

Post-doctoral position

Representing and enriching events in knowledge graphs

Key words: machine learning on graphs, knowledge graph, event prediction

Context: XP-event project (2021-2024)

An event is defined as “anything that happens, anything that fits over time”: meetings, phone calls, purchases, but also business buyouts, change of management, health crises, etc. The events are shared at through various communication channels that can be private (internal documentation, emails, Slack, Teams, phone, etc.) or public (press, Twitter, Facebook, etc.). Knowledge of these events is essential for humans to make decisions which themselves will have an impact on future events. Many innovative applications can benefit or even emerge from a technology capable of extracting events from various sources, representing them, aggregating them and exploiting them to predict future events. We can for example cite: anticipating demand for sanitary products, the supervision of cultural, advertising or festive events, but also the study of competition, the study of commercial markets, etc. .

One of the main obstacles to the deployment of these applications is the excessively high cost of their development when it is carried out on an ad hoc basis by competing players. The XP-Event project proposes to respond to this difficulty by setting up a common base for all the applications organized around the notion of event. This project is led by a consortium naturally formed by two companies (GeoTrend and Emvista) and a research team from the IRIT laboratory sharing this vision and each having significant scientific and technological heritage in the field.

Position Description

The candidate will contribute to the tasks in which IRIT is involved, and will be more particularly in charge of realizing and implementing the proposed solutions. The first task concerns the representation of event graphs. The first task will be to define an adapted ontology and a process allowing to exploit it to access or represent in RDF the graphs of the industrial partners of the project, which are graphs of quite different nature. The second task aims at defining a process to evaluate the quality of the event graphs. Evaluation will be based on the ontology structure as well as on reasoning from the knowledge graph. The third task concerns the enrichment of these graphs. Two types of approaches will be implemented in the project, and for each of them research will be needed to advance the state of the art. The first approach consists in extracting information from texts. Each of the industrial partners already has its own processing chain that it will improve and unify. The second approach consists in exploiting the current state of a graph but also the structure of an event in the ontology to suggest the addition of new nodes or new relations to the graph. This approach will be implemented through learning algorithms from graph.

Requirements for this position

Applicants are required to have a PhD in computer science, and strong background ideally in two areas of artificial intelligence: semantic web technologies (ontology engineering, linked data management and querying, SPARQL, SHACL, RuleML, ...), and machine learning from graphs and vector representations, recursive neural networks, etc. Good programming skills (Python, OWL API) and experience in participating in collaborative projects is required. In addition, the candidate must have a taste for innovation, and the ability to dialogue and collaborate with industrial partners. Experience in managing graph warehouses (Virtuoso, Strabon, Neo4j...) is desired. Fluency in written / spoken English is required too. Fluency in French language will be a plus.

Work environment

Location : Institut de Recherche en informatique de Toulouse (IRIT) - UPS, 118 Route de Narbonne F-31062 Toulouse Cedex - France and et UT2J, 5 allées Antonio Machado 31300 Toulouse

Duration: 24 months – starting on January 1st, 2022

Host team: MELODI <https://www.irit.fr/en/departement/dep-artificial-intelligence/melodi-team/>

The candidate will work with four academic researchers from MELODI (F. Benamara, Ph. Muller, N. Aussenac-Gilles and N. Hernandez). He will collaborate with the partner companies in the project, namely Geotrend, located in Toulouse, and Emvista, located in Montpellier.

Income: between 2300 and 2800 euros before taxes (brut) monthly according to past experience

How to apply?

Applicants should upload their application files before November 1st 2021 on the following web page : . Application files should contain at least a full Curriculum Vitae including a complete list of publications, a cover letter indicating their research interests, achievements to date and vision for the future, as well as either support letters or the name of 2 persons that have worked with them.

Candidates should contact: N. Aussenac-Gilles <aussenac@irit.fr> and N. Hernandez <hernande@irit.fr>

Poste de chercheur en CDD (24 mois)

Représentation et enrichissement de graphes d'événements

Mots clés : Apprentissage à partir de graphes, Graphes de connaissances, prédiction d'événements

Contexte : [Projet XP-Event \(2021-2024\)](#)

Un événement est défini comme "tout ce qui se produit, tout fait qui s'insère dans la durée" : réunions, appels téléphoniques, achats, mais aussi rachat d'entreprises, changement de direction, crises sanitaires, ... Les événements sont partagés à travers divers canaux de communication privés (documentation interne, e-mails, Slack, Teams, téléphone, etc.) ou publics (presse, Twitter, Facebook, etc.). La connaissance de ces événements est indispensable à l'humain pour prendre des décisions qui auront un impact sur les événements à venir. De nombreuses applications innovantes peuvent bénéficier ou même émerger d'une technologie capable d'extraire des événements de sources variées, de les représenter, les agréger et les exploiter pour prédire les événements à venir. Nous pouvons par exemple citer : l'anticipation de la demande de produits sanitaires, l'encadrement de manifestations culturelles, publicitaires ou festives, mais aussi l'étude de la concurrence, l'étude de marchés commerciaux, etc. .

Un des principaux freins au déploiement de ces applications est le coût trop élevé de leur développement lorsqu'il est mené de façon ad hoc par des acteurs en concurrence. Le projet XP-Event propose de répondre à cette difficulté par la mise en place d'un socle commun à toutes les applications organisées autour de la notion d'événement. Ce projet est porté par un consortium naturellement formé de deux entreprises (GeoTrend et Emvista) et d'une équipe de recherche du laboratoire IRIT partageant cette vision et disposant chacun d'un patrimoine scientifique et technologique important dans le domaine.

Description du poste

Le candidat contribuera aux tâches dans lesquelles l'IRIT est impliqué, et sera plus particulièrement chargé de réaliser et implémenter les solutions proposées. La première tâche concerne la *représentation de graphes d'événement*. Il s'agira d'abord de définir une ontologie adaptée puis un processus permettant de l'exploiter pour accéder à ou représenter en RDF les graphes des partenaires industriels du projet, qui sont des graphes de nature assez différente. La deuxième tâche consiste à définir un processus pour *évaluer la qualité des graphes d'événement*. On s'appuiera sur la structure de l'ontologie ainsi que sur des raisonnements à partir du graphe de connaissances. La troisième tâche porte sur *l'enrichissement de ces graphes*. Deux types d'approches seront mises en œuvre dans le projet, et pour chacune d'elle des recherches seront nécessaires pour avancer l'état de l'art. La première approche consiste à extraire de l'information à partir de textes. Chacun des partenaires industriels a déjà sa propre chaîne de traitement qu'il va améliorer et unifier. La deuxième approche consiste à exploiter l'état courant d'un graphe mais aussi la structure d'un événement dans l'ontologie pour suggérer l'ajout de nouveaux nœuds ou de nouvelles relations dans le graphe. Cette démarche sera mise en œuvre grâce à des algorithmes d'apprentissage à partir du graphe.

Exigences pour ce poste

Les candidats doivent être titulaires d'un doctorat en informatique, avoir une solide expérience idéalement dans deux domaines de l'intelligence artificielle : les technologies du web sémantique, (ingénierie des ontologies, gestion et interrogation des données liées, SPARQL, SHACL, RuleML, ...), et apprentissage automatique à partir de graphes et de représentations vectorielles, réseaux de neurones récurrents, etc. Une bonne maîtrise en programmation (Python, API OWL) et une expérience en participation à des projets collaboratifs est requise. De plus, le candidat devra avoir un goût pour l'innovation, des capacités de dialogue et collaboration avec les partenaires industriels. Une expérience dans la gestion d'entrepôts de graphes (Virtuoso, Strabon, Neo4j...) est souhaitée.

Environnement de travail

Localisation : Institut de Recherche en informatique de Toulouse (IRIT) - UPS, 118 Route de Narbonne F-31062 Toulouse Cedex et UT2J, 5 allées Antonio Machado 31300 Toulouse

Durée : 24 mois - débutant le 1er janvier 2022

Equipe d'accueil : MELODI <https://www.irit.fr/en/departement/dep-artificial-intelligence/melodi-team/>

Le.la candidat.e travaillera avec les permanents de MELODI impliqués dans le projet (F. Benamara, Ph. Muller, N. Aussenac-Gilles et N. Hernandez). Il collaborera avec les entreprises partenaires Geotrend, localisée à Toulouse, et Emvista, située à Montpellier.

Salaire: entre 2300 et 2780 euros brut mensuels selon expérience

Comment candidater ?

Les candidat.e.s doivent transmettre leur dossier avant le 1 novembre 2021 par email aux adresses ci-après. Le dossier doit comprendre au moins les documents suivants : CV complet avec une liste de publications, lettre de motivation précisant les intérêts, résultats et perspectives de vos recherches, ainsi que des lettres de soutien ou le nom de trois référents parmi vos collaborateurs. Contact pour toute information supplémentaire : N. Aussenac-Gilles <aussenac@irit.fr>, N. Hernandez <hernande@irit.fr>.