INTELLIGENCE INFO

Knowledge Engineering in Intelligence Gathering

Nicolae Sfetcu

Publicat online: 18.10.2023

Knowledge Engineering in Intelligence Gathering

Ing. fiz. Nicolae Sfetcu, MPhil[1]

nicolae@sfetcu.com

Abstract

Intelligence gathering is an essential function for governments, security agencies, and organizations worldwide. It involves the collection, analysis, and dissemination of information critical to decision-making, security, and strategic planning. In this age of information abundance, the field of intelligence gathering has evolved, with knowledge engineering playing a pivotal role in ensuring efficient and effective information management and analysis. A process of intelligence gathering begins when a user enters a query into the system. In knowledge engineering, intelligence gathering consists in finding information from structured and unstructured sources in a way that must represent knowledge in a way that facilitates inference. This essay explores the significance of knowledge engineering in intelligence gathering, highlighting its applications, challenges, and future prospects.

Keywords: knowledge engineering, intelligence gathering, HUMINT, OSINT, intelligence agencies

Ingineria cunoașterii în colectarea informațiilor

Rezumat

Colectarea informațiilor este o funcție esențială pentru guverne, servicii de informații și organizații din întreaga lume. Aceasta implică colectarea, analiza și diseminarea informațiilor esențiale pentru luarea deciziilor, securitate și planificare strategică. În această epocă a abundenței informațiilor, domeniul culegerii de informații a evoluat, ingineria cunoașterii jucând un rol esențial în asigurarea unei gestionări și analize efective și eficiente a informațiilor. Un proces de colectare a informațiilor începe atunci

---

[1] Romanian Academy - Romanian Committee for the History and Philosophy of Science and Technology (CRIFST), History of Science Division (DIS)

când un utilizator introduce o interogare în sistem. În ingineria cunoașterii, colectarea de informații constă în găsirea de informații din surse structurate și nestructurate într-un mod care trebuie să reprezinte cunoașterea într-un mod care să faciliteze inferența. Acest eseu explorează importanța ingineriei cunoașterii în culegerea de informații, evidențiind aplicațiile, provocările și perspectivele de viitor ale acesteia.

Cuvinte cheie: ingineria cunoașterii, colectarea informațiilor, HUMINT, OSINT, servicii de informații

Introduction

Intelligence gathering is an essential function for governments, intelligence agencies, and organizations worldwide. It involves the collection, analysis, and dissemination of information critical to decision-making, security, and strategic planning. In this age of information abundance, the field of intelligence gathering has evolved, with knowledge engineering playing a pivotal role in ensuring efficient and effective information management and analysis. This essay explores the significance of knowledge engineering in intelligence gathering, highlighting its applications, challenges, and future prospects.

Knowledge engineering was defined by Edward Feigenbaum, and Pamela McCorduck as follows: (Feigenbaum and McCorduck 1984)

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

Currently, knowledge engineering refers to building, maintaining and developing knowledge-based systems. Knowledge engineering is related to mathematical logic, and

heavily involved in cognitive sciences and socio-cognitive engineering where knowledge is produced by socio-cognitive aggregates (especially human) and is structured according to our understanding of how human rationality and logic work.

Intelligence gathering

A process of intelligence gathering begins when a user enters a query into the system. Several objects can match the result of a query with different degrees of relevance. Most systems estimate a numeric value about how well each object matches the query and classify objects according to this value. Many researches have focused on practices of intelligence gathering. Much of this research was based on the work of Leckie, Pettigrew and Sylvain, who in 1996 carried out an extensive review of the information science literature on the search for information by professionals. The authors have proposed an analytical model of the behavior of search professionals seeking to be generalizable across the profession, thus providing a future research platform in the field. The model was designed to "prompt new insights... and give rise to more refined and applicable theories of information seeking." (Leckie, Pettigrew, and Sylvain 1996, 188) The distinctive sign of the intelligence activity is to find the type of information others want to conceal.

Investigations in the data collection process are aimed at enriching information, eliminating some doubts, or solving problems.

The process of intelligence gathering from people (abbreviated HUMINT) is achieved through interpersonal contacts. NATO defines HUMINT as "a category of intelligence derived from information collected and provided by human sources." (NATO 2018) Typical HUMINT activities consist of queries and conversations with people who have access to information. The way HUMINT operations are conducted is dictated by both the official protocol and the nature of the information source.

Sources may be neutral, friendly or hostile and may or may not be aware of their involvement in intelligence gathering.

The HUMINT gathering process involves selecting source people, identifying them and conducting interviews. The analysis of information can help with biographical and cultural information. Lloyd F. Jordan recognizes two forms of culture study, both of which are relevant to HUMINT. (Jordan 2008)

Information published in media around the world can be classified and treated as secret when it becomes an intelligence product. All sources are secret, and intelligence is defined to exclude open sources. (Robertson 1987)

Closed or secret sources involve "special means" to reach information, and the technique may include manipulation, interrogation, the use of technical devices, and extensive use of criminal methods. These techniques are costly, time consuming and labor intensive compared to open-source methods (OSINT). In some cases, hidden collection methods have a strong association with the criminal world. Noam Chomsky noted that there are good reasons why intelligence services are so closely linked to criminal activities. "Clandestine terror," he argued, "requires hidden funds, and the criminal elements to whom the intelligence agencies naturally turn expect a quid pro quo." (Chomsky 1992)

Coverage methods are complicated and dangerous but raise ethical and moral questions as well. A well-known technique, for example, is the manipulation of human agents to obtain information. The process, known as "the development of controlled sources," may involve extensive use of psychological manipulation, blackmail, and financial rewards. (Godfrey 1978) Intelligence gathering applying these techniques work in hostile environments. But intelligence, Sherman Kent argued, could be likened to familiar means of seeking the truth. (Kent 1966) Intelligence, unlike any other profession, does not work according to known moral or ethical standards. Some of these standards tend to be, at best, cosmetic. The argument is that anything vital to national survival is acceptable in any situation, even when the method provokes everything that is democratic. Clandestine operations remain unclear in international law and there is very little scientific research to cover the subject.

Knowledge Engineering

In knowledge engineering, intelligence gathering consists in finding information from structured and unstructured sources in a way that must represent knowledge in a way that facilitates inference. The result of the extraction goes beyond establishing structured information or transforming it into a relational scheme, requiring either reuse of existing formal knowledge (identifiers or ontologies) or generating a system based on source data. (Sfetcu 2016)

The role of knowledge engineering in intelligence fathering

- *Data Collection and Integration*: Knowledge engineering involves designing systems that can gather, integrate, and organize data from various sources, including open-source information, classified documents, social media, and human intelligence. Automated tools and algorithms assist in extracting relevant information from unstructured data, helping intelligence analysts access a comprehensive set of data to work with.
- *Knowledge Representation*: Intelligence agencies need to structure and represent knowledge in a format that is easily understandable and searchable. Knowledge engineering helps create knowledge graphs and ontologies that organize data, relationships, and entities. This structured information allows analysts to navigate and connect pieces of information effectively, uncovering patterns and trends.
- *Pattern Recognition and Analysis*: Knowledge engineering supports the development of machine learning and AI algorithms for pattern recognition. These algorithms can sift through vast datasets to detect anomalies, predict future events, and identify potential threats. This is especially crucial in counterterrorism, cybersecurity, and law enforcement operations.
- *Decision Support*: Intelligence professionals rely on knowledge engineering systems for decision support. These systems provide recommendations, based on historical data and real-time information, to assist analysts and decision-makers in making informed choices.

Challenges

Traditional information extraction is a natural language processing technology that extracts information from language texts and their typically natural structures in an appropriate way. The types of information to be identified must be specified in a model before the process starts, so the entire process of extracting traditional information is domain dependent. The extraction of intelligence is divided into the following five secondary tasks: (Cunningham 2006)

- *Named Entity Recognizing (REN)* - Recognizing and classifying all named entities contained in a text, using grammar-based methods or statistical models.
- *Coreference resolution (CO)* - identifies equivalent entities that have been recognized by REN in a text.
- *Construction of the template element (TE)* - identifies the descriptive properties of the entities, recognized by REN and CO.
- *Construction of the template relationship (TR)* - identifies the relationships that exist between the template elements.
- *Production of the script template (ST)* - will be identified and structured according to entities recognized by REN and CO and relationships identified by TR.

Challenges in knowledge engineering for intelligence gathering:

- *Data Overload*: The volume of data available for intelligence purposes has grown exponentially, making it difficult to manage, process, and analyze. Knowledge engineering must address the challenge of sifting through enormous amounts of data to find actionable intelligence.
- *Data Quality and Reliability*: The accuracy and reliability of data sources vary, making it essential to create robust systems for data validation. False or misleading information can have severe consequences in intelligence operations.
- *Security and Privacy Concerns*: Intelligence gathering often involves sensitive and classified information. Knowledge engineering systems must prioritize security and privacy to protect the integrity of collected data and the identities of sources.
- *Human-Machine Collaboration*: Balancing the roles of human intelligence analysts and knowledge engineering systems is a challenge. While automation is crucial for processing vast datasets, human expertise is irreplaceable in assessing the context and nuance of information.

In ontology-based information mining, at least one ontology is used to guide the process of extracting information from the text in natural language. The OBIE system uses traditional information extraction methods to identify concepts, cases and relationships of ontologies used in the text, which will be structured in an ontology after the process. Thus, entering ontologies is the model of information to be extracted. (Wimalasuriya and Dejing Dou 2010, 306–23) Ontology learning automates the process of constructing ontologies in natural language.

The Future

The future of knowledge engineering in intelligence gathering holds immense potential:

- *Enhanced Automation*: As AI and machine learning technologies continue to advance, knowledge engineering systems will become more sophisticated, allowing for greater automation in data collection, analysis, and decision support.
- *Improved Predictive Analytics*: With access to historical data and advanced algorithms, intelligence agencies will enhance their ability to predict future threats, aiding in proactive security measures.
- *Interoperability*: Efforts to improve the interoperability of different intelligence systems and agencies will enable more effective sharing of information while respecting privacy and security concerns.
- *Ethical Considerations*: As the field of knowledge engineering progresses, there will be increasing emphasis on ethical considerations, including data privacy, bias, and transparency in intelligence operations.

The discovery of knowledge involves an automatic process of searching for large volume data, using data mining, and based on similar methodologies and terminologies. (Wimalasuriya and Dejing Dou 2010, 306–23) Data mining creates abstractions of input

data, and the knowledge gained through the process can become additional data that can be used later. (Cao 2010)

Conclusion

Knowledge engineering is a cornerstone of modern intelligence gathering, aiding in data collection, integration, analysis, and decision support. While it comes with its challenges, such as data overload and security concerns, the future prospects for knowledge engineering in intelligence operations are promising. As technology continues to evolve, the role of knowledge engineering in intelligence gathering will be indispensable in addressing global security threats and maintaining the safety of nations and organizations.

Bibliography

Cao, Longbing. 2010. "(PDF) Domain-Driven Data Mining: Challenges and Prospects." ResearchGate. 2010. https://www.researchgate.net/publication/220073304_Domain-Driven_Data_Mining_Challenges_and_Prospects.

Chomsky, Noam. 1992. *Deterring Democracy*. Reissue edition. New York: Hill and Wang.

Cunningham, Hamish. 2006. "Information Extraction, Automatic." ResearchGate. 2006. https://www.researchgate.net/publication/228630298_Information_Extraction_Automatic.

Feigenbaum, Edward A., and Pamela McCorduck. 1984. *The Fifth Generation: Artificial Intelligence and Japan's Computer Challenge to the World*. New American Library.

Godfrey, E. Drexel. 1978. "Ethics and Intelligence." *Foreign Affairs*, 1978. https://www.foreignaffairs.com/articles/united-states/1978-04-01/ethics-and-intelligence.

Jordan, Lloyd F. 2008. "The Arab Mind by Raphael Patai. Book Review by Lloyd F. Jordan — Central Intelligence Agency." 2008. https://web.archive.org/web/20080213114422/https://www.cia.gov/library/center-for-the-study-of-intelligence/kent-csi/docs/v18i3a06p_0001.htm.

Kent, SHERMAN. 1966. *Strategic Intelligence for American World Policy*. Princeton University Press. https://www.jstor.org/stable/j.ctt183q0qt.

Leckie, Gloria J., Karen E. Pettigrew, and Christian Sylvain. 1996. "Modeling the Information Seeking of Professionals: A General Model Derived from Research on Engineers, Health Care Professionals, and Lawyers." ResearchGate. 1996. https://www.researchgate.net/publication/237440858_Modeling_the_Information _Seeking_of_Professionals_A_General_Model_Derived_from_Research_on_Eng ineers_Health_Care_Professionals_and_Lawyers.

NATO. 2018. "AAP-6 - NATO Glossary of Terms and Definitions." https://apps.dtic.mil/dtic/tr/fulltext/u2/a574310.pdf.

Robertson, K. G. 1987. "Intelligence Requirements for the 1980s." *Intelligence and National Security* 2 (4): 157–67. https://doi.org/10.1080/02684528708431921.

Sfetcu, Nicolae. 2016. *CunoaștereșiInformații*. Nicolae Sfetcu.

Wimalasuriya, Daya C., and Dejing Dou Dejing Dou. 2010. "(PDF) Ontology-Based Information Extraction: An Introduction and a Survey of Current Approaches." ResearchGate. 2010. https://www.researchgate.net/publication/220195792_Ontology-based_information_extraction_An_Introduction_and_a_survey_of_current_approaches.