Chapitre 2 Étude Préalable / Concepts Théorique

Dans ce chapitre, nous nous intéressons aux concepts de base liés à notre travail. Par la suite, nous ferons une étude de l'existant en le critiquant. Enfin, nous allons nous intéresser à la présentation de la solution proposée.

2.1 Concepts General

2.1.1 Business Intelligence

« L’informatique décisionnelle (En anglais business intelligence (BI) ou decision support system (DSS)) est l’informatique à l’usage des décideurs et des dirigeants d’entreprises.

Elle désigne les moyens, les outils et les méthodes qui permettent de collecter, consolider, modéliser et restituer les données, matérielles ou immatérielles, d’une entreprise en vue d’offrir une aide à la décision et de permettre à un décideur d’avoir une vue d’ensemble de l’activité traitée. » [12]

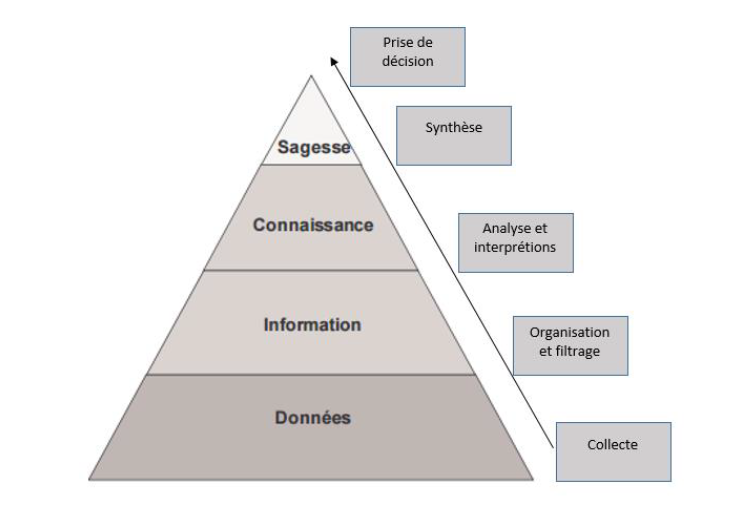
« La Business Intelligence permet aux entreprises de mieux comprendre et d’analyser voire même d’anticiper ce qui se passe en leur sein et dans l’environnement économique en général. La BI permet de transformer des données en informations utiles et pertinentes, puis de les diffuser en temps et lieu voulus, pour faciliter la prise de décisions opportunes en toute connaissance de cause. Elle combine les données issues de sources diverses pour offrir une vue d’ensemble unifiée et actualisée. » [2]

**Avantages**



À la suite de cette définition, nous introduisons les concepts clés qui forment la chaîne appelée Modèle DIKW.

Il met en œuvre les notions de bases expliquées ci-dessous et leur hiérarchie comme les montre la figure suivante.



§ Données : Ce sont les faits bruts. C’est le résultat direct d’une mesure qui peut être obtenu par un outil de mesure

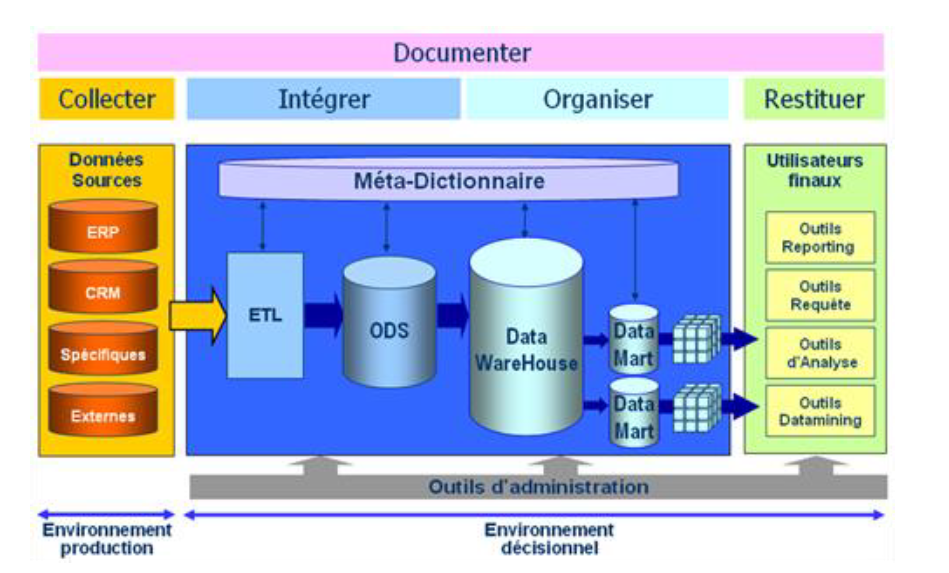
§ Information : c’est une donnée avec une signification, liée à un contexte.

§ Connaissance : c’est le résultat des analyses et des réflexions sur les informations obtenues. Elle se base sur le suivi et l’interprétation des informations et des expériences.

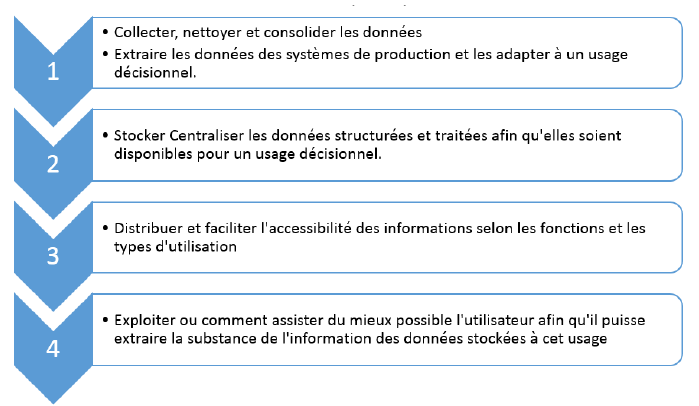
§ Sagesse : c’est savoir utiliser les connaissances, c’est donc l’art de la prise de décisions.

**Les phases de l’analyse décisionnelle**

L’informatique décisionnelle désigne les moyens, les outils et les méthodes qui permettent de collecter, intégrer, organiser et restituer les données immatérielles d’une entreprise en vue d’offrir une aide à la décision et de permettre aux responsables de la stratégie d’entreprise d’avoir une vue d’ensemble de l’activité traitée. Pour la réussite d’un projet BI il est nécessaire de mettre en place différentes étapes dans une chaine décisionnelle afin de profiter pleinement et efficacement d’une plateforme de business intelligence.

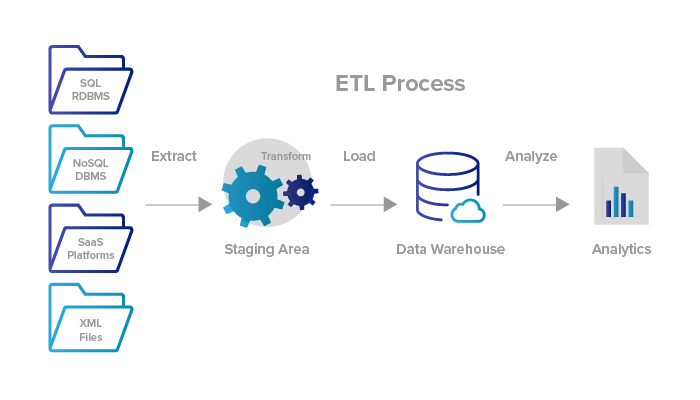


Ce schéma résume les 4 phases de l’analyse décisionnelle :



**ETL (Extract, Transform, Load)**

Signifie Extract-Transform-Load. C’est un processus d'entreposage de données chargé de retirer des données des systèmes source et de les placer dans un entrepôt de données. L’ETL comprend aussi une étape distincte de nettoyage. La séquence est alors Extract-Clean-Transform-Load



Extract

L'étape d'extraction comprend l'extraction de données à partir du système source et le rend accessible à un traitement ultérieur. L'étape de nettoyage est l'une des plus importantes car elle garantit la qualité des données dans l'entrepôt de données.

Transform

L'étape de transformation applique un ensemble de règles pour transformer les données de la source vers la cible. Cela comprend la conversion de toutes les données mesurées à la même dimension en utilisant les mêmes unités pour qu'elles puissent ensuite être rejointes. L'étape de transformation exige également de joindre les données provenant de plusieurs sources, générant des agrégats, des clés de substitution, le tri, dériver de nouvelles valeurs calculées et l'application des règles de validation avancées. [7]

Load

Lors de l'étape de chargement, il est nécessaire de veiller à ce que le chargement soit exécuté correctement et avec le moins de ressources que possible. La cible du processus Load est souvent une base de données.

Data Warehouse

Le Data Warehouse est une base de données recueillant et gérant toutes les données collectées au sein de l'organisme, dans le cadre de la prise de décision. Son but principal est de soutenir l'analyse et faciliter le processus de prise de décision.

Data Warehouse présentent quatre caractéristiques spécifiques.

* Orienté sujet : les données sont organisées par thème.
* Intégré : les données provenant de sources hétérogènes, elles utilisent chacune un type de format. Elles doivent donc être intégrées avant d'être proposées à utilisation.
* Non volatile : les données ne disparaissent pas et ne changent pas au fil des traitements, au fil du temps.
* Historisé : les données sont horodatées, afin de visualiser l'évolution dans le temps d'une valeur donnée.

OLAP

OLAP est un acronyme Online Analytical Processing. OLAP effectue une analyse multidimensionnelle des données et offre la possibilité de calculs complexes, d'analyse des tendances et de modélisation sophistiquée des données. OLAP permet aux utilisateurs finaux d'effectuer une analyse des données en plusieurs dimensions, fournissant ainsi la compréhension dont ils ont besoin pour une meilleure prise de décision.



Le tableau de bord

Le tableau de bord appartient à l’ensemble des outils de pilotage qui permettent aux gestionnaires de prendre connaissance de l’état de l’évolution des systèmes qu’ils pilotent et d’identifier les tendances C’est un outil qui nous renvoi une synthèse du passé et des prévisions sur le futur en intégrant des alertes aux décideurs pour raffiner au meilleur la prise de décision. Ce fut un outil d’évaluation de la performance par le biais des indicateurs sur des périodes précises.

Les tableaux de bord jouent un rôle essentiel dans le suivi des performances de l’organisation.

Ils remplissent plusieurs fonctions.

* Responsabilisation : Implication des équipes dans l’identification, la réalisation et le suivi des performances.
* Mesure du progrès : Information sur le degré de réalisation des objectifs, mise en évidence des écarts entre prévisions et réalisations grâce aux indicateurs de performance
* Anticipation : Alertes dès lors qu’une tendance met en évidence un risque ou un problème
* Aide à la décision : Diagnostic de la situation et analyse de l’information pour faire des choix pertinents.

Le reporting

C’est la préparation des rapports en se basant sur des activités et des objectifs souhaités. C’est une présentation périodique sous une structure prédéterminée telles que des présentations graphiques et des grandeurs de calculs afin d’avoir des bilans pertinents.

L’objectif du reporting est d’informer les chargés de contrôle et de supervision en interne comme en externe, ou bien tout simplement ceux qui sont concernés par ces activités ou ces résultats, on parle alors de dirigeants et décideurs

Data minning

 Data Mining, également appelée découverte de données ou de connaissances, consiste à analyser des données et à les transformer en un aperçu qui informe les décisions d'affaires.

C’est ainsi le procédé qui permet d’extraire des corrélations, dégager et détecter des tendances, découvrir des patterns entre les données permettant d’expliquer le passé ou prédire le futur ainsi que de dégager de la valeur à partir de données auparavant inexploitables.

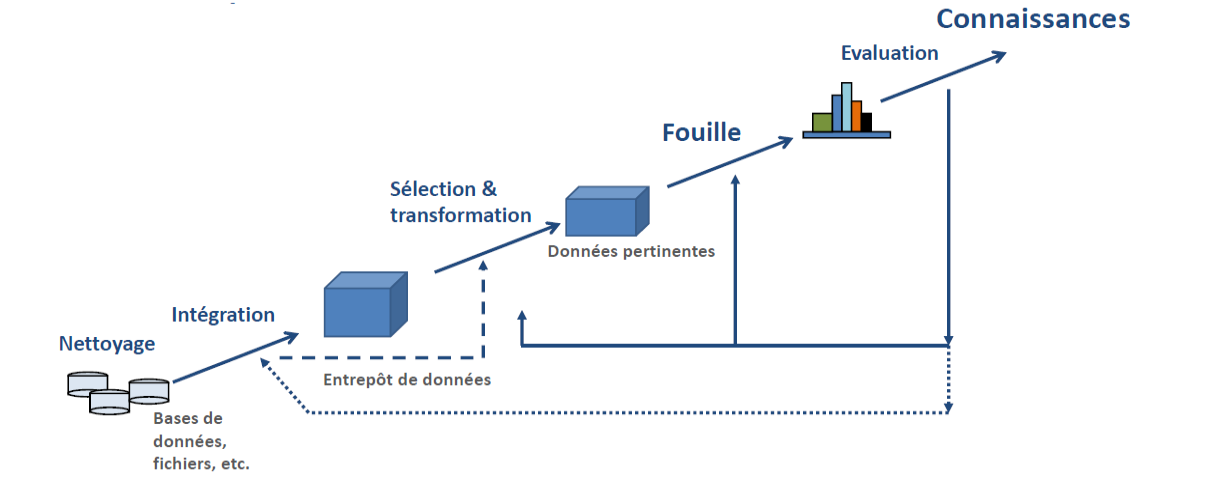
Ce que l'on peut traduire par la transformation des données en connaissances.

Parmi les objectifs du data mining :

* Prédire les tendances futures
* Déterminer les habitudes de la clientèle
* Aide à la prise de décision
* Analyse basée sur le marché
* Détection rapide de la fraude

**Processus d’Extraction de Connaissances à partir des Données**

L'exploration de données est un processus itératif qui implique généralement les phases suivantes :



**1. Développer et comprendre le domaine de l’application**

C’est le premier pas dans ce processus. Il prépare la scène pour comprendre et développer les buts de l’application.

**2. Sélection des données**

La sélection et la création d’un ensemble de données sur lequel va être appliqué le processus d’exploration.

**3. Le prétraitement et le nettoyage des données**

Cette étape inclut des opérations comme l’enlèvement du bruit et des valeurs aberrantes si nécessaire, des décisions sur les stratégies qui vont être utilisées pour traiter les valeurs manquantes.

**4. La transformation des données**

Cette étape est très importante pour la réussite du projet et doit être adaptée en fonction de chaque base de données et des objectifs du projet.

Dans cette étape nous cherchons les méthodes correctes pour représenter les données. Ces méthodes incluent la réduction des dimensions et la transformation des attributs.

**5. Choisir la meilleure tâche pour Data mining**

Nous devrons choisir le type de Datamining à utiliser et en décidant le but du modèle.

Par exemple : classification, régression, regroupement...

**6. Choisir l’algorithme de Data mining**

Dans cette étape nous sommes amenés à choisir la méthode spécifique pour faire la recherche des motifs, en décidant quel modèle et des paramètres sont appropriés.

**7. Implémenter l’algorithme de Data mining**

Dans cette étape nous procéderons à l’implémentation des algorithmes de Data mining choisis dans l’étape antérieure.

Il se peut qu’on applique l’algorithme plusieurs fois pour avoir le résultat attendu.

Obtenir le data patterns.

**8. Evaluation**

Evaluation et interprétation des patterns découverts.

Cette étape donne la possibilité de :

* Retourner à une des étapes précédentes.
* Avoir une représentation visuelle des motifs.
* Enlever les motifs redondants ou non-représentatifs.
* Les transformer dans des termes compréhensibles pour l’utilisateur.

**9. Utiliser les connaissances découvertes**

Incorporation de ces connaissances dans d’autres systèmes pour d’autres actions.

Nous devrons aussi mesurer l’effet de ces connaissances sur le système, vérifier et résoudre les conflits possibles avec les connaissances antérieures.

2.2 Etude et critique de l’existant

Actuellement, OdasLab dispose d’une application qui permet de gérer les activités quotidiennes de transférer, retirer et déposer l’argent.

Certes les outils de l’application bancaire sont satisfaisants pour les transactions mais d’un autre côté ces outils présentent plusieurs déficiences, comme :

**-Absence de Business Intelligence :** A l’heure actuelle, l’application bancaire ne possède pas un outil BI qui permet la génération des rapports et tableau de bord et qui faciliteront la prise de décision et offriront une vision globale et détaillée. Les rapports et les tableaux de bord sont générés à l’aide de EXCEL.

**-Difficulté dans l’élaboration des rapports d’activité :** L’élaboration des rapports d’activité fait intervenir plusieurs intermédiaires. C’est à dire qu’à chaque fois qu’il est nécessaire d’établir les rapports d’activité, il faudra procéder à l’extraction des données à partir de plusieurs sources de données qui sont réparties sur plusieurs serveurs pour les diriger vers une structure centralisée.

**-Absence de moyen pour découvrir les transactions frauduleuses :** lors qu’il y a des transactions frauduleuses c'est-à-dire en cas d'opérations pour lesquelles vous n'avez pas donné votre consentement l’application bancaire ne permet pas de les détectées

**-Lenteur de la procédure de Reporting :** Le processus de Reporting actuel confronte certaines difficultés. Les décideurs ont besoin des rapports dans les brefs délais. Vu la charge du travail et la lourdeur de la procédure, ce processus prend plus du temps qu’il le faut.

2.3 Solutions proposées

**Objectifs à réaliser**

Ce projet a pour objectif de concevoir et développer une application bancaire qui permet les activités quotidiennes de transférer l’argent, retirer et déposer muni d’un tableau de bord et des rapport basé sur l’informatique décisionnelle qui vont aider à la prise de décision et offrir une vision globale et détaillée, ainsi la détection des fraudes grâce aux fouilles de données « data mining ».

En effet, le travail consiste à concevoir un « système centralisé », un entrepôt de données rassemblant toutes les données relatives. L’idée est d’arriver, à la fin du processus, de générer des rapports de qualité ainsi que la construction d’un tableau de bord destiné aux décideurs pour avoir une vision profonde sur l’activité de l’entreprise et qui mène à des décisions innovantes pour bien améliorer le domaine d’activité de la banque.

La solution décrite ci-dessus doit permettre :

* La conception d’un Data Warehouse de données centralisé, intégré, non volatile et historisé.
* L’extraction des données depuis le Data Warehouse pour construire des cubes qui sont des éléments clés de l’analyse multidimensionnelle.
* L’utilisation des cubes et du Data Warehouse dont la mise à jour est automatique pour créer des rapports et des tableaux de Bord.
* Déploiement tableaux de bord et des rapports et leur envoie automatique par courrier électronique « email ».
* Une analyse fine des données en utilisant la technique de datamining afin de détecter des activités frauduleuses

2.4 Conclusion

Durant ce chapitre, nous avons des Concepts de base liés à notre travail. Ensuite, nous avons étudié l’existant en soulignant ses limites, Ensuite, nous avons proposé une solution qui répond aux besoins des décideurs de la banque. Le prochain chapitre sera consacré à l’analyse et la spécification des besoins.