

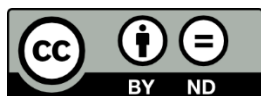
Epistemologia colectării informațiilor

Nicolae Sfetcu

4februarie2019

Sfetcu, Nicolae, "Epistemologia colectării informațiilor", SetThings (4februarie 2019), URL = <https://www.telework.ro/ro/epistemologia-colectarii-informatiilor/>

Email: nicolae@sfetcu.com



Acest articol este sub licență Creative Commons Attribution-NonCommercial 4.0 International. Pentru a vedea o copie a acestei licențe, vizitați <http://creativecommons.org/licenses/by-nd/4.0/>.

Extras din:

Sfetcu, Nicolae, "Epistemologia serviciilor de informații", SetThings (4februarie 2019), MultiMedia Publishing (ed.), DOI: 10.13140/RG.2.2.19751.39849, ISBN 978-606-033-160-5, URL = <https://www.telework.ro/ro/e-books/epistemologia-serviciilor-de-informatii/>

Epistemologia colectării informațiilor

Un proces de colectare a informațiilor începe atunci când un utilizator introduce o interogare în sistem. Mai multe obiecte se pot potrivi rezultatului unei interogări, cu diferite grade de relevanță. Cele mai multe sisteme estimează o valoare numerică despre cât de bine se potrivește fiecare obiect cu interogarea, și clasifică obiectele în funcție de această valoare. Multe cercetări s-au concentrat pe practicile de căutare a informațiilor. O mare parte din aceste cercetări s-au bazat pe activitatea desfășurată de Leckie, Pettigrew (acum Fisher) și Sylvain, care în 1996 au efectuat o revizuire extensivă a literaturii științei informațiilor privind căutarea informațiilor de către profesioniști. Autorii au propus un model analitic al comportamentului profesioniștilor care caută informații, destinat să fie generalizabil în întreaga profesie, oferind astfel o platformă de cercetare viitoare în domeniu. Modelul a fost destinat să "descopere noi perspective... și să

dea naștere unor teorii mai rafinate și aplicabile ale căutării informațiilor".(Leckie, Pettigrew, and Sylvain 1996, 188) Semnul distinctiv al activității de informații este acela de a găsi tipul de informații pe care alții vor să le ascundă.

Ingenieria cunoașterii a fost definită de Edward Feigenbaum, și Pamela McCorduck astfel: (Feigenbaum and McCorduck 1984)

"Ingenieria cunoașterii este disciplină în inginerie care implică integrarea cunoașterii într-un sistem informatic pentru a rezolva probleme complexe care necesită în mod normal un înalt nivel de expertiză umană."

În prezent, ingineria cunoașterii se referă la construirea, menținerea și dezvoltarea unor sisteme bazate pe cunoaștere. Ingineria cunoașterii este legată de logica matematică, și puternic implicată în științele cognitive și ingineria socio-cognitivă unde cunoașterea este produsă de agregate socio-cognitive (în special umane) și este structurată conform cu înțelegerea noastră despre cum lucrează raționalitatea și logica umană.

În ingineria cunoașterii, extragerea cunoașterii constă în stabilirea de cunoașteri din surse structurate și nestructurate într-un mod care trebuie să reprezinte cunoaștere într-un mod care să faciliteze inferența. Rezultatul extracției merge dincolo de stabilirea unor informații structurate sau transformarea într-o schemă relațională, impunând fie reutilizarea cunoștințelor existente formale (identificatori sau ontologii) fie generarea unui sistem bazat pe datele sursă. (Sfetcu 2016)

Extragerea de informații tradițională este o tehnologie de procesare a limbajului natural, care extrage informații din texte de limbaj și structurile acestora în mod tipic natural, într-un mod adecvat. Tipurile de informații care urmează să fie identificate trebuie să fie specificate într-un model înainte de a începe procesul, deci întregul proces de extragere a informațiilor tradițională este dependentă de domeniu. Extragerea de informații este împărțită în următoarele cinci sarcini secundare: (Cunningham 2006)

- *Recunoașterea entității numite (REN)* - recunoașterea și clasificarea tuturor entităților numite conținute într-un text, prin metode bazate pe gramatică sau modele statistice.
- *Rezoluția coreferenței (CO)* - identifică entități echivalente, care au fost recunoscute de către REN, într-un text.
- *Construcția elementului șablon (ES)* - se identifică proprietățile descriptive ale entităților, recunoscute de REN și CO
- *Construcția relației șablon (RS)* - se identifică relațiile care există între elementele șablon.
- *Producția scenariu șablon (SS)* - vor fi identificate și structurate în funcție de entități, recunoscute de REN și CO și relații, identificate prin RS.

În extracția de informații bazată pe ontologie, cel puțin o ontologie este utilizată pentru a ghida procesul de extragere a informației din textul din limbajul natural. Sistemul OBIE utilizează metode de extragere a informației tradiționale pentru a identifica concepte, cazuri și relațiile ontologiilor utilizate în text, care vor fi structurate într-o ontologie după proces. Astfel, intrarea ontologiilor constituie modelul de informații care trebuie extrase. (Wimalasuriya and Dejing Dou 2010, 306–23) Învățarea ontologiei automatizează procesul de construcție a ontologiilor din limbajul natural.

Informațiile publicate în mass-media din întreaga lume pot fi clasificate și tratate ca fiind secrete în momentul în care devin un produs de informații. Toate sursele sunt secrete, iar activitatea de informații este definită astfel încât să excludă sursele deschise. (Robertson 1987)

Sursele închise sau cele secrete implică "mijloace speciale" pentru a ajunge la informații, iar tehnica poate include manipularea, interogatorii, utilizarea dispozitivelor tehnice, precum și o utilizare extensivă a metodelor criminale. Aceste tehnici sunt considerate a fi costisitoare,

consumatoare de timp și laborioase comparativ cu metodele în sursă deschisă. În unele cazuri, metodele de colectare ascunse au o asociere puternică cu lumea criminală. Noam Chomsky a observat că există motive întemeiate pentru care serviciile de informații sunt atât de strâns legate de activitățile criminale.

”Teroarea clandestină”, a susținut el, ”necesită fonduri ascunse și elemente criminale, la care agențiile de informații se transformă în mod natural, se așteaptă la un quid pro quo”. (Chomsky 1992)

Descoperirea cunoașterii implică un proces automat de căutare a informațiilor în volume mari de date, folosind mineritul de date, și bazându-se pe metodologii și terminologii asemănătoare. (Wimalasuriya and Dejing Dou 2010, 306–23) Mineritul de date creează abstracții ale datelor de intrare, cunoștințele obținute prin proces putând deveni date suplimentare care pot fi utilizate ulterior. (Cao 2010)

Investigațiile în procesul de colectare a datelor au scopul de îmbogățire a informațiilor, eliminarea unor dubii, sau rezolvarea unor probleme.

Procesul colectării informațiilor de la oameni (abreviat HUMINT) se realizează prin intermediul contactelor interpersonale. NATO definește HUMINT ca "o categorie de informații derivată din informațiile colectate și furnizate de surse umane". (NATO 2018) Activitățile tipice HUMINT constau în interogări și conversații cu persoane care au acces la informații. Modul în care se desfășoară operațiunile HUMINT este dictat atât de protocolul oficial, cât și de natura sursei informațiilor.

Sursele pot fi neutre, prietenoase sau ostile și pot sau nu să fie conștiente de implicarea lor în colectarea de informații.

Procesul colectării HUMINT implică selectarea persoanelor sursă, identificarea lor și efectuarea de interviuri. Analiza informațiilor poate ajuta prin informații biografice, și culturale.

Lloyd F. Jordan recunoaște două forme de studiu al culturii, ambele fiind relevante pentru HUMINT. (Jordan 2008)

Metodele de acoperire sunt complicate și pline de primejdie, dar ridică și întrebări etice și morale. O tehnică bine cunoscută, de exemplu, este manipularea agenților umani pentru a obține informațiile. Procesul, cunoscut sub numele de "dezvoltarea surselor controlate", poate implica o utilizare extensivă a manipulării psihologice, a șantajului și a recompenselor financiare. (Godfrey 1978) Colectorii de informații care aplică aceste tehnici acționează în medii ostile. Dar activitatea de informații, a argumentat Sherman Kent, ar putea fi asemănată cu mijloacele familiare de căutare a adevărului. (Kent 1966) Activitatea de informații, spre deosebire de orice altă profesie, nu funcționează conform unor standarde morale sau etice cunoscute. Unele dintre aceste standarde tind să fie, în cel mai bun caz, cosmetizate. Argumentul este că orice lucru vital pentru supraviețuirea națională este acceptabil, în orice situație, chiar și atunci când metoda provoacă tot ceea ce este democratic. Operațiunile clandestine rămân în continuare neclare în dreptul internațional și există foarte puțină cercetare științifică pentru acoperirea subiectului.

Bibliografie

- 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_Engineers_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. *Cunoașterea și Informații*. 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.