

Analyse Ontologique de scénario dans un contexte Big Data

Marwan Batrouni*, Aurélie Bertaux*
Christophe Nicolle*

*Laboratoire électronique, Informatique et image (Le2i) - FRE CNRS 2005
Bât. I3M - Rue Sully - 21000 Dijon - France
prenom.nom@u-bourgogne.fr,
<http://checksem.fr>

Notre analyse du domaine prédictif a établi deux constats : il n'existe pas de définition universelle de scénario et l'analyse de scénarios ne prend pas en compte suffisamment de contraintes pour établir des prédictions de qualité prouvée. Nous proposons donc une définition universelle des scénarios pour répondre au premier point. Pour répondre au second nous nous sommes intéressés au fait que les prédictions sont plus précises si le volume et l'origine des données étaient conséquent et hétérogène. Ce constat nous a porté dans le domaine du Big Data auxquels doivent être ajoutés la sémantique, la Valeur et la Véracité des données pour une prédiction fiable et sensée. Nous montrons que les ontologies permettent de répondre à l'ensemble de ces critères. Par ailleurs, nous avons identifié que les machines d'états répondent aux besoins (en terme d'outil et de formalisation) de l'analyse de scénario. Nous proposons donc à partir de la définition uniformisée des scénarios, de les analyser par le biais des machines d'états basées sur des ontologies. Cet article ébauche les prérequis de formalisation par la combinaison de 3 domaines qui apportent chacun les réponses aux différents éléments de la définition : l'analyse de scénario, les ontologies et le Big Data.

1 Analyse de scénarios dans un environnement Big Data

Dans le contexte du Big Data, l'analyse de scénarios doit exploiter ses outils pour améliorer son efficacité. Pour cela deux technologies (parmi les 3 ci-dessous) dans la sphère Big Data sont particulièrement intéressantes, à savoir l'analyse prédictive et l'analyse prescriptive.

- **Analyse descriptive : ce qui s'est passé.** ETL, MapReduce, les lacs de données et les algorithmes de clustering...
- **Analyse prédictive : ce qui se passerait.** L'exploration de texte notamment via CRISP-DM.
- **Analyse prescriptive : ce qui devrait arriver.** L'apprentissage machine, la recherche opérationnelle, l'analyse statistique, et la modélisation et la simulation.

Le Big Data fournit des outils capables d'améliorer l'efficacité de l'analyse de scénarios. Cependant les prédictions devant être crédibles et pertinentes il faut prendre en compte la Valeur et la Véracité des données. Ces deux aspects sont traités par l'utilisation d'ontologies.