International Journal of Knowledge and Language Processing
Volume 9, Number 3, **2018**

www.ijklp.org
KLP International ©2018 ISSN 2191-2734
pp.16—29

A Literature Review on Science and Technology Intelligence Technology

Jing Zhao[1] and Xin Tong[2]

[1]Institute of Scientific and Technical Information of China
No. 15 Fuxing Road, Haidian District
Beijing, 100038, China
zhaojing@istic.ac.cn
[2]Peking University
No. 5 Yiheyuan Road, Haidian District
Beijing, 100871, China

Received Sep 2018; revised Sep 2018

ABSTRACT. *Based on the meaning of scientific and technological intelligence technology in the current research, this paper elaborates on the scope, development process and evaluation method of scientific and technological information technology. At the same time, this paper summarizes the classification of information technology under various standards, and classifies and elaborates the current stage of science and technology information technology according to specific technology, in order to help readers better understand the connotation of the proper term of science and technology information technology.*

Keywords: Science and Technology Intelligence Technology, Intelligence technology, automatic classification, Gathering Technology, Ontology

1. **Introduction.** In 1983, Professor Qian Xuesen proposed that “Intelligence is the knowledge needed to solve a specific problem." "The first thing here is knowledge, and second, it is timely and targeted.”[1]

Scientific and technological intelligence is the knowledge needed to solve scientific and technological problems. China's science and technology information industry began in 1956, and its research has gone through five stages: translation report, subject professional research, intelligence review and review research, decision support research, and

knowledge service involved in decision-making.[2]

2. **Intelligence analysis methods.** Traditionally, the methods of scientific and technological information research include joint analysis, content analysis, technology roadmap, Delphi method, scenario analysis, etc.[3] With the advent of the era of big data, information science research has begun to be affected by methods of automation, intelligence and visualization, thus making network analysis, model simulation, bibliometric analysis, patent map analysis, Web text mining, Data mining, knowledge discovery and technology foresight have been developing rapidly in recent years.[4]

The results of scientific and technological information research are research reports, which are mainly divided into summary type, commentary type, prediction type, evaluation type and background type. Since the reform and opening up, science and technology information work has extended from the field of science and technology to the main battlefield of economy, emphasizing the strategic decision-making service for government's science and technology economy and the information support services for enterprise market.

He Defang[5] proposed a scientific and technological information research method based on factual data. He pointed out that many institutions at home and abroad have attached great importance to fact-based research, and the United States, Japan, Canada, Germany and other countries have made great progress in this field. In China, the T21 model developed by the China Institute of Science and Technology Information and the Millennium Research Institute of the United States can be used to fit the data related to sustainable development (such as population data, environmental data, etc.) and find the correlation of various factors, so as to provide a basis for policy adjustment.

Professor Boranscomb[6] , former US President's Scientific Advisor and Harvard University, defined Technology Roadmapping (TRM) as "a consensus on technology prospects based on scientific knowledge and insights". Zhang Tie-nan and others considered that the compilation of technical roadmap and science and technology intelligence research has commonality in research subjects (experts), research objects (existing data and information), research processes (both generating new knowledge and high intellectual creative labor), and thus can introduce technology roadmaps into scientific and technological information research. It provides a practical analytical framework for scientific and technological information research in theory.

3. **Science and Technology Intelligence Technology.** Professor Qian Xuesen believes that scientific and technological information work can be summarized into two aspects: first, collecting data, establishing database, and establishing a retrieval system to facilitate use; and second, activating these data into information. He Defang[7] pointed out that the current scientific and technological information industry is mainly engaged in information collection and service, intelligence research, competitive intelligence, scientific and technological evaluation and technological statistics, theoretical research and education and training. Since China's intelligence technology was used to solve the related problems in

the field of science and technology at the beginning of its development, the technology of scientific and technological information is intelligence technology. Zhao Zongren[8] and others think that intelligence technology is the general name of the methods and equipment used to acquire, process, store, retrieve and transmit text, digital, image and sound information. It mainly includes computer, communication, high-density storage, audio-visual, copying, printing and other technical fields and their applications in information work. Among them, computer technology is the core of intelligence technology. Wu Heng[9] believes that intelligence technology mainly refers to various technologies closely related to intelligence work, such as computer technology, microfilm technology, audio-visual technology and communication technology, and thinks that compared with information technology, intelligence technology lays more emphasis on application level. Ye Bingbing[10] et al. believe that intelligence technology is a variety of means used to achieve the goal of mechanization and automation of intelligence work. Laixinxia[11] defines intelligence technology as the main and specific technology that be used and mastered in the process of production, collection, processing and utilization of information, including computer processing technology, modern communication technology of text and image information transmission and miniaturization technology of compressing the volume of information carrier. In English, the term "intelligence technology" means the general term of technology for producing, storing, exchanging and using various forms of information. The most common classification of information technology in the intelligence field is according to the process of intelligence work, which can be divided into: intelligence input technology, intelligence storage technology, intelligence processing technology (intelligence analysis technology), intelligence output technology and intelligence transmission technology.[12-14] According to the field involved, it can be divided into computer technology, microfilm technology, copying technology and communication technology. According to the times it belonged, it can be divided into traditional intelligence technology and modern intelligence technology. According to the development stage of computer network, it can be divided into stand-alone intelligence technology, LAN-based intelligence technology and WAN-based intelligence technology.[15] With the trend of interconnection and integration of professional technologies such as integrated circuit technology, computer technology, communication technology, and broadcasting and television technology, it is obviously difficult to classify the intelligence technology by classification methods such as related technical fields, automation degree, times and computer network differentiation, and the information work processes in the digital age. With the trend of mutual integration, some technologies aiming at the whole or multiple processes cannot be classified into one kind of technology. Therefore, it is difficult to fully and accurately reflect the current intelligence technology system and its characteristics according to a single classification standard.[16] Therefore, this paper reviews intelligence technology directly from the perspective of specific technology implementation.

4. **Specific Intelligence Technology.** This section will describe the technology used in the

field of intelligence from the following aspects, including related technologies for information acquisition, intelligence analysis and other steps.

4.1. **Intelligence technology based on automatic classification.** In order to search, filter and manage massive and complex intelligence data more effectively, intelligence personnel must classify it. The efficiency of manual classification is low, which is not enough to meet the requirements of intelligence processing at this stage. Therefore, it is necessary to apply various advanced theories and technologies comprehensively to research and realize automatic classification of information.

Zhao Tianyun[17] proposed to use multi-class SVM for automatic classification of enterprise competitive intelligence. The specific steps are as follows: firstly, Web pages are preprocessed to transform the content into text, and text will be segmented, text features while be established, and then feature extraction of text is carried out by mutual information method. This paper proposes a method of combining SVM with decision tree, and uses a binary tree structure to divide each decision into a class. In order to prove the feasibility of this technology, he applied this technology to the analysis of the list of enterprise competitive intelligence resources websites, and obtained a Web page classification method combining SVM with multi-classification of decision tree, which was applied to enterprise competitive intelligence and its classification. The results show that support vector machine is a better classification method and can satisfy the Web knowledge discovery of a large number of professional websites demand.

Sun Chao[18] et al. put forward that SVM technology should be used to construct the model of enterprise competitive intelligence automatic classification system. Based on the idea of automatic classification of information and multi-agent of support vector machine, they constructed an intelligence classification system. It can be divided into three steps: kernel function selection, support vector machine building and competitive intelligence classification. At the same time, it needs auxiliary work such as intelligence preprocessing, feature extraction and quality assessment. In the competitive intelligence classification system, each agent completes the entire task by sharing resources, cooperating and serving each other, so that the whole system is autonomous, initiative and collaborative, so as to achieve the best results.

Bai Guanghui[19] et al. put forward the automatic classification technology of support vector machine (SVM) method for enterprise competitive intelligence system. They proposed to use support vector machine to form a classifier. The specific implementation process is as follows: firstly, establish a digital calculation model of training text; secondly, use support vector machine learning algorithm for classification training to establish a classifier; thirdly, use classifier to automatically classify unknown information. Subsequently, they selected 300 articles from the Internet including wireless communication, wired communication and mobile communication. The SVM classifier was used to test the article. The results show that the average accuracy of automatic classification using SVM algorithm can reach more than 80%, which can meet the practical requirements of enterprise intelligence and its automatic classification.

Xue Yanbo[20] proposed that Web text categorization technology should be used in

enterprise competitive intelligence analysis, using VSM (Vector Space Model) to extract features of web documents, and using K-nearest neighbor classification algorithm and Bayesian classification algorithm to classify web documents, which can effectively improve the effectiveness of enterprise competitive intelligence work.

Peipei Sun[21] et al. proposed using text classifier to segment scientific and technological information text more effectively. They use three word segmentations respectively based on Naive Bayesian algorithm, J48 and SMO algorithm to segment scientific and technological intelligence text. The specific steps are: obtaining the original data, preprocessing the date, selecting feature words, building a word segmentation device based on the algorithm and segmenting words. As a result, all three word breakers can effectively segment the scientific and technical intelligence texts. Among them, the SMO tokenizer takes the shortest time to process scientific and technological information texts. The word segmenter based on Naive Bayesian algorithm has the highest efficiency in the segmentation of scientific and technological information text, and can help the intelligence staff to better complete the information processing work.

4.2. **Web-based Intelligence Gathering Technology.** Web text mining is the discovery of implicit patterns from a large collection of Web documents. The process of Web text mining is the mapping from input to output. Web-based intelligence gathering technology can better complete the intelligence gathering work and adapt to the reality of the explosive development of the Internet.

Chen Duoling[22] et al. put forward that Web-based text mining technology would be used in the study of enterprise competitive intelligence. Enterprise competitive intelligence system based on Web text mining technology is composed of several subsystems: intelligence automatic search subsystem, intelligent data mining subsystem and intelligence service subsystem. They propose to use Vector Space Model (VSM) method to extract features of web text, classify documents in specific fields, and divide clustering methods into hierarchical clustering method and plane clustering method. Each text in the document is traversed many times, which is equivalent to constructing a spanning tree. Web mining technology is one of the main technologies of automation and intellectualization of enterprise competitive intelligence system, which can significantly improve the efficiency and effectiveness of enterprise competitive intelligence collection and analysis.

Yufeng Zhang[23] et al. proposed to automatically collect competitive intelligence based on Web semantic mining. They believe that the accuracy and relevance of information provided by search engines cannot meet the requirements of competitive intelligence collection. Web Semantic Mining technology can discover the hidden information in Web information resources and build a Web mining framework to better complete the collection of competitive intelligence.

Tian Xueyun[24] proposed using Naive Bayesian algorithm to predict the theme of the collected page content in the topic-based network competitive intelligence collection model. In order to realize the experiment of the subject's intelligence gathering technology, it selects China Food Safety Network (http://foodsafety.ce.cn/), Sina Network (http://www.sina.com.cn), Sohu (http://www.sohu.com), Netease (http://www.163.com) as

seed URLs, with the theme of "food safety". Each time a certain number of web pages are collected, the harvest rate is counted, so as to grasp the relationship of the harvest rate and collection rate. In the experiment, the system is compared with the traditional network acquisition (breadth first algorithm), content evaluation based topic crawler (Best first search algorithm), link relationship based topic crawler (PageRank algorithm), standard keyword algorithm. It is concluded that the improved Naive Bayesian algorithm has higher accuracy and can better complete the intelligence collection.

4.3. **Semantic-based intelligence technology.** Scientific and technological literature is the carrier of scientific and technological information. Sentences are the basic unit of expressing the semantics of scientific and technological information. Analyzing and interpreting scientific and technological literature at the sentence level can effectively capture the semantics information of scientific and technological innovation. Semantic-based information technology includes ontology-based information technology, semantic annotation-based information technology and semantic association-based information technology.

4.3.1. **Ontology-based intelligence technology.** The concept of ontology was proposed by Gruber of Stanford University in 1993 as an agreement on shared concepts. Sharing concepts include a conceptual framework for domain knowledge modeling, interoperable system communication protocols and domain-specific theory representation protocols. In a knowledge sharing environment, ontology is described by defining vocabulary expressions.[25]

Since advanced information resources integration technologies cannot effectively solve the semantic heterogeneity problem between different data sources and traditional competitive intelligence technology cannot obtain high-quality deep intelligence knowledge. Zhang Yufeng[26] et al. constructed an ontology-based competitive intelligence semantic integration and analysis model to solve the problems of semantic heterogeneity between different data sources. The model uses ontology theory and method to organically integrate network easy-to-purchase information, and solves the problem of information isolation and we can realize competitive intelligence self-delivery analysis and intelligent analysis at the semantic level, which is conducive to improving the accuracy and efficiency of intelligence analysis.

Xu Ping[27] et al. proposed to use ontology to extract competitive intelligence in depth from the shortcomings of the accuracy and coverage of information extraction techniques used in competitive intelligence and processing. Ontology is applied to the whole process of competitive intelligence information extraction, such as text and processing, sentence segmentation, entity extraction, integration and entity relationship extraction. At the same time, the results of information extraction are used to enrich the existing ontology to better extract information in competitive intelligence.

Tian Lei[28] used ontology ideas and methods to provide semantic knowledge by using domain ontology to construct an enterprise competitive intelligence mining model based on domain ontology. The model mainly includes data acquisition and extraction module of enterprise competitive intelligence source, source data filtering and information processing

module, enterprise competitive intelligence analysis and mining module, enterprise competitive intelligence query retrieval and visualization display module. The automatic collection technology of source data is used to collect intelligence data. Intelligent intelligence analysis and mining methods are used to analyze and process intelligence data, and the mining results are visually displayed. It is concluded that enterprise competitive intelligence mining based on domain ontology can effectively utilize domain ontology to provide semantic knowledge such as shared concepts set and relationship between concepts in the domain, upgrade current information collection to semantic information mining and knowledge extraction, and realize automated, intelligent and efficient information mining and analysis.

Jiao Li[29] et al. proposed an ontology-based excavator to explore competitive intelligence in neuroscience. It is able to search for new discoveries in neuroscience through semantics while simultaneously pursuing new trends in data to support decision makers.

4.3.2. **Intelligence Technology Based on Semantic Association.** The semantic relevance can be calculated by the algorithm to calculate the semantic relevance, which can be better used for intelligence analysis.

He Chao[30] et al. integrated semantic association analysis into business intelligence analysis method and designed a business intelligence analysis algorithm based on semantic association analysis to conduct business intelligence association analysis and information knowledge discovery at the semantic level. They verify the superiority and validity of the algorithm by contrastive experiments, and draw a conclusion that the algorithm of business intelligence analysis based on semantic association analysis can significantly improve the performance of business intelligence mining and association analysis at the semantic level by making full use of the hierarchical conceptual structure and semantic reasoning function of ontology, and by using ontology to provide search constraint mechanism for top-down multi-level knowledge mining and analysis. The quality and efficiency of business intelligence analysis, the analysis results have stronger practicability and potential value.

Zhang Yufeng[31] et al. integrated problems and context knowledge into enterprise competitive intelligence analysis, designed an inductive learning algorithm based on semantic decision tree for information analysis and information knowledge discovery at the semantic level. They selected three data sets of Monk Balance Scale and Bresat Cancer provided by UCI website as the original data for experimental analysis to verify the superiority of the algorithm. It is concluded that the fusion of ontology and context knowledge for the semantic level of enterprise competitive intelligence analysis is to improve intelligence mining and Gaining efficiency and accuracy, as well as effective methods for specific application environments and objects, can lead to the creation of deeper intelligence content that is more potentially valuable and useful.

4.3.3. **Intelligence Technology Based on Semantic Annotation.** The process of semantic annotation is actually the process of similarity calculation. The similarity between concepts in the domain of extracting information language ontology is calculated. The concept corresponding to the maximum similarity is the concept corresponding to extracting

information.[32] Using semantic annotations in the field of intelligence analysis can improve the efficiency of intelligence processing.

Zhu Na[33] et al. proposed the use of semantic annotation technology for scientific and technological intelligence analysis, which is to use semantic role labeling technology to penetrate into the knowledge level of scientific and technological intelligence resources data, mainly aiming at the mining of scientific and technological innovation knowledge, discovering the semantic association between potential scientific and technological innovation knowledge and identifying the valuable scientific and technological intelligence. Specific processes are as follows: building professional corpus, comprehensively interpreting scientific and technological literature based on semantic role annotation, constructing mathematical models to describe scientific and technological innovation text vectors, and further purifying the text vectors through features, building machine learning classification model to calculate semantic similarity, and feeding back the relevant semantic content of scientific and technological innovation knowledge to the user in a specific form. Semantic role labeling can improve the ability of natural language processing in the process of scientific and technological intelligence analysis, and excavate valuable scientific and technological innovation knowledge hidden in massive scientific and technological data and their semantic connection, and then provide knowledge services for scientific researchers, and provide basis for scientific and technological strategic information decision-making.

Feng Xiangyun[34] put forward the technology of automatic document indexing and retrieval based on the semantic structure of word dependence, also known as "latent semantic indexing". The technology of word frequency statistics and singular value decomposition is used to capture the semantic structure of documents, and the vector representation of indexed words, questions and documents is obtained. The retrieval system can predict the correlation between documents and questions, so as to achieve the purpose of retrieval. LSI latent semantic indexing method first preprocesses and indexing documents, then uses singular-value decomposition (SVD) to deal with the semantic relationship between indexed documents, and finally calculates the similarity, which can effectively improve the information retrieval ability.

Wang Wei[35] proposed using semantic mining to intelligently collect, analyze and process unstructured data and information，which can help enterprises quickly establish an economic and effective competitive intelligence process. Firstly, semi-structured and unstructured information is transformed into structured data through semantic doll aggregation. Then, the semantic mining technology based on data warehouse is used to realize the information processing of competitive intelligence data. Finally, the competitive intelligence is analyzed. Introducing semantic data mining technology into the construction of competitive intelligence system will improve the intelligence level of enterprise competition and promote its development from information management to knowledge management.

4.3.4. **Intelligence Technology Based on LDA Model.** The Latent Dirichlet Allocation (LDA) model is an unsupervised learning model with text topic representation ability

proposed in recent years. It was proposed by David M. Bei in 2003. The LDA model mines a number of potential topic models from a data set, and then uses these topic models to represent a text, so as to achieve the goal of feature dimension reduction. [36] The LDA model can be used in intelligence collection and analysis to process intelligence better.

Liu Qihua[37] proposed to build a competitive intelligence collection system based on LDA (Latent Dirichlet Allocation) model and domain ontology, which includes web content analysis module, web link analysis module and theme correlation degree calculation module. Subsequently, Liu Qihua carried out the experiment of parameter determination in LDA model and topic correlation formula, and tested the performance of competitive intelligence acquisition system. The experimental results show that the competitive intelligence acquisition system based on LDA model and domain ontology effectively improves the subject harvest rate and achieves good results.

Guan Peng[38] et al. proposed a method of scientific and technological intelligence analysis based on LDA theme model. From CNKI, he selected 1018 papers in the field of new energy in China from 1994 to 2000. He built a database of titles, authors, institutions, abstracts and key words, and then used 10% of the documents in the corpus as test set evaluation model. The remaining papers were used to train LDA model. It is concluded that in the knowledge mining of scientific literature, using the method of scientific and technological information analysis based on LDA topic model can effectively determine the number of topics, obtain better results of topic extraction, help information analysts to extract significant topics from a large number of scientific and technological documents, and improve the recommendation effect of scientific and technological documents based on similarity.

Wang Bo[39] et al. proposed a patent analysis method based on topic model (LDA). Through the analysis of several topic models of extracting noun phrases from patent documents, the content of patent documents can be quantitatively studied in patent texts, and then the research hotspots and directions of patent technology in a certain field can be revealed, and the structure and intrinsic relationship between patent knowledge can be excavated. Taking LTE (Long Term Evolution), the fourth-generation mobile key communication technology in the field of mobile communication technology, as the research object, they chose Derwent Innovation Index (DII) in Web of Science of the American Institute of Scientific Information as the data source. According to the noun extraction rules, the downloaded patents are extracted from the feature words, and the LAD full probability generation model is used to classify the LTE-related patents. By improving the existing LDA model and adding the agency-topic probability level, we can clearly identify the main competitive enterprises in each hot topic, which is beneficial to the analysis of patent competitive intelligence.

4.4. **Intelligence Technology Based on Automatic Summary.**The so-called automatic summarization (abstract) is the use of a computer to automatically extract abstracts from the original documents. Abstracts are concise and coherent essays that accurately and comprehensively reflect the content of a document. The research of automatic abstracting began in 1952 and was first proposed by H.P. Luhn. The study of automatic abstracting

technology can improve the effectiveness of abstract magazine.[40]

Li Nianfeng[41] proposed a network intelligence collection system model based on automatic summary to better complete the intelligence collection. The network intelligence collection system mainly depends on automatic summarization technology and network security technology. The specific steps of using automatic summarization technology to collect intelligence in the network are as follows: firstly, the structure of network documents is analyzed, and then the key words and key sentences of documents are eliminated by using statistical methods and heuristic rules. Finally, the honor of key sentences is eliminated by using semantic distance and the document summary is generated, which is of guiding significance for improving other information systems.

Wang Yanhong[42] et al. proposed to build a technical framework for news information analysis based on automatic summary technology, so as to achieve timely and convenient presentation of important information in news intelligence to users, and to provide users with an efficient experience of news information analysis. The implementation of the framework requires three phases: the analysis of the original text, the transformation of presentation and the generation of abstracts. Because of the semi-structured characteristics of news information data, the most frequently used method in the process of automatic generation of news intelligence abstracts is to construct a vector space model.

4.5. **Visualization-based Intelligence Analysis Technology.**The interaction between visual representation tools, using human visual system to help people form a certain thinking model, aims to improve the "depth" and "range" of interaction between researchers and resources, and inspire users thinking. Visualization can be divided into Radial Graph Visualization, Theme Graph Visualization and Structured Data Visualization.

Dong Xianzhou[43] et al. put forward a river model to help intelligence analysts determine the main content and relationship of time. The river model belongs to thematic visualization structure. In addition, it puts forward an association analysis model to help intelligence analysts quickly find the main factors in information.

Johannes Kehrer[44] et al. used visual analysis to study multi-dimensional scientific data. Wu Ming[45] used visualization method and SCIE analysis to visualize the literature in the field of nanobiotechnology. Hou Xiaorong[46] et al. used visual analysis method to analyze the patents of medical endoscope and therapeutic ultrasound invention in China, and drew patent citation map, which is helpful for the objective evaluation of patent.

4.6. **OLAP-based Intelligence Analysis Technology.**OLAP is a software technology, which enables analysts to observe information from all aspects quickly, consistently and interactively, so as to achieve in-depth understanding of data.[47]

Matteo Golfarelli[48] proposed to apply OLAP analysis technology to the analysis of user's original text, so that decision makers can analyze their business according to the trends perceived from the environment. He proposed to build a social business intelligence model, including crawling components, semantic enrichment components, ETL components. Using OLAP tools in a multi-dimensional way uec, can greatly improve the flexibility of the model structure, thus better support the decision-making process.

Ipek Deveci Kocakoç[49] et al. proposed building a software based on OLAP and data

mining to analyze business intelligence (mainly for retail). The application includes the following aspects: building data warehouse, building OLAP cube, and applying data mining algorithm on OLAP cube. OLAP database increases the extensibility of relational database and can effectively form a competitive advantage.

4.7. **Intelligence Technology Based on Cluster Analysis.**Cluster analysis has a long history and is one of the important research directions in data mining and pattern recognition. [50] Using cluster analysis in the intelligence field can improve the efficiency and accuracy of intelligence processing.

Quweiqun[51] proposed to use grey clustering analysis to evaluate the results of intelligence research. Grey clustering is to sum up the whitening number of clustering objects for different clustering indicators according to several grey classes in order to determine which kind of clustering objects belong to. The concrete steps are as follows: constructing the cluster whitening number d_{ij}, determining the whitening function of grey class, calculating the cluster weight, calculating the cluster number δ_{ik}, constructing the cluster vector and clustering six steps. The results of intelligence research are evaluated quantitatively.

Zhang Yufeng[52] et al. introduced cluster analysis into competitive intelligence analysis, and constructed a clustering analysis model of competitive intelligence based on domain ontology. This model uses the domain ontology of competitive intelligence to know the text for analysis and annotation, extracts the relationship between feature concepts and concepts, and realizes the text representation at the semantic level. Aiming at the puzzlement that traditional clustering analysis algorithm cannot carry out deep clustering analysis, a clustering analysis algorithm based on semantic core function is designed to mine and analyze, and realize deep clustering analysis. They chose Reuters-21578, a standard text data set commonly used in text clustering research, to conduct experiments. The experimental results show that the model achieves good expected results, and significantly improves the accuracy and efficiency of competitive intelligence analysis.

Tang Rui[53] et al. proposed the application of natural clustering algorithm to network intelligence data processing. Alfred Mahr[54] et al. used cluster analysis to predict some cases in the field of clinical medicine. C. C. Nwangburuka[55] et al. used cluster analysis to evaluate the genetic diversity of seeds in different agro-ecological areas in Nigeria.

5. **Conclusions.** In the early stage, the development of intelligence technology mainly relies on manual work. The sources of information collection are newspapers, magazines in related fields, etc. It is necessary to collect relevant information manually and classify it manually. In this process, a lot of manpower and material resources are needed. With the further improvement of information science and the continuous development of the Internet, information technology began to be used in the field of information, and information technology was divided according to the process of information processing and the era of information technology. It mainly includes automatic classification, automatic indexing, automatic summary, semantic analysis and annotation, content analysis, OLAP on-line analysis and processing technology for information processing, information retrieval and

service technology for information application services, artificial intelligence (natural language processing, machine translation, machine learning) and human-computer interaction technology for information collection, storage and transmission. And so on. With the further development of computer technology, the future science and technology information technology will depend more on the development of new technologies such as artificial intelligence. How to better combine new technologies with information processing is also a problem that the staff in the field of information need to consider.

REFERENCES

[1] Qian Xuesen. Science and Technology of Science and Technology Information Work[J]. Information Theory and Practice, 1983(6): 3-10.

[2] Liang Zhanping. Exploration and Development of Science and Technology Information Research in China[J]. Information Exploration,2007(7):3-7.

[3] Zhao Fan. Comparative Study on Strategic Intelligence Analysis Methods of Foreign Consulting Information Organizations[J]. Journal of Information,2008(3):132. 134,141

[4] Liang Zhanping. Thoughts on the Innovation and Development of Information Science Subjects[J]. China Information Guide,2007(1):6-7,15.

[5] He Defang. Thinking on the Research of Scientific and Technological Information Based on Factual Data[J]. Journal of the China Society for Scientific Information, 2009, 28(5): 764-770.

[6] Zeng Lu, Sun Yongming. Principles and Formulation of Industrial Technology Roadmap. Guangzhou: South China University of Technology Press, 2007: 2-3

[7] He Defang. Discussion on the Development Direction of China's Science and Technology Information Industry[J]. Journal of the China Society for Scientific Intelligence, 2008, 27(4): 483-489.

[8] Zhao Zongren. China Encyclopedia Publishing House. China Encyclopedia Online Edition. [2006-07-28], http://210.37.23.25:918.

[9] Wu Heng. China Encyclopedia Publishing House. China Encyclopedia Online Edition. [2006-07-28], http://210.37.23.25:918.

[10] Ye Bingbing, Wu Qizhi. Library and Information Handbook. Beijing: Xueyuan Publishing House, 1988

[11] Lai Xinxia. A Concise Dictionary of Library and Information Science Archives. Tianjin: Nankai University Press, 1993

[12] Liang Zhanping, Liang Jian. An Analysis of the Development Trend of Information Science in the New Century. Intelligence Theory and Practice, 2005(5): 25-229

[13] Jing Jipeng, Zheng Rong, Research on the Innovation Mechanism and Innovation Field of the Development of Information Science in China. Journal of the China Society of Information Science, 2005, 24(3): 268-276

[14] Wang Zhijin, Zheng Hongjun. Development of Information Science and Construction of Discipline System in the Digital Age. Journal of Zhengzhou University (Philosophy and Social Sciences Edition), 2005(4): 9~11

[15] Information Science. [2018-07-24], https://en.wikipedia.org/wiki/Information_science.

[16] Li Guangjian, Huang Yongwen, Kong Jing, et al. Information Technology in the Digital Age[J]. Digital Library Forum, 2006(10): 61-71.

[17] Zhao Tianqi. Application of Multi-class SVM in Automatic Classification of Enterprise Competitive Intelligence[J]. Modern Information, 2008, 28(10): 184-186.
[18] Sun Chao, Zhang Yufeng. Model of Enterprise Competitive Intelligence Automatic Classification System Based on SVM[J]. Journal of Information, 2007, 26(3): 40-43.
[19] Bai Guanghui, Lian Hao, Xu Hongbo, et al. Application of Automatic Classification Technology in Enterprise Competitive Intelligence System[J]. Computer Engineering and Applications, 2005, 41(18): 212-214.
[20] Xue Yanbo. Application of WEB Text Classification Technology in Enterprise Competitive Intelligence Analysis[J]. Information Science, 2004, 22(3): 378-380.
[21] Sun P P, Zhu Q Y, Zhou L, et al. Comparative Analysis of Text Categorizer on Science and Technology Intelligence [J]. Applied Mechanics and Materials, 2014, 530-531: 502-505.
[22] Chen Duoling, Hu Xiaofeng. Research on Enterprise Competitive Intelligence System Based on WEB Text Mining Technology[J]. Journal of Information, 2005, 24(6): 22-24.
[23] Yufeng Z, Jinhong W, Cuibo W . Automatic Competitive Intelligence Collection Based on Semantic Web Mining [C]// International Conference on Wireless Communications. IEEE, 2007.
[24] Tian Xuewei. Research on the Collection Technology of Network Competitive Intelligence Subjects[J]. Library and Inforamtion, 2014(5):132-137.
[25] MIKE U, MICHAEL G. Ontologies: principles, methods and applications [J]. Knowledge Engineering Review, 1996 (2) : 6-11.
[26] Zhang Yufeng, He Chao. Research on Semantic Integration and Analysis of Competitive Intelligence Based on Ontology[J]. Information Theory & Practice, 2011, 34(7): 64-66.
[27] Xu Ping, Shao Bo. Pre-processing Analysis of Competitive Intelligence Based on Ontology Information Extraction[J]. Journal of Information, 2008, 27(9): 33-35.
[28] Tian Lei. Research on Enterprise Competitive Intelligence Mining Model Based on Domain Ontology[J]. Information Science, 2014(6): 43-46.
[29] J, Huang M, Zhu X. An Ontology-Based Mining System for Competitive Intelligence in Neuroscience [M]// Web Intelligence Meets Brain Informatics. Springer Berlin Heidelberg, 2007.
[30] He Chao, Zhang Yufeng. Research on Business Intelligence Analysis Algorithm Based on Semantic Correlation Analysis[J]. Journal of Information, 2013(4): 134-137.
[31] Zhang Yufeng. Research on Enterprise Competitive Intelligence Analysis Algorithm Based on Ontology and Context Knowledge[J]. Information and Documentation Services, 2012, 33(2): 48-52.
[32] XU Bin, LI Yi-zi, WANG Ke-hong. Semantic annotation method for Web services[J]. Journal of Tsinghua University(Science and Technology), 2006, 46(10): 1784-1787.
[33] Zhu Na, Wang Xiaoyue, Bai Rujiang. Semantic Role Labeling and Its Application in Scientific and Technological Information Analysis[J]. Journal of Information Theory and Practice, 2015, 38(1): 98-103.
[34] Xiangyun F, Feng Xiangyun. Application of LSI Latent Semantic Indexing Method in Information Retrieval[J]. Modern Library and Information Technology, 1998, 14(4):19-21.
[35] Wang Wei. Research on Intelligent Competitive Intelligence System Based on Semantic Mining[J]. Information Theory & Practice, 2008, 31(5): 773-776.
[36] Zhang Xiaoyan, Wang Ting, Liang Xiaobo. Application of LDA Model in Topic Tracking[J]. Computer Science, 38(Z10).

[37] Liu Qihua. Research on Competitive Intelligence Collection Based on LDA and Domain Ontology[J]. Information Science, 2013(4): 53-57+64.
[38] Guan Peng, Wang Rifen. Research on the Method of Determining the Optimal Topic Number of LDA Topic Model in Scientific and Technological Information Analysis[J]. Modern Library and Information Technology, 2016, 32(9): 42-50.
[39] Wang Bo, Liu Shengbo, Ding Wei, et al. Analysis of Patent Competitive Intelligence Based on Topic Model--Taking the Field of LTE Technology in Communication Industry as an Example [C]// Seminar on Scientometrics and Information Metrology in the Mainland and Taiwan. 2013.
[40] Li Aihong. On Automatic Abstracting Technology[J]. Library and Information Service, 2000(4): 40-42.
[41] Li Nianfeng. Research on Network Information Collection System Based on Automatic Abstract[J]. Modern Intelligence, 2007, 27(11):161-163.
[42] Wang Yanhong, Zhou Jun. Research on Automatic Abstract Technology for News Intelligence Analysis[J]. Journal of Hunan Industrial Polytechnic College, 2014(4):18-20.
[43] Dong Xianzhou, Hu Xiaofeng, Si Guangya. Application of Information Visualization Technology in Intelligence Analysis[J]. Computer Engineering and Applications, 2006, 42(34): 175-177.
[44] Kehrer J, Hauser H. Visualization and visual analysis of multifaceted scientific data: A survey[J]. Visualization and Computer Graphics, IEEE Transactions on, 2013, 19(3): 495-513.
[45] Wu Ming. Visualization Analysis of Subject Domain Information——Taking the Field of Nanobiotechnology as an Example[J]. Library and Information Service, 2010, 54(6): 44-47.
[46] Hou Yurong, Zhao Dechun, Hu Hong. Visualization Analysis of Patent Citation Type[J]. Science and Technology Management Research, 2011, 31(17): 142-144.
[47] Han Peiyou, Dong Guiyun, Liu Hong. Database Technology: Northwestern Polytechnical University Press, 2008. 12
[48] Golfarelli M. Social business intelligence: OLAP applied to user generated contents[C]// International Conference on E-business. IEEE, 2016.
[49] Ipek Deveci Kocakoç, Sabri Erdem. Business Intelligence Applications in Retail Business: OLAP, Data Mining & Reporting Services [J]. Journal of Information & Knowledge Management, 2010, 9(02):
[50] Sun Jigui, Liu Jie, Zhao Lianyu. Research on Clustering Algorithm[J]. Journal of Software, 2008, 19(1):48-61.
[51] Qu Weiqun. Application of Grey Clustering Analysis in the Evaluation of Intelligence Research Results[J]. Journal of Information, 1997(5): 30-31.
[52] Zhang Yufeng, He Chao. Research on Clustering Analysis of Competitive Intelligence Based on Domain Ontology[J]. Information Science, 2011(11):1613-1615.
[53] Rui T, Fong S, Yang X S, et al. Nature-Inspired Clustering Algorithms for Web Intelligence Data [C]// IEEE/WIC/ACM International Conferences on Web Intelligence & Intelligent Agent Technology. IEEE, 2012.
[54] Mahr A, Katsahian S, Varet H, et al. Revisiting the classification of clinical phenotypes of anti-neutrophil cytoplasmic antibody-associated vasculitis: a cluster analysis[J]. Annals of the rheumatic diseases, 2013,72(6): 1003- 1010
[55] Nwangburuka CC, Kehinde OB, Ojo DK, et al. Morphological classification of genetic diversity in cultivated okra, Abelmoschus esculentus (L) Moench using principal component analysis (PCA) and single linkage cluster analysis (SLCA) [J]. African Journal of Biotechnology, 2013,10(54): 11165-11172.