

Vers la Reconstruction des Connaissances Agricoles: Perspectives de Détection des Risques Naturels à partir de Sources de Données Hétérogènes

Shufan Jiang^{*,**}, Rafael Angarita^{*},
Raja Chiky^{*}, Stéphane Cormier^{**}, Francis Rousseaux^{**}

^{*}Institut Supérieur d'Electronique de Paris, LISITE, Paris, France
name.lastname@isep.fr,

^{**} Université de Reims Champagne Ardenne, CReSTIC, Reims, France
name.lastname@univ-reims.fr

Le progrès des technologies de l'information et de la communication (TIC) visent à relever certains des défis les plus importants de l'agriculture Cox (2002). L'utilisation de ces TIC pour améliorer les processus agricoles est connue sous le nom d'*agriculture intelligente* Wolfert et al. (2017). Des chercheurs ont appliqué un large éventail de technologies pour s'attaquer à certains objectifs spécifiques liés à l'agriculture : des modèles de simulation Hammer et al. (2001), la vision par ordinateur et l'intelligence artificielle Patricio et Rieder (2018), des drones Tripicchio et al. (2015), et le paradigme de l'IoT Patil et al. (2012). Les **appareils IoT** collectent des informations et produisent des *données structurées*; cependant, des informations importantes liées à l'agriculture peuvent également provenir de différentes sources telles que des rapports périodiques officiels comme les *Bulletin de Santé du Végétal*(BSV)¹ et les réseaux sociaux tels que Twitter. Le BSV et les autres rapports formels sont des *données semi-structurées*. Les publications sur les réseaux sociaux sont des *données non-structurées*.

À partir de ce scénario, la question à laquelle nous voulons répondre est la suivante Jiang et al. (2020) : comment tirer parti de ces sources de données **dynamiques** et **hétérogènes** pour détecter des risques pour l'agriculture? Pour répondre à cette question, l'objectif de cette recherche est de soutenir l'intégration dynamique de sources de données agricoles hétérogènes dans une base de connaissances complète qui peut soutenir les agriculteurs dans leurs activités, présenter des informations globales et en temps réel aux chercheurs et aux parties intéressées et soutenir les systèmes de détection des aléas. Pour cela, nous divisons ce travail en éléments suivants : (i) modéliser une ontologie des aléas naturels et leurs relations avec les cultures ; (ii) la classification automatique des sources de données suivantes selon l'ontologie des aléas naturels : le BSV, les tweets, les données IoT ; et (iii) enrichir l'ontologie des aléas naturels avec de nouvelles informations extraites des différentes sources de données.

Dans l'état de l'art sur l'intégration des données agricoles, le Web de données et les technologies du Web sémantique sont mobilisés. Les chercheurs d'Irstea ont conduit le Projet VESPA qui facilite la recherche documentaire par région, par période et par type

1. <https://agriculture.gouv.fr/bulletins-de-sante-du-vegetal>