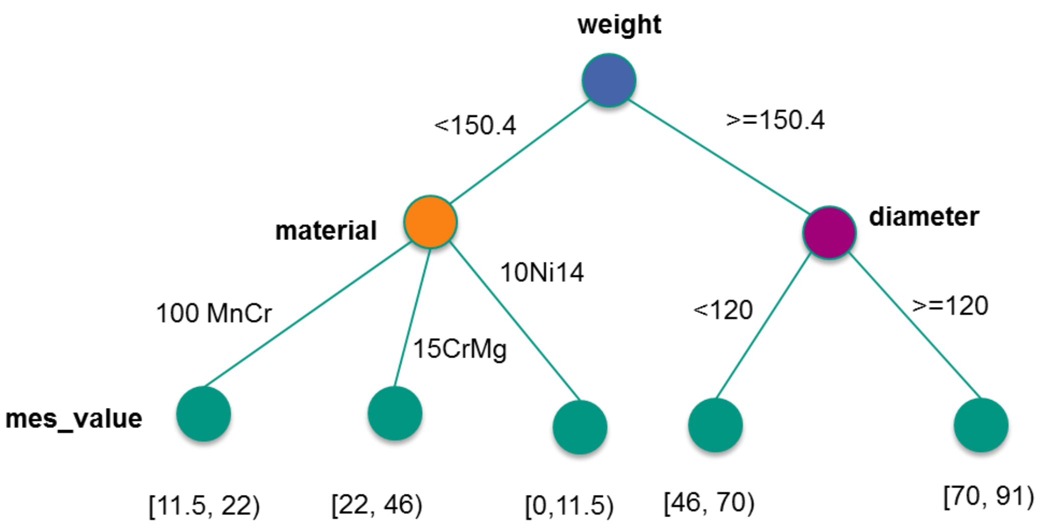


Figure 4.19 A decision tree representing a set of rules for finding the power consumption pattern of a rolling machine

As explained in Section 2.3.5, a path from the root to a leaf in a decision tree indicates a classification rule. From the decision rules illustrated in Figure 4.19, the classification rules listed in Table 4.6 can be generated. The rules describe the common relations of products, manufacturing processes, and energy consumptions in the company. By using the rules as references, anomalies or inconsistent energy performances could be detected.

Table 4.6 Classification rules resulted from a decision tree for predicting typical power consumption

No	Rule
1	`IF weight < 150.4 kg AND material='100MnCr' THEN mes_value=[11.5,22)`
2	`IF weight < 150.4 kg AND material='15CrMg' THEN mes_value=[22,46)`
3	`IF weight < 150.4 kg AND material='10Ni14' THEN mes_value=[0,11.5)`
4	`IF weight ≥ 150.4 kg AND diameter < 120 cm THEN mes_value=[46,70)`
5	`IF weight ≥ 150.4 kg AND diameter ≥ 120 cm THEN mes_value=[70,91)`

Example:

A product P1 has the weight of 100 kg and material type of '100MnCr'. At T=T1, a power meter attached to the rolling machine shows the value of 25 kW, while the machine is processing P1. By considering the rule no 1 in Table 4.6, it can be concluded that there is an inconsistent performance on the machine since the typical power consumption of the machine when processing the similar product is lower than 22 kW. Therefore, the employee should examine the machine.

At T=T2, the power meter shows the power consumption of 11.5 kW, while the process processing P1 is still running. The process is consuming less power than in the typical situation. It is not considered as an inconsistent performance, but rather an energy efficient state. The employee could find out that such a process should be reproduced in the future.

Decision tree is a popular technique in data mining for predicting and describing the correlation between values of attributes and the target values in order to segment a dataset into subsets (classes) [Lebe-08]. This is due to the simple and self-descriptive representation. Thus, it is easy to understand. It allows the processing of nominal attributes. Decision tree can deal well with errors and missing values. It only needs the values of the attributes that are queried along the paths. In the case of many attributes per record, it has the benefits in allowing quick classification. However, it has the drawback due to the incapability to predict numerical values. As seen in Figure 4.17, the values of the class attribute representing power consumption have to be discretized before the algorithm is executed.

Another approach for predicting numerical values is linear regression. It is a statistical technique that allows classification of instances through linear functions. It is able to predict numerical values. The linear functions model the mathematical interrelationships among attributes. A linear function can be represented in (4.24). x denotes the class, a_i the attribute values, and w_i the weight of a_i, and w_0 is a constant [WiFH-05].

$$x = w_0 + w_1 a_1 + w_2 a_2 + \cdots + w_k a_k \tag{4.24}$$

The aim of linear regression is to find the weight w_i that minimize the sum of the square of the difference between the actual value $c^{(i)}$ and the predicted value, as presented in (4.25).

$$\min(\sum_{i=1}^{n} \left(c^{(i)} - \sum_{j=0}^{k} w_j a_j \right)^2) \tag{4.25}$$

The weights are searched using common search algorithms, such as greedy search, which have available implementation software. Linear regression is a simple technique for numerical prediction and has been applied to solve different problems for many years. However, it can only handle numerical values.

M5P Tree is a data mining technique that can solve the problem. It is also called "Model Tree" method which is typically used for numerical forecasts. The M5P model is similar to decision tree model, except the leaves contain linear models, which enable to classify instances reaching the leaves. M5P creates the initial tree through the decision tree algorithm. The difference between a model tree and the typical decision tree is that for the model tree, each node is replaced by a regression plane instead of a constant value. The split attribute is determined by a criterion called standard deviation reduction (SDR) of a node. The splitting process terminates when the standard deviation of the class values that reach a node is only a small fraction (e.g. < 5%) of the standard deviation of the original instance set, or when just a few instances are left (e.g. < 4) [WiFH-05, p. 245]. If an instance is classified through M5P tree, similar to the one in the decision tree, it follows the path accordingly until it reaches a leaf. The leaf contains a linear model representing raw predicted value of the instance. Before pruning, each internal node must have linear model with the following format:

$$w_0 + w_1 a_1 + w_2 a_2 + \cdots + w_k a_k \tag{4.26}$$

where a_1, a_2, …, a_k are the attribute values and w_1, w_2, …, w_k indicate the weight calculated through standard regression technique. From the same example explained in Figure 4.19, the M5P tree for predicting power consumption value is illustrated in Figure 4.20. The corresponding rules are displayed in Table 4.7.

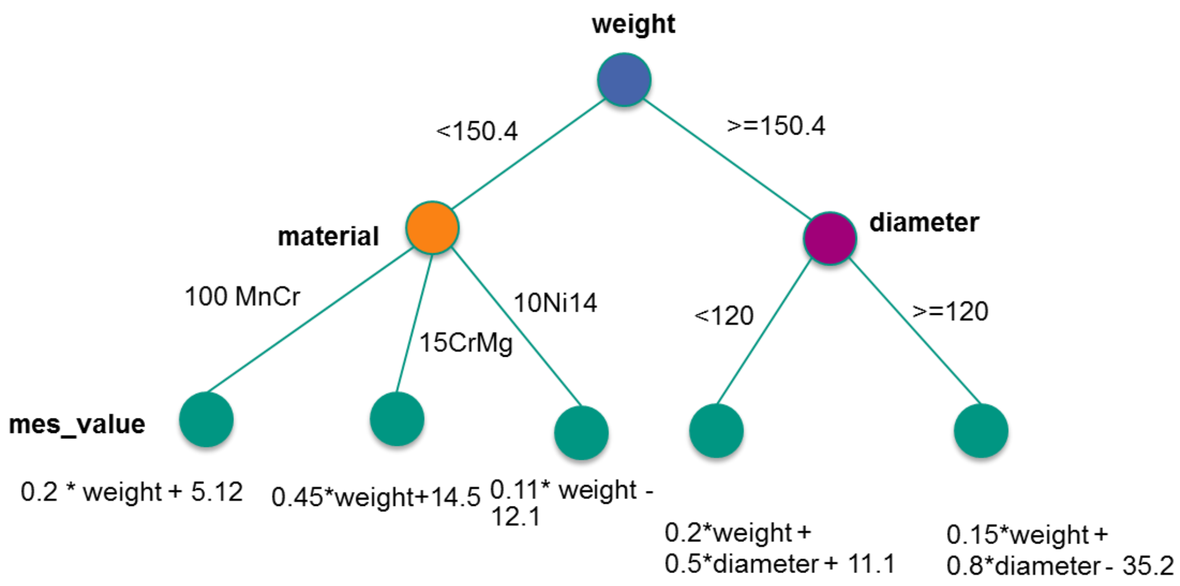


Figure 4.20 An M5P tree representing for the prediction of power consumption pattern

Table 4.7 Rules resulted from an M5P tree for predicting typical power consumption of a machine

No	Rule
1	`IF weight < 150.4 kg AND material='100MnCr' THEN mes_value=0.2*weight+5.12`
2	`IF weight < 150.4 kg AND material='15CrMg' THEN mes_value=0.45*weight+14.5`
3	`IF weight < 150.4 kg AND material='10Ni14' THEN mes_value=0.11*weight-12.1`
4	`IF weight ≥ 150.4 kg AND diameter < 120 cm THEN mes_value=0.2*weight+0.5*diameter+11.1`
5	`IF weight ≥ 150.4 kg AND diameter ≥ 120 cm THEN mes_value=0.15*weight+0.8*diameter-35.2`

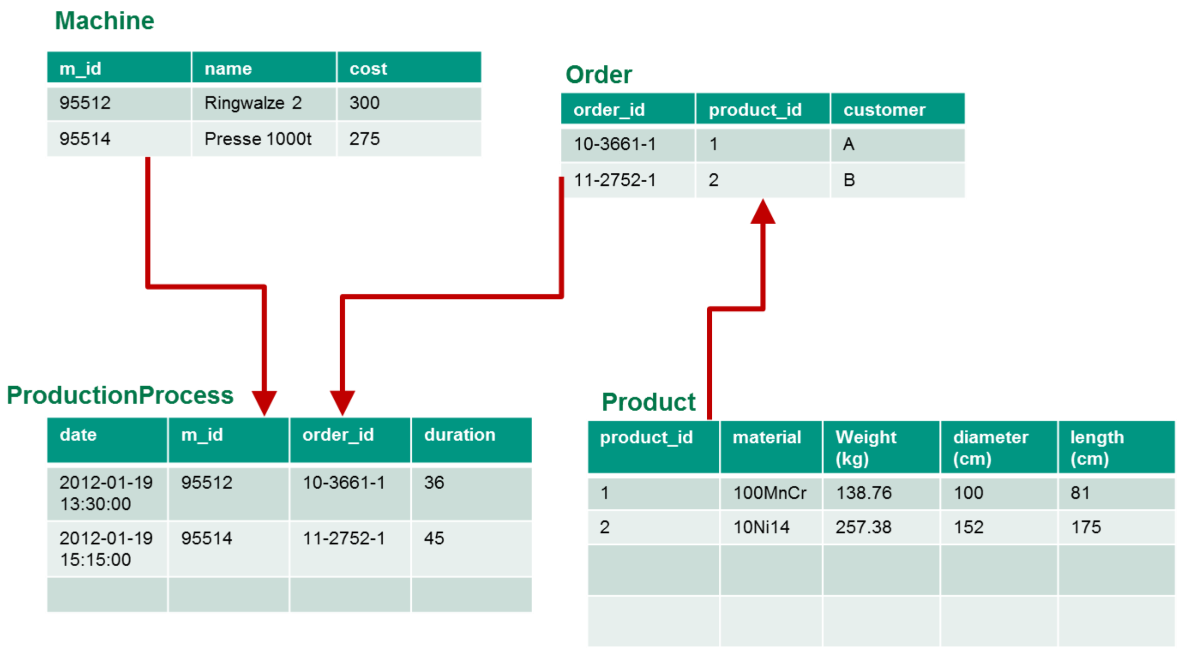


Figure 4.21 Preprocessed data to predict process duration

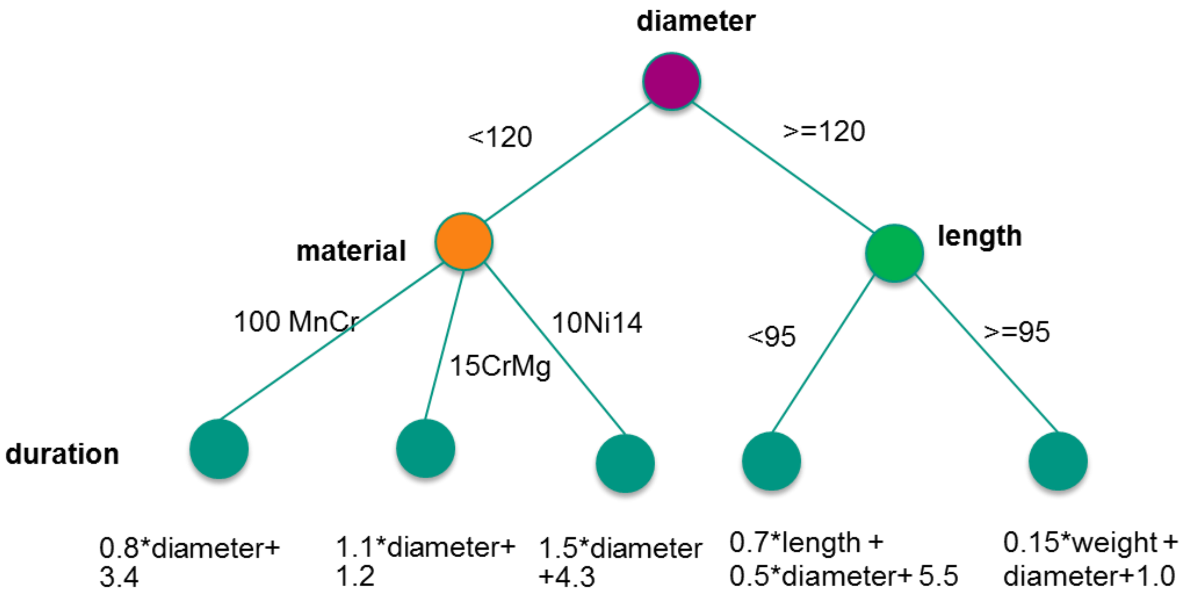


Figure 4.22 An M5P tree representing for the prediction of process duration of a machine

Then, the process cost can be calculated using equation (4.27), where $C(k)$ denotes the cost of process k, $d(k)$ is the predicted duration of process k on machine m_i and Cm_i is the machine cost per time unit. Similarly, equation (4.28) calculates the energy cost of a process which is carried out on a certain machine. $C_E(k)$ represents the energy cost of process k, $P(k)$ is predicted power consumption and c_E denotes the energy cost per kWh.

$$C(k) = d(k) * Cm_i \tag{4.27}$$

$$C_E(k) = d(k) * P(k) * c_E \tag{4.28}$$

Example:

Given the energy tariff (c_E) is 8.6 cents per kWh and the cost of machine M (C_M) is 75 €/hour. The following table lists two products with their properties. The power consumptions, durations, and costs of the process carried out on the machine M are to be predicted.

Table 4.8 Inputs for prediction of typical power consumption, process duration, and costs

Product	Properties			
	material	**weight**	**Diameter**	**length**
P1	100MnCr	100	120	50
P2	15CrMg	200	150	100

The typical power consumption of the process to produce P1 is predicted using the rule no.1 in Table 4.6.

$P(P1) = 0.2 \times 100 + 5.12 = 25.12\ kW$

The time required by M to process P1 is predicted using the rule generated from the decision tree in Figure 4.22 which is as follows:

$d(P1) = 0.7 \times 50 + 0.5 \times 150 + 5.5 = 115.5\ minutes = 1.925\ hours$

The machine cost to process P1 on M is calculated as follows:

$C(P1) = 1.925\ h\ \times 75\ €/h = 144.38\ €$

The energy cost to carry out the process on M can be estimated as follows:

$C_E(P1) = 1.925\ h \times 25.12\ kW \times 0.086€/kWh = 4.15\ €$

The calculation of the typical power consumption of P2 uses rule no. 5 of Table 4.6.

$P(P2) = 0.15 \times 200\ + 0.8 \times 150 - 35.2 = 114.8\ kW$

Similarly, the duration and costs to process P2 on M can be estimated as follows:

$d(P2) = 0.15 \times 100 + 150 + 1.5 = 166.5\, minutes = 2.775\, hours$
$C(P2) = 2.775\, h\ \times 75\, €/h = 208.13\, €$
$C_E(P1) = 2.775\, h \times 114.8\, kW \times 0.086€/kWh = 27.40\, €$

Data mining can also extract knowledge to support the machine selection based on the experience reflected in the historical data. Figure 4.23 gives an example of a decision tree that models a machine selection for forging process. It is shown that the machine selection depends mainly on length, weight, and diameter of the product. It can be seen that the heavier and bulkier (having bigger volume) products should be forged on the 3500-ton press while the lighter ones should be forged on the 800-ton press or the drop hammer. The decision tree is built on relational data containing interrelationships between products, orders, machines and production processes similar to the ones depicted in Figure 4.21.

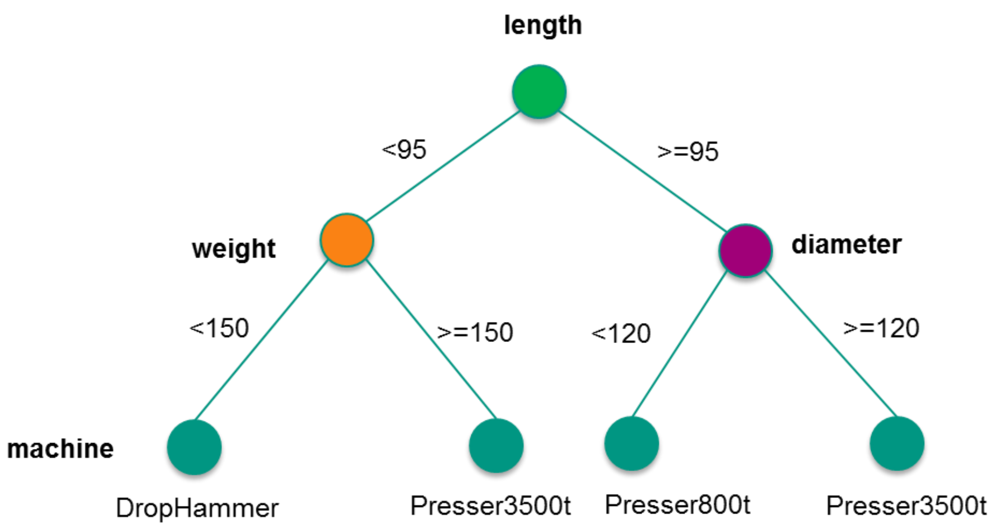


Figure 4.23 A decision tree modeling a machine selection for forging process

Example:

Consider the products P1 and P2 in Table 4.8. The system will recommend the forging process of P1 on DropHammer, whereas the forging process of P2 should be carried out on Press3500t

4.3.4 Integration into Ontology

In order to make the extracted knowledge available for the evaluation towards energy efficiency, it has to be incorporated into the ontology knowledge base (see Figure 4.1, step 4). The extracted rules are converted into SWRL. The rules to detect inconsistencies or anomalies of the power consumption pattern are derived from the rules listed in Table 4.6. An inconsistent power consumption or anomaly is detected if the power consumption value greater than the values within the range indicated in the consequent part of a rule.

Table 4.9 SWRL rules for detecting inconsistent power consumption derived from rules in Table 4.6

No	Rule
1	`Product(?p) ∧ hasWeight(?w) ∧ swrlb:lessThan(?w,150.4) ∧ hasMaterial(?p, '100McCr') ∧ RollingMachine (?m) ∧ ManufacturingJob(?j) ∧ operates (?m, ?j) ∧ produces(?j, ?p) ∧ hasPowerConsumption(?j,?e) ∧ swrlb:greaterThanOrEqual (?e,22) → ProductionFacilityInconsistentEnergyPerformance(?m)`
2	`Product(?p) ∧ hasWeight(?w) ∧ swrlb:lessThan(?w,150.4) ∧ hasMaterial(?p, '15CrMg') ∧ RollingMachine (?m) ∧ ManufacturingJob(?j) ∧ operates (?m, ?j) ∧ produces(?j, ?p) ∧ hasPowerConsumption(?j,?e) ∧ swrlb:greaterThanOrEqual (?e,46) → ProductionFacilityInconsistentEnergyPerformance(?m)`
3	`Product(?p) ∧ hasWeight(?w) ∧ swrlb:lessThan(?w,150.4) ∧ hasMaterial(?p, '10Ni14') ∧ RollingMachine (?m) ∧ ManufacturingJob(?j) ∧ operates (?m, ?j) ∧ produces(?j, ?p) ∧ hasPowerConsumption(?j,?e) ∧ swrlb:greaterThanOrEqual (?e,11.5) → ProductionFacilityInconsistentEnergyPerformance(?m)`
4	`Product(?p) ∧ hasWeight(?w) ∧ swrlb:greaterThanOrEqual(?w,150.4) ∧ hasDiameter(?p, ?d) ∧ swrlb:lessThan(?d, 120) ∧ RollingMachine (?m) ∧ ManufacturingJob(?j) ∧ operates (?m, ?j) ∧ produces(?j, ?p) ∧ hasPowerConsumption(?j,?e) ∧ swrlb:greaterThanOrEqual (?e,70) → ProductionFacilityInconsistentEnergyPerformance(?m)`
5	`Product(?p) ∧ hasWeight(?w) ∧ swrlb:greaterThanOrEqual(?w,150.4) ∧ hasDiameter(?p, ?d) ∧ swrlb:greaterThanOrEqual(?d, 120) ∧ RollingMachine (?m) ∧ ManufacturingJob(?j) ∧ operates (?m, ?j) ∧ produces(?j, ?p) ∧ hasPowerConsumption(?j,?e) ∧ swrlb:greaterThanOrEqual (?e,91) → ProductionFacilityInconsistentEnergyPerformance(?m)`

Consider the rule no. 1 in the table, the statement `weight < 150.4 kg` is converted to the `Product(?p) ∧ hasWeight(?w) ∧ swrlb:greaterThanOrEqual(?w,150.4)`. The variable *?p* denotes a product instance which is evaluated by the rule. The SWRL built-in `swrlb:greaterThanOrEqual` expresses the operator ≥. Similarly, the statement `material='100MnCr'` is expressed as `hasMaterial(?p, '100McCr')`. The expression `RollingMachine (?m) ∧ ManufacturingJob(?j) ∧ operates (?m, ?j) ∧ produces(?j, ?p) ∧ hasPowerConsumption(?m,?e) ∧ swrlb:greaterThanOrEqual (?e,22)` corresponds to `mes_value ≥22`. From the expression, it can be understood that in order to get the power consumption value of the product, the relations between machine and process and between process and product have to be established through the object properties `operates` and `produces` respectively. The value of the object property `hasPowerConsumption` associated to the machine is then evaluated. If the evaluation of the antecedent expression returns true, the machine instance *?m* is assigned as an instance of the class representing inconsistent energy performance through the conclusion expression `ProductionFacilityInconsistentEnergyPerformance(?m)`.

Through a similar procedure, the M5P rules listed in Table 4.7 are also converted to SWRL. Data properties `hasPowerConsumptionFormula` and `hasDurationFormula` are introduced to hold the linear equation for predicting power consumption of a machine and duration of a process executed on a machine by means of product properties (see Figure 4.19 and Figure 4.20). The results of the rule conversion from Table 4.7 are listed in Table 4.10. The table only gives examples of rules for the prediction of power consumption values. Analogously, the rules for predicting process duration can be determined. The data type of the formulas is string because the calculations of the power consumption and duration values have to be performed by not using the rule engine. The drawback of using a string representation is that the attributes involved in the formulas have to be written exactly the same as the corresponding object properties, for example, `hasWeight` and `hasDiameter` in the rules shown in Table 4.10.

Table 4.10 SWRL rules for determining the formula to predict the power consumption derived from rules in Table 4.7.

No	Rule
1	`Product(?p) ∧ hasWeight(?w) ∧ swrlb:lessThan(?w,150.4) ∧ hasMaterial(?p, '100McCr') ∧ RollingMachine (?m) ∧ ManufacturingJob(?j) ∧ operates (?m, ?j) ∧ produces(?j, ?p)→ hasPowerConsumptionFormula(?m,´0.2*hasWeight´)`
2	`Product(?p) ∧ hasWeight(?w) ∧ swrlb:lessThan(?w,150.4) ∧ hasMaterial(?p, '15CrMg') ∧ RollingMachine (?m) ∧ ManufacturingJob(?j) ∧ operates (?m, ?j) ∧ produces(?j, ?p)→ hasPowerConsumptionFormula(?m, ´0.45*hasWeight´)`

3	`Product(?p) ∧ hasWeight(?w) ∧ swrlb:lessThan(?w,150.4) ∧ hasMaterial(?p, '10Ni14') ∧ RollingMachine (?m) ∧ ManufacturingJob(?j) ∧ operates (?m, ?j) ∧ produces(?j, ?p)→ hasPowerConsumptionFormula(?m,´0.11*hasWeight´)`
4	`Product(?p) ∧ hasWeight(?w) ∧ swrlb:greaterThanOrEqual(?w,150.4) ∧ hasDiameter(?p, ?d) ∧ swrlb:lessThan(?d, 120) ∧ RollingMachine (?m) ∧ ManufacturingJob(?j) ∧ operates (?m, ?j) ∧ produces(?j, ?p) →hasPowerConsumptionFormula(?m,´0.2*hasWeight + 0.5*hasDiameter´)`
5	`Product(?p) ∧ hasWeight(?w) ∧ swrlb:greaterThanOrEqual(?w,150.4) ∧ hasDiameter(?p, ?d) ∧ swrlb:greaterThanOrEqual(?d, 120) ∧ RollingMachine (?m) ∧ ManufacturingJob(?j) ∧ operates (?m, ?j) ∧ produces(?j, ?p) → hasPowerConsumptionFormula(?m, ´0.15*hasWeight + 0.8*hasDiameter´)`

4.4 Development of Energy Performance Indicators

In this section, the developed concept of energy performance indicators (EnPI) to assess energy efficiency in manufacturing is described. The aim of the concept development is to fulfill the requirements formulated in Section 3.3. Figure 4.24 depict the methodological steps to determine the EnPI.

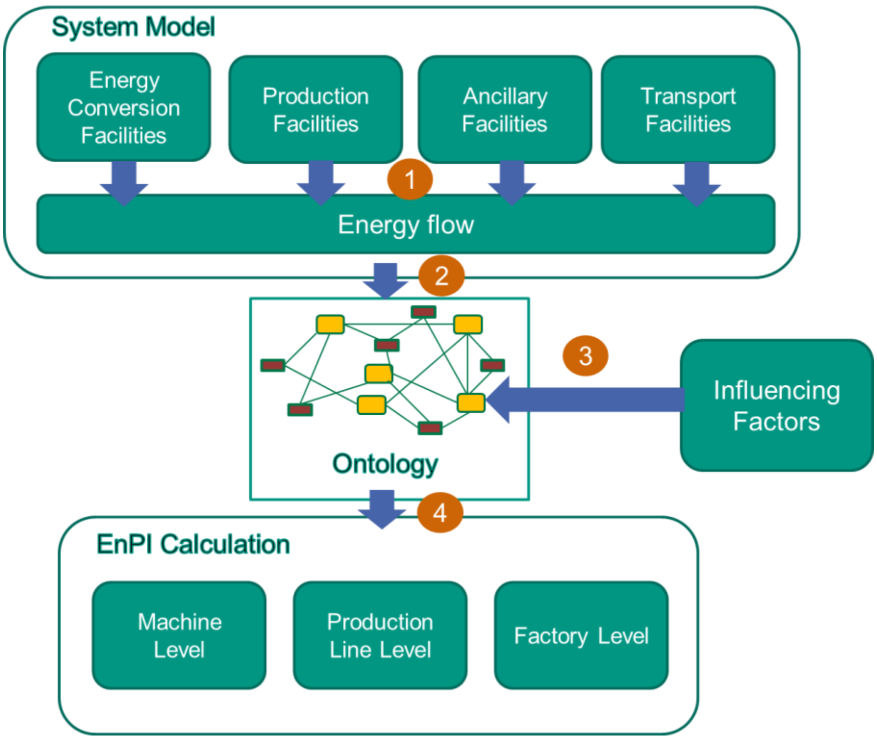


Figure 4.24 Overview of EnPI development approach

According to Figure 4.24, energy conversion, production, ancillary, transport facilities and the energy flowing from and to them are modeled first. This will result in an energy flow model of the production system. Section 4.4.1 elaborates the approach in developing the model. In the second step, the model is represented in the ontology knowledge base which has been described in section 4.2. This reflects the interaction between ontology and quantitative evaluation (EnPI) module depicted in Figure 4.1 step 7. The third step is to define influencing factors in the ontology which contribute to the energy inputs and outputs modeled in the energy flow. The factors are listed in section 4.4.2. As seen in the last step in Figure 4.24, the EnPI calculation uses the system model located in the ontology to calculate the EnPI at the machine, production line, and factory level. The calculation approaches for those different levels are presented in Section 4.4.4, 4.4.5, and 4.4.6 respectively.

4.4.1 System Modeling

In order to develop the EnPI of a single manufacturing company, it is important to define the system boundaries and a model that covers the different organization levels of the manufacturing organization first. According to the requirements formulated in Section 3.3, the developed model has to consider the application in the machine or production facility, production line, and factory level. In this thesis, UPNT model is developed. The UPNT model is an extension of the UPN model developed by Schieferdecker and Fünfgeld et al. [Fünf-00] [ScFB-06]. The basic elements of the UPNT model are energy conversion (German: *Umwandlungsandlage*, U), production (German: *Produktionsanlage*, P), ancillary (German: *Nebenanlage*, N), and transport (German: *Transportanlage*, T) facilities. Only through the interaction of these facilities production on an industrial scale is possible. In this thesis, an additional component namely transport facilities (T) is introduced, to represents the transport means for intralogistics. These facilities are also in principle ancillary facilities, but they do not directly serve other facilities. They transport goods between production facilities. Therefore, they are considered separately. The following sections explain each considered facility including energy flow of each facility and methods to calculate the energy balance.

Energy Conversion Facility (U)

Energy conversion facilities convert energy and provide it in a processed form to the other facilities (N, T, P, as well as U). The energy conversion facilities most likely do not directly involved in production processes. Examples of energy conversion facilities include transformators, air compressors, steam generators, and power plants. On an energy conversion facility, an energy balance can be considered. The efficiency can be specified or calculated. An energy conversion facility can receive energy from external sources such as utility company or other installed conversion facilities. It gives output energy that can be consumed by production, ancillary, transport, as well as other energy conversion facilities. The energy loss emerged from the energy conversion facility is also considered in the balance calculation. Figure 4.25 depicts energy balance of an energy conversion facility. It should be noted that the energy balance model is valid for different energy form, for instance, electrical energy, gas, pressure, etc.

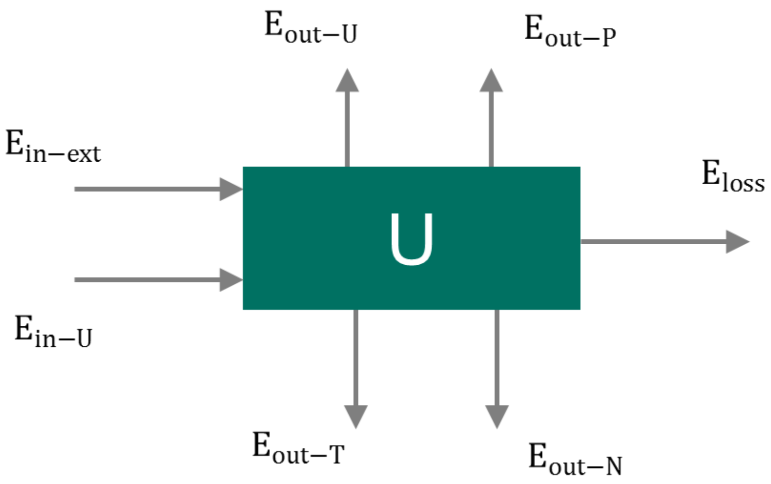


Figure 4.25 Energy balance of an energy conversion facility

Energy balance for an energy conversion facility is mathematically expressed in (4.29).

$$E_{in,ext} + E_{in,U} = E_{out,U} + E_{out,P} + E_{out,N} + E_{out,T} + E_{loss} \quad (4.29)$$

where

$E_{in,ext}$	=	total input energy from external sources, such as from utility companies
$E_{in,U}$	=	total input energy from internal energy conversion facilities
$E_{out,U}$	=	total output energy consumed by other connected conversion facilities
$E_{out,P}$	=	total output energy consumed by connected production facilities
$E_{out,N}$	=	total output energy consumed by connected ancillary facilities
E_{outT}	=	total output energy consumed by transport facilities
E_{loss}	=	energy loss of the conversion facility

The total input or output energy of each facility $E_{in,U}, E_{out,U}, E_{out,P}, E_{out,N}, E_{out,T}$ can be calculated using equation (4.30). α_i denotes the ratio of energy received/given by the facility i, and $E_{i,X}$ represent the output/input energy of facility i.

$$E_{in/out,X} = \sum_{i=1}^{n} \alpha_i E_{i,X} \ , X \in \{U, P, N, T\} \quad (4.30)$$

Example:

Let us assume that the electrical energy distribution of a factory electrical energy network can be measured and regulated. In a certain time, for example, 8 hours, a transformator U1 as an energy conversion facility takes 200 kWh of electrical energy from a utility company and 30 kWh from solar energy generation system. Table 4.11 shows the facilities connecting to U1 and their input energy.

Table 4.11 Consumption distribution of energy generated by U1

Facility f	Total input energy $E_{in,f}$ (kWh)	Energy supply α_f (%)	
		From U1	From other sources
P1	125	80	20
P2	200	40	60
N1	100	20	80
T1	70	30	70

The total input and output energy of U1 can be calculated as follows:

$$E_{in} = 200\ kWh + 30\ kWh = 230\ kWh$$

$$E_{out} = 80\% \times 125\ kWh + 40\% \times 200\ kWh + 20\% \times 100\ kWh + 30\% \times 70\ kWh = 221\ kWh$$

The difference of 9 kWh between $E_{in,U1}$ and $E_{out,U1}$ is the energy loss (E_{loss}).

Production Facility (P)

Production facilities are the most important entities in the production because they are used to manufacture the product and generate the value. The production facilities depend on functioning conversion, ancillary, and transport facilities to work. A production facility consumes energy directly from external sources and energy given by conversion facilities. The indirect energy given to the production facility is also considered in defining the energy balance. It includes energy coming from ancillary facilities and transport facilities. For example, if lighting is used to support the functionality of production process executed on the production facility or workstation, the energy effort to generate the light is considered in the balance calculation. This also applies to transport facilities that carry goods to the production facility. The input energy is converted to another form of energy that is used to execute production processes. Furthermore, energy loss should be taken into consideration. The energy balance modeling of a production facility is shown in Figure 4.26.

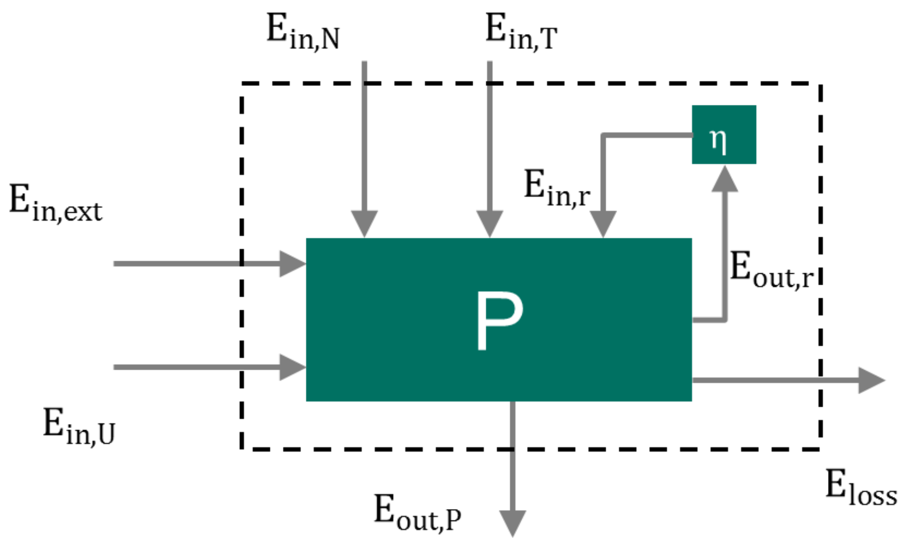


Figure 4.26 Energy balance of a production facility

Mathematically, the energy balance of a production facility is formulated with the following equation:

$$E_{in,ext} + E_{in,U} + E_{in,N} + E_{in,T} + (E_{in,r}) = E_{out,P} + (E_{out,r}) + E_{loss} \tag{4.31}$$

where

$E_{in,ext}$ = total input energy from external sources, such as from utility companies

$E_{in,U}$ = total input energy from internal energy conversion facilities

$E_{in,N}$ = total energy received from ancillary facilities indirectly

$E_{in,T}$ = total energy indirectly received from transport facilities carrying goods to the production facility

$E_{in,r}$ = total recovered energy

$E_{out,P}$ = total energy converted and used to execute production process

$E_{out,r}$ = recoverable energy

E_{loss} = energy loss of the conversion facility

Example:

Similar to previous example, the energy consumption and distribution in a factory energy network can be measured and regulated. For the measurement period of 8 hours, the production facility P1 gets the input energy of 140 kWh from an external source and 50 kWh from the energy conversion facility U1. The total estimated energy consumption of lighting needed to make sure that activities related P1 work properly is 20 kWh. The total required energy to transport goods to P1 within the period is 30 kWh. Therefore, the total energy efforts of P1 within the period can be computed as follows:

$$E_{in} = 140\ kWh + 50\ kWh + 20\ kWh + 30\ kWh = 240\ kWh$$

It can be seen that the energy efforts of P1 do not come only from external sources i.e. utility companies, but also from transport and ancillary facilities i.e. lighting systems that support the production processes on P1.

Each input and output energy can be obtained by aggregation of different facilities and sources that contribute to the facility. It can be calculated using (4.30). The energy coming from utility companies and conversion facilities is mainly transformed to energy utilized for producing the product. The potentially harvestable output energy might also be exploited by the production facility to get additional energy. In this work, the energy harvesting in a production facility is viewed as a black box. Because the calculation of EnPI requires an energy flow model of in the system and involves different relations between conversion, production, ancillary, and transport facilities, the energy recovery is not considered in the calculation of EnPI.

Ancillary Facility (N)

Ancillary facilities provide the necessary conditions to achieve an unobstructed production. They serve the production directly or operating condition generally. Unlike in the UPN model, in UPNT model, ancillary facilities are not only understood as lighting, air conditioning, and similar systems but also as all facilities that allow the execution of production processes. The examples include coolant pumps, filtration systems, chip conveyors, exhaust systems. Facilities that are not located directly in production, such as kitchen appliances in the canteen are also considered as ancillary facilities, but these are not always the focus in the EnPI calculation.

Similar to the production facility, an ancillary facility obtains energy from external sources such as utility companies and from installed energy conversion facilities. The output energy is given to production facilities in order to ensure the functionality of the executed production process, and to other facilities that are not related to the production process. Figure 4.27 illustrates energy flow model of an ancillary facility to determine the energy balance. The corresponding energy balance equation is expressed in (4.32).

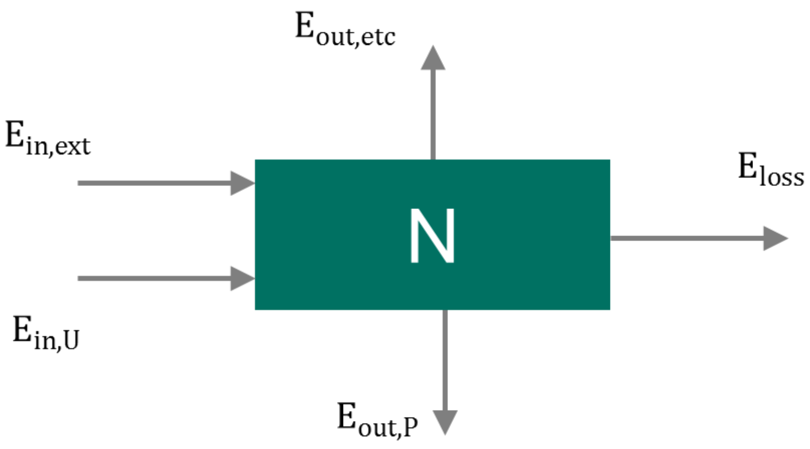


Figure 4.27 Energy balance of an ancillary facility

$$E_{in,ext} + E_{in,U} = E_{out,p} + E_{out,etc} + E_{loss} \tag{4.32}$$

where

$E_{in,ext}$	=	total input energy from external sources, such as from utility companies
$E_{in,U}$	=	total input energy from internal energy conversion facilities
$E_{out,P}$	=	total output energy used to support production process
$E_{out,etc}$	=	total output energy used for other purposes
E_{loss}	=	energy loss of the conversion facility

The energy outputted by an ancillary facility can be distributed to several production facilities. For example, a lighting system installed in an area of production hall illuminates the production facilities located nearby. Like in production and energy conversion facility, the total input and output energy can be determined using (4.32).

Example:

The lighting system of a production hall has total input energy of 100 kWh over an 8 hours period. The total input energy is 80% from the installed solar cell system ($E_{in,U}$) and 20% from external sources ($E_{in,ext}$). The estimated total energy used to illuminate the production area containing production facilities is 70 kWh. Around 30 kWh from the lighting is estimated for illumination of the offices and rooms for monitoring the production processes.

Ancillary facilities most likely do not have the capability to recover energy [Schi-87] [Stor-89] [Fünf-00]. Therefore, the energy recovery is not considered in the calculation of energy balance of an ancillary facility. The usage of ancillary facilities is very specific based on the facilities that are associated with them. Therefore, it is not possible to describe them in a general way. For an individual ancillary facility, it is possible to measure the energy consumption. A generally ascertainable figure is, however, difficult to define. Apart from that, ancillary facilities do not affect the production process directly. Therefore in this thesis, no ancillary facility related figure is developed. It makes more sense to ascribe the energy effort of ancillary facilities to energy conversion and production facilities. If several facilities operate by taking benefit of an ancillary facility together, then the cost corresponding to the proportion of the benefits will be distributed to those facilities. The purpose here is to represent the energy effort with cost and not with an energy quantity.

Estimating Energy Consumption of Lighting

One of the most important ancillary facilities in a factory is lighting. It determines the quality of the production processes, especially production process involving personnel. Without lights, the workers are not able to operate the machine. The energy consumption of the lighting that is used further by production or conversion facilities depends on the area that is illuminated by the lighting and the distance where the lighting is installed. It also depends on the type of the lighting and the desired illumination level. It is difficult and costly to measure the energy consumption of lighting. However, it is possible to estimate the power consumption of lighting using (4.33) [Pist-09, p.K19].

$$P_{lighting} = \kappa_t \, j_{\widetilde{E_m},h} \, A \tag{4.33}$$

where

$P_{lighting}$ = power consumption lighting [W]

κ_t = adjustment factor of lighting type t

$j_{\widetilde{E_m},h}$ = electrical power of lighting with desired luminance level $\widetilde{E_m}$ and installation height h [W/m^2]

A = illuminated area [m^2]

Example:

A halogen bulb used to illuminate an area of 100 m^2 in production hall containing the production facility P1. The bulb is placed 4 m height from the ground. The desired illumination for the area ($\widetilde{E_m}$) is 500 lx. According to Table C.1 in Appendix C, the corresponding electrical power of lighting for the 4 m height and desired illumination of 500 lx is 13 W/m^2. A halogen bulb has adjustment factor κ_t of 5.0 (see Table C.2 in Appendix C). The power consumption of the halogen bulb can be estimated as follows:

$$P_{halogenbulb} = 5 \times\ 13\frac{W}{m^2} \times 100\ m^2 = 6500\ W$$

If the bulb operates 24 hours non-stop, it will consume 156 kWh energy.

Estimating Energy Consumption of Heating

The next important ancillary facility in manufacturing is heating system. Compared to the heating system in household and office buildings, the production hall has particular characteristics as follows [BDEW-04]:

- Production halls normally contain a large room without any partitions and basements. This lead to low temperature on the outer walls and higher heat losses
- Production halls normally have large doors, windows, and a high roof that cause heat losses. In addition, some old production halls are leaky and have inadequate insulations.
- Production processes, facilities, and storages require special heating conditions, such as particular humidity and air exchange rates
- Production processes could emit extra heat
- Production processes could produce harmful gas emission that requires a good ventilation

It is not always possible to measure the energy consumption of heating systems. Nevertheless, there are several methods to estimate the energy consumption of heating systems in manufacturing. The heating load of production hall is calculated according to the DIN 4701 and DIN EN 12 831. The heating load is understood as the amount of heat required to produce a standard indoor temperature. It is used to measure the maximum heating power [BDEW-04, p.20].

The heating load depends on the characteristics of the production hall. It depends on the dimension of the hall (length, width, and height), surfaces of different building elements, such as outer wall surface, outer windows, outer doors and roof, and the heat transfer coefficient. In addition, it also depends on the average inside temperature, outside temperature, the floor temperature, and the required air exchange rate. The heating load φ_{HL} is calculated as the sum of transmission heat loss φ_T and ventilation heat loss φ_V (see (4.34)). Transmission heat losses are caused by the heat conduction through surfaces that connect a closed room to the outside or to the neighboring rooms having a lower temperature. The ventilation heat losses are made through ventilation to the outside space or by infiltration through the building envelope between the heated space and adjacent rooms [Pist-09, p. H22]. The heat delivered by the production machines and humans who work in the hall could also be considered to contribute to the heating load. A hard-working person could produce heat up to 500W [Pist-09, p.H6].

$$\varphi_{HL} = \varphi_T + \varphi_V \tag{4.34}$$

The transmission heat loss that goes outside through walls, doors, windows, etc. is calculated using the equation (4.35). Further equations to calculate heat losses from floors to adjacent rooms, etc. in a more advanced way can be found in [Pist-09, p. H26-30].

$$\varphi_T = A\,U(T_i - T_o) \tag{4.35}$$

where

A	=	surface area conducting the heat [m^2]
U	=	thermal transmittance coefficient [W/m^{2}K]
T_i	=	inside temperature [K]
T_o	=	outside temperature or temperature of neighbouring rooms [K]

Example:

A production hall has four well-insulated walls (U = 0.25 W/m^{2}K) having a total surface area of 600 m^2 and triple glazed windows (U = 0.8 W/m^{2}K) with the total surface are of 40 m^2. The thermometer placed inside the hall shows the value of 20 C or 293.15 K, whereas outside temperature is 10 C or 283.15 K. The transmission heat loss of the walls and windows can be calculated

$$\varphi_T = \left(600\ m^2 \times 0.25\frac{\mathrm{W}}{\mathrm{m^2K}} + 40\ \mathrm{m}^2 \times 0.8\frac{\mathrm{W}}{\mathrm{m^2K}}\right) \times (293.15\ K - 283.15\ K) = 1.82\ kW$$

The transmission heat losses through the roof and ground have to be calculated separately.

Equation (4.36) estimates the amount of ventilation heat loss.

$$\varphi_V = n\,V\,\rho\,c\,(T_i - T_o) \tag{4.36}$$

where

n	=	air exchange rate [1/h]
V	=	volume of heated room [m^3]

ρ = air density [kg/m^3]

c = specific heat capacity of air [J/kg K]

T_i = inside temperature [K or C]

T_o = outside temperature or temperature of neighboring rooms [K or C]

and table

$\rho\ c \approx 0.34$ Wh/m^3.K

Example:

Consider the production hall from the previous example (room volume 6400 m^3) having an air exchange rate of 1.2/h. The inside and outside temperatures are 20 C and 10 C respectively. The ventilation heat loss can be estimated as follows:

$$\varphi_V = 1.2/h \times 6400\ m^3 \times 0.34\ Wh/m^3.K\ \times (293.15\ K - 283.15\ K) = 26.11\ kW$$

The energy consumption of heating system in production hall depends on the heating consumption period and the efficiency degree of the heating system itself. It can be estimated with equation (4.37).

$$E = \frac{\varphi_{HL}\, t}{\eta} \tag{4.37}$$

where

φ_{HL} = heating load [W]

t = heating consumption period [h]

η = efficiency of heating system

Example:

The heating system of the production hall has a total efficiency of 0.85. By using the results of both previous examples, the energy consumption of the heating for 24 hours operation excluding transmission loss through the floor, roof and doors can be estimated as follows:

$$E = \frac{(1.82\ kW + 26.11\ kW) \times 24h}{0.85} = 788.61\ kWh$$

Furthermore, according to [Pist-09][VDI-07], the fuel demand to produce the heat can be estimated by taking into account the energy source specific coefficient. For example to produce 1 kWh heat energy, it needs 1 kWh electrical energy, and to make 10 kWh heat energy, it requires 1 Litre fuel oil. It is expressed in (4.38).

$$D_F = \frac{E}{H_i} \tag{4.38}$$

where

D_F	=	number of required fuel [unit, unit ∈ $\{l, kg, m^3, kWh, \ldots\}$]
E	=	total required energy for heating [kWh]
H_i	=	energy source specific coefficient [kWh/unit]

Transport Facility (T)

Principally, transport facilities can be included as ancillary facilities. However, they differ, because although they are directly related to the production process, they are not used directly by any production facilities. Rather, they play roles between two production facilities by transporting products to the next process step in order to be processed further. It is very important, in particular, if the manufacturing company has different sites. Transport facilities are one of the essential factors to be considered for energy efficiency. Although the main objective is to minimize the energy consumption of the transportation, it is still important to keep a "just in time" production, in which the material is delivered in the required quantities and at the right time, as it is used in the production process. As shown in Figure 4.28, an external energy source and connected energy conversion facilities feed energy to a transport facility, for example, an electrical fork-lift truck. The obtained energy is transformed into energy used for transporting goods to production facilities. The energy is considered as energy effort or input of production facilities. The energy balance of a transport facility is determined using equation (4.39).

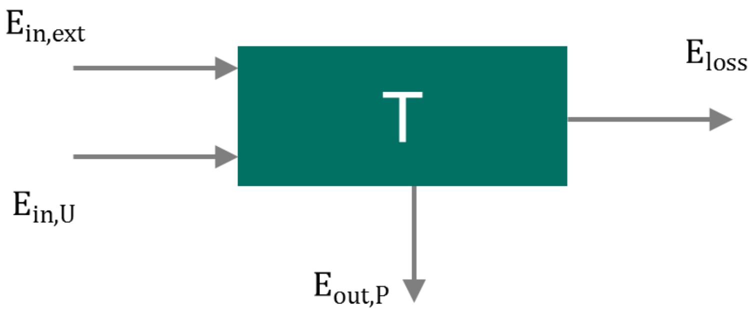


Figure 4.28 Energy balance of a transport facility

$$E_{in,ext} + E_{in,U} = E_{out,P} + E_{loss} \tag{4.39}$$

where

$E_{in,ext}$	=	total input energy from external sources, such as from utility companies
$E_{in,U}$	=	total input energy from internal energy conversion facilities
$E_{out,P}$	=	total output energy used to transport goods to production facilities P
E_{loss}	=	energy loss of the transport facility

Example:

An electric forklift T1 transports the materials from a container to several production facilities P1, P2, and P3. The energy demand of T1 is measured in its charging station. After 60 minutes charging, T1 has a maximum battery power of 35 kW. The battery is then used to operate the forklift for 24 hours. The charging station receives 15 kW from an internal source ($E_{in,U1}$) and 20 kW from an external network ($E_{in,ext}$). The total input energy of T1 is then

$$E_{in} = 35\ kW \times 24\ h = 840\ kWh$$

Within 24 hours, T1 travels around 50% to P1, 40% to P2, and 10% to P3. The energy efforts of P1, P2, and P3 from transport tasks are estimated to 420 kWh, 336 kWh, and 84 kWh respectively.

The energy consumption is affected by the distance between the origin and destination. A forklift is mostly used to overcome the different height of origin and target facilities. Additionally, the energy is lost due to the heat created by friction and deceleration (braking). Due to the greater energy loss, the efficiency degree of transport facility becomes lower. The energy recovery technology can be applied to overcome this, but it does not always give economic benefit. The objective of using transport facilities is to overcome distances. It is not meant to create an adding value, but to transport goods to the next production facilities where the adding-value production processes take place. Therefore, in this thesis, the energy efficiency figure of transport facility is not calculated. The energy cost of the facility is ascribed as energy effort of the target production facility. From a certain origin, the transport effort depends solely on the position of target production facility. The closer the destined production facility, the lower the energy required for the transport.

Like ancillary facilities, it is not always possible to measure the energy consumption of manufacturing transport facilities, but the energy consumption can still be estimated. The most common transport facility in manufacturing is forklift truck. A forklift truck is considered as non-continuous conveyor, because in transporting the goods, it travels both loaded to carry the goods and unloaded to pick up the goods. It is considered to be flexible because its tasks and services can be customized. However, human labor is still needed, which results in higher operating costs. A forklift truck might have either an electric motor or operate using a combustion engine. For indoor usage, an electric forklift is well suited since it emits no pollutants. There are assumptions that should be made to estimate the energy consumption of a forklift truck. For example, the braking force that has to be used to brake the truck from its constant velocity to zero is ignored. In addition, the energy consumed for lowering the load is negligibly small compared to the energy needed for lifting. The truck is assumed performing four steps, where each step has a different method to estimate the energy consumption. The steps are (1) lifting the goods; (2) starting to move and carrying the goods until it stops; (3) releasing the goods; (4) moving without carrying the goods, running with constant speed until it stops

Energy Flow of a Production System

Supposed a production system consists of a production facility, an energy conversion, transport, and an ancillary facility. Based on energy balance model described in previous sections and production system UPN model developed by Schieferdecker and Fünfgeld et al. [Fünf-00] [ScFB-06, p.71], the energy balance of the production system can be illustrated in Figure 4.29. The dashed line represents the production system boundary. The energy given to the system is from external sources, such as utility company (UC). The energy that released from the system is the energy used to perform production process and energy loss. Equation (4.40) shows the method to assess the energy balance of production system defined in Figure 4.29.

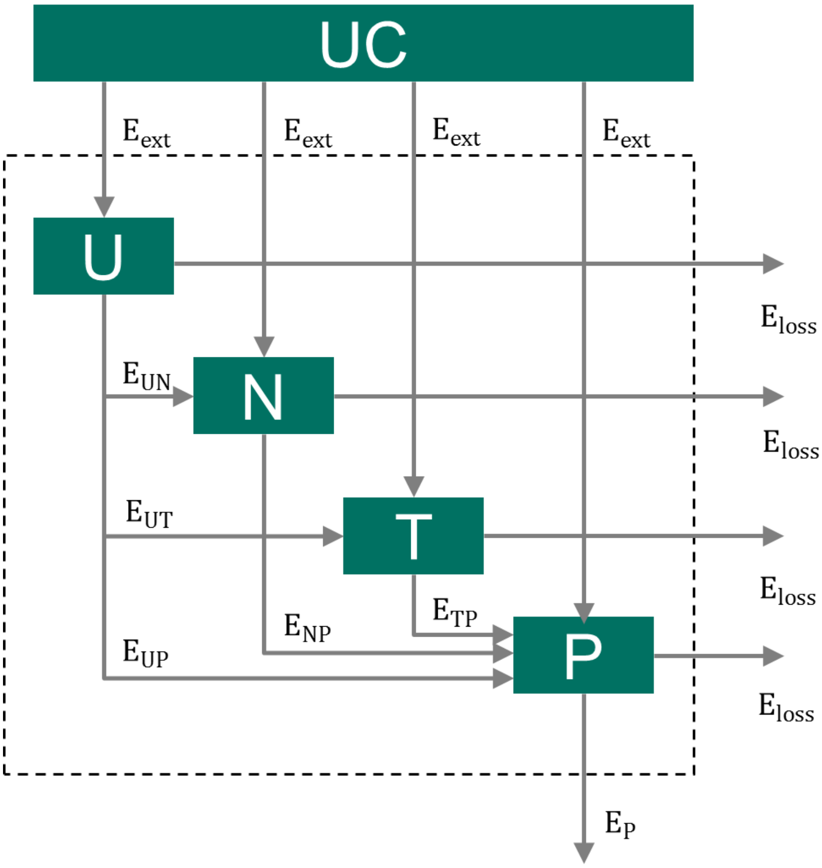


Figure 4.29 Energy balance of a production system

$$\sum_{i=1}^{m} E_{ext,i} = E_P + \sum_{j=1}^{m} E_{loss,j} \tag{4.40}$$

The method as described in Figure 4.29 and equation (4.40) is basically applicable for any production systems. The precondition is that the system boundary has to be clearly defined first. Therefore, the method is promising to be applied at different levels of production system organization. As shown in the Figure 4.29, it can be concluded that the energy consumption of an organization unit of production, e.g. factory, production line, etc., does not depend on only the energy consumption of production facilities or machines. It also considers the energy conversion, ancillary, and transport facilities included in the organization unit.

4.4.2 Influencing Factors

As the basic in calculating energy efficiency degree, the standard definition of efficiency described in (2.1) is used in this thesis. After building the energy flow model and energy balance calculation method, the factors influencing each of output and input energy have to be found out.

In this work, there are many factors that are considered influencing the input and output in our energy flow model. They are listed and described in the following tables. Table 4.12 describes the factors that influence the output of the system used in our work, whereas Table 4.13 shows the factors that are considered in determining the input of the system.

Table 4.12 Factors influencing the output

Influencing Factor	**Description**
Product quantity and characteristics	The influencing factors are not only product quantity, but also the qualitative characteristics of the products. Both of the influencing factors should be integrated as a unified parameter.
Added value	The generated added value may not be known for each stage of production directly. It can be calculated by considering the values of incoming to outgoing products.
Output energy	The output of conversion facility is the amount of generated energy that is used in the production.
Service and maintenance	The ancillary facilities are used to keep the production process working correctly.
Distance traveled between production facilities	In production, transport facilities are needed to carry the products or materials between production facilities. Transport facility itself is not an adding value since it does not change the product value.

Table 4.13 Factors influencing the input

Influencing Factor	Description
Energy demand	The energy demand corresponds to the energy effort in executing a production process. It can also be easily measured. It is calculated at a certain time as the integral of power over time.
Energy form	Various energy forms have different characteristics, application areas, and costs.
Energy price	The energy price takes into account all costs incurred in the generation of every form of energy.
Average energy load	It plays a role in calculating the peak load if only the additional load over the average load is taken into account. It can also be calculated as the ratio between the total energy consumption and the measurement time.
Energy peak load	The peak energy load plays a role in the costs incurred in providing the energy to the company. It corresponds to which amount of power should be provided to the company in a certain time in order to ensure that the production activities are still running properly
Emission certificate efforts	Companies are demanded to get emission certificate because of the pollution and climate change issue. The costs to achieve emission certificate is considered in the calculation.

4.4.3 Ontology Model for EnPI Calculation

To represent the energy flow model and influencing factors described in section 4.4.1and 4.4.2, the instances of the following classes are created in the ontology knowledge base. The ancillary facilities are represented as instances of the class `AncillaryFacility` which is a sub-class of `BuildingElement`. For the production and transport facilities, instances of the classes `ProductionFacility` and `TransportFacilities` are created. Both classes are subclasses of `ManufacturingResource`. To model the energy conversion facilities instances of the class `ConversionFacility` are generated. `ConversionFacility` is a subclass of `EnergySource`.

For the different production organization levels, where the EnPI calculation takes place, instances for the classes `Workplace` (refers to workstation or machine level), `ProductionLine` and `Factory` are created. They are subclasses of `Enterprise`. Additionally, the class `Flow` and the subclass `EnergyFlow` are created to model any kind of flows including energy flow. The corresponding instances are created to represents the energy flow.

The following properties are set in the created instances in order to represent the relations, for instance, energy flow, and enterprise structure.

- `hasSubNodes` relates the instances of `Enterprise` to build an enterprise hierarchy, for example, a factory consists of several production lines, Each production line is a set of workstations.
- `hasEnergySources` connects the instances of `Enterprise` to `EnergySource` instances, including instances of `ConversionFacility`. It means that the facilities are located in the enterprise.
- `hasBuildingElements` links the instances of `Enterprise` to the instances of `BuildingElement` which includes instances of `AncillaryFacility`.
- `hasResources` connects an `Enterprise` instance to instances of `ManufacturingResources` including production and transport facilities.
- `hasTransportTarget` represents the targets of a transport facility which are production facilities
- `supplies` represents the connection reflecting energy usage and supply relationships resulting the energy flow

4.4.4 EnPI Calculation on Machine Level

The basic principle to develop the energy efficiency figure is as in (2.1) that defines the efficiency as the ratio of output and input. The purpose of the EnPI calculation in machine or workstation level is to support the production planning activities in choosing the most efficient workstations. A workstation is defined as the unit in a production system that includes a set of production resources (personals, materials, machines or production facility, and information). A workstation is generally assigned to a manufacturing operation or process [Neuh-01]. In this work, only machine/production facility is considered to calculate the EnPI since it is only possible to collect the energy consumption of machines.

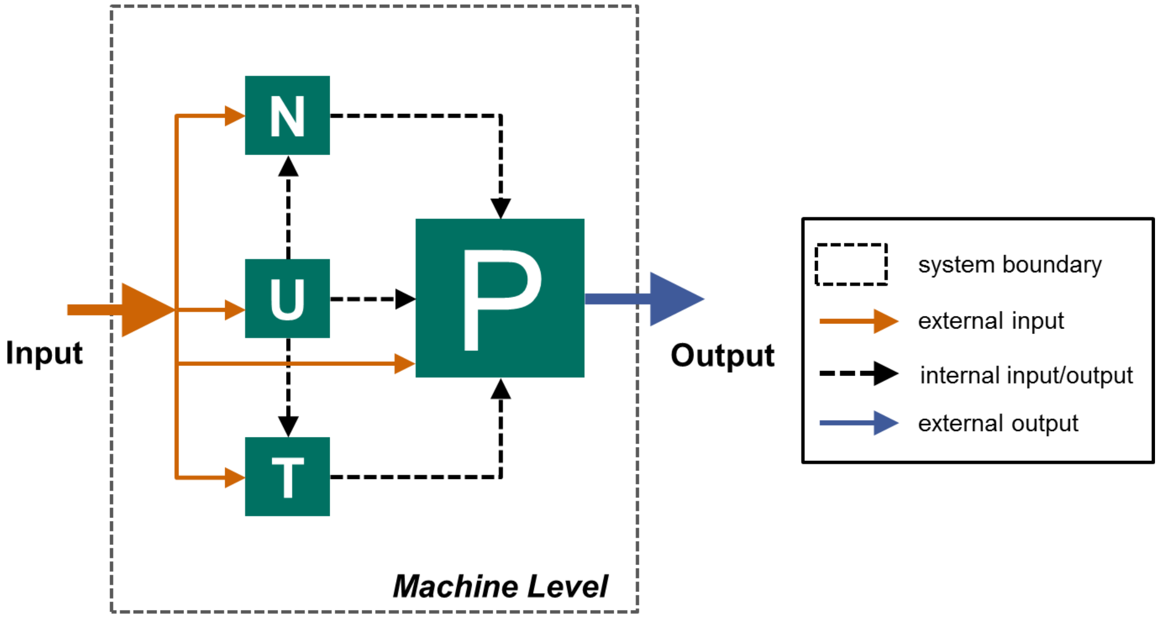


Figure 4.30 Energy flow representing input and output in machine level

As shown in Figure 4.30 the production facility (P) depends on other facilities. The inputs from other facilities (U, N, and T) are taken into consideration in calculating the energy efficiency of production facility since overall energy input and output need to be examined. First, the output and input of the energy conversion facility (U) are examined. After that, the energy flow related to the transport facility (T) and ancillary facility (N) is analyzed. Finally, the efficiency of production facility can be calculated.

Energy conversion facility converts energy to make it available to other facilities. This means that it receives energy as input, converts it, and gives it as output. Therefore, the figures to determine the efficiency are efficiency and utilization ratio as described in section 2.4.1. For further consideration, it is necessary to calculate the costs of input and output energy (C_{Ei} and C_{Eo}). As shown in Figure 4.31 the costs of energy consumption should be fully apportioned to the output. The total value of all inputted energy and the total value of all outputted energy must be equal. It is also valid for the costs incurred from energy load. Equations (4.41) and (4.42) calculate the cost balance of inputted and outputted energy.

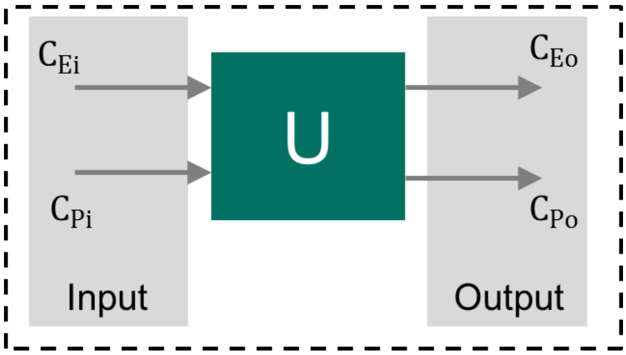


Figure 4.31 Cost balance in an energy conversion facility

$$E_i \times c_E = E_o \times c_E \tag{4.41}$$

$$P_i \times c_P = P_o \times c_P \tag{4.42}$$

where

E_i = input energy [kWh]

E_o = output energy [kWh]

P_i = input load [kW]

P_o = output load [kW]

c_E = energy consumption cost per unit [€/kW]

c_P = energy load per unit [€/W]

It is obvious that an efficiency degree should be less than 1, which is manifested in the form of losses. It causes the cost per unit of outputted energy greater than inputted energy. However, the cost of inputted energy is lower if the conversion facility uses surrounding free energy, such as heat pump. In this work, various generated energy sources/ types are assigned to the costs accordingly and proportionally.

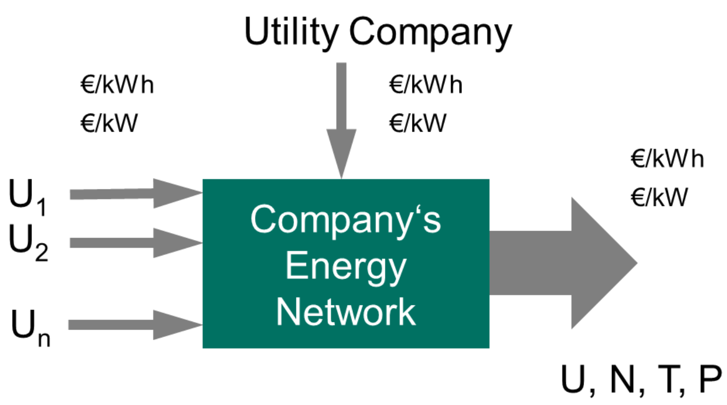


Figure 4.32 Method for calculating energy costs in company's internal energy network

If an energy conversion facility is connected to internal energy network, where other sources including the ones from utility companies are connected, mix-costs from different sources is resulted in the network. The cost is calculated proportionally based on these different sources. Figure 4.32 illustrates the method to calculate the mix costs in the company's internal energy network. The energy mix-cost is calculated using equation (4.43).

$$c_{mix} = \frac{\sum_{i=1}^{n} E_{Ui}\, c_{Ui} + E_{ext} c_{ext}}{\sum_{i=1}^{n} E_{Ui} + E_{ext}} \tag{4.43}$$

where

E_{Ui} = energy generated by energy conversion facility *i* [kWh]

E_{ext} = energy taken from external sources (utility company) [kWh]

c_{Ui} = energy cost per unit of energy conversion facility *i* [€/kWh]

c_{ext} = cost per unit of energy coming from utility company [€/kWh]

c_{mix} = mix cost [€/kWh]

Example:

The internal electricity network of a manufacturing company gets 350,000 kWh from a utility company (E_{ext}), 50,000 kWh and 100,000 kWh from internal energy source U1 (E_{U1}) and U2 (E_{U2}) in a period of 1 April 2014 to 31 March 2015. The tariff of the energy from the utility company (c_{ext}) is 8.6 cents per kWh. The costs per kWh of U1 ($c_{U,i}$) and U2 (c_{Ui}) are 5.1 cents and 4.9 cents respectively. The mix energy cost of the internal network can be calculated as follows:

$$c_{mix} = \frac{50{,}000\ kWh \times 5.1\ cents/kWh + 100{,}000\ kWh \times 4.9\ cents/kWh + 350{,}000 \times 8.6\ cents/kWh}{50{,}000\ kWh + 100{,}000\ kWh + 350{,}000\ kWh} = \mathbf{7.51\ cents/kWh}$$

The main function of production facilities or machines is to create a new product or product part from source materials. As mentioned in Section 4.4.2, it is not appropriate only to consider products for evaluating energy efficiency. The strong consideration of product type is shown in EnergyStar EnPI (see Section 2.4.2). Each product type, for example, cement, car, paper, etc., has its own method to calculate the EnPI. By using EnergyStar EnPI, it is only possible to compare different factory with the similar manufacturing process. In this thesis, an EnPI that does not depend on the product type and manufacturing process is developed. This can be achieved by describing the output of the production facility as the economic value generated by it. The value takes into account the source materials that are fed to the production facilities, the maintenance and personal costs associated with the production facility, as well as the type and characteristics of the product. As an example, if particleboard manufactured from waste wood, the output value is greater than the value when they are made from fresh wood. The characteristic of the product is also taken into account in the value. So, for example, the defined dimensional tolerances and surface quality directly affects the value of the product and thus the value of the process.

As illustrated in Figure 4.30, the input of a production facility is composed of direct and indirect input. The direct input consists of energy [kWh] and power load [kW] that is fed to the production facility from external sources or conversion facilities. The indirect input of a production facility originates from ancillary facilities, from where the production facility takes the benefit to support its manufacturing operations, and from transport facilities that are used to transport the goods to it. The total input and output are shown in the Figure 4.30 as bold arrows coming to and leaving from defined system boundaries. The input of a production facility is summarized as follows:

- Direct energy consumption cost
- Load cost
- Proportion of energy consumption and load costs of ancillary facilities that support the manufacturing operation
- Proportion of energy consumption and load costs of used transport facilities
- Input and conversion efficiency degree of energy conversion facility
- Energy form

Equation (4.44) calculates the energy performance index of machine/production facility *p*.

$$EnPI_k = \frac{V_k}{P_k + \sum_{i=1}^{m} \alpha_i A_i + \sum_{j=1}^{n} T_j} \tag{4.44}$$

where

$EnPI_k$	=	energy performance indicator of production facility k
V_k	=	output value generated by production facility k [€]
P_k	=	direct energy effort/ input of production facility k [€]
α_i	=	ratio of energy usage from ancillary facility i = [0,1]
A_i	=	energy effort of ancillary facility i [€]
T_j	=	energy effort of transport facility j whose destination is production facility k [€]

and

$$P_k = \sum_{p=1}^{m} (E_{k,p} C_{E,p} + L_{k,p} C_{L,p}) \tag{4.45}$$

$$A_i = \sum_{p=1}^{m} (E_{i,p} C_{E,p} + L_{i,p} C_{L,p}) \tag{4.46}$$

$$T_j = \sum_{p=1}^{m} (E_{j,p} C_{E,p} + L_{j,p} C_{L,p}) \quad (4.47)$$

where

$E_{k,p}$ = energy consumption in form p of production facility k

$E_{i,p}$ = energy consumption in form p of ancillary facility i

$E_{j,p}$ = energy consumption in form p of transport facility j

$C_{E,p}$ = consumption cost per unit of energy p

$L_{k,p}$ = load of production facility k, energy form p.

$L_{i,p}$ = load of ancillary facility i, energy form p

$L_{j,p}$ = load of transport facility j, energy form p

$C_{L,p}$ = load cost per unit, energy form p

To avoid ambiguities of the variable P, L is used to represent the load.

Example:

A production facility PF processes a product and generates a value of 45 € within an hour. The energy meter attached to PF shows the energy consumption of 30 kWh. The mix energy cost of the internal network is 7.6 cents/kWh. The energy required to illuminate and to cool PF are approximately 2 kWh and 15 kWh respectively. The total energy required for a forklift to transport materials to PF is around 7 kWh. The charging station of the forklift uses only external energy cost from the utility company which costs 8.6 cents/kWh. The EnPI of PF can be calculated as follows:

$$EnPI_{PF} = \frac{45€}{0.076 \frac{€}{kWh} \times (30\ kWh + 17\ kWh) + 0.086 \frac{€}{kWh} \times 7kWh} = 10{,}78$$

The generated output value V_k in a certain time period can be broken down into production costs, such as personnel costs, material costs, and auxiliary material costs. Equation (4.48) determines the output value of production facility k.

$$V_k = C_{mat} + C_{aux} + \sum_{i=1}^{n} \beta\ W_i t_{ik} + C_{mai} + C_{misc} \quad (4.48)$$

where

C_{mat} = material costs processed by production facility *k*

C_{aux} = costs of auxiliary material used by production facility *k*

W_i = wage of worker *i* of certain period of time

t_{ik} = duration of worker *i* operating production facility *k*

β = surcharge factor; $\beta=1$ implies a normal shift work

C_{mai} = maintenance costs

C_{misc} = miscellaneous costs

Materials costs of production facility are ascribed by summarizing expenses from purchasing raw materials, supplies, auxiliary materials, and components or modules together. The type of required materials depends on the characteristics of the processed product. For this reason, such material costs are assigned to workstation level. In some special conditions, the material cost may vary although the same good is processed.

The wages of employees, who are in their jobs directly involved in the production, such as operator or retooling personnel, are considered to contribute workstation or production facility related personal costs. The costs are calculated by multiplying the wage cost per time unit with the duration of work. Additionally, the used shift work model reflected by surcharge factor affects the personnel costs. The surcharge factor is calculated using equation (4.49) [Roga-09].

$$\beta = \frac{t_n W_n + t_s W_s}{(t_n + t_s) W_n} \tag{4.49}$$

where

t_n = working time without surcharge

t_s = working time with surcharge

W_n = wage without surcharge

W_s = wage with surcharge

Example:

A production facility PF processes materials having a total cost of 210 € within 24 hours. The material processing needs water which costs approximately 10 €. Within this 24 hours, there is three working shift work with an operator in each shift to operate PF. The morning and afternoon shifts are from 6 a.m. to 10 p.m. The night shift is carried out from 10 p.m. to 6 a.m. The operator's wage is 15 €/hours. The operator working in night shift get 50% more wage. The total maintenance and miscellaneous cost of PF is approximately 9 €/hour. The surcharge factor can be calculated as follows:

$$\beta = \frac{16\,h \times 15\frac{€}{h} + 8\,h\ \times 22.5\,€/h}{24\,h \times 15\,€/h} = 1.17$$

The production value generated by PF in an hour is then estimated in the following:

$$V_k = \frac{210}{24}€ + \frac{10}{24}€ + 1.17 \times 15€ + 9€ = 35.72\,€$$

The equipment costs on production facility level can be ascribed in two methods. First, it can be ascribed individually on the workstation, and second, it ascribed in the form of depreciation. To keep a production facility or workstation functioning and to avoid operational failures, maintenance activities are required. The activities include periodic inspection and restoration activities caused by faults or damage.

The extension costs assigned to workstation or production facility are related to the technical and organizational activities in changing the workstation in order to increase the productivity and process efficiency. They incur most likely within a project that aims to improve the production performance. Due to this feature, this kind of costs is separated from others in accounting. The example concerns the cost of an additional machine module or tool to process other product variations. The costs commonly incur in both planning and implementation phase of extension. However, they include costs for coordination and organization after the implementation phase, such as costs for recruitment new personnel to work on the extended workstation, training expenses, etc. Because the costs are associated with long-term planning, they are then evenly distributed over the intended measurement period.

4.4.5 EnPI Calculation on Production Line Level

On the production line level, the EnPI calculation is in principle similar to the machine level, but with extended system boundary. It includes other ancillary and conversion facilities which are not considered at the machine level. Figure 4.33 illustrates energy flow representing this relationship.

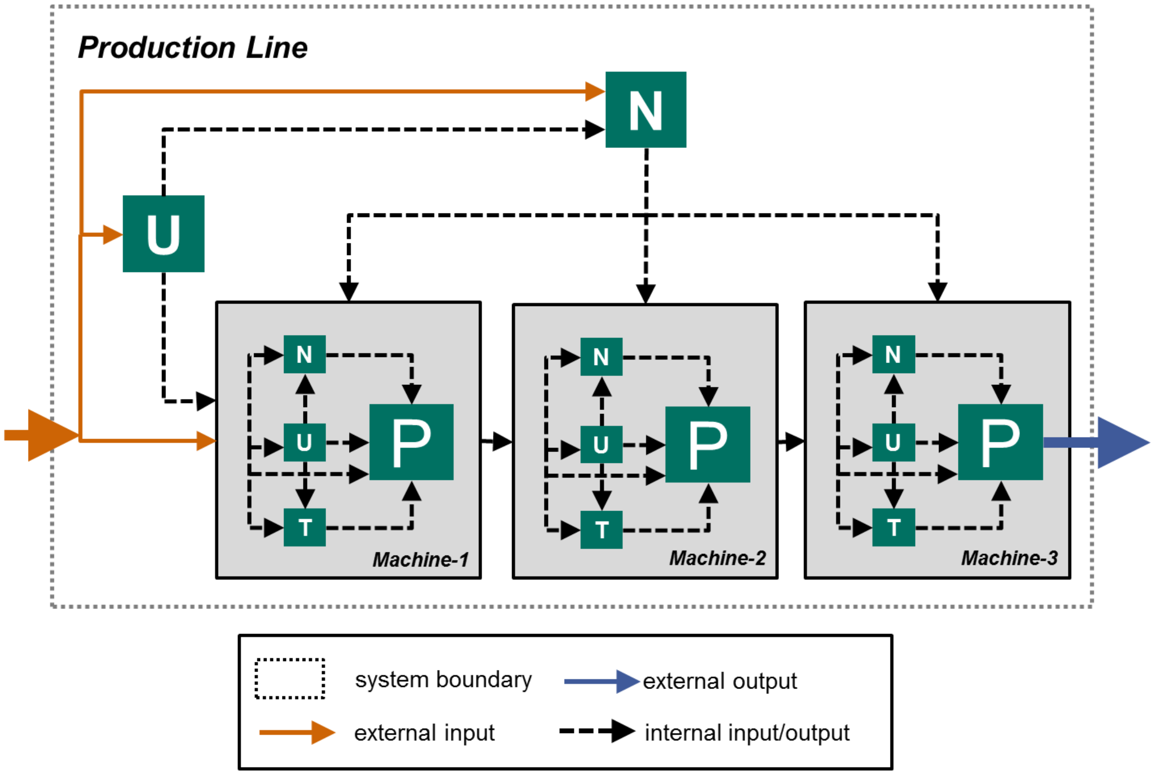


Figure 4.33 Energy flow model for production line level

The facilities that are taken into account cannot be defined in a general way. They must be decided individually. A question that leads to the decision, for example, which facilities should be shut down if the entire production line is shut down? Examples of such facilities are as follow:

- **Lighting**: The lighting systems that are considered here is the ones which directly affect the functionality and material flow in a production line. If the lighting systems are switched off, it will cease the activities in the production line. The energy effort of the lighting systems is distributed proportionally to the production line. If the production facilities are located spatially distributed, an assignment of the lighting system to production line is not possible. Thus, the energy effort of such lighting system cannot be ascribed from the energy effort of the whole production line.

- **Transport of small parts**: The efforts to transport small parts, such as screws, that assure the continuance of production processes.

- **Information systems**: Computer systems that control the processes of a production line or display panels showing targets and actual figures of the production line.

Justification and documentation of the facility selection are required to ensure traceability and repeatability of efficiency measure. Equation (4.50) defines the energy performance indicator considering additional ancillary facilities affecting the whole production line.

$$EnPI_l = \frac{V_l}{\sum_{k=1}^{p} P_k + \sum_{i=1}^{m} \alpha_i A_i + \sum_{j=1}^{n} T_j + \sum_{v}^{q} A_v} \quad (4.50)$$

where

$EnPI_l$ = energy performance indicator of production line *l*

V_l = output value generated by production line *l* [€]

P_k = direct energy effort/ input of production facility *k* involved in production line *l* [€]

α_i = ratio of energy usage from ancillary facility *i* = [0,1]

A_i = energy effort of ancillary facility *i* connected to production facility *k* [€]

T_j = used transport facility *j* [€]

A_v = energy effort from ancillary facility *v* connected to production line *l* [€]

V_l is calculated by using equation (4.51).

$$V_l = \sum_{k=1}^{p} V_k + C_l \quad (4.51)$$

where

V_k = Value generated by production facility *k*

C_{mai} = Costs related to production line *l*

Material costs assigned at the production line level are to maintain the production line specific functionalities. These include costs of fuel, for line specific for operating transport systems/medium or heating/cooling and lighting of the production environment. Included here are office materials such as paper, pencils, notebooks and other materials that provide some form of operational readiness of the production line.

In addition to workers who work in production workstations or operate production facilities, there are also personnel resources who deal with the planning, organization, and control of the production line. The costs ascribed from these personnel resources are considered as production line related costs. In contrast to personnel costs assignment on workstation level, the personnel costs on the production line are considered as production independent. Thus, they are summed up as fixed costs. The costs incurred by the depreciation of equipment are considered at line level as well. They are independent of the type of production and quantity. They include the depreciation costs of transportation systems and costs for continuous and discontinuous connectors between production lines. Examples of the connectors include belt conveyors, roller or chain conveyors, vehicle systems, as well as containers. Production line costs also include depreciation of office and other equipment which are impossible to be broken down to workstation level.

In the same way as the production facility, line-level requires maintenance activities which relate primarily in keeping the transportation systems and material flows in their operational capabilities. Technical and organizational changes that are required to increase productivity do not have to be restricted to the workstation level. Activities such as the construction of new workstations could extend the production line. Like at the workstation level, this requires proper planning and organization to carry out such extension measures. The resulting financial burden incurred from the extension is attributed to production line costs.

4.4.6 EnPI Calculation on Factory Level

Similar to machine and production line level, the energy efficiency indicator on factory level is calculated with extended system boundary. Only ancillary facilities that are used by production processes are considered in the calculation. The energy efforts of facilities belonging to other business departments such as sales, logistics is not taken into account. But industrial automation systems, social rooms, canteen, etc. are taken into consideration since they are used by the production personnel. Similar to line level, the system boundary is widened again including the considered conversion and ancillary facilities. Figure 4.34 depicts the energy flow model on the factory level.

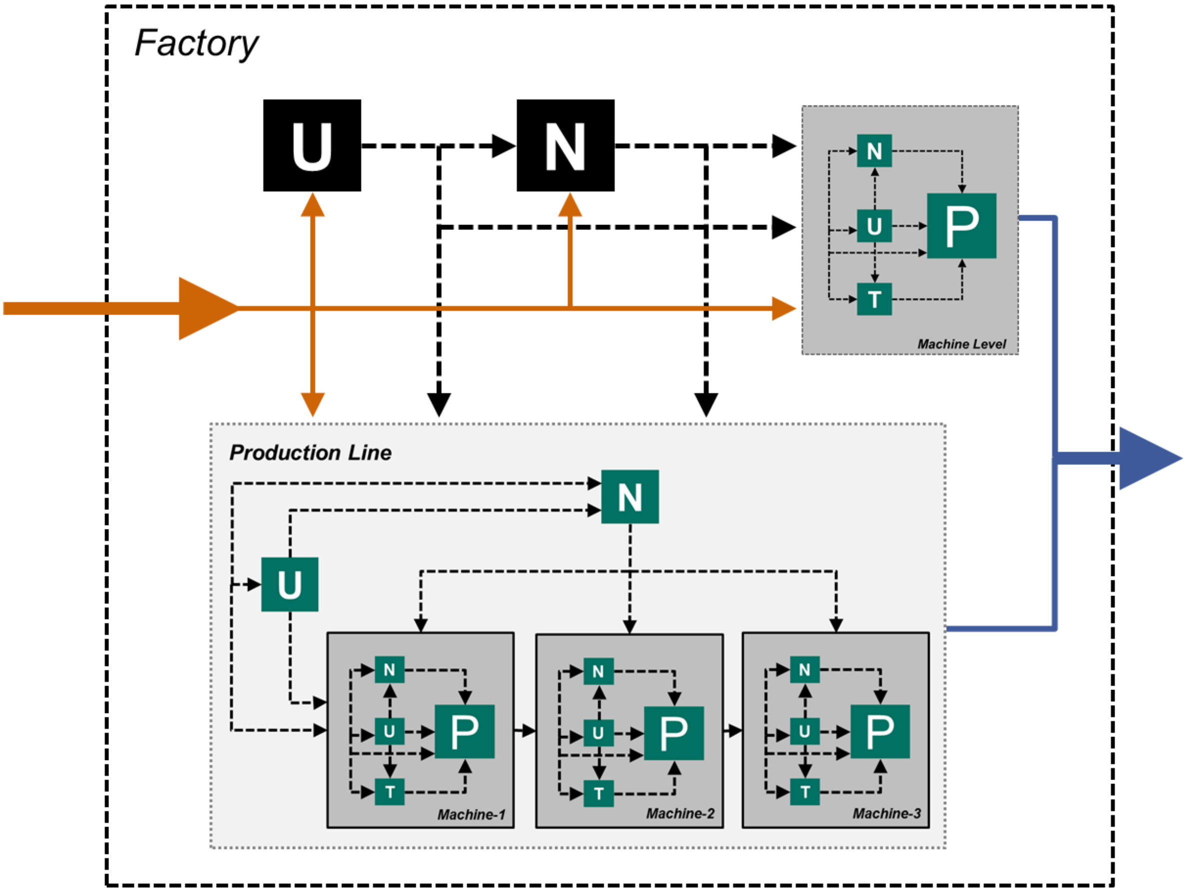


Figure 4.34 Energy flow model for factory level

To calculate the energy performance indicator on factory level, equation (4.52) is used. The equation includes the energy effort of further ancillary facility A_w.

$$EnPI_f = \frac{V_f}{\sum_k^p P_k + \sum_{i=1}^m \alpha_i A_i + \sum_{j=1}^n T_j + \sum_v^q A_v + \sum_w^r A_w} \tag{4.52}$$

where

$EnPI_f$ = energy performance indicator of factory f

V_f = output value generated by factory f [€]

P_k = direct energy effort/ input of production facility k in factory f[€]

α_i = ratio of energy usage from ancillary facility i = [0,1]

A_i = energy effort of ancillary facility i connected to production facility k [€]

T_j = used transport facility j [€]

A_v = energy effort from ancillary facility v connected to production line l [€]

A_w = energy effort from ancillary facility w associated with factory f [€]

and V_f is calculated through the following equation:

$$V_f = \sum_{l=1}^{q} V_l + C_f \tag{4.53}$$

where

V_l = Value generated by production line l

C_f = Costs related to factory f

In the factory level, there are also material costs that are considered separately. The costs are resulted from the central activities of the company, such as accounting human resources and logistics. These activities require material in substantial quantity, for example, office material, lights, and central cooling. All factory-related material costs cannot be broken down, but they are considered viable because they vary depending on the production volume and complexity. The personnel costs considered in factory level are those which relate to production-independent activities as well.

The costs generated by equipment in fulfilling the factory-related costs are assigned to factory level. Like in the production line level, the equipment does not take part directly in the production process. The costs are calculated by taken into account product independent amortization. Maintenance activities to be carried out directly related to the equipment at the factory level are directly included here. These include maintenance and service work on factory related building and site improvements. Costs of expansion projects which do not refer specifically to a workstation or production facility, a line or segment, but affect the production system as a whole, by its nature can also be summed up as factory level costs. It includes engineering, organizational, and start-up costs that result from the production of the technical and organizational transformation of a production system, such as the construction of a new manufacturing segment. Since such costs are considered as long term costs, they are evenly distributed over the expected period of use.

4.5 Energy Optimized Production Planning and Scheduling

This section explains the concept of energy optimized planning and scheduling to meet the requirements described in Section 3.5. The section begins with the formal description of the scheduling problem based on the scheduling framework developed by Blazewicz’s et al. Some additional problem specific factors added to the framework, for instance, lot size, retooling, shift-work model, precedence relations, and multiplicity of the operations. The section continues with the description, how the parameters represented in the problem definitions relate, resulting in a mathematical model. The overall structure of the approach to find the solution is then introduced in the following sub-section. The section ends with the explanation of the main heuristics applied to manage the tasks in scheduling, i.e. dispatching and dismantling heuristics, and finally, with the description of the hyper-heuristic approach considering adaptable and different strategies and constraints.

4.5.1 Problem Description

Blazewicz's et al. developed a scheme to classify scheduling problems [BlEt-07]. This section describes the scheduling problem using the scheme. Some parameters need to be separately specified since they are not covered by the Blazewicz's schema. The problem described in the thesis represents the worst case scenario of the problem. Some specific instances could formulate simpler problems, such as the single machine case. Nevertheless, the algorithm is required to handle the larger scale problems.

The scheme developed by Blazewicz et al. comprises three fields α, β, and χ, with

$$\alpha = \{\alpha_1, \alpha_2\}$$

$$\beta = \{\beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6, \beta_7, \beta_8\}$$

$$\chi = \text{the objective function}$$

where α denotes the machine environment and β describes tasks and resources. Based on the notation above, the scheduling problem in this thesis can be described in Table 4.14.

Table 4.14 Machine environment based on Blazewicz's scheduling problem framework

Variable	**Variable's meaning**	**Possible values**		**Value in this thesis**	**Description**
α_1	Characteristic of shop floor	∅ O F J	Single processor Open shop environment Flow shop environment Job shop environment	J	Orders are divided into tasks. The tasks are executed on the different machines/ workstations in different sequences.
α_2	Number of machines	∅ $k \in \mathbb{N}$ *no-wait*	Number of machines is variable. Number of machines equals to k. After finishing process on a machine, the process on the next machine is immediately started.	k	The number of machines/ workstations is greater than zero and is defined differently depending on each problem instance. It is defined as an exact value, not a variable value.

Table 4.15 Tasks and resources based on Blazewicz's scheduling framework

Variable	Variable's meaning	Possible values		Value in this thesis	Description
β_1	Pre-emption of tasks	∅	Pre-emption is not allowed.	*pmtn*	Pre-emption is allowed, but it has to wait after the number of goods corresponding to the defined lot size being produced. While retooling is performed on a machine, pre-emption is not allowed.
		pmtn	Pre-emption is allowed.		
β_2	Resource constraints	∅	Resource constraint exists.	∅	The model allows the representation of the resource constraints but the algorithm does not handle resource constraints for raw materials. However, the model offers the possibility for extension.
		res	Resource constraint does not exist.		
β_3	Precedence constraints of the tasks	∅	The tasks are independent.	*tree*	The precedence can be roughly considered as a tree. A product is manufactured through different operations/tasks with complex sequencing. An operation has multiple outputs.
		prec	General precedence constraints		
		uan	Unconnected activity network		
		tree	Precedence constraint is represented as a tree.		
		chains	Precedence constraint forms a set of chains.		
β_4	Ready times	∅	All ready times are zero		

		r_j	Ready times differ per task.	r_j	Order specific release time occurs.
β_5	Task processing time	$\emptyset$ $p_j = p$ $a \leq p_j \leq b$	Arbitrary processing time All tasks have process time *p* units. All tasks have process time between *a* and *b*	$\emptyset$	The process time is arbitrary depending on the machine, operation, and processed product.
β_6	Deadlines of tasks	$\emptyset$ $\tilde{d}$	No deadline are assumed, but due dates may be defined. Deadlines are defined for tasks.	$\emptyset$	The model puts in deadlines for orders, but the deadlines are loosened to due dates.
β_7	The maximal number of tasks forming a job in case $\alpha_1 = J$	$\emptyset$ $n_j \leq k$	The maximal number of tasks forming a job is arbitrary, or the scheduling problem is not JSSP. The number of tasks of each job is less than or equals to *k*.	$\emptyset$	There is no fixed definition of number of tasks.
β_8	Time buffers between machines	$\emptyset$ *no-wait*	Buffers between machines are unlimited. After finishing process on a machine, the process on the next machine is immediately started.	$\emptyset$	Buffers in between machines are infinite.

The optimal criteria depend on the problem instances. For example, a problem instance focuses on energy efficiency, and others aim to achieve the minimum production costs. In this thesis, three objective functions are taken into account. They are as follow:

- **Energy efficiency [η(s)]:** The production time influences the energy efficiency, for example producing at night is considered more energy efficient.

- **Cost[c(s)]:** The cost of a scheduling solution s depends on different parameters, for example, the shift work schedule selection, machine selection, etc.
- **Tardiness t(s):** In this work, the deadline is loosened. However, the delays are penalized for achieving feasible schedule.

The objective function is weighted in order to offer the flexibility of the model depending on the problem instance. The objective function considers three optimization criteria as follows:

$$\gamma = F(s) = c(s) + p\,\eta(s) + q\,t(s), \qquad p, q \in \mathbb{R} \tag{4.54}$$

The changes of the objective function do not affect the functionality of the algorithm since they follow the minimization principle. All criteria can be derived from s.

In addition to the parameters included in Blazewicz's scheduling framework, this thesis also considers the following parameters:

Lot Size

Supposed a few small components e.g. screws are required to assemble an end product. Most likely, the company produces the components in larger lot size. Each time the company receives a new order of the end product, it does not have to carry out the sub operation of producing the components. Therefore, intermediate component stocks are considered in order to decide whether the production operation that needs the components should be started.

Retooling

The retooling process on a machine cannot be pre-empted. It will make no sense if the retooling is stopped half way through, just to produce different products. However, once shift models are introduced, it might be reasonable to start retooling in the evening and continue the next morning. In this thesis, this is considered in the model instead of in the algorithm.

Operation with Multiple Input and Output

In this work, the *tree* constraint of Blazewicz's scheduling framework is extended, so that one operation can receive multiple inputs and it can produce multiple outputs as well. For example, in stamping production, a number of different products that tightly fit on one sheet are combined in one operation to get the stamping pieces. This is a common practice to reduce the scrap rate. Figure 4.35 illustrates the tree precedence relation where every operation produces only a single product. The multiple input-output relation precedence relation model allows an operation to have multiple output products. As shown in Figure 4.36, the operation O2 produces multiple products P5 and P6.

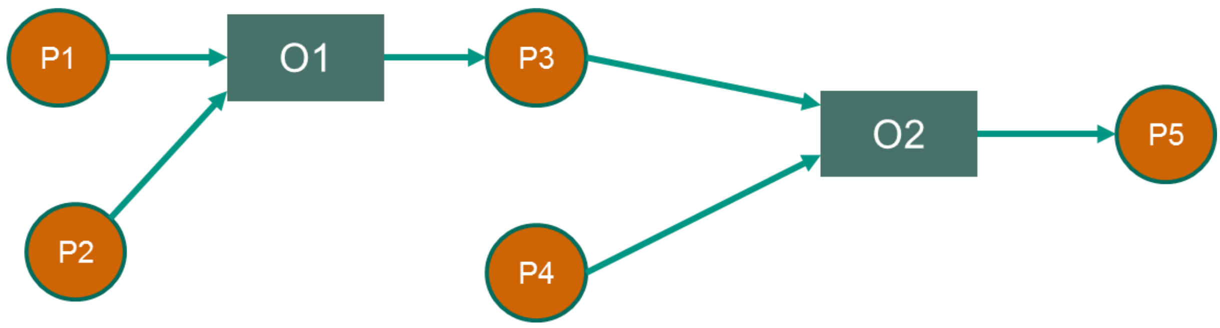


Figure 4.35 Tree precedence relations

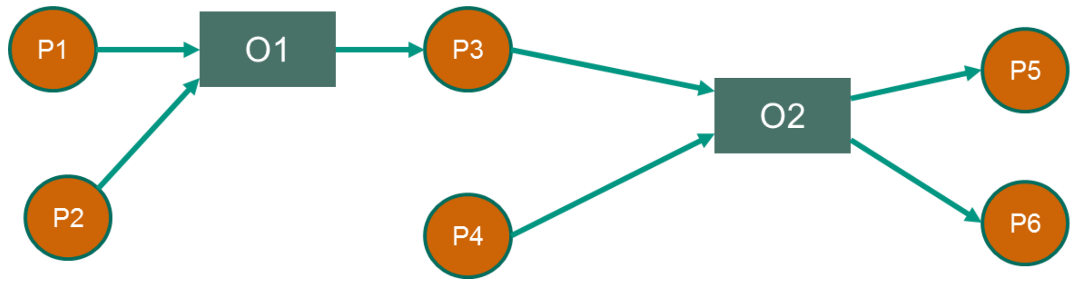


Figure 4.36 Multiple input-output precedence relations

Multiplicity of Operation-Product Relation

In this thesis, it is considered that the same product can be produced by different operations that might be executed on different machines. For example, a stamping piece that is regularly stamped in large lot sizes could also be cut by a laser in times of urgent need. The stamping operation requires a large sheet of metal while the laser needs a much smaller one. After choosing one of the two possible operations, the next operation, for example on a rolling machine, has to be scheduled accordingly. Consequently, a set of dynamic tasks has to be scheduled.

Shift Work Model

The costs and energy consumption of production also depend on the shift-work model. The various operational shifts provide the company more flexibility to process the orders. The developed concept creates a machine scheduling with an instance specific shift-work model which are defined in the company.

4.5.2 Mathematical Model of the Scheduling problem

Table 4.16 depicts the necessary parameters that are used to solve the scheduling problem with their units and value assignments.

Table 4.16 Necessary fundamental parameters of the scheduling problem

Parameter	**Unit**	**Description**
W		A set of machines i.e. workstations, where $W_i \in W = \{W_1, W_2, \dots, W_m\}$
P		A set of products including intermediate and end products, where $P_j \in P = \{P_1, P_2, \dots, P_n\}$
A		A set of customer orders, where $A_k \in A = \{A_1, A_2, \dots, A_o\}$
S		A set of shift works, where $S_l \in S = \{S_1, S_2, \dots, S_p\}$
O		A set of operations, where $O_{ij} \in O = \{O_{11}, O_{21}, \dots, O_{mn}\}$
lot_{ij}	units	Lot size of operation O_{ij} to produce product P_j on workstation W_i
t_{ij}	s	Required time to execute operation O_{ij}
$c_{ij}\ (S_l)$	€	Cost to execute operation O_{ij} if shift-work models S_l is chosen
Rij		A set of retooling operation, where $R_{ij} \in R = \{R_{11}, R_{21}, \dots, R_{ij}\}$ A retooling is a step that has to be accomplished before an operation O_{ij} produces a new kind of product on a workstation W_i.
tr_{ij}	s	Duration of retooling R_{ij}
cr_{ij}	€	Cost of retooling R_{ij}

An order can be specified as 3-tuple as follows

$$A_k = (P_k, Z_k, TR_k, TD_k) \tag{4.55}$$

where $P_k \in P$, Z_k denotes the amount of P_k specified in the order, and TR_k and TD_k are the defined deadline and release time. These parameters are instantiated each time the orders are received from customers. Due to the relaxation of deadline, a due date TD_k is used. It implies that late completion of operation entail penalization but not invalidness. Order's tardiness T_k does not consider the completion time TC_k prior to the due date TD_k. The tardiness is defined in (4.56):

$$T_k = \max(0, TC_k - TD_k) \tag{4.56}$$

The retooling R_{ij} is represented as 2-tuple as follow:

$$R_{ij} = (cr_{ij}, tr_{ij}) \tag{4.57}$$

The required time t_{ij} to execute an operation depends strongly on the characteristics of the processed product. In most cases t_{ij} is determined by process planner based on knowledge and their experience. This often leads to inaccurate production cost calculation. In this thesis, the required operation time is approximated by considering the relation between the characteristics of processed product. Equation (4.58) estimates the operation execution time t_{ij} on machine W_i to process product P_j. f_j denotes a property of product P_j and v_j is the weight of the property.

$$t_{ij}(f_1, f_2, \dots f_n\ ) = v_1 f_1 + v_2 f_2 + \cdots + v_n f_n + \gamma \qquad v_1, v_2, \dots, v_n \in \mathbb{R} \tag{4.58}$$

The equation calculates the operation time without considering the different operation states as described in Table 4.17

Table 4.17 Energy-related parameters of the scheduling problem

Parameter	Unit	Description
$\sigma(O_{ij})$		State of operation O_{ij} based on different energy consumption level, for example, STARTING, RUNNING, STOPPING
$P(\sigma)$	Watt	Power consumed during the execution of operation if the operation state is σ
P_{max}	Watt	The power threshold of the solution. The peak power load generated by the schedule may not exceed the E_{max}
EnPI		Energy Performance Indicator

4.5.3 Overall Structure of the Solution

The job scheduling algorithm tries to solve the shop floor scheduling problem when the company receives a set of orders A. The orders have to be split into a set of tasks in order to be executed on the available machines. Derivation from a set of orders to a set of tasks is not trivial, due to the extension factors of the scheduling framework such as lot sizes and multiplicity of operations.

As mentioned in Section 2.1.4, a combination of priority dispatching and local search allows implementation of autonomous search algorithms that explore solution spaces. The priority dispatching algorithm is considered as low-level heuristic, and the local search is the actual hyper-heuristic. With given a problem instance, a low-level heuristic h_1 will always point to one possible solution space $s \in S$. By perturbing h_1, a further dispatching h_2 that points at a different solution space s' could be created. The low level heuristic can be formulated with equation (4.59), where A is a set of parameter in parameter space, and $h(\alpha)$, $\alpha \in A$ is parameterized heuristic [WiPO-13].

$$h: A \rightarrow S \tag{4.59}$$

The objective function that evaluates the solution can be described as follow:

$$\gamma: S \rightarrow R \tag{4.60}$$

R denotes the ranked solutions that are in the solution space. Thus, the minimization problem can be formulated in (4.61).

$$\min \gamma(h(\alpha)) \tag{4.61}$$

Figure 4.37 depicts the overall structure of developed approach for solving the scheduling problem in production that is formulated in Section 4.5.1 and Section 4.5.2. The set of orders A is transformed into a set of tasks by dismantling heuristic. After that, the set of tasks is assigned to a schedule by low-level heuristic. The dispatching heuristic will be executed iteratively and perturbed each time by different parameter set from hyper-heuristic. The hyper-heuristic contains high-level heuristic which is implemented as hill climbing algorithm.

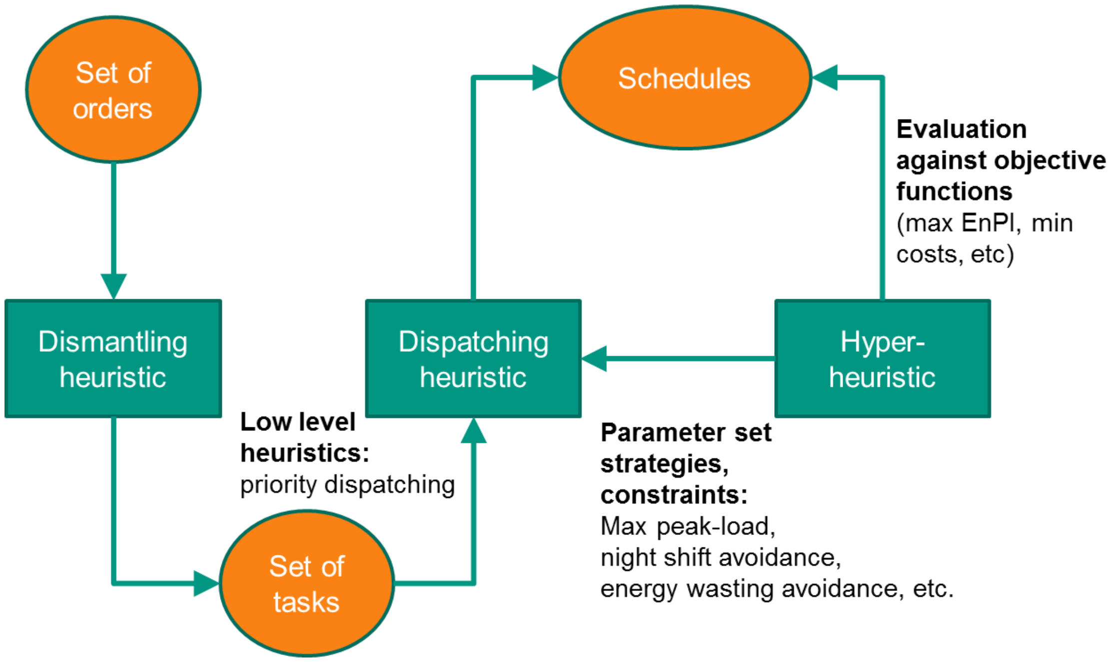


Figure 4.37 Overall structure of the hyper-heuristic [WIPO-13]

4.5.4 Dismantling and Dispatching Heuristics

One of the challenges in dismantling the orders into tasks is to consolidate the orders with different release times. The important start point is to decide how to consider the release time in the schedule model. It depends strongly on the knowledge of the upcoming orders. If it is assumed that there is no knowledge about the upcoming orders prior to release time of the considered orders, consolidation would not be possible. The consolidation is only possible, if at all time, the upcoming orders are known prior to the release time of the considered orders. In this case, it would be possible to postpone one order to consolidate it with another that has later release-time. The dismantling heuristic should consider the overproduction by looking at the residual production amount induced by lot sizes. In order to do this, the concept of *imaginary stocks* is introduced to ensure that no overproduction occurs.

The multiplicity of operation as explained in Section 4.5.1 is the next challenge to perform the dismantling. It leads to a dynamic set of tasks. However, in the hyper-heuristic architecture, the tasks have to be defined as static to perform the dispatching. Therefore, only usable operations are considered to create the tasks. The concept symmetrical operations can enforce the staticness. Two symmetrical operations mean that both operations have the same lot size, output products, input products, and quantities of input and output products. The operations may still have different costs, duration and workstation, where the operations are executed. During the dismantling process, a group of symmetrical operations is defined, so that later the dispatching can choose among these operations.

An interchangeable strategy to choose the group of tasks is implemented to maintain the flexibility of dismantling heuristic. By using this kind of flexibility, the dismantling heuristic is adaptable based on the problem instances, such as strategies to choose the smallest lot sizes and to ensure inclusion of the cheapest operations. In this thesis, a time progressive dispatching heuristic is developed. Tasks are scheduled at a certain point of time. If there is no task to be assigned anymore, the time pointer is moved forward, and more tasks will be assigned at the new time point.

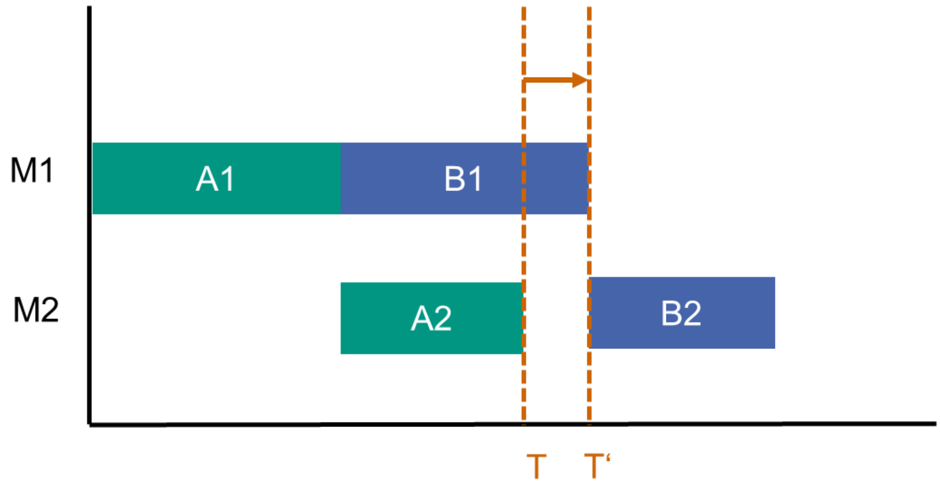


Figure 4.38 Dispatching at a time point T by considering the stock [WiPO-13]

Figure 4.38 depicts the assignment of task *B2* to the schedule. Information on all scheduled tasks prior to the point of time *T* is reflected by the stock information at *T*. In Figure 4.38, it can be seen that the task *B1* is not completed at *T*. Thus, the execution of task *B2* is not possible due to the insufficient stock. Therefore, the time pointer is moved to *T'*, where the task *B1* finishes, and there is still enough required stock to execute *B2*.

The parallel stock access that is when two output products require the same input product at the same time could cause problems. As an illustration, consider the task assignment in Figure 4.39. To produce *P1* and *P2* through task *B* and *C*, it requires double amount of *P3* which is produced by task *A*. At time point *T*, where the task *A* finishes, it seems to be possible to assign both tasks *B* and *C*, since each task has only information about the scheduled tasks before *T*. However, the stock of *P3* will be insufficient because it is consumed at twice parallel, so that it becomes used up on half way through. A taboo list for stocks of intermediate products is implemented to solve this problem. As long as the stock is being accessed by a task, it is no longer possible to assign another task that requires the same product.

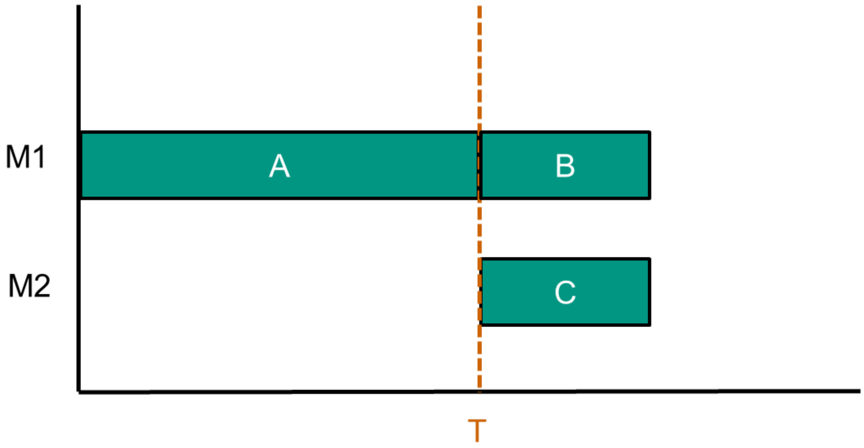


Figure 4.39 Parallel stock access [WiPO-13]

The movement of the time pointer has to consider the task completion and order release. Every time that a task is completed, the machine is ready to accept a new task. Thus, the minimum of completion times of the currently running machines is considered. Every time a new order is released, new tasks are ready to be scheduled. They are executed on machines that are already idle and ready to accept tasks.

4.5.5 Hyper-Heuristic

The hyper-heuristic in this thesis considers two tuples to be direct neighbors if they differ in a certain number of strategies at maximum. A hill climbing algorithm is implemented for the hyper-heuristic. The neighborhood relation is parameterized so that neighbors down to the depth i can be explored. Therefore, solutions that differ in up to i strategies can still be considered adjacent. However, the increase of i means the longer runtime. The evaluation of one hyper-heuristic solution will then initiate the dispatching heuristic that will return a schedule. The produced solutions are saved and mapped to the used strategy tuples and will be revisited by the hill climbing algorithm [RuNo-03].

Different strategies are implemented in this thesis in order to consider different knowledge corresponding to specific problem instance. The strategies have to regard the objective function. Thus, introducing a new objective function means that strategies have to be changed, or new strategies have to be added. In this thesis, Earliest Due Date (EDD) and Deepest Task First (DTF) strategies are implemented.

In EDD, the dispatching heuristic the task dismantled from the orders having earliest due date are assigned to the schedule first.

Example:

Figure 4.40 illustrates an example of EDD strategy for the given data listed in Table 4.18.

Table 4.18 Problem data for hyper-heuristic

Order	Machine sequence (machine:duration)	Deadline
1	(M1:2)	2
2	(M2:4) → (M1:2) → (M2:2)	10
3	(M1:1) → (M2:2)	6

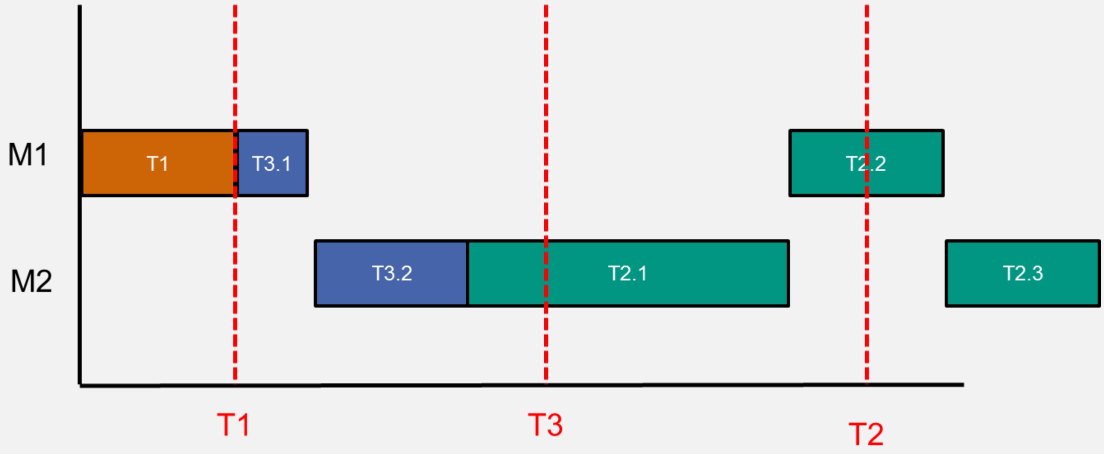


Figure 4.40 Gantt chart illustrating EDD strategy

Since there are no direct task dependencies in this thesis, it is not possible to implement the Longest Finishing Time (LFT) strategy. Therefore, DTF strategy is implemented. DTF is similar to LFT, but the deepest instead of longest finished tasks have the higher priority to be assigned to the schedule. A task is considered deep if the corresponding order needs more operations to get the end product.

Example:

Considering the problem data presented in Table 4.18, Figure 4.41 shows DTF strategy in a Gantt chart.

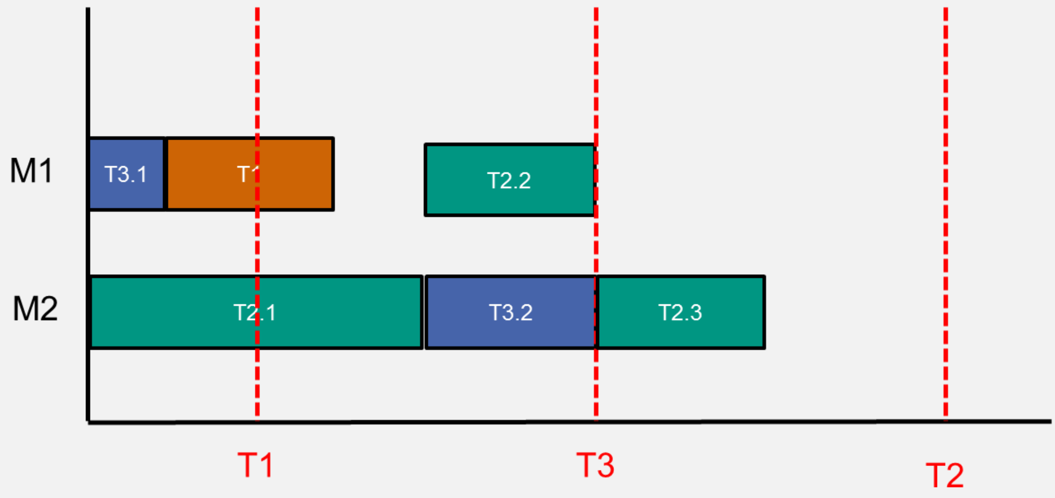


Figure 4.41 Gantt chart illustrating DTF strategy

As described in Section 4.5.1 and Section 4.5.2, the scheduling model considers the shift work model. It affects the costs and energy efficiency since the personal and energy costs depend on the time. In the model, a day is divided into a number of shifts where each of shifts has different cost and energy factor. Therefore, additional constraints are added to the strategy, for example, EDD using all shifts, EDD avoiding night shift, or EDD avoiding day shift.

The dispatching heuristic does not move the time pointer to avoided time interval. It always assigns the tasks the next feasible point of time. If the duration of a task is greater than a possible interval, the task will be disrupted, and it will be continued in the next feasible interval. The disrupted task will still be considered as a single task. If the time pointer *T* lays in an avoided interval, the task should not be considered as scheduled in the future, but as scheduled now with a disruption in the beginning.

The design choice of the hyper-heuristic level is determined by a set of dispatching strategies. In first implementations one strategy was assigned for each dispatching decision, as it was first proposed in [StEt-93]. For example, Figure 4.40 illustrates a tuple containing six strategies [EDD, EDD, EDD, EDD, EDD, EDD] representing the dispatching heuristic. The solution is searched by generating neighboring tuples. A neighboring tuple is a tuple having an altered strategy comparing to the original one.

Example:

The neighboring tuples of [EDD, EDD, EDD, EDD, EDD, EDD] are [EDD, EDD, EDD, EDD, EDD, DTF], [EDD, EDD, EDD, EDD, DTF, EDD], [EDD, EDD, EDD, DTF, EDD,EDD], [EDD, EDD, DTF, EDD, EDD, EDD], [EDD, DTF, EDD, EDD, EDD, EDD], [DTF, EDD, EDD, EDD, EDD, EDD]. Figure 4.42 depicts a Gantt chart representing a neighboring tuple [DTF, EDD, EDD, EDD, EDD, EDD] which leads to a solution. As seen in the figure, the production plan meets all deadlines T1, T2, T3.

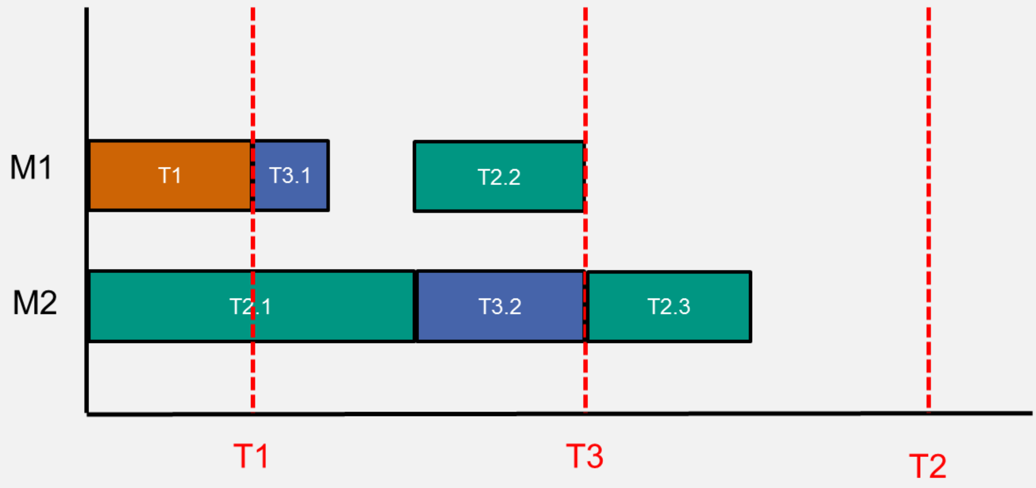


Figure 4.42 Gantt chart illustrating the tuple [DTF, EDD, EDD, EDD, EDD, EDD]

Shift-Work-Strategies

Since energy consumption and production costs depend on whether the tasks performed at night or day, the shift-work model has to be considered in choosing the strategies. The shift-work model consists of a number shifts. Each shift has weight factors corresponding to energy and personal costs. For example, a night shift has higher personal cost factor than a day shift. In this work, the EDD and DTF strategies are extended by adding the shift strategies, resulting in combinations of shift-work and task allocation strategies, for example, EDD with all shift, EDD avoiding weekend, DTF avoiding night shift, DTF avoiding Sunday work, etc. Combinations of DTF or EDD avoiding day shift do not correspond to reality. Therefore, they are not considered as strategies which lead to a solution. The dispatching heuristic will schedule the task in the next possible time interval if the dispatching time *T* lies in an avoided interval. The tasks which take more than 24 hours will be disrupted at the avoided interval, and they will continue in the next feasible interval. The task segments due to the interval avoidance like this are still considered as one task.

Power Peak-load Avoidance Strategies

The dispatching strategies are the configuration elements of the framework since they allow the users to define optimization goals and algorithm constraints. New strategies, such as power peak-load avoidance can be introduced in the framework by configuring the goals and constraints. When introducing a new strategy, such as peak load avoidance, it is inevitable to alter the objective function as well. Only if peaks in energy consumption are punished by the objective function, applying the PLA strategy will be reasonable. In many cases, the objective function will need to be modified with respect to energy pricing manually. On the implementation level, a strategy needs to provide a method that will return a single task (or no task), when given a set of tasks. As the first step, all unfeasible tasks need to be excluded for reasons like "*input products are not available*" or "*task has not been released yet.*" In the second step, the remaining feasible tasks are ordered, and the best one among them is returned. In order to aid reusability among strategies, it is possible to break them down into single *constraints*. Each constraint monitors one property of a task in a percentage value. Values less or equal to 100% represent feasibility, and a feasible value close to 100% is considered favorable [WIPO-13].

The peak load avoidance strategy consists of four following constraints in evaluation a task before it is assigned to the schedule: the maximum allowed power level, the availability of input products stock states), the availability of the machine required by the task, and the state of the task whether it has been released. Figure 4.43 depicts the energy level and stock state constraints for evaluation a task.

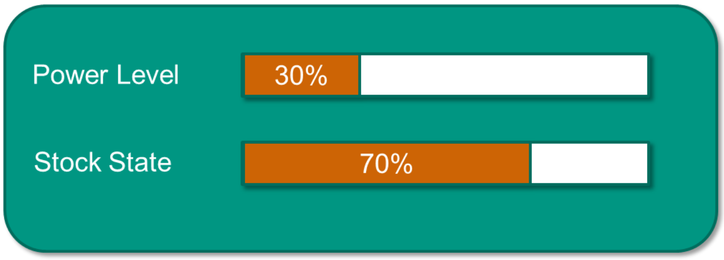


Figure 4.43 Power level and stock state constraints for task evaluation

Figure 4.43 shows the stock state of 70%. It means, more products is available in the stock than the amount required by the task. A stock state more than 100% indicates unavailability of the stock. The figure indicates that adding the task to the schedule would reach 30% of the maximum energy level allowed. So far, this section has described how the maximum power load constraint is incorporated into the framework. The first step is to implement the peak power load avoidance strategy multiple times at different power levels. Hyper-heuristic will select the lowest possible power level where all tasks are still able to finish before the deadlines. As a consequence, the complexity of the problem raises. The number of neighbors will increase due to the combination with different power levels. If *n* different power levels are provided and a solution tuple contains *m* strategies, the number of neighbors would be $2 * m * (n - 1)$. The factor 2 indicates the considerations of both EDD and DTF strategies for each power level. At least, the final iteration of the Hill Climber will need to explore all of those neighbors. A strategy alignment will help to limit the number of neighbors. If the hyper-heuristic has evaluated one energy level, it would be reasonable only to explore the next lower energy level in the next iteration [WIPO-13].

Example:

Figure 4.44 illustrates the strategy alignment in a simplified version.

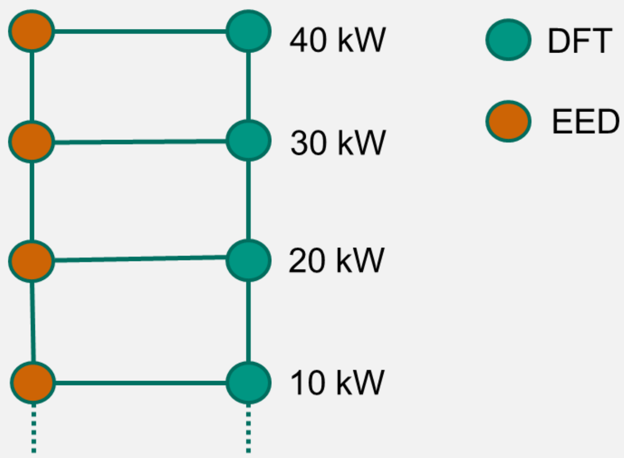


Figure 4.44 A simplified example of strategy alignment [WIPO-13]

Consider a tuple [EDD-30kW, EDD-30kW, EDD-30kW], according to the alignment strategy depicted in Figure 4.44, the neighbors would be [EDD-20kW, EDD-30kW, EDD-30kW] and [DTF-30kW, EDD-30kW, EDD-30kW], but not [EDD-10kW, EDD-30kW, EDD-30kW], [EDD-40kW, EDD-30kW, EDD-30kW] because of higher power level (40kW), or [DTF-20kW, EDD-30kW, EDD-30kW].

For larger problem instances the strategy alignment takes the form of a hyper-cuboid of *k* dimensions. *k* is determined by the number of objectives in the given problem, provided that all strategies can be ordered according to each objective. Additional strategies following an additional objective would cause additional dimensions in strategy alignment. For a tuple containing m strategy alignment will result to have $m * 2^k$ neighbors. For a large number of power levels, it can be assumed that $2 \ll n - 1$ and thus that strategy alignment has significantly reduced the complexity of the search neighborhood [WIPO-13].

5 Verification and Validation of the Concepts

"A theory must be tempered with reality."
~ Jawaharlal Nehru (1889-1964)

This chapter deals with the verification and validation of the integrated concept consisting of the ontological knowledge base including the ontology model and knowledge acquisition through KDD, the energy performance indicators, and the energy optimized production scheduling developed in this thesis, which has been elaborated in Chapter 4. The verification aims at checking and evaluating the developed concepts, whether they satisfy the requirements formulated in Chapter 3. In this thesis, software prototypes are implemented to enable the verification of the functionalities concerned in the requirements. This chapter begins by describing the implementation of the prototypes.

However, the implementation of prototypes is not sufficient to prove the feasibility of the developed concepts to be applied in real environments of manufacturing companies. The second section of this chapter reports the application and validation of the concepts in different use case scenarios from several types of manufacturing companies. The section describes the scenarios and how the functionalities of concepts and implemented prototypes are applied related to those scenarios. The applications are made more clearly by displaying the screenshots of the prototypes. Finally, this chapter presents the evaluation of the developed concepts and implanted prototypes in details, whether they comply the requirements formulated in Chapter 3.

5.1 Implementation of Software Prototypes

The tasks for implementing the software prototypes are in accordance with the requirements of software architecture specified in Section 3.5. The prototypes are implemented by considering the modern technologies which give the technical advantages, such as simplicity in software development, extension possibilities and applicability in manufacturing companies with different ICT environments. Furthermore, the development of the software prototypes have to consider the advantages for both users and developers, for example, user-friendliness of graphical user interface (GUI) and modularity or extensibility of the software.

5.1.1 Software Architecture

As presented in Chapter 4, this thesis develops four main concepts, i.e. the manufacturing energy management knowledge base, the semi-automatic knowledge generation using data mining, the EnPI, and the energy optimized production planning and scheduling. Those concepts have to be incorporated in an integrated system comprising hardware and software. The system is called Smart and Integrated Energy Efficiency Evaluation (SERENE). The developed prototype concerns only the software part of the system. Figure 5.1 depicts the software architecture that implements the developed concepts.

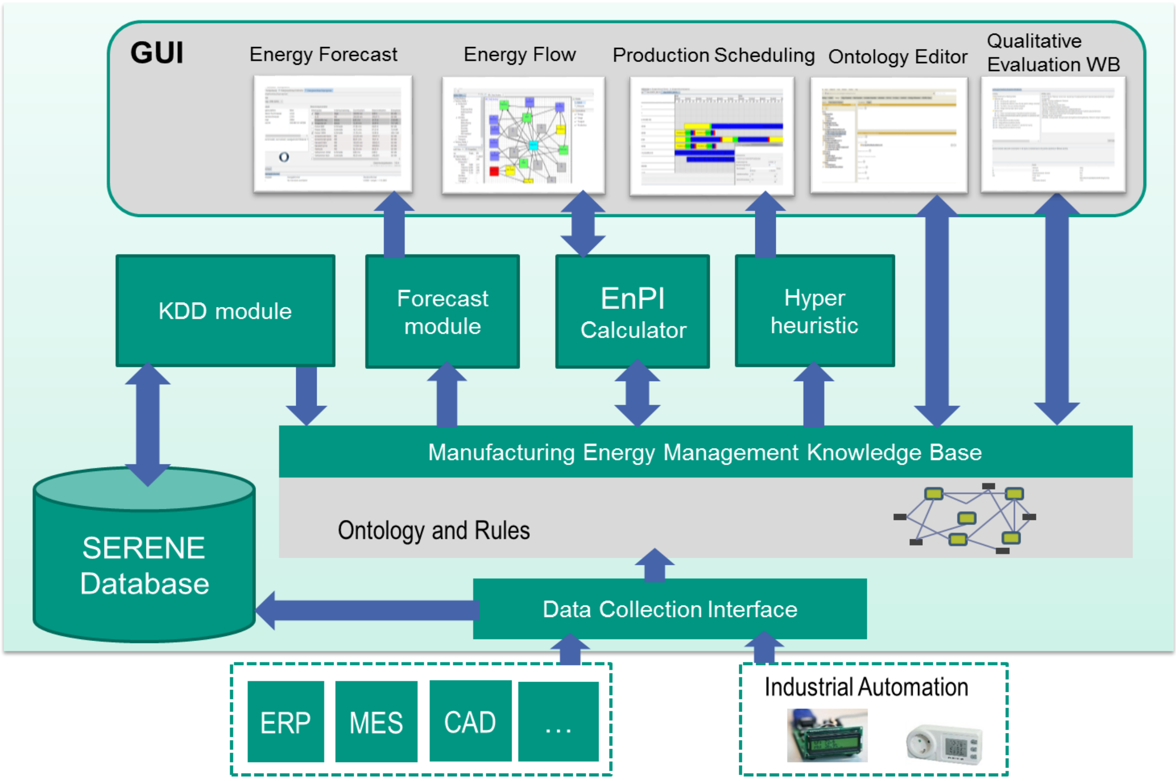


Figure 5.1 Software architecture of SERENE

The architecture consists of eight main components as follows:

1. ***Data Collection Interface***. Most of the data required for the developed system exist already in ICT environment of a typical manufacturing company, for example, resource data from ERP system, product data from PDM system, process data from ERP and MES, energy metering, and sensor data from data acquisition system and industrial automation. As seen in Figure 5.1, the Data Collection Interface is responsible for transferring data from those systems, to structure the data following the defined schema, and to store the data into the SERENE database. Those existing systems provide data communication interface in order to enable the access from external systems. The interfaces are for example CSV export, SOAP7 web service, REST web service, etc.

2. ***SERENE Database***. The SERENE database contains a relational model of all entities, e.g. products, processes, resources, orders, and energy consumptions required for the developed concept of manufacturing energy management system. Moreover, the data collected by the

[7] https://www.w3.org/TR/soap/

Data Collection Interface can be stored in SERENE database permanently. SERENE database also serves as the storage of pre-processed data required by KDD module.

3. ***Manufacturing Energy Management Knowledge Base.*** This component serves as integration point of different modules in a higher layer, i.e. Forecast Module, EnPI Calculator, and Hyper-Heuristic and simultaneously as a semantic representation of data models required by the upper layer modules. The knowledge base comprises (i) OWL that model the hierarchies and relations of semantic concepts and (ii) a set of rules represented as SWRL.

4. ***KDD Module.*** This component is responsible for the implementation of KDD algorithms including pre-processing, i.e. selection, data cleaning, transformation, and discretization, and a set of data mining algorithms, e.g. decision trees, M5P classification rules, and linear regression. It also contains a software module to integrate the resulting rules and equations into the knowledge base through the conversion to SWRL.

5. ***Forecast Module***. This module performs the prediction of power consumption and duration if a machine executes an operation. The prediction uses the rules and equations resulted from data mining which is already stored in the knowledge base. The module assigns the variables in the rules and equations with instances of product properties, machines, and processes which corresponds to a customer order. It fires then the rules resulting in the predicted duration and power consumption. The module implements the equations (4.27) and (4.28) as software to calculate the energy consumption and cost of the whole product.

6. ***EnPI Calculator.*** This module is responsible for the energy flow modeling including the definition of system boundaries and interconnections among facilities as the implementation of the concepts presented in Section 4.4.1 and Section 4.4.2. The energy flow can be modeled manually through GUI or loaded from a model that is already stored in the ontology knowledge base. The module also implements the EnPI calculation at the machine, production line, and factory level as designed in Section 4.4.4, Section 4.4.5, and Section 4.4.6. It then presents the results in the GUI,

7. ***Hyper-Heuristic.*** It implements hyper-heuristic concept and algorithms for the energy optimized production scheduling as described in Section 4.5. The objects required for the execution of the algorithms are loaded from the ontology, for example, machines, processes, and products. The hyper-heuristics objective and constraints, such as maximum power load and shift-work model, are defined by the user through GUI. Moreover, the resulting production schedule, the corresponding power consumption plot, and figures, such as total cycle time and costs, are presented through GUI

8. ***Graphical User Interface (GUI).*** It serves as the interaction interface between the user and the system. This thesis develops three GUI to display the energy and cost forecast results, the energy flow and EnPI, and the production schedule resulted from hyper-heuristic. The GUI is implemented as desktop based application. In this work an existing ontology editor Protégé is used to create the main classes, properties, and rules in the ontology.

5.1.2 Software Implementation

The software implementation of most of the architecture components explained in the Section 5.1.1 is based on Java programming language. This work employs the Java version *"Java Standard Edition Development Kit 7 (Java SE 7)* and *Eclipse IDE version 3.7* named *Indigo*[8] and *version 3.8* named *Juno*[9]. The following sections explain the software implementation of each module. The software is developed under *Windows 7* and *Windows 8* desktop environment.

Data Collection Interface

In order to transfer and transform the data from the interfaces of the existing IT systems into the SERENE database, the following technologies are applied for the implementation of data collection interface.

- *Java Database Connectivity (JDBC) version 4.1* to import data from CSV files or other SQL databases into SERENE database. The JDBC v4.1 is included in Java SE 7. The technology imports the product, process, and resource data.

- *Apache Axis2*[10] to implement the web service client for connecting to SOAP web services interfaces of existing IT systems. The web service client transfers and parses the energy and sensor data, and stores in SERENE database. The implementation of the web service client considers the security of data communication since the data are transferred over the internet. A security mechanism using authentication header for each connection and query is implemented for that.

SERENE Database

The SERENE database is implemented using MySQL version 5.1.41. MySQL is the world's most popular open-source Relational Database Management System (RDBMS) [ORA-16]. It contains main tables, which stores product, process, resource and energy consumption data, and auxiliary tables to help the data processing. Figure 5.2 depicts a screen shot of SERENE database for a make-to-order manufacturing producing stainless steel. The screenshot shows the database displayed using web-based database management application phpmyadmin[11].

---

[8] http://wiki.eclipse.org/Indigo

[9] http://wiki.eclipse.org/Juno

[10] https://axis.apache.org/

[11] http://www.phpmyadmin.net/

Tabelle	Aktion	Einträge [1]	Typ	Kollation	Größe	Überhang
1_datetime_transf		118,677	MyISAM	utf8_unicode_ci	3,6 MiB	-
2_discretization		118,677	MyISAM	utf8_unicode_ci	5,2 MiB	-
3_granulation		1,916,321	MyISAM	utf8_unicode_ci	106,3 MiB	-
4_merge		136,826	MyISAM	utf8_unicode_ci	7,9 MiB	-
5_test		121,083	MyISAM	utf8_unicode_ci	6,6 MiB	-
cons_perv_temp		2,576,121	MyISAM	utf8_unicode_ci	53,8 MiB	-
customer_order		9,260	MyISAM	utf8_unicode_ci	329,8 KiB	-
energy_consumption		0	InnoDB	latin1_swedish_ci	16,0 KiB	-
energy_type		14	InnoDB	latin1_swedish_ci	16,0 KiB	-
heattreatment_category		7	InnoDB	latin1_swedish_ci	16,0 KiB	-
machine_activity		~63,863	InnoDB	latin1_swedish_ci	7,5 MiB	-
material		254	InnoDB	latin1_swedish_ci	16,0 KiB	-
material_category		6	InnoDB	latin1_swedish_ci	16,0 KiB	-
measurement_point		26	InnoDB	latin1_swedish_ci	16,0 KiB	-
oldproduct		27	InnoDB	latin1_swedish_ci	16,0 KiB	-
plastics_product		0	InnoDB	latin1_swedish_ci	16,0 KiB	-
product_category		6	InnoDB	latin1_swedish_ci	16,0 KiB	-
ring_product		3,676	InnoDB	latin1_swedish_ci	1,5 MiB	-
temp_bau_order		9,360	InnoDB	latin1_swedish_ci	1,5 MiB	-
temp_betriebsdaten		~64,475	InnoDB	latin1_swedish_ci	7,5 MiB	-
temp_erw		9,260	InnoDB	latin1_swedish_ci	3,5 MiB	-
temp_operation		~74,937	InnoDB	latin1_swedish_ci	8,5 MiB	-
temp_perv		~3,263,012	InnoDB	latin1_swedish_ci	258,8 MiB	-
temp_raw_order		6,209	InnoDB	latin1_swedish_ci	2,5 MiB	-
temp_wst		1,204	MyISAM	latin1_swedish_ci	51,1 KiB	-
workplace		43	InnoDB	latin1_swedish_ci	16,0 KiB	-
workplace_category		5	InnoDB	latin1_swedish_ci	16,0 KiB	-
27 Tabellen	Gesamt	~8,493,349	MyISAM	utf8_unicode_ci	475,5 MiB	0 Bytes

Figure 5.2 Screen shot of SERENE database

Manufacturing Energy Management Knowledge Base

The ontology in manufacturing energy management knowledge base is mainly constructed using ontology editor Protégé. The language used by the ontology is OWL-DL in RDF/XML format. Figure 5.3 depicts the class hierarchy of developed manufacturing energy management ontology SERENE which is created using the ontology editing tool Protégé[12]. A total of 163 classes has been created. Ten of those classes are top classes located directly under the root class `Thing`. Figure 5.4 shows the data object and data type properties of manufacturing energy management ontology. There are 40 object and 52 data type properties. Furthermore, this thesis has modeled 49 SWRL rules in the ontology for the references to perform the qualitative evaluation of energy efficiency. They do not include the rules generated by the data mining.

12 http://protege.stanford.edu/

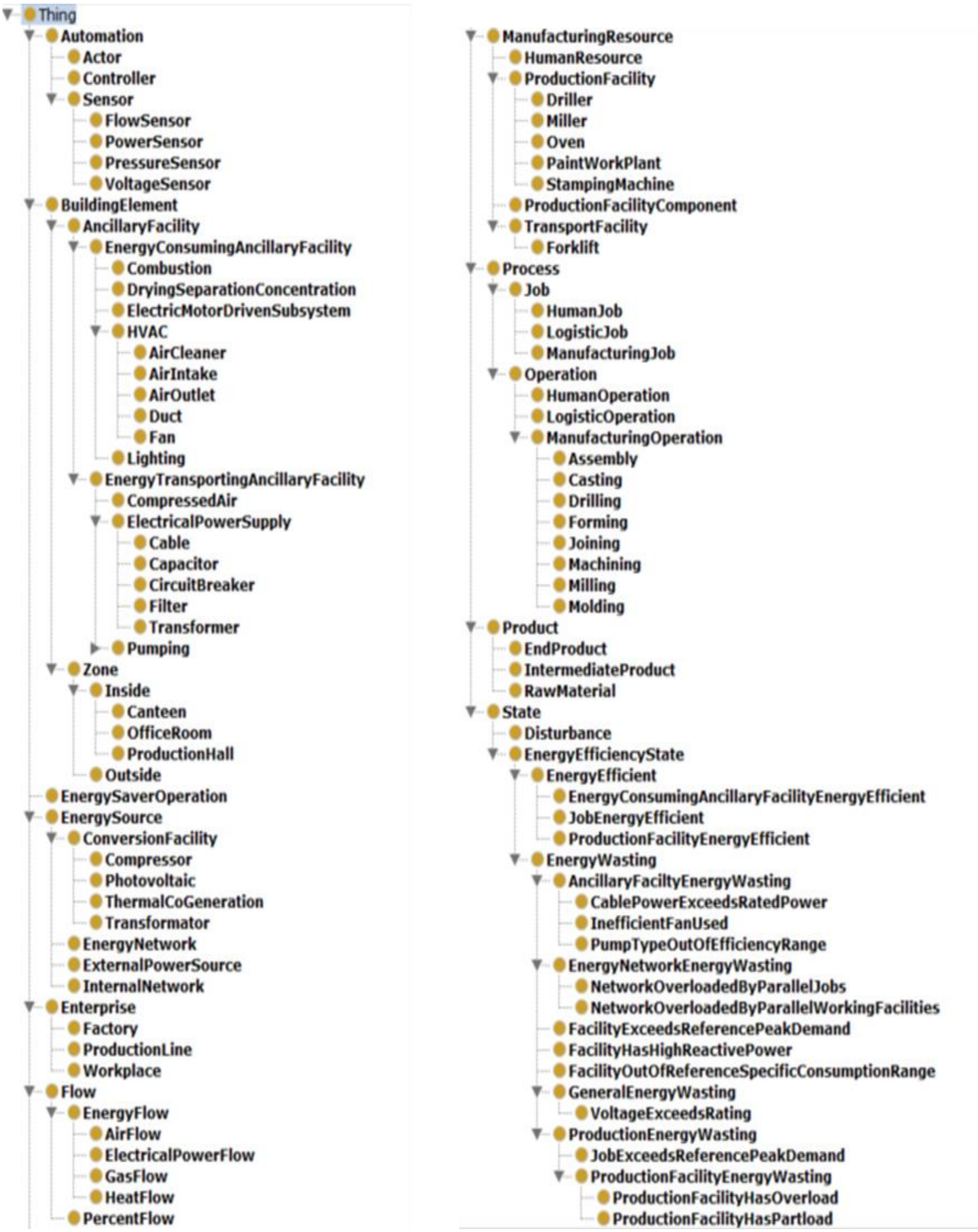


Figure 5.3 The class hierarchy of manufacturing energy management ontology SERENE

Figure 5.4 The object type and data type properties of manufacturing energy management ontology SERENE

```
Rules
ProductionFacility(?x), hasCurrentType(?x, "DC"^^string), hasRatedEnergyConsumption(?x, ?raec), hasRatedEnergyIntensity(?x, ?raei), hasRatedPowerInput(?x, ?rapi), hasReferenceMaxEnergyIntensity(?x, ?rmaei), hasReferencePeakPowerInput(?x, ?rppi), hasReferenceSpecificEnergyConsumption(?x, ?rsec), lessThanOrEqual(?raec, ?rsec), lessThanOrEqual(?raei, ?rmaei), lessThanOrEqual(?rapi, ?rppi) ->
(BuildingElement or ManufacturingResource)(?x), hasRatedEnergyConsumption(?x, ?raec), hasReferenceSpecificEnergyConsumption(?x, ?rsec), greaterThan(?raec, ?rsec) -> FacilityOutOfReferenceSpecificConsumptionRange(?x)
(BuildingElement or ManufacturingResource)(?x), isActive(?x, false) -> hasActualState(?x, "Off"^^string)
EnergyNetwork(?n), (BuildingElement or ManufacturingResource)(?x), isInNetwork(?x, ?n), hasReferencePeakPowerInput(?n, ?rppi) -> hasReferencePeakPowerInput(?x, ?rppi)
Controller(?c), ManufacturingJob(?j), ProductionFacility(?pf), hasController(?pf, ?c), operates(?pf, ?j), hasActualTime(?c, ?at), hasEndTime(?j, ?et), greaterThan(?at, ?et) -> hasActualStatus(?j, "done"^^string)
EnergyNetwork(?n), ManufacturingJob(J1-02), ManufacturingJob(J3-02), isInNetwork(?pf1, ?n), isInNetwork(?pf2, ?n), operatedOn(J1-02, ?pf1), operatedOn(J3-02, ?pf2), hasEndTime(J1-02, ?endtime1), hasRatedPowerInput(J1-02, ?rpi1), hasRatedPowerInput(J3-02, ?rpi2), hasStartTime(J1-02, ?starttime1), hasStartTime(J3-02, ?starttime2), add(?rtpd, ?rpi1, ?rpi2), greaterThanOrEqual(?starttime2, ?starttime1), lessThanOrEqual(?starttime2, ?endtime1) -> hasRatedOperationalTotalPowerDemand(?n, ?rtpd)
(EnergyConsumingAncillaryFacility or ProductionFacility)(?a), hasRatedMinPowerFactor(?a, ?rampf), lessThan(?rampf, 0.9f) -> FacilityHasHighReactivePower(?a)
EnergyConsumingAncillaryFacility(?x), hasCurrentType(?x, "DC"^^string), hasRatedEnergyConsumption(?x, ?raec), hasRatedPowerInput(?x, ?rpi), hasReferencePeakPowerInput(?x, ?rppi), hasReferenceSpecificEnergyConsumption(?x, ?rsec), lessThanOrEqual(?raec, ?rsec), lessThanOrEqual(?rpi, ?rppi) -> EnergyConsumingAncillaryFacilityEnergyEfficient(?x)
hasRatedTotalPowerDemand(?n, ?pd), hasReferencePeakPowerInput(?n, ?ppi), greaterThan(?pd, ?ppi) -> NetworkOverloadedByParallelWorkingFacilities(?n)
Fan(?f), hasRatedMaxEfficiency(?f, ?e), lessThan(?e, "0.6"^^double) -> InefficientFanUsed(?f)
ManufacturingJob(?j), ProductionFacility(?pf), operates(?pf, ?j), hasRatedPowerInput(?j, ?rpi) -> hasRatedPowerInput(?pf, ?rpi)
ProductionFacility(?f1), ProductionFacility(?f2), isEqualOperation(?op1, ?op2), operableOn(?op2, ?f2), operates(?f1, ?j1), operates(?f2, ?j2), operationMoreEfficient(?op2, ?op1), performsOperation(?j1, ?op1), hasEndTime(?j1, ?e), hasStartTime(?j2, ?s), lessThan(?e, ?s) -> ProductionFacilityEnergyWasting(?f1), moreEfficientOperatedOn(?j1, ?f2)
hasRatedEnergyIntensity(?z, ?rei), hasReferenceMagnitude(?z, ?rm), divide(?sec, ?rei, ?rm) -> hasRatedSpecificEnergyRequirement(?z, ?sec)
ManufacturingJob(?j), ProductionFacility(?pf), operatedOn(?j, ?pf), hasRatedEnergyIntensity(?j, ?raei) -> hasRatedEnergyIntensity(?pf, ?raei)
ProductionFacilityEnergyWasting(?x), hasRatedCapacityUtilisation(?x, ?rcu), hasReferenceMinCapacityUtilisation(?pf, ?rmicu), lessThan(?rcu, ?rmicu) -> ProductionFacilityHasPartload(?x)
hasRatedSpecificEnergyRequirement(?z, ?ser), hasReferenceSpecificEnergyConsumption(?z, ?rsec), divide(?eel, ?rsec, ?ser) -> hasEnergyEfficiencyLevel(?z, ?eel)
ProductionFacility(?x), hasCurrentType(?x, "AC"^^string), hasRatedEnergyConsumption(?x, ?raec), hasRatedEnergyIntensity(?x, ?raei), hasRatedMinPowerFactor(?x, ?rmpf), hasRatedPowerInput(?x, ?rapi), hasReferenceMaxEnergyIntensity(?x, ?rmaei), hasReferencePeakPowerInput(?x, ?rppi), hasReferenceSpecificEnergyConsumption(?x, ?rsec), greaterThan(?rmpf, 0.9f), lessThanOrEqual(?raec, ?rsec), lessThanOrEqual(?raei, ?rmaei), lessThanOrEqual(?rapi, ?rppi) -> ProductionFacilityEnergyEfficient(?x)
VoltageSensor(?s), (BuildingElement or ManufacturingResource)(?x), attachedSensor(?x, ?s), hasRatedVoltage(?x, ?vr), hasValue(?s, ?vs), lessThan(?vr, ?vs) -> VoltageExceedsRating(?x)
ManufacturingJob(?j), hasRatedPowerInput(?j, ?raec), hasReferencePeakPowerInput(?j, ?rpec), greaterThan(?raec, ?rpec) -> JobExceedsReferencePeakDemand(?j)
```

Figure 5.5 Examples of SWRL rules incorporated into the ontology

KDD Module

The data pre-processing algorithms are implemented as a Java project named *datapreprocessing*. The project contains the following Java classes:

- `DataCleaning` for removing NULL values, negative values, and outliers
- `Discretization` to split numeric values (e.g. power consumption) into a set of discrete ranges for each machine
- `FineGranulation` to granulate the metering data to 1-minute interval resolution since the metering data are collected periodically in different intervals ranging from 5 to 15 minutes
- `Transformation` to transform data to suit the schema
- `Merge` to merge process and metering data by relating the time stamps (see Figure 4.18)

This work implements the data mining algorithms as a Java project named *datamining* and utilizes a library called *Waikato Environment for Knowledge Analysis (WEKA)*[13]. WEKA is a Java based machine learning software suite and library developed at the University of Waikato in New Zealand. It contains different implementations of data mining techniques, including decision tree, M5P rules, and linear regression. For example, there are 11 implementations of decision tree based classification. This thesis performs a comparison to choose the most suitable decision tree implementation. The comparison is performed by testing the algorithm implementations with the real data coming from the manufacturing stainless steel manufacturing company. The decision tree implementations are compared based on the following criteria:

- Time to build a decision tree model
- The accuracy of the classifier to predict the class of given instances
- Tree visualization possibility to track the splits and the instances assignments

Table 5.1 presents the comparison results of different decision tree algorithms following the criteria above. It shows that the J48 requires relatively shorter time to build the decision tree model of the given data and good classification accuracy. It also provides visualization of the decision tree. Furthermore, J48 that implements the C4.5 algorithm is mostly used for the similar case in other work and offers generally effective learning algorithms [KiSK-11] [ShMe-12] (see Table 2.4 and Table 2.4). The M5P tree is implemented by utilizing the `M5P` class of WEKA library, whereas the linear regression uses `LinearRegression` class.

Table 5.1 Comparison results of different decision tree algorithms

Algorithm	Time to Build Model (seconds)	Accuracy (%)	Visualization of Decision Tree
BFTree	1.15	64.54	No
DecisionStump	0.08	63.68	No
FT	55.73	75.08	Yes
J48	0.71	75.71	Yes
J48graft	1.47	75.71	Yes
LADTree	12.07	67.34	Yes
LMT	103.26	75.63	Yes
NBTree	20.33	75.70	Yes
RandomForest	2.67	75.85	No
REPTree	0.76	75.51	Yes
SimpleCart	5.21	75.82	No

[13] http://www.cs.waikato.ac.nz/~ml/weka/

Forecast Module

The forecast module implements the prediction of power consumption and process duration by evaluating the SWRL rules stored in the knowledge base. The module employs Jess[14] rule engine to evaluate the SWRL rules. The calculations of predicted values based on outputs from rule engine are implemented as a Java program. The module applies Protégé-OWL API[15] to access the ontological elements in the knowledge base.

EnPI Calculator

The EnPI calculator uses Protégé-OWL API to load the production model from the ontological knowledge base. The calculation formulas are implemented as a Java program

Graphical User Interface (GUI)

This work develops GUIs for production scheduling, energy forecasting, and energy flow modeling including the EnPI calculation results. The following technologies are employed to develop the GUIs:

- *Eclipse Rich Client Platform (RCP)*[16] version 3.7.2 for building the desktop based GUIs including their basic and commonly used elements, such as frames, text editors, buttons, menu bars, drop-down lists, etc. Eclipse RCP allows building a professional-looking GUI quickly with operating system native look-and-feel and multiple platforms. Eclipse RCP works based on UI standard user interface frameworks *Standard Widget Toolkit (SWT)*[17] and *JFace*[18].

- *Eclipse Graphical Editing Framework (GEF)*[19] for the Energy Flow Modeler GUI. GEF allows graphical representation and visualization of an arbitrary model. The framework provides a basic structure for building graphical editor in a simple way. By means of the editor, it is possible to modify the model, for example editing the attributes or manipulating the connection among model elements. The framework also allows user interactions with the model, such as drag and drop and copy-paste, which can be performed through menu or toolbars.

14 http://www.jessrules.com/

15 http://protegewiki.stanford.edu/wiki/ProtegeOWL_API_Programmers_Guide

16 https://wiki.eclipse.org/Rich_Client_Platform

17 http://www.eclipse.org/swt/

18 http://wiki.eclipse.org/index.php/JFace

19 http://www.eclipse.org/gef/

- *Jaret timebars* version 1.49[20] for displaying the Gantt chart in production planning and scheduling GUI. The Jaret timebars is a Java package that provides a viewer for data model containing rows and intervals. It allows intuitive user interaction with the Gantt chart, for instance, drag and drop, resizing the intervals, changing the orientation, etc. The Jaret timebars works on SWT technology.

This work uses an existing open source ontology editor Protégé[21] to construct and manipulate the ontology knowledge base. It is not necessary to develop a new GUI since the functionalities in editing ontologies are already covered by Protégé.

5.1.3 Functionalities of the Prototypes in Terms of General Use Cases

In order to assure the functionalities, the prototypes require a PC with the 64-bit operating system, minimum 1 GHz processor and minimum 1 GB memory. The prototypes also require Java Runtime Environment (JRE) and an internet connection in order to communicate with the web services. The following sections elaborate the functionalities of the prototypes from user interaction perspective.

Order and Product Management

When the SERENE GUI is started, the user is asked to load the production system data. The data is loaded from the knowledge base. It also allows data loading from files. Then, the window for managing the products and orders will appear as shown in Figure 5.6. The main part of the window is a table containing a list of incoming orders. The table contains the information about the ordered products, the quantities of ordered products, start date to process the order, and the deadline of the orders (see Figure 5.6, no.1). The right part of the window shows the structures of products and the steps which have to be carried out to produce the products. The structures are displayed as trees (see Figure 5.6, no.2). The Figure 5.6 no.3 depicts the properties of the product, such as name, the machine where the product will be processed, the costs, etc., if the user selects a product in the right frame (no 2). The window shown in Figure 5.6 no.4 allows the user to create a new order and edit an existing order. He can specify the product quantity, the processing start date, and the deadline. The entries are blurred due to the data privacy of the pilot company.

20 http://jaret.de/timebars/

21 http://protege.stanford.edu/

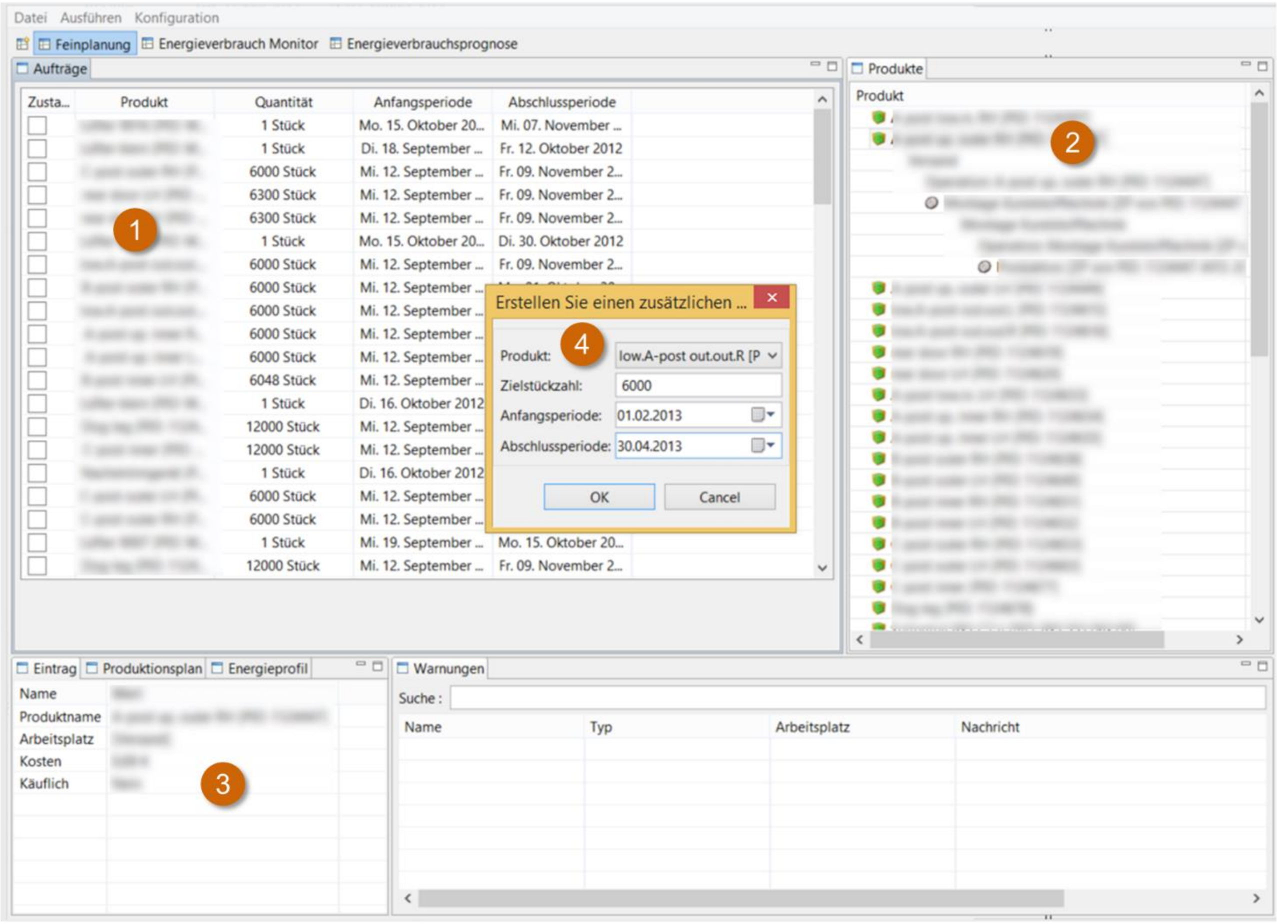


Figure 5.6 GUI managing the orders

Energy Forecast of Configurable Product

Once the production system and order data are loaded, the user can proceed to the next steps, for instance predicting the energy consumption and costs of the whole process to produces the ordered product. Figure 5.7 shows the screen shot of a GUI to perform the prediction. The GUI window consists of five main parts. Through the drop-down list shown in no. 1, the user can choose one of the incoming orders listed in Figure 5.6. The product properties are then displayed in the left part of the GUI (see no 2). The screen shot gives some example of energy and operation time influencing properties of a ring stainless steel, e.g. outer diameter, inner diameter, length, weight, and material type. The product image is also displayed in the part. The main part on the right side of the window presents a list of machines and workstations available on the shop floor for processing products. The predicted operation duration, machine cost, and power values to process the selected product for each machine are shown in this part. The different color highlights different machine groups. A group consists of a set of machines or workstations to execute a process type. The forecast module automatically selects the most efficient machine from each group. Nevertheless, the user is still allowed to select the machines that are not the most efficient ones. The energy and order processing costs are then calculated by the forecast module. The results are then displayed under the table (see no 4). Figure 5.7 no

5 gives an overview of rules and equations to the user which are used to perform prediction for each machine. They are loaded from the ontology knowledge base.

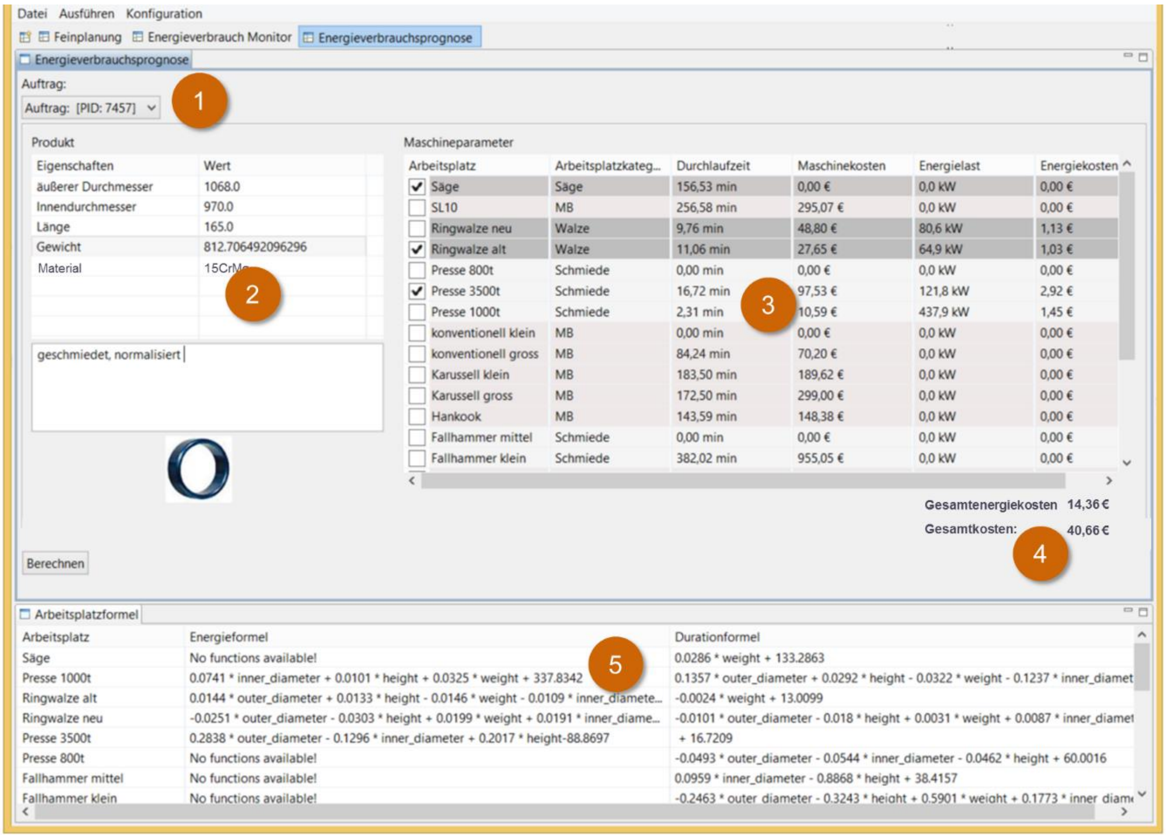


Figure 5.7 GUI for energy and cost forecasting

Energy Efficiency Qualitative Evaluation Workbench

The energy efficiency qualitative evaluation workbench is a GUI to helps an energy manager or production planner to evaluate the energy efficiency of the manufacturing system qualitatively by using the best practices and rules stored in the ontology knowledge base. By using the GUI depicted in Figure 5.8, the user can consult the knowledge base which processes or facilities in the manufacturing system have inconsistent or inefficient energy performance, and which ones are currently inefficient or consistent state.

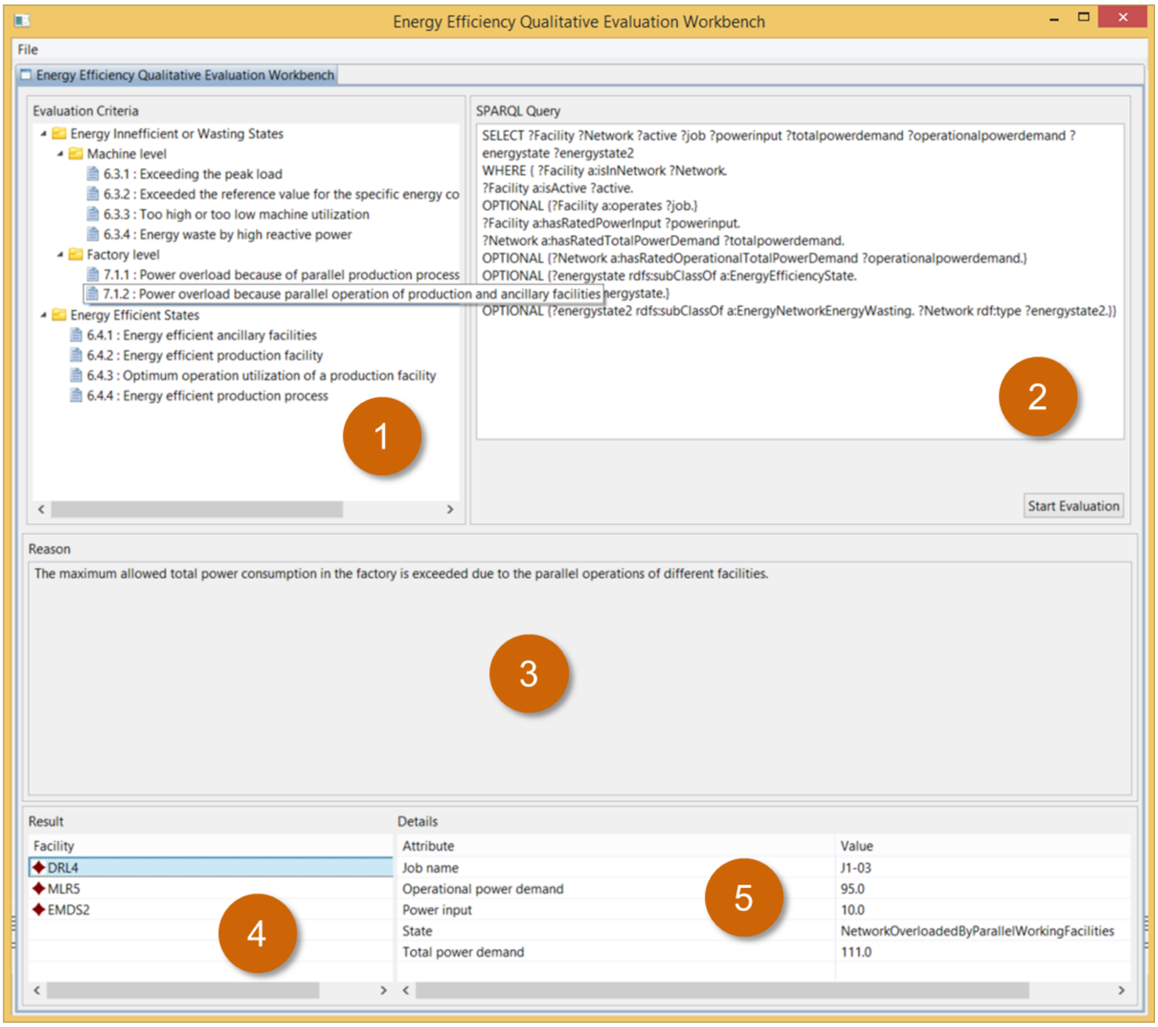


Figure 5.8 GUI of energy efficiency qualitative evaluation workbench

The tree structure shown in Figure 5.8 no 1 represents a catalog of evaluation criteria, for example, energy-inefficient and efficient criteria. Examples of the energy-inefficient criteria are the parallel execution of some processes and the excess of the maximum allowed power. The energy efficient criteria are for instance optimum operation utilization of production facilities and efficient operation of ancillary facilities. Each evaluation criterion corresponds to a SPARQL query which is displayed in the text area on the top right (see Figure 5.8, no 2). It also permits the user to edit the SPARQL query. If the user presses the button “Start Evaluation”, the objects that fulfill the selected evaluation criterion will be displayed in a table on the bottom left of the GUI (see Figure 5.8, no 4). Figure 5.8 shows the facilities that are inefficient states because they operate simultaneously and cause a greater power consumption than the maximum allowed power consumption. This reason is presented in the middle panel (see Figure 5.8, no 3). If the user selects one of the objects in the result table, the properties of the object are displayed in the bottom right table (see Figure 5.8, no 5)

EnPI Calculation

The EnPI Calculator loads the energy flow model from the ontology knowledge base and displays it in the GUI as illustrated in Figure 5.9. The GUI comprises four main frames. The top left frame presents the organization structure of an evaluated company as a tree (see Figure 5.9, frame no 1). It lists production lines and facilities involved in the production lines, i.e. production, ancillary, energy conversion, and transport facilities. The main part is the editor in the middle of the GUI (Figure 5.9, frame no 2). The facilities are represented as squares with different colors. The green, purple, yellow, gray squares indicate the production, ancillary, energy conversion, and transport facilities consecutively. The red square represents the external energy network from a utility company, whereas the cyan square indicates the internal energy network. The user can perform drag and drop on objects from the factory tree view to the editor to define the connections between them.

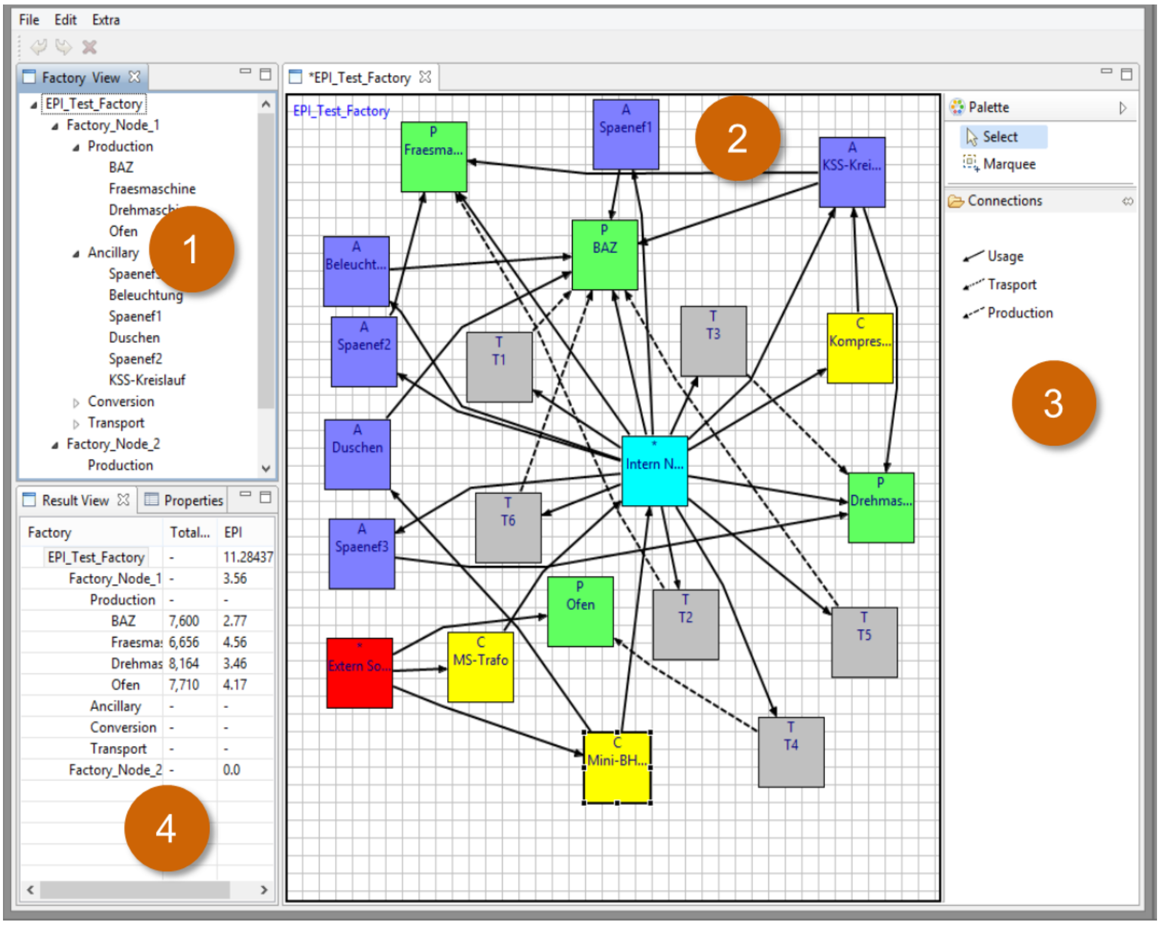


Figure 5.9 GUI for EnPI calculation

The palette on the right-hand side (see Figure 5.9, frame no 3) contains standard GEF selection tools, such as point selection and marquee. The connection tool allows the user to define three following connection types between objects in the editor:

- *Energy usage connection* to connect all type of facilities and networks resulting an energy flow model of the company.

- *Transport connection* to define the paths of transport facilities. It connects a transport facility to origin and destination production facilities.

- *Production connection* to specify preceding relations among production facilities reflecting a material flow

The bottom left frame (Figure 5.9, no 4) presents EnPI calculation results. It displays the company production organization as a tree and EnPI of each production facility, production line and the whole factory.

Production Planning and Scheduling

The GUI depicted in Figure 5.6 allows the user to select orders that will be processed on the shop floor. Once the orders are selected, the user can run the hyper-heuristic to create a schedule to process the selected orders. Figure 5.10 no 6 shows the GUI to start the hyper-heuristic. The GUI enables the user to specify different parameters of hyper-heuristic, such as the start date to process the selected orders, the maximum allowed power load involved in the schedule, and some additional constraints, for example, the importance weights (0 to 100) for avoiding peak load and night shift work. In order to find the optimum production schedule, hyper-heuristic could take a considerable amount of time, especially for large problems. Therefore, through the GUI, the user can specify the maximum duration of the hyper-heuristic to perform the search. If the search reaches the specified duration, the hyper-heuristics will stop and present the last found optimal schedule. The schedule is still a good solution, but it is not the global-optimum one.

The resulting schedule is also presented in the GUI depicted in Figure 5.10. The GUI consists of five parts. The top left part (Figure 5.10, no 1) lists the machines or workstations available on the shop floor. The top middle frame (Figure 5.10, no 2) displays the resulting schedule in a Gantt chart. The horizontal bars represents the operations executed on each machine. The length of a bar reflects the operation duration. The different colors in a bar represent different phases of the corresponding operation. The yellow, green, blue, and red parts indicate retooling, start, operating, and shut-down phases of the operation consecutively. The top right part of the GUI (Figure 5.10, no 3) displays the structure of available products as a tree. The bottom left (Figure 5.10, no 4) shows some information corresponding to the resulting schedule, such as the costs of all operations involved in the schedule, the total energy consumption and costs, the total duration of all operations, start and end time of the schedule, and EnPI of the schedule.

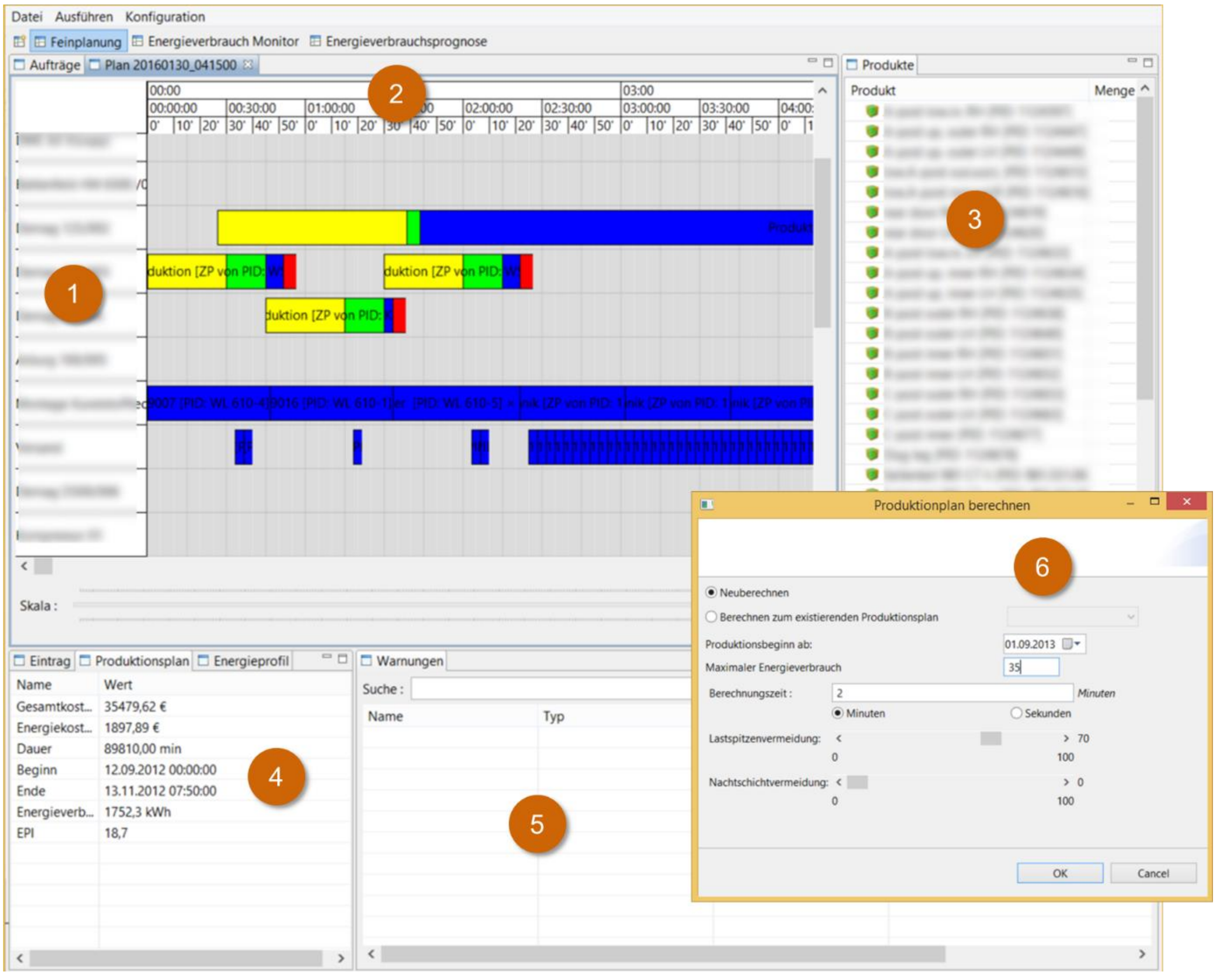


Figure 5.10 GUI of production planning and scheduling

The GUI allows the user to modify the schedule by performing drag and drop on products from the tree (Figure 5.10, no 3) to the Gantt chart (Figure 5.10, no 2), It will create an operation bar in the Gantt chart. The user can also resize a bar which means to change the operation duration. Furthermore, the user can also move a bar to another row to move the operation execution to another machine. The GUI evaluates the changes performed by the user and gives warnings displayed on the bottom right (Figure 5.10, no 5), if the changes are not valid, for example not enough stock, wrong machine, etc. The power consumption plot corresponding to the schedule is displayed in a window as seen in the screenshot in Figure 5.11. The user can select the plot of a machine selected from the drop-down list displayed on the top left.

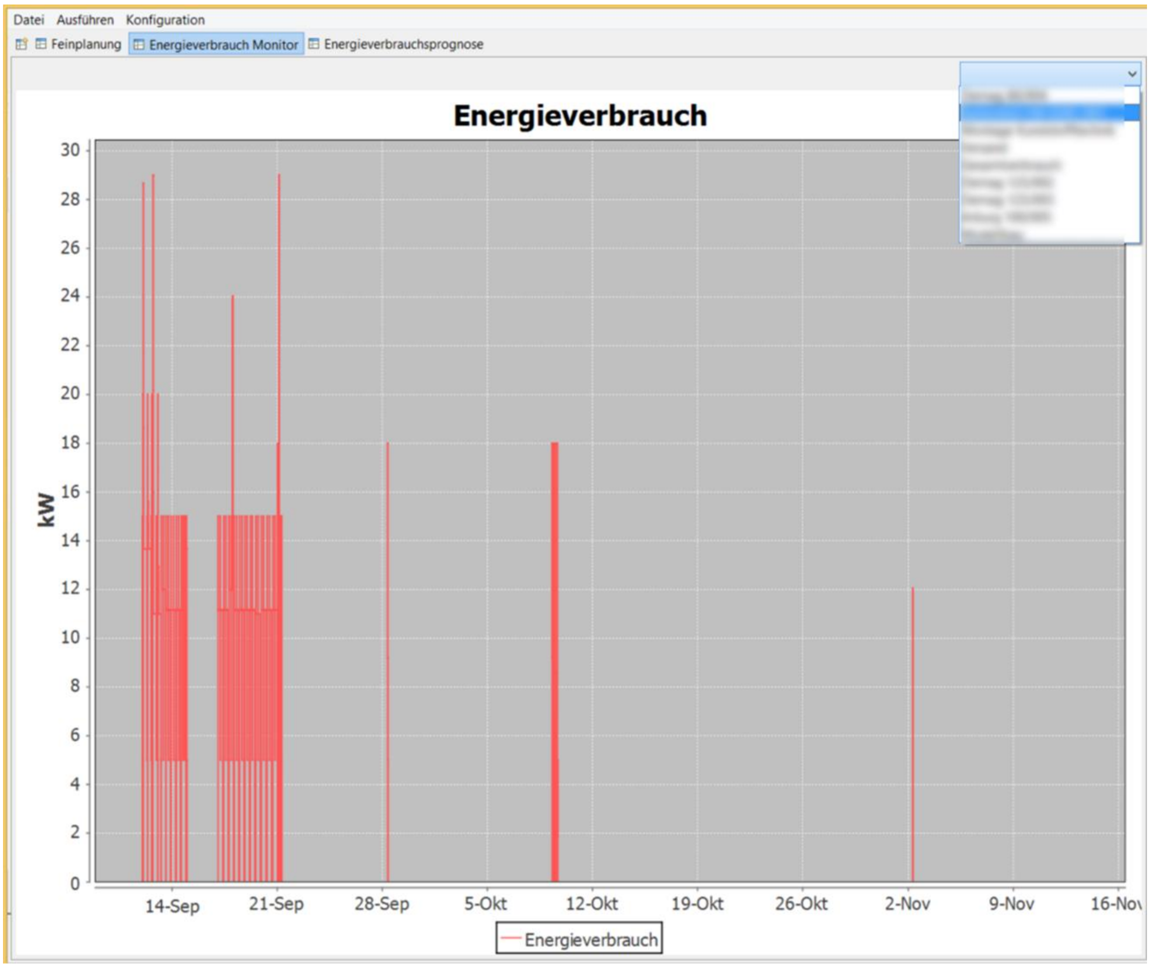


Figure 5.11 GUI showing corresponding predicted power consumption of a production plan

5.2 Case Examples in Manufacturing Companies

This section reports experiences gained from the use of the software in several types of manufacturing, i.e. make-to-order manufacturing, serial manufacturing, and a manufacturing environment using assessable ancillary, transport, and energy conversion facilities. Due to data privacy, some of the names and IDs of the objects involved in the scenarios are altered.

5.2.1 Use Case of a Production Having Energy Conversion, Transport, and Ancillary Facilities

EnPI Calculation

The aim of the first use case is the validation of the developed EnPI by applying the energy flow modeler prototype. A factory part of an automotive manufacturing company is chosen to gather application experience with the prototype. The factory consists of several production lines. However, this thesis uses only one production line for validation. The production line comprises energy conversion, ancillary, production, and transport facilities that make up a system model for EnPI calculation (see section 4.4.1). Table 5.2 describes the facilities in the factory, the energy form they consume or produce, the origin of their inputs, and the destination and values of their outputs. The table lists the percentages of energy generated by ancillary facilities that are consumed by production facilities. It also indicates the destination to which goods are delivered by transport facilities. As seen in the table, the total production value and the production value per hour are specified as well.

The facilities are interconnected using the EnPI Calculation GUI. Figure 5.12 shows the screen shot of the editor of the EnPI Calculation GUI containing the interconnections among the facilities. The interconnections reflect the energy flow model of the factory. In addition to the energy conversion (yellow), ancillary (purple), production (green), and transport (gray) facilities, the external energy source (red), and the internal energy network (cyan) are modeled. Most of the energy conversion facilities supply the energy to other facilities indirectly through the internal energy network. Some of them, for example, the compressor transfers the energy directly to a facility. Most of the facilities get their energy from the internal energy network as well. The production line indicates the production facility precedence relation Oven – Cutting Machine – Rolling Machine – Drilling Machine.

Table 5.2 List of facilities in the factory and their inputs and outputs

		Input		Output	
	Machine	**Energy Form**	**Origin**	**Energy Form**	**Destination**
Energy Conversion Facilities	Transformator	electricity	external network	electricity	internal network
	BHKW	gas	external network	electricity heat	internal network Water Heater
	Compressor	electricity	internal network	compressed air	Cooling Circuit
	Machine	**Energy Form**	**Origin**	**Percentage**	**Destination**
Ancillary Facilities	Cooler1	electricity	internal network	100	Drilling Machine
	Cooler2	electricity	internal network	100	Cutting Machine
	Cooler3	electricity	internal network	100	Rolling Machine
	Cooling Circuit	compressed air electricity	Compressor internal network	30 30 40	Drilling Machine Cutting Machine Rolling Machine
	Lighting System	electricity	internal network	100	Drilling Machine
	Water Heater	heat	BHKW	100	Drilling Machine

	Machine	Energy Form	Origin	Destination	
Transport Facilities	Forklift1	electricity	internal network	Drilling Machine	
	Forklift2	electricity	internal network	Cutting Machine	
	Forklift3	electricity	internal network	Rolling Machine	
	Forklift4	electricity	internal network	Oven	
	Forklift5	electricity	internal network	Drilling Machine	
	Forklift6	electricity	internal network	Drilling Machine	
	Machine	**Energy Form**	**Origin**	**Production Value [€]**	**Production Value per Hour [€]**
Production Facilities	Oven	gas	external network	600	100
	Cutting Machine	electricity	internal network	300	75
	Rolling Machine	electricity	internal network	300	75
	Drilling Machine	electricity	internal network	450	90

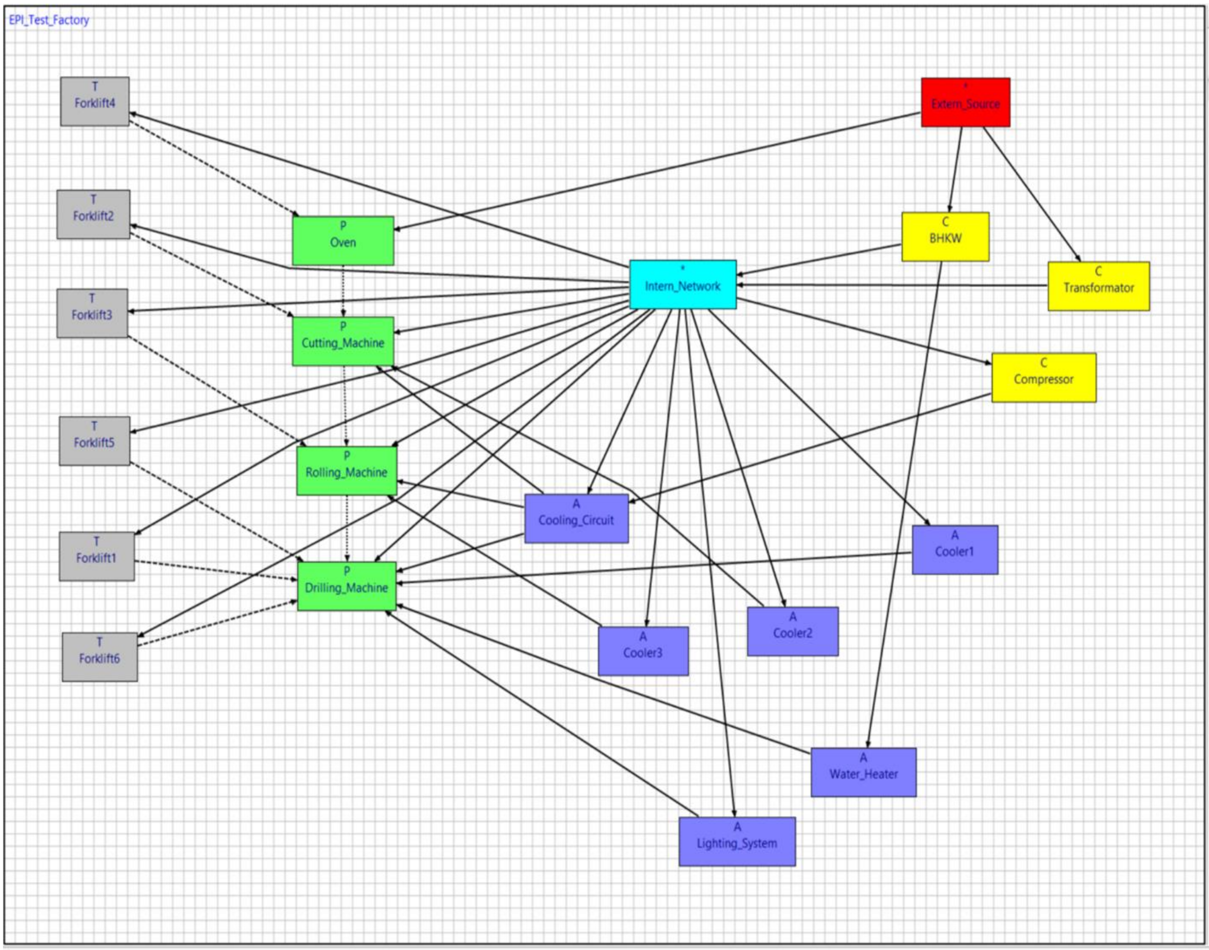


Figure 5.12 Energy flow model of the pilot manufacturing

The EnPI calculator module calculates the EnPIs based on the energy flow of the factory depicted in Figure 5.12. The EnPIs are computed for the factory, the production line, and the four machines, i.e. Oven, Cutting Machine, Rolling Machine, Rolling Machine, and Drilling Machine. The resulting EnPI are shown in the GUI and depicted in Figure 5.13.

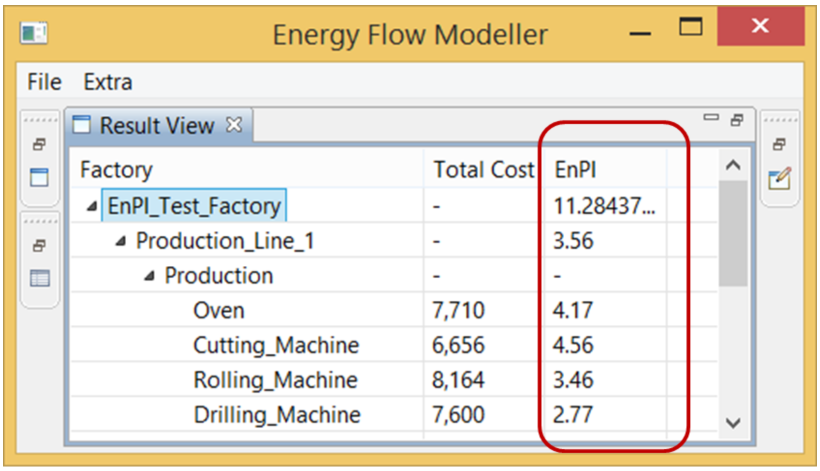


Figure 5.13 EnPI values of the factory, production line and machines

By using the EnPI, energy performance between the machines can be benchmarked. As seen in Figure 5.13, the Cutting Machine has the highest EnPI, whereas the Drilling Machine has the lowest one. The production line has a fair EnPI value. It can be improved by replacing some facilities by more energy-efficient ones or by reconfiguring the factory layout to reduce the distances between production facilities. Table 5.3 gives some examples of changes that are considered to improve energy efficiency.

Table 5.3 List of measures to improve the energy performance of the factory

Change Description	Changed Parameter	Old Value	New Value
Substitute the transformator with the new one having higher efficiency	Input power	5000 kW	4950 kW
	Output power	4750 kW	4900 kW
Use energy efficient cooling circuit	Max input power	16 kW	2 kW
Substitute the drilling machine with a more efficient one	Max input power	50 kW	10 kW
	Energy consumption	150 kWh	30 kWh
Change the factory layout	Destination of Forklift4 to Oven	30 m	10 m
	Destination of Forklift5 to Drilling Machine	20 m	10 m

After the changes have been carried out, the EnPIs are improved compared to the original ones which illustrated in Figure 5.14. The replacement of the drilling machine causes the EnPI on the corresponding machine level to be doubled. The change of factory layout makes the travel distance to the oven shorter. This reduces the energy required for transporting the goods to the oven and the effort to process the products is also lower. As a result, the EnPI of the oven is increased. The introduction of a new energy-efficient cooling circuit slightly improves the EnPI of the cutting machine and rolling machine, as the cooling circuit is used for both machines and the drilling machine (see Figure 5.12). The exchange of the transformer improves the energy performance of the production line since the output energy flows to the internal network. All of these measures improve the EnPI of the production line and the factory by up to 40% and 27%, respectively. The difference between the EnPI of the factory level and the production line is due to the contribution of the facilities in the factory that is not included in the evaluation of the production line.

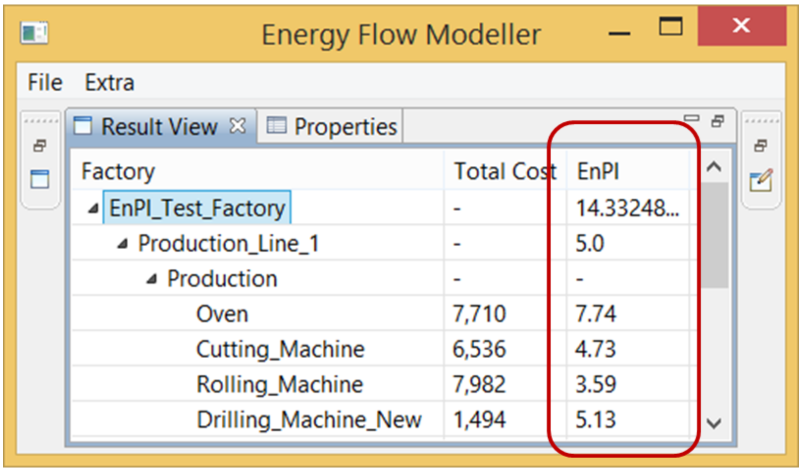


Figure 5.14 EnPI values of the factory, production line and machines after the changes

An EnPI value indicates the ratio of the generated value of a machine, production line, or factory to energy costs. A figure exceeding one does not automatically mean a positive contribution margin, as there are other costs that must be covered in addition to the energy costs.

Theoretically, the EnPI figures are always positive rational numbers, if it is assumed that both production value and energy costs cannot be negative. A value below one means that the energy cost exceeds the value and should be avoided in practice. Otherwise, it would be very critical since the process is extremely inefficient. In the validation example, the calculated EnPIs range from 2 to 15. Higher values are conceivable when a process generates very high value and/or requires much less energy.

The EnPI developed in this thesis is not aimed at measuring the economic performance of a factory, production lines, or machines. It is very useful to simulate and compare the energy performances of different configuration alternatives of production, ancillary, conversion, and transport facilities used in manufacturing. The energy costs are distributed to the average operation duration per calculation or measurement period. The calculation is made once in a certain period and is independent of operation duration. The assessment of highly utilized facilities, hence, is more accurate than that of the facilities having a short running time. This kind of correlation seems to make sense. Thus, the planning of machine usage has to consider high machine utilization in order to reach a positive impact on the energy efficiency. In order to perform the EnPI calculation, the average operating time per calculation period should be estimated based on the state of the selected machine. Simultaneously, the selection of different machines improves the EnPI value of the factory or the production line [WiEt-13].

Knowledge-based Qualitative Evaluation

This use case validates the concept of qualitative evaluation using ontology knowledge based which incorporates a set of rules formalizing best practices in energy efficiency. Four production facilities and two ancillary facilities of the factory configuration depicted Figure 5.12 in are put into focus in this use case. At a certain point of time, each production facility executes a job as illustrated in Figure 5.15. The cutting, rolling, drilling machines, and the two coolers are connected to the internal electricity network.

Table 5.4 shows the ontological relationships between the facilities, jobs, and the internal network that are represented as OWL individuals. Each time the measurement values are retrieved from the industrial automation and manufacturing IT systems through the Data Collection Interface, the data property values of those individuals are updated. Table 5.5 illustrates the data property values at two different points of time (*T1* and *T2*). For the sake of simplicity, Table 5.5 presents only the most relevant values for the thesis validation.

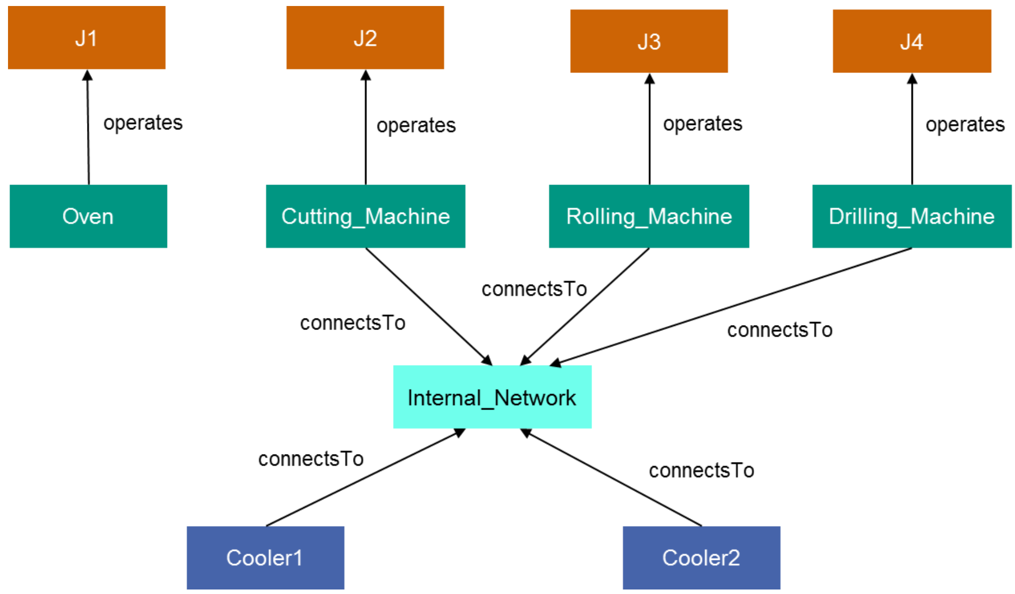


Figure 5.15 Relationships between objects in the manufacturing system of the pilot

Table 5.4 List of relations among objects in the manufacturing

Subject		Predicate	Object	
Instance	Class	ObjectTypeProperty	Instance	Class
Oven	ProductionFacility	operates	J1	ManufacturingJob
Cutting_Machine	ProductionFacility	operates	J2	ManufacturingJob
		connectsTo	Intern_Network	InternalNetwork
Rolling_Machine	ProductionFacility	operates	J3	ManufacturingJob

		connectsTo	Intern_Network	InternalNetwork
Drilling_Machine	ProductionFacility	operates	J4	Manufacturing Job
		connectsTo	Intern_Network	InternalNetwork
Cooler1	AncillaryFacility	connectsTo	Intern_Network	Internal Network
Cooler2	AncillaryFacility	connectsTo	Intern_Network	Internal Network

Table 5.5 List of property values at two different measurement times (T1 and T2)

Name	**Property**	**Value at t=T1**	**Value at t=T2**
J1	hasPowerConsumption	102.4 kW	**95.4 kW**
	hasReferenceMaxPowerConsumption	85 kW	85 kW
Oven	hasPowerConsumption	102.4 kW	**95.4 kW**
	hasReferenceMaxPowerConsumption	100 kW	100 kW
	hasCapacityUtilization	0.45	**0.65**
	hasReferenceMinCapacityUtilization	0.60	0.60
Cutting_Machine	hasEnergyConsumption	4.5 kWh	**6.1 kWh**
	hasReferenceSpecificEnergyConsumption	5 kWh	5 kWh
Rolling_Machine	hasEnergyIntensity	15 kWh/prod	**9.8 kWh/prod**
	hasReferenceMaxEnergyIntensity	10 kWh/prod	10 kWh/prod
	hasPowerConsumption	2 kW	**11 kW**
	hasReferenceMaxPowerConsumption	20 kW	20 kW
Drilling_Machine	hasEnergyIntensity	8.5 kWh/prod	**21 kWh/prod**
	hasReferenceMaxEnergyIntensity	10 kWh/prod	10 kWh/prod
	hasPowerConsumption	21.5 kW	**84 kW**
	hasReferenceMaxPowerConsumption	50 kWh	50 kWh
	hasMinPowerFactor	0.94	**0.75**
J4	hasPowerConsumption	21.5 kW	**84 kW**
	hasReferenceMaxPowerConsumption	35 kW	35 kW
	hasEnergyIntensity	8.5 kWh/prod	**21 kWh/prod**
	hasReferenceMaxEnergyIntensity	9 kWh/prod	9 kWh/prod
Cooler1	hasPowerConsumption	1.7 kWh	**1.8 kWh**

	hasReferenceMaxPowerConsumption	3 kWh	3 kWh
Cooler2	hasPowerConsumption	5.2 kWh	**2.2 kWh**
	hasReferenceMaxPowerConsumption	3 kWh	3 kWh
InternalNetwork	hasReferenceMaxPowerConsumption	100 kW	100kW

To perform the qualitative evaluation of the energy efficiency, the user uses the Energy Efficiency Qualitative Evaluation Workbench (see Section 5.1.2). The Jess rule engine executes the SWRL rules resulting in the actual states of energy efficiency. By using the workbench, the user can find which objects are evaluated by the knowledge base as energy efficient or inefficient. Table 5.6 demonstrates the states of the manufacturing objects at *T1* that are retrieved through the workbench. It can be seen that several facilities are in energy-inefficient states due to exceeding the maximum allowed power consumption and in a production overload state, for instance, the oven and rolling machine. However, some of the facilities still operate in normal or energy efficient modes, e.g. cutting and rolling machines. As seen in the table, it is also possible to evaluate the energy efficiency of manufacturing jobs by using the workbench.

Table 5.6 Evaluation given by the knowledge base at t=T1

Manufacturing Objects	Evaluated State	Reasons
J1	JobExceedsReferencePeakDemand; EnergyWasting	The current power consumption of the job equals 102.4 kW. It exceeds the maximum allowed power consumption of 85 kW.
Oven	FacilityExceedsReferencePeakDemand; EnergyWasting; ProductionFacilityIncon-sistentEnergyPerformance	The current power consumption of the machine is 102.4 kW. It exceeds the maximum allowed power consumption of 100 kW.
Cutting_Machine	ProductionFacilityEnergyEfficient	The measured energy consumption of the facility (4.5 kWh) is lower than the defined facility specific energy consumption (5 kWh).
Rolling_Machine	ProductionFacilityHasOverload	The energy intensity of the production facility (15 kWh/prod unit) exceeds the reference maximum energy intensity (10 kWh/prod).
Drilling_Machine	ProductionFacilityEnergyEfficient	The measured energy intensity of the production facility (8.5 kWh/prod unit) is less than the maximum allowed

		energy intensity (10 kWh/prod unit); the measured energy consumption of the production facility (21.5 kWh) is less than the defined facility specific energy consumption (50 kW); the measured power factor of the production facility (0.94) is greater than the reference power factor cos φ = 0.9 [Ries-12].
`J4`	`JobEnergyEfficient`	The measured energy intensity of the job (8.5 kWh/prod) is less than or equal to the maximum allowed energy intensity (9 kWh/prod unit) and the power consumption of the job (11.5 kW) is lower than or equal to the maximum allowed power consumption 15 kW).
`Cooler1`	`EnergyConsumingAncillaryFacilityEnergyEfficient`	The measured power consumption of the ancillary facility (1.7 kW) is less than the defined maximum specific power consumption (3 kW).
`Cooler2`	`EnergyConsumingAncillaryFacilityEnergyWasting`	The measured power consumption of the ancillary facility (5.2 kW) is greater than the defined maximum specific power consumption (3 kW).

At *t=T2*, the property values change as depicted Table 5.5. And again, the rules are evaluated and resulting in new states. Table 5.7 presents the states evaluated using the Energy Efficiency Qualitative Evaluation Workbench at *t=T2*. It shows that the oven operates in energy efficient state, but the power consumption value of *J1* is still above the allowed threshold. The drilling machine and its executed job *J4* also experience exceeding peak load state. Additionally, the total power consumption of facilities connected to the internal network exceeds the maximum allowed power consumption of the network.

Table 5.7 Evaluation given by the knowledge base at t=T2

Manufacturing Objects	**Evaluated State**	**Reasons**
`J1`	`JobExceedsReferencePeakDemand; EnergyWasting`	The current power consumption of the job is 95.4 kW. It exceeds the maximum allowed power consumption of 85 kW.
`Oven`	`ProductionFacilityInconsistentEnergyPerformance`	The current power consumption of the machine is 95.4 kW. It is less than the maximum allowed power consumption of 100 kW, but it exceeds the usual power consumption in processing the current product based on the rules from KDD.

Cutting_Machine	FacilityExceedsReferencePeakDemand; EnergyWasting; NetworkOverloadedByParallelOperatingFacilites	The measured energy consumption of the facility (6.1 kWh) is greater than the defined facility specific energy consumption (5 kWh).
Rolling_Machine	ProductionFacilityEnergyEfficient; NetworkOverloadedByParallelOperatingFacilites	The energy intensity of the production facility (9.8 kWh/prod unit) is less than maximum energy intensity (10 kWh/prod). The total power consumption (105.1 kW) exceeds the maximal allowed power consumption in the internal network (100 kW) due to the parallel operations of different facilities.
Drilling_Machine	FacilityExceedsReferencePeakDemand; NetworkOverloadedByParallelOperatingFacilites; FacilityHasHighReactivePower	The measured energy intensity of the production facility (21 kWh/prod unit) exceeds the maximum allowed energy intensity (10 kWh/prod unit); the measured energy consumption of the production facility (84 kW) exceeds the defined facility specific energy consumption (50 kW); the measured power factor of the production facility (0.75) is less than the reference power factor cos φ = 0.9 [Ries-12]. The total power consumption (105.1 kW) exceeds the maximal allowed power consumption in the internal network (100 kW) due to the parallel operations of different facilities.
J4	JobExceedsReferencePeakDemand; EnergyWasting	The measured energy intensity of the job (21 kWh/prod) is greater than the maximum allowed energy intensity (9 kWh/prod unit) and the power consumption of the job (84 kW) exceeds the maximum allowed power consumption (35kW).
Cooler1	EnergyConsumingAncillaryFacilityEnergyEfficient; NetworkOverloadedByParallelOperatingFacilites;	The measured power consumption of the ancillary facility (1.8 kW) is less than the defined maximum specific power consumption (3 kW). The total power consumption (105.1 kW) exceeds the maximal allowed power consumption in the internal network (100 kW) due to the parallel operations of different facilities.
Cooler2	EnergyConsumingAncillaryFacilityEnergyEfficient;	The measured power consumption of the ancillary facility (2,2 kW) is less than the defined maximum specific power consumption (3 kW). The total power

	`NetworkOverloadedByParallelOperatingFacilites;`	consumption (105.1 kW) exceeds the maximal allowed power consumption in the internal network (100 kW) due to the parallel operations of different facilities.

The validation in the manufacturing system has shown the positive contributions of the ontological knowledge base in supporting the qualitative evaluation of the manufacturing objects. By consulting the rules in the ontological knowledge base through the Energy Efficiency Qualitative Evaluation Workbench, the energy efficiency states of the objects can be automatically evaluated. This can accelerate the decision-making process to implement the measures to improve the energy efficiency. The references for energy efficiency evaluation are stored and always available in the knowledge base. They can be retrieved anytime for the comparison to the actual states of the manufacturing objects.

Anyone who has access to the Energy Efficiency Qualitative Evaluation Workbench can choose the evaluation criteria and retrieve the facilities or processes which are in energy-inefficient states, for example having power consumption exceeding the maximum allowed peak load, overloaded jobs, etc. However, using the developed prototype, it is not possible to make changes to the inefficient facilities or processes since the industrial automation system does not provide control interfaces. The changes have to be performed manually. After changes are made to the facilities and processes, a quantitative evaluation using the EnPI can be carried out to assess whether the changes give positive consequences at the machine, production line, and factory level of the manufacturing system.

5.2.2 Use Case in a Make-to-Order Manufacturing

The concepts developed in this thesis are validated at several German medium-sized manufacturing companies. One of them is a stainless steel manufacturer. The company is a make-to-order manufacturer. In order to apply the concept, it is necessary to understand the product ranges and production processes of the company. Figure 5.16 depicts the general steps to manufacture the stainless steel products. As shown in the figure, the customers are allowed to specify the ordered products. The specifications include the material type, the dimension of the products, i.e., height, inner and outer diameter, and at which temperature the product should be heated. There are three geometrical categories of products, i.e. disks, rings, and bars [WiOv-12].

The raw material has to be processed in several steps to produce a final product. First, the raw material is sawed to get the rough dimension of the final product. Then, a forging process is performed on the material. If the desired product is not a bar, i.e. a disk or ring, it has to go through a rolling process afterward. Otherwise, it goes through a heat treatment process in an oven. Next, after the material is taken out from the oven, a machining process has to be carried out. Less than 1% of the ordered products do not go through the forging, rolling, and heat treatment processes, but directly to the machining process. Two final steps that have to be performed are quality assurance and finalizing.

As depicted in Figure 5.16, there are several choices of machines, on which a production process can be carried out, in each step. The machine chosen depends strongly on the product specifications. For example, forging, there are five different machine options, i.e. small drop hammer, medium drop hammer, press 800 tons, press 1000 tons, and press 3500 tons. The data mining algorithm automatically selects the optimal machine for each process step (see Section 4.3.3).

Table 5.8 presents the list of available machines on the shop floor, categorized by the process step. It shows that the IT system of the company is unable to provide machine and process log data for sawing and the heat treatment process. The reason is that an oven simultaneously processes several products for different orders, which is why it is difficult to record the data. There also is no system at the company that is able to register the sawing processes. As seen in the table, it is not necessary to measure the energy consumption of manual processes, such as sawing, quality assurance, and finalizing.

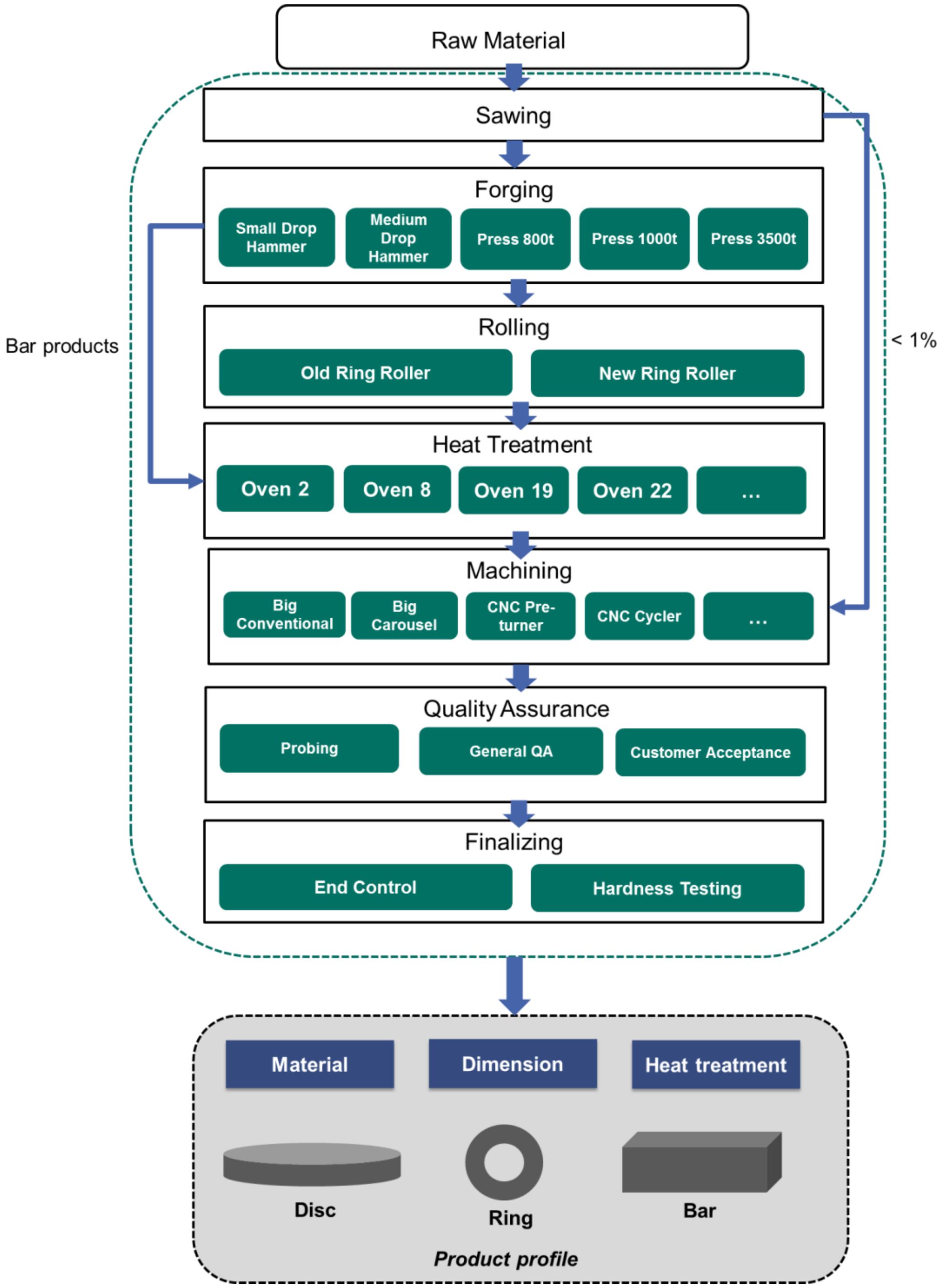


Figure 5.16 Product and process profile of the pilot company [WiOv-12]

Table 5.8 List of available machines categorized by operation step

Operation step	Machine or Workstation	Machine cost per hour	Number of personnel	Number of shift work per day	Process logging	Energy metering
Sawing	n.a.	n.a.	n.a.	n.a.	no	no
Forging	Small Drop Hammer	150 €	3	1	yes	yes
	Medium Drop Hammer	175 €	3	1	yes	yes
	Press 800t	175 €	3	2	yes	yes
	Press 1000t	275 €	3	1	yes	yes
	Press 3500t	350 €	3	1	yes	yes
Rolling	New Ring Roller	300 €	3	1	yes	yes
	Old Ring Roller	150 €	3	1	yes	yes
Heat treatment	Oven 2	n.a.	n.a.	n.a.	no	yes
	Oven 3	n.a.	n.a.	n.a.	no	yes
	Oven 4	n.a.	n.a.	n.a.	no	yes
	Oven 6	n.a.	n.a.	n.a.	no	yes
	Oven 8	n.a.	n.a.	n.a.	no	yes
	Oven 18	n.a.	n.a.	n.a.	no	yes
	Oven 19	n.a.	n.a.	n.a.	no	yes
	Oven 20	n.a.	n.a.	n.a.	no	yes
	Oven 21	n.a.	n.a.	n.a.	no	yes
	Oven 22	n.a.	n.a.	n.a.	no	yes
Machining	Big Conventional Machine	50 €	1	1	yes	yes
	Small Conventional Machine	50 €	1	1	yes	yes
	Big Carousel	104 €	1	2	yes	yes
	Small Carousel	62 €	1	1	yes	yes
	Hankook	62 €	1	2	yes	yes
	CNC Cycler	65 €	1	1	yes	yes
	CNC Billower	76 €	1	1	yes	yes
	CNC Pre-turner	54 €	1	2	yes	yes
	CNC Post-turner	81 €	1	2	yes	yes
	Large Bohrwerk	107 €	1	1	yes	yes
	Manual Processing	81 €	1	1	yes	no
Quality Assurance (QA)	Probing	300 €	n.a.	n.a.	yes	no
	General QA	50 €	n.a.	n.a.	yes	no
	Customer acceptance test	1 €	n.a.	n.a.	yes	no
Finalizing	End control	25 €	n.a.	n.a.	yes	no
	Hardness testing	90 €	n.a.	n.a.	yes	no

Because of lacking data, it is not possible to relate the process and energy consumption data for sawing, heat treatment, quality assurance, and finalizing processes. This work only validates the concepts developed for forging, rolling, and machining processes. It only considers ring products for concept validation due to the available data and support by the company. Table 5.9 lists the results of the energy and machine cost prediction of eight orders.

Table 5.9 Results of energy and process costs prediction in the make-to-order stainless steel manufacturing company

Order No		**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Product Profile	**Outer Diameter (cm)**	620.00	710.00	223.00	520.00	648.00	1042.00	335.00	308.00
	Inner Diameter (cm)	535.00	603.00	199.00	360.00	533.00	1001.00	312.00	222.00
	Length (cm)	46.00	60.00	18.00	176.00	68.00	83.00	18.00	80.00
	Weight (kg)	111.37	207.88	4.50	611.13	227.76	171.45	6.61	89.93
	Material	St 37-2	St 37-2	A Gr. 5	St 37-2	St 37-2	St 37-2	St 37-2	St 37-2
Forging	**Machine**	Press 1000t	Press 3500t	Small Drop Hammer	Press 3500t	Press 3500t	Press 1000t	Small Drop Hammer	Press 1000t
	Duration (minutes)	18.30	16.72	20.41	17.41	16.72	13.20	21.54	12.53
	Machine Cost (€)	83.88	97.53	51.03	101.56	97.53	60.50	53.85	57.43
	Avg. Power (kW)	381.10	46.60	52.50	66.70	39.70	418.50	53.22	358.00
	Energy Cost (€)	10.00	1.12	1.54	1.66	0.95	7.92	1.64	6.43
Rolling	**Machine**	Old Ring Roller	Old Ring Roller	New Ring Roller	Old Ring Roller	Old Ring Roller	New Ring Roller	New Ring Roller	Old Ring Roller
	Duration (minutes)	12.76	12.51	11.70	9.16	12.46	9.76	12.99	12.79
	Machine Cost (€)	31.90	31.28	58.50	22.90	31.15	24.40	32.48	31.98
	Avg. Power (kW)	72.70	71.10	75.40	78.30	70.80	71.50	71.20	71.40
	Energy Cost (€)	1.33	1.27	1.26	1.03	1.26	1.00	1.33	1.31
Machining	**Machine**	CNC Pre-Turning	CNC Pre-Turning	Big Bohrwerk	CNC Pre-Turning	CNC Pre-Turning	CNC Cycler	Big Bohrwerk	CNC Cycler
	Duration (minutes)	388.20	422.50	65.40	455.00	431.50	190,00	66,40	190,00
	Machine Cost (€)	349.38	380.25	116,63	409,50	388.35	205.83	118.41	205.83

Avg. Power (kW)	14.20	15,80	13,50	14,30	15,30	16,90	15.40	15.12
Energy Cost (€)	7.90	9.57	1.27	9.33	9.46	4.60	1,47	4,12
Energy costs	**19.23**	**11.96**	**4.07**	**12.02**	**11.68**	**13.52**	**4.43**	**11.86**
Machine costs	**465.16**	**509.06**	**226.16**	**533.96**	**517.03**	**290.73**	**204.74**	**295.24**

As seen in Table 5.9, the machine selection for forging depends on the weight. The light products are predicted to be forged using Small Drop Hammer, whereas the products having medium and heavy weight are predicted to be processed with Press 1000t and Press 3500t consecutively. Although the machine cost per hour of the Small Drop Hammer is the lowest, it takes more time to forge a product. The machine cost per hour of Press 3500t has the highest machine cost per unit, but it carries out the forging process fastest. It consumes less energy than the Press 1000t.

The thin rings i.e. the rings having lower differences between outer and inner diameters, will be rolled with the New Ring Roller. The difficulty level of rolling a thinner ring is higher so that the new ring roller equipped with more advanced technology can perform the task more efficient than the older one. The power consumptions of both rollers are similar and do not depend much on the product properties. The machining process depends on the dimension and the weight of the ring products. Large and heavy products have to be processed using CNC Pre-Turning, and it will take relatively long. Thus, it will consume more energy. The smaller and light rings will be processed by CNC Cycler or Big Bohrwerk. The lightest products are predicted to go through Big Bohrwerk and to consume the least energy. As seen in the table, the total energy and machines costs depend strongly on the process duration. The products which will be processed by the CNC Pre-Turning machine and take the longest time, have the highest predicted machine and energy cost.

At a make-to-order manufacturing company, such as a stainless steel manufacturer, there are many ways to process a customer order. The application experience of the company shows that the developed prototype helps the production planner to decide how to process the order such that energy consumption and costs are lowest. The prototype gives the production planner recommendations for the selection of machines for each step. Moreover, it is also difficult to estimate the costs of a make-to-order product since they are influenced by many variables. The prototype assists the company to estimate the costs involved in processing the order so it can determine the price of the product more exactly. Additionally, energy consumption of each process step and of the whole cycle can be predicted. However, due to the unavailability of some data and support from existing IT systems in the company, the prediction cannot be performed for all process steps. For this reason, the total energy and machine costs cannot be calculated accurately.

5.2.3 Use Case in a Serial Manufacturing

The third use case is a validation of the developed approach at a serial manufacturing company. The company is a German SME and produces synthetic material and plastics for automobile components. The manufacturing process at the company is quite simple and shown in Figure 5.17. When the company receives an order, it will be checked whether the same or similar products were produced previously. If the ordered product is a new one, it has to go through the construction step. For example, the product has to be sketched with CAD first. Since the shape and size of a product might be different, the production machine processing the product has to be retooled before the main operation starts. The product is then assembled in an assembly workstation. Finally, the product goes to the shipping center and is ready for delivery. The construction, production, and assembly operations can be executed on different machines or workstations. Table 5.10 presents a list of machines available for each operation step. For example, the production step can be performed on eleven machines that are associated with different costs. The machine names are changed for the sake of data privacy.

The production planner has to schedule the jobs on the shop floor in order to process the orders following the defined deadline. This is however not an easy task, especially if he also has to consider to minimize costs and maximize the energy efficiency. The aim of applying the developed prototype in the company is to help the production planner to find the cost and energy optimized production schedule by still considering the available required stocks and different shift-work model. It also intends to create a schedule with the constraint of the maximum allowed power load.

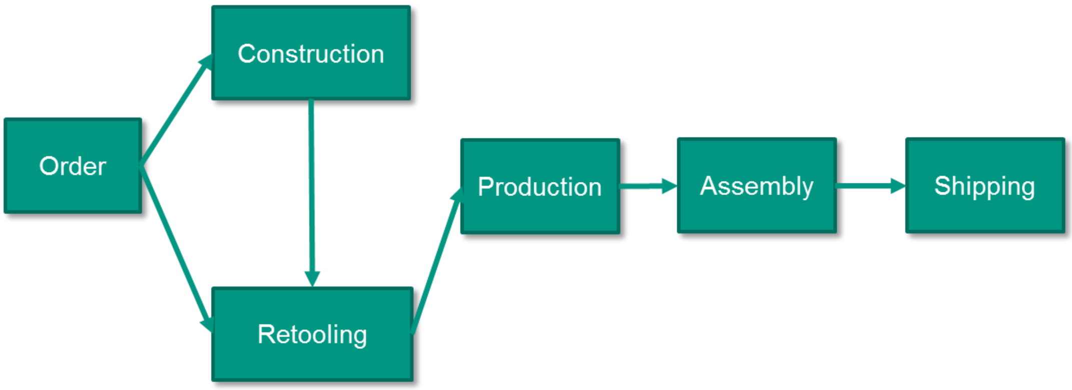


Figure 5.17 Typical operation sequence of the production

Table 5.10 List of available machines and corresponding operation types

Operation step	Machine or workstation	Cost per hour
Construction	M7	40 €
	M8	60 €
Production	M9	54 €
	M10	32 €
	M11	32 €
	M12	45 €
	M1	30 €
	M2	30 €
	M3	30 €
	M4	30 €
	M13	40 €
	M14	45 €
	M15	40 €
Assembly	M17	40 €
	M5	18 €
	M18	40 €
Shipping	M6	20 €

When the production planner starts the prototype, the GUI depicted in Figure 5.6 will appear and all incoming orders are loaded from the knowledge base. Then, he defines the parameters for hyper-heuristics and executes a scheduling algorithm using the GUI depicted in Figure 5.10. Figure 5.18 shows the screenshot of the resulting Gantt chart of a schedule for processing all incoming orders starting from September 12th, 2012. For the schedule, the production planner does not define any shift work and maximum power load constraints. As seen in Figure 5.18 the weights of both constraints are set to zero.

Figure 5.19 depicts the power load plot corresponding to the schedule. It can be seen that there is a peak load of around 53 kW at the beginning of September 14th. The schedule causes total costs of 35,719.66 Euros (energy cost not included) and 4,268.62 Euros of energy costs. All jobs in the schedule are finished in 1,496.83 hours. The total energy consumption is 1,747.2 kWh. The calculation of EnPI yields 8.3.

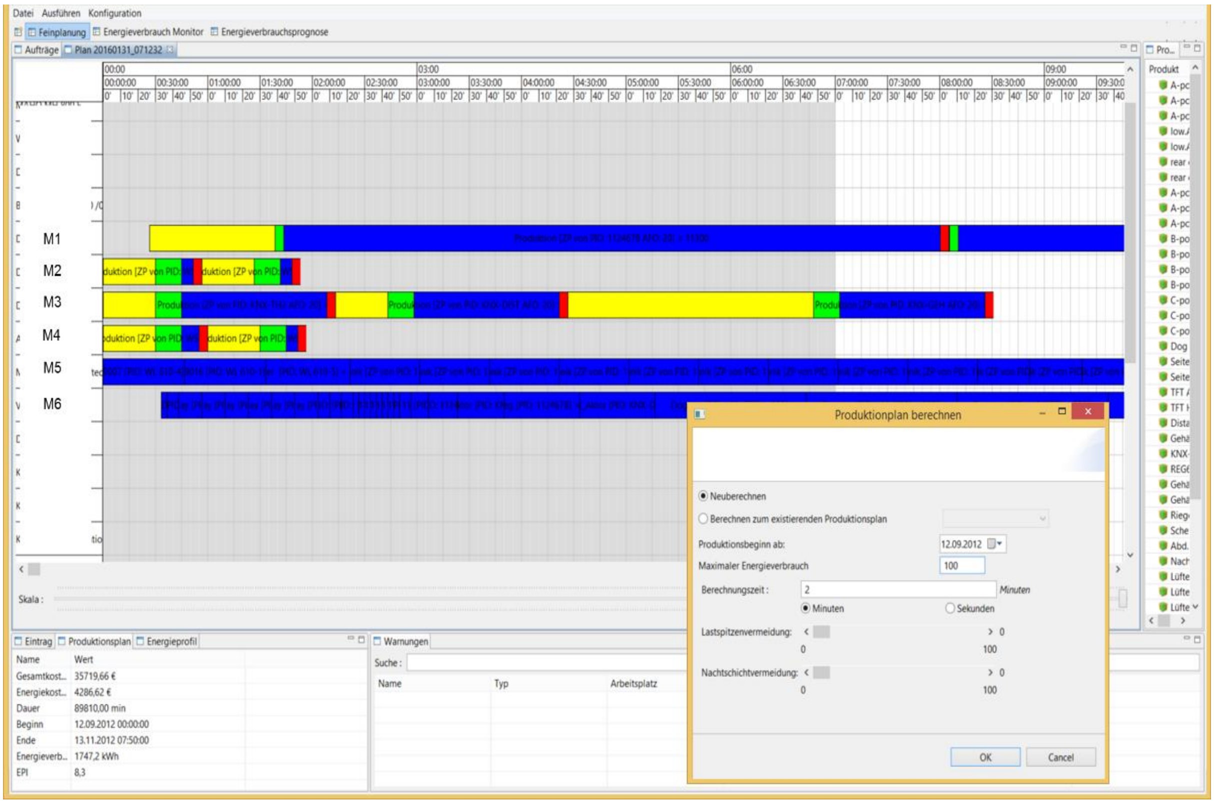


Figure 5.18 Gantt chart illustrating the optimized process plan without peak load avoidance

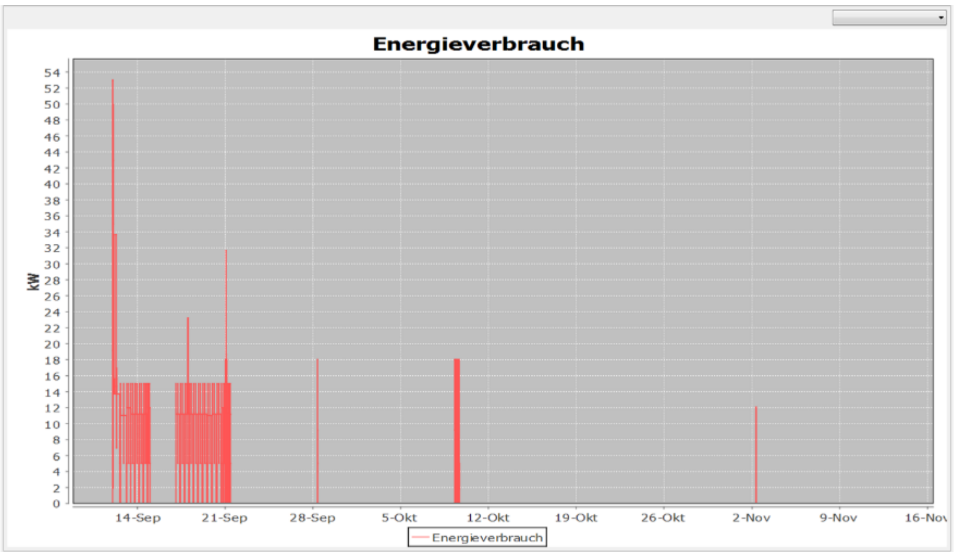


Figure 5.19 Resulting energy profile of a process plan without peak load avoidance

Figure 5.20 displays the resulting schedule if the production planner specifies a maximum power load of 35 kW. As seen in the figure, the weight of peak load avoidance is set to maximum (100). If the jobs are executed following the schedule, the costs are 35,608.43 Euros, which is slightly less than in the previous schedule.

Figure 5.21 depicts the power load profile of the schedule. It is shown that the power load has values lower than or equal to 35 kW. The energy cost is 2,484.24 Euros which is much lower compared to the previous schedule due to the better load distribution. The total energy consumption and duration of the schedule are the same than in the previous case. Because of the lower maximum power load, the EnPI of the production line is higher accordingly, which is 14.3. Compared to the schedules in Figure 5.18, the job on the machine M4 in Figure 5.20 starts around 45 minutes later, whereas in Figure 5.18, three jobs on M2, M3 and M4 are started simultaneously. The job on the machine M1 is also executed around 15 minutes later. It can be concluded that simultaneous start of several machines can cause a peak power load.

Figure 5.22 shows the schedule that avoids night shift-work, while Figure 5.23 illustrates the schedule without avoiding night shift work. The gray area in the Gantt chart represents the night shift interval. The energy cost of the schedule with night shift avoidance is higher than that without night shift avoidance because the dense schedule resulting from night shift avoidance requires parallel operations of several machines. This makes the EnPI of the production line with night shift avoidance lower.

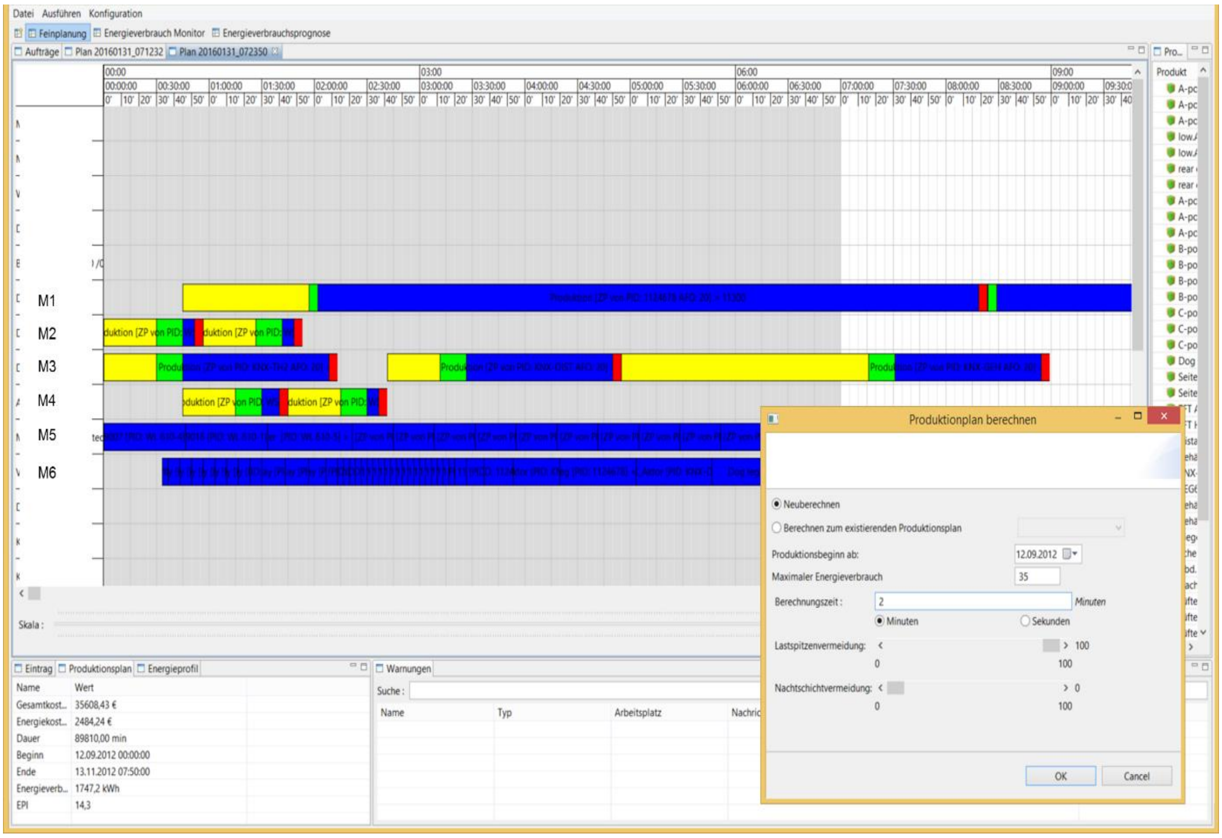


Figure 5.20 Gantt chart illustrating the optimized process plan with peak load avoidance

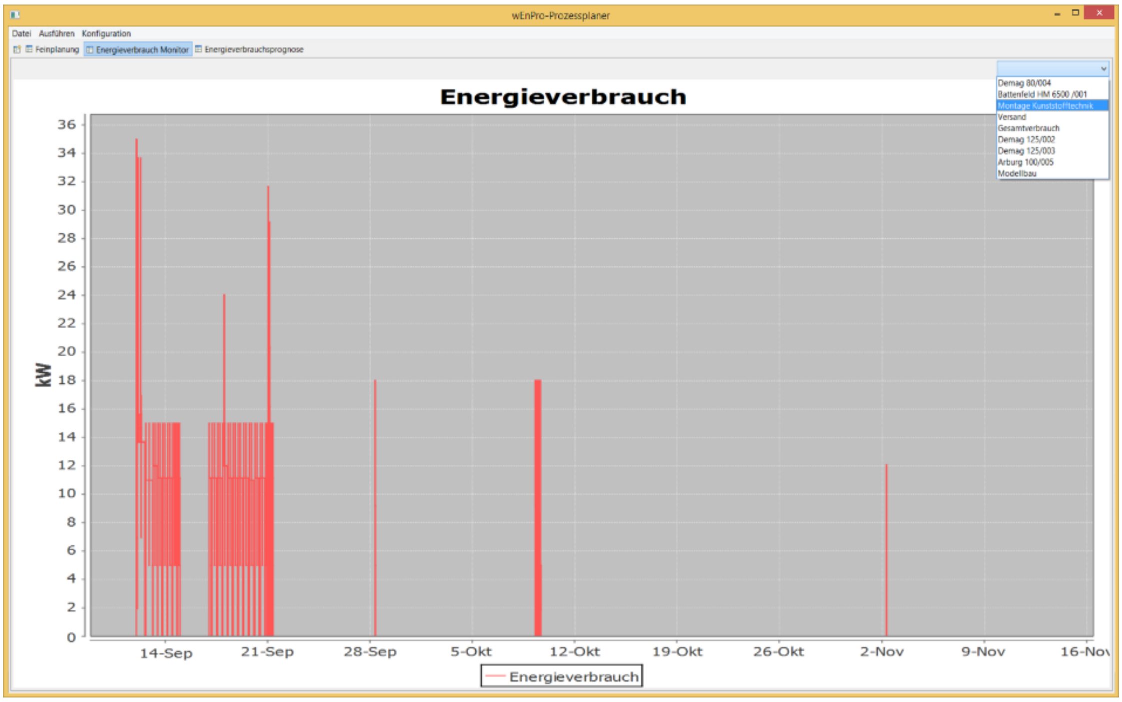


Figure 5.21 Resulting energy profile of a process plan with peak load avoidance

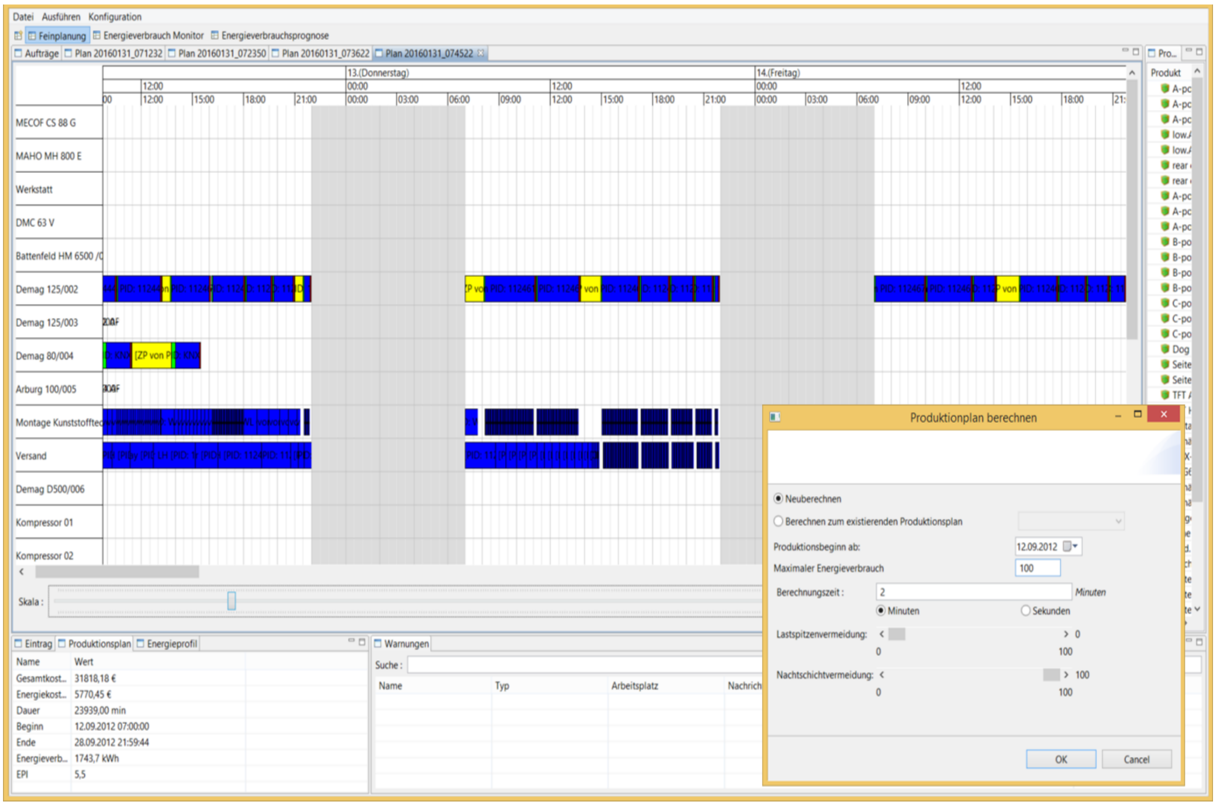


Figure 5.22 Gantt chart showing a schedule avoiding night work-shift

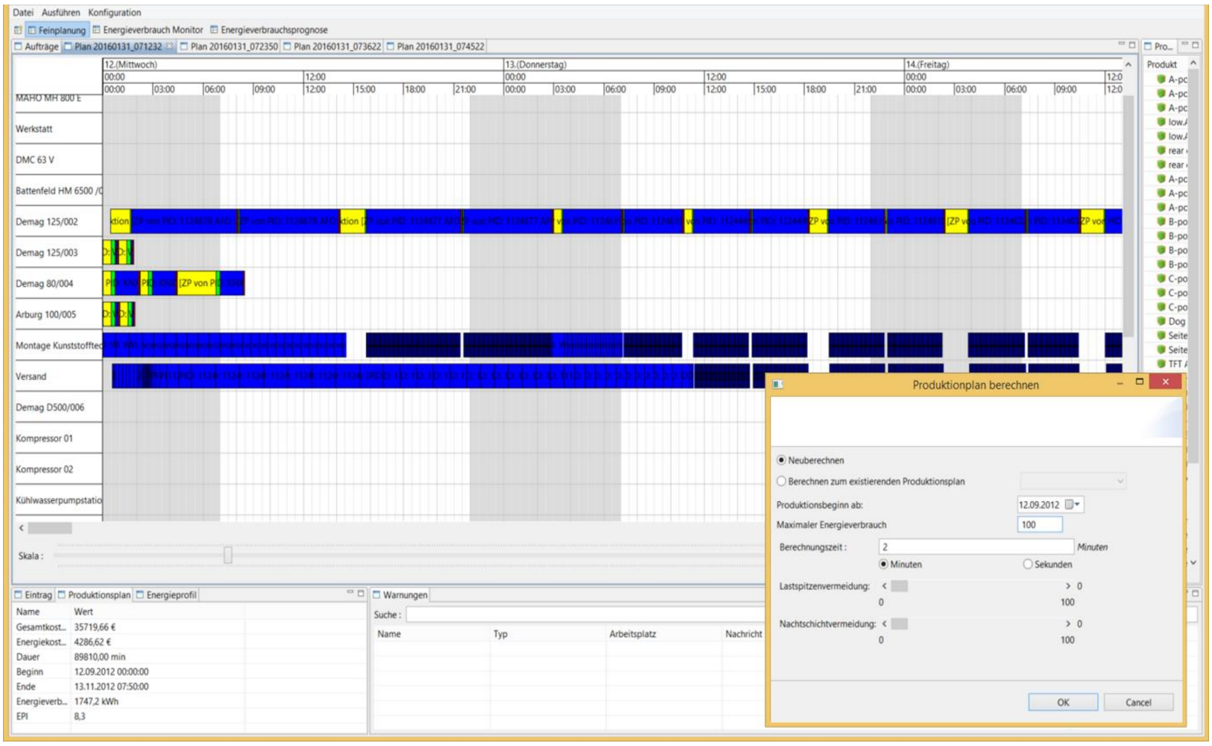


Figure 5.23 Gantt chart showing a schedule without night work-shift avoidance

As learned from the experience, the developed prototype can assist the production planner in creating a production schedule in a flexible way. The production planner is allowed to decide which aspects are to be considered for production scheduling, such as power load distribution and shift work model. The prototype creates the schedule accordingly, but the production planner can still perform changes on the automatically generated schedule. However, the machine and energy costs calculated by the prototype do not represent the real cost of shop floor work, as some workstation data and energy data of ancillary facilities, such as cooling and lighting systems, are lacking.

5.3 Verifications Based on the Requirements

The previous section has shown the positive experiences on the application of the software that implements the developed concepts. However, the developed concepts still need to be verified, whether they correspond to the requirements formulated in Chapter 3. It is essential as the reference to assure the exploitation and benefits of the developed concepts in terms of further development and applications in wide range of discrete manufacturing. This section discusses the evaluation of the developed concepts and implemented prototypes for each category in more detail. It begins with the verification of requirements for the three main concepts, the architecture and functionalities of the prototypes, and followed by the verification of the general requirements as the conclusion of the verifications of the individual elements. It concludes with a table summarizing the evaluation results.

5.3.1 Verification of the Ontology Knowledge Base

The ontology knowledge base plays an essential role in the concept. It serves as the integration point of the solution (see Figure 4.1 and Figure 5.1). The knowledge base that has been developed in the concept phase is proved to be applicable. All necessary objects representing the entities involved in energy management have been modeled in the knowledge base. They have been applied in the industrial use cases according to the requirements defined in section 3.2.

Interoperability and knowledge gap among the entities and stakeholders is one of the main challenges in implementing an energy management system. Ontologies have been proved as an approach for providing a common, understandable representation to address the problem. By using OWL, the shared information or knowledge can be represented in an expressive way. This thesis has modeled the main entities, e.g. building elements, building control systems including sensor and networks, human actors, actor and system behaviors, states, etc., involved in an energy management system using OWL. The ontology serves as shared vocabularies describing those entities. The OWL represented model also comprises complex interrelationships between those entities in forms of hierarchies, property relationships, and rules. As elaborated in section 4.2.1, the developed approach allows the construction of the knowledge with multiple levels of abstractions. A manager can express a rule in a high level of abstraction, whereas a technical person can extend the rule with a more technical model. Nevertheless, both of them have to use the same vocabulary which is defined using OWL. Hence, the approach facilitates a collaborative knowledge base construction involving multiple stakeholders. Figure 5.3 and Figure 5.4 show the screen shots of the ontology displayed with the ontology editing tool Protégé. Furthermore, chapter 4.2 explains the ontology model in more detail. This, therefore, corresponds to the **Requirement R2.1.**

The development of the knowledge base is intended to support qualitative evaluation on the environment of a manufacturing system containing complex interrelationships between entities. It is difficult for a human being to consider all of them. The SWRL rules allow the representation of the complex relationships. The qualitative evaluation is performed by utilizing human knowledge as the evaluation objectives. The SWRL rules allow formalized representation of the human knowledge about the energy efficiency best practices and measures. Furthermore, the rules extracted through KDD process on the historical data are also converted into SWRL. This allows incorporation both human knowledge and automatically extracted knowledge as a basis to perform the qualitative evaluation. Figure 5.5 depicts a screen shot of Protégé that displays the SWRL rules. This enrichment of the ontology with a set of rules satisfies the **Requirement R2.2.**

A systematic approach for developing ontology has to be considered to assure the quality of the ontology. This thesis has analyzed existing well-established ontology development approaches. After analyzing based on a set of criteria, the ontology development approach developed by Noy and McGuinness is chosen (see Table 4.1). The approach offers better adaptation and further development since the widely used ontology editor tool Protégé was implemented based on this approach. Because of this reason, the **Requirement R2.3** is met.

One of the measures to improve the energy efficiency in manufacturing is the prediction of energy consumption of processes, resources e.g. machines, facilities, and incoming orders by means of the product characteristics. Relying on the human knowledge is not adequate to perform this tasks since the extensive knowledge about the historical behaviors and all aspects of the manufacturing environment. Data mining allows the automatic extraction of knowledge from the different data sources. Section 4.2.3 explained the development of KDD technique in this thesis that includes data mining. The data mining generates a set of rules that are converted into SWRL in order to integrate them into the ontology knowledge base. This shows an automatic generation of the knowledge. The integration technique is explained in section 4.3.4. Therefore, it indicates that the developed concepts correspond to the **Requirement R2.4.**

Software modules responsible for KDD and the forecasting have been developed and are described in Section 5.1.3. Figure 5.7 shows the screen shot of GUI that interfaces the user with the KDD module through the forecasting module (see the architecture depicted in Figure 5.1). The GUI enables the user to predict the energy and costs of each machine and incoming order. Moreover, an experience has shown that the prototype can help a make-to-order manufacturing company to predict the energy and costs, as well as to improve the production plan, to process orders. This shows the fulfillment of **Requirement R2.5**.

By consulting the knowledge base, production planners or energy managers can identify the energy efficient states of the objects in the manufacturing system and detect energy inconsistent performance or wasting. The SPARQL queries on knowledge base make this possible. They are implemented as the Energy Efficiency Qualitative Evaluation Workbench GUI (see section 5.1.2). The GUI assists the user in performing the qualitative evaluation. The user can view a list of objects that are in inefficient or efficient states and the reasons why the states occur. This is in conformity with **Requirement R2.6**.

5.3.2 Verification of the Energy Performance Indicator

The EnPI should allow quantification of energy performance so that the quantitative evaluation of processes, products, resources, and organization units in the manufacturing system is possible.

The EnPI calculation considers the production value of the outputs of each process. All of the production outputs having different units are converted into the currency unit, i.e. Euro. This is also easier for the manufacturing companies to supply the developed prototype with currency unit since the production value is usually derived from the costs to produce the product which are ascribed from the observed process. Table 5.2 shows that the outputs of production facilities that carries out the production process are represented in currency values. The corresponding

EnPI results are displayed in Figure 5.13 and Figure 5.14. This is in accordance with the **Requirement R3.1**.

The application in a manufacturing environment containing energy conversion, production, ancillary, and transport facilities has shown the ability of the software prototype to calculate EnPI on different levels of a manufacturing organization i.e. machine or production facility level, production line level, and factory level. It is proven by the screenshots depicted in Figure 5.13 and Figure 5.14. Therefore, the **Requirement R3.2** is considered to be successfully satisfied. The EnPI permits the energy performance benchmarking between different machines that are possibly able to execute the same operation. This helps the production planner to select the most energy efficient machine to execute a certain operation. Moreover, the company is able to simulate whether efforts to invest new machines or to change factory layout could improve the energy efficiencies in the factory by comparing the EnPIs in the simulated configuration to the original one. The changes of EnPIs depicted in Figure 5.13 and Figure 5.14 due to the measures listed in Table 5.3, indicates these benefits. Thus, the **Requirement R3.4** is also met. Furthermore, the EnPI differences show the evidence that not only energy consumption contributes to the EnPIs, but also the peak load influenced by the maximum input power and the contribution of energy conversion facilities. The contributions are represented in a monetary unit. Thus, it does not depend on energy forms. Additionally, the EnPI deviation shown in

Figure 5.20 compared to Figure 5.18 strengthens the evidence that the fair distribution of power load increases the EnPI. It means the EnPI complies the **Requirement R3.3.**

As the conclusion, through the use case described in section 5.2.1, the cause-effect relation has been proven as the basis of the EnPI. This is shown by the increase in the EnPI value if the cause (i.e. energy efforts) decreases. This is in accordance with the **Requirement R3.5**.

5.3.3 Verification of the Energy Optimized Production Planning and Scheduling

The production planning and scheduling developed in this thesis addresses the problem in a real manufacturing environment. However, it can be still formalized using standard representation developed by Blazewicz et al. resulting in an instance of his model. Table 4.14 and Table 4.15 describe the real problem instance considered in this thesis complying with Blazewicz's approach. This proves the fulfillment of **Requirement R4.1**.

In order to represent the real manufacturing environment, the standard scheduling model has to be extended, for example with shift work model, retooling on machines before executing jobs, multiple inputs multiple outputs precedence relations, etc. Section 4.5.1 gives more details the concept that addresses this aspect. The Gantt chart depicted Figure 5.10 illustrates that a retooling is performed before a machine starts. It is indicated by the yellow part of the bar. Moreover, Figure 5.22 and Figure 5.23 give the evidence that the shift-work model influences the resulting schedule. This corresponds to the **Requirement R4.2.**

The developed scheduling approach considers both energy efficiency and costs as the objectives and constraints. By using the developed prototype, the user can define the maximum power load allowed on the shop floor and also indirectly the personal costs determined by the shift-work model (see Figure 5.10, Figure 5.18, and

Figure 5.20). The hyper-heuristic performs a search by achieving objectives composed of both energy and costs as formulated in (4.54). This satisfies the **Requirement R4.3**.

5.3.4 Verification of Software Architecture and Implementation

The minimal adaptation and deployment efforts in the real manufacturing environments are essential to ensure the usability of the developed prototype. The data collection interface has been developed to achieve it. It allows the communication and data transfer between the prototype and the existing ICT systems. Figure 5.1 shows the role of data collection interface in the prototype architecture. This thesis has implemented both export-import interface to databases and a client to communicate with web services. This allows flexibility of the prototype in the communication with any systems and data sources located both on the local network and on the internet. This flexible interface corresponds to the **Requirement R4.1**.

The GUIs of the prototype are implemented using advanced desktop based GUI technologies, such as RCP, Jaret, and GEF. Those technologies provide advanced visualization and user interactions, such as drag and drop, resizing or zooming, and graph editing. The screen shots in section 5.1.3 give the impressions, how a user interacts with the GUI prototypes and takes benefits from the functionalities. This regards to the fulfillment of the **Requirement 4.2**.

As depicted in Figure 5.1 the architecture of the solution is designed in a modular way where the ontology representing manufacturing energy management knowledge base serve as the integration point. The modular solution allows a manufacturing company selects only modules that they need. The experiences presented in section 5.2 indicate that a company does not have to deploy all of the modules. It leads to the fewer efforts in taking the benefits of the system functionalities. Furthermore, the flexible, adaptable, and extendable nature of the ontology knowledge model permits the company to react more flexible in dealing with the rapid growth of ICT technologies in manufacturing. This, therefore, satisfies the **Requirement 4.3.**

5.3.5 Verification of General Requirements

The general requirements are the requirements which are directly derived from the research questions stated in Section 1.2. Those requirements are then broken down into lower-level requirements for the specific concepts as defined in Section 3.2, Section 3.3, and Section 3.4. Consequently, the verification of the general requirements is a conclusion of the verifications of specific concepts discussed in Section 5.3.2, Section 5.3.3, and Section 5.3.4.

The interoperability and knowledge sharing problems are addressed through ontology represented knowledge base aligned with existing models. The knowledge base covers both informational and organizational infrastructures that build an energy management system. This corresponds to **Requirement 1.1**.

The best practices in energy management, which are human knowledge, are formalized and stored as hierarchies and rules in the ontology knowledge base to allow easy access from and reuse by different stakeholders and systems. It makes the extensive evaluation on raw data is not necessary anymore. The deductive reasoning on the rules and knowledge model are performed to get the qualitative evaluation results. Therefore, it satisfies the **Requirement 1.2.**

The effectiveness of energy efficiency evaluation strongly depends on the understanding of the company or case-specific data because it is not possible only to rely on expert knowledge. The KDD approach allows learning and extracting knowledge from the company or use-case-specific data sets. The use case example in a make-to-order manufacturing presented in Section 5.2.1 gives the proof of the benefits of the approach in a real manufacturing environment as required through **Requirement 1.3**.

The experience in applying the prototype in the use case presented in Section 5.2.1 has shown the benefits of the developed EnPI for supporting the quantitative evaluation of energy performance and benchmarking among different processes, facilities, systems, and organization units in a manufacturing company. The validation results in real manufacturing are promising. Therefore, the **Requirement 1.4** is met.

The developed energy optimized production planning and scheduling has been proven to give benefits, especially for energy intensive serial manufacturing (see Section 5.2.3) since it can help the production planner to accomplish the difficult task to schedule the order processing by considering both economical and energy efficiency aspects. This indicates the accordance of the developed concept and implanted prototypes with **Requirement 1.5.**

5.3.6 Conclusions

The verification results of the developed concepts and prototypes against the defined requirements are summarized in Table 5.11.

Table 5.11 Summary of verification results in accordance with requirements

Requirement No.	Requirement	Satisfied?	Remark
General Requirements			
R1.1	Interoperability and knowledge sharing	Yes	Representation of the knowledge base with OWL
R1.2	Intelligent analysis and qualitative evaluation	Yes	Deductive reasoning using rules to persist human knowledge; qualitative evaluation using SPARQL implemented in Energy Efficiency Qualitative Evaluation Workbench
R1.3	Capability to learn the knowledge	Yes	KDD technique to extract knowledge from data
R1.4	Quantitative evaluation using performance indicator	Yes	Straightforward EnPI for benchmarking of different processes, facilities, systems, and organization levels/units; application experience in complex manufacturing system
R1.5	Applicability in production planning and control	Yes	Hyper-heuristic energy optimized production scheduling; application experience in serial manufacturing
Requirements for knowledge base to support qualitative evaluation			
R2.1	Expressive knowledge base representation	Yes	OWL representation of different entities in class hierarchies, interrelationships of properties, rules
R2.2	Integration of rules	Yes	Introduction of SWRL rules for both human and automatically created rules
R2.3	Systematic ontology creation	Yes	Ontology development through Noy and McGuinness approach; creating ontology with widely used tool Protégé
R2.4	Data mining knowledge generation	Yes	Development of KDD technique and implementation of KDD software module
R2.5	Energy and costs prediction	Yes	Implementation of the forecast module; application experience in a make-to-order company

R2.6	Energy efficiency qualitative evaluation based on best practices	Yes	SPARQL query to retrieve the object states; implementation of Qualitative Energy Efficiency Evaluation Workbench
Requirements for EnPI to support quantitative evaluation			
R3.1	Product type independent	Yes	Conversion of all values to the same unit
R3.2	Measurement of different levels of production organization	Yes	EnPI for machine, production line and factory level; application experience of prototype
R3.3	Consideration of energy costs and balance by including internal energy generation/conversion; independent from energy forms	Yes	Consideration of all energy efforts including peak load, certifications; application experience in substituting machines and energy conversion facilities and also changing layouts
R3.4	Transparency for machine selection	Yes	Straightforward and direct comparable EnPI value
R3.5	Cause-effect or input-output representation of EnPI	Yes	The straight-forward representation; application experience
Requirements for energy optimized production planning and scheduling			
R4.1	Problem description that compatible with standards	Yes	Blazewicz's problem description
R4.2	Consider both energy and costs	Yes	Configurable and composition of objectives and constraints covering both energy and costs
R4.3	Consider parameters in real problems	Yes	Retooling, shift-work model, multiple inputs multiple outputs precedence relations; representation in GUI
Requirements for software architecture and implementation			
R5.1	Flexible communication and data transfer interface	Yes	Implementation of database and web service interfaces

R5.2	User-friendly, intuitive and flexible interaction	Yes	Advanced GUI technologies: RCP, GEF, Jaret, etc.
R5.3	Minimal adaptation, changes and extension in deployment	Yes	Modular architecture, ontology integration point

5.4 Further Exploitation and Applications of the Developed Concepts

This section describes some further exploitation and application possibilities of each developed concepts, i.e. manufacturing energy management ontological knowledge base, KDD module including data mining algorithms, the EnPI, and hyper-heuristic approach for solving optimization problems.

5.4.1 Further Applications of Knowledge Base and Data Mining

This thesis develops ontology knowledge base for manufacturing energy management domain. The ontology model covers the objects involved in manufacturing energy management activities. However, it does not keep out the possibilities of the association of those objects to other domains. Additionally, ontology offers flexible extensions of the model. Thus, further adaptation and applications of the knowledge base are imaginable. Similarly, data mining approach, which extracts the knowledge from data, work independently from the type and contents of the data. Therefore, the applications of data mining are possible to solve diverse problems. The following sections elaborate some possible further adaptations and applications of the knowledge base and data mining approach.

Building Energy Management

Building energy management requires knowledge about building envelopes and physical elements and sensor networks. Some of the models in the ontology, such as building element, building control including sensor networks for energy metering, actors, and state representations, can be directly applied to build a knowledge model for building energy management system. However, rules have to be replaced with the ones representing building energy management best practices. The data mining algorithms can be directly applied to predict energy consumption of building elements based on the parameters existed in building environments, e.g. occupant behaviors, dry-bulb temperature, building envelope, etc., but still, the input data and pre-processing steps have to be adapted.

Energy Efficient Smart Cities

A smart city necessitates the integration of different technologies to manage assets including building and manufacturing facilities. The developed ontology can be slightly extended to meet the requirements for smart city knowledge model by adding the city or district perspective.

Nevertheless, it only covers smart energy and smart building aspects. Another possible approach is to develop a linked data to interlink the developed ontology with other existing models applied for smart cities [McEt-16].

Big Data in Industry 4.0

The developed KDD approach has built several foundations for big data analytics that is one of the key technologies in industry 4.0 following the 5C architecture [LeEt-15]. The KDD can play a significant role on data-to-information conversion layer in providing the multi-dimensional data correlation, smart analytic, and performance prediction. It serves as a basis for the decision and diagnostics elements incorporated in cognition layer. Furthermore, the clustering tasks in cyber level can be taken over by the developed data mining algorithms. However, this thesis validated the data mining concept in a moderate IT infrastructure. Therefore, the scalability of the developed algorithms has to be tested in big data infrastructure.

Process Optimization across Product Life Cycle

The ontology containing product-process-resource (PPR) model can be applied as a shared knowledge model and vocabulary across the product life cycle. The PPR model plays the central role in most of the processes involved in different product life cycle phases. However, the model needs to incorporate concepts representing customers and their product usage context in order to allow the evaluation of the product from pre-order and usage phase perspectives [ScWR-11] [WiEt-11].

5.4.2 Further Applications of EnPIs

In this work, the focus is on the production planning and control. Further applications of the proposed EnPI are also conceivable.

Optimization of Operation Modes

If the EnPI is calculated and recorded in real time, the efficiency of the production facilities and processes can be observed in real time. In this case, it is worth to look at the respective operating states in more details where the energy performance is high or low. If necessary, it can be derived, which modes of operation (speed, power expended, stop-and-go operation) are especially energy-intensive, thus in the future, they should be avoided or paid more attention in order to save more energy.

Identification of Defects and Wastage

Another possible application of calculated EnPI is the detection of defects or wastage. If the EnPI is getting worst during unchanged operating conditions, it could be an indication that system is torn apart or for example lubrication must be greased. If certain patterns are identified, an alarm function can be programmed to detect wastage or defects, thus warning messages are generated.

Assessment of Machine Tool Level

A machine can be equipped with different tools to execute different operations. Each operation creates a product having a certain value. The EnPI can be applied to assess the energy efficiency of each machine tool, but it requires the registration of each machine tooling or retooling process. Principally, the EnPI calculation is similar to the machine level, except the additional logging of (re) tooling process.

Investment planning

The EnPI can be used to compare alternatives, in case that the company plans to invest a new production facility. The EnPI allows the evaluation of new production facility individually as well as the evaluation how the production facility works together in the existing systems. This also allows interactions and interdependencies between plants to be considered in the investment decision. Another possible benefit of the developed EnPI is the consideration which existing facility should be replaced with the more energy efficient new facility. If the EnPI is used in process planning, usually the most inefficient facilities are used least frequently. With appropriate utilization, this kind of facilities is nevertheless still regularly used. Based on the EnPI it can be seen, which facilities are highly energy intensive, and investment for replacement should be considered.

Logistics

The EnPI considers the transport facilities in the production already. It is conceivable, therefore, that the metric could be used in the routing and scheduling for logistics. Using the metric could be determined, which means of transport for a given transport task has the lowest energy cost. It would allow for a task to be related with the preceding and following tasks, so empty trips could be taken into account.

5.4.3 Further Applications of Hyper-Heuristic Optimization Approach

Through the developed hyper-heuristic approach, an optimization problem can be solved in a flexible way. It allows a configuration of optimization objectives and constraints and the introduction of user-defined or well-known heuristics, for example, the EDD and DTF strategies in production scheduling. This section elaborates further possible applications of the hyper-heuristic approach developed in this thesis.

Logistics and Transport Optimization

There are classic optimization problems in logistics and transport domain, for example traveling salesman problem, vehicle routing problem, and their variations. By applying the hyper-heuristic approach, a generic model for the problem does not have to be developed. The user-customized or existing standard strategies, such as the path through highway first, or taking the longest street first can be employed. Furthermore, different objectives or their combinations can be defined, for example, the cheapest travel costs and the shortest travel time.

Smart Grids

One of typical optimization problems in a smart grid environment is the scheduling for demand response management. Demand response management aims to shift users' demand to off-peak hours in order to have a fair load distribution. The problem is very similar to the production scheduling problem attempted to be solved in this thesis. The smart grid specific strategies can be defined through the hyper-heuristic framework, for example, avoiding some time range, setting higher priorities to some areas or users, etc.

6 Conclusions and Outlook

"The science of today is the technology of tomorrow."
~Edward Teller (1908-2003)

Global climate change and rapid growth of industrialization have led to a significant increase in energy demand that results in constantly increasing prices of electricity, gas, and oil. Meanwhile, the changes of social, technical, and economic conditions in the market have forced manufacturers to focus on the requirements for various and complex products. Hence, energy efficiency in accordance with the economization of production costs is an important competitive factor in the energy-intensive industry. A common approach to improve the energy efficiency is the corporate energy management system which defines all processes and measures to ensure minimal energy consumption based on given requirements, involving organization, information structure, and tools. ISO 50001 describes the requirements to be met by energy management systems in industrial companies. However, most of the manufacturing companies have problems in implementing them due to the distributed and unstructured energy-related information within the company. This information is often difficult to access by all stakeholders in the company, which is why it is not used effectively to support timely operational and strategic decision-making. Furthermore, there is a knowledge gap among staff members who sometimes are unaware of best practices to reach energy efficiency. There are also interoperability problems among the applied technologies and stakeholders due to the lack of common vocabularies and shared knowledge structures. Moreover, the existing energy-related KPIs are mostly inaccurate. They do not provide any exact statements about energy performances of processes, resources, and organization units within the company and are often unmanageable and difficult to use.

This thesis overcomes these problems by developing an approach which utilizes information and communication technologies and supports a holistic evaluation of energy efficiency in discrete manufacturing. The evaluation is performed both qualitatively and quantitatively to achieve accurate energy efficiency measures. This thesis focuses mainly on improving "Check" activities in the PDCA cycle of an ISO 50001 energy management system. Qualitative evaluation concentrates on the criteria that can be observed but not measured in figures. Thereby, the correlations between object states, such as energy-wasting situations or energy-efficient best practices, are assessed. Those correlations are modeled in a knowledge base. Quantitative evaluation, by contrast, is performed by assessing energy efficiency using measurements of the energy performance indicator (EnPI). In this way, it is possible to determine the degree of energy efficiency of different processes, resources, and organizational units within a company.

6.1 Thesis Contributions

This section summarizes the contributions of this thesis from the scientific, technological, and industrial points of view. The contributions address the problems and answer the research questions formulated in Section 1.1 and Section 1.2. By analyzing the fundamental concepts and state of the art, the relevance of this research to the real problem, the position of this research compared to other studies as well as the novelty of this research, are illustrated. The analysis of state of the art shows that no appropriate manufacturing energy management approach exists which is based on integrated ICT and can answer the research questions. Compared to other related work, the innovation of this thesis consists in the *integrated method applying information and communication technologies (ICT)* to support both *quantitative and qualitative evaluation of energy efficiency in manufacturing by utilizing an expressive and adaptive ontology knowledge base that is capable of learning as well as a sector-independent, straightforward, and directly usable energy performance indicator (EnPI)* for evaluating and benchmarking different processes and organization levels within a company. Furthermore, this thesis utilizes the ontology knowledge base and EnPI in production planning and control or, more specifically, in production scheduling by means of a *hyper-heuristics-based energy-optimized, flexible, and user-configurable scheduling* approach. These approaches help the companies to improve the check phase in the implementation of energy management system.

Review of the research questions and state of the art analysis yield a list of requirements that are fulfilled by the developed concepts. The requirements cover the knowledge base for qualitative evaluation, EnPI, production schedule, and prototype implementation. The developed concepts also produce scientific findings which are summarized in Section 6.1.1. The concept implementation and validation results reflect the technical and industrial contributions of this thesis which are summarized in Section 6.1.2 and Section 6.1.3.

6.1.1 Scientific Contributions

This thesis contributes to developing the approach to *qualitative evaluation by creating the manufacturing energy management knowledge base* which is represented using an expressive ontology representation language, called OWL. The proposed knowledge base serves as a shared vocabulary and knowledge structure that is accessible to all staff members and systems in the company. The knowledge base contains explicit formulated best practices and knowledge about energy-efficient and energy-wasting activities. The analysis of state of the art reveals that researchers and industries have been applying ontologies (OWL) to represent knowledge in various domains with the aim of eliminating common understanding problems among people and even software, but ontology development for holistic manufacturing energy management has hardly been studied so far. Existing research approaches use OWL as knowledge representation of either building energy management or the manufacturing domain. This thesis *combines both energy management and manufacturing domains and yields an OWL knowledge base for manufacturing energy management.*

Unlike other approaches, the knowledge base in this thesis is equipped with SWRL rules for intelligent reasoning through a rule engine. These rules that are stored in the knowledge base make the knowledge about best practices available throughout the company so that every staff member or system can access them thus solving the interoperability problems and bridging the knowledge gap. The approach developed in this thesis allows *a collaborative construction of the rules and knowledge stored in the knowledge base*. It enables people having different knowledge backgrounds and levels of abstraction to construct the knowledge base. The object state correlations in the rules are used as references to perform the qualitative evaluation. The rule engine is capable of giving answers about current states of different objects in the manufacturing system based on the conditional correlations represented by the rules so that the *qualitative evaluation of energy efficiency can be carried out automatically*.

Knowledge base construction and instantiation are very complex tasks involving experts and manufacturing personnel. An expert often has to analyze heterogeneous and large amounts of data before he constructs and instantiates the knowledge base. In order to support these tasks, this thesis contributes to developing methods to generate the knowledge base by *learning from historical manufacturing and energy data to extract the knowledge*. The approach utilizes the knowledge discovery in database (KDD) technique which includes a set of data mining algorithms. The KDD is based on a decision tree combined with linear equations, as this is an easy-to-understand model representation and capable of handling different types of data. The KDD allows a company to estimate the duration and power consumption for processing the ordered products in every production step more accurately.

Furthermore, this thesis contributes to the development of an energy efficiency measurement parameter called *energy performance indicator (EnPI) for the quantitative evaluation and benchmarking of the degrees of energy efficiency in different organizational parts, processes, and resources within a company***.** According to the state of the art analysis, there is a lack of approaches that yield the energy efficiency of an object directly and sector-independently. This thesis provides an EnPI calculation by considering energy flow and balance models. Energy flow and balance modeling takes into account positive and negative contributions, not only of production facilities, but also of other energy consuming and generating facilities, i.e. energy conversion, transportation, and ancillary facilities. The *EnPI representation also is straightforward and directly usable for benchmarking on different objects, so that the user does not have to use different figures* to measure the energy efficiency of an object. The figure does not depend on the process, product, and sector.

This thesis also addresses energy performance optimization of a concrete field in production planning and control, i.e. production scheduling. In the context of energy efficiency, the scheduling approach has to consider both economic and energetic optimization. The state of the art analysis has shown the lack of scheduling approaches that correspond to real problems. The existing exact scheduling approaches cannot give solutions for large problems within reasonable periods of time. Meta-heuristics, by contrast, work on general models that do not correspond to reality. For this reason, this thesis develops a *hyper-heuristic scheduling approach yielding a flexible, but still real problem-relevant model.* The hyper-heuristic approach allows for the *incorporation of different user-defined heuristics strategies, optimization objectives, and constraints.* In this thesis, the optimization objectives and constraints consider both economic aspects, e.g. lowest costs and avoiding certain shift work, and energy efficiency, e.g. lowest energy consumption, and avoiding peak power load. This scheduling approach uses a model from the entities in the knowledge base. The resulting schedule can be assessed using the EnPI and allows for user modifications so that the user is enabled to benchmark across different schedule alternatives.

6.1.2 Technical Contributions

In order to ensure applicability of the scientific findings described in Section 6.1.1, this thesis develops *software prototypes and a system integration architecture.* The architecture contains the knowledge base as the middleware or integration point of different application modules, e.g. production scheduling, EnPI calculation, forecast module, KDD module, etc. Through the data collection interface, the knowledge base also serves as an integration model of different schemas and data from heterogeneous data sources in the area of ICT systems for manufacturing.

The software prototypes are implemented in a modular manner using popular and low-cost software technologies, such as Java, Eclipse (RCP, GEF), and MySQL, which facilitates software development and improves extension and applicability in manufacturing companies having different ICT environments, especially SMEs. Moreover, the flexible data transfer interface for communication with databases and web services allows the prototypes to be integrated into the existing ICT environment. The development of the software prototypes in this thesis validates the concepts. The prototypes are validated and demonstrated in several manufacturing companies in different sectors. Thus, they reach the technology readiness level (TRL) 6-7 according to the European Commission's Horizon 2020 Work Programme [EuCo-15].

6.1.3 Contribution to Industries

This thesis has *validated* the developed concept in a *real manufacturing environment and at companies. These include a make-to-order manufacturing company, a serial manufacturing company, and a production system involving energy conversion, transport, ancillary, and production facilities.* The validation has produced promising results. The concepts implemented as software prototypes can help production planners and operators perform the quantitative and qualitative evaluation of products, processes, and organization states in their manufacturing systems. The utilization of knowledge stored the knowledge base that was extracted from historical production and metering data by means of the KDD technique, helps the company estimate the costs and energy consumption associated with the processing of the order by looking at the features of the ordered products. They can decide how to process the orders such that energy consumption and costs are lowest.

The application of the developed hyper-heuristic scheduling approach in the serial manufacturing company has shown that it can assist the production planner in creating a production schedule in a flexible way. The planner can decide which aspects are important for production scheduling. This is an improvement compared to classical scheduling approaches since user-defined energy efficiency-related strategies, such as power peak load avoidance and time interval avoidance, can be incorporated.

This thesis has validated the developed concepts in a manufacturing company owning energy conversion, transport, ancillary, and production facilities. The EnPI measures the energy performance of workstation, production line, and factory level. The knowledge base enables qualitative evaluation of manufacturing objects, such as facilities, jobs, and machines. Validation has shown that the EnPIs and knowledge base help the company evaluate and benchmark the entities in their companies. The EnPIs and ontology knowledge base are used as the basis for decision-making in optimizing the factory layout, processes, and machines in energy–efficient manner.

6.2 Outlook

The outlook of this thesis is given from different perspectives. It begins with further research work needed to improve the functionalities of the currently developed concepts. Then, potential technical improvements of the prototypes to increase industry-wide applicability and further exploitation in the industry are highlighted.

6.2.1 Methodological and Research Outlook

The results of this thesis can be used and further developed in a number of ways, as is presented in Section 5.4. Even though the knowledge model in the manufacturing energy management ontology is extendable and adaptable, there are still challenges associated with populating the ontology. Building elements of the factory can be populated using the OntoCAD tool, but the population of products and machines mostly has to be performed manually. Hence, semantic interpretation of the geometry of products and machines and the mapping of the result into the ontology classes or properties would be an interesting subject of further research. It is also interesting to align the ontology to existing industrial standards and schemas, such as IFC4 and AutomationML, in order to increase the applicability of the ontology in industry. Automatic alignment of the manufacturing energy management ontology to those schemas would also be a potential subject of research. The KDD approach to estimating energy consumption and costs based on product characteristics results in the prediction of a single value (i.e. duration or average power load) by using decision trees and linear equations. The accuracy of the prediction can be improved by the incorporation of non-linear functions in the decision tree and prediction of a series of power values. These are challenges that should be addressed by further research.

The EnPI calculation approach assumes constant energy prices for energy balance modeling. Electrical power must be generated on demand at any time. This means that the price strongly depends on the relationship between supply and demand at that time. Following the integration of renewable energy and development of smart grids, the EnPI will have to take the variable and more granular tariffs into account.

The hyper-heuristic scheduling approach developed in this thesis proved to be suitable for industrial applications due to its flexible and configurable features. For more complex and large problems, integration of a learning mechanism would help accelerate the search for the solution. Moreover, the incorporation of offline learning allows the hyper-heuristics to learn the strategies from the company-specific environment. As a result, the schedule would be more accurate and correspond better to the real problems.

The developed SERENE ontology, EnPI, and hyper-heuristic scheduling can be integrated into digital factory systems which offer more advanced geometrical representations and visualizations. The thesis results will extend the function of the digital factory systems in terms of energy-efficiency-oriented production planning. They will also make the digital factory system more intelligent due to the automatic and advanced decision support capabilities.

6.2.2 Technical Development Outlook

Since the developed software prototypes reach TRL 6-7, there is much potential to increase further their technology readiness. Although the prototypes are implemented using platform-independent Java technology, user access to the systems is still limited. This is due to the desktop-based GUI that requires the user to install the GUI first before he uses the applications. Therefore, a web-based GUI would be a potential improvement. By using web applications, the user can access the system from everywhere using any device, as long as a web browser is installed. Usage of qualitatively proved and widely spread proprietary technologies could improve the readiness of the system. Moreover, modern software architectures, such as the service-oriented architecture (SOA), for enhanced flexibility and adaptability of the system might be considered an improvement, although implementation always entails considerable effort and requires company resources.

6.2.3 Industrial Application Outlook

Each concept and the overall solution presented in this thesis have a great potential of being applied in industrial contexts, from the micro level, such as factories, buildings, and their elements, to the meso level involving the symbiosis among industries, such as logistics and smart grids, to the macro level of e.g. smart cities. Furthermore, the developed concepts could be extended to meet the requirements of process industries. Some concepts can be applied directly, such as EnPI, while others, such as hyper-heuristic production scheduling, have to be adapted since the products and processes in process and chemical industries cannot be as clearly differentiated as in discrete manufacturing. However, in order to obtain an idea as to whether the developed concepts can be transferred to other companies, a replicability study covering, at least, an iteration of a PDCA cycle has to be performed. By means of the replicability study, an exact energy efficiency improvement degree can be obtained for each company.

7 Bibliography

[AdMF-06] Addicott, R.; McGivern, G.; Ferlie, E.: Networks, Organizational Learning and Knowledge Management: NHS Cancer Networks, in: Public Money and Management, vol. 26, no.2, pp. 87–94, 2006

[Aggt-90] Aggteleky, B.: Fabrikplanung – Werksentwicklung und Betriebsrationalisierung. Band 2, 2. Auflage, Carl Hanser Verlag, München, 1990

[AlSe-12] Altuntas, S.; Selim, H.: Facility layout using weighted association rule-based data mining algorithms, in: Expert System Application, vol. 39, no. 1, pp. 3-13, 2012

[AlVy-10] Alsafi, Y.; Vyatkin, V.: Ontology-based reconfiguration agent for intelligent mechatronic systems in flexible manufacturing, In: Journal Robot. Computer Integrated Manufacturing, vol. 26, no. 4, pp. 381-391, 2010

[amim-08] N.N.: AmI-MoSES Final Public Report, 2008, http://www.ami-moses.eu/fileadmin/templates/amimoses/files/AmI-MoSES_D8.5_Final%20Public%20Report_v1.0.pdf, access date 12/08/2012

[BaFG-91] Bauer, G.; Friedrich L.; Goos, G.: Informatik 1-4, Springer-Verlag Heidelberg, 1991

[BaHS-08] Baader, F.; Horrocks, I.; Sattler, U.: Description Logics, In: Frank van Harmelen, Vladimir Lifschitz, and Bruce Porter, editors, Handbook of Knowledge Representation, pp. 135-180. Elsevier, 2008

[BDEW-04] Bundesverband der Energie- und Wasserwirtschaft e.V.: *Warmluftheizung*. Berlin, 2004

[BeCM-04] Bellinger, G.; Castro, D.; Mills, A.: Data, Information, Knowledge, and Wisdom, In: System Thinking, 2004, online http://www.systems-thinking.org/dikw/dikw.htm, access date 05/01/2013

[BeFS-15] Bergmann, S.; Feldkamp, N.; Strassburger, S.: Approximation of dispatching rules for manufacturing simulation using data mining methods, in: Proceeding Winter Simulation Conference (WSC) 2015, pp. 2329-2340, 2015

[BEKG-13] Bremer-Energie-Konsens GmbH: Internet-Energie-Check, http://www.internet-energie-check.de, access date 03/04/2013

[BiKl-93] Birmingham, W.; Klinker, G. : Knowledge acquisition tools with explicit problem-solving models, In: The Knowledge Engineering Review, vol. 8, no.1, pp. 5-25, 1993

[BlEt-07] J. Blazewicz, K. Ecker, E. Pesch, G. Schmidt, and J. Weglarz. Handbook on Scheduling: From Theory to Applications. Springer Berlin Heidelberg, 2007

[BMU-10] Bundesministerium für Umwelt, Naturschutz und Reaktorsicherheit: DIN EN 16001: Energiemanagementsysteme in der Praxis, 2010

[BMU-11] Umwelthauptstadt, Deutschland bei Klimaschutz-Zielen 2020 auf Kurs 2011, http://www.umwelthauptstadt.de/nachhaltigkeit/deutschland-bei-klimaschutz-zielen-2020-auf-kurs, access date 14/04/2012

[BMU-12] Bundesministerium für Umwelt, Energiemanagementsysteme in der Praxis ISO 50001 - Leitfaden für Unternehmen und Organisation, 2012

[BoLe-04] Borgo, S.; Leitão, P.: The role of foundational ontologies in manufacturing domain applications. In: On the Move to Meaningful Internet Systems 2004: CoopIS, DOA, and ODBASE, pp. 670–688, 2004

[Boyd-05] Boyd, G.A.: Development of a Performance-Based Industrial Energy Efficiency Indicator for Automobile Assembly Plants / Argonne National Laboratory, Decision and Information Sciences Division, 2005

[Boyd-06] Boyd, A.: Development of a Performance-Based Industrial Energy Efficiency Indicator for Cement Manufacturing Plants / Argonne Nationa Laboratory, Decision and Information Sciences Division, 2006

[Boyd-09] Boyd, A.: Development of a Performance-based Industrial Energy Efficiency Indicator for Glass Manufacturing Plants / Duke University, Department of Economics, 2009

[BoZh-11] Boyd, G. A. ; Zhang, G.: Measuring Improvement in the Energy Performance of the U.S. Cement Industry / Duke University, Department of Economics, 2011

[Bram-13] Bramer, M.: Principles of Data Mining, Second Edition. London: Springer, 2013

[BrGu-14] Brickley, D.; Guha, R.V.: RDF Schema 1.1, W3C Recommendation 25 February 2014, http://www.w3.org/TR/2014/REC-rdf-schema-20140225/, access date 25/11/2015

[BuEt-03] Burke, E. K.; Hart, E.; Kendall, G.; Newall, J.; Ross, P.; Schulenburg, S.: Hyper-heuristics: An emerging direction in modern search technology, In: Handbook of Metaheuristics (F. Glover and G. Kochenberger, eds.), Kluwer, pp. 457–474, 2003

[BuEt-10] Burke, E.K.; Hyde, M.; Kendall, G.; Ochoa, G.; Ozcan, E.; Woodward, J.R.: A Classification of Hyper-heuristic Approaches, In: International Series in Operations Research and Management Science, pp. 449-468, 2010

[BuEt-11] Bunse, K.; Vodicka, M.; Schönsleben, P.; Brülhart, M.; Ernst, F.O.: Integrating energy efficiency performance in production management – gap analysis between industrial needs and scientific literature, in Journal of Cleaner Production, Volume 19, Issues 6–7, pp. 667-679, April–May 2011

[BUN-07] Die Bundesregierung.: G8: Verantwortung für Klima und Umwelt, 2007, http://bit.ly/qFSOtc, access date 22/03/2012

[Bunk-11] Bunk, M: Energiemanagementsysteme ab 2013 Pflicht, Essen: ener|gate Verlag, 2011

[Busi-12] Business Dictionary, http://www.businessdictionary.com/, access date 22/03/2012

[Buss-98] Bussmann, S.: Autonome und Kooperative Produktionssysteme. In: it+ti 4, Nr. August 1998, pp. 34–39, 1998

[CaEt-11] Cai, M.; Zhang, W.Y.; Zhang, K: ManuHub: A Semantic Web System for Ontology-Based Service Management in Distributed Manufacturing Environments, in: IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans, pp. 574-582, 2011

[CaEt-12] Cardenas, J.J.; Romeral, L.; Garcia, A.; Andrade, F.: Load forecasting framework of electricity consumptions for an Intelligent Energy Management System in the user-side, in: Journal Expert Systems with Applications, vol. 39, no. 5, pp. 5557–5565, 2012

[ChEt-07] Chapurlat, V.; Diep, D.; Kalogeras; A.; Gialelis, J.: Building and validating a Manufacturing Ontology to achieve Interoperability, in: Gonçalves, R.J., Müller, J.P., Mertins, K.,Zelm, M. Enterprise Interoperability II, pp. 261-272, 2007

[ChGT-96] Cheng, R.; Gen, M.; Tsujimura, Y.: A Tutorial Survey of Job-Shop Scheduling Problems Using Genetic Algorithms, I. Representation, n: Computers and Industrial Engineering, Vol. 30, No. 4, pp. 983-997, 1996

[CoEt-05] Corcho, O.; Fernández-López, M.; Gómez-Pérez, A.; López-Cima, A.: Building legal ontologies with METHONTOLOGY and WebODE. In: Lecture Notes in Computer Science 3369, pp. 142-157, 2005

[Coll-01] Collins, H.M.: Tacit knowledge, trust, and the Q of sapphire, In: Journal Social Studies of Science, vol. 31, no.1, pp.71-85, 2001

[DaEt-10] Daouadji, A.; Nguyen, K.-K.; Lemay, M.; Cheriet, M.: Ontology-based Resource Description and Discovery Framework For Low Carbon Grid Networks, In: First IEEE International Conference on Proceeding Smart Grid Communications (SmartGridComm), pp.477-482, 2010

[Dang-01] Dangelmaier, W.: Fertigungsplanung – Planung von Aufbau und Ablauf der Fertigung, Springer-Verlag 2001

[DaOS-03] Daconta, M.C.; Obrst, L.J.; Smith, K.T.: The Semantic Web: A Guide to the Future of XML, Web Services, and Knowledge Management, Indianapolis: Wiley Publishing, Inc, 2003

[DaSS-93] Davis, R.; Shrobe, H.; Szolovits, P.: What is a Knowledge Representation, An introductory critical paper by in AI Magazine, 1993

[Demi-10] DEMI consortium: DEMI - Product and Process Design for AmI Supported Energy Efficient Manufacturing Installations, Project Leaflet, 2010

[DENA-11] Deutsche Energie-Agentur: Initiative Energie Effizienz http://www.stromeffizienz.de/dienstleister-oeffentliche-hand/energiemanagement/ziele-und-vorteile.html, access date 11/10/2011

[DG-08] DG-Information Society and Media: ICT for Energy Efficiency, 2008

[DiEt-99] Diekmann, J. ; Eichhammer, W.; Neubert, A.; Rieke, H. ; Schlomann, B.; Ziesing, H.J. ; Müller, Werner A. (Hrsg.) ; Bihn, M. (Hrsg.): Energie-Effizienz-Indikatoren/ Statistische Grundlagen, theoretische Fundierung und Orientierungsbasis für die politische Praxis. Heidelberg : Physica-Verlag, 1999

[DoEt-14] Dogdu, E.; Murat Ozbayoglu, A.; Benli, O.; Akinc, H.E.; Erol, E.; Atasoy, T.; Gurec, O.; Ercin, O.: Ontology-centric data modelling and decision support in smart grid applications a distribution service operator perspective, in: Proceeding IEEE International Conference on Intelligent Energy and Power Systems (IEPS), pp. 198-204, 2014

[Dona-08] Donath, D.: Bauaufnahme und Planung im Bestand, Vieweg+Teubner Verlag, Wiesbaden, pp. 35-36, 2008

[DoSV-93] Domschke, W.; Scholl, A.; Voß, S.: Produktionsplanung, Springer-Verlag, Berlin; Heidelberg; New York, 1993

[ECIS-09] European Commission Information Society and Media: ICT and Energy Efficiency The Case for Manufacturing, European Communities, 2009, ftp://ftp.cordis.europa.eu/pub/fp7/ict/docs/micro-nanosystems/smart-manufacturing_en.pdf, Access Date: 15.11.2015

[EEA-13] European Environment Agency: Final energy consumption by sector, http://www.eea.europa.eu/data-and-maps/indicators/final-energy-consumption-by-sector-5/assessment, 2013, access date 03/12/2015

[EEP-13] Institut für Energieeffizienz in der Produktion (EEP) Universität Stuttgart: Energieeffizienz-Index der deutschen Industrie 1. Auswertung, access date 20/12/2013

[EEP-15] Institut für Energieeffizienz in der Produktion (EEP) Universität Stuttgart: Der Energieeffizienz-Index: Inhalte und Ziele, http://www.eep.uni-stuttgart.de/eei/, access date 24/11/2015

[Ener-10] Enerdata: Definition of ODEX indicators in ODYSSEE data base, Version: March 2010. http://www.odyssee-indicators.org/registred/definition_odex.pdf, access date 16/05/2011

[Ener-11] Enerdata: Energy Efficiency Indicators | Odyssee | Energy efficiency in Europe, http://www.odyssee-indicators.org/, access date 12/11/2011

[enhi-12] N.N.: EnHiPro project website, http://www.enhipro.de/de/home, access date 12/08/2012

[ENRW-13] EnergieAgentur.NRW: KURZ.energiecheck: www.energieagentur.nrw.de/kurzenergiecheck, access date 03/04/2013

[EO-11] Energy Office: Tools Energiemanagement: Definition http://www.energyoffice.org/english/tools/emanagement/definition/main.html, access date 11/10/2011

[ErWe-09] Erlach, K.; Westkämper, E.: Energiewertstrom. Der Weg zur energieeffizienten Fabrik, Fraunhofer-Verlag, 2009

[ErWe-09] Erlach, K.; Westkämpfer, E. (Hrsg): Energiewertstrom Der Weg zur energie-effizienten Fabrik, Stuttgart : Fraunhofer Verlag, 2009

[esav-13] N.N.: Project website e-Save, http://www.e-save.eu/, access date 24/01/2013

[esim-12] N.N.: e-SimPro project description, http://www.effizienzfabrik.de/sites/effizienzfabrik/files/Projektseiten%20e-simpro%20FINAL.pdf, access date 12/08/2012

[esto-12] N.N.: ESTOMAD project website, http://www.estomad.org/, access date 15/09/2012

[EuCo-15] European Commission: Horizon 2020 Work Programme 2014-2015 General Annexes, Consolidated version following European Commission Decision C, 4 December 2015

[Euro-03] European Commission. European Guide to Pump Efficiency for Single Stage Centrifugal Pumps. Technical Report. European Commission, pp. 1–8, 2003

[Euro-09] European Commission: Reference Document on Best Available Techniques for, Technical Report, 2009

[euro-13] N.N.: Project website Euronegest, http://www.euroenergest.eu/, access date 23/01/2013

[FaIr-93] Fayyad, U.M.; Irani, K.B.: Multi Interval Discretization of Continuous-Valued Attributes for Classification Learning, In: Proceeding of the International Joint Conference on Uncertainty in AI, pp. 1022-1027, 1993

[FaPS-96] Fayyad, U.; Piatestky-Shapiro, G.; Smyth,P.: From Data Mining to Knowledge Discovery in Databases, American Association for Artificial Intelligence, 1996

[Fark-11] Farkisch, K.: Data-Warehouse-Systeme Kompakt – Aufbau, Architektur, Grundfunktionen, Springer Heidelberg Dordrecht London New York, 2011

[FeEt-15] Ferrer, B.R.; Ahmad, B.; Lobov, A.; Vera, D.A.; Martinez Lastra, J.L.; Harrison, R.: An approach for knowledge-driven product, process and resource mappings for assembly automation, in: Proceeding IEEE International Conference on Automation Science and Engineering (CASE), pp. 1104-1109, 2015

[FeGJ-97] Fernández-López, M.; Gómez-Pérez, A.; Juristo, N.: METHONTOLOGY: From Ontological Art Towards Ontological Engineering. In: AAAI-97 Spring Symposium on Ontological Engineering, pp. 33–40, 1997

[FeGJ-97] Fernández, M.; Gómez-Pérez, A.; Juristo, N.: METHONTOLOGY: From Ontological Art Towards Ontological Engineering. Symposium on Ontological Engineering of AAAI. Stanford (California). March 1997

[FeMc-83] Feigenbaum, E.A.; McCorduck, P.: The fifth generation, 1st edition, Addison-Wesley, Reading, MA,1983

[Fens-00] Fensel, D.: "Ontologies: Silver Bullet for Knowledge Management and Electronic Commerce", Springer-Verlag, 2000

[Fern-99] Fernández López, M.: Overview Of Methodologies For Building Ontologies. In: Proceedings of the IJCAI-99 workshop on Ontologies and Problem-Solving Methods (KRR5) Bd. 17. Stockholm, Sweden, 1999

[FeSi-98] Ferstl, O. K.; Sinz, E. J.: Grundlagen der Wirtschaftsinformatik, vol. 1, Oldenbourg Verlag, 1998

[Frak-02] Frako Kondensatoren- und Anlagenbau: Handbuch der Blindstrom-Kompensation.

http://www.schmachtl.at/fileadmin/img/produkte/Energietechnik/124_Frako/Handbuch_95-00135_08_02_7121.pdf., online 2002, access date 12/05/2015

[Frau-08] Fraunhofer Gesselschaft: Energieeffizienz in der Produktion Untersuchung zum Handlungs- und Forschungsbedarf, 2008

[Fünf-00] Fünfgeld, C.: Energiekosten im Betrieb. Solar Promotion GmbH Verlag, 2000

[Geba-01] Gebauer, M.: Kooperative Produktentwicklung auf Basis verteilter Anforderungen, Dissertation Universität Karlsruhe; Shaker Verlag, Aachen, 2001

[GeBo-11] Gerdes, J.; Boonekamp, P.G.M.: Uncertainty in Odyssee indicators and energy savings – Development of a methodology and first results, technical report, ECN – The Netherland, 2011

[Geer-13] Geerts, F.: Competitiveness of the European machine tool industry, EPC Industrial Policy Task Force meeting, 27 June 2013

[Gere-66] Gere, W.S.: Heuristics in Job Shop Scheduling, In: Management Science, Vol. 13, No. 3, pp. 167-190, 1966

[Glov-89] Glover, F.: Tabu Search - Part 1. In: ORSA Journal on Computing, Vol. 1, No. 3, pp. 191-206, 1989

[GoKo-11] Goffin, K.; Koners, U.: Tacit Knowledge, Lessons Learnt, and New Product Development, In: Journal Product Innovation Management, vol. 28, pp. 300-318, 2011

[Góme-98] Gómez Pérez, A.: Knowledge Sharing and Reuse. In J. Liebowitz (Editor) Handbook of Expert Systems. CRC. 1998

[Gour-02] Gourlay, S.: Tacit knowledge, tacit knowing, or behaving? In: Proceeding 3rd European Organizational Knowledge, Learning and Capabilities Conference, Athens, Greece, 5-6 April 2002

[GrAF-00] Grüninger, M.; Atefi, K.; Fox, M. S.: Ontologies to Support Process Integration in Enterprise Engineering. In: Computational & Mathematical Organization Theory 6, pp. 381-394. 2000

[GrFo-95] Gruninger, M; Fox, M.S.: Methodology for the Design and Evaluation of Ontologies. In: Proceedings of the Workshop on Basic Ontological Issues in Knowledge Sharing, IJCAI-95, Montreal, 1995

[GrNP-11] Grassi, M.; Nucci, M.; Piazza, F.: Towards an ontology framework for intelligent smart home management and energy saving, In: Proceeding IEEE

International Symposium on Industrial Electronics (ISIE), pp. 1753-1758, June 2011

[Grub-07] Gruber, T.: Ontology - Definition in Encyclopedia of Database Systems, 2007. Available from World Wide Web: http://tomgruber.org/writing/ontology-definition-2007.htm

[Grub-93] Gruber, T. R.: "A Translation Approach to Portable Ontology Specifications". In: Knowledge Acquisition, 5(2), 1993

[GuTD-06] Gutowski, T.; Dahmus, J.; Thiriez, A.: Electrical Energy Requirements for Manufacturing Processes, in: Proceeding 13th CIRP International Conference on Life Cycle Engineering, Leuven, 2006

[HaEt-06] Harding, J. A.; Shahbaz, M.; Srinivas; Kusiak, A.: Data Mining in Manufacturing: A Review, in: Journal of Manufacturing Science and Engineering, ASME, vol. 128, no. 4, pp. 969-976, 2006

[HäEt-13] Häfner, P.; Häfner, P.; Wicaksono, W.; Ovtcharova, J.: Semi-automated Ontology Population from Building Construction Drawings, in: Proceedings of the International Conference on Knowledge Engineering and Ontology Development (KEOD), Vilamoura, Algarve, Portugal, 19-22 September, pp. 379-386, 2013

[HaJL-11] Han, J.; Jeong, Y.-K.; Lee, I.: Efficient Building Energy Management System Based on Ontology, Inference Rules, and Simulation, In: Proceeding International Conference on Intelligent Building and Management. Singapore: IACSIT, 2011

[HaJL-15] Han J.; Jeong J.K.; Lee, I.: A Rule-Based Ontology Reasoning System for Context-Aware Building Energy Management, in: Proceeding IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing (CIT/IUCC/DASC/PICOM), pp. 2134-2142, 2015

[HaKa-01] Han, J.; Kamber, M.: Data Mining Concept and Techniques, Second Edition, Morgan Kaufmann Publisher, 2001

[HaTM-11] Hai, R.; Theißen, M.; Marquardt, W.: An ontology-based approach for operational process modeling, In: Journal Advanced Engineering Informatics, vol. 25, no.4, pp.748-759, 2011

[HaVr-15] Harcuba, O.; Vrba, P.: Ontologies for flexible production systems, in: Proceeding IEEE 20th Conference on Emerging Technologies Factory Automation (ETFA), pp. 1-8, 2015

[HaWL-83] Hayes-Roth, F.; Waterman, D.A.; Lenat, D.B.: Building Expert Systems, Addison-Wesley, New York, 1983

[HeEt-10] Heilala, J.; Klobut, K.; Salonen, T.; Järvinen, P.; Siltanen, S.; Shemeikka, J.: Energy Efficiency Enhancement in Discrete Manufacturing Process with Energy Use Parameters, In: Proceeding 2010 International Conference on Advances in Production Management Systems, Competitive and Sustainable Manufacturing Products and Service (APMS), 2010

[HeEt-11] Heilala, J.; Klobut, K.; Salonen, T.; Siltanen, P.; Ruusu, R.; Armijo, A.; Sorli, M.; Urosevic, L.; Reimer, P.; Fatur, T.; Gantar, Z.; Jung, A: Ambient Intelligence based monitoring and energy efficiency optimisation system, In: Proceeding 2011 IEEE International Symposium on Assembly and Manufacturing (ISAM), pp. 1-6, 2011

[HiKR-10] Hitzler, P.; Krötzsch, M.; Rudolph, S.: Foundation of Semantic Web Technologies. Boca Raton: Chapman & Hall/CRC, 2010

[HoEt-04] Horrocks, I.; Patel-Schneider, P.F.; Boley, H.; Tabet, S.; Grosof, B.; Dean, M.: SWRL: A Semantic Web Rule Language Combining OWL and RuleM, W3C Member Submission 21 May 2004

[HoSa-12] Horrocks, I.; Sattler, U.: Description Logics: Basics, Applications and More, http://www.cs.manchester.ac.uk/~horrocks/Slides/ecai- handout.pdf, access date 27/12/2012

[HuGa-08] Husain, B.; Ganz, C.: ICT Contribution to Energy Efficiency in Manufacturing Industries, ABB Corporate Research, 2008

[IEA-12] International Energy Agency: Global carbon-dioxide emissions increase by 1.0 Gt in 2011 to record high. http://www.iea.org/newsroomandevents/news/2012/may/name,27216,en.html, access date 15/10/2012

[IrTh-06] Irrek, W.; Thomas, S.: Der EnergieSparFonds für Deutschland, Hans-Böckler-Stiftung, Düsseldorf, 2006

[ises-13] N.N.: Project website ISES, http://ises.eu-project.info, access date 24/01/2013

[ISO-11] International Organization for Standardization: Win the energy challenge with ISO 50001, ISO 2011-06/3000, 2011

[JaZP-99] Jacobsen, C.; Zscherpel,U.; Perner, P.: A Comparison between Neural Networks and Decision Trees, Lecture Notes in Computer Science: Machine Learning and Data Mining in Pattern Recognition, Springer Berlin Heidelberg, pp.144-158, 1999

[Jenn-01] Jenne, F.: PDM-basiertes Entwicklungsmonitoring Ein Beitrag zur Planung und Steuerung von Entwicklungsprozessen, Aachen: Shaker, 2001

[JoBV-98] Jones, D.; Bench-Capon, T.; Visser, P.: Methodologies for ontology development. In: Proceedings of the 15th IFIP World Computer Congress, 1998

[KaEt-10] Kahlenborn, W., Kabisch S., Klein J., Richter I., Schürmann S.: DIN EN 16001: Energy Management Systems in Practice - A Guide for Companies and Organisations. Bundesministerium für Umwelt, Naturschutz und Reaktorsicherheit, Berlin, Germany, 2010

[KaEt-12] Kahlenborn, W., Kabisch S., Klein J., Richter I., Schürmann S.: Energiemanagementsysteme in der Praxis ISO 50001: Leitfaden für Unternehmen und Organisationen. Bundesministerium für Umwelt, Naturschutz und Reaktorsicherheit, Berlin, Germany, 2012

[Kals-10] Kals, J.: Betriebliches Energiemanagement. Neustadt an der Weinstraße/Ludwigshafen: Kohlhammer GmbH, 2010

[Kals-10] Kals, J.: Betriebliches Energiemanagement - Eine Einführung. Stuttgart : Verlag W. Kohlhammer, 2010

[Kama-11] Kamal, A.M.M.: A Data Mining Approach for Improving Manufacturing Processes Quality Control, In: Proceeding 2nd International Conference in Next Generation Information Technology (ICNIT), pp. 207-212, 2011

[Kamp-99] Kamp, J.A.L.M.: Knowledge-based systems: from research to practical application: pitfalls and critical success factors, In: Journal Computers and Electronics in Agriculture, vol. 22, no. 23, pp. 243-250, 1999

[KeCr-07] Kendal, S.L.; Creen, M.: An introduction to knowledge engineering, London: Springer, 2007

[KEHL-11] N.N: KEHL (Kontrollierte Energie Haushalt Lösungen) Project Homepage, http://kehl.actihome.de/, access date 16/02/2011

[Kern-80] Kern, W.: Industrielle Produktionswirtschaft. 3. überarbeitete Auflage, Poeschel Verlag, Stuttgart, 1980

[KhEt-15] Khan, A.R.; Schioler, H.; Knudsen, T.; Kulahci, M.: Statistical data mining for efficient quality control in manufacturing, in: Proceeding IEEE 20th Conference on Emerging Technologies Factory Automation (ETFA), pp. 1-4, 2015

[KiGV-83] Kirkpatrick, D.; Gelatt, C.D.; Vecchi, M.P.: Optimization by Simulated Annealing. In: Science, Vol. 220, No. 4598, pp. 671-680, 1983

[Kink-05] Kinkel, S.: Anforderungen an die Fertigungstechnik von morgen – Wie verändern sich Variantenzahlen, Losgrößen, Materialeinsatz, Genauigkeitsanforderungen und Produktlebenszyklen tatsächlich? Fraunhofer Institut ISI, Mitteilungen aus der Produktionsinnovationserhebung Nr. 37, September 2005

[KiSK-11] Kim, H.; Stumpf, A.; Kim, W.: Analysis of an energy efficient building design through data mining approach, In: Journal Automation in Construction, vol. 20, no. 1, pp. 37-43, 2011

[KlCM-14] Klyne, G.; Carroll, J.J.; McBride, B.: RDF 1.1 Concepts and Abstract Syntax, W3C Recommendation 25 February 2014, http://www.w3.org/TR/rdf11-concepts/, access date 25/11/2015

[Klim-03] Klimesch, C.: Ein Beitrag zur prozessgetriebenen Informationslogistik durch kontextorientiertes domänenübergreifendes Wissensmangement, Karlsruhe, Shaker- Verlag, 2003

[KöBT-11] Köksal, G.; Batmaz, I.; Testik, M.C.: A review of data mining applications for quality improvement in manufacturing industry, In: Journal Expert Systems with Applications, vol. 38, no. 10, pp. 13448 – 13467, 2011

[Kram-94] Kramer, D.: Kostenorientierte Reihenfolgeplanung, Verlag Josef Eul, Bergisch Gladbach, Köln, 1994

[KrEt-09] Krahtov, K.; Rogalski, S.; Wacker, D.; Gutu, D.; Ovtcharova, J.: A Generic Framework for Life-Cycle-Management of heterogenic Building Automation Systems, Proceedings to Flexible Automation and Intelligent Manufacturing, 19th International Conference (FAIM 2009), July 6th-8th, 2009

[Krie-12] Kriesel, D.: A Brief Introduction to Neural Networks, online book, http://www.dkriesel.com/_media/science/neuronalenetze-en-zeta2-2col-dkrieselcom.pdf, access date 22/11/2012

[Lam-00] Lam, A.: Tacit knowledge, organizational learning and societal institutions: An integrated framework, In: Organization Studies, Vol. 21, No. 3, pp. 487-513, 2000

[Lanz-11] Lanza, G. Prof. Dr.-Ing.: Integrierte Produktionsplanung, Karlsruher Institut of Technology, Institute of Production Science, 2011

[Lebe-08] S. Lebert: Gruppen segmentieren mit Entscheidungsbäumen: Ein praxisorientierter Methodenvergleich am Beispiel von PC-Nutzung, Vdm Verlag Dr. Müller, April 2008

[LeEt-15] Lee, J.; Bagheri, B.; Kao, H.: A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems, Journal Manufacturing Letters vol. 3, 2015

[Leut-02] Leutsch, M.: Unterstützung des Konstruktionsprozess durch integration von prozeduralem Wissen, PhD thesis, vol.8, Shaker Verlag, 2002

[life-12] N.N.: Project website Lifesaver, http://www.lifesaver-fp7.eu/, access date 15/09/2012

[LiHa-07] Lin, H.; Harding, J.: A manufacturing system engineering ontology model on the semantic web for inter-enterprise collaboration. In: Computers in Industry vol. 58, Nr. 5, pp. 428–437, 2007

[Lu-91] Lu, S: Die Generation nach regelbasierten Expertensystemen – Stand er Forschung in technischen Anwendungen der künstlichen Intelligenz in den USA, In: VDI Report 903 – Proceeding „Wissensbasierte System“, Heidelberg: VDI-Verlag, 1991

[Lunz-94] Lunze, J.: Künstliche Intelligenz für Ingenieure, vol. 1: Methodische Grundlagen und Softwaretechnologie. München, Wien: Oldenbourg Verlag, 1994

[Mahl-09] Mahl, A: Ein Beitrag zur wissensbasierten Unterstützung unternehmensübergreifender interkultureller Kooperationen auf der Grundlage eines generischen Baukastensystems, Dissertation, Karlsruhe Insitute of Technology, 2009

[Mand-07] Mandil, C.: Making the difference: energy efficiency. In: G8 Summit 2007 Growth and Responsibility, Buxton Press, London, 2007

[Mann-07] Manninen, J., VTT: Achievable Gains in Energy Consumption of Manufacturing Processes; Bernstein, L. et al.: Industry, in: Metz, O. R. et al. (eds.): Climate Change 2007: Mitigation. Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NU, USA; Gupta, S. et al.: Policies, Instruments and Co-operative Arrangements, in: Metz, O. R. et al., 2007

[Manu-04] Manufuture: A vision for 2020. Assuring the future of manufacturing in Europe, 2004, http://www.manufuture.org, access date 11/10/2012

[McEt-00] McGuinness, D.L.; Fikes, R.; Rice, J.; Wilder, S.: An Environment for Merging and Testing Large Ontologies. In: Principles of Knowledge Representation and Reasoning Proceedings of the Seventh International Conference (KR2000). A. G. Cohn, F. Giunchiglia and B. Selman, editors. San Francisco, CA, Morgan Kaufmann Publishers, 2000

[McEt-16] McGlinn, K.; Wicaksono, H; Lawton, W.; Weise, M.; Kaklanis, N.; Petri, I.; Tzovaras, D.: Identifying Use Cases and Data Requirements for BIM Based Energy Management Processes, CIBSE Technical Symposium, Heriott Watt University, Edinburg, 2016

[McVa-04] McGuinness, D.L.; van Harmelen, F.: OWL Web Ontology Language Overview, W3C Recommendation 10 February 2004, http://www.w3.org/TR/owl-features/, access date 06/01/2013

[Mer-09] Merkel, A. Regierungserklärung zu den Ergebnissen des Europäischen Rates sowie zur UN-Klimakonferenz von Bundeskanzlerin Merkel, 2009, http://bit.ly/nthEC6, access date 17/10/2011

[Meye-11] Meyer, L.: Automated Surveillance, Management and Optimisation of Energy Utilisation. Efficiency Days Track, Hannover Messe April 8th, 2011

[Moos-10] Moosmayer, V.: Erfüllung der Anforderungen der DIN EN 16001 Energiemanagementsysteme durch EMAS III, Berlin, 2010

[Mott-99] Motta, E.: Reusable Components for Knowledge Modelling, IOS Press, 1999

[Müet-09] Müller, E.; Engelmann, J; Löffler, T; Strauch, J: Energieeffiziente Fabriken planen und betreiben. Springer Verlag, 2009

[Neap-04] Neapolitan, R.E.: Learning Bayesian Networks. New Jersey: Pearson Prentice Hall, 2004

[Negn-05] Negnevitsky, M.: Artificial intelligence: A guide to intelligence systems, Addison Wesley, 2005

[Neuh-01] Neuhausen, J.: Methodik zur Gestaltung modularer Produktionssysteme für Unternehmen der Serienproduktion. Dissertation RWTH Aachen, 2001

[Neuh-01] Neuhausen, J.: Methodik zur Gestaltung modularer Produktionssysteme für Unternehmen der Serienproduktion, Dissertation, RTWH Aachen, November 2001

[Neum-96] Neumann, K.: Produktions- und Operations-Management. Springer Verlag, Berlin, 1996

[NN-04] N.N.: Regular Expression, The Open Group Base Specifications Issue 6, IEEE Std 1003.1, 2004, access date 20/01/2013

[NoMc-01] Noy, N.; McGuinness, D. Ontology development 101: A guide to creating your first ontology. 2001

[Ovtc-11] Ovtcharova, J.: Product Lifecycle Management, official summary 2011- IMI Insitute for Information engineering at Karlsruhe Institute of Technology, 2011

[OWL-12] W3C Semantic Web OWL Working Group: Web Ontology Language (OWL), published 2012, https://www.w3.org/OWL/, access date 25/05/2016

[Oxfo-12] Oxford Dictionaries, http://english.oxforddictionaries.com/, access date 22/03/2012

[OzBK-08] Ozcan, E.; Bilgin, B.; Korkmaz, E.E.: A Comprehensive Analysis of Hyper-heuristics, In: Intelligent Data Analysis, vol. 12 no.1, pp. 3-23, 2008

[PaDT-12] Panetto, H.; Dassisti, M.; Tursi, A.: ONTO-PDM: Product-driven ONTOlogy for Product Data Management interoperability within manufacturing process environment, In: Journal Advanced Engineering Informatics, vo.26, no. 2, pp.334-348, 2012

[PaMo-08] Panetto, H.; Molina, A.: Enterprise Integration and Interoperability in Manufacturing Systems: trends and issues, In: Special issue on Enterprise Integration and Interoperability in Manufacturing Systems, A. Molina and H. Panetto (Eds). Computers In Industry, 59/5, May, Elsevier, 2008

[Pars-12] Parsia, B.: What is a Knowledge Representation?, COMP34512, University of Manchester, 2012

[Patt-96] Patterson, G.: What is energy efficiency?: Concepts, indicators and methodological issues, in: Journal of Energy Policy, vol. 24, no.5, pp. 377-390, 1996

[Peme-00] Pezzella, F.; Merelli. E.: A tabu search method guided by shifting bottleneck for the job shop scheduling problem, In: European Journal of Operational Research, No. 120, pp. 297-310, 2000

[Pete-05] Petersohn, H.: Datamining : Verfahren, Prozesse, Anwendungsarchitektur. München: Oldenbourg Wissenschaftsverlag, 2005

[PhBW-97] Phylipsen, G.J.M. ; Blok, K.; Worrell, E.: International comparisons of energy efficiency methodologies for the manufacturing industry, in Energy Policy, 25, pp. 715–725, 1997

[Pist-09] Postohl, W.: Handbuch der Gebäudetechnik- Heizung, Lüftung, Beleuchtung, Energiesparen- Band 2. Werner Verlag. 7. Auflage, August, 2009

[Pocs-00] Pocsai, Z.: Ontologiebasiertes Wissensmanagement für die Produktentwicklung. Aachen: Shaker, 2000

[Pola-66] Polanyi, M.: The tacit dimensions, University Of Chicago Press, 1966

[PrEt-00] Preece, A.; Flett, A.; Sleeman, D.; Curry, D., Meany N., Perry, P.: Better Knowledge Management through Knowledge Engineering: A Case Study in Drilling Optimisation. IEEE Intelligent Systems, vol. 2001, pp. 36-42, 2000

[PrSe-08] Prud'hommeaux, E.; Seaborne, A.: SPARQL Language for RDF, W3C Recommendation, published 2008 https://www.w3.org/TR/rdf-sparql-query/, acces date 25/05/2016

[Qzuin-93] Quinlan, J.R.: C4.5: Programs for Machine Learning, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1993

[RDF-14] W3C Semantic Web RDF Working Group: Resource Description Framework (RDF), published 2014, https://www.w3.org/RDF/, access date 25/05/2016

[reBO-12] reBOP consortium: reBOP- Bewertung und Optimierung von Produktionsprozessen, Effizienzfabrik, 2012

[ReWe-10] Rey, G.D.; Wender, K.F.: Neuronale Netze – Eine Einführung in die Grundlagen, Anwendungen und Datenauswertung, Huber Bern, 2010

[Ries-12] Riese, P.: Handbuch der Blindleistungs-Kompensation, http://www.frako.com/fileadmin/pdf/Downloads/Handbuch/95-00135_11_13_9066_handbuch_blk.pdf, access date 02/12/2015

[RIG-11] N.N.: Research in Germany - EWOTeK, 2011, http://bit.ly/rfe0T6, access date 20/09/2011

[RoEt-11] Rosello-Busquet, A.; Brewka, L. J.; Soler,J.; Dittman, L.: OWL Ontologies and SWRL Rules Applied to Energy Management, In: Proceeding UKSim 13th International Conference on Modelling and Simulation. Cambridge, 2011

[Roga-09] Rogalski, S.: Entwicklung einer Methodik zur Flexibilitätsbewertung von Produktionssystemen, Dissertation, Karlsruhe Institute of Technology, May 2009

[Ropo-99] Ropohl, G.: Allgemeine Technologie- Eine Systemtheorie der Technik. 2. Edition, Carl Hanser Verlag, München, Wien, 1999

[Ross-05] Ross, P.: Hyper-Heuristics, In: E. K. Burke and G. Kendall, Eds., SearchMethodologies, Springer US, pp. 529-556, 2005

[Rude-98] Rude, S.: Wissensbasiertes Konstruieren, Habilitation, University of Karlsruhe, Shaker Verlag, Aachen, 1998

[RuHV-13] Rumph, F.J.; Huitema, G.B.; Verhoosel, J.: Towards interoperability of energy management systems using flexibility through conceptual modeling, In: Proceeding of 2013 IFIP/IEEE International Symposium on Integrated Network Management (IM 2013), pp. 744-747, 2013

[RuNo-03] Russell, S.J.; Norvig, P.: Artificial Intelligence: A Modern Approach, Upper Saddle River, New Jersey: Prentice Hall, 2003

[Rzeh-98] Rzehorz, C.: Wissensbasierte Anforderungsentwicklung auf Basis eines integrierten Produktmodells, Dissertation Universität Karlsruhe; Shaker Verlag, Aachen, 1998

[ScEt-00] Schreiber, G.; de Hoog, R.; Akkermans, H.; Anjewierden, A.; Shadbolt, N. van de Velde, W.: Knowledge Engineering and Management, MIT Press, 2000

[ScEt-06] Schuh, G. (Hrsg.) et al: Produktionsplanung und –steuerung, Grundlagen, Gestaltung und Konzepte, Springer-Verlag, 2006

[ScFB-05] Schieferdecker, B.; Fünfgeld, C.; Bonneschky, A.: Energiemanagement-Tools- Anwendung im Industrieunternehmen, Springer, 2006

[Schi-87] Schieferdecker, B.: Rationelle Energieanwendung und Energie-Recycling im Industriebetrieb zur Sicherung einer hohen ökonomischen Effektivität unter den Bedingungen der intensive erweiterten Reproduktion, VEB, Leipzig, 1987

[Schi--96] Schiemenz, B.: Komplexität von Produktionssystemen. In: Kern, W.; Schröder, H.-H.; Weber, J. (Hrsg.): Handwörterbuch der Produktionswirtschaft, vol. 2, Stuttgart 1996

[ScMe-09] Schilberg, D.; Meisen, T.: Ontology Based Semantic Interconnection of Distributed Numerical Simulations for Virtual Production, In: Proceeding 16th International Conference on Industrial Engineering and Engineering Management (IEEM '09), pp. 1789 -1793, 2009

[ScWR-11] Schubert, V.; Wicaksono, H.; Rogalski, S.: A Product Life-Cycle Oriented Approach for Knowledge-Based Product Configuration System, Global Product Development, Bernard, Alain (Ed.), Springer Verlag, Heidelberg-London-New York, 2011

[Shie-06] Schieferdecker, B.: Energiemanagement-Tools. Heidelberg: Springer-Verlag, 2006

[ShMe-12] Shahzad, A.; Mebarki, N.: Data mining based job dispatching using hybrid simulation-optimization approach for shop scheduling problem, In: Journal Engineering Applications of Artificial Intelligence, vol. 25, no. 6, pp. 1173 – 1181, 2012

[ShWB-09] Schröter, M.; Weißfloch, U.; Buschak, D.: Energieeffizienz in der Produktion – Wunsch oder Wirklichkeit?, Fraunhofer-Institut für System- und Innovationsforschung ISI, 2009

[SiEt-06] Siadat, A.; Dantan, J.-y.; Semenenko, A.; Fresnel, A.: MASON : A Proposal For An Ontology Of Manufacturing Domain. In: Proceedings of the IEEE Workshop on Distributed Intelligent Systems: Collective Intelligence and Its Applications (DIS’06), 2006. – ISBN 076952589X, S. 6–11

[SiEt-11] Silva, M.; Elias, E.; Costa, E.; Bittencourt, I. I.; Barros, H.; Dias, L.; Pedro, A.; Véras, D.: Combining Methontology and Ontology-Driven Approach to Build an Educational Ontology. In: IEEE MULTIDISCIPLINARY ENGINEERING EDUCATION MAGAZINE 6, 2011

[SpEt-01] Speel, P.H.; Schreiber A. Th.; van Joolingen, W.; van Heijst, G.; Beijer, G.J.: Conceptual Modelling for Knowledge-Based Systems, In: Encyclopedia for Computer Science and Technology, New York: Marcel Dekker, Inc, 2001

[StBF-98] Studer, R., Benjamins V. R., Fensel, D: Knowledge Engineering: Principles and Methods. Institute AIFB, University of Karlsruhe, 1998

[StEt-93] Storer, R. H.; Wu, S. D.; Vaccari, R.: New Search Spaces for Sequencing Problems with Application to Job Shop Scheduling, Management Science, Vol. 38, No. 10, pp. 1495-1509, 1993

[StMa-02] Studer, R.; Maedche, A..: "Interoperable Anwendungssysteme.", Vorlesung. AIFB, Universität Karlsruhe (TH), Karlsruhe, 2002

[Stor-10] Storey, R.: ISO 50001 An International Energy Management Effort, PowerLogic ION Users Group Conference, Toronto, October 2010

[Stor-89] Storm, S.; Schieferdecker, B.: Verbesserte Erfassung und Abrechnung der betrieblichen Energiekosten, Energietechnik, vol. 39, no. 10, 1989

[StWV-92] Storer, R.H.; Wu, S.D.; Vaccari R.: New Search Spaces for Sequencing Problems with Application to Job Shop Scheduling, In: Management Science, Vol. 38, No. 10, pp. 1495-1509, 1992

[SyEt-13] Symens, W. and ESToMaD partners: Energy efficient machines through model-based design - Public summary of what the ESToMaD achievements bring to the machine construction industry, Deliverable EU STREP ICT for Mobility, Environmental Sustainability and Energy Efficiency project ESToMaD, 2013

[Tana-08] Tanaka, K.: Assessing Measures of Energy Efficiency Performance and their Application in Industry, IEA Information Paper in Support of the G8 Plan of Action, International Energy Agency, 2008

[TeEt-11a] Fundación Tecnalia Research & Innovation; ATB - Institute for Applied Systems Technology Bremen GmbH; IVL Swedish Environmental Research Institute Ltd.; Lithuanian Energy Institute: DEMI Methodology Specification, Deliverable D200.2, 2011

[TeEt-11b] Fundación Tecnalia Research & Innovation; ATB - Institute for Applied Systems Technology Bremen GmbH; IVL Swedish Environmental Research Institute Ltd.; Lithuanian Energy Institute: DEMI RTD Methodology, Deliverable D300.1, 2011

[Teic-98] Teich, T.: Optimierung von Maschinenbelegungsplänen unter Benutzung heuristischer Verfahren., Josef Eul Verlag, Köln, 1998

[Trnk-10] Trnka, A.: Classification and Regression Trees as a Part of Data Mining in Six Sigma Methodology. WCECS 2010: World Congress on Engineering and Computer Science, International Association of Engineers, pp. 449–453, 2010

[Trös-05] Tröster, F.: Steuerungs- und Regelungstechnik für Ingenieure. 2. Auflage, Oldenbourg Wissenschaftsverlag, München, Wien, 2005

[TuAL-05] Turban, E.; Aronson, J.E.; Liang, T.-P.: Decision support systems and intelligent systems, Pearson/Prentice Hall, 2005

[USEI-13] U.S. Energy Information Administration: Energy Consumption and Efficiency, http://www.eia.gov/consumption/ access date 26/04/2013

[UsGr-96] Uschold, M.; Gruninger, M. (1996). Ontologies: Principles, Methods and Applications. Knowledge Engineering Review 11(2), 1996

[UsKi-95] Uschold, M.; King, M.: Towards a Methodology for Building Ontologies. Workshop on Basic Ontological Issues in Knowledge Sharing. 1995

[VaLu-06] Van Dam, K.H.; Lukszo, Z.: Modelling Energy and Transport Infrastructures as a Multi-Agent System using a Generic Ontology, In: Proceeding IEEE International Conference on Systems. Man, and Cybernetics, Taipei, 2006

[VaNC-08] Vanlande, R.; Nicolle, C.; Cruz, C.: IFC and building lifecycle management, Automation in Construction, Volume 18, Issue 1, December 2008, pp. 70-78., 2008

[VaTK-11] Vazan, P.; Tanuska, P., Kebisek, M.: The Data Mining usage in Pro-duction System Management, World Academy of Science, Engineering and Technology 77, 2011

[VDI-03] Verein Deutscher Ingenieure: VDI 4661 Energetic characteristics definitions- terms- methodology. Berlin: Beuth Verlag GmbH, 2003

[VDI-07] VDI Verein Deutscher Ingenieure: VDI Richtlinie 4602/ Energiemanagement – Begriffe, Oktober 2007

[VeDa-12] Veleva, S.; Davcev, D.: Data Mining as an Enabling Technology for Home Energy Management System, In: Innovative Smart Grid Technologies (ISGT), 2012 IEEE PES. Washington D.C., 2012

[W3C-13] World Wide Web Consortium: Semantic Web Activity 2013, http://www.w3.org/2001/sw/, access date 25/11/2015

[WEC-11] World Energy Council: Energy Efficiency Policies around the World: Review and Evaluation, 2011, http://bit.ly/qi0zBo, access date 13/02/2012

[WeEt-15] Weise, M.; Villalón, M.P.; Castro, R.G.; Euzenat, J.; Priego, L.M.; Fies, B.; Cavallaro, A.; Peters-Anders, J.; Tsagkari, K.Z.: Ontologies and datasets for energy management system interoperability, Deliverable D2.3 Ready4SmartCities, 2015

[WiAR-12] Wicaksono, H.; Aleksandrov, K.; Rogalski, S.: An Intelligent System for Improving Energy Efficiency in Building Using Ontology and Building Automation Systems, in: Kongoli, F. (Ed.), Automation, Publisher: InTech, Open Access Publisher, pp. 531-548, 2012

[WiEt-11] Wicaksono, H.; Schubert, V.; Rogalski, S.; Laydi, Y.A; Ovtcharova, J.: Ontology-driven Requirements Elicitation in Product Configuration Systems, Enabling Manufacturing Competitiveness and Economic Sustainability, ElMaraghy, H. A. (Ed.), Springer Verlag, Berlin, Heidelberg, pp. 63-67, 2011

[WiEt-12] Wicaksono, H.; Aleksandrov, K.; Rogalski, S.; Ovtcharova, J.: An ICT Supported Holistic Approach for Qualitative and Quantitative Energy Efficiency Evaluation in Manufacturing Company, In: Proceeding International Conference Advances in Production Management Systems (APMS 2012), Rhodes Island, Greece, 2012

[WiEt-13] Wicaksono, H.; Belzner, T.; Ovtcharova, J.: Efficient Energy Performance Indicators for Different Level of Production Organizations in Manufacturing Companies, In: Proceeding International Conference Advances in Production Management Systems (APMS 2013), State College, PA, United States, 2013

[WiEt-14] Wicaksono, H.; Jost, F.; Rogalski, S.; Ovtcharova, J.: Energy Efficiency Evaluation in manufacturing through an Ontology-Represented Knowledge Base, In: Intelligent System in Accounting, Finance and Management 21(1): pp. 59-69, 2014

[WiEt-15] Wicaksono, H.; Dobreva, P.; Häfner, P.; Rogalski, S.: Methodology to Develop Ontological Building Information Model for Energy Management

System in Building Operational Phase, In: Knowledge Discovery, Knowledge Engineering and Knowledge Management Communications in Computer and Information Science Volume 454, pp 168-181, 2015

[WiFH-05] Witten, I.H.; Frank, E.: Data Mining Practical Machine Learning Tools and Techniques, San Francisco: Elsevier, 2005

[WiFH-11] Witten, I.H.; Frank, E.; Hall, M. A.: Data Mining: Practical Machine Learning Tools and Techniques. Morgan Kaufmann, 2011

[WiHü-11] Windt, K., Hütt, M.-T.: Exploring due date reliability in production systems using data mining methods adapted from gene expression analysis, Journal CIRP Annals - Manufacturing Technology, vol. 60, no. 1, pp. 473-476, 2011

[WiOv-12] Wicaksono, H.; Ovtcharova, J.: Energy Consumption Regulation and Optimization in Discrete Manufacturing through Semi-automatic Knowledge Generation Using Data Mining, In: Proceeding The 10th Global Conference on Sustainable Manufacturing (GCSM), Istanbul, Turkey, 2012

[WIPO-13] Wicaksono, H.; Prohl, E.V.; Ovtcharova, J.: Hyper-heuristic based production process scheduling to improve productivity in sustainable manufacturing, 22nd International Conference on Production Research (ICPR 22), Iguassu Falls, Brazil, 28 Jul – 01 Aug 2013

[WiRK-10] Wicaksono, H.; Rogalski, S.; Kusnady, E.: Knowledge-based Intelligent Energy Management Using Building Automation System, In: Proceeding IPEC, pp. 11040-11045, 2010

[WiRo-10a] Wicaksono, H.; Rogalski, S.: Ontology Supported Intelligent Energy Management System in Buildings, In: Proceeding IEEE International Conference on Industrial Engineering and Business Management (ICIEBM), Yogyakarta, Indonesia, 12-13 October 2010

[WiRo-10b] Wicaksono, H.; Rogalski, S.: Wissensbasierte Energieanalyse – Verbesserte Energieeffizienz in Gebäuden des Privat- und Geschäftsbereichs, Zeitschrift für wirtschaftlichen Fabrikbetrieb (ZWF), 105. Jahrgang, Hanser Verlag, Heft 06/2010, pp. 551-555, 2010

[WiRO-12] Wicaksono, H.; Rogalski, S.; Ovtcharova, J.: Ontology-Driven Approach for Intelligent Energy Management in Discrete Manufacturing, In: Proceedings of the International Conference on Knowledge Engineering and Ontology Development (KEOD), Barcelona, Spain, 4 - 7 October 2012

[YuEt-12] Yu, Z.; Haghighat, F.; Fung, B.C.M.; Zhou, L.: A novel methodology for knowledge discovery through mining associations between building operational data, In: Journal Energy and Building, vol. 47, pp. 430-440, 2012

[Zäpf-07] Zäpfel, G.: Taktisches Produktionsmanagement. 2. Auflage, Oldenbourg Wissenschaftsverlagsverlag, München, 2007

[ZaTI-08] Zobolas, G.I.; Tarantilis, C.D; Ioannou, G.: Exact, Heuristic and Metaheuristic Algorithms for Solving Shop Scheduling Problems, in Xhafa, Abraham (eds.), Metaheuristics for Scheduling in Industrial and Manufacturing Applications, Berlin, 2008

[ZuEt-14] Zucker, G.; Malinao, J.; Habib, U.; Leber, T.; Preisler, A.; Judex, F.: Improving energy efficiency of buildings using data mining technologies, in IEEE 23rd International Symposium on Industrial Electronics (ISIE) pp.2664-2669, 2014

Appendix A Ontology Languages

A.1 Resource Description Framework (RDF)

Resource Description Framework (RDF) is a subset of XML with a standardized model for data interchange on the web. RDF allows structured and semi-structured data to be mixed, exposed, and shared across different applications [RDF-14]. RDF describes a directed graph, where a set of nodes are connected with directed edges. In RDF any object is described as a resource. A resource is either a blank node or a node identified by Uniform Resource Identifier (URI). Blank nodes are also called anonymous resource. The URI also acts as a unique identifier of the resource. Property is a type of resource that describes the relation between resources. Data value in RDF is represented as literal. The interpretation of string in literal is defined by datatype. The datatype of XML schema are used in RDF, for example, `xsd:string`, `xsd:Boolean`, `xsd:date`, etc.

RDF defined a triple consisting subject, predicate, and object. The triple is called statement or RDF-Triple *T(subject, predicate, object).* A predicate relates a subject to an object. A resource can occur a subject, a predicate or an object, whereas a literal can only represent an object. Figure A.1 depicts an example of a RDF statement "cutting machine produces metal ring." Resources occur as the subject, the predicate, and the object.

Figure A.1 Example of a RDF statement having resources as subject, predicate, and object

There are several possibilities to express the RDF graph, for instance, RDF/XML, Turtle, and Notation3 (N3). Figure A.2 and Figure A.3 illustrate the RDF statement in Figure A.1, expressed in Turtle and RDF/XML.

```
@prefix imi:  <http://www.imi.kit.edu/rdf/#> .

imi:cutting_machine imi:produces imi:metal_ring .
```

Figure A.2 Example of an RDF statement having resources as subject, predicate, and object expressed in Turtle

```
<?xml version="1.0"?>

<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:imi="http://www.imi.kit.edu/rdf/#">

  <rdf:Description
     rdf:about="http://www.imi.kit.edu/rdf/#cutting_machine">
     <imi:produces>
         <rdf:Description
             rdf:about=”http://www.imi.kit.edu/rdf/#metal_ring”>
         </rdf:Description>
     </imi:produces>
  </rdf:Description>
</rdf:RDF>
```

Figure A.3 Example of an RDF statement having resources as subject, predicate, and object expressed in RDF/XML

Figure A.4 gives an example of a RDF statement containing subject and predicate resources and object literal. The corresponding Turtle and RDF/XML notations are illustrated in

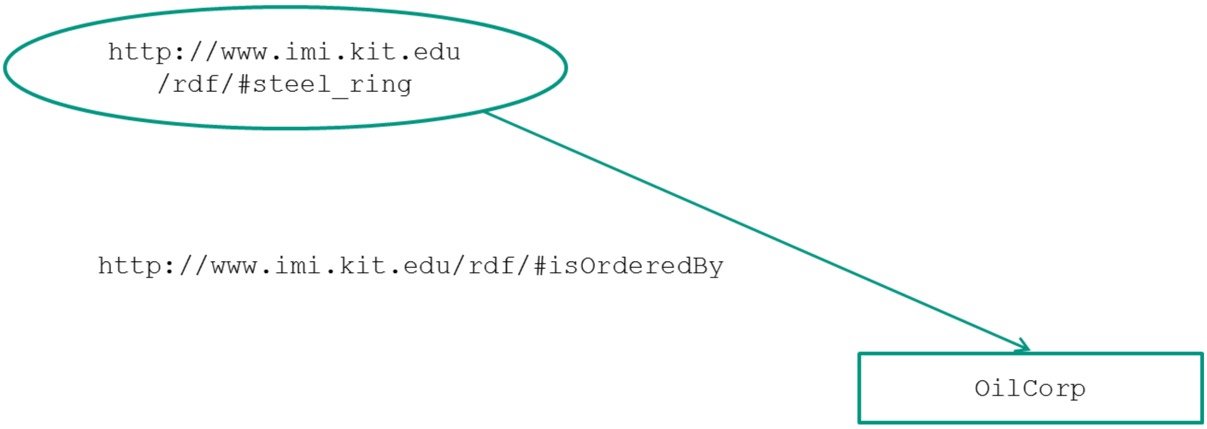


Figure A.4 Example of an RDF statement having resources as subject and predicate, and literal as object

```
@prefix imi:  <http://www.imi.kit.edu/rdf/#> .

imi:steel_ring imi:isOrderedBy “OilCorp” .
```

Figure A.5 Example of an RDF statement having resources as subject, predicate, and object expressed in Turtle

```
<?xml version="1.0"?>

<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:imi="http://www.imi.kit.edu/rdf/#">

  <rdf:Description
     rdf:about="http://www.imi.kit.edu/rdf/#cutting_machine">
     <imi:isOrderedBy>OilCorp</imi:isOrderedBy>
  </rdf:Description>
</rdf:RDF>
```

Figure A.6 Example of an RDF statement having resources as subject, predicate, and object expressed in RDF/XML

A.2 RDF Schema

RDF Schema provides a data-modelling vocabulary for RDF data and mechanism for describing groups of related resources and the relationships between them. RDFS class and property is similar to the class and property concepts of object-oriented programming. RDFS differs from object-oriented programming in that RDFS defines the properties in terms of classes through the mechanism of domain and range, whereas object-oriented programming defines classes in terms of properties [BrGu-14]. RDFS defines the type or class of RDF resources. Table A.1 lists the main RDFS constructs.

Table A.1 Main RDFS constructs

Construct	Description
`rdfs:Class`	Declares a resource as a class for other resource
`rdf:type`	Relate a resource/instance to its class
`rdfs:subClassOf`	Defines subclass relation between classes. It can build class hierarchies.
`rdfs:subPropertyOf`	Builds hierarchies between properties.
`rdfs:domain`	Declares the class of the subject in a triple
`rdfs:range`	Declares the class of the object in a triple

RDFS is able to express hierarchies of classes and properties, and also relations between entity or classes (a simple Entity-Relation-Model). However, it cannot express the complex semantics for example equivalent classes, transitivity, negations, inverse relations, etc. OWL is required to express that semantics.

A.3 Web Ontology Language (OWL)

Web Ontology Language (OWL) is semantic web language to represent rich and complex knowledge about things, group of things, and relations between things based on computational logic. The current version of OWL is developed by W3C and referred as “OWL 2” [OWL-12]. OWL is an extension of RDFS meaning and is able to describe the semantic of RDF as well. OWL uses open world assumption where the truth of a statement depends on whether it is known as true by the observer. A statement that is not known as true, it does not have to be wrong. The logical formalism of OWL is defined based on description logic.

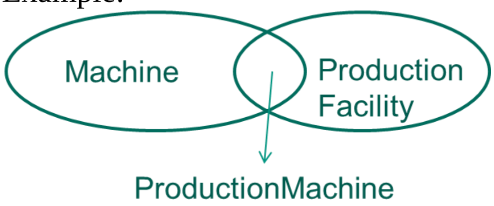
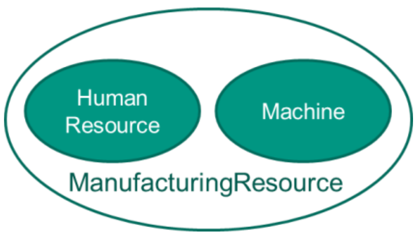
There are three main elements of OWL, i.e. classes, individuals (instances), and roles (properties). Figure A.7 depicts a part of an RDF/XML document giving an example of OWL class, individual, and role declarations. The class `HumanResource` is a subclass of `ManufacturingResource`. `M1` is an individual of the class `Machine`. The role `operates` relates the `HumanResource` as subject class (domain) and `Machine` as object class (range).

```
<owl:Class rdf:about="HumanResource"/>
      <rdfs:subClassOf rdf:resource="ManufacturingResource"/>
</owl:Class>

<owl:NamedIndividual rdf:about="M1">
      <rdf:type rdf:resource="Machine"/>
</owl:NamedIndividual>

<owl:ObjectProperty rdf:about="operates">
      <rdfs:domain rdf:resource="HumanResource"/>
      <rdfs:range rdf:resource="Machine"/>
</owl:ObjectProperty>
```

Figure A.7 Example of OWL class, individual, and role in RDF/XML

OWL is able to express complex classes and relation between classes. Table A.2 lists the OWL constructs implementing those capabilities.

Table A.2 OWL constructs expressing complex classes and relation between classes

Construct	Description
`owl:disjointWith`	There are no individuals of the class that are declared as disjoint.
`owl:equivalentClass`	Expresses the equivalence of two or more named classes. The classes have the same intentional meaning. The members of the classes do not refer to the same individuals.
`owl:intersectionOf`	Expresses logical conjunction of two or more classes. It is analog to the multiple inheritances in object-oriented programming. Example: 
`owl:unionOf`	Expresses logical disjunction of two or more classes. Example: 

OWL is also capable to describe the relations between individuals (see Table A.3).

Table A.3 OWL constructs expressing identity of individuals

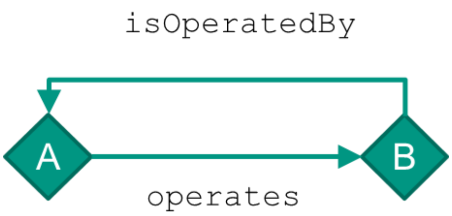
Construct	Description
`owl:sameAs`	Describes that two URI refer to the same individual.
`owl:differentFrom`	Describes that two URI refer to different individuals.
`owl:AllDifferent`	Expresses that a list of individuals refers to different individuals.

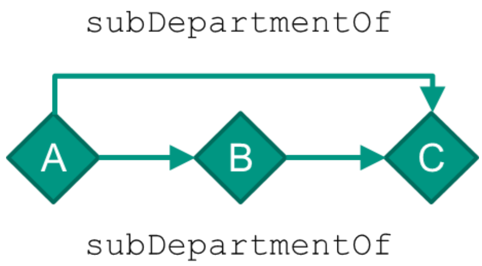
Table A.4 and Table A.5 show OWL constructs to express property value constraints and relations between properties.

Table A.4 OWL constructs expressing property value constraints

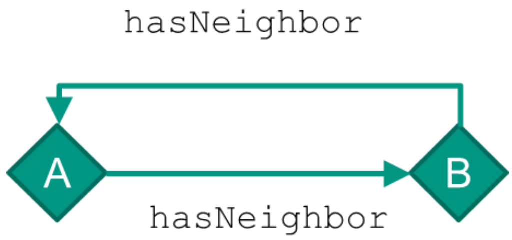
Construct	Description
`owl:allValuesFrom`	All individuals for the range of the property having the constraint `allValuesFrom` are members of a certain class.
`owl:someValuesFrom`	At least one individual for the range of the property having the constraint `someValuesFrom` are members of a certain class.
`owl:cardinality`	Expresses the cardinality of the range of a restricted property
`owl:mincardinality` `owl:maxcardinality`	Expresses the minimum and maximum cardinality of the range of a restricted property.
`owl:hasValue`	Declares that the restricted property has a certain individual for its range.

Table A.5 OWL constructs expressing relations between properties

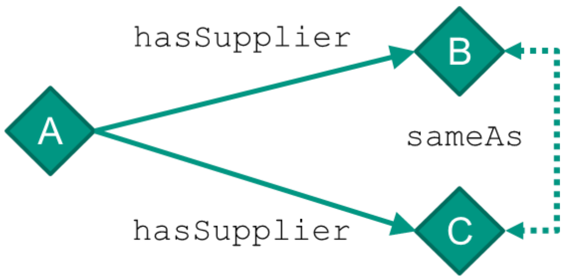
Construct	Description
`owl:inverseOf`	Declares inverse relation between properties, for example if the property `isOperatedBy` is declared as `inverseOf` operates and individual A `operates` individual B, then B `isOperatedBy` A is automatically set. 
`owl:TransitiveProperty`	Expresses the transitivity relation between properties, for example `subDepartmentOf` is declared as transitive property. If A `subDepartmentOf` B and B `subDepartmentOf` C, then A `subDepartmentOf` C.



`owl:SymmetricProperty` Expresses the symmetricity of a property. The domain and range of a symmetric property are the same. For example, if the `hasNeighbor` is declared as a symmetric property, and A `hasNeighbor` B, then B `hasNeighbor` A.



`owl:FunctionalProperty` A functional property is a property that can have only one (unique) value for each instance. For example, if `hasSupplier` is declared as a functional property, A `hasSupplier` B, and A `hasSupplier` C, then B and C can be concluded as the same individual.



Appendix B Data Mining Algorithms

B.1 Association

Association is a data mining technique to find correlations among different attributes in a dataset. Market basket analysis is a common application of the algorithm that creates association rules. The algorithm evaluates the frequency of each attribute in the dataset. It then generates rules in form $A \rightarrow B$ with a certain probability, based on the common occurrence of items. This means the probability that in the case of the occurrence of *A* (antecedent), product *B* (consequent) is also acquired. Association rules have two important parameters. These are "Support" and "Confidence" [Pete-05]. The support of an association rule refers to how frequently a specific set of items occurs in the observed dataset. Confidence or accuracy is the number of instances that is correctly predicted, expressed as a proportion of all instances to which it applies [WiFH-05].

$$Support\ (A \rightarrow B)\ =\ \frac{count\ (A \rightarrow B)}{number\ of\ transactions} \tag{B.1}$$

$$Confidence(A \rightarrow B) = \frac{count\ (A \rightarrow B)}{count\ (A)} \tag{B.2}$$

The equation (B.1) indicates the support of the rule and the equation (B.2) calculates the confidence of the rule. The methods used in association analysis can be found mostly in E-Commerce. For example in online shopping website, it is common that customers individually are directed to interesting offers. It can be identified from a set of customer orders, who buy the product *A* will also buy the product *B* with probability *p*.

The apriori algorithm is a classic algorithm that generates association rules in a reasonable timescale, at least for relatively small databases. The apriori algorithm is an algorithm to find a set of items (itemsets) that occur together with maximum support.

B.2 Neural Network

Neural network is a technique of data mining that inspired by the mechanism of the human brain. The technique uses a human neuron model to perform learning process based on predetermined input ("Input Units"). The learning process is divided to supervised and unsupervised, depending on whether the right results (output) can be predetermined. Figure B.1 shows the structure of neural network with distribution of neurons in different layers [ReWe-11]. Input neurons are neurons that accept the input signals or impulses from outside world, whereas output neurons give output signals or impulses to the outside world. Hidden neurons are internal representations of the outside world which are located between input and output units.

A neural network learns from the data using algorithms that change the structure of the network in different ways. The structure of a neural network consists of the following basic elements [Krie-12]:

- *Neurons* are the node of in the network and represent the simple processing units
- *Connections* that connect between neurons
- *Weights*, the strength of the connections
- *Processing functions* that are depicted in Figure B.2 and comprise the following components:
    - *Propagation function* that transforms the outputs of other neurons into the network input in order to be processed by an activation function. It converts vector inputs to scalar network inputs. It is typically implemented using weighted sum.
    - *Activation function* which transforms network input and previous activation state into new activation state based on threshold values. The reaction if the neuron to the input value is determined by the activation state. Therefore, the activation function gives the switching status of the neuron.
    - *Output function* that calculates the output values of the neurons from their activation states

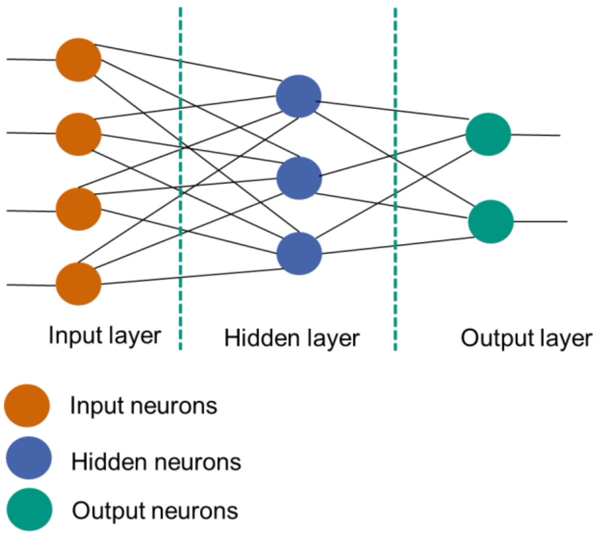


Figure B.1 Structure of a neural network [ReWe-11]

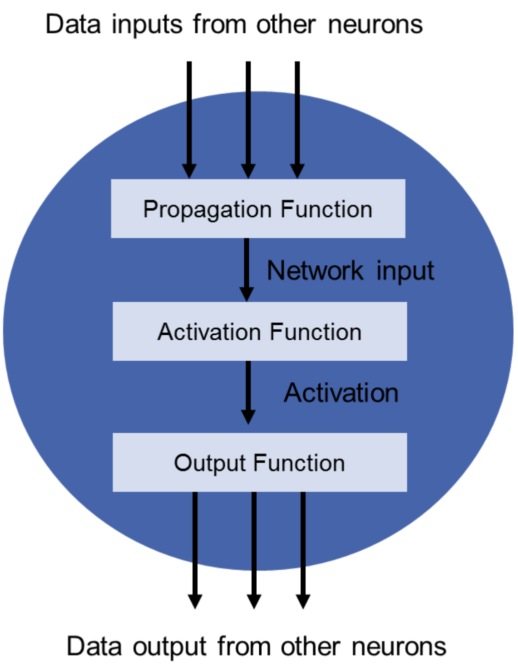


Figure B.2 Processing functions in a neuron [Krie-12]

The learning process in a neural network is performed by changing the network structure. It can be carried out through the following methods:

- Changing the connection weight
- Adapting the transfer functions
- Changing the threshold values of activation functions
- Adding or removing the connections or neurons

A neural network model is relatively more complex and quite more difficult to understand compared to a decision tree model. Neural network also has better performance in handling large datasets. In smaller datasets, it has slightly better accuracy. However, the neural network cannot handle discrete or categorical values [JaZP-99] [WiFH-05].

Appendix C Estimation of Values Related to Lighting

Table C.1 Estimated power of lighting depending on desired average illumination and installation height [Pist-09]

Average illumination $\widetilde{E_m}$ [lx]	**Electrical power of lighting $j_{\widetilde{E_m},h}$ [W/m^2] with installation height of**		
	2 m	**3 m**	**4 m**
1000	50	60	64
750	38	45	48
500	25	30	32
300	15	17	19
200	10	11	13
100	5	6	6
50	3	3	4

Table C.2 Adjustment factors of different lamp types [Pist-09]

Lamp type	**Adjustment factor**
Filament bulb	6.0
Halogen bulb	5.0
Fluorescent lamp EVG	1.0
Fluorescent lamp KVG	1.24
Compact fluorescent lamp with integrated EVG	1.6
Compact fluorescent lamp with external KVG	1.2
LED	1.1
LED retrofit	1.5