

## ■ Apprentissage Artificiel et Représentation des Connaissances

LERIA/ARC  
Université d'Angers  
[www.leria.univ-angers.fr](http://www.leria.univ-angers.fr)

**Béatrice DUVAL**

[beatrice.duval@univ-angers.fr](mailto:beatrice.duval@univ-angers.fr)

**Stéphane LOISEAU**

[stephane.loiseau@univ-angers.fr](mailto:stephane.loiseau@univ-angers.fr)

### Membres

- Touria AÏT EL MEKKI, MCF
- Tassadit AMGHAR, MCF-HDR
- Benoît DA MOTA, MCF
- Béatrice DUVAL, PR
- David GENEST, MCF
- Olivier GOUDET, MCF
- Nicolas GUTOWSKI, MCF
- Olivier HÛ, MCF
- Stéphane LOISEAU, PR
- Aïcha BAKKI, ECER
- Antoine JAMIN, ECER
- Nathan FRADET, Doctorant
- Jules LEGUY, Doctorant
- Alexandre LETARD, Doctorant

### Autres participants

- Laurent GARCIA, MCF
- Jin-Kao HAO, PR

### Les thèmes de recherche du LERIA

Le Laboratoire d'Étude et de Recherche en Informatique d'Angers (LERIA) mène des recherches de nature fondamentale et appliquée dans deux domaines connexes de l'informatique : l'intelligence artificielle et l'optimisation combinatoire. Les sujets abordés forment un continuum que l'on peut décliner en quelques grandes thématiques :

- Méthodes de résolution approchée pour l'optimisation combinatoire issues du calcul évolutionnaire ;
- Techniques de modélisation, reformulation et résolution exacte fondées sur la programmation

par contraintes et la logique propositionnelle ;

- Langages et algorithmes pour le raisonnement non-monotone fondées sur la programmation par ensembles réponses et autres formalismes de contraintes quantifiées ;
- Méthodes d'apprentissage artificiel en appui de techniques de résolution ou portées à d'autres champs disciplinaires tels la bio-informatique et la chimie quantique ;
- Méthodes d'interrogation de bases de connaissances fondée sur des modèles de représentation graphiques.

À ces travaux théoriques s'ajoutent le développement d'approches à visée applicative. Ces travaux, souvent menés dans le cadre de collaborations industrielles ou de projets pluridisciplinaires, recouvrent des domaines variés. Citons entre autres exemples la résolution de problèmes logistiques par optimisation combinatoire, l'analyse logique de données fondée sur le raisonnement symbolique, et le développement de systèmes de recommandation par apprentissage artificiel.

Le LERIA est organisé en 3 thèmes d'activité :

- Thème **MOC** : Méta-heuristiques et Optimisation Combinatoire
- Thème **RIC** : Raisonnement dans l'Incertain et Contraintes
- Thème **ARC** : Apprentissage Artificiel et Représentation des Connaissances.

## Le thème ARC

Les travaux du thème ARC proposent des solutions à base de modèles de représentation des connaissances et de méthodes d'apprentissage artificiel pour l'analyse et la compréhension de problèmes scientifiques ou applicatifs. Concernant la représentation des connaissances, les apports concernent notamment les données liées (linked data) et les modèles graphiques pour l'élicitation et le raisonnement sur les connaissances métiers d'un domaine. Concernant l'apprentissage, les travaux portent, d'une part, sur les systèmes de recommandation, et d'autre part sur l'inférence de connaissances à partir de données dans le domaine de la biologie et de la chimie moléculaire. En collaboration avec le thème MOC, des travaux sont également menés pour intégrer des méthodes d'apprentissage et de fouille de données au sein des recherches méta-heuristiques

### Représentation et interrogation dans les modèles du web des données et des cartes cognitives [25, 26, 27, 24, 23, 28, 15, 22]

Deux types de travaux ont été menés autour de ces problématiques. D'une part, le web des données est aujourd'hui particulièrement populaire pour stocker des connaissances avec RDF et OWL. Les contributions de nos travaux ont consisté d'abord à fournir une approche plus performante pour interroger en SPARQL des bases (par l'usage d'aperçus des bases interrogées), ensuite à proposer une méthodologie permettant de faire évoluer les bases tout en garantissant la non-dégradation de celles-ci, enfin en proposant une nouvelle approche visuelle et algorithmique pour construire des requêtes SPARQL en s'inspirant des travaux de graphes conceptuels. Ces travaux ont été présentés dans les conférences EGC (2015) et RFIA (2016). D'autre part, les cartes cognitives sont un modèle visuel de représentation des connaissances d'influence. Les contribu-

tions des travaux menés ont porté sur l'extension du modèle à des connaissances temporelles de type intervalles périodiques, et sur la proposition d'un langage et d'un moteur de requêtes des cartes qui permettent d'analyser les cartes. Une partie de ce travail a été menée à partir d'une application réelle de 50 cartes produites par des marins pêcheurs de la région pays de la Loire.

### Systèmes de recommandation [19, 21, 20, 7, 8, 5, 6, 4]

Les travaux sur ce sujet sont consacrés à deux problèmes, la recommandation contextuelle de services d'une part, et l'apprentissage multi-critères guidé par les besoins des utilisateurs d'autre part. Les algorithmes de bandits-manchots pour les systèmes de recommandation sensibles au contexte font l'objet de nombreuses études. Afin de répondre aux enjeux de cette thématique, nos contributions se sont organisées autour de 3 axes : les systèmes de recommandation, les algorithmes de bandits-manchots (contextuels et non contextuels), et le contexte. La première partie des contributions a porté sur les algorithmes de bandits-manchots pour la recommandation. Elle aborde la diversification des recommandations visant à améliorer la précision individuelle. La seconde partie a porté sur la capture de contexte, le raisonnement contextuel pour les systèmes de recommandation d'événements culturels dans la ville intelligente, et l'enrichissement dynamique de contexte pour les algorithmes de bandits-manchots contextuels. Ces travaux ont été réalisés dans le cadre d'une thèse et ont donné lieu à une preuve de concept dans le cadre d'un projet financé par la région. Concernant l'apprentissage multi-critères guidé par les besoins des utilisateurs, ces travaux visent la construction d'un système de navigation de plaisance intelligent et se positionnent dans le cadre des systèmes de recommandation. Ils portent sur l'utilisation de bandits man-

chots combinatoires (COM-MAB) pour lesquels nous avons proposé de pouvoir utiliser un vecteur de récompenses partiel quand l'acquisition d'un vecteur complet nécessitant alors des sollicitations directes et trop fréquentes de l'utilisateur, semble difficile. C'est le cas dans un environnement maritime. L'approche que nous avons proposée permet de maintenir une précision globale proche de celle obtenue par les méthodes utilisant des vecteurs complets. Ces travaux font l'objet d'une collaboration industrielle.

### **Intelligence artificielle et chimie moléculaire**

**[16, 2, 3, 1, 17, 18]** La découverte de molécules possédant des propriétés particulières est un objectif majeur en chimie, et depuis quelques années, de nombreux travaux ont étudié l'intérêt de guider ce processus par des méthodes d'intelligence artificielle. Concernant la génération de molécules prometteuses pour les matériaux moléculaires, qui intéresse notre partenaire chimiste, de nombreux verrous scientifiques sont à lever. En effet, l'espace chimique est gigantesque et peu connu alors même que certains îlots ont été largement explorés. De plus, le calcul des propriétés des molécules est un processus très coûteux; certains travaux essaient donc d'estimer les propriétés cibles par apprentissage artificiel. Après avoir proposé une méthode d'apprentissage supervisé pour prédire des distances interatomiques, nous avons démontré que les jeux de données disponibles, rares et peu diversifiés, ne permettent pas à ces méthodes de correctement généraliser. Nous avons alors défini une mesure de diversité, chimiquement pertinente, afin de générer de nouveaux jeux de données d'apprentissage. Une plateforme de calcul collaboratif et un système d'information de type Big Data permettent de calculer et partager nos résultats, de les curer et les mettre en forme. La génération de molécules est abordée comme un problème d'opti-

misation de la propriété cible sous contraintes (stabilité, synthétisabilité et préférences du chimiste). A l'opposé de certaines méthodes de la littérature basées sur des architecture complexes d'apprentissage profond, nous proposons un générateur moléculaire, flexible et interprétable, dont le cœur est un algorithme évolutionnaire opérant sur une représentation graphe des molécules. Les résultats obtenus sur des benchmarks de natures diverses, mais explorant un espace encore réduit, sont comparables à l'état de l'art, et permettent donc d'envisager l'hybridation de ce générateur avec des méthodes d'apprentissage lui permettant une exploration plus large de l'espace moléculaire. Ces travaux ont été financés par la commission recherche de l'Université d'Angers et par la région Pays de la Loire et son dispositif RFI Atlantic 2020 (projet AIQU) dans le cadre de la bourse postdoctorale de Marta Glavatskikh en 2019 et par une bourse doctorale ministérielle dans le cadre de la thèse de Jules LEGUY (2019-2022). Ces travaux sont réalisés en collaboration avec l'UMR de chimie MOLTECH-Anjou. Des outils libres sont mis à disposition de la communauté scientifique et le projet communique avec le public et l'implique à travers sa plateforme de calculs collaboratifs quchempedia@home.

### **Inférences de connaissances biologiques**

**[11, 14, 13, 12, 10]** Un des problèmes actuels en bio-informatique est de comprendre les mécanismes de régulation au sein d'une cellule ou d'un organisme en exploitant les données issues de plateformes de biologie moléculaire. En collaboration avec l'Institut de Recherche en Horticulture et Semences (IRHS), nous avons étudié les réseaux de co-expression de gènes chez le pommier avec la particularité d'y intégrer les transcrits anti-sens qui sont des ARN généralement non-codants, dont les différents modes d'action sont encore mal connus. Dans

notre étude exploratoire du rôle des anti-sens, nous avons proposé d'une part une analyse fonctionnelle différentielle qui met en évidence l'intérêt de l'intégration des données anti-sens en transcriptomique. D'autre part, concernant les réseaux de gènes pour lesquels de nombreuses méthodes d'apprentissage ont été développées, nous avons proposé de limiter l'inférence à un cœur de réseau qui identifie les interactions majeures au sein du réseau. De plus, nous avons introduit une méthode d'analyse différentielle de réseaux qui a fourni des pistes pertinentes pour l'étude de voies de signalisation impactées par les anti-sens. Ces travaux ont été portés par le projet régional GRIOTE

### Autres contributions

- *Inférence causale* [9]  
En collaboration avec l'équipe TAU de l'INRIA à Paris Saclay, nous avons développé des méthodes d'inférences de réseaux de causalité à partir de données d'observation dont la méthode SAM (Structural Agnostic Model). SAM met en jeu des outils d'apprentissage profond pour apprendre un modèle fonctionnel causal qui a vocation à être un outil d'aide à la décision afin de comprendre et prédire les effets d'interventions sur le système étudié. Ces méthodes ont été intégrées dans une nouvelle librairie pour Python, CausalDiscoveryToolbox, disponible en ligne.
- *Évaluation des dispositifs numériques de médiation culturelle* Les outils de médiation numériques sont en pleine expansion tant dans le domaine culturel que touristique. Nos travaux visent entre autres la constitution d'une grille multi-critères d'évaluation prenant en considération les dimensions d'accessibilité, d'utilisabilité, d'émotion et de persuasion des interfaces, et de corrélérer cette évaluation avec l'étude de l'expérience vécue par les usagers. Ces travaux ont

été financés par la région Pays-de-la-Loire et ses dispositifs RFI TourismLab et RFI Ouest Industrie Créative et ont donné lieu au financement d'une bourse post-doctorale en 2018/2019. Ces travaux ont fait l'objet d'une publication dans le colloque international "Méthodes visuelles dans les recherches sur la communication" en 2018, et dans différents colloques comme "Innovation dans le patrimoine" en 2018.

- *Intelligence artificielle pour l'archivistique* La messagerie électronique est une archive au sens mémoriel comme au sens informatique et l'ensemble des messageries d'un organisme sont une archive de cet organisme. L'objet de ce projet est de tester des stratégies d'évaluation archivistique sur les messageries électroniques, pour aboutir à une méthodologie d'exploration et d'aide à la décision de sélection des messages sur la base de critères de pertinence et de représentativité. Le projet fait appel à des compétences croisées en matière d'évaluation archivistique et d'intelligence artificielle (TALN et apprentissage artificiel). Ce travail mené en collaboration avec le laboratoire Temos UMR CNRS 9016-Université d'Angers et le centre Jean Mabillon (École nationale des Chartes) fait l'objet du projet ANR BALise dans le cadre de l'AAP 2020 Révolution numérique : rapports au savoir et à la culture. Le projet est en phase deux de réponse aux experts.

### Références

- [1] Thomas Cauchy and Benoit Da Mota. Prédire une géométrie moléculaire convergée par des modèles d'apprentissage automatique. In *Workshop " L'intelligence artificielle pour la chimie des matériaux "*, Thiais, France, September 2018.
- [2] Marta Glavatskikh, Jules Leguy, Benoit Da Mota, and Thomas Cauchy. Machine learning in prediction of DFT-derived geo-

- metries and energies. Journées Scientifiques SCF-BPL 2019, April 2019. Poster.
- [3] Marta Glavatskikh, Jules Leguy, Gilles Hunault, Thomas Cauchy, and Benoit Da Mota. Dataset's chemical diversity limits the generalizability of machine learning predictions. *Journal of Cheminformatics*, 11(1), December 2019.
- [4] Nicolas GUTOWSKI, Tassadit Amghar, Olivier Camp, and Fabien Chhel. Global Versus Individual Accuracy in Contextual Multi-Armed Bandit. In *The 34th ACM/SIGAPP Symposium on Applied Computing (SAC '19)*, pages 1647–1654., LIMASSOL, Cyprus, April 2019.
- [5] Nicolas GUTOWSKI, Tassadit Amghar, Olivier Camp, and Fabien Chhel. Gorthaur : A Portfolio Approach for Dynamic Selection of Multi-Armed Bandit Algorithms for Recommendation. In *IEEE International Conference on Tools with Artificial Intelligence (ICTAI)*, pages 1164–1171, Portland, Oregon, United States, November 2019.
- [6] Nicolas Gutowski, Tassadit Amghar, Olivier Camp, and Fabien Chhel. Gorthaur-EXP3 : Bandit-based selection from a portfolio of recommendation algorithms balancing the accuracy-diversity dilemma. *Information Sciences*, 546 :378–396, February 2021.
- [7] Nicolas GUTOWSKI, Olivier Camp, Tassadit Amghar, and Fabien Chhel. Using Individual Accuracy to Create Context for Non-Contextual Multi-Armed Bandit Problems. In *IEEE Research, Innovation and Vision for the Future*, pages 1–6, Danang, France, March 2019. IEEE.
- [8] Nicolas GUTOWSKI, Olivier Camp, Fabien Chhel, Tassadit Amghar, and Patrick Albers. Improving Bandit-Based Recommendations with Spatial Context Reasoning : An Online Evaluation. In *IEEE International Conference on Tools with Artificial Intelligence (ICTAI)*, pages 1366–1373, Portland, Oregon, United States, November 2019.
- [9] Diviyam Kalainathan, Olivier Goudet, and Ritik Dutta. Causal Discovery Toolbox : Uncovering causal relationships in Python. *Journal of Machine Learning Research*, 21 :1–5, March 2020.
- [10] Marc Legeay, Sébastien Aubourg, Jean-Pierre Renou, and Béatrice Duval. Large scale study of anti-sense regulation by differential network analysis. *BMC Systems Biology*, 12(S5) :95, 2018.
- [11] Marc Legeay, Béatrice Duval, and Jean-Pierre Renou. Context-specific co-expression networks to explore the impact of anti-sense transcription. 13th conference on Computational Methods for Systems Biology (CMSB'15), September 2015. Poster.
- [12] Marc Legeay, Béatrice Duval, and Jean-Pierre Renou. Differential Functional Analysis and Change Motifs in Gene Networks to Explore the Role of Anti-sense Transcription. In *Bioinformatics Research and Applications*, pages 117–126. Springer International Publishing, 2016.
- [13] Marc Legeay, Béatrice Duval, and Jean-Pierre Renou. Inference and Differential Analysis of Extended Core Networks : a way to study Anti-Sense Regulation. In *IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, page 1939 p., Shenzhen, China, 2016. IEEE Computer Society.
- [14] Marc Legeay, Béatrice Duval, and Jean-Pierre Renou. Differential Network Analysis of Anti-sense Regulation. In *Bioinformatics and Biomedical Engineering*,

- volume 10209 of *Lecture Notes in Computer Science*, pages 277–288. Springer International Publishing, 2017.
- [15] Marc Legeay, David Genest, and Stéphane Loiseau. Un langage d’interrogation à la SPARQL pour les graphes conceptuels. In *Extraction et Gestion des Connaissances (EGC 2015)*, In : *Revue des Nouvelles Technologies de l’Information*, volume E-28, pages 227–238, Non spécifié, France, 2015. Cépaduès.
- [16] Jules Leguy, Thomas Cauchy, Béatrice Duval, and Benoit Da Mota. Des réseaux de neurones pour prédire des distances interatomiques extraites d’une base de données ouverte de calculs en chimie quantique. In *Extraction et Gestion des connaissances, EGC 2019*, Metz, France, January 2019.
- [17] Jules Leguy, Thomas Cauchy, Marta Glavatskikh, Béatrice Duval, and Benoit Da Mota. EvoMol : a flexible and interpretable evolutionary algorithm for unbiased de novo molecular generation. *Journal of Cheminformatics*, 12(1), December 2020.
- [18] Jules Leguy, Marta Glavatskikh, Thomas Cauchy, and Benoit Da Mota. Scalable estimator of the diversity for de novo molecular generation resulting in a more robust QM dataset (OD9) and a more efficient molecular optimization. *J. Cheminformatics*, 13(1) :76, 2021.
- [19] A Letard, T Amghar, O Camp, and N Gutowski. Bandits-Manchots Combinatoires : du retour utilisateur à la recommandation. In *CNIA 2021 : Conférence Nationale en Intelligence Artificielle*, pages pp 52–59, Bordeaux, France, June 2021.
- [20] Alexandre Letard, T Amghar, O Camp, and N Gutowski. Bandit et Semi-Bandit avec Retour Partiel : Une Stratégie d’Optimisation du Retour Utilisateur. In *5ème Conférence Nationale sur les Applications Pratiques de l’Intelligence Artificielle (APIA)*, Angers, France, July 2020.
- [21] Alexandre Letard, Tassadit Amghar, Olivier Camp, and Nicolas Gutowski. Partial Bandit and Semi-Bandit : Making the Most Out of Scarce Users’ Feedback. In *32th International Conference on Tools with Artificial Intelligence (ICTAI)*, 2020.
- [22] Pierre Maillot, David Genest, Stéphane Loiseau, and Thomas Raimbault. Diagnostic pour le maintien de la qualité des bases du Web des données. In *Reconnaissance des Formes et Intelligence Artificielle 2016*, Clermont-Ferrand, France, 2016.
- [23] Adrian Robert, David Genest, and Stéphane Loiseau. A Query Language for Cognitive Maps. In *18th International Conference Artificial Intelligence : Methodology, Systems, and Applications (AIMSA 2018)*, volume 11089 of *Lecture Notes in Computer Science*, pages 218–227, Varna, Bulgaria, 2018.
- [24] Adrian Robert, David Genest, and Stéphane Loiseau. The Taxonomic Cognitive Map Query Language : A General Approach to Analyse Cognitive Maps. In *2018 IEEE 30th International Conference on Tools with Artificial Intelligence (ICTAI)*, pages 999–1006, Volos, Greece, November 2018. IEEE.
- [25] Adrian Robert, David Genest, and Stéphane Loiseau. Cartes cognitives temporelles. In *Extraction et Gestion des Connaissances (EGC 2020)*, volume E-36 of *Revue des Nouvelles Technologies de l’Information*, pages 309–316, Bruxelles, Belgium, 2020.
- [26] Adrian Robert, David Genest, and Stéphane Loiseau. Temporal Cognitive

- Maps. In *12th International Conference on Agents and Artificial Intelligence*, pages 58–68, Valletta, Malta, February 2020. SCITEPRESS - Science and Technology Publications.
- [27] Adrian Robert, David Genest, Stephane Loiseau, Thomas Raimbault, and Brice Trouillet. Les cartes cognitives temporelles : modélisation et interrogation. In *Extraction et Gestion des Connaissances (EGC 2019)*, volume E-35 of *Revue des Nouvelles Technologies de l'Information*, pages 369–370, Metz, France, 2019.
- [28] Hugo Traverson, David Genest, and Stephane Loiseau. Improving Feedbacks for ITS Assessment of Concept Maps. In *30th International Florida Artificial Intelligence Research Society (FLAIRS 2017)*, pages 146–151, Marco Island, United States, 2017. AAAI Press.