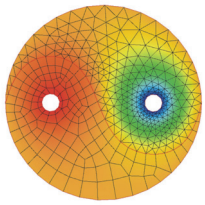
Publishing House
AKAPIT

COMPUTER METHODS IN MATERIALS SCIENCE

Informatyka w Technologii Materiałów

Vol. 13, 2013, No. 2



A USER-INSPIRED KNOWLEDGE SYSTEM FOR THE NEEDS OF METAL PROCESSING INDUSTRY

STANISŁAWA KLUSKA-NAWARECKA[1]*, ZENON PIROWSKI[1], ZORA JANČÍKOVÁ[2], MILAN VROŽINA[2], JIŘÍ DAVID[2], KRZYSZTOF REGULSKI[3], DOROTA WILK-KOŁODZIEJCZYK[4]

[1] Foundry Research Institute, Zakopiańska 73, Cracow, Poland
[2] Department of Automation and Computer Application of Metallurgy, VŠB – Technical University Ostrava 17. listopadu 15, Ostrava-Poruba, 708 33 CZECH
[3] AGH University of Science and Technology, Mickiewicza 30, Cracow, Poland
[4] Andrzej Frycz Modrzewski University, Herlinga-Grudzińskiego ,1 Cracow, Poland
**Corresponding author: nawar@iod.krakow.pl*

Abstract

This article describes the works related with the development of an information platform to render accessible the knowledge on casting technologies. The initial part presents the results of a survey on the preferences of potential users of the platform regarding areas of the used knowledge and functionalities provided by the system. The second part contains a presentation of selected modules of the knowledge with attention focussed on their functionalities targeted at the user needs. The guide facilitating the use of the platform is a "virtual handbook". The System is used as a coupling link in the diagnosis of defects in castings, while ontological module serves the purpose of knowledge integration when different sources of knowledge are used.

Key words: artificial intelligence, knowledge, distributed and heterogeneous sources, technology platforms, databases integration, ontologies

1. INTRODUCTION

Currently, the software market offers “knowledge" systems for computer-aided design and simulation processes (CAD / CAM) and also knowledge management tools and industrial information of the ERP / MRPII type. On the other hand, still very poorly developed area remains that of the technological decision support tools, i.e. expert systems, technology platforms to share domain knowledge, tools for integration of knowledge from distributed and heterogeneous sources.

In recent years, the interest in expert systems supporting diagnosis and technological decision-making process has been subject to some fluctuations . The observed disappointment in this class of systems was due to some difficulties related with collection and application of a sufficient number of rules in knowledge bases. Currently, the development of algorithms for automated knowledge acquisition has aroused a new interest of science centres in inference systems what has been described (Adrian et al., 2012; Kluska-Nawarecka et al., 2011a; Kluska-Nawarecka et al., 2011c; Mrzygłód et al., 2007; Zygmunt et al., 2012). Studies are carried out on possibilities to apply modern knowledge engineering formalisms as it has been written in (Jančíková et al., 2011; Kluska-Nawarecka et al., 2009; Spicka et al., 2010; Švec et al., 2010), including fuzzy logic, rough sets, decision tables and the

 ISSN 1641-8581

use of multimedia techniques to render this knowledge accessible to users.

With the current rush of knowledge and data, unavoidable is the research on the methods of knowledge acquisition and integration, including ontologies enabling modelling of domain knowledge, and ultimately the creation of semantically structured systems.

The article outlines the future plans and gives selected results of work aimed at building a system platform with the task of creating and sharing the knowledge from the area of foundry technologies.

2. ANALYSIS OF USER NEEDS

When the work was started on the development of a concept of the structure, and on the substantive content of the system with determination of the functionality of each of its modules, it was considered necessary to refer to the needs of potential users of the system.

Therefore, the main objective of the first stage of the work was considered to be an interview with the industry and scientific communities in Poland and abroad dealing in some way with the casting practice, to determine the need for different types of functionality of the future system rendering the knowledge accessible. The interview was conducted in the form of questionnaires and discussions carried out with the representatives of industry and research centres. Surveys covered a specific range of the system utility, namely the area rendering available the knowledge components.

were selected from the circles of the scientists cooperating with plants processing different types of metals.

Respondents were asked to indicate the types of knowledge they believe are most commonly demanded by the manufacturing plants. In the questionnaire they were given the following options:

- *literature knowledge about processes, descriptions of processes and applied technologies, handbooks including descriptions of the possible types of treatment*
- *norms, certifications, standards, Polish Standards, ISO, etc.*
- *visuals: pictures, computer simulations,*
- *photographs of castings and defects, and microstructures*
- *reports and studies of own production, document templates, balance sheets, ready compilations*
- *specification of requirements and properties, and chemical compositions of materials.*
- *tables, databases*
- *expert knowledge in case of defects - irregularities in the process, request for expert opinion*
- *Branch stats, volume of production etc.*
- *market analysis*
- *marketing data, studies, foresight*
- *industry statistics, statistical yearbooks, data on production volume in a given sector, data from Chief Statistical Office (CSO)*

The results of survey are presented in the form of a diagram in figure 1.

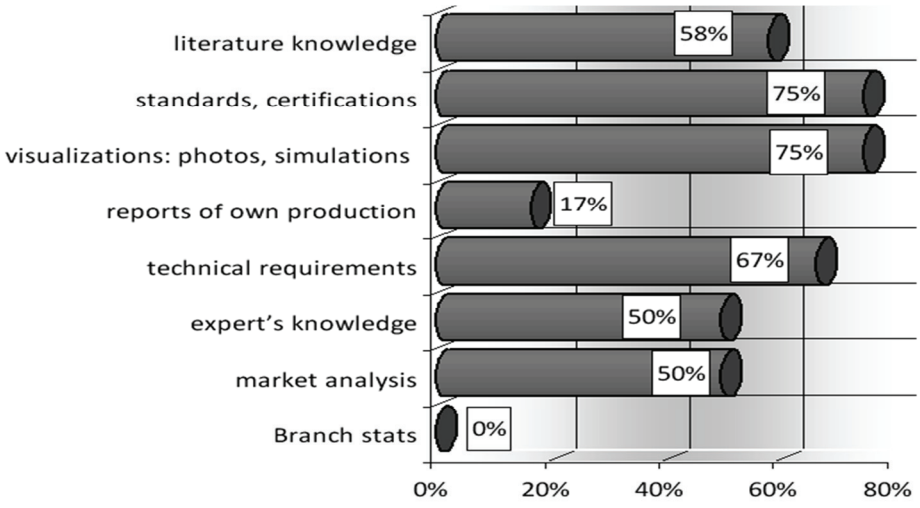


Fig. 1. *User preferences on the types of shared knowledge.*

The selected companies represented the large, small and medium-sized enterprises (joint stock companies, limited liability companies); experts

It is clear that professionals reach most often for multimedia resources in the form of photographs and simulations, also for diagrams and visuals in the

form of charts and diagrams. As important the following ones were also identified:

- norms, standards and certifications,
- literature knowledge and handbooks,
- specification of requirements and properties, chemical composition of materials, tables, databases,
- market analysis and marketing studies,
- expert knowledge when it is necessary to have an expertise performed.

Of minor importance was considered industry statistics, and reports and studies of own production. Probably the reason is effective circulation of such information and needs satisfied by the already existing tools and management systems.

In order to clarify the need for different types of knowledge, the respondents were requested to assess the individual potential functionalities of the future information system operating in their companies. The proposed list of functional features is presented in table 1.

Table 1. *List of potential functionalities of an information system.*

Potential functionalities of information system
virtual handbook descriptions of processes and applied technologies, handbooks including description of possible treatments, access to publications
visuals: pictures, simulations photographs of castings and defects, microstructures
electronic standards, certifications Polish standards, ISO standards, etc.
databases, catalogues specifications of requirements and properties, chemical compositions of materials, catalogues of materials
e-learning virtual training in advanced manufacturing technologies, interactive courses in casting techniques
expert systems discovering the causes of defects, detection of process irregularities
tools for classification determination of defect types and class/grade of material from which the product should be made, etc.
tools to make reports based on production data, statements and reports on e.g. production volume, consumption of materials, costs, severity of defects

The list of potential functionalities of the system was based on the conducted research and currently available capabilities of a system developed as a result of cooperation between the team of workers from the Foundry Research Institute and Knowledge Engineering Team from the Department of Applied Informatics and Modelling at the Faculty of Metals Engineering and Industrial Computer Science, University of Science and Technology, Cracow, Poland.

Definitely the highest rated was the proposed "virtual handbok" - a platform to share the collections of documents with descriptions of processes and technologies, and handbooks and publications in electronic form.

Also highly rated were the visuals: photographs, simulations, pictures, photographs of castings and defects, and microstructures, as well as databases, catalogues, the specifications of requirements and properties, chemical compositions of materials, catalogues of materials – all of them reflecting the most common needs. The responses obtained allowed establishing the following ranking of other functionalities:

1. Databases, catalogues;
2. Expert systems;
3.4 Classification tools, reporting tools;
5. Electronic standards, certifications;
6. E-learning.

As a general conclusion from the survey and discussions held, it can be stated that such knowledge sharing is needed that, while giving the user an easy to handle interface, will also ensure a constant supply of current information and knowledge from the area of the casting practice, not only derived from the literature, but also from all other sources such as databases, or knowledge obtained algorithmically from the process data. To achieve this, the above mentioned sources will have to be integrated and made ready for processing (e.g. indexed for convenient search).

At the same time, the need arises to design a knowledge base in such a way as to make it interactive, to enable user to get through a dialogue with the system just this knowledge that is necessary for solving of a problem.

3. KNOWLEDGE-SHARING PLATFORM

When the concept of a knowledge-sharing platform on casting technologies has been created, it was assumed that the platform should include all major solutions developed in the course of previous works on the computer-aided manufacturing processes. At the same time it should be enriched with new modules and functionalities, targeted at meeting the user's preferences as regards the application of new trends and opportunities that arise from the

development of methods and technologies based on computer science what has been said in (David et al., 2011; Olejarczyk et al., 2010).

Consequently, the platform has a multi-module structure, where individual modules can operate independently, and the results of their actions are subjected (if necessary) to the process of integration.

The degree and manner of this integration depends mainly on the scenario of actions taken by the user (when using the system in an interactive mode).

Among the modules already existing and available on the Internet, one can mention the Infocast system (including databases on publications, standards, and catalogues) and the Castexpert system designed to serve as a tool for the diagnosis of casting defects assisted with knowledge presented in the form of multimedia.

Below are outlined the results of the implementation of additional modules specific for the operation of the whole platform and which received most attention from the users.

3.1. Virtual handbook

The virtual handbook, whose schematic diagram is shown in figure 2, is a kind of clause, linkining together the functionalities of other modules of the platform. Using this tool, the user gets a general idea about the type of knowledge provided by the platform, and as a result can find out which variant in the scenario will lead him to a solution of the problem.

A casual scenario of the use of the handbook is presented as an example in table 2.

3.2. System for decision support and classification of defects in castings

The RoughCast system allows the use of approximate logic to enable a classification of casting defects according to international standards: Polish, Czech and French. The databases can be expanded in the future to include other standards.

The approximate logic is a tool to model the uncertainty arising from incomplete knowledge resulting from the granularity of information. The main application of approximate logic is in classification,

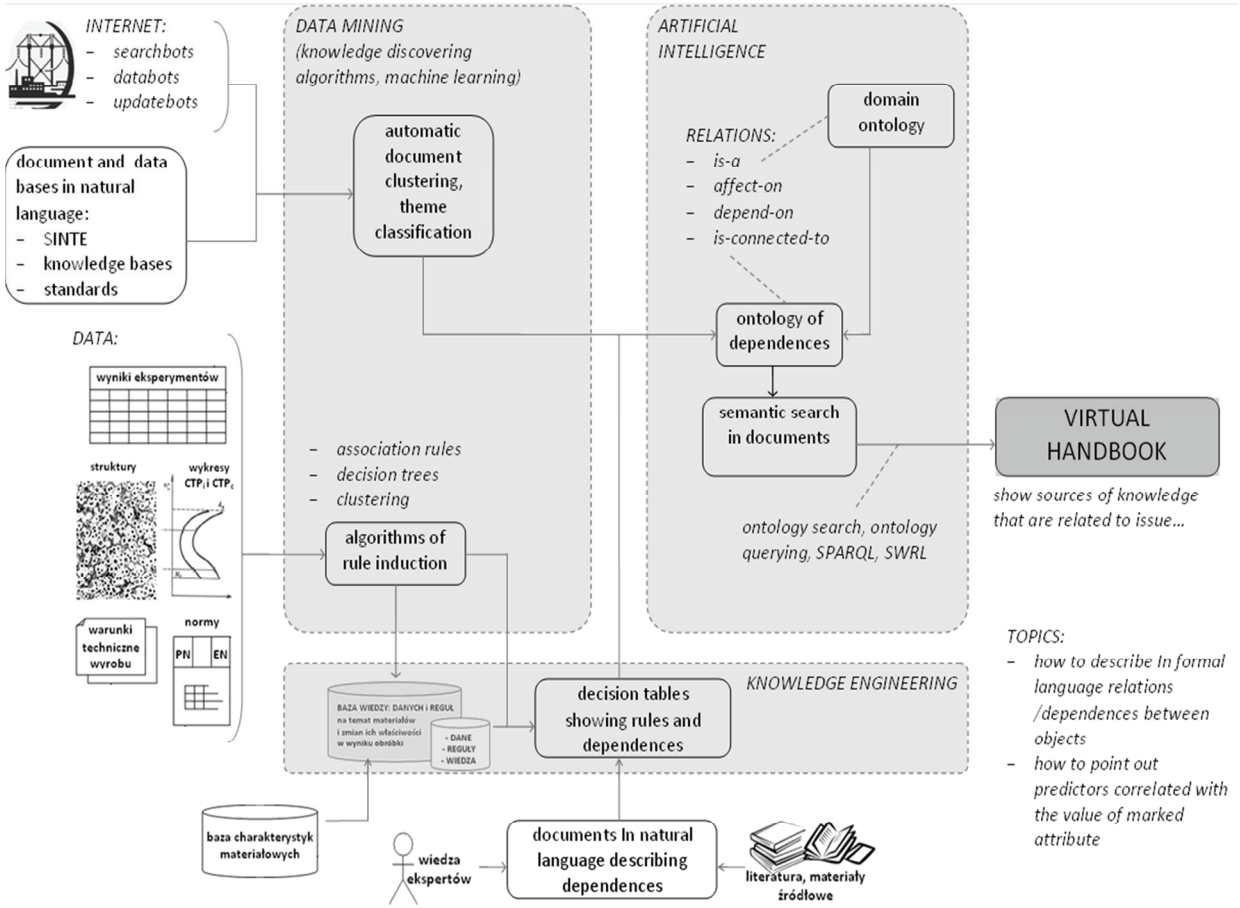


Fig. 2. *Specification of functional requirements for the knowledge module „Virtual Handbook".*

Table 2. *An example of the use of virtual handbook.*

Name:	Virtual handbook
Actors	End User, Expert, Knowledge Engineer
Sharehold-ers/Stakeholders:	– Artificial intelligence – Data – Knowledge engineering – Sources – Data Mining – Internet
Short description	Preparing a specialised Virtual Handbook.
Preliminary conditions	The user must have access to a computer and specified topic of the handbook
Final conditions	Handbook ready to display. Database updated and saved .
The main flow of events	1. The user opens the Virtual Handbook interface (on-line) 2. The user writes in the subject 3. The system analyses the subject a) Finds data on the Internet and in the documents and databases in natural language b) Collects statistical data c) Searches alternative sources of knowledge (literature, source materials, expert knowledge, etc.) 4. Cataloguing of data: a) Taken from the Internet by means of the Data Mining methods b) Statistical data using rule induction algorithms c) Data from sub-item 4.b and alternative sources of knowledge 3.c using decision tables 5. Saving in XML files 6. Algorithms of artificial intelligence are preparing relevant data for display in the handbook (semantic analysis with the use of ontologies) 7. Displaying the appropriate page of the handbook
Special requirements	Device with access to Internet.

Table 3. *Fragment of decision table for defects in steel castings*

Attribute symbol	a1	a2	a3	a4	a6	a7	a8	a9	a12
Symbol of object in array	**Damage symbol**	**Damage name**	**Standard**	**damage type**	**distribution**	**location**	**occurrence**	**damage shape**	**technological operation**
x1	341	COLD LAPS	CZ	wrinkles, scratch, erosion scab	local	insert wall, chaplet, surface	numerous	narrow, rounded edges	casting design, pouring, cooling
x2	W207	COLD LAP	PL	fissure, scratch	local	surface	single	narrow, rounded edges	gating system design, pouring
x3	W407	COLD SHOTS	PL	metal beads		interior		spherical	gating system design, pouring
x4	C311	COLD LAP, COLD SHOTS	FR	discontinuity, fissure	widespread	surface, subsurface area	numerous	rounded edges, narrow	feeding system, design, pouring
x5	C331	COLD LAP NEAR CORE OR OTHER METALLIC PART	FR	discontinuity	local	near inserts	data not available	curved walls	pouring, solidification

as with this logic it is possible to build models of approximation for a family of the sets of elements for which the membership in sets is determined by the attributes. The conducted research allowed developing a methodology for the creation of decision tables to serve the knowledge of casting defects.

Using this methodology, a decision table was developed for the selected defects in steel castings. A fragment of the array is presented in table 3.

Based on rough set theory, elementary sets can be determined in the array. Sets determined in this way represent the division of the universe in terms of the indiscernibility relations for the attribute *distribution*.

The most important step is to determine the upper and lower approximations in the form of a pair of precise sets. Abstract class is the smallest unit in the calculation of rough sets. Depending on the que-

ry, the upper and lower approximation is calculated by summing up the respective elementary sets.

The system operates on decision-making tables - this structure of the data allows the use of an inference engine based on approximate logic. The system maintains a dialogue with the user asking questions about the successive attributes. The user can select the answer (the required attribute) in an intuitive manner. Owing to this method of the formulation of queries, there is no need for the user to know the syntax and semantics of queries in the approximate logic. However, to make this dialogue possible without the need to build a query by the user in the language of logic, the system was equipped with a query interpreter in a semantics much more narrow than the original Pawlak semantics. It has been assumed that the most convenient way to build queries in the case of casting defects is by selection from a list of attributes required for a given object (defect). The user chooses which characteristics (attributes) the selected defect has. This approach is consistent with the situation occurring every day when the user has to deal with specific cases of defects, and not with the hypothetical tuples. Thus set up queries are limited to conjunctions of attributes, and therefore the query interpreter has been equipped with only this one logical operator. The knowledge base created for the needs of a RoughCast system is the embodiment of an information system in the form of decision-making tables. It contains tabulated knowledge on the characteristics of defects in steel castings taken from Polish Standards, Czech studies, French directory of casting defects, and German textbook of defects.

Using this formalism it becomes possible to solve a number of difficulties arising from the foundry knowledge granularity in the form of indistinguishable descriptions created with attributes and inconsistent classifications from various sources (as in the case of standards for steel castings).

3.3. Cluster Analisys

General design requirements apply to the data mining system, whose main objective is to classify documents (articles) by a thematic classification. The implementation of task module was developed based on the full-text clustering method, supported by the use of a thesaurus. The process of full-text cluster analysis was used to create the task category (conceptual clustering) as a method of unsupervised learning. The aim was to design a system operating efficiently, which, based on the documents provided (in the correct format and in accordance with the established standards and norms), will carry out the task of clustering the documents by thematic classification based on data mining methods - cluster analysis. This module is fully compatible with the directly cooperating document repository systems and databases, and should be, to the greatest extent possible, susceptible to subsequent modifications or development.

In the task of conceptual grouping, the training set Ω is a collection of articles provided with a compatible system in the form of a knowledge base, while the task of the aggregation analysis is to split these objects into categories (aggregations) and construct a description of each category (aggregation centroids), enabling the efficient classification of new articles. As a result of this process, each article is included into one of the resulting aggregations. Each of the resulting aggregations has its centroid, which represents a concept associated with this aggregation.

The classification of text documents is a very complex problem. The main reason for the difficulties is the semantic complexity of natural language. The difficulties associated with the classification of documents written in natural language are related to, among others, polysemy, i.e. terms having many different meanings. For example, the term 'table' may refer to either a piece of furniture, or it may also mean a set of specifically arranged data, numbers, etc. The dimension of feature space in the document classification tasks, related to the number of possible words in a natural language (usually the order of tens of thousands of words), is also a difficulty. On the other hand, the representation of documents using the selected (small) number of words will reduce the quality of classification.

Therefore, the first step of the task of conceptual clustering is to reduce the articles from the knowledge base to the basic grammatical forms. The process of cluster analysis will compare the sets of articles (with each other) in search for common words excluding irrelevant words (or, also, but, etc.). Articles will be grouped in clusters on the basis of the probability of adjustment determined by the number of occurrences of the common words, and on this basis the aggregation centroid will be created, as presented in figure 3.

To improve the quality of cluster formation and classification of new articles, these processes will be supported by a thesaurus (a structured set of key-

words), which will also eliminate the problem when compared to the sets that do not have the same keywords in the text, and yet belong to the same category.

Constraints can be formulated in OCL for the class Aggregation_Centroid as exemplified is given below.

```
context Aggrega-
tion_Centroid::Generate_Aggregation_Centroid() :
Boolean
pre: self.Aggregation_Centroid_Path = ''
post self.Aggregation_Centroid_Path<>''
context Aggregation_Centroid::get_ID : Inteager
post rezult = self.ID_Centroid
context Aggregation_Centroid::get_PATH : String
post: rezult self.Aggregation_Centroid_Path
```

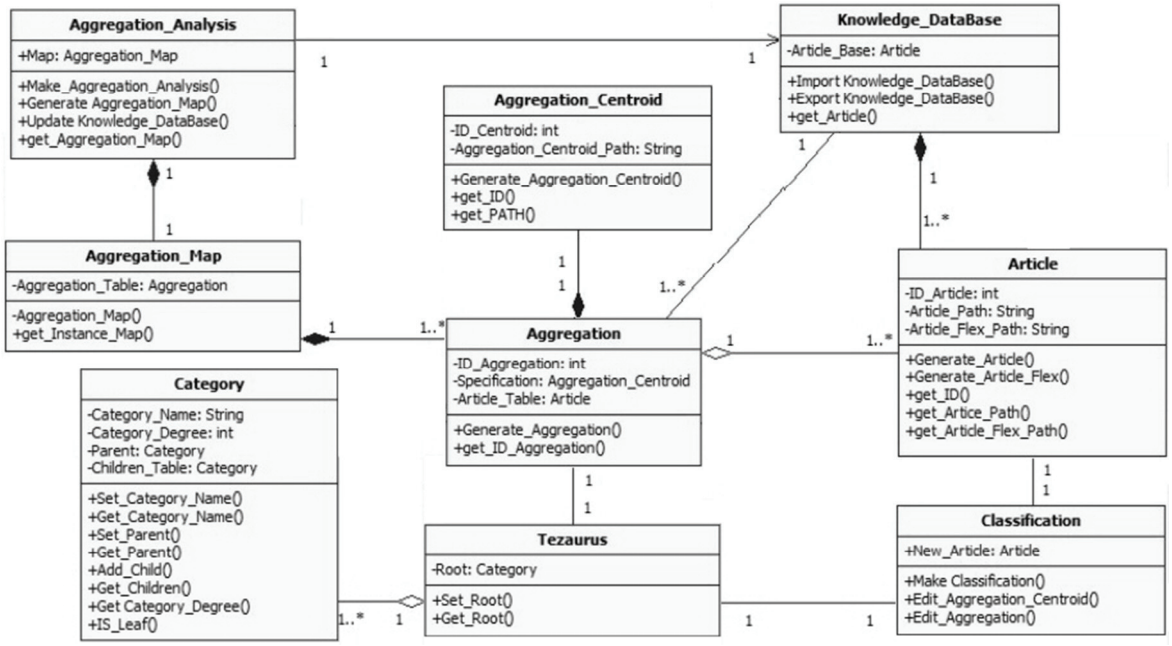


Fig. 3. *Diagram of classes in the module of document aggregation analysis.*

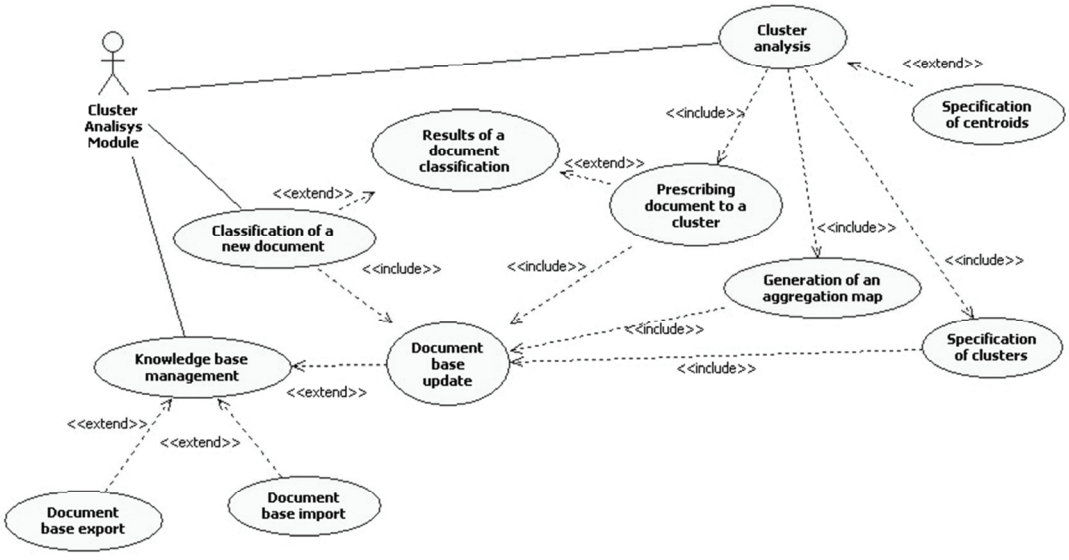


Fig. 4. *Use Case Diagram.*

```
context Aggregation_Centroid inv
self.ID_Centroid >=1 and
self.Aggregation_Centroid_Path<>''
and Aggregation_Centroid.allinstance ->forall
(p1,p2) p1<>p2 implies p1.ID_Centroid <>
p2.ID_Centroid
and Aggregation_Centroid.allinstance ->forall
(p1,p2) p1<>p2 implies
p1.Aggregation_Centroid_Path<>
p2.Aggregation_Centroid_Path
```

The use case diagram is shown in figure 4.

In table 4 there is a description of one of possible utilizations discussed algorithm.

Table 4. *Description of the case of use.*

Name:	Automatic clustering of documents
Actors:	System
(Stakeholders/Interests)	The system manages the service module of the operations of the database. It performs import operations on documents from a knowledge base and export operations on the resulting database to ontology classes. The system also manages the module of document clustering by the method of Data Mining.
Short description:	Automatic clustering of documents in terms of thematic fit.
Pre-conditions:	Obtaining edited documents and information in the form of a knowledge base
Post-conditions:	Thematically grouped forwarded database of articles.
Main flow of events:	1. Receiving knowledge base 2. System performs cluster analysis for the resulting knowledge base. 3. System creates a new database of articles grouped thematically based on cluster analysis carried out. 4. System sends the created database.
Alternative flow of events:	1. When new article is downloaded, the cluster analysis classifies it into one of the topics. 2. When there is a limit to the number of articles classified to various topics, and all these articles are characterised by a high degree of fit with each other, then a new topic will be created to which these articles will be assigned.
Special requirements:	1. For the main flow of events to occur, the knowledge base must be delivered. 2. For an alternative flow of events to occur, the artificial intelligence system must initiate the delivery of a new article.

4. FINAL REMARKS

The article describes solutions developed for selected modules of a platform for sharing the knowledge of foundry technologies in the context of the preferences expressed by users. It seems that studies carried out to create this context make an interesting contribution to the contents of this article, since the results of such surveys are not often disclosed in the presentations of different expert and decision-making systems.

Selecting modules of the platform described in the article, it was attempted, on the one hand, to show their diversity and, on the other, to expose the way by which they will be adapted to the declared user needs. It has been the intention of the creators and promoters of the platform to offer a system that will have an evolving nature, and will be gradually enriched with new modules, according to the emerging needs. Currently work is underway on the implementation of modules of the automatic acquisition of knowledge from the Internet and on the analysis and classification of text documents which is presented (Kluska-Nawarecka et al., 2011b).

Scientific work financed from funds for the scientific research as an international project. Decision No. 820/N-Czechy/2010/0.

REFERENCES

Adrian, A., Mrzygłód, B., Durak, J., 2012, Model strukturyzacji wiedzy dla systemu wspomagania decyzji, Hutnik - Wiadomości Hutnicze, 1, 67-70 (in Polish).

David, J., Vrožina, M., Jančíková, Z., 2011, Determination of crystallizer service life on continuous steel casting by means of the knowledge system, *Transactions On Circuits And Systems*, 10, 359-369.

Jančíková, Z., Ružiak, I., Kopal, I., Jonšta, P., 2011, The Influence of Rubber Blend Aging and Sample on Heat Transport Phenomena, *Defect and Diffusion Forum*, 312-315, 183-186.

Kluska-Nawarecka, S., Wilk-Kołodziejczyk, D., Dobrowolski, G., Nawarecki, E., 2009, Structuralization of knowledge about casting defects diagnosis based on rough sets theory, *Computer Methods in Materials Science*, 9, 341-346.

Kluska-Nawarecka, S., Wilk-Kołodziejczyk, D., Regulski, K., 2011a, Practical aspects of knowledge integration using attribute tables generated from relational databases, Semantic Methods for Knowledge Management and Communication, eds, *Katarzyniak, R., Chiu, T.F., Hong, C.F., Nguyen, N.T.,* Springer, Berlin, Heidelberg, 13-22.

Kluska-Nawarecka, S., Wilk-Kołodziejczyk, D., Dziaduś-Rudnicka, J., Smolarek-Grzyb, A., 2011b, Acquisition of technology knowledge from online information sources, *Archives of Foundry Engineering*, 11, 107-112.

Kluska-Nawarecka, S., Wilk-Kołodziejczyk, D., Regulski, K., Dobrowolski, G., 2011c, Rough Sets Applied to the RoughCast System for Steel Castings, Inteligent Information and Database Systems, *Proc. Conf. Third International Conference, ACIIDS 2011*, eds, *Nguyen, N.T., Kim, C.G., Janiak*, A., Daegu, Korea, 52-61.

Mrzygłód, B., Adrian, A., Kluska-Nawarecka, S., Marcjan, R., 2007, Application of Bayesian network in the diagnosis of hot-dip galvanising process, *Computer Methods in Materials Science*, 7, 317-323.

Olejarczyk, I., Adrian, A., Adrian, H., Mrzygłód, B., 2010, Algorithm for controling of quench hardening process of constructional steels, *Archives of Metallurgy and Materials*, 55, 171-179.

Spicka, I., Heger, M., Franz, J., 2010, The mathematical-physical models and the neural network exploitation for time prediction of cooling down low range specimen, *Archives of Metallurgy and Materials*, 55, 921-926.

Švec, P., Jančíková, Z., Melecký, J., Koštial, P., 2010, Implementation of neural networks for prediction of chemical composition of refining slag, *Proc. Conf. Metal 2010*, *International Conference on Metallurgy and Materials*, Tanger, 155-159.

Zygmunt, A., Koźlak, J., Nawarecki, E., 2012, Analiza otwartych zrodel internetowych z zastosowaniem metodologii sieci spolecznych, Bialy wywiad: otwarte zrodla informacji – wokol teorii i praktyki, eds: W. Filipkowski, W. Mądrzejowski, C.H. Beck, Warszawa, 197-221 (in Polish).

SYSTEM UDOSTĘPNIANIA WIEDZY INSPIROWANY POTRZEBAMI UŻYTKOWNIKA Z ZAKRESU PRZEMYSŁU PRZETWÓRSTWA METALI

Streszczenie

Artykuł dotyczy prac związanych z realizacją platformy informatycznej, służącej od udostępnienia wiedzy z zakresu technologii odlewniczych. W części początkowej przedstawiono rezultaty sondażu dotyczącego preferencji potencjalnych użytkowników platformy odnośnie obszarów wykorzystywanej wiedzy oraz funkcjonalności udostępnianych przez system. Część druga zawiera prezentację wybranych modułów wiedzy ze zwróceniem uwagi na ich funkcjonalności ukierunkowane na potrzeby użytkowników. Rolę przewodnika ułatwiającego korzystanie z platformy pełni „wirtualny poradnik", system Rought Cast służy do sprzęgania diagnostyki wad odlewów, zaś moduł ontologiczny, pozwala na integrację wiedzy pochodzącej z różnych źródeł.

Received: October 16, 2012
Received in a revised form: November 28, 2012
Accepted: December 7, 2012