

Construction, personnalisation et enrichissement d'ontologies pour l'analyse et l'exploitation des données des points de vente de Bimedia

Laboratoire : Laboratoire d'Informatique et d'Automatique pour les Systèmes (LIAS)

Entreprise : Bimedia

Localisation : Poitiers (LIAS)

Équipe encadrante : Mickael BARON, Brice CHARDIN et Stéphane JEAN (*L'encadrant dont le nom est souligné est l'encadrant référent, contact : brice.chardin@ensma.fr*)

Mots-clés : ontologies, modélisation, analyse prédictive, personnalisation, bases de données, données ouvertes et liées

Bimedia est une société qui propose ses services aux points de vente (bureaux de tabac, boulangeries, etc.) sous la forme d'une offre composée de matériels (caisses enregistreuses), d'un logiciel d'encaissement, de solutions dématérialisées (carte prépayées) et de dispositifs d'animation à travers des publicités fournies par des annonceurs. Dans le but d'améliorer son offre de service, Bimedia souhaite analyser un ensemble de données issues des points de vente et des fournisseurs comme, par exemple, les ventes de produits ou l'évolution des stocks. Ceci pose différents verrous scientifiques dans deux domaines principaux : la modélisation des données à analyser et l'exploitation de ces modèles.

Modélisation des données

La manière de percevoir les données à analyser et en particulier leur catégorisation ainsi que le vocabulaire utilisé pour les décrire ne sont pas explicitement modélisés. Ces éléments ne sont actuellement que partiellement représentés dans des sources de données diverses et notamment des bases de données. Il y a donc besoin de définir des modèles formels, nommés ontologies, pour expliciter cette connaissance.

Les différents points de ventes et partenaires de Bimedia ont leur propre façon de percevoir les données à analyser et de les catégoriser. Afin de laisser une grande autonomie à chaque acteur et de pouvoir restituer les analyses faites avec le vocabulaire adapté, il est nécessaire que chacun puisse avoir sa propre ontologie. Des équivalences entre les différentes déclinaisons de modélisation devront donc être définies, afin de pouvoir exploiter les données sous-jacentes selon le point de vue de l'analyste (buraliste, fournisseur, prestataire, etc.).

Exploitation des modèles

Les modèles définis devront permettre à Bimedia de gagner en expertise sur le comportement de ses points de ventes et des clients finaux. Pour cela, les modèles dont le contenu pourra être exploré et exploité (dits *boîtes blanches*) par les analystes de Bimedia seront préférés.

Un besoin important d'analyse automatique porte sur un ensemble de prédictions à réaliser, que ce soit pour affiner les commandes de renouvellement du stock ou encore pour recommander de nouveaux produits. Les modèles de recommandation devront pouvoir argumenter leurs



prédictions afin qu'elles soient admises par le point de vente, à qui revient la décision finale.

Les analyses s'appuyant sur les ontologies définies devront répondre aux besoins opérationnels de Bimedia (renouvellement des stocks, agencement du rayonnement, publicité). Deux cas d'application spécifiques ont été identifiés.

Modélisation des annonceurs et des points de vente Bimedia communique aux annonceurs publicitaire des informations liées à l'activité des points de vente, et cherche à améliorer cette offre en sélectionnant ou en élaborant des jeux de données plus complets et plus pertinents issus des points de vente et de sources externes (données ouvertes et liées).

L'objectif est de pouvoir fournir aux annonceurs des informations sur leurs produits, croisées aux informations démo-socio-géographiques des points de vente. Les besoins peuvent être variés car les axes d'analyse de la donnée restituée—enrichie et consolidée—peuvent différer selon les annonceurs.

Modélisation des produits Les produits d'un point de vente sont catégorisés au niveau de la caisse. Chaque commerçant peut ainsi organiser ses produits en groupes et sous groupes, indépendamment de l'organisation des autres points de vente et de la catégorisation interne de Bimedia. Ces catégories servent principalement au commerçant pour obtenir des résultats d'analyse et des indicateurs qui lui soient pertinents : ses performances économiques (quantités et chiffres d'affaires) sont ainsi structurées selon sa catégorisation.

L'objectif de Bimedia est de pouvoir raisonner sur les différentes déclinaisons de catégorisations et de proposer, orienter et guider les nouveaux points de vente pour aboutir à une catégorisation la plus pertinente possible.

De plus, la Loi de modernisation de la distribution de la presse—effective au 1er janvier 2023—autorise désormais les points de vente à avoir un contrôle sur les titres de presse mis en vente. Un système devra être mis en place pour intégrer cette nouvelle catégorie de produit et proposer aux points de vente une sélection de titres et les quantités associées, en s'appuyant à la fois sur les historiques de commandes et les profils de clients identifiés.

Profil du candidat

Le candidat devra être inscrit en master 2 ou équivalent et posséder des connaissances en développement logiciel, systèmes d'information, statistiques et analyse de données. Un bon niveau en français et en anglais est nécessaire.

Documents à fournir :

- Curriculum Vitae;
- lettre de motivation;
- notes de Master ou équivalent;
- tout autre document jugé nécessaire par le candidat pouvant enrichir le dossier de candidature.

