

UN CANEVAS LOGICIEL POUR INFORMATIQUE DECISIONNELLE

INTRODUCTION

L'informatique décisionnelle (Business intelligence en anglais ou BI) se réfère principalement aux techniques informatiques utilisées pour l'identification, l'extraction, et l'analyse des données d'affaires, tels que le chiffre d'affaires par produit et/ou département, ou par les coûts associés et les revenus. Les solutions de Business Intelligence proposent généralement des fonctionnalités dans les domaines suivants : rapports, tableaux de bord, exploration de données et analyse prédictive (Moss & Atre, 2003).

Avec les progrès constants en informatique et le développement des applications Web et mobiles, l'amélioration des interfaces utilisateurs est nécessaire pour traiter différents types d'utilisateurs métier. Un des facteurs clés de la réussite de l'évolution du Web 2.0 était l'accent mis sur l'expérience utilisateur. Depuis lors, la convivialité et l'intuitivité sont devenues une exigence clé pour n'importe quelle application.

CONTEXT DESCRIPTION

En BI, les utilisateurs métier effectuent généralement quatre types de processus: sélection de données, manipulation de données, analyse des données et visualisation de données. Après un cycle complet, l'utilisateur a la possibilité d'effectuer ces processus de nouveau, nous appelons cela le cycle de BI. Il est nécessaire de présenter des données détaillées pour ces processus (Celino & Corcoglioniti, 2010). Les interfaces utilisateur doivent être visuellement optimisées pour les présenter de la meilleure façon, même si les données sont en grande quantité. Les utilisateurs doivent pouvoir y accéder facilement, les visualiser, les manipuler, chercher d'autres données, extraire des informations, et découvrir des relations entre les données (Goldstein & Roth, 1994). En général, ils veulent éviter la charge cognitive de chercher ce qu'ils désirent; ils veulent des outils simples mais puissants. La simplicité est cependant quelque peu absente de tous les différents processus.

Les interactions avec les outils logiciels BI, tels que les graphiques, les outils d'analyse ou tableaux de bord, doivent tenir compte de nombreux facteurs. Tout d'abord, la diminution rapide des coûts de la bande passante réseau et la puissance de traitement et de stockage ont rendu les outils BI disponibles à un large et diversifié ensemble d'utilisateurs. Les outils de BI doivent donc être utilisables par des non experts en informatique. Deuxièmement, les outils doivent faire face à différentes sources de données et à des formats de données différents et doivent cacher cette hétérogénéité aux utilisateurs. Troisièmement, les outils de BI doivent fournir des interfaces simples permettant aux utilisateurs de créer de nouveaux contenus BI en mixant des données de l'entreprise et des données externes. Enfin, l'analyse des données doit être effectuée en tenant compte de la capacité de calcul de la plateforme sous-jacente. Par conséquent, capturer les données les plus pertinentes et les délivrer aux bonnes personnes au bon moment et de la meilleure façon est essentiel.

Nous vivons maintenant dans l'ère de l'informatique mobile. Nous ne sommes plus liés à un navigateur sur une machine sédentaire. Les utilisateurs doivent accéder aux données clés de BI depuis leurs périphériques mobiles lors de leurs déplacements. Cela implique que les interfaces doivent communiquer les résultats sur des écrans de taille limitée. Cette évolution radicale exige de nouveaux paradigmes d'interaction en gardant à l'esprit les attentes des utilisateurs, comme l'interaction multi-touch sur des tablettes, les gestes naturels et les interfaces en langage naturel tels que les assistants personnels.

RESEARCH OBJECTIVES

L'objectif de cette recherche est de fournir un canevas logiciel qui permet aux utilisateurs métier d'avoir une meilleure interaction avec les différents processus BI en s'attaquant aux restrictions qui influent sur

l'expérience utilisateur, et en fournissant des recommandations concises et relatives à chaque processus. Cela permettra aux utilisateurs d'avoir un meilleur aperçu et une analyse de leurs données, ce qui conduira à des décisions stratégiques plus efficaces et plus précises. Ce canevas logiciel interviendra sur le cycle complet de BI, de la sélection, la manipulation et la transformation des données à leur visualisation. La figure 1 illustre les défis qui se posent pour chaque processus.

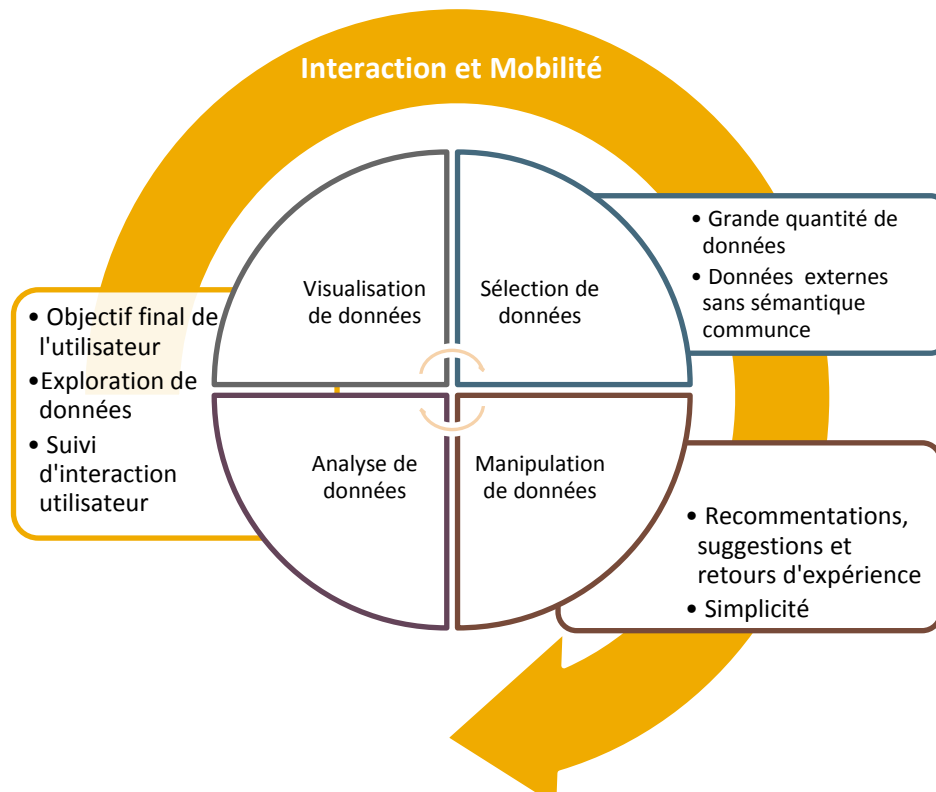


Figure 1 – Processus BI and leurs défis

Les défis dans le processus de Sélection des Données proviennent du besoin de travailler avec des ensembles de données potentiellement très importants. Par exemple, comment pouvons-nous décider quelle partie des données doit être présentée aux utilisateurs et comment pouvons-nous afficher cette portion de données? Travailler avec les sources de données externes est un autre défi. Comment pouvons-nous présenter les différentes sources de façon intuitive?

Dans le processus de Manipulation de Données, d'autres défis apparaissent. L'objectif principal est la simplicité dans le traitement des données. Comment pouvons-nous concevoir une interface intuitive qui permet aux utilisateurs de manipuler des données (analyse des valeurs comme des chaînes d'adresses, analyse et nettoyage des données ou des valeurs temporelles, les dédoublant ou les agrégeant, etc.)? En outre, il est très important à ce stade de présenter des suggestions et des recommandations aux utilisateurs pour qu'ils obtiennent de meilleurs résultats. Recueillir les expériences et tracer les interactions des utilisateurs sont nécessaires pour améliorer la qualité des futures recommandations.

Enfin, l'Analyse des Données et la Visualisation sont étroitement liés dans les défis qu'ils posent et de leur importance dans le cycle de BI. Le traçage des interactions et l'analyse des données présentées contribueront à fournir des suggestions de visualisation plus précises. En outre, si les utilisateurs expriment leur objectif final, de meilleures recommandations et de suggestions à chaque étape du cycle de BI doivent être fournies. Enfin, toute visualisation doit être en mesure de permettre une interaction directe et de communiquer à des

utilisateurs là où ils sont dans les processus, où ils pourraient aller et où ils ont été. Un ensemble de questions découlent de cette exigence en particulier lorsqu'il s'agit de visualisations complexes:

- Comment les visualisations vont être connectées aux données sous-jacentes?
- **Est-ce que l'utilisateur pourra agir directement sur les visualisations sans connaissance a priori des données sous-jacentes?**
- **Comment le canevas peut-il recommander des visualisations?**
- **Quels sont les besoins pour une interface interactive qui supportera la sélection, manipulation et transformation de données ?**
- Comment contrôler la portée et la granularité des données présentées ?
- Comment identifier les interactions intéressantes à enregistrer ? Comment peut-on définir et présenter une interaction? Comment l'enregistrer ? Comment montrer les interactions sauvegardées à l'utilisateur ?
- Est-ce que des interactions erronées peuvent être capturées par la visualisation and peut-on fournir des recommandations pour les corriger ? Comment savoir si une interaction est erronée ou peut mener à une interaction qui l'est ?

Est-il possible pour les visualisations de partager les résultats et les expériences des utilisateurs? Cela en soi ouvre un grand potentiel pour analyser les habitudes d'utilisation et présenter des recommandations. Alors que les interfaces et l'interactivité des visualisations est important, une méthodologie appropriée pour évaluer leurs coûts est nécessaire. Les interactions doivent être formulées et le retour d'expérience des utilisateurs doit être collecté.

De plus, avec le succès des dispositifs mobiles, il est utile d'examiner comment les différents résultats de la recherche peuvent avoir un impact sur l'expérience de l'utilisateur sur les appareils mobiles et comment nous pouvons tirer parti de la plateforme mobile pour fournir une façon plus intuitive d'interagir avec des données et des visualisations.

RESEARCH METHODOLOGY

Les travaux de recherche sont organisés en plusieurs étapes. D'abord, il y aura une évaluation de l'état de l'art. Il s'agira notamment d'une étude approfondie de la littérature et des tests d'outils existants de BI. Le développement de prototypes, la familiarisation avec l'analyse de données et les techniques de visualisation auront également lieu.

Deuxièmement, la modélisation du canevas d'interaction aura lieu et le système sera implémenté. La modélisation prendra en compte les résultats de l'état de l'art.

Troisièmement, le canevas logiciel sera évalué à la fois du point de vue fonctionnel et du point de vue utilisateur. Côté fonctionnel, les besoins en processeur et mémoire, les résolutions d'écran et la mobilité seront mesurés, ainsi que la réactivité du système. Côté utilisateur, des questionnaires, entretiens et groupes de discussions seront utilisés pour obtenir une évaluation qualitative. Les données d'interaction seront inspectées pour analyser le comportement des utilisateurs. Les résultats seront comparés à l'état de l'art.

Enfin, les six derniers mois seront consacrés à la rédaction de la thèse.

BIBLIOGRAPHY

Celino, I., & Corcoglioniti, F. (2010). Towards the Formalization of Interaction Semantics. *I-SEMANTICS*. Graz, Austria: ACM.

- Cockburn, A., Karlson, A., & Bederson, B. (2007). A Review of Overview+Detail, Zooming, and Focus+Context Interfaces. *ACM*.
- Cyganiak, R., Reynolds, D., & Tennison, J. (2010, July 14). *The RDF Data Cube vocabulary*. Retrieved June 30, 2011, from <http://publishing-statistical-data.googlecode.com/svn/trunk/specs/src/main/html/cube.html>
- Di Lorenzo, G., Hacid, H., Paik, H.-y., & Benatallah, B. (2009). Data Integration in Mashups. *SIGMOD Record*, Vol 38, No. 1, 59-66.
- Evelson, B. (2011). *Trends 2011 And Beyond: Business Intelligence*. Forrester Research, Inc.
- Goldstein, J., & Roth, S. F. (1994). Using Aggregation and Dynamic Queries for Exploring Large Data Sets. *CHI'94 Human Factors in Computing Systems*. ACM.
- Golfarelli, M., Maio, D., & Rizzi, S. (1998). The Dimensional Fact Model: a conceptual model for data warehouses. *International Journal of Cooperative Information Systems*, 215-247 (Volume 7).
- Horner, J., Song, I.-Y., & Chen, P. P. (2004). An Analysis of Additivity in OLAP Systems. *Proceedings of the 7th ACM international workshop on Data warehousing and OLAP* (pp. 83-91). Washington, DC, USA: ACM New York, NY, USA.
- Jankun-Kelly, T., Ma, K.-L., & Gertz, M. (n.d.). *A Model and Framework for Visualization Exploration*. Mississippi State University, Department of Computer Science and Engineering.
- Johnson, J. (2010). *Designing with the - Mind in Mind Simple Guide to Understanding User Interface Design Rules*. E lsevier Inc.
- Keefe, D. (2010). Integrating Visualization and Interaction Research to Imporve Scientific Workflows. *IEEE Computer Society*.
- Lam, H. (2008). A Framework of Interaction Costs in Information Visualization. *IEEE TRANSACTIONS ON VISUALIZATION AND COMPUTER GRAPHICS*, 14(6), 1149-1157.
- Lam, H., & Munzner, T. (2010). *A Guide to Visual Mutli-Level Interface Design From Synthesis of Emperical Study Evidence*. UBC Computer Science Technical Report.
- Landauer, T., & Prabhu, P. (1997). *Handook of Human-Computer Interaction*. Elsevier Science B.V.
- Lenz, H.-J., & Shoshani, A. (1997). Summarizability in OLAP and Statistical Data bases. *SSDBM '97 Proceedings of the Ninth International Conference on Scientific and Statistical Database Management* (pp. 132-143). IEEE Computer Society Washington, DC, USA.
- Malinowski, E., & Zimanyi, E. (2004). OLAP Hierarchies: A Conceptual Perspective. In E. Malinowski, & E. Zimanyi, *Lecture Notes in Computer Science* (pp. 19-35).
- Moss, L., & Atre, S. (2003). *Business Intelligence Roadmap: The Complete Project Lifecycle for Decision-Support Applications*. Addison-Wesley Professional.
- O'Reilly, T. (2007). *Google Admits "Data is the Intel Inside"*. O'Reilly.
- Thor, A., Aumueller, D., & Rahm, E. (2007). Data Integration Support for Mashups. *Proceedings of the Sixth International AAAI Workshop on Information Integration on the Web* (pp. 104-109). Citeseer.