Strategic information perception methods and practices in the open source intelligence[①]

Zeng Wen(曾　文)[*], Li Hui[②][**]

(* Institute of Scientific and Technical Information of China, Beijing 100038, P. R. China)
(** Beijing Institute of Science and Technology Information, Beijing 100044, P. R. China)

Abstract

Open source intelligence is one of the most important public data sources for strategic information analysis. One of the primary and core issues of strategic information research is information perception, so this paper mainly expounds the perception method for strategic information perception in the open source intelligence environment as well as the framework and basic process of information perception. This paper argues that in order to match the information perception result with the information depiction result, it conducts practical exploration for the results of information acquisition, perception, depiction and analysis. This paper introduces and develops a monitoring platform for information perception. The results show that the method proposed in this paper is feasible.

Key words: open source intelligence (OSINT), open source information (OSIF), information perception, methods and practices, information, science and technology

0　Introduction

With the spread of Internet and development of science and technology, information is becoming more and more extensive and complex. Occupying an increasingly important position in national security, strategic intelligence has become one of the important factors that affects national security. The value of open source information (OSIF) to supplement classified intelligence has long been recognized, but the growing pervasiveness of the Internet and the rise of social media and big data analysis in the past two decades have revolutionized open source intelligence (OSINT). The RAND Corporation defines OSINT as publicly available information that has been discovered, determined to be of intelligence value, and disseminated by a member of the Intelligence Community (IC)[1]. Chinese military master Sun Tzu said: 'Know the enemy and know yourself, you can fight a hundred battles with no danger of defeat.' The fundamental purpose of strategic information pre-judgment is to maintain national security and interests and be in an invincible position in the face of a complex, varied and fiercely competitive global pattern of intense rivalry. A wide range of strategic information can be divided into political, economic, military, cultural, scientific and tech-

nological, and other types of strategic information from the perspective of national economic sector. The focus hereof is to explore and study strategic information methods in the network environment. The study of strategic information refers to a process integrating the analysis, comparison and fusion of large amounts of information, the selection and identification of various action plans and strategies, and the prediction of the future development trends as well as the impact on politics, economy, technology, and society through scientific methods based on the mastery of relevant information, with the focus on the long-term and overall issues concerning politics, economy, military, culture, and scientific and technological development[2]. From the perspective of current practice of information services at home and abroad, strategic information research can be divided into three types based on the needs of intelligence users: 1) Dynamic monitoring of strategic information, which refers to a strategic information research method that an information agency collects, compiles, sorts and processes the latest global and strategic dynamic information about the target field in a timely, periodic, and continuous manner and provides it to information users (i. e. higher management authorities) for use in a certain form of information report; 2) Situation analysis of strategic information, which refers to that the current status and trend of target areas are rec-

---

ognized, the factors influencing the development are analyzed, and the actual intention and purpose are inferred to prepare for further forward-looking forecasts, through the analysis and research of strategic information; 3) Forward-looking forecasting of strategic information, which refers to the prediction of the development status and situation of the research object in a future long period of time through systematic analysis of the content of strategic information. Perception is one of the primary and core issues in strategic information research.

1 Related work

1.1 Classification and characteristics about open source intelligence

Information produced from publicly available information is collected, exploited, and disseminated in a timely manner to an appropriate audience for the purpose of addressing a specific intelligence requirement. In 2011, in a document issued by the Office of the Director of National Intelligence, OSINT was defined as "intelligence produced from publicly available information that is collected, exploited, and disseminated in a timely manner to an appropriate audience for the purpose of addressing a specific intelligence requirement". Since its emergence, open source intelligence has been a service at the level of national security and national strategy. So far, open source intelligence in the United States can be divided into three generations.

1) Open source intelligence of the first-generation

The open source intelligence of the first-generation emerged with the establishment of the Foreign Broadcasting Information Service Office in the United States, which is characterized by dependence on translation professionals, physical information acquisition, regular publication, and focus on analysis and distribution.

2) Open source intelligence of the second-generation

Second-generation open source intelligence emerged with the establishment of DNI Open Source Center in 2005, which is characterized by dependence on technical experts, virtual information acquisition, continuous acquisition, and focusing on utilization and production. Emergence of Web 2.0 caused the evolution of open source intelligence, and Internet context was turned to dynamic web pages and user-generated contents.

3) Open source intelligence of the third-generation

For more than a decade, technical experts have been talking about Web 3.0 ('Semantic Web'). The third-generation open source intelligence will be characterized by direct and indirect machine data processing, machine learning, and automatic reasoning through the intelligent technology.

1.2 The concept of information perception

As direct reflection of objective things in the human brain through sensory organs, perception can be understood literally as being perceived by senses. However, from the perspective of cognition, the difference between sensory sensitivity and cognitive understanding creates differences in the perceptual ability. In the context of information perception, 'sense' is no longer a complete passive response. Instead, professional information personnel actively uses a variety of methods and tools to process data and information. 'Knowledge' is no longer confined to vague learning and knowing, but clear understanding, assessment, analysis and expectation. Bao Changhuo said that perception was part of the information work just as 'eyes and ears'. Information perception is a process of collecting, selecting, and interpreting information in the information work to gain an understanding of the current situation of stakeholders and to predict future development[3]. Hua[4] argued that as a combination of senses and perceptions of information, information perception was the starting point for understanding and absorbing information. Documentary information is one of the sources and raw materials for information production. Different from traditional literature information services, information perception is a process of analyzing information and eventually forming value-added products that serve for decision-making. For the information work, importance should be attached to conversion of information into information and strategy as well as information perception, which is the core task of intelligence work and research. Heuer believes that one of the limitations in intelligence work is the restriction of perception on the cognitive accuracy and timeliness of intelligence analysis, which reflects the fact that intelligence perception is a key part of information work[5]. In the process of information perception, although more information and analysis do not ensure better or more successful decisions, logic and experience show that decisions based on in-depth understanding of related factors and careful analysis are more likely to get the desired result than those simply based on information assessment. The perceptions of professional analysts who are familiar with the needs of decision makers can increase the correctness of judgment description and strengthen the recognition and implementation confidence in the overall strategy[6,7].

Therefore, strategic information refers to that information professionals achieve the understanding, interpretation and expression of strategic information needs, information target content and information task organization in the course of information collection, processing and analysis by comprehensive use of various knowledge tools to help strategic decision makers make a judgment when information is incomplete.

1.3 Target content of information perception research

The 'perception' of information perception does not mean learning or knowing, but understanding, assessment, analysis and expectation. Research on information perception should focus on perceptual awareness, perceptual readiness, and perceptual implementation. For perceptual awareness, it should be made clear that the main target of information perception is the objects that are unknown or not clearly known, and that perceiving the unknown is not equivalent to exploring the unknown. Making judgments and descriptions based on the understanding of the current or prevailing situation is a typical task of information perception. The preparation of information perception mainly refers to situation perception. Perceptual implementation does not simply organize several task processes. The relationship among the task environment in which information is perceived, the task objects that are confronted, and the task methods that are held usually needs to be

sorted out based on the research perspective of complex systems[7]. For the exploration of field technology, information technology is taken as an example. If the goal of information science projects focuses on 'fast' and 'precise' calculation, that of information science projects should include 'early' and 'far-reaching' understanding. For implementation, it should be known that the implementation of information perception is not simply to organize several task processes[8-10].

2 Research on strategic information perception

2.1 Framework of strategic information perception

The basic framework of strategic information perception research mainly includes (see Fig. 1):

1) Data source of information perception

The basic conditions of information resources are analyzed, and the rules and ways of information data fusion are explored based on the special requirements of professional fields and specialized tasks.

2) Core methods of information perception

Full-spectrum scanning is adopted for purposes of sorting, evaluating, designing and creating the methods, technologies and corresponding combinations suitable for perception of open source strategic information and exploring relevant methodological elements.
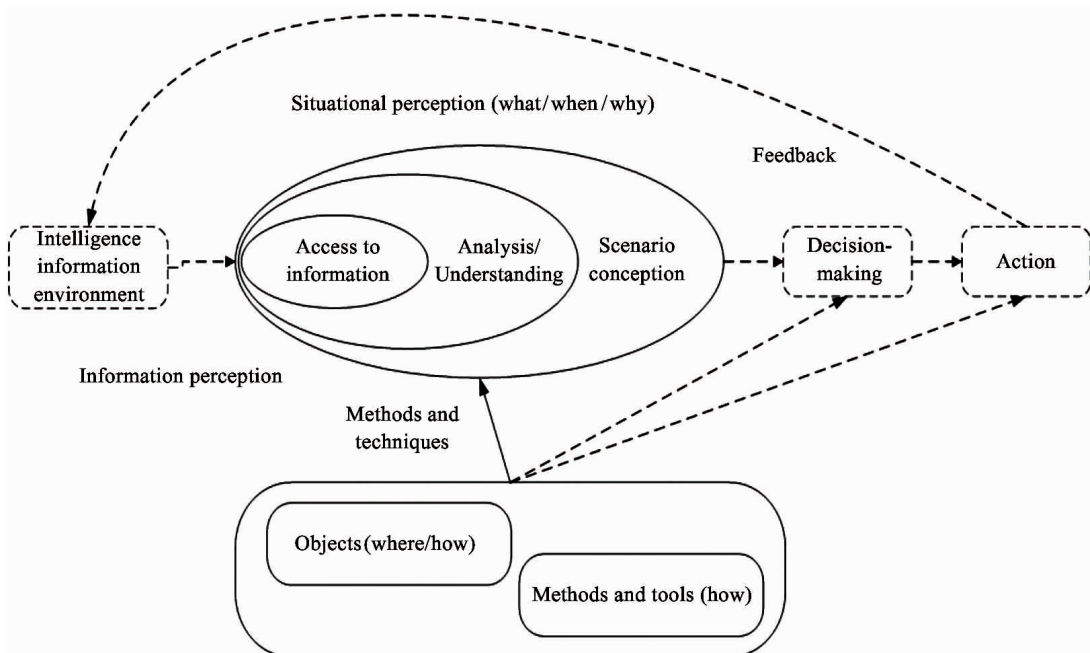

Fig. 1   The framework for strategic information perception

3) Depiction of information perception results

Information analysis products are described and presented by comprehensive use of various means to improve the efficiency of information delivery and use.

2.2   The basic process of information perception

The basic process of strategic information perception mainly includes (see Fig. 2):

1) Information perception

Long-term monitoring and comprehensive scanning of perceptual objects are conducted. The scanned results are changed into an understandable and usable state through pre-processing. On this basis, preliminary analysis of the developmental trend of perceptual objects is made, and the applicable field of perceptual results and the value of the field are pre-judged.

2) Information depiction

By understanding, analyzing, and portraying the characteristics and prototypes of different types of information users, needs and tasks, the objects of informa-

tion service and the scope of decision support are more focused.

3) Information response

Upon information perception and information depiction, the latest report on political, economic, military, and scientific and technological aspects, that is, information response, is obtained. The results of information response are information products, including dynamic reports and research reports.

Information perception and information depiction are not two phases completely severed, but two steps are closely linked. With respect to information perception, without initial cognition and understanding of intelligence users, needs and tasks, it is impossible to effectively predict and evaluate the value of perceptual objects; with respect to information depiction, without an accumulation of the corresponding sensory abilities, it is impossible to accurately understand and depict intelligence users, needs and tasks.
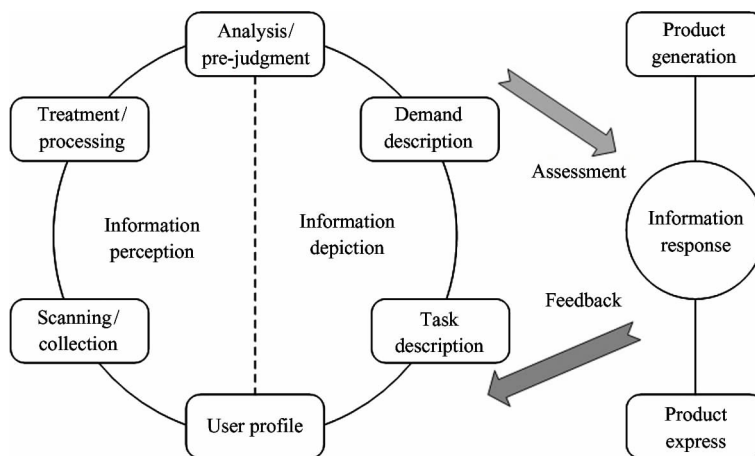

Fig. 2   The basic process of information perception

3   Strategy information analysis practices

The most important feature of Internet information is the richness and complexity of data content. Data is flatly transmitted at a fast speed, and everyone is both the receiver and provider of information. Data is not restricted by country and geographic space. The Internet is the source through which intelligence analysts mine information.

3.1   Network-monitoring system platform for strategy information analysis

In the Internet era, open source intelligence is the main source of data for information analysis[11-14]. With the public information on the official website of RAND

Corporation, taking NII (National Institute of Information), and JST(Japan Science and Technology Agency) as examples, this paper illustrates the process and methods of strategic information perception in the network environment. For the purpose of convenient acquisition, processing and analysis of network information, the authors of the paper have developed a network-monitoring platform to achieve information perception. The monitoring platform regularly grabs the content of specified websites and supports the automatic translation of Chinese, English and Japanese. Automatic statistical analysis of the terms on the web is made. The monitoring of hot topics based on 'hot words' as well as the automatic retrieval of their content is provided. Fig. 3 is the basic processing flow.
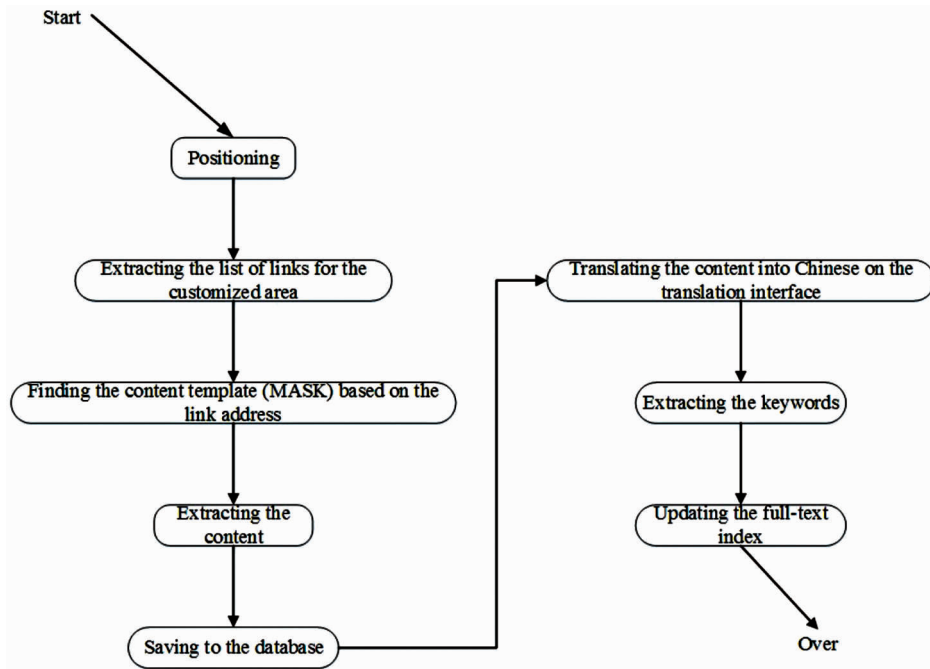
Fig. 3   The basic process of network data acquisition

3.2   Design of monitoring system model

Design of monitoring system model is given in Fig. 4. The main fields and descriptions of the program code are as follows. And the monitoring system interface is given in Fig. 5.

Nocontent Paper: document without content

(1) Id: document Id

(2) title: document title

(3) poster: publisher name

(4) postTime: release time;

format: yyyy-MM-dd HH:mm:ss

(5) sourceUrl: source link address

(6) tagArray: tag array

(7) fromPaperInfoEntity(): read data from PaperInfoEntity

FullPaper: document with content

(1) Id: document Id

(2) title: document title

(3) poster: publisher name

(4) postTime: release time;

format: yyyy-MM-dd HH:mm:ss

(5) content: document content

(6) origContent: original document content

(7) sourceUrl: source link address

(8) tagArray: tag array

(9) fromPaperInfoEntity(): read data from PaperInfoEntity

(10) fromPaperContentEntity(): read data from PaperContentEntity

PaperInfoEntity: Document information entity

(1) Id: document Id

(2) title: document title

(3) poster: publisher name

(4) postTime: release time

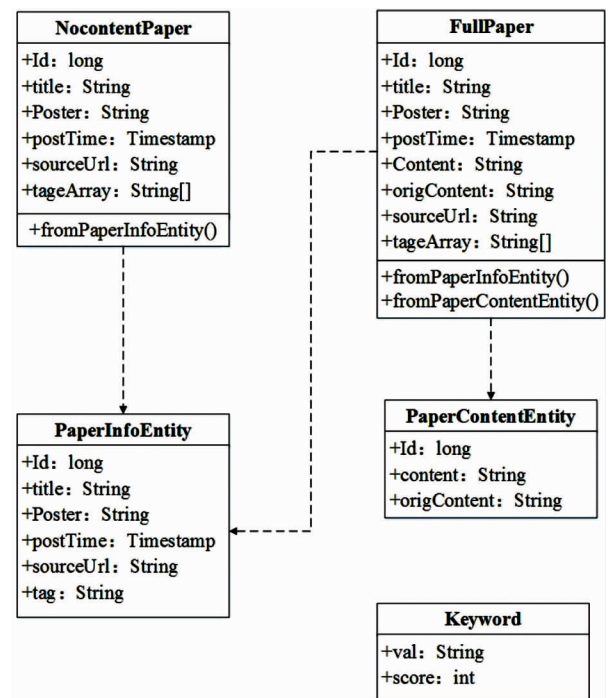(5) sourceUrl: source link address

(6) tag: tag array string

Fig. 4   Design of monitoring system model

Fig. 5 Interface of the monitoring system platform

PaperContentEntity：Document content entity
(7) Id：document Id
(8) content：document content
(9) origContent：original document content
Keyword：keywords
(1) val：keyword string
(2) score：score

3.3 Information perception results

RAND Corporation is used as an example to illustrate the collation and analysis of the information perception results. As a non-profit international research organization and the most important comprehensive military matters focused strategy research institution in the United States，RAND Corporation is famous for the research on advanced military science and technologies and major military strategies and then expands its research to internal and external policies. Now it is gradually developing into a comprehensive think tank studying politics，military matters，economy，technology，society and so on and thus is honored as a 'brain concentration camp' and 'super military college' of the modern brain trust，as well as the pioneer and spokesman of the world brain trust. RAND Corporation is the most prestigious decision-making consulting institution in the United States or even the world. Through the automatic monitoring of the system，the authors of this paper find the contents related to national security and scientific research in RAND Corporation's strategic research，as shown in Table 1.

Table 1   Classification of strategic studies of RAND corporation (related to national security，science and technology research)

Key research topics	Subtopics (bold fields are hot topics)
International affairs	**Developing countries**；arms proliferation and control；APEC；**EU**；**global security**；**Globalization**；**international diplomacy**；international education；international humanitarianism assistance；international laws；North Atlantic Treaty Organization；war laws；sea laws，etc.
National security	Air warfare；aircraft carriers；anti-satellite systems；anti-submarine warfare；biological weapons and warfare；chemical weapons and warfare；combat service support；counter terrorism；network warfare；national defense infrastructure；domestic intelligence；electronic intelligence；intelligence analysis；intelligence gathering；intelligence community；**military planning**；military communications system；military intelligence；**military procurement**；missile defense；terrorism and homeland security；national security legislation；nuclear weapons and warfare；special operations forces；UAVs；**weapons proliferation and control**；US troops；US marine corps；US Department of Defense；**war and military operations**；**threat assessment**，etc.
Science and Technology	Aerospace technology；agricultural sciences；astronomy；biology and life sciences；chemistry；artificial intelligence；automatic driving；communication technology；**computer and information science and technology**；big data；computer viruses；data mining；electronic cases；**emerging technologies**；global positioning system；information security；Earth science；mathematics；military technology；physics；**STEM education**；**science**，**technology and innovation policy**；**space science and technology**；UAVs，etc.
Terrorism and homeland security	Al-qaida；**border and port security**；biomonitoring；**bioterrorism**；chemical terrorism；infrastructure and transport；civil defense；civil-military relations；**anti-terrorism**；critical infrastructure protection；network warfare；domestic intelligence；domestic terrorism，homeland security，homeland security legislation，**terrorism risk assessment**；**terrorist organizations**；Islamism，etc.

4　Conclusion

　　The Internet has a history of more than 30 years. Throughout the Internet, the total number of web pages has exceeded 10 billion and a mass of web pages are added every day, of which an important proportion is about the strategic dynamics of science & technology and economy, and the content reports on scientific achievements. If these contents can be collected and traced in time, it is clearly understood the dynamics of the world, which can guide us in making reasonable attempts or responses[15]. This paper proposes some constructive research contents and methods, and carries out preliminary research and exploration work. Although the research work is still in its infancy, these research contents and methods will be the basis for further improvement of the research work. This paper argues that in order to match the information perception result with the information depiction result, it is necessary to further clarify the value of different types of information from different sources and their supporting roles in different categories and combine the matching results of information perception and information depiction with certain situations to provide reference for the decision makers and give full play to the initiative of the information work, or make supplement and adjustment according to specific task to achieve response in the shortest time, shorten the waiting time of users, and improve the information work and decision-making efficiency. Decision makers' further request for information products is fed back to the information perception and depiction links to generate new information products through re-matching. Decision makers' evaluation or adoption of information products may trigger new information perception concerns and make for more detailed information depiction.

References

[1] Heather J W, Ilana B. Defining second generation open source intelligence for the defense enterprise[EB/OL]. https://www. rand. org/pubs/research _ reports/RR1964. htm:RAND, 2018

[2] Zeng W. Research on scientific and technological information servicing for national science and technology security [J]. *Information studies*: *Theory & Application*, 2018,06:34-38(In Chinese)

[3] Li H. Research on the capability of science and technology information in the intelligence 3. 0 [J]. *Information studies*: *Theory & Application*, 2017, 40 (3) :1-4 (In Chinese)

[4] Hua X J. Introduction to Information Science[M]. Guangzhou: Sun Yat-sen University Press, 1990. 8-15 (In Chinese)

[5] Chen M H, Wang Y F. Analysis on conditions of information awareness [J/OL]. *Information Studies*: *Theory & Application*. http://kns. cnki. net/kcms/ detail/11. 1762. G3. 20180517. 0931. 006. html:CNKI, 2018 (In Chinese)

[6] Zhao K R, Wan Y F. Analysis on information awareness methods [J/OL]. *Information Studies*: *Theory & Application*. http://kns. cnki. net/kcms/detai l/11. 1762. G3. 201805220. 16350. 008. html:CNKI, 2018(In Chinese)

[7] Wang Y F, Zhao K R, Chen M H, et al. Analysis on the research of information awareness [J/OL]. *Information Studies*: *Theory & Application*. http://kns. cnki. net/kcm/detail/11. 1762. G3. 20180517. 0921. 002. html: CNKI, 2018(In Chinese)

[8] Gu J Z. Context awareness calculation[J]. *Journal of East China Normal University*: *Natural Science Edition*, 2009(5):1-20,145(In Chinese)

[9] Gartner. Context-aware security [EB/OL]. [2018-03-03]. https://www. gartner. com/it- glossary/ context-a-ware-security:GARTER, 2018

[10] Yan W Y, Guo L, Liao P, et al. Construction of the context awareness based active defense system of information safety[J]. *Electric Power Information and Communication Technology*, 2016, 14(1): 28-32

[11] Tian X J. Research on context awareness based mobile electronic resource recommendation technology[J]. *Information Studies*: *Theory & Application*, 2015, 38 (05): 86-89, 104(In Chinese)

[12] Hu G Y, Zhang B C, Zhou Z J, et al. Network Security Situation Awareness Based on Confidence Rule Base [M]. Beijing: Science Press, 2017: 1-2(In Chinese)

[13] Gong J, Yu X D, Su Q, et al. The overview of network security situation awareness [J]. *Journal of Software*, 2017, 28(4): 1010-1026(In Chinese)

[14] An J, Li X H, You C L, et al. The research of cyber situation awareness model[C]. In: Proceedings of the International Conference on Intelligent Systems and Interactive Applications, Shanghai, China, 2017. 232-238

[15] Richards J H. Psychology of Intelligence Analysis[M]. Washington: United States Government Printing Office, 1999. 8-25

　　Zeng Wen, born in 1973. She received her Ph. D degree in Shenyang Institute of Automation, Chinese Academy of Sciences in 2009. She also received her M. S. degree from the College of Information Science and Engineering at Northeastern University in 2003. Her research interests include information analysis technology, information theory and method.