



CONVENTION CIFRE¹ N° 2012/0305

ENTREPRISE : SAP Labs France SAS

DOCTORANT : Ahmad ASSAF

RAPPORT D'ACTIVITES A L'ISSUE DE LA 1^{ère} ANNEE DE THESE²

**Si ce n'est déjà fait, pensez à joindre
la dernière attestation d'inscription en doctorat (2^{ème} ou 3^{ème} année)**

Merci de présenter succinctement les travaux réalisés au cours de l'année écoulée. Préciser l'avancement des recherches par rapport au planning initial, l'évolution par rapport au projet initial et les éventuels problèmes rencontrés (intégration à l'équipe de recherche, adaptation aux contraintes du milieu industriel, ...).

Durant cette première année, j'ai tout d'abord effectué une revue de la littérature couvrant différents domaines allant de la Business Intelligence (BI), de l'analyse des données à l'intégration de données dans le Web sémantique. Cela a contribué à enrichir mes connaissances globales dans ces domaines et m'a permis de mieux définir le champ de mes recherches.

J'ai participé à un projet interne à SAP appelé Remix. Remix est une solution de BI qui permet aux utilisateurs d'entreprise de créer rapidement des rapports d'analyse. Il permet notamment la composition d'objets BI existants avec de nouvelles données de l'entreprise et à partir de sources de données externes. Il s'agit d'un outil de BI en libre-service destiné aux utilisateurs professionnels.

Le focus des recherches a évolué depuis l'idée initiale, qui était l'exploitation des interactions des utilisateurs de BI, vers mon centre d'intérêt actuel qui est l'enrichissement de données de l'entreprise avec de la sémantique qui permet une meilleure intégration des données, entraînant ainsi des décisions d'affaires plus éclairées. J'ai proposé dans ce cadre plusieurs idées innovantes que j'ai prototypées afin de valider leur faisabilité.

Merci de préciser la répartition du temps passé en entreprise et au laboratoire.

Cette répartition vous convient-elle ? Va-t-elle évoluer ?

Rappeler les interruptions longues de travail (ex : congé maternité, accident, ...)

J'ai passé la plupart de mon temps dans SAP, tout en allant travailler à EURECOM quand le besoin se faisait ressentir (en moyenne 1 jour par semaine). Je pense que cette répartition restera similaire pour l'année à venir. Cependant, quand je vais commencer à me concentrer sur l'écriture de ma thèse au cours de la dernière année, je vais devoir interagir plus directement avec mon directeur de thèse à EURECOM.

Avez-vous bénéficié de formation(s) au cours de l'année écoulée ? Merci d'en préciser le thème et l'organisateur.

J'ai reçu une formation dans les technologies SAP HANA et les méthodes de développement Agile dans SAP.

¹ Rappeler le numéro de convention (année de la demande / numéro de dossier)

² Rayer la mention inutile. Attention, dans certains cas particuliers, les conditions générales d'octroi stipulent qu'un rapport doit être fourni tous les 6 mois.

Vos travaux ont-ils donné lieu à publication, présentation à un congrès, poster ou encore à dépôt de brevet au cours de l'année écoulée ? Merci de préciser notamment le caractère international des communications.

- Ahmad Assaf, Aline Senart, Raphaël Troncy: An Objective Assessment Framework for Linked Data Quality. Special Issue on Web Data Quality of the International Journal on Semantic Web and Information Systems (IJSWIS) (submitted)
- Ahmad Assaf, Aline Senart, Raphaël Troncy: SNARC - An Approach for Aggregating and Recommending Contextualized Social Content. Extended Semantic Web Conference (ESWC), pages 319-326, May 2013
- Ahmad Assaf, Aline Senart, Raphaël Troncy : SNARC - A Semantic Social News Aggregator, AI Mashup Challenge, Extended Semantic Web Conference (ESWC), Montpellier, France, May 2013.
- Ahmad Assaf, Aline Senart: Data Quality Principles in the Semantic Web. IEEE International Conference on Semantic Computing (ICSC), pages 226-229, Palermo, Italy, September 2012.
- Ahmad Assaf, Eldad Louw, Aline Senart, Corentin Follenfant, Raphaël Troncy, David Trastour: RUBIX: A framework for improving data integration with linked data. First International Workshop on Open Data (WOD), pages 13-21, <http://dx.doi.org/10.1145/2422604.2422607>
- Ahmad Assaf, Eldad Louw, Aline Senart, Corentin Follenfant, Raphaël Troncy, David Trastour: Improving Schema Matching with Linked Data. First International Workshop on Open Data (WOD), Nantes, France, May 2012

Quelles sont vos perspectives pour l'année à venir ?

J'ai l'intention d'affiner davantage le sujet de thèse en mettant l'accent sur l'intégration de données (entreprises et données publiques) et leurs visualisations (e.g. tableau de bord), de poursuivre les développements engagés dans le cadre de ma thèse, de publier des résultats issus des expérimentations et d'intégrer mon travail de thèse au sein de SAP.

Dans l'état actuel d'avancement de vos travaux, envisagez-vous de soutenir votre thèse dans les délais impartis ?

Oui, je m'attends à finir dans les temps.

Commentaires et appréciations du responsable scientifique dans l'entreprise

Dr. Aline Senart

Ahmad a fait une très bonne année. Il a acquis une connaissance approfondie du domaine, a su identifier les besoins et problèmes spécifiques et proposer des solutions innovantes. Il a déjà publié un nombre important de résultats intermédiaires dans des workshops et conférences renommées. Il lui faudra l'année prochaine réussir à regrouper toutes ses idées sous un canevas commun et poursuivre ses travaux afin de les formaliser et les valider par des expérimentations poussées.

Commentaires et appréciations du directeur de thèse

Dr. Raphaël Troncy

Ahmad a fait une excellente première année de thèse. Ses travaux ont déjà donné lieu à plusieurs publications (et d'autres sont en préparation) et à un prix (pour le meilleur mashup sémantique) confirmant son aptitude à développer des applications innovantes. La thèse d'Ahmad couvrera sans aucun doute plusieurs domaines (intégration de données, business intelligence, fouille donnée, visualisation de données) et Ahmad a fait un travail important de bibliographie durant cette première année. L'objectif de cette deuxième année est clairement d'approfondir certaines pistes de recherche qui ont déjà été esquissés et de continuer à raffiner ce qui sera la contribution originale de cette thèse.

Nom du doctorant Ahmad Assaf	Nom et fonction du tuteur industriel Dr Aline Senart	Nom et titre du tuteur de thèse Dr Raphaël Troncy
Date et signature 8/10/2013 	Date, cachet et signature ³ 8/10/2013 	Date, cachet et signature ³ 08/10/2013 

Ce document est à retourner à : ANRT – Service CIFRE, 41 bd des Capucines, 75002 Paris.

³ Conformément à l'article 7 des conditions générales d'octroi d'une CIFRE, si le rapport n'est pas signé par les 2 tuteurs, il ne sera pas valide. En conséquence, le paiement de la subvention sera interrompu.

