|  |
| --- |
| **Projet de Parcours SIE** |
| Extraction Automatique de Données |
|  |
| OpenStreetMap est une carte du monde modifiable. La communauté des utilisateurs intervient donc dans la création/modification des éléments cartographiés. Il est de plus, extrêmement facile d’extraire des parties de la carte et de les exporter sous différents formats, notamment en XML. |
|  |
| **BELLUOT Vincent, LORIOT Sébastien, MONTOLIU Teddy** |
| **14/06/2013** |
|  |

Table des matières

Introduction……………………………………………………………………………………………………………………………………….4

Conception………………………………………………………………………………………………………………………………………...5

Etude statistique……………………………………………………………………………………………………………………………….. ?

Travail de groupe………………………………………………………………………………………………………………………………. ?

Conclusion………………………………………………………………………………………………………………………………………… ?

**Introduction**

 Après avoir fait plusieurs projets (Génie Logiciel 1 & 2, Projet Différencié) durant notre première année d’ingénieur et avant de la terminer, nous avons un dernier projet, le projet de parcours. Cela consiste à réaliser un projet dont le thème est fortement orienté vers notre parcours, ici SIE (Système d’Information d’Entreprise). Nous allons utiliser tous les outils appris durant l’année pour concevoir ce projet.

OpenStreetMap est une carte du monde modifiable. La communauté des utilisateurs intervient donc dans la création et/ou dans la modification des éléments cartographiés. Il est de plus, extrêmement facile d’extraire des parties de la carte et de les exporter sous différents formats, notamment en XML.

 Le projet consiste à modéliser une base de données qui va servir de sauvegarde des données définies dans le fichier XML d’implémenter cette base de données, de l’alimenter et de faire des traitements statistiques sur les données ainsi remplies.

**Conception**

 Avant de faire la partie de code JAVA que l’on a demande, il a fallu traiter toutes les données que l’on avait dans le fichier XML qui était fournie avec notre projet. Ainsi, en ouvrant le fichier **cergy.osm.xml** et en le comprenant, nous avons réussi à en sortir un MCD adapté. Pour rappel, le modèle entité-association est un modèle de données pour des descriptions de haut niveau de modèles conceptuels de données. Cela permet d’avoir une description graphique pour représenter notre base de données.

Voici le MCD que l’on obtient :



 Ainsi, toutes les entités que l’on a dans notre XML comme les nœuds ou les tags y sont représentées avec leurs cardinalités respectives. Grâce au logiciel AnalyseSI, avec lequel nous fait notre MCD, nous pouvons en faire ressortir nos tables SQL.

 Le problème avec le logiciel AnalyseSI est que les tables SQL que l’on obtient sont erronées. En effet, les types des différents ID ne sont pas toutes cohérentes. Par exemple, pour l’idU, on obtient sur AnalyseSI comme type un « AUTO\_INCREMENT » alors que l’on devrait avoir un « INT ». Ainsi, dans un deuxième cas, nous devions modifier le résultat des tables que l’on avait avec le logiciel.

 Une fois cette étape faite, nous allons transformer notre XML en un format XSL. Cette étape est importante puisque le XML est un langage de structuration des données, et non de représentation des données. Le XSL est un langage recommandé par le W3C (organisme de normalisation) pour effectuer la représentation des données de documents XML. Ce fichier est nommé **eisti.osm.little.xsl**.

 Une fois que nous avons obtenu tout cela, nous avons donc essayé de voir si toutes nos tables étaient bonnes et si la relation entre le XML et les tables se faisaient bien. Nous avons eu un problème lors de la compilation. Nous avons du modifier le fichier XML qui était fournie avec notre travail car, il comportait des accents qui n’étaient pas lu.

 Le fichier XML modifié, nous avons pu finir la partie base de données de notre projet car, avec les tables, le XML et le XSL, nous avons donc pu obtenir notre base de donnés.

 Ainsi, nous avons commencé à faire la partie JAVA avec une IHM