



Université Claude Bernard



Lyon 1

DIPLÔME NATIONAL DE DOCTORAT

(Arrêté du 25 mai 2016)

Date de la soutenance : **30 août 2018**

Nom de famille et prénom de l'auteur : **OMRI Asma**

Titre de la thèse : « Modélisation et utilisation de ressources et services Web et indexation de données dans un contexte d'incertitude ».



Résumé

Au cours des dernières années, il y a eu une énorme explosion dans la production de données en raison de la multiplication des nouvelles technologies telles que les réseaux sociaux, les nouveaux et les plus puissants appareils mobiles, les compteurs intelligents, les capteurs et le cloud computing. Cette explosion des données devrait se poursuivre et même accélérer. Comme les données sont de plus en plus varié, plus complexe et moins structuré, il est devenu impératif de les traiter rapidement. DaaS peut être définie comme l'approvisionnement, la gestion et la fourniture de données fournies dans un format immédiatement consommable aux utilisateurs professionnels des organisations en tant que service. Les données retournées par ces services sont incertaines, hétérogènes, grand volume et grande vitesse d'évolution.

Plusieurs approches ont traité les différentes phases de cycle de vie de service web tel que la création, la sélection, la découverte, la modélisation, l'invocation et la composition des services soit pour résoudre le problème de volume de données, d'hétérogénéité de données ou la vitesse d'évolution de données. Mais la qualité de données n'a pas reçu une grande attention et très peu d'approches qui ont traitée l'incertitude de données dans le web.

Cette thèse s'intéresse à la problématique des services Web dans un contexte de systèmes distribués et hétérogènes. La principale contribution à apporter dans le cadre de cette thèse est d'étudier la composition de services et/ou de ressources Web et indexation de données dans un contexte de présence d'informations incertaines.

Dans un premier temps, cette thèse s'intéresse à la définition de concept de service DaaS incertain en adoptant la théorie possibiliste. Nous nous intéressons au problème de la composition des services Web de données et étudions l'impact de l'incertitude qui peut être associée à la sortie d'un service sur les processus de sélection et de composition des services. Nous avons proposé une approche possibiliste pour modéliser l'incertitude des données renvoyées par des services de données incertains. Plus précisément, nous avons étendu les normes de description de service Web (par exemple, WSDL) pour représenter les degrés d'incertitude des sorties. Nous avons également étendu le processus d'invocation de service pour prendre en compte l'incertitude des données d'entrée. Cette extension est basée sur la théorie des mondes possibles utilisée dans les bases de données possibilistes. Nous avons également

proposé un ensemble d'opérateurs de composition sensibles aux valeurs d'incertitude, nécessaires pour orchestrer des services de données incertains.

Dans un deuxième temps, nous avons étudié l'impact de l'incertitude sur la représentation et la manipulation des ressources Web. Nous avons défini le concept de ressource Web incertaine et nous avons proposé des mécanismes de composition de ressources. Nous avons proposé un modèle de description de l'incertitude à travers le concept de ressource Web incertaine. Celui-ci est basé sur un modèle probabiliste où chaque ressource peut avoir plusieurs représentations possibles, avec une certaine probabilité. En nous basant sur ce modèle, nous proposons une approche permettant d'évaluer des requêtes dans un contexte incertain.

Enfin, nous avons proposé des méthodes d'indexation documentaire de données de Big Data. Tout d'abord, nous avons proposé une approche d'indexation syntaxique de données incertaines, ensuite, nous avons proposé une méthode d'indexation sémantique incertaine et nous avons finalisé notre thèse par la proposition d'une méthode hybride d'indexation dans un contexte incertain.