

Offre de thèse	
Titre	Recommandation guidée par des connaissances métier en business intelligence
Entreprise	INTESCIA group 8 rue Rouget de Lisle 92130 Issy-les-Moulineaux
Laboratoire	LAB-STICC, IMT Atlantique campus de Brest Equipe MOTEL Technopôle Brest-Iroise CS 83818 29238 Brest cedex 03
Direction / encadrement	Grégory SMITS (direction) Fahima DJELIL, Jean-Marie GILLIOT, Laurent BRISSON (encadrement)

Contexte

La société INTESCIA¹ est un éditeur de logiciels spécialisé dans la Business Intelligence (BI) et la science des données. Créée en avril 2013, INTESCIA GROUP a pour ambition de devenir un acteur majeur de la Data Economy. Son cœur de métier est l'exploitation de l'information afin de lui donner de la valeur et sa mission dans l'environnement Smart Data : Offrir aux entreprises des outils et des services business performants et en avance, adaptés à leurs exigences.

Le Lab-STICC est une Unité Mixte de Recherche rattachée à 6 tutelles, à savoir le CNRS, l'ENIB, l'ENSTA Bretagne, IMT Atlantique, l'Université de Bretagne Occidentale, l'Université de Bretagne Sud. Ce laboratoire de Recherche pluridisciplinaire est implanté sur plusieurs sites géographiques, en l'occurrence sur Brest (campus d'IMT Atlantique, de l'ENSTA Bretagne, de l'ENIB et de l'UBO), sur Quimper (IUT de quimper / UBO) et sur Lorient/Vannes (Campus de l'UBS).

Au sein du pôle SHM (Sciences Humaines et Sociales), l'équipe MOTEL vise à proposer des outils, des méthodes et des modèles centrés humains pour faciliter les interactions entre un utilisateur et un système informatique, d'apprentissage en ligne généralement.

Le projet de thèse proposé s'inscrit dans le cadre d'un partenariat entre INTESCIA et l'équipe MOTEL dont l'objectif est de concevoir de nouvelles fonctionnalités analytiques et prédictives à l'intersection entre IA et BI.

Problématique

DoubleTrade est une solution logicielle de veille business permettant aux utilisateurs (sociétés clientes) d'explorer les offres de marchés publics et privés, en France et à l'étranger.

Un enjeu majeur pour l'entreprise INTESCIA est d'augmenter la valeur ajoutée de sa solution *DoubleTrade* avec des fonctionnalités analytiques et prédictives qui soient novatrices et pragmatiques. L'objectif est à la fois de fournir des vues concises et interprétables des marchés en cours et de leur évolution, mais surtout d'être en mesure de recommander à un client les marchés les plus adaptés à ses activités.

Pour atteindre cet objectif fonctionnel, quatre verrous scientifiques seront adressés lors de cette thèse :

¹ <https://corporama.fr/blog/interview-alexandre-sidommo/>

1. Définir des algorithmes de résumé des appels d'offre [Smits et al., 2018] et des mesures de comparaison entre résumés pour mettre en exergue les évolutions des caractéristiques de ces appels d'offre [Dong and Li, 1999].
2. Enrichir un modèle métier (type graphe de connaissances [Ji et al., 2021]) défini a priori à l'aide de connaissances acquises automatiquement des données reflétant les dépendances entre offres de marchés, les liens entre les différents acteurs (entreprises, collectivités, décisionnaires), etc.
3. Concevoir une stratégie de recommandation d'appels d'offres guidée par le modèle métier [Guo et al., 2020] ainsi que des techniques d'explication des raisons de la recommandation. Cette dernière fonctionnalité s'appuiera sur les travaux existants en intelligence artificielle explicable et notamment la génération d'explications contrastives entre la classe des appels d'offres adaptés à un profil d'entreprise et ceux qui ne le sont pas [Došilović et al, 2018].
4. Proposer et développer des tableaux de bord combinant des vues « classiques » de métriques, des visualisations des résumés de données ainsi que des recommandations d'appels d'offre et leurs explications associées.

Déroulement de la thèse

La thèse pourra se dérouler soit dans les locaux de l'entreprise INTESCIA (Issy-les-Moulineaux) soit au Lab-STICC (IMT Atlantique campus de Brest) avec des possibilités de télétravail partiel. L'ensemble des personnes impliquées dans le projet (INTESCIA + Lab-STICC) se réunira mensuellement pour s'assurer de la bonne conduite du projet.

Formation et compétences requises

Le candidat doit posséder un diplôme de Master recherche ou un diplôme d'ingénieur en informatique. Outre le fait de disposer d'une capacité à travailler en équipe et d'un réel goût pour la recherche scientifique, on attend du candidat une formation et une expérience en sciences de données et en apprentissage automatique. Des compétences en business intelligence seraient un plus très apprécié.

Candidature

Les candidatures (CV et lettre de motivation) doivent être envoyées à Grégory SMITS (gregory.smits@imt-atlantique.fr). Offre valable jusqu'à ce qu'un.e candidat.e soit retenu.e.

Références

- [Dong and Li, 1999] Dong, G., & Li, J. (1999, August). Efficient mining of emerging patterns: Discovering trends and differences. In *Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 43-52).
- [Došilović et al, 2018] F. K. Došilović, M. Brčić and N. Hlupić, "Explainable artificial intelligence: A survey," 2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), 2018, pp. 0210-0215, doi: 10.23919/MIPRO.2018.8400040.
- [Guo et al., 2020] Guo, Q., Zhuang, F., Qin, C., Zhu, H., Xie, X., Xiong, H., & He, Q. (2020). A survey on knowledge graph-based recommender systems. *IEEE Transactions on Knowledge and Data Engineering*.
- [Ji et al, 2021] Ji, S., Pan, S., Cambria, E., Marttinen, P., & Philip, S. Y. (2021). A survey on knowledge graphs: Representation, acquisition, and applications. *IEEE Transactions on Neural Networks and Learning Systems*, 33(2), 494-514.
- [Smits et al., 2018] G. Smits, P. Nerzic, O. Pivert and M.J. Lesot. Efficient Generation of Reliable Estimated Linguistic Summaries. In Proc. of the 27th IEEE International Conference on Fuzzy Systems (Fuzz-IEEE'18), Rio de Janeiro, Brazil, 2018.

PhD offer	
Title	Corporate knowledge-guided recommendation in a BI context
Company	INTESCIA group 8 rue Rouget de Lisle 92130 Issy-les-Moulineaux
Laboratory	LAB-STICC, IMT Atlantique campus de Brest Equipe MOTEL Technopôle Brest-Iroise CS 83818 29238 Brest cedex 03
Supervisors	Grégory SMITS (direction) Fahima DJELIL, Jean-Marie GILLIOT, Laurent BRISSON (encadrement)

Context

INTESCIA² is a French software editor that provides Business Intelligence (BI) solutions that embed predictive solutions for business opportunities based on Artificial Intelligence to validate, link and interpret the millions of data (sometimes source of anxiety) in order to allow company management to see more clearly and make the right decisions. Founded in April 2023, INTESCIA GROUP aims at being a major actor of the Data Economy. Its main activity is the management and the enrichment of the data so as to provide companies with efficient and cutting-edge business tools and services that are in adequacy with their expectations.

The Lab-STICC, with its double affiliation to the INS2I and INSIS institutes of the CNRS, is a research unit historically recognized in Brittany and in France in the field of ICTS. It has a proven capacity to cover a broad scientific spectrum around digital sciences, and in particular with this ability to address various disciplinary fields (Information Theory, Waves & Materials, Embedded Electronics and Computing, Data Sciences, Communication and Signal Detection, Human-Machine Interfaces,...) following multiple themes/application sectors: maritime environment, communicating objects, defense, space, health, security, robotics...

The MOTEL « MODels and Tools for Enhanced Learning » team takes part of the research community EIAH (in French, Environnements Informatiques pour l'Apprentissage Humain") - Computing Environments for Human Learning. MOTEL works on several subjects that contribute with computing human-centred tools, methods and models for Education using an experimental approach.

This PhD offer takes part in a collaborative project between INTESCIA and the MOTEL research team. The goal of this collaboration is to conceive novel analytic and predictive functionalities at the intersection between AI and BI.

Scientific issues

DoubleTrade is an ultimate software solution to the monitoring of public sector tender that helps their French and foreign customers find the best professional opportunities.

A crucial issue for INTESCIA is to integrate cutting-edge analytic and predictive functionalities in the *DoubleTrade* solution that are both new and pragmatic. A possible

² <https://corporama.fr/blog/interview-alexandre-sidommo/>

direction for the new-generation of BI solutions is to provide users with concise and interpretable views of the current offers and their evolution. But most of all, the added value of such solutions lies in their capacity of identifying and recommending the most appropriate offers for a given company, and in fine to explain the reasons of each recommendation.

To reach this goal, four main scientific issues have to be addressed during the proposed PhD:

1. To define efficient algorithms to summarize the available offers [Smits et al., 2018] as well as comparison measures between such summaries. Comparisons should emphasize on the emerging trends [Dong and Li, 1999].
2. To enrich an *a priori*-defined corporate model (as a knowledge graph for instance) [Ji et al., 2021]) with knowledge automatically extracted from the data that model links between invitation to tender, actors (company, public services, decision makers), etc.
3. To conceive a recommendation system of invitations to tender guided by the corporate model [Guo et al., 2020] completed with explanation strategies that point out the reason of the recommendation. This last functionality will rely on recent work about explainable AI et especially the generation of appropriate contrastive explanations between relevant and irrelevant offers [Došilović et al, 2018].
4. To propose and develop novel dashboards that combine ``classical'' views about static metrics, data summaries and recommendations of relevant offers associated with explanations about their provenance.

PhD organization

The PhD student will work in the INTESCIA building (Issy-les-Moulineaux) or at the Lab-STICC (IMT Atlantique - Brest) with possibilities to work remote. Monthly meetings will regroup all the members (INTESCIA + Lab STICC) involved in the project.

Required training and skills

The PhD candidate should possess a Master degree or be an engineer in computer science. In addition to be able to work in team and to be strongly interested in scientific research, we expect from the candidate to been trained on data science and machine learning. Skills in BI are very welcome.

How to apply?

A CV and a motivation letter have to be sent to Grégory SMITS (gregory.smits@imt-atlantique.fr). The offer is available until an adequate candidate is found.

References

- [Dong and Li, 1999] Dong, G., & Li, J. (1999, August). Efficient mining of emerging patterns: Discovering trends and differences. In Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 43-52).
- [Došilović et al, 2018] F. K. Došilović, M. Brčić and N. Hlupić, "Explainable artificial intelligence: A survey," 2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), 2018, pp. 0210-0215, doi: 10.23919/MIPRO.2018.8400040.
- [Guo et al., 2020] Guo, Q., Zhuang, F., Qin, C., Zhu, H., Xie, X., Xiong, H., & He, Q. (2020). A survey on knowledge graph-based recommender systems. IEEE Transactions on Knowledge and Data Engineering.
- [Ji et al, 2021] Ji, S., Pan, S., Cambria, E., Martinen, P., & Philip, S. Y. (2021). A survey on knowledge graphs: Representation, acquisition, and applications. IEEE Transactions on Neural Networks and Learning Systems, 33(2), 494-514.
- [Smits et al., 2018] G. Smits, P. Nerzic, O. Pivert and M.J. Lesot. Efficient Generation of Reliable Estimated Linguistic Summaries. In Proc. of the 27th IEEE International Conference on Fuzzy Systems (Fuzz-IEEE'18), Rio de Janeiro, Brazil, 2018.