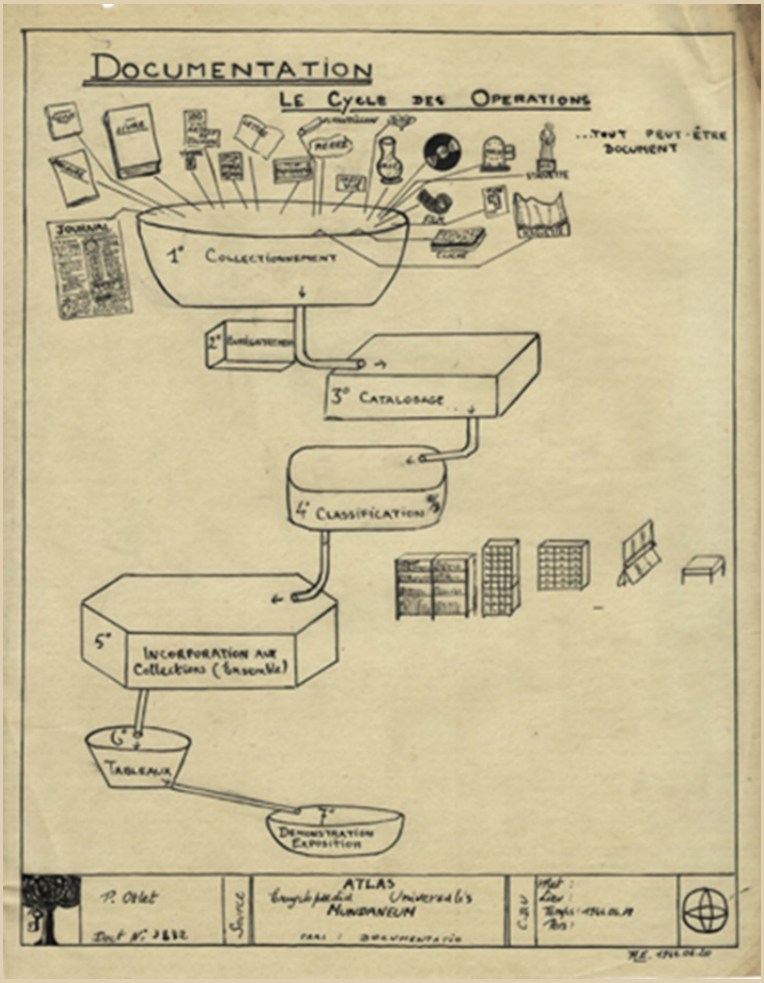
AIDAinformazioni

RIVISTA SEMESTRALE DI SCIENZE DELL'INFORMAZIONE

NUMERO 3-
ANNO 43
LUGLIO-DICEMBRE 2025



CACUCCI EDITORE
BARI

AIDAinformazioni

Rivista semestrale di Scienze dell'Informazione

Fondata nel 1983 da Paolo Bisogno

Proprietario della rivista:
Università della Calabria

Direttore Scientifico:
Roberto Guarasci, *Università della Calabria*

Direttore Responsabile:
Fabrizia Flavia Sernia

Comitato scientifico:
Anna Rovella, *Università della Calabria*;
Maria Guercio, *Sapienza Università di Roma*;
Giovanni Adamo, *Consiglio Nazionale delle Ricerche* †;
Claudio Gnoli, *Università degli Studi di Pavia*;
Ferruccio Diozzi, *Centro Italiano Ricerche Aerospaziali*;
Gino Roncaglia, *Università della Tuscia*;
Laurence Favier, *Université Charles-de-Gaulle Lille 3*;
Madjid Ihadjadene, *Université Vincennes-Saint-Dénis Paris 8*;
Maria Mirabelli, *Università della Calabria*;
Agustín Vivas Moreno, *Universidad de Extremadura*;
Douglas Tudhope, *University of South Wales*;
Christian Galinski, *International Information Centre for Terminology*;
Béatrice Daille, *Université de Nantes*;
Alexander Murzaku, *College of Saint Elizabeth*, USA;
Federico Valacchi, *Università di Macerata.*

Comitato di redazione:
Antonietta Folino, *Università della Calabria*;
Erika Pasceri, *Università della Calabria*;
Maria Taverniti, *Consiglio Nazionale delle Ricerche*;
Maria Teresa Chiaravalloti, *Consiglio Nazionale delle Ricerche*;
Assunta Caruso, *Università della Calabria*;
Claudia Lanza, *Università della Calabria.*

Segreteria di Redazione:
Valeria Rovella, *Università della Calabria*

Editrice: Cacucci Editore S.a.s.
Via D. Nicolai, 39 – 70122 Bari (BA)
www.cacuccieditore.it
e-mail: riviste@cacuccieditore.it
Telefono 080/5214220

AIDAinformazioni

Rivista semestrale di Scienze dell'Informazione

«AIDAinformazioni» è una rivista scientifica che pubblica articoli inerenti alle Scienze dell'Informazione, alla Documentazione, all'Archivistica, alla Gestione Documentale e all'Organizzazione della Conoscenza ma amplia i suoi confini in ulteriori campi di ricerca affini quali la Terminologia, la Linguistica Computazionale, la Statistica Testuale, ecc. È stata fondata nel 1983 quale rivista ufficiale dell'Associazione Italiana di Documentazione Avanzata e nel febbraio 2014 è stata acquisita dal Laboratorio di Documentazione dell'Università della Calabria. La rivista si propone di promuovere studi interdisciplinari oltre che la cooperazione e il dialogo tra profili professionali aventi competenze diverse, ma interdipendenti. I contributi pubblicati affrontano questioni teoriche, metodologie adottate e risultati ottenuti in attività di ricerca o progettuali, definizione di approcci metodologici originali e innovativi, analisi dello stato dell'arte, ecc.

«AIDAinformazioni» è riconosciuta dall'ANVUR come rivista di Classe A per l'Area 11 – Gruppo Scientifico Disciplinare 11/HIST-04 – Scienze del libro, del documento e storico-religiose e come rivista scientifica per le Aree 10 – Scienze dell'antichità, filologico-letterarie e storico-artistiche; 11 – Scienze storiche, filosofiche, pedagogiche e psicologiche; 12 – Scienze giuridiche; 14 – Scienze politiche e sociali. È anche annoverata dall'ARES (Agence d'évaluation de la recherche et de l'enseignement supérieur) tra le riviste scientifiche dell'ambito delle Scienze dell'Informazione e della Comunicazione. La rivista è, inoltre, indicizzata in: ACNP – Catalogo Italiano dei Periodici; BASE –Bielefeld Academic Search Engine; ERIH PLUS – European Reference Index for the Humanities and Social Sciences – EZB – Elektronische Zeitschriftenbibliothek – Universitätsbibliothek Regensburg; Gateway Bayern; KVK – Karlsruhe Virtual Catalog; The Library Catalog of Georgetown University; SBN – Italian union catalogue; Ulrich's; Union Catalog of Canada; LIBRIS – Union Catalogue of Swedish Libraries; Worldcat.

I contributi sono valutati seguendo il sistema del *double blind peer review*: gli articoli ricevuti sono inviati in forma anonima a due referee, selezionati sulla base della loro comprovata esperienza nei topics specifici del contributo in valutazione.

AIDAinformazioni

Anno 43

N. 3-4 – luglio-dicembre 2025

CACUCCI EDITORE

BARI

PROPRIETÀ LETTERARIA RISERVATA

© 2025 Cacucci Editore – Bari
Via Nicolai, 39 – 70122 Bari – Tel. 080/5214220
http://www.cacucciedi tore.it e-mail: info@cacucci.it

Ai sensi della legge sui diritti d'Autore e del codice civile è vietata la riproduzione di questo libro o di parte di esso con qualsiasi mezzo, elettronico, meccanico, per mezzo di fotocopie, microfilms, registrazioni o altro, senza il consenso dell'autore e dell'editore.

Sommario

Giovanna Aracri, Antonietta Folino, Maria Teresa Guaglianone, Marco Manna, Identification of thesaurus relationships with the support of ChatGPT and Gemini 9

Alessandra Bonsignori, Terminologia giuridica e organizzazione della conoscenza 51

Maria Teresa Chiaravalloti, Maria Teresa Guaglianone, Elena Cardillo, Metadatazione delle tipologie di documenti clinici: un approccio alla standardizzazione mediante LOINC 73

Carla Beatriz Marques Felipe, Carla Maria Martellote Viola, Raimunda Fernanda dos Santos, Domain Analysis on Data Management 91

Claudio Gnoli, Emanuela Valenzano, Che faccetta sei? 111

Manuela Grillo, Strumenti di convergenza tra archivi e biblioteche: i thesauri 121

Bianca La Manna, Rappresentare gli atti notarili medievali in Linked Open Data (LOD) 141

Claudia Lanza, L'univocità e la variabilità terminologica a supporto della rappresentatività delle risorse semantiche di organizzazione e rappresentazione della conoscenza specialistica 163

Larissa Moraes Martins, Carlos Eduardo da Silva Carvalho, Ana Cristina de Albuquerque, Mario Barité, Possibilities for using cultural warrant to complement literary work 181

Bianca Therezinha Carvalho Panisset, Lucia Maria Velloso de Oliveira, Leticia Souza da Costa Sampaio, Camilla Campoi de Sobral, Identification of types of documents and knowledge organization: transparency and access in records management 197

SoYoung Yoon, Claudio Gnoli, Ziyoung Park, Applying Integrative Levels Classification to Korean KOS Registry 211

Emanuela Nicole Donato, Grazia Serratore, Maria Taverniti, Maria Teresa Chiaravalloti, L'evoluzione della struttura documentale nelle cartelle cliniche psichiatriche dell'archivio storico del manicomio di Girifalco (1881-1978) 237

Lorenza Talarico, Manuscripts in the Digital Age 263

Maria Lígia Triques, Ananda Fernanda de Jesus, Mario Barité, José Eduardo Santarem Segundo, Ana Cristina de Albuquerque, Vocabularies in the digital context 275

Marcin Trzmielewski, Context-Driven Use of Automatic Tools in Medical Ontology Design: A Case Study in Allergology 291

Carla Maria Martellote Viola, Carla Beatriz Marques Felipe, Exploring the relationships between the Cross-Cutting Women's Agenda - Multi-Year Plan 2024-2027 and the Knowledge Organization of legislative proposals under processing on women in Brazil 313

Elena Cencetti, Camillo Carlo Pellizzari di San Girolamo, Elisabetta Viti, Modellizzazione delle conoscenze: Thesaurus Nuovo soggettario e Wikidata, due strumenti a confronto 325

Pontish Yeramyan, Alexander Murzaku, Nick Uva, Curt Anderson, The transferability of brilliance 353

Ana Flávia Dias Zammataro, Ana Cristina de Albuquerque, Giulia Crippa, Lucia Sardo, Carlo Bianchini, Maria Lígia Triques, ISKO Italy in perspective 375

Contributi

AIDAinformazioni
ISSN 1121–0095
ISBN 979-12-5965-639-1
DOI 10.57574/5965639115
pag. 291-312 (luglio-dicembre 2025)
Open access article licensed under CC-BY-NC-SA

Context-Driven Use of Automatic Tools in Medical Ontology Design: A Case Study in Allergology

Marcin Trzmielewski*

Abstract: The aim of this paper is to show a context-driven use of automatic tools for medical ontology design. Following an analysis of the limitations of the content-driven and quantitative approaches currently used by knowledge organization specialists, we propose a contextualized and qualitative approach. This approach enables the definition of user-oriented resources and criteria for conducting semi-automatic terminology extraction and annotation tasks, ultimately supporting the development of a useful allergology ontology. At the same time, it reveals a paradox: although automatic tools are intended to reduce time-consuming processes, qualitative methods require considerable time and effort to collect and analyze contextual data.

Keywords: Medical Ontologies, Terminology Extraction, Document Annotation, Document Indexing, Contextualized Approach.

1. Introduction

Due to the evolution of documentary environments and the information practices of actors in health, the use of automatic tools to design medical knowledge organization systems (MKOSs) appears to be a necessity. In numerous cases, automatic processes are *content-driven*, which can be reductive when developing systems intended to anticipate the needs of future users and effectively support their activities. Therefore, our aim is to show a *context-driven* use of extraction and annotation tools for processing allergology documentation, as part of the ontology design process for the Allergy Unit of Montpellier University Hospital. Our previous work has shown that the conceptualization and terminology of existing MKOSs in this domain – The Allergy Ontology, The Allergy Information Ontology for Enlightening People, The AllergyIntolerance, and The Allergy Detector II – do not adequately

* Laboratoire d'Etudes et de Recherches Appliquées en Sciences Sociales (LERASS), Université de Montpellier Paul-Valéry, France. marcin.trzmielewski@univ-montp3.fr. ORCID: 0000-0001-7436-2846.

2025, The Author(s). This is an open access article, free of all copyright, that anyone can freely read, download, copy, distribute, print, search, or link to the full texts or use them for any other lawful purpose. This article is made available under a Creative Commons Attribution Share-Alike 4.0 International License. AIDAinformazioni is a journal of the Department of Culture, Education and Society, University of Calabria, Rende (CS), Italy, published by Cacucci Editore, Bari, Italy.

represent the reality and knowledge of allergology, nor are they sufficiently adapted to the daily practices of specialists (Trzmielewski 2023; 2024). Thus, the need to design a new MKOS has become evident.

The present article focuses on the use of automatic tools in the design of MKOSs, discussing current methods and their limitations, as well as the theoretical and methodological approaches applied in the design of the ALLERGIDOC Ontology. It also presents the results of our study. Finally, we discuss the contributions and limitations of the proposed approach and offer perspectives for the representation and organization of allergology knowledge.

2. The use of automatic tools in the design of medical knowledge organization systems

Following the integration of information and communication technologies into various professional and everyday-life social contexts, both human information interactions and health-related documentation have grown increasingly complex (Trzmielewski 2023). Consequently, information specialists must consider how to design MKOSs capable of representing and organizing vast and heterogeneous bodies of knowledge, while simultaneously supporting diverse information practices across multiple contexts and user groups (Trzmielewski 2024). Although intellectual techniques such as interviews, surveys, thematic analysis, user-oriented approaches, and concept theory-based analyses are widely employed for this purpose (Almeida and Farinelli 2017; Iyer and Raghavan 2018; Oh et al. 2015; Raghavan and Sajana 2010), the abundance and heterogeneity of the available data have led to the consideration of automated techniques – especially in ontology design – aimed at minimizing tedious, repetitive, time-consuming, and costly tasks (Hassan 2017; Richard 2017; Toro et al. 2024). For these reasons, over the past three decades, natural language processing, text mining, and machine learning have been employed to extract, compare, map, and structure medical terminology – i.e., candidate and validated terms and definitions – from large textual corpora, as well to match terms with corresponding words and definitions (Baneyx and Charlet 2006; Choi 2016; Lopes et al. 2009; Muresan and Klavans 2013; Soergel et al. 2004; Tapi Nzali et al. 2018). More recently, large language models and retrieval-augmented generation techniques have begun to be integrated into medical ontology editing tasks, including relationship prediction, term definition generation, and the formulation of logical and conceptual definitions (Liu et al. 2024; Toro et al. 2024).

However, the use of automatic tools is frequently associated with ontological (empirical) (Tennis 2012) and literary warrant (Barité 2018) approaches to knowledge organization and representation. These approaches typically in-

volve the design of MKOSs under laboratory conditions and methods, often based on limited samples that are not drawn from the actual document collections managed by identified organizations and intended to be organized and sought by intended users (Trzmielewski 2024). It also tends to focus on testing only the internal structure of the developed systems. These methods tend to simplify reality and overlook the human and socio-organizational aspects related to the use of documentation resources and info-communication tools in concrete situations of use. Designers provide a set of a priori, content-related quantitative assessment criteria, such as the relevance of extracted terms – measured by frequency, inverse document frequency (IDF), and C-value – or the performance of MKOSs in accomplishing document processing tasks, evaluated through recall and precision scores (Baneyx and Charlet 2006; Lopes et al. 2009; Tang et al. 2017; Tapi Nzali et al. 2018). Even though specialists still involve domain experts – most often physicians – and information professionals to validate certain content-related aspects of terminology and conceptualization (through consensus and quality assessments), potential and end users are consistently not included in this process, largely due to the assumption that they lack the necessary informational competencies (Baneyx and Charlet 2006; Tapi Nzali et al. 2018; Toro et al. 2024). As a result, little is known about their needs, and their criteria of usefulness are not taken into consideration.

Current practices appear to be reductive, and the use of automatic tools should be aligned with epistemological approaches to knowledge organization (Tennis 2012), as well as with qualitative methodology (Cisek 2013). Accordingly, a contextual and qualitative approach was adopted for the design of the ALLERGIDOC Ontology discussed in this paper. Instead of relying on a priori, content-driven quantitative indicators, we aimed to inductively identify criteria that were "external to the corpus" to propose representations that are both meaningful and useful to the actors involved (Clavier 2010, 111; Tennis 2012). However, such criteria are not available in advance. It was therefore necessary to actively explore the contextual environment and interact with the actors to find relevant external elements, to perform what Tennis (2012, §64) refers to as the "right action" – in this case, applying automatic techniques appropriately within the ontology design process. We therefore hypothesized that a contextual approach can assist knowledge organization specialists in identifying relevant criteria for the effective use of automatic tools in the development of MKOSs.

3. Theoretical and methodological framework

3.1. Production of contextual resources

We developed a utility-based contextualized approach (Trzmielewski 2024) by analyzing and actively participating in the interactions between allergy professionals and their informational and socio-organizational environments. This approach relied on three main components: (1) the analysis of the context of use of allergy knowledge, conducted through the study of the informational practices of allergy professionals who seek, produce, and mobilize knowledge in their domain; (2) the automatic analysis of a document corpus regularly used by these professionals in their daily activities; and (3) our involvement as a designer in interactions with actors, identified and considered both as experts and final users. We did not apprehend *context* as a reality out there, waiting to be analyzed and fully described, but rather as a theoretical and empirical construct shaped by the researcher's specific purpose in interaction with scientific object (Mucchielli and Noy 2005; Talja et al. 1999). Our aim was to produce contextual resources and criteria that would contribute to the development of a useful ontology to effectively support the work and information activities of allergy professionals. To this end, we sought to identify and construct criteria of *usefulness* by capturing the value of use that professionals assign to the ontology tool and related actions, through their interpretation and use (Mayère 1997; Simonnot 2002), as well as through our engagement with these actors, involving their interaction with the constructed ontology within the context of their practice. This approach aligns with human-centered design principles, which advocate for involving end users as the most qualified actors to assess and appropriate the developed systems, and for grounding the design process in the *context of use* – considered as a «combination of users, goals and tasks, resources, and environment» – rather than in technical or technological possibilities (ISO 2019, 3.1).

We began by studying the information practices of professionals working in the unit, as previously described in Trzmielewski (2024) and illustrated in Figure 1. We conducted 16 participant observations and 20 interviews with supervising and apprentice doctors, researchers, clinical study coordinators, nurses, childcare workers, and secretaries. This led to the construction of an informational corpus composed of observation reports and interview transcripts. We used this corpus to extract candidate terms and create allergy user-oriented facets, which were subsequently discussed and tested during validation meetings with professionals to assess their usefulness for document indexing and searching of allergy professionals' information targets.

The informational corpus also helped us to identify the documents used daily by professionals, which enabled the development of a documentary cor-

pus containing 4,363 documents (608,910 words). All documents were less than three years old, reflecting the professionals' preference for recent documentation, and were entirely written in French – the working language of the unit. This documentary corpus was divided into three subcorpora – scientific[1], generalist[2], and clinical[3] – and analyzed using automatic tools. Our goal was to extract allergy-related terminology as well as to evaluate the ALLERGIDOC Ontology as a tool supporting document processing tasks.

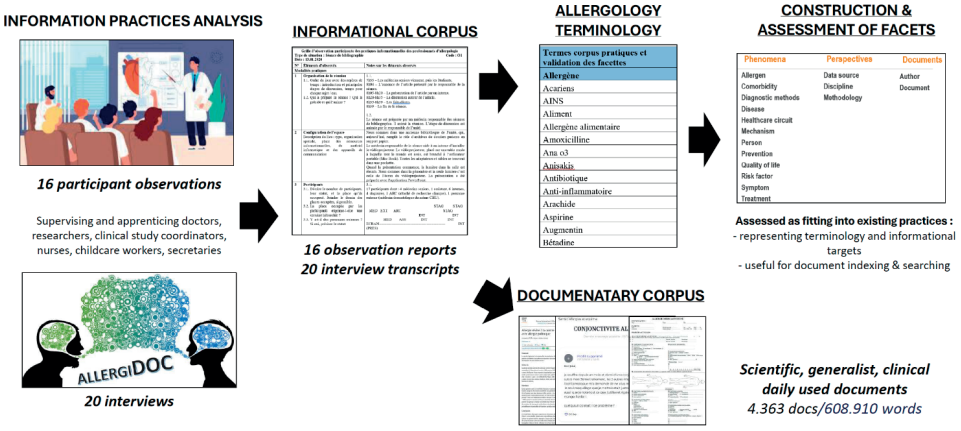


Figure 1. Process of producing contextual resources.

3.2. Ethical and legal obligations

The clinical and generalist documents from our corpus contained sensible data on people health condition, as well as personal data such as name, surname, age, sex, etc. We had to carry out few steps to use these data according to the French *Information Technology and Liberties Law* (Présidence de la République Française 1978) and the European *General Data Protection Regulation* (European Parliament, and Council of the European Union 2016, arts. 14-20). The collection of generalist data was approved by the university hospital's ethics committee and registered on *ClinicalTrials.gov* and the *French*

[1] The scientific subcorpus was composed of 864 titles and 761 scientific articles abstracts from the *Revue Française d'Allergologie* journal. As shown by the analysis of information practices, titles and abstracts are useful to professionals for representing and understanding the subject of documents (Trzmielewski 2023).

[2] The generalist subcorpus contained 403 titles of threads and 2,175 messages from *Doctissimo* generalist online health forum.

[3] The clinical subcorpus included 70 clinical activity documents, including models of standardized protocols, medical reports, standardized questionnaires, check and patients lists, 89 anonymized medical reports, and 1 text document containing terms used in a document management system to categorize documents.

Health Data Hub. Considering this legislation, all personal data that appeared in clinical documents had to be erased. Although we erased personal data, we collected sensitive data, particularly related to individuals' health. Even though sensitive data are considered as a special category of personal data, the public interest inherent in our research justified their use (Cnil 2023). Moreover, the data were not collected directly from the individuals concerned, but rather secondarily. In this context, according to the French and European Law, we were required to inform *Doctissimo* (Doctissimo 2025) users, as well as Allergy Unit patients about the purpose of the data processing, the types of personal data collected, and the conditions of data storage. We were also obligated to guarantee their rights to rectification, erasure, etc., and to explain how to exercise them.

Doctissimo, the source of the collected generalist documents, is a public forum that does not require prior registration to access discussion threads. The content is clearly made public, and we are therefore authorized to reuse it for our research (Cnil 2023). Following the advice of a hospital methodologist and in accordance with the various legislative frameworks previously mentioned, we collectively informed forum users about our study. We posted a message on the first discussion thread in the "Allergies and Eczéma" section, titled *Modération sur Doctissimo*, which is the most visible thread in this forum category. On the forum, user profiles had already been anonymized, and all personal data had been removed by administrators in compliance with data protection regulations. Account names were pseudonymized, i.e., replaced by automatically generated alphanumeric strings, such as "sev93ydy". We did not collect any data on names, first names, pseudonyms, age, phone numbers, and email addresses that may have appeared in the message content.

Regarding the clinical corpus, to prevent the identification of individuals in medical records, a supervising doctor who assisted us with data collection removed the names of physicians and patients, hospital identification numbers, date of birth and consultation, as well as sex and geographical location-related data. He also used a tool available at the hospital to send individual notifications to patients whose data were included in the collection. No objections were received in response. At our end, we anonymized the documents sourced from document management systems used in the unit by removing the names of physicians, phone numbers, and email addresses.

The data, prepared accordingly, were securely stored on a university server at the University of Montpellier Paul-Valéry.

3.3. Use of semi-automatic tools

To collect data relevant to ontology design, we used the terminology extractor TermoStat Web 3.0 (Drouin n.d.) and the text annotation tool Anno-

tator (Whetzel et al. 2011) to process documents from the documentary corpus. TermoStat is an automatic term acquisition tool developed and hosted at the University of Montreal. Its mechanism uses a method based on contrasting specialized and non-specialized textual corpora, combined with text tagging, character string extraction rules, and weighting and selection scores – such as frequency, specificity, log-likelihood, and weirdness – to identify terms (Drouin 2010). The tool takes a text as input and outputs a ranked list of candidate terms extracted from the text. Terms can be simple or complex, with each term assigned a score based on its frequency both in the analyzed corpus and in a reference corpus (Figure 2). The process of allergy terminology extraction (carried out in 2022) required converting the content into raw text and processing the three subcorpora separately. From each subcorpus, we extracted the 497 most frequently occurring candidate terms, corresponding to the number of terms manually extracted from the informational corpus. We compared the automatic extraction results for each subcorpus with the facets and terms identified through manual analysis of the informational corpus, to identify new useful terms and eliminate irrelevant or redundant ones. Allergy facets were particularly helpful in this task, as they provided a contextual semantic frame within which candidate terms could be used. The extraction results were validated by an allergist from the unit during a 1 hour and 30 minutes interview. Further validation was conducted by 10 other allergy professionals during the assessment of the 1.0 prototype of the ontology.

Liste des termes | Nuage | Statistiques | Structuration | Bigrammes

Candidat de regroupement	Fréquence
patient	1851
cas	1071
allergie	1031
étude	864
résultat	827
asthme	784
réaction	766
test	745
enfant	734
méthode	585
conclusion	569
traitement	546
p	510
introduction	490
ige	410
allergène	404
anaphylaxie	396
diagnostic	365
sensibilisation	349
prise	342
âge	314

Figure 2. A partial list of the terminological extraction of the scientific corpus with TermoStat.

Then, we assessed the utility of 3.0 version of the ALLERGIDOC Ontology to support automatic document processing – a need identified through the analysis of information practices (Trzmielewski 2024) – in situation of use. Seven professionals, i.e., one medical secretary, four attending doctors and two residents participated in meetings held in the Allergy Unit in 2025. Most of them reported having little or no prior experience with document annotation or indexing, as they are only accustomed to organizing work-related digital documents into folders. Some of them knew indexing principles from publicly available tools such as the PubMed search browser, consulted for scientific research.

Participants analyzed six selected daily-used digital documents, chosen among 15 pre-selected by the design team from the documentary corpus. The first participant – a medical secretary – chose two titles and abstract of scientific articles, two medical reports, and two messages from the generalist health

forum, as documents corresponding to their daily practice. The selected documents were then processed by the remaining participants. After careful and aloud reading and analysis, professionals manually extracted and proposed keywords[4] on paper forms for each document that, in their opinion, were useful to represent and search for the processed documents (Figure 3). This was followed by a semi-automatic annotation[5] of the same documents with Annotator, performed by the participants themselves, to obtain the corresponding ontology terms (Figure 3). Annotator is a biomedical annotation tool provided by the American National Center for Biomedical Ontology (NCBO). The tool is publicly and freely available, though it imposes a limit on the maximum number of words that can be processed in a text document.

To generate annotations, participants copied and pasted textual content of documents in the online box, selected the ALLERGIDOC Ontology as dictionary[6], and pressed the submit button. Annotator carried out text pre-processing task (such as eliminating stop words), matched words in the text to terms in the ontology by performing an exact string comparison (a "direct" match) between the text and ontology preferred and non-preferred terms (NCBO 2025). To perform the term recognition by using string matching on the dictionary, Annotator uses the Mgrep tool (Dai et al. 2008).

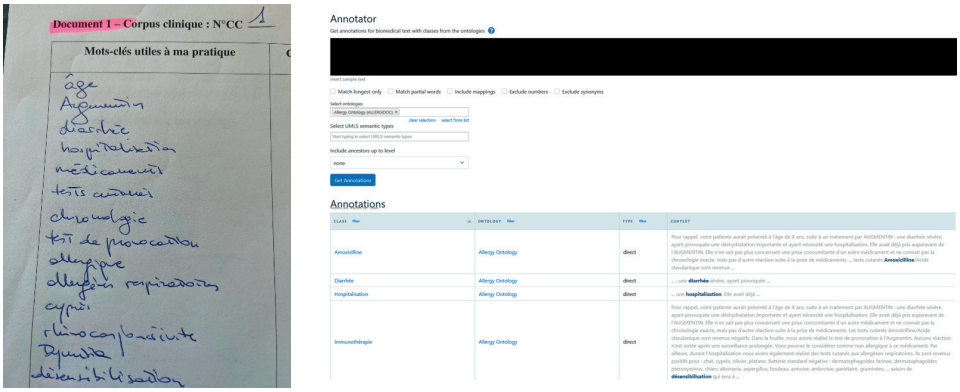


Figure 3. Example of keywords assigned by an allergy professional (left) and terms from the ALLERGIDOC Ontology proposed by the Annotator tool (right) to represent the content of a medical report.

To perform the assigned task, participants used hospital computers and spent an average of 47 minutes. The results from each step were compared and discussed by participants, while the design team interviewed them, observed

[4] A keyword should be considered here as a word that «from some perspective provides a useful description of a document» (Lardera and Hjørland 2021).

[5] An annotation is a specific note attached to a document for various purposes – in our case, to represent its content (Bringay et al. 2005).

[6] The ALLERGIDOC Ontology is hosted and available on BioPortal (Whetzel et al. 2011).

their informational, bodily and verbal behavior, and took notes. The content of the paper forms, as well as the observation and interview notes, were analyzed. We calculated an average number of keywords proposed by participants and the annotation tool, and we assessed their matching score and related patterns. Furthermore, we conducted a thematic analysis of their discourses on the proposed keywords and ontology terms, as well as on the perceived utility of the ALLERGIDOC Ontology to support document processing.

4. Results

Our study reveals that the data on the context of use of knowledge in the allergy unit, obtained through qualitative methods, leads to the identification of a set of useful human, informational, and socio-organizational elements. These elements guided the automatic extraction and annotation processes carried out by the design team and final users.

4.1. The semi-automatically extracted terminology

The specialized vocabulary used by professionals, which was extracted from the content of the observation reports and interview transcripts, as well as the user-oriented facets and validation meetings with professionals, guided the process of semi-automatic extraction of useful candidate terms, the discrimination of irrelevant ones, and the assessment of the performance of the extraction software. Almost all selected terms were accepted by the allergist during validation meetings. The involved professional excluded only 12 candidate terms that did not represent his practice. Ten other employees, participating in the assessment of the first version of the ALLERGIDOC Ontology, highlighted the relevance and utility of the extracted and further organized terminology. They emphasized that the ontology reflects "well", "largely" and "completely" their professional reality and that the represented terms are used by them "almost every day". The selected terms made sense to them, and they suggested that such terms could be used to index and search for allergy documentation, as well as to help to introduce new professionals to domain terminology.

Furthermore, we could observe above-average performance of TermoStat in relation to our extraction aims. From the total of 1,491 of the most frequently occurring candidate terms extracted to construct the first version of the ontology, on average, 63% were actually useful, while 37% were out of context (Table 1). This performance score demonstrates the relevance of the TermoStat tool for extracting medical terminology. At the same time, it highlights the need for contextual resources – i.e., daily-used documentation to

be represented and organized, contextual semantic frames, experts, and final users – to create a favorable environment and intellectual resources to guide the automated extraction process. However, the extraction tool also had its limitations, especially in processing complex terms, composed of two or three lexical units. We had to complete several single terms to adapt them to the allergy domain context. For example: "Arrêt" → "Arrêt cardio-respiratoire", "Eosinophile" → "Œsophagite à éosinophiles", "Anticorps" → "Anticorps monoclonaux". To do so, we consulted the occurrence of terms in their phrase context within TermoStat.

	Scientific corpus	Clinical corpus	Generalist corpus	Total of extracted terms	Performance of the tool
Useful terms	192 (39%)	138 (27%)	240 (48%)	570 (38%)	**63% of useful terms**
Useful but redundant	119	171	77	367	
Not useful	186 (37%)	188 (38%)	180 (36%)	554 (37%)	**37% terms out of context**
Total	497	497	497	1491	

Table 1. Results of the semi-automatic term extraction supported by TermoStat, after assessment using contextual resources.

4.2. The utility of the ontology to support automatic document processing

The assessment meetings on the utility of the ALLERGIDOC Ontology version 3.0 to perform automatic document annotation revealed low performance. Among the average of 11 keywords proposed by professionals per document (Table 2), we noted only 2 exact matches per document with the ontology terms automatically proposed by Annotator (Table 3). This low efficiency is mostly due to three factors: 1) missing terms or terms that were not specific enough in the ontology; 2) the absence of contextual knowledge in the tools – knowledge that matters to allergy professionals and goes beyond string matching-based document processing; 3) the limitations of the annotation tool itself.

	Doc 1	Doc 2	Doc 3	Doc 4	Doc 5	Doc 6	Total	Mean
Participant 1	5	5	3	3	4	6	26	4,33
Participant 2	3	5	3	6	8	8	35	5,83
Participant 3	14	15	13	10	19	15	86	14,33

Participant 4	4	24	25	10	21	11	95	15,83
Participant 5	11	8	7	8	14	9	57	9,5
Participant 6	6	14	15	15	21	18	79	13,16
Participant 7	12	14	11	8	13	15	73	12,16
TOTAL	55	85	77	60	100	82	451	X
Mean (participants)	7,85	12,14	11	8,57	14,28	11,71	64,42	**10,79**
Mean (Annotator)	19	37	16	13	1	12	98	16,33

Table 2. The number of descriptors assigned by participants (green) and by annotation tool (orange) to represent the content of documents.

	Doc 1	**Doc 2**	**Doc 3**	**Doc 4**	**Doc 5**	**Doc 6**	**Total**	
Participant 1	2	0	0	0	0	1	3	
Participant 2	1	1	1	2	0	2	7	
Participant 3	5	2	4	5	0	4	20	
Participant 4	2	13	5	4	0	1	25	
Participant 5	0	5	0	1	0	1	7	
Participant 6	0	4	2	4	0	2	12	
Participant 7	4	3	1	6	0	3	17	
TOTAL	14	28	13	22	0	14	91	
Mean	2	4	1,86	3,14	0	2	X	**2,16**

Table 3. Scores showing the number of exact matches between keywords assigned by professionals and ontology terms proposed by the Annotator.

At the semantic level, the ontology terms lacking specificity included document section titles (e.g., “Introduction”, “Résultats”, “Conclusion”), dynamic phenomena (e.g., “Maladie”, “Traitement”, “Sensibilisation”, “Hospitalisation”), individuals (e.g., “Médecin”, “Patient”, “Allergologue”), and quantities (e.g., “Dose”, “Heure”), among others.

At the lexical level, employees suggested several keywords that were missing in the ontology: commercial drug names (e.g., “Zyrtec”, “Telfast”, “Wystam”, “Atarax”), Latin binomial names (e.g., *Solenopsis invicta*), names of events leading to hypersensitivity reactions (e.g., “Envenimation”, “Morsure”), as well as metonymic expressions such as “Tryptase” (a substance) used to designate “Dosage de la tryptase” (a blood test).

At the morphological level, professionals frequently used acronyms, such as “TC” for “Test cutané”, “AO” for “Angioœdème”, “AHI” for “Antihistaminique”, “ETP” for “Education thérapeutique”; abbreviations (e.g., “Allergo”, “Dermato”); hyphenated terms (e.g., “Angio-œdème”); unaccented terms

(e.g., “Eczema”); and plural forms designating allergen sources, diseases, and symptoms – many of which are currently absent from the ontology.

Also, many keywords proposed by professionals were present in the developed MKOS but literally absent from the corpus. This was frequently the case of hypernyms – such as the keyword “Pneumallergène” which was used to group various standard aeroallergens tested in patients: “Chat”, “Cyprès”, “Olivier”, “Platane”, “Ambroisie”, “Pariétaire”, etc., as reported in medical documents. The same observation applied to other allergens, diseases, and drugs classes. Some terms proposed by the employees were also semantically related to words appearing in the text, for example, the keyword “Pollen” was associated with terms referring to plants such as “Cyprès”, “Olivier”, “Platane”. Moreover, the annotation tool did not match words in the text when a plural form was missing in the ontology (e.g., “Hyménoptère” instead of “Hyménoptères”), or when ontology terms were morphologically close but either too broad (“Anaphylaxie” instead of “Choc anaphylactique”) or too narrow (“Eczéma atopique” instead of “Eczéma”).

Some issues were difficult to identify because, although the ontology terms corresponded to the keywords, they were not extracted by the tool. We assumed this could result from the absence of “rdfs:label” on some class names, missing underscores between words in certain complex terms within the ontology (e.g., “Syndrome oral”), typographical issues (e.g., “Angioedème” instead of “Angioœdème”); the presence of an apostrophe (as in “Trousse d’urgence”), or even the tool’s inability to extract complex lexical units at all. In addition, one document received only a single matching term, possibly due to a formatting problem. Therefore, we can question whether we were truly assessing the ontology itself, the annotator tool, or the combination of both as a single tool. Indeed, without the use of Annotator, the number of exact matches between keywords assigned by professionals and terms present in the ontology increased. There were, on average, more than 3 exact matches per document between the keywords assigned by professionals and the terms present in the ontology (Table 4).

	Doc 1	**Doc 2**	**Doc 3**	**Doc 4**	**Doc 5**	**Doc 6**	**Total**	
Participant 1	2	3	1	0	0	2	8	
Participant 2	2	4	3	4	2	2	17	
Participant 3	5	6	7	4	2	5	29	
Participant 4	0	14	9	5	7	2	37	
Participant 5	3	5	2	2	5	3	20	
Participant 6	1	7	5	6	5	2	26	
Participant 7	3	4	2	5	2	2	18	
TOTAL	16	43	29	26	22	18	154	
Mean	2,28	6,14	4,14	3,71	3,14	2,57	X	**3,66**

Table 4. Number of exact matches between keywords assigned by professionals and terms present in the ontology.

Nevertheless, professionals had no issues manipulating the annotation tool supported by the 3.0 version of the ontology, and most of them found the tool "hyper advanced", "top", "smart", and stated that it "would make life easier". The tool provided them with "interesting" descriptor suggestions they had not previously considered, such as those related to alternative medicine or disease stages. Professionals also shared with the design team some suggestions regarding potential applications of such a tool. All of them found the tool useful for annotating and indexing medical reports in electronic health records system, documents on appointment platforms, as well as social media messages in databases for their research–types of informational resources that currently lack adequate knowledge organization tools in both the hospital and the research laboratory where the employees work. They also shared with us their informational targets, such as common and rare skin tests (for allergens such as mites, drugs by pharmaceutical class and commercial name, exotic foods, etc.), blood tests, allergies, symptoms as expressed by specialists and patients (e.g., "Démangeaisons" vs. "Grattage"), affected body parts, alternative therapies, emotions/suffering, and photos – particularly emphasizing the need for detailed instructions on indexing and searching for allergens, which remains one of their most important informational priorities. In fact, they most frequently search for medical reports containing positive test results to support their clinical and research practices – for example to identify which types of patients reacted to a particular allergen – and make decisions on how to treat another patient with a similar medical history, or to compile statistics on the characteristics of patients allergic to the same agent. The medical secretary highlighted the need for interoperability between the Allergy Unit and the Pneumology Unit information systems, as these units are located close to each other and may share the same patients.

The annotation of allergens contributed to the gap between the number of annotations generated by Annotator – 16 terms per document – and the nearly 11 keywords proposed by the allergy professionals (Table 2). To improve this aspect, participants suggested including ancestor classes to group classes of allergens discussed in processed documents and prioritizing the indexing of only positive allergen test results. The potential broader classes already exist in the ontology as subclasses of the separated class "Allergen", but they are not defined as direct broader classes of specific plants or foods, which are currently classified under "Living beings" and "Agricultural or manufactured products", respectively. In addition, when (most) patients in the Allergy Unit undergo testing with standard baseline series – for example for aeroallergens or food allergens – the names of all allergens, whether negative or positive as results, are documented in medical reports. This distinction was not recognized by the annotation tool.

Moreover, professionals suggested introducing coordination rules to represent and search for complex subjects, particularly those involving static, dynamic, and abstract phenomena – such as agents, objects, events, processes, temporal aspects – and various attributes, especially within medical reports:

- allergy agent-disease, e.g., "SOA-fruits", "Anaphylaxie au blé",
- allergen-allergen cases of cross-reactivity, e.g., "Acariens-Crevette",
- allergic test-allergic agent, e.g., "Test-Pneumallergène"
- symptom-severity (e.g., "Diarrhée-Sévère"), event (or disease)-severity ("Réaction allergique-Sévère", as well as body part-clinical characteristics ("Main-Rouge").

Accordingly, professionals also recommended introducing temporality-based indexing (e.g., allergy test-allergen-date) to reconstruct events from patients' allergy histories.

Finally, during the tool assessment, professionals give us many suggestions regarding new terms (preferred terms, synonyms) and relationships to be added to the ontology, as well as terms and relationships to be removed.

5. Conclusions

This study shows that the effective use of automatic tools in medical ontology design may rely on resources and criteria that are external to the documentary corpus itself. In our case, such elements emerged from the analysis of the *context of knowledge use* by social actors – composed of human, informational, and socio-organizational environments – as well as through active interactions between the designer and the actors involved in validating and potentially using the system. By adopting an ecological and situated approach, we were able to observe and engage with allergy professionals in their natural work environments, analyze their information practices, and select documents representative of their daily tasks – all while respecting ethical and legal obligations related to the use of personal and sensitive data. This approach enabled us to assess and identify new elements requiring adjustment concerning the content, intentionality, and operationality of the ALLERGIDOC Ontology. It also allowed us to identify new challenges, such as representing and searching for information related to allergens, temporality, and complex informational subjects within allergology documentation, as well as ensuring interoperability between hospital internal information systems. Such an approach goes beyond data-driven (Brewster et al. 2004) and satisfaction-based user evaluations (Sure et al. 2009), fostering instead *intelligent action* grounded in *search for intervention*, according to constructivist, complex systems modeling (Le Moigne 1999). Rather than predefining problems to solve, the contextual ontology design process itself became a space for formulating new, relevant problems

to address – to ultimately support allergology professionals' practices more *effectively*.

Thus, a contextual approach can assist knowledge organization specialists in identifying relevant elements for using automatic tools in the development of MKOSs. In this framework, the context of use becomes a dynamic "frame of reference" (Vakkari 1997, 8), constructed throughout the research process, within which designers can interpret situations and take appropriate actions (Broudoux 2012), while leveraging automatic tools to build and evaluate useful MKOSs for intended environments. Our approach is illustrated in Figure 4 through the Utility in Context Model (UCM), a framework that can be applied to design complex and useful medical ontologies, as well as other types of classificatory and terminological systems.

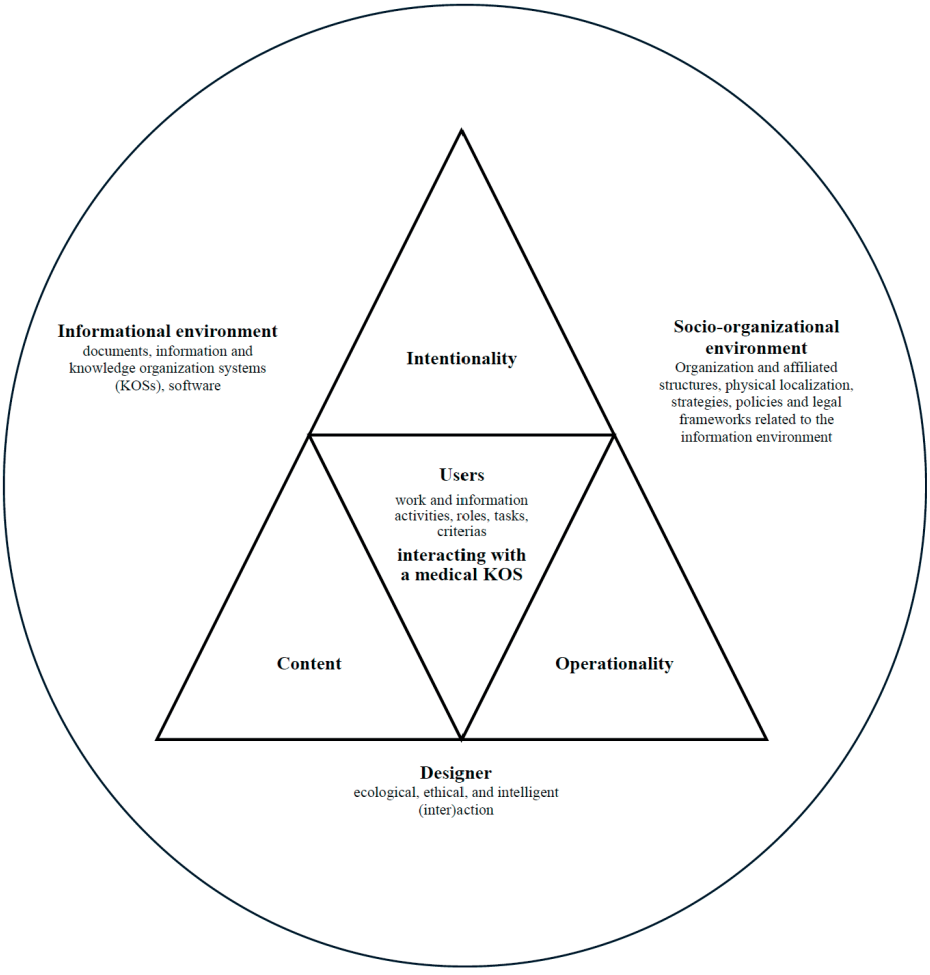


Figure 4. Utility in Context Model (UCM).

Nevertheless, our approach also reveals an inherent paradox. While automatic tools are generally employed to save time by automating tasks, their use within a qualitative, context-driven framework demands significant time investments for data collection, analysis, and validation. Furthermore, although we aimed to simulate real-world conditions, the experimental setting differed from the actual document organization environment of the Allergy Unit. Currently, professionals organize work-related digital documents in digital folders, with no established indexing system. Additional limitations stem from our reliance on external tools, such as the Annotator (which experienced some technical issues during our sessions) and TermoStat, which were not locally hosted – even though we used secure, university-provided services and excluded personal and sensitive data from the study. Moreover, we primarily evaluated the ontology representational capabilities, with limited assessment of its search functionalities in operational settings. A key limitation lies in the experimental nature of the study: it remains still a simulation rather than a full evaluation of the ontology integrated into a fully operational system. Assessment through the annotation tool was also constrained by the tool's limitations, such as its inability to extract complex terms from the ontology. The same issue arose during terminology extraction, suggesting that issues with complex term extraction may be common with certain automatic processing tools.

This research opens several perspectives for future work. Refining the ontology based on user feedback will be essential. Developing approaches to represent allergens, temporality, and complex subjects also remains a significant challenge. Ultimately, this study reinforces the need for more work emphasizing that medical ontologies for information processing and searching should not only be technically functional and logically consistent, but also meaningful and useful to their intended users.

Acknowledgments

I would like to thank the allergy professionals from the Allergy Unit of the Montpellier University Hospital – Selsabil Ayeb, Yanis Bouderbala, Jean-Luc Bourrain, Waffa Benderdour, Davide Caimmi, Pascal Demoly, and Snedjana Tintari – for their involvement and feedback regarding the present study. Special thanks also to the Information-Documentation students (referred to as "the design team") at the University of Montpellier Paul-Valéry – Nada Bouhsina, Claire Casasso, Sabrina Duc, and Anne Guerre – for their valuable contribution to data collection and analysis.

References

Almeida, Mauricio Barcellos, and Fernanda Farinelli. 2017. "Ontologies for the representation of electronic medical records: The obstetric and neonatal ontology." *Journal of the American Society for Information Science and Technology* 68 (11): 2529-42. https://doi.org/10.1002/asi.23900.

Baneyx, Audrey, and Jean Charlet. 2006. "Évaluation, évolution et maintenance d'une ontologie en médecine : état des lieux et expérimentation." *Revue I3-Information, Interaction, Intelligence* special issue on *Ressources termino-ontologiques*: 147-73. Cépaduès Éditions.

Barité, Mario. 2018. "Literary warrant." *Knowledge Organization* 45 (6): 517-36. https://doi.org/10.5771/0943-7444-2018-6-517.

Brewster, Christopher, Harith Alani, Srinandan Dasmahapatra, and Yorick Wilks. 2004. "Data-Driven Ontology Evaluation." In *Proceedings of the Fourth International Conference on Language Resources and Evaluation (LREC'04)*, edited by Maria Teresa Lino, Maria Francisca Xavier, Fátima Ferreira, Rute Costa and Raquel Silva. European Language Resources Association (ELRA). http://www.lrec-conf.org/proceedings/lrec2004/pdf/737.pdf.

Bringay, Sandra, Catherine Barry, and Jean Charlet. 2005. "Les annotations pour gérer les connaissances du dossier patient." In *Actes de la conférence IC 2005 – 16es journées francophones d'Ingénierie des Connaissances – 1-3 juin 2005 Nice*, edited by Marie-Christine Jaulent. https://inria.hal.science/hal-01025416/document.

Broudoux, Evelyne. 2012. "Vers l'objet documentaire (re)contextualisé," paper presented at *9ème congrès des enseignants documentalistes de l'Éducation Nationale : Objets documentaires numériques : nouvel enseignement ?*, Paris, March 22. https://archivesic.ccsd.cnrs.fr/sic_00715868/document.

Choi, Yunseon. 2016. "Supporting better treatments for meeting health consumers' needs: extracting semantics in social data for representing a consumer health ontology." *Information Research* 21 (4). https://informationr.net/ir/21-4/paper731.html.

Cisek, Sabina. 2013. "Metodologia jakościowa we współczesnej informatologii: wybrane aspekty." *Przegląd Biblioteczny* 81 (3): 299-310. https://doi.org/10.36702/pb.248.

Clavier, Viviane. 2010. "Indexer des parcours thématiques pour valoriser les collections de presse numérisée." In *Document numérique, entre permanence et mutations. Actes du 13ème Colloque International sur le Document Électronique (CIDE.13), 16-17 décembre 2010*, *INHA, Paris*, edited by Madjid Ihadjadene, Manuel Zacklad and Khaldoun Zreik. Europia Productions.

Cnil. 2023. "Données sensibles." Last modified August 2023. https://www.cnil.fr/fr/definition/donnee-sensible.

Dai, Manhong, Nigam H. Shah, W. Xuan, et al. 2008. "An Efficient Solution for Mapping Free Text to Ontology Terms." In *American Medical Informatics Association Symposium on Translational BioInformatics, AMIA-TBI'08 12-15 March 2008 San Francisco, United States.*

Doctissimo. 2025. "Forum santé – Allergies et eczéma." Last modified July 3. https://forum.doctissimo.fr/sante/allergies-eczema/liste_sujet-1.htm.

Drouin, Patrick. 2010. "TermoStat – Guide de l'utilisateur." Last modified December 2010. https://termostat.ling.umontreal.ca/doc_termostat/doc_termostat_en.html.

Drouin, Patrick. n.d. "TermoStat Web 3.0." Last access October 2025. https://termostat.ling.umontreal.ca/.

European Parliament, and Council of the European Union. 2016. *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)*. Official Journal of the European Union. Approved 2016. https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng.

Hassan, Mohsen. 2017. "Knowledge Discovery Considering Domain Literature and Ontologies: Application to rare Diseases." PhD diss., Université de Lorraine.

International Organization for Standardization. *Ergonomics and human-system interaction – Part 210: Human-centred design for interactive systems*. ISO 9241-2010:2019. ISO. Published 2019; confirmed 2025. https://www.iso.org/standard/77520.html.

Iyer, Hemalata, and K. S. Raghavan. 2018. "Medical ontology: Siddha System of Medicine." In *Challenges and Opportunities for Knowledge Organization in the Digital Age: Proceedings of the Fifteenth International ISKO Conference 9-11 July 2018 Porto, Portugal*, edited by Fernanda Ribeiro and Maria Elisa Cerveira. Ergon. https://doi.org/10.5771/9783956504211-347.

Lardera, Marco, and Birger Hjørland. 2021. "Keyword". *Knowledge Organization* 48 (6): 430-56. https://doi.org/10.5771/0943-7444-2021-6-430.

Le Moigne, Jean-Louis. 1999. *La modélisation des systèmes complexes*. Dunod.

Liu, Hao, Shuxin Zhou, Zhehuan Chen, Yehoshua Perl, and Jiayin Wang. 2024. "Using Generative Large Language Models for Hierarchical Relationship Prediction in Medical Ontologies." In *2024 IEEE 12th International Conference on Healthcare Informatics (ICHI)*, Orlando, FL, USA, June 03-06. Institute of Electrical and Electronics Engineers (IEEE). https://doi.org/10.1109/ICHI61247.2024.00040.

Lopes, Lucelene, Renata Vieira, Maria José Finatto, Daniel Martins, Adriano Zanette, and Luiz Carlos Ribeiro Jr. 2009. "Automatic extraction of composite terms for construction of ontologies: an experiment in health care area." *Electronic Journal of Communication, Information & Innovation in Health* 3 (1): 72-84. https://doi.org/10.3395/RECIIS.V3I1.244EN.

Mayère, Anne. 1997. "Produits et services d'information : proposition de dépassement de la théorie standard de l'information." In *Société informationnelle : enjeux sociaux et approches économiques*, edited by Anne Mayère. L'Harmattan.

Mucchielli, Alex, and Clair Noy. 2005. *Etudes des communications – Approches constructivistes*. Armand Collin.

Muresan, Smaranda, and Judith L. Klavans. 2013. "Inducing terminologies from text: A case study for the consumer health domain." *Journal of the American Society for Information Science and Technology* 64 (4): 727-44. https://doi.org/10.1002/asi.22787.

NCBO. 2025. "BioPortal Help – Annotator." Last modified January 7 2005. https://www.bioontology.org/wiki/BioPortal_Help#Annotator.

Oh, Kyong Eun, Soohyung Joo, and Eun-Ja Jeong. 2015. "Online Consumer Health Information Organization: Users' Perspectives on Faceted Navigation." *Knowledge Organization* 42 (3): 176-86. https://doi.org/10.5771/0943-7444-2015-3-176.

Présidence de la République Française. 1978. *Loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés*. In *Journal officiel de la République française*. Published 1978; confirmed July 2, 2025. https://www.legifrance.gouv.fr/loda/id/JORFTEXT000000886460.

Raghavan, K. S., and C. Sajana. 2010. "NeurOn. Modeling ontology for neurosurgery." In *Paradigms and conceptual systems in knowledge organization: Proceedings of the Eleventh International ISKO Conference 23-26 February 2010 Rome, Italy*, edited by Claudio Gnoli and Fulvio Mazzocchi. Ergon.

Richard, Marion. 2017. "Apport de la modélisation ontologique pour le partage des connaissances en psychiatrie." PhD diss., Université Pierre et Marie Curie (Paris 6).

Simonnot, Brigitte. 2002. "De la pertinence à l'utilité en recherche d'information : le cas du Web." In *Recherche récentes en sciences de l'information. Convergences et dynamiques – Actes du colloque MICS-LERASS 21-22 mars 2002 Toulouse*, edited by Viviane Couzinet and Gérard Régimbeau. ADBS Editions.

Soergel, Dagobert, Tony Tse, and Laura Slaughter. 2004. "Helping Healthcare Consumers Understand: An 'Interpretative Layer' for Finding and Making Sense of Medical Information." In *MEDINFO Proceedings of the 11th World Congress on Medical Informatics*, edited by Marius Fieschi, Enrico Coiera and Yu-Chan Jack Li. IOS Press. https://ebooks.iospress.nl/publication/21125.

Sure, York, Steffen Staab, and Rudi Studer. 2009. "Ontology Engineering Methodology." In *Handbook on Ontologies*, edited by Steffen Staab and Rudi Studer. Springer. https://doi.org/10.1007/978-3-540-92673-3_6.

Talja, Sanna, Heidi Keso, and Tarja Pietiläinen. 1999. "The production of 'context' in information seeking research: a methodological view." *Information Processing and Management* 35 (6): 751-63. https://doi.org/10.1016/S0306-4573(99)00024-2.

Tang, Xiao-Bo, Wei Wei, Guang-Chao Liu, and Juan Zhu. 2017. "An Inference Model of Medical Insurance Fraud Detection: Based on Ontology and SWRL." *Knowledge Organization* 44 (2): 84-96. https://doi.org/10.5771/0943-7444-2017-2-84.

Tapi Nzali, Mike Donald, Jérôme Azé, Sandra Bringay, Christian Lavergne, Caroline Mollevi, and Thomas Optiz. 2018. "Reconciliation of patient/doctor vocabulary in a structured resource." *Health Informatics Journal* 25 (4): 1219-31. https://doi.org/10.1177/1460458217751014.

Tennis, Joseph T. 2012. "Le poids du langage et de l'action dans l'organisation des connaissances : Position épistémologique, action méthodologique et perspective théorique." *Etudes de communication* 39: 15-40. https://doi.org/10.4000/edc.4010.

Toro, Sabrina, Anna V. Anagnostopoulos, Sue Bello, et al. 2024. "Dynamic Retrieval Augmented Generation of Ontologies using Artificial Intelligence (DRAGON-AI)." *Journal of Biomedical Semantics* 15. https://doi.org/10.1186/s13326-024-00320-3.

Trzmielewski, Marcin. 2023. "Vers l'élaboration d'un système d'organisation des connaissances en allergologie : l'analyse des documents et des pratiques informationnelles des acteurs." PhD diss., Université Paul-Valéry Montpellier 3.

Trzmielewski, Marcin. 2024. "Towards an Allergy Ontology: A Utility-Based Contextual Approach." *Knowledge Organization* 51 (3): 151-68. https://doi.org/10.5771/0943-7444-2024-3-151.

Vakkari, Pertti. 1997. "Information seeking in context: A challenging metatheory." In *ISIC '96: Proceedings of an international conference on Information seeking in context*, edited by Pertti Vakkari, Reijo Savolainen and Brenda Dervin. Taylor Graham. https://dl.acm.org/doi/10.5555/267190.267221.

Whetzel, Patricia L., Natalya F. Noy, Nigam H. Shah, et al. 2011. "BioPortal: enhanced functionality via new Web services from the National Center for Biomedical Ontology to access and use ontologies in software applications." *Nucleic Acids Res* 39 (Web Server issue): W541-5. https://doi.org/10.1093/nar/gkr469.

AIDAinformazioni

Rivista semestrale di Scienze dell'Informazione

Anno 43

N. 3-4 – luglio-dicembre 2025

Contributi

Giovanna Aracri, Antonietta Folino, Maria Teresa Guaglianone, Marco Manna, Identification of thesaurus relationships with the support of ChatGPT and Gemini

Alessandra Bonsignori, Terminologia giuridica e organizzazione della conoscenza

Maria Teresa Chiaravalloti, Maria Teresa Guaglianone, Elena Cardillo, Metadatazione delle tipologie di documenti clinici: un approccio alla standardizzazione mediante LOINC

Carla Beatriz Marques Felipe, Carla Maria Martellote Viola, Raimunda Fernanda dos Santos, Domain Analysis on Data Management

Claudio Gnoli, Emanuela Valenzano, Che faccetta sei?

Manuela Grillo, Strumenti di convergenza tra archivi e biblioteche: i thesauri

Bianca La Manna, Rappresentare gli atti notarili medievali in Linked Open Data (LOD)

Claudia Lanza, L'univocità e la variabilità terminologica a supporto della rappresentatività delle risorse semantiche di organizzazione e rappresentazione della conoscenza specialistica

Larissa Moraes Martins, Carlos Eduardo da Silva Carvalho, Ana Cristina de Albuquerque, Mario Barité, Possibilities for using cultural warrant to complement literary work

Bianca Therezinha Carvalho Panisset, Lucia Maria Velloso de Oliveira, Leticia Souza da Costa Sampaio, Camilla Campoi de Sobral, Identification of types of documents and knowledge organization: transparency and access in records management

SoYoung Yoon, Claudio Gnoli, Ziyoung Park, Applying Integrative Levels Classification to Korean KOS Registry

Emanuela Nicole Donato, Grazia Serratore, Maria Taverniti, Maria Teresa Chiaravalloti, L'evoluzione della struttura documentale nelle cartelle cliniche psichiatriche dell'archivio storico del manicomio di Girifalco (1881-1978)

Lorenza Talarico, Manuscripts in the Digital Age

Maria Lígia Triques, Ananda Fernanda de Jesus, Mario Barité, José Eduardo Santarem Segundo, Ana Cristina de Albuquerque, Vocabularies in the digital context

Marcin Trzmielewski, Context-Driven Use of Automatic Tools in Medical Ontology Design: A Case Study in Allergology

Carla Maria Martellote Viola, Carla Beatriz Marques Felipe, Exploring the relationships between the Cross-Cutting Women's Agenda - Multi-Year Plan 2024-2027 and the Knowledge Organization of legislative proposals under processing on women in Brazil

Elena Cencetti, Camillo Carlo Pellizzari di San Girolamo, Elisabetta Viti, Modellizzazione delle conoscenze: Thesaurus Nuovo soggettario e Wikidata, due strumenti a confronto

Pontish Yeramyan, Alexander Murzaku, Nick Uva, Curt Anderson, The transferability of brilliance

Ana Flávia Dias Zammataro, Ana Cristina de Albuquerque, Giulia Crippa, Lucia Sardo, Carlo Bianchini, Maria Lígia Triques, ISKO Italy in perspective

mundaneum

In copertina
Disegno di Paul Otlet, Collections Mundaneum, centre d'Archives, Mons (Belgique).

ISBN 979-12-5965-639-1

ISSN 1121-0095