

## Proposition de sujet de Master de Recherche

### Titre : Recherche d'Information Agrégative sur les Workflows

**Contexte:** Projet ANR CAIR 2014-2017

La recherche d'information agrégée [1] vise à évaluer des requêtes utilisateurs en assemblant des fragments d'information provenant de plusieurs sources. Ce nouveau type de requêtes agrégatives permet de créer des objets qui n'existent pas en tant que tels dans les sources interrogées, mais sont construits à partir de fragments issus de ces sources. La recherche d'information agrégée a le potentiel de **révolutionner** l'accès à l'information distribuée en permettant de produire, de façon dynamique, des **associations entre fragments d'information publiés séparément** et qui se rapportent à la même entité/événement/sujet. Les résultats ainsi produits sont des objets à **forte valeur ajoutée** qui pourront être utilisés dans l'aide à la décision, le data journalisme ou l'acquisition de connaissances sur des phénomènes ou processus particuliers. Dans le cadre de projet, nous nous intéressons en particulier aux requêtes visant de récupérer (et agréger) des informations sur des processus (que ce métier [2] ou scientifique [3]).

### Travail demandé

Etant donnée une source contenant des processus nous souhaitons extraire des informations complémentaires à partir d'autres sources de données sur les données utilisées ou produites par les processus, les acteurs humains qui ont des connaissance sur le processus, ainsi que de la provenance décrivant l'origine de ces informations complémentaires. Ces sources complémentaires fournissent des informations de différents types, par exemple, google scholar, facebook in peuvent être utilisé pour extraire des informations sur les acteurs (auteurs) d'un processus, linkedIn peut fournir des informations sur les relations entre acteurs ainsi que les compagnies.

**Laboratoire d'accueil :** Lamsade

**Encadrants :** Daniela Grigori ([daniela.grigori@dauphine.fr](mailto:daniela.grigori@dauphine.fr)), Khalid Belhajjame ([khalid.belhajjame@dauphine.fr](mailto:khalid.belhajjame@dauphine.fr))

*Ce sujet est susceptible d'être poursuivi en these.*

### Références

- [1] Lalmas, M. (2011). Aggregated search. In *Advanced Topics in Information Retrieval* (pp. 109-123). Springer Berlin Heidelberg.
- [2] Van Der Aalst, W. M., Ter Hofstede, A. H., & Weske, M. (2003). Business process management: A survey. In *Business process management* (pp. 1-12). Springer Berlin Heidelberg.
- [3] Ewa Deelman, Dennis Gannon, Matthew S. Shields, Ian Taylor: Workflows and e-Science: An overview of workflow system features and capabilities. *Future Generation Comp. Syst.* 25(5): 528-540 (2009)