

La collecte du renseignement

Nicolae Sfetcu

12.10.2019

Sfetcu, Nicolae, « La collecte du renseignement », SetThings (12 octobre 2019), URL = <https://www.setthings.com/fr/la-collecte-du-renseignement/>

Email: nicolae@sfetcu.com



Cet article est sous licence Creative Commons Attribution-NoDerivatives 4.0 International. Pour voir une copie de cette licence, visitez <http://creativecommons.org/licenses/by-nd/4.0/>.

Une traduction partielle de :

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

Un processus de collecte du renseignement commence lorsqu'un utilisateur introduit une requête dans le système. Plusieurs objets peuvent être associés au résultat d'une requête, avec différents degrés de pertinence. La plupart des systèmes estiment une valeur numérique indiquant dans quelle mesure chaque objet correspond à la requête, et classent les objets en fonction de cette valeur. De nombreuses recherches se sont concentrées sur les pratiques de recherche du renseignement. Une grande partie de ces recherches reposait sur les travaux de Leckie, Pettigrew (maintenant Fisher) et Sylvain, qui, en 1996, ont procédé à un examen approfondi de la littérature scientifique sur la recherche du renseignement par des professionnels. Les auteurs ont proposé un modèle analytique du comportement des professionnels en quête du renseignement, destiné à être généralisable dans l'ensemble de la profession, offrant ainsi une future plate-forme de recherche sur le terrain. Le modèle visait à « découvrir de nouvelles perspectives ... et à donner naissance à

des théories plus raffinées et applicables de la recherche du renseignement ». (Leckie, Pettigrew, and Sylvain 1996, 188) Le signe distinctif de l'activité du renseignement est la recherche du type de renseignement que d'autres veulent cacher.

Edward Feigenbaum et Pamela McCorduck ont défini l'ingénierie de la connaissance comme suit : (Feigenbaum and McCorduck 1984)

« L'ingénierie de la connaissance est la discipline de l'ingénierie qui implique l'intégration des connaissances dans un système informatique afin de résoudre des problèmes complexes qui nécessitent normalement un niveau élevé d'expertise humaine. »

À l'heure actuelle, l'ingénierie de la connaissance fait référence à la construction, à la maintenance et au développement de systèmes basés sur les connaissances. L'ingénierie de la connaissance est liée à la logique mathématique et fortement impliquée dans les sciences cognitives et l'ingénierie socio-cognitive. La connaissance est produite par des agrégats socio-cognitifs (en particulier humains) et structurée en fonction de notre compréhension du fonctionnement de la rationalité et de la logique humaines.

Dans l'ingénierie de la connaissance, l'extraction des connaissances consiste à établir des connaissances à partir des sources structurées et non structurées d'une manière qui doit représenter les connaissances de manière à faciliter l'inférence. Le résultat de l'extraction va au-delà de l'établissement d'informations structurées ou de leur transformation en un schéma relationnel, nécessitant soit la réutilisation des connaissances formelles existantes (identificateurs ou ontologies), soit la création d'un système basé sur les données source. (Sfetcu 2016)

La collecte traditionnelle du renseignement est une technologie de traitement du langage naturel, qui extrait les informations des textes de langage et de leurs structures typiquement naturelles de manière appropriée. Les types d'informations à identifier doivent être spécifiés dans un modèle avant de démarrer le processus, de sorte que l'ensemble du processus de collecte

traditionnelle du renseignement dépend du domaine. La collecte du renseignement est divisée en cinq tâches secondaires suivantes : (Cunningham 2006)

- *Reconnaissance des entités nommées* - reconnaissance et classification de toutes les entités nommées contenues dans un texte, à l'aide de méthodes basées sur des modèles grammaticaux ou statistiques.
- *Résolution de la coréférence* - identifie les entités équivalentes, qui ont été reconnues par la reconnaissance des entités nommées, dans un texte.
- *Construction de l'élément modèle* - les propriétés descriptives des entités, reconnues par la reconnaissance des entités nommées et la résolution de la coréférence, sont identifiées
- *Construction de la relation de modèle* - identifie les relations existantes entre les éléments de modèle.
- *Production du scénario modèle* - sera identifié et structuré en fonction des entités, reconnu par des entités nommées et résolution de la coréférence et des relations identifiées par la construction de la relation de modèle.

Dans la collecte du renseignement basée sur une ontologie, au moins une ontologie est utilisée pour guider le processus de collecte du renseignement à partir de texte en langage naturel. Le système OBIE utilise des méthodes de collecte traditionnelles du renseignement pour identifier les concepts, les cas et les relations des ontologies utilisées dans le texte, qui seront structurés dans une ontologie après le processus. Ainsi, la saisie d'ontologies est le modèle d'information à extraire. (Wimalasuriya and Dejing Dou 2010, 306–23) L'apprentissage d'ontologies automatise le processus de construction d'ontologies en langage naturel.

Les informations publiées dans les médias du monde entier peuvent être classées et traitées comme des secrets lorsqu'elles deviennent un produit d'information. Toutes les sources sont

secrètes et l'activité de renseignement est définie de manière à exclure les sources ouvertes. (Robertson 1987)

Les sources fermées ou secrètes impliquent des « moyens spéciaux » pour accéder à l'information, et la technique peut inclure la manipulation, l'interrogation, l'utilisation des dispositifs techniques et l'utilisation extensive des méthodes criminelles. Ces techniques sont considérées comme coûteuses, fastidieuses et laborieuses par rapport aux méthodes open source. Dans certains cas, les méthodes de collecte cachées sont fortement associées au monde criminel. Noam Chomsky a indiqué qu'il existait de bonnes raisons pour lesquelles les services de renseignement sont si étroitement liés aux activités criminelles.

« La terreur clandestine », a-t-il déclaré, « nécessite des fonds cachés et des éléments criminels, les agences de renseignement se tournant naturellement vers le quid pro quo ». (Chomsky 1992)

La découverte de la connaissance implique un processus automatique de recherche d'informations dans de grands volumes de données, à l'aide de l'exploration de données, sur la base de méthodologies et de terminologies similaires. (Wimalasuriya and Dejing Dou 2010, 306–23)

La collecte du renseignement crée des abstractions des données d'entrée. Les connaissances acquises au cours du processus peuvent devenir des données supplémentaires pouvant être utilisées plus tard. (Cao 2010)

Les enquêtes en cours de collecte du renseignement visent à enrichir le renseignement, à éliminer les doutes ou à résoudre les problèmes.

Le processus de collecte du renseignement auprès de personnes (abrégié HUMINT) se fait par le biais des contacts interpersonnels. L'OTAN définit HUMINT comme « une catégorie de du renseignement tirées de l'informations recueillies et fournies par des sources humaines ». (NATO 2018) Les activités HUMINT typiques consistent en des requêtes et des conversations avec des

personnes ayant accès à des informations. La manière dont les opérations HUMINT sont effectuées est dictée à la fois par le protocole officiel et par la nature de la source de l'information.

Les sources peuvent être neutres, amicales ou hostiles et peuvent ne pas être conscientes de leur implication dans la collecte du renseignement.

Le processus de collecte HUMINT consiste à sélectionner les personnes source, à les identifier et à mener des entretiens. L'analyse du renseignement peut aider avec des informations biographiques et culturelles. Lloyd F. Jordan reconnaît deux formes d'étude de la culture, toutes deux pertinentes pour HUMINT. (Jordan 2008)

Les méthodes de couverture sont compliquées et très risquées, mais elles soulèvent également des questions éthiques et morales. Une technique bien connue, par exemple, consiste à manipuler des agents humains pour obtenir des informations. Le processus, connu sous le nom de « développement de source contrôlé », peut impliquer un recours intensif à la manipulation psychologique, au chantage et à des récompenses financières. (Godfrey 1978) Les collecteurs de renseignement qui appliquent ces techniques opèrent dans des environnements hostiles. Selon Sherman Kent, l'activité de renseignement pourrait s'apparenter à un moyen familier de recherche de la vérité. (Kent 1966) L'activité de renseignement, contrairement à toute autre profession, ne fonctionne pas selon des normes morales ou éthiques connues. Certaines de ces normes ont tendance à être, au mieux, cosmétiques. L'argument est que tout élément vital à la survie nationale est acceptable dans n'importe quelle situation, même lorsque la méthode provoque tout ce qui est démocratique. Les opérations clandestines restent floues en droit international et très peu de recherches scientifiques couvrent ce sujet.

Bibliographie

- 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. *Cunoaștere 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.