

Technology of Intelligent Service for Energy Technology Forecasting

Alex N. Kopaygorodsky ^{1,2}

¹ Melentiev Energy Systems Institute of Siberian Branch of the Russian Academy of Sciences,
Lermontov str., 130
Irkutsk, Russia
E-mail: kopaygorodsky@isem.irk.ru

² Irkutsk National Research Technical University,
Lermontov str., 83
Irkutsk, Russia
E-mail: kopaygorodsky@istu.edu

Abstract

This article reports the approach and software tools for decision-making support in forecasting the energy infrastructure development. The author considers the problem of searching for information from various open sources, technology of information searching, knowledge detection and classification. The author describes architecture of Intelligent information system. For experts this classification and integrated warehouse simplify search for the knowledge required.

Keywords: Intelligent service, knowledge management, expert decision support, energy infrastructure development forecasting.

1. Introduction

Future-oriented research of industry development is becoming particularly relevant because of complicated economic conditions worldwide. This is necessitated by the task (a) to predict potential results of innovative projects with high reliability and (b) to save money. Because the Energy industry is basic production of final products largely depends on its development. In addition, in the recent years, alternative energy technologies have been actively developed, for instance those related to ecology and energy conservation, e.g. Green Energy.

The technological forecasting of the RF energy sector up to year 2035 considers three different world energy development scenarios [1]. Fidelity of each scenario cannot be properly estimated. This leads to the high uncertainty in innovative development strategic planning because of high investment risks in the energy sector. Improving evaluation of technological development and rolling forecasts will let to focus on the amplification of certain technologies with better effect in future. Thus, providing researchers with topical information and knowledge is an important part of development of expert assessment of new technological solutions in energy infrastructure.

2. State of Art

Since 2003, the international workgroup of leading scientists from the USA, Europe and East Asia coordinates research on Future Oriented Technology Analysis [2].

Researchers are carried out on such closely related areas as Technology Monitoring, Watch and Alerts, Technical and Competitive Intelligence, Technological Forecasting, Strategic Technology Assessment, development Technology Roadmapping, Technology Foresight and other.

Forecasting as a method of research is used in the domain of Energy Infrastructure to study the development and the functioning. In [3], authors were considering the technological prospects of various directions of decisions of the problem of resource restrictions of the development of wind and solar energy.

Recently, more and more often analysis results of Big Data and Linked Open Data as information source has been used. For example, the Big Data analysis was used for Electricity Load Forecasting [4]. This model of Electricity Load was designed to make predictions for time series with specific properties (strong seasonal dependence and concept drift). The prediction results

via Big Data analysis described the behavior of the real Energy System in the near future very well.

Text Mining is the process of deriving information from text. Some applications of this technique are Social Media Monitoring and Scientific Discovery. Authors of [5] propose a conceptual approach to the research into customer satisfaction based on a detailed analysis of consumer reviews written in natural languages.

Artificial Intelligence techniques successfully are used for forecasting the conduct of separate energy technologies. In [6], authors use patent indicators to predict the technological advances in Hydrogen Storage Materials (HSM). Patent analysis was been carried out using bibliometrics and Text Mining approaches in order to forecast the future trend of development. Authors were evaluating the technological life cycle stage, HSM class prominence and the role of different countries in HSM patenting.

The team employed at the Melentiev Energy Systems Institute, Siberian Branch, Russian Academy of Sciences (ESI SB RAS) has been developing the intelligent information resources and support software tools for decision-making and collective expert activity for innovative development forecasting of energy infrastructure [7, 8]. The latter is truly important for work at the ESI. The experts from various fields such as Energy, Mathematics, Economy, Ecology and Informatics are involved in research.

3. Energy Infrastructure Development Forecasting

The ESI has its own data sources and previous results, but it is not enough to fulfill high-quality forecasts. Researchers should attract additional knowledge and data to improve the quality and accuracy of forecasts.

Researchers use two classes of models to forecast: (1) Qualitative models show the relationship and influence of various factors; they are employed to determine some trends; (2) Quantitative model. The qualitative analysis as the first stage of forecasting development makes it possible to quickly and cheaply solve the problem. In some cases, quantitative modeling can be used as a more expensive approach, but giving better results in addition to the qualitative solution obtained. In any case, researchers need to make relevant changes in the model to obtain adequate results, when researching forecast of energy development.

Knowledge for energy infrastructure development forecasting arrives from various sources. All knowledge sources can be conditionally divided into two classes:

(1) thematic databases and (2) free format sources. The thematic databases contain knowledge in structured or semistructured form from various areas. Examples of such information systems are government information systems (CITIS, RFBR, FIPS and others) and various commercial systems (SCOPUS, Web of Science, RSCI, Science Index and others). Free format sources are pure Internet resources. Each of them may be contain any relevant and irrelevant information in any form. Figure 1 exhibits ways to enrich knowledge warehouse for energy infrastructure development forecasting.

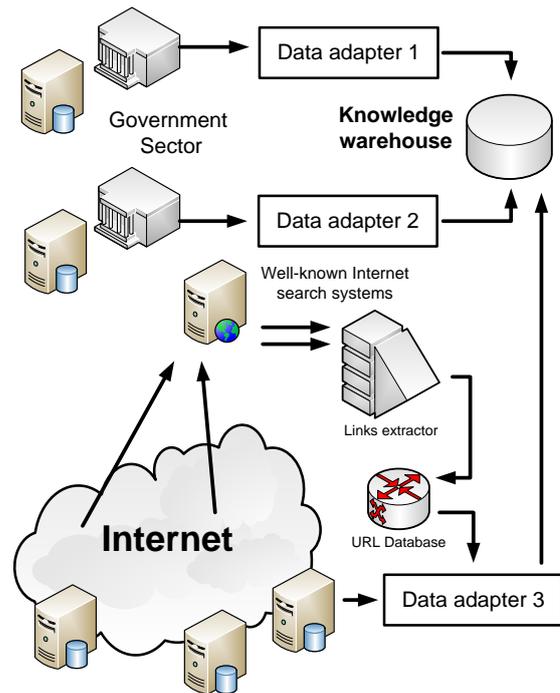


Fig. 1. Ways to enrich knowledge warehouse.

There are resources data adapters for each type of resource. The main task of data adapter is extracting and converting data to common format and prepare data for processing. Well-known Internet search systems (e.g. Google, Yandex, Bing et al.) may be used to collect initial URL database by keywords and search queries.

On the next step, web-crawler performs scanning linked pages and extracts text fragments for analyzing and classification.

Figure 2 exhibits the process of extracting interesting knowledge from the general information flow and their subsequent processing. Each information flow is unified and the text fragments are extracted. Some elements may contain keywords, abstracts, descriptions in explicit forms. The other elements may

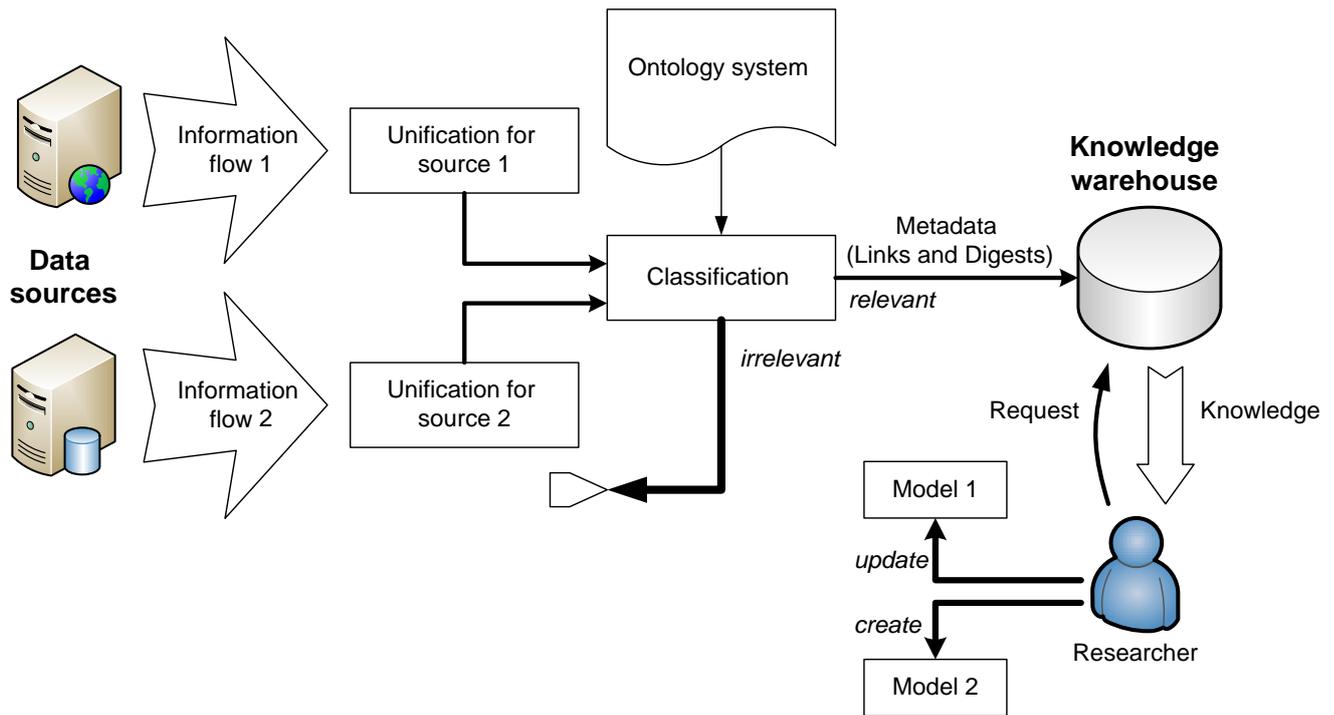


Fig. 2. The process of extracting knowledge from the sources, transformation and using for updating research models.

contain only raw text. At the stage of classification two important tasks are performed: (a) definition of word-set and (b) mapping with concepts of the Ontology system. The result of this step is the set of normalized word-sets (patterns). The groups of patterns are mapped to keywords and concepts of ontology space. Thus, the text fragments of elements are mapped to some set of concepts and numerical characteristics. As a result of classification at the last step the digest contains keywords related to the Ontology system concepts only. This element holds information about the meaning of the text fragment of the knowledge element. The Knowledge warehouse keeps digests, links on raw documents and ontologies.

The texts of scientific and specialized literature were used for learning and testing the classification subsystem of text. Besides, the other texts were used to determine common phrases of the Russian language (words of general vocabulary).

For creation information flow specialized data source adapter should be developed. In case government information systems this is not so difficult task, because information already semistructured. For implementing knowledge extraction in pure Internet should be used crawlers as a part of information search system. But this way may be very expensive by time and money. Using a database of well-known Internet search systems allows to us get access to links on related Internet resources.

For each search query we can get several hundred links and then inspect and classify each.

Technology of application consists of several steps:

- (i) designing the ontology system;
- (ii) machine learning on the texts (extracting, definition and binding word-sets with ontology concepts);
- (iii) preparation information flows;
- (iv) automatic classification and knowledge collection from information flows;
- (v) using knowledge warehouse for information search and forecasting.

Steps (i) and (iv) may be combined to enrich ontology system of a new concepts and word-sets. In this case a newest unusual word-set generate an ontology concept and probabilistic connection with subject matter of neighbor word-sets of current text fragment. In the future this probability of connection may be increased or decreased also automatically by analyzing other related texts.

4. Intelligent Service for Energy Technology Forecasting

To support collaborative expert work and substantiate recommended solutions the author develops Intelligent Information Environment of Energy Innovative Development Forecasting (IIEIDF). IIEIDF supplies interconnections between applications and stores and

transforms the data and knowledge. Figure 3 exhibits the architecture of Intelligent Information Environment

of Energy Innovative Development Forecasting.

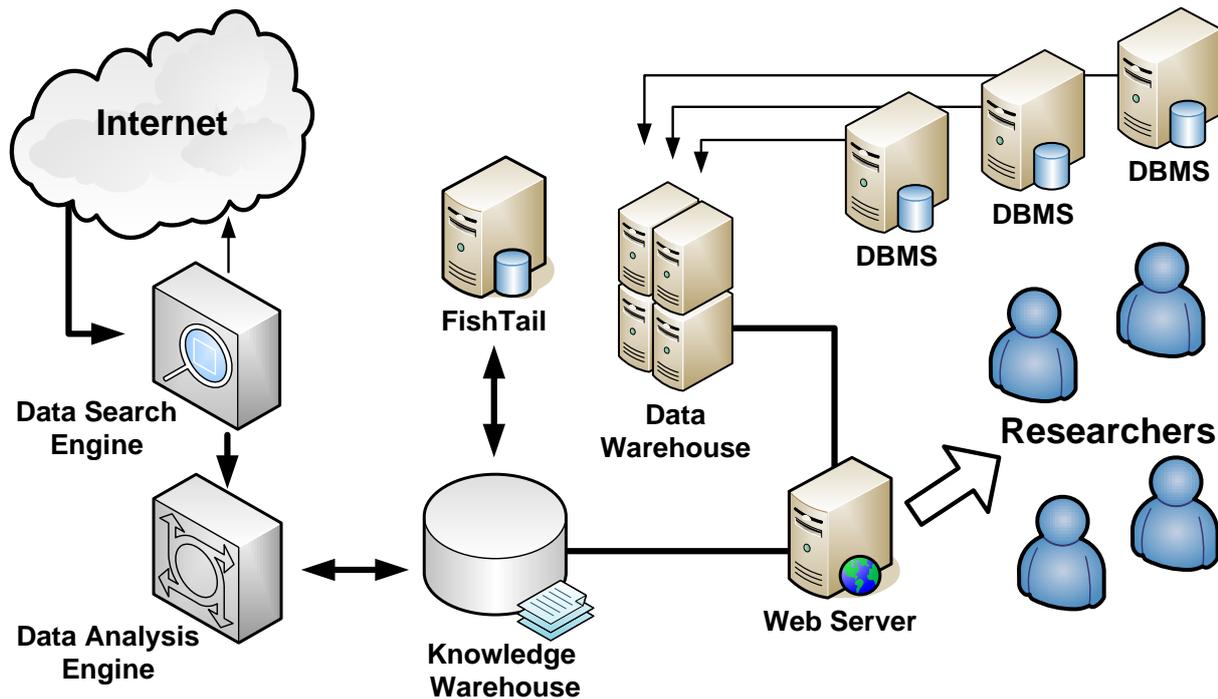


Fig. 3. Architecture of Intelligent Information Environment of Energy Innovative Development Forecasting.

Intelligent service consists of several components: data search and data analysis engines, knowledge and data warehouses; FishTail server; Web-server supporting the server-side components of Web-applications; data search and analysis engines. IIEIDF back-end contains applications for interacting with researchers.

Methodical approaches to knowledge collection from text sources are employed to solve some tasks:

- to obtain raw semi-structured text data;
- to clean, identify and extract the text;
- to classify the text; draft the digest for each obtained element.

As a result of classification at the last step the digest contains keywords related to the Ontology system concepts only. This element holds information about the meaning of the text fragment of the knowledge element.

Figure 4 exhibits the process of extracting and analyzing data from various sources. The data search engine automatically performs scheduled searching, obtains knowledge elements and then extracts the text content. Data adapters are plug-ins of data search engine and realize algorithms of clearing, identification and text extraction. Thus, data search engine may works as distributed system to improve performance. The data

analysis engine performs lexical and morphological analysis of the text fragments and mapping to concepts.

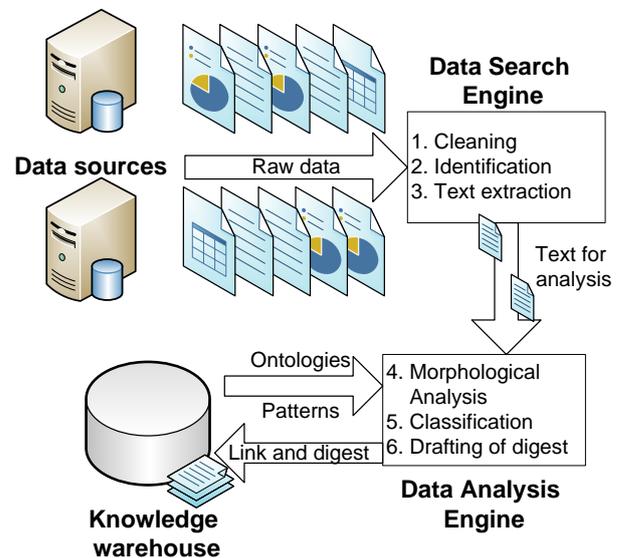


Fig. 4. The process of extracting and analyzing data.

5. Implementation of Intelligent Service

Implementation of Web-based application components IIEIDF are performed in a Private Cloud on the basis of approach of Rich Internet Applications (RIA). In RIA the partial implementation of the user interface is performed on tough standardized software tools providing better control over the user interface and capable to bypass many incompatibilities with configurations of Web-browsers.

RIA Web-application is a full functionality of traditional desktop applications. The implementation of the RIA client Web-based applications is performed on the Java platform as the most suitable and convenient one for development. The technology of Java Web Start (JWS) and Java Network Launching Protocol (JNLP) lets downloading and running quazilocal applications that are automatically updated through the Internet and work outside of the Web-browser environment.

The FishTail Server developed at ESI SB RAS is applied to store and manage RDF-triples associated with some knowledge and concepts of Ontology space. The FishTail is employed for enrichment by the new RDF-triples based on some predefined rules and existing RDF-triples.

Researchers commonly apply relevant modeling tools to support the proposed methodical approach, e.g. OntoMap, CognitiveMap, EventMap and BayNet. The modeling software tools for qualitative analysis are based on graphical core GrModeling. This approach provides shared space for building and transforming models.

Implementation of native Web-application that contained not so many various functions is produced with such technologies as JavaScript (jQuery, Bootstrap et al.) and HTML5. Server-side components are developed in PHP and Java.

6. Conclusion

The Open Linked Data from government sector and other commercial systems may be employed as the additional knowledge sources for the purpose of the energy infrastructure development forecasting.

The analysis of text fragments and creation of ontology-based digest improve the efficiency and quality the forecast estimation via simplified access to knowledge sets.

These techniques surely improve quality of forecast results in the studies of energy infrastructure development.

Acknowledgements

This paper was supported by the Russian Foundation for basic Research (RFBR) via projects 17-07-01341, 16-07-00569. The results were obtained during the implementation of the basis scientific project of the fundamental research programs of SB RAS III.17.2.1, reg. № AAAA-A17-117030310444-2, III.17.1.4, reg. № AAAA-A17-117030310436-7.

References

1. *The technological forecast of sectors development of the Russian fuel and energy complex for period till 2035. Ministry of energy RF*, (available at <http://minenergo.gov.ru/system/download-pdf/6365/66647>, February 2018).
2. Cagnin C., Keenan M., Johnston R., Scapolo F., Barre R., *Future-oriented technology analysis: strategic intelligence for an innovative economy* (Springer, 2008, 170 p., DOI: 10.1007/978-3-540-68811-2).
3. Nizhegorodtsev R. M., Ratner S. V., Trends in the Development of Industrially Assimilated Renewable Energy: The Problem of Resource Restrictions, *Thermal Engineering* (Springer, 2016, DOI:10.1134/S0040601516030083) 197–207.
4. Grmanova G., Laurinec P., Rozinajova V. et al., Incremental Ensemble Learning for Electricity Load Forecasting, *Acta Polytech Hungarica*, vol. 13, no. 2, (2016) 97–117. DOI: 10.12700/APH.13.2.2016.2.6.
5. Yussupova N., George Kovacs G., Maxim Boyko M. et al., Models and Methods for Quality Management Based on Artificial Intelligence Applications, *Acta Polytech Hungarica*, vol. 13, no. 3, (2016) 45–60. DOI: 10.12700/APH.13.3.2016.3.3.
6. Chanchetti L. F., Diaz S. M. O., Leiva D. R. et al., Technological Forecasting of Hydrogen Storage Materials using Patent Indicators, *International Journal of Hydrogen Energy*, vol. 41, (Elsevier, 2016, DOI:10.1016/j.ijhydene.2016.08.137) 18301–18310.
7. Kopaygorodsky A. N., Searching for information and development of expert assessment of new technological solution in technical infrastructure, in *Proceedings of the 19th International Workshop on Computer Science and Information Technologies. Germany, Baden-Baden*. (USATU, Vol. 1. 2017, ISBN 978-5-1030-8, ISBN 978-4-4221-1031-5), pp. 71–76.
8. Alexey V. Mikheev, Feasibility of tech mining for forecasting innovation pathways in energy sector, *Information and mathematical technologies in science and management*, vol. 4(8), (2017) 166–176 (In Russian).