

Graphes sémantiques flous et raisonnement approximatif pour la découverte de connaissances

Projet de thèse CIFRE

Encadrement	LRASC	Thales TRT	Claire Laudy
	LIP6	Sorbonne Université	Marie-Jeanne Lesot
Dates	2022 - 2025		
Mots-clés	apprentissage automatique - <i>soft computing</i> - fouille de graphes - représentation de connaissances - graphes conceptuels - fusion et agrégation - données structurées - données imparfaites		

1 Partenaires impliqués dans le projet de recherche

Le projet de thèse CIFRE se place dans le cadre d'une collaboration entre l'entreprise Thales et le Laboratoire d'Informatique de Paris 6 (LIP6). Ils sont brièvement présentés ci-dessous.

1.1 Thales - TRT

Thales est le leader mondial des systèmes d'information critique avec des activités dans trois principaux domaines : aérospatial, défense et sécurité. Thales emploie 68000 personnes dans le monde et développe des composants stratégiques, logiciels, systèmes et architectures via son organisation de R&T.

La première mission de Thales Research & Technology (entité centrale de recherche de l'entreprise Thales) est de faire le lien entre les principaux acteurs de la recherche pour engendrer des ruptures technologiques pour Thales.

Le groupe de recherche en Sciences et Techniques de l'Information de Thales Research & Technology TRT se compose de plusieurs laboratoires dont un spécialisé dans l'analyse et le raisonnement dans les systèmes complexes (LRASC). Le laboratoire LRASC recherche, pour un problème donné à Concevoir et développer des solutions algorithmiques de représentation et de raisonnement à base d'IA, ayant pour objectif principalement d'aider à la construction de sens et l'amélioration de la conscience de situation en adéquation avec des exigences de qualification des solutions.

Au sein du laboratoire, divers travaux sont menés dans le domaine de l'analyse des données et informations, afin d'expliquer des situations observées. Une partie des informations manipulées sont représentées sous forme de graphes sémantiques. le laboratoire développe des algorithmes pour la manipulation et l'analyse de ces graphes sémantiques.

1.2 Laboratoire d'Informatique de Paris 6 (LIP6)

Le LIP6 est un laboratoire de recherche en informatique placé sous la tutelle du CNRS et de l'université Sorbonne Université. Il regroupe 528 personnes, dont 219 permanents et 188 doctorants. Il se consacre à la modélisation et la résolution de problèmes fondamentaux motivés par les applications, ainsi qu'à la mise en œuvre et la validation des solutions au travers de partenariats académiques et industriels.

L'équipe LFI impliquée dans l'encadrement de ce sujet de thèse développe des recherches dans le cadre de l'intelligence computationnelle, elle étudie et propose de nouvelles approches pour la prise en compte et le traitement de données et de connaissances imparfaites et subjectives. Elle est spécialisée dans la gestion de telles données, en particulier pour les tâches de raisonnement et d'apprentissage automatique, en exploitant les cadres des logiques non classiques et du soft computing.

1.3 Collaborations antérieures

Au niveau institutionnel d'abord, l'entreprise Thales et l'Université Pierre et Marie Curie (UPMC) ont un riche historique de collaborations, qui les a conduites à établir un accord-cadre qui facilite leur travail commun. L'université Sorbonne Université, créée en janvier 2018 par fusion de l'UPMC et de l'université Paris IV, bénéficie de cet historique.

De plus, l'équipe LFI et le laboratoire LRASC ont l'habitude de collaborer et de co-encadrer des thèses : actuellement, ils sont partenaires dans le cadre du projet FUI FAUCON et co-encadrent deux thèses en cotutelle avec la Pologne.

En outre, Claire Laudy, encadrante pour Thales, a réalisé sa thèse dans le département DAPA (département de rattachement de l'équipe LFI) dans le contexte d'un contrat CIFRE entre Thales et le LIP6.

Enfin, Marie-Jeanne Lesot a dirigée la thèse CIFRE d'Adam Faci co-encadrée par Claire Laudy. Ce nouveau projet de thèse est proposé dans la continuité des travaux réalisés par Adam Faci, Marie-Jeanne Lesot et Claire Laudy dans le cadre de cette précédente collaboration.

2 Contexte du projet de recherche

Contexte opérationnel Afin de proposer des solutions pour améliorer les produits THALES en intégrant des outils issus de l'intelligence artificielle, le laboratoire LRASC s'intéresse en particulier à développer des méthodologies, algorithmes et outils pour la gestion des connaissances au niveau sémantique. L'objectif est de permettre la représentation, le stockage, la manipulation et le raisonnement sur des informations qui sont compréhensibles et interprétables et exploitables à la fois par des humains et aussi par des machines (ou algorithmes).

En effet, Thales comme tous les autres acteurs de l'industrie, se trouve confronté à la nécessité de prendre en compte des informations qui ne sont plus uniquement produites par des opérationnels formés et dans un cadre directement lié à une mission particulière. On cherche actuellement à prendre en compte le plus d'informations et donc de sources d'informations possibles afin d'aider les opérateurs dans leur différentes tâches de prise de décision. Pour cela, les informations utilisées par les systèmes pour raisonner ne sont plus uniquement celles des capteurs physiques et observateurs humains entraînés à l'acquisition de connaissances. Dans le cadre du renseignement par exemple, il est maintenant nécessaire de prendre en compte les informations circulant sur des sources diverses et issues de citoyens, comme les réseaux sociaux par exemple. Ces sources, de par leur déploiement très large, permettent d'acquérir de l'information sur des situations qui n'auraient pas été observées auparavant. Cependant, l'exploitation de ces informations pose de nouveaux challenges pour les outils d'aide à la décision.

Dans le cadre de cette thèse, nous nous intéressons au verrou scientifique que représente l'intégration d'une dimension d'imperfection dans les outils Thales. En effet, ces outils requièrent aujourd'hui d'intégrer des sources d'informations dites OSINT¹ et HUMINT². Les informations provenant de ces sources sont imparfaites de par leur nature et provenance. La capacité des humains à échanger et interpréter des informations imparfaites, en utilisant cette imperfection pour obtenir une vision subtile d'une situation qui leur est rapportée reste un challenge pour les algorithmes d'intelligence artificielle.

1. Open Source INtelligence - informations issues de sources ouvertes.

2. HUMain INtelligence - renseignement d'origine humaine (a contrario de capteurs physiques).

Il s’agira donc d’intégrer la dimension imparfaite de l’information, d’une part dans sa représentation, via un formalisme adapté, et d’autre part prendre en compte cette imperfection dans les mécanismes de manipulation de cette information. On s’intéressera en particulier, au cours de cette thèse, aux opérations de fusion et de recherche de motifs fréquents dans les bases d’informations sémantiques.

Contexte formel

Le cadre général précédent soulève le problème essentiel de la représentation des informations considérées, qui a des implications majeures sur les tâches de fusion et de recherche de motifs fréquents citées, à la fois en termes de formulation et de résolution.

Représentation de la connaissance incertaine Pour stocker des connaissances, les ontologies sont actuellement très populaires, car elles permettent de définir les idées et les concepts manipulés dans un domaine spécifique, puis de raisonner sur ces concepts afin d’inférer de nouvelles connaissances à leur sujet.

Les graphes conceptuels (voir par exemple Chein & Mugnier, 2008) constituent un cadre de représentation des connaissances plus riche que les ontologies, qui distingue explicitement les connaissances ontologiques des connaissances factuelles : les premières définissent un vocabulaire, sous la forme d’une terminologie hiérarchisée, utilisé pour construire les secondes, en la forme de graphes bipartis dont les nœuds et les arêtes sont étiquetés avec des éléments du vocabulaire. Les graphes conceptuels offrent l’avantage de combiner une sémantique logique et des outils de manipulation efficaces basés sur la théorie des graphes. Le sujet de thèse proposé considère pour cela le formalisme des graphes conceptuels, introduit par Sowa (1984), largement étudié d’un point de vue théorique et exploité d’un point de vue pratique (voir par exemple Chein et Mugnier (2009) ou Laudy (2010)) : celui-ci est en effet particulièrement approprié pour les informations sémantiques, pour lesquelles il offre un cadre formel flexible, riche et très expressif, qui est, de plus, muni d’outils de manipulation performants.

Le cadre théorique des graphes conceptuels a été appliqué à de nombreux domaines (Rajangam & Annamalai, 2016) tels que la gestion de dossiers médicaux (Delamarre et al., 1995), la gestion documentaire (Genest & Chein, 2005), la gestion de crise (Delavallade et al., 2015), la détection de rumeurs sur les réseaux sociaux (Laudy, 2017) ou la cybersécurité (Aguessy et al., 2015; Laudy & Fontarensky, 2017) par exemple.

Fusion d’informations incertaines Pour des informations représentées sous la forme de graphes conceptuels, les problématiques de filtrage et d’agrégation peuvent être vues dans le cadre de la théorie des graphes et résolues selon plusieurs approches (Laudy, 2011). Le filtrage, qui porte le nom de *fusion*, peut être exprimé comme une tâche d’isomorphisme de graphes : le schéma général que l’on cherche à instancier, représentant un modèle de situation d’intérêt, est représenté par un graphe conceptuel dont les nœuds conceptuels sont étiquetés par la constante générique. La mise en correspondance de ce graphe de référence avec les graphes conceptuels représentant les informations disponibles, ou observations, doit spécifier cette constante générique et l’instancier. Une extension de cette opération de filtrage au cas d’informations incertaines a été proposée (Fossier et al., 2013) : elle construit une distribution de probabilité sur un ensemble de graphes conceptuels candidats, résultant de plusieurs instanciations possibles, selon la certitude associée aux informations fusionnées.

Recherche de motifs fréquents L’extraction de connaissances à partir d’un graphe conceptuel ou d’un ensemble de graphes conceptuels peut prendre la forme de l’identification de motifs fréquents, c’est-à-dire de sous-graphes qui se produisent fréquemment et peuvent donc être interprétés comme des régularités pertinentes. Cela concerne le domaine du Frequent Subgraph Mining, pour lequel de nombreuses approches ont été proposées (voir par exemple Jiang et al., 2013). Des algorithmes dédiés ont été proposés pour le cas particulier des graphes étiquetés taxonomiques (Inokuchi et al., 2000; Cakmak & Ozsoyoglu, 2008; Petermann et al., 2017), qui

exploitent les caractéristiques particulières de ces types de graphes afin d'améliorer à la fois leur efficacité et la pertinence des motifs extraits.

Des travaux récents menés dans l'équipe du LFI se sont focalisés plus avant sur le cas des graphes conceptuels : l'algorithme cgSpan (Faci et al., 2021) exploite l'information sur l'arité des relations pour diminuer efficacement le nombre de motifs candidats explorés, les signatures de relation pour élaguer les motifs candidats redondants avec peu d'informations ainsi que des règles d'inférence pour étendre les candidats plus rapidement.

3 Problématique de la thèse : Raisonnement sur les graphes sémantiques flous pour la découverte de connaissances

La thèse proposée s'inscrit dans une problématique générale de permettre la représentation et la manipulation des connaissances et observations manipulées habituellement par des humains. Ces informations et connaissances, contiennent naturellement différents niveaux de précision et d'imperfections, en fonction des compétences et expertises des humains qui les transmettent et les manipulent.

Un certain degré d'imperfection dans les informations acquises n'empêche cependant pas les opérateurs humains de raisonner et déduire des connaissances. Nous souhaitons dans les travaux proposés ici, aller vers de telles capacités pour des algorithmes d'intelligence artificielle.

La problématique étudiée dans le cadre de la thèse s'articule ainsi autour des 2 axes suivants :

- l'extension du formalisme des graphes conceptuels avec des pondérations,
- la prise en compte de cette extension dans les opérations de fusion.

Extension du formalisme des graphes conceptuels avec des pondérations Comme évoqué précédemment, des travaux antérieurs existent, pour intégrer des notions modélisées dans le cadre de la théorie des sous-ensembles flous dans les graphes conceptuels. L'objectif ici est d'étudier la nécessité d'intégrer d'autres types d'imperfection. Il s'agira aussi de caractériser la sémantique associée à chaque forme d'imperfection introduite dans le formalisme, afin de pouvoir les associer aux différents besoins potentiellement exprimés, aux niveaux opérationnels et applicatifs.

Prise en compte de l'imperfection de l'information dans les opérations de fusion Que ce soit pour l'opération de fusion de deux graphes conceptuels en un seul, ou pour la recherche de motifs fréquents dans une base de graphes conceptuels, l'enjeu des travaux de thèse est de prendre en compte l'extension qui aura été proposée dans la première phase, dans les algorithmes existant de fusion et recherche de motifs fréquents.

4 Méthodologie envisagée

Cette section présente les outils et approches envisagés pour aborder la problématique présentée dans la section précédente et précise le travail à réaliser.

4.1 Outils et approches considérés

Extensions pondérées des graphes conceptuels La première composante du travail de thèse portera sur la définition de cadres étendus par rapport aux graphes conceptuels classiques, et notamment l'introduction de pondérations, selon diverses sémantiques. En effet, comme mentionné précédemment, il pourra être intéressant d'associer les nœuds ou les arcs du graphes à des coefficients numériques : ceux-ci pourront par exemple autoriser des appartenances partielles au graphe complet, afin de représenter des composantes optionnelles du modèle de situation, ou des degrés de certitude, associés à des connaissances incertaines. Une autre extension pourra s'intéresser à des pondérations à l'intérieur des nœuds, par exemple pour formaliser des distributions sur les valeurs associées.

S'il existe des travaux antérieurs sur les graphes conceptuels pondérés, notamment flous (Thomopoulos et al., 2003; Faci et al., 2021), le cadre de la thèse proposera d'étudier les différentes sémantiques associées aux propositions, ainsi que l'étude de leur adéquation avec les différents types de besoins opérationnels en termes de représentation de connaissances imparfaites, et raisonnement.

D'autres approches (voir par exemple Couchariere (2010)) pourront également être explorées, en raison des liens entre graphes conceptuels et logiques de description (Sattler et al., 2003).

Nouvelles méthodes d'agrégation Une deuxième composante de la thèse portera sur la tâche d'agrégation de graphes conceptuels et visera à proposer des extensions des approches existantes : d'une part, il s'agira de définir des opérateurs exprimant des sémantiques variées, en transposant par exemple les catégories conjonctive, disjonctive, de compromis ou avec renforcement qui existent pour les opérateurs d'agrégation numériques (voir par exemple Detyniecki (2000)). D'autre part, il s'agira d'exploiter les extensions pondérées envisagées, pour la forme du résultat fourni ainsi que pour la forme des graphes considérés en entrée de l'agrégation, par exemple pour la prise en compte des degrés de criticité exprimés par les utilisateurs.

4.2 Travail à réaliser

D'un point de vue pratique, le travail à réaliser comportera les composantes classiques d'une thèse en informatique, constituées notamment d'une étude bibliographique, la formalisation des problèmes considérés, la proposition de solutions originales, leur étude expérimentale et leur validation.

L'étude bibliographique, la formalisation et la proposition de solutions originales se placeront dans le cadre décrit ci-dessus et porteront sur les multiples sujets à l'intersection desquels se place la thèse, au delà des quelques références indicatives citées précédemment. L'étude expérimentale et la validation constitueront une partie importante du travail de thèse. Elles s'appuieront sur des implémentations des propositions et le développement de prototypes avancés.

5 Enjeux

Comme mentionné à la fin de la section 2, le cadre théorique des graphes conceptuels a été à la base de nombreuses applications aussi variées.

Dans le cadre de ses travaux, le laboratoire LRASC de Thales a développé plusieurs outils afin de permettre la représentation de situations à un niveau sémantique. L'intérêt de ce niveau de représentation est qu'il rend les situations compréhensibles et interprétables par des opérateurs humains, ainsi que par des programmes informatiques, permettant ainsi l'utilisation d'algorithmes d'intelligence artificielle pour aider l'opérateur humain dans l'analyse de grandes quantités d'informations.

Parmi ces outils nous pouvons citer :

InSyTo. La boîte à outils InSyTo (Delavallade et al., 2015) met à disposition des algorithmes de fusion d'information sémantique, et de requête dans les graphes sémantiques. Elle est utilisée dans des applications aussi variées que la gestion de crise, l'analyse et le renseignement stratégique, la détection de rumeurs, la cybersécurité, la gestion des ressources maritimes et la surveillance de frontières. Ces applications vont de la preuve de concept théorique au prototypage en vue d'intégration dans des outils déployés opérationnellement.

Knous. Knous est un outil centralisant plusieurs algorithmes de recherche de motifs fréquents dans une base de graphes conceptuels. Il comprend notamment l'algorithme cgSpan (Faci et al., 2021), développé lors d'une précédente thèse co-encadrée par le LIP6 et Thales.

Ces deux outils sont utilisés dans des applications de différentes unités du groupe Thales, entre autres, pour des applications de C2 (Command & Control), de surveillance maritime ou encore de maintenance en conditions opérationnelles. Dans ces applications, des besoins de représentation

et d'analyse des dimensions incertaine et floue des informations ont été mises en exergues au fil des projets développés. Les exemples ci-dessous illustrent l'intérêt et les gains attendus des résultats de la thèse pour ces applications.

Command & Control Dans cette application, l'objectif est de soutenir, à l'aide d'analyse de textes courts, un opérateur de renseignement dans sa tâche de collecte et d'analyse d'indices de changement de situation géopolitique. L'enjeu de l'intégration des résultats de la thèse dans la réponse apportée est l'intégration de la dimension pondérée des indices trouvés. En effet, les textes produits par des humains, contiennent de nombreuses imprécisions quant à la nature des faits observés, ainsi que leur intensité. L'application actuellement développée ne permet que de représenter des indices entièrement ou pas du tout trouvés dans les textes. Le fait d'utiliser des représentations sémantiques classiques, binaires et non floues (aussi appelées crisp), pour des sources d'informations issus d'observations et rapports d'observateurs humains prive de la subtilité naturelle exprimée et peut avoir de lourdes conséquences sur les décisions prises. Intégrer la dimension imparfaite de l'information, à la fois dans la représentation acquise informatiquement et dans les processus d'analyse est donc crucial.

Surveillance maritime Le cas considéré, dans une application de surveillance maritime, est la recherche d'événements d'intérêt à aller surveiller, en marge d'une mission de blanchiment de zone. En d'autres termes, l'équipage a pour mission de vérifier qu'il n'y a pas de problème dans une zone donnée. La mission est a priori sans risque, et le pilote peut, si les conditions s'y prêtent, en "profiter" pour aller prendre des informations sur des événements ayant lieu à proximité de sa zone de mission, en cas d'événements qui auraient débuté après le début de la mission, donc non inclus initialement. Pour détecter de tels événements, on s'appuie sur des informations rapportées par ces citoyens, notamment au travers de réseaux sociaux. Il s'agit donc de collecter et interpréter des informations issues de rapports faits par des observateurs potentiellement non professionnels du renseignement. Comme dans le cas précédent, il est donc primordial de pouvoir ajouter, à la réponse actuellement fournie par les systèmes Thales, une dimension de représentation et d'analyse de l'imperfection de l'information.

Maintien en conditions opérationnelles. Dans le cadre du maintien en conditions opérationnelles des systèmes déployés par Thales, nous proposons d'utiliser les algorithmes de InSyTo et Knous afin de rechercher, dans des bases de données d'historique de pannes, des causes probables à des pannes observées. Pour ce faire, nous proposons de rechercher, dans l'historique, des symptômes proches des ceux nouvellement observés, ainsi que les actions de réparation qui ont été effectuées pour le maintien en conditions opérationnelles. En l'état actuel des outils InSyTo et Knous, nous ne sommes en mesure que de travailler sur des descriptions complètes et certaines des symptômes observés et des pannes passées. Il est nécessaire de pouvoir relâcher les contraintes d'observations de symptômes strictement identiques, afin de proposer une réponse qui s'adapte aux nouvelles conditions d'utilisation et donc de pannes. Les résultats des travaux de thèse permettront de fournir ces outils de recherche de motifs de pannes prenant en compte un degré d'incertitude, par exemple, dans les nouveaux symptômes observés.

Les trois exemples décrits ci-dessus ne sont pas exhaustifs des applications Thales qui bénéficieront des résultats de la thèse, mais ils mettent en évidence le caractère nécessaire des travaux qui permettront d'inclure dans les produits Thales une représentation et une gestion des informations imparfaites.

L'objectif opérationnel pour Thales est d'agrandir les perspectives d'intégration des outils InSyTo et Knous dans les outils déployés opérationnellement.

6 Calendrier prévisionnel

Publications Tout au long de la thèse, des articles seront soumis à des conférences, pour valider scientifiquement l'avancement ; des articles de journaux seront également considérés, à la fin de

la seconde année ainsi que durant la troisième année.

Le choix des conférences et journaux se fera selon le contenu envisagé pour chaque article, les possibilités suivantes ont été identifiées à titre préliminaire : Int. Conf. on Information Fusion (FUSION), IEEE Int. Conf. on Fuzzy Systems (FuzzIEEE), Int. Conf. on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU), Conf. of the European Society for Fuzzy Logic and Technology (EUSFLAT), Int. Conf. on Scalable Uncertainty Management (SUM), European Conf. on Symbolic and Quantitative Approaches to Reasoning and Uncertainty (ECQ SARU) et, en ce qui concerne les journaux, Fuzzy Sets and Systems, IEEE Trans. on Fuzzy Systems, IEEE Trans. on Knowledge and Data Engineering et ISIF Journal of Advances in Information Fusion (JAIF) par exemple.

Organisation du travail Pendant les trois années de thèse, il est prévu que le doctorant passe 60% de son temps dans l'entreprise et 40% au sein du laboratoire. Des réunions hebdomadaires auront lieu avec tous les encadrants, en plus de points fréquents selon les besoins d'avancement.

7 Profil recherché et candidature

Le/la candidat(e) doit posséder un master en intelligence artificielle/apprentissage automatique, avec des connaissances en logique formelle et représentation des connaissances. Le langage de programmation privilégié pour les travaux qui seront réalisés pendant la thèse est python.

Pour postuler : Envoyer un mail à Claire.Laudy@thalesgroup.com et marie-jeanne.lesot@lip6.fr, en mettant dans le sujet [Candidature thèse] et en joignant un CV, une lettre de motivation et des relevés de notes des formations antérieures.

Références

- Aguessy, F.-X., Bettan, O., Dobigny, R., Laudy, C., Lortal, G., & Faure, D. (2015). Adjustable fusion to support cyber security operators. *Human Aspects of Information Security, Privacy, and Trust* (pp. 143–153). Cham : Springer International Publishing.
- Cakmak, A., & Ozsoyoglu, G. (2008). Taxonomy-superimposed graph mining. *Proc. of the 11th Int. Conf. on Extending database technology : Advances in database technology* (pp. 217–228).
- Chein, M., & Mugnier, M.-L. (2008). *Graph-based knowledge representation : Computational foundations of conceptual graphs*. Springer.
- Chein, M., & Mugnier, M.-L. (2009). *Graph-based knowledge representation : Computational foundations of conceptual graphs*. Springer.
- Couchariere, O. (2010). *Logique de description ALC possibiliste pour les systèmes d'information opérationnels : théorie, algorithme et applications*. Thèse de doctorat, Université Pierre et Marie Curie, Paris VI.
- Delamarre, D., Burgun, A., Seka, L. P., & Beux, P. L. (1995). Automated coding of patient discharge summaries using conceptual graphs. *Methods Inf Med*.
- Delavallade, T., Fossier, S., Laudy, C., & Lortal, G. (2015). On the challenge of using social media for crisis management. In G. Rogova and P. Scott (Eds.), *Fusion methodologies in crisis management*. Springer International Publishing.
- Detyniecki, M. (2000). *Mathematical aggregation operators and their application to video querying*. Thèse de doctorat, Université Pierre et Marie Curie, Paris VI.
- Faci, A., Lesot, M.-J., & Laudy, C. (2021). cgspan : Pattern mining in conceptual graphs. *Proc. of the Int. Conf. on Artificial Intelligence and Soft Computing* (pp. 149–158). Springer, LNCS12855.
- Faci, A., Lesot, M.-J., & Laudy, C. (2021). cgSpan : Pattern Mining in Conceptual Graphs. *Proc. of the Int. Conf. on Artificial Intelligence and Soft Computing* (pp. 149–158).

- Faci, A., Lesot, M.-J., & Laudy, C. (2021). Fuzzy conceptual graphs : a comparative discussion. *Proc. of the Symposium Series on Computational Intelligence, Foundation of Computational Intelligence (SSCI-FOCI21)*.
- Fossier, S., Laudy, C., & Pichon, F. (2013). Managing uncertainty in conceptual graph-based soft information fusion. *Proc. of the 16th Int. Conf. on Information Fusion, FUSION13* (pp. 930–937).
- Genest, D., & Chein, M. (2005). A content-search information retrieval process based on conceptual graphs. *Knowledge and Information Systems*, 8, 292–309.
- Inokuchi, A., Washio, T., & Motoda, H. (2000). An apriori-based algorithm for mining frequent substructures from graph data. *Proc. of the European Conf. on principles of data mining and knowledge discovery* (pp. 13–23).
- Jiang, C., Coenen, F., & Zito, M. (2013). A survey of frequent subgraph mining algorithms. *The Knowledge Engineering Review*, 28, 75–105.
- Laudy, C. (2010). *Fusion multi-sources d'informations de haut niveau : introduction de connaissances sémantiques pour la gestion des incohérences*. Thèse de doctorat, Université Pierre et Marie Curie, Paris VI.
- Laudy, C. (2011). *Semantic knowledge representations for soft data fusion : efficient decision support systems. practice and challenges from current to future*. Chiang Jao publisher.
- Laudy, C. (2017). Rumors detection on social media during crisis management. *14th International Conference on Information Systems for Crisis Response And Management*.
- Laudy, C., & Fontarensky, Y. (2017). Leak, data privacy and ... hacking. *Computer & Electronics Security Applications Rendez-vous*.
- Petermann, A., Micale, G., Bergami, G., Pulvirenti, A., & Rahm, E. (2017). Mining and ranking of generalized multi-dimensional frequent subgraphs. *Proc. of the 12th Int. Conf. on Digital Information Management (ICDIM)* (pp. 236–245).
- Rajangam, E., & Annamalai, C. (2016). Graph models for knowledge representation and reasoning for contemporary and emerging needs - a survey. *Int. J. Information Technology and Computer Science*, 2, 14–22.
- Sattler, U., Calvanese, D., & Molitor, R. (2003). Relationships with other formalisms. In F. Baader, D. Calvanese, D. L. McGuinness, D. Nardi and P. F. Patel-Schneider (Eds.), *The description logic handbook : Theory, implementation and applications*, chapter 4, 137–177. Cambridge University Press.
- Sowa, J. F. (1984). *Conceptual structures : information processing in mind and machine*. Addison-Wesley.
- Thomopoulos, R., Buche, P., & Haemmerlé, O. (2003). Representation of weakly structured imprecise data for fuzzy querying. *Fuzzy Sets and Systems*, 140, 111–128.