

Business Intelligence et KPI

1. Outils Conceptuels

Salim Lardjane
Université de Bretagne Sud

Définition

- Le première référence à la Business Intelligence date de 1958, dans un article de H. P. Luhn.
- Il définit le terme anglais « intelligence » comme la « *capacité à appréhender les inter-relations entre des faits de façon à guider l'action vers un objectif souhaité* ».
- Il définit le terme « business », que nous traduirons par **entreprise**, comme « *une collection d'activités effectuées en vue d'un objectif quelconque, qu'il soit scientifique, technique, commercial, industriel, juridique, politique, militaire, etc* ».

Définition

- Il définit les **Systemes de Business Intelligence** comme « *des systèmes automatiques développés pour disséminer de l'information vers différentes parties d'une organisation industrielle, scientifique ou gouvernementale* ».
- Cette définition a précédé l'émergence de la BI de 30 ans et c'est en 1989 que Howard Dresner proposa d'utiliser terme BI à nouveau.
- Il entendait par **BI** un *ensemble de concepts et méthodes utilisées pour améliorer la prise de décision en se basant sur des faits*.

Définition

- De nombreuses définitions analogues de la BI ont été données depuis.
- Par exemple, chez Negash (2004), des aspects importants de la BI sont mis en évidence dans le paragraphe suivant : « *...les systèmes de business intelligence fournissent de l'information actionnable au bon moment, au bon endroit et sous la bonne forme pour assister les décideurs.* »

Définition

- Plus récemment, chez Rausch et al. (2013), la BI est définie comme « *une approche globale, intégrée, spécifique à chaque entreprise, basée sur les technologies de l'information, de l'aide à la décision managériale* ».
- Wikipedia définit la BI comme « *un ensemble de théories, méthodologies, processus, architectures et technologies qui transforment des données brutes en information utile et signifiante pour les objectifs de l'entreprise* ».

Définition

- On peut résumer ces différentes définitions de la BI à l'aide des caractérisations suivantes :
- **Tâches de la BI** : fournir une **aide à la décision** en vue d'objectifs spécifiques, définis dans le contexte des activités de l'entreprise dans différentes domaines, et prenant en compte le cadre organisationnel et institutionnel.
- **Fondement de la BI** : C'est **l'information empirique** obtenue à partir des données. En plus de cette base empirique, la BI utilise différents types de connaissances et de théories pour générer de l'information.

Définition

- **Réalisation de la BI** : La BI se réalise dans un **systeme d'aide à la décision** basé sur les Technologies de l'Information et de la Communication (TIC) les plus récentes.
- **Livrables de la BI** : Un système de BI doit fournir l'information au bon moment, aux bonnes personnes et sous la bonne forme.

Historique

- L'évolution de la BI suit celle des TIC et de la disponibilité des données.
- La préhistoire de la BI peut être résumée par l'acronyme **DSS** (Decision Support Systems, Systèmes d'Aide à la Décision) et est décrite dans un article de Power (2007).
- Elle couvre la période s'étendant des années 1960 à la fin des années 1980. Au cours de cette période, on a développé des DSS basés sur les modèles, les données, les communications, les documents et les connaissances.

Historique

- Suite à la définition de Dresner (1989), le terme BI est devenu populaire dans les années 1990, mais n'était compris que comme une aide à la décision basée sur les données, étroitement liée au développement des entrepôts de données (**data warehouses**), à l'utilisation du Traitement Analytique en Ligne (**OLAP** : On Line Analytical Processing) et d'outils de **reporting**.
- Par la suite, et parallèlement au développement du **data management**, d'autres outils d'analyse, comme le **data mining** ou la **modélisation prédictive** sont devenus populaires en BI.

Historique

- On a parfois utilisé le terme **Business Analytics** pour regrouper les outils de modélisation, de prévision et de reporting utilisés en BI.
- La BI a pu donner alors l'impression d'être un ensemble hétérogène d'outils peu liés les uns aux autres, utilisés pour faire différentes choses au sein d'une entreprise.
- Mais ces différents axes de développement ont été ensuite consolidés au cours des années 2000 en mettant de nouveau l'accent sur **la problématique de l'aide à la décision**.

Défis actuels

- Au cours des dernières années, la disponibilité des données et les capacités d'analyse ont augmenté de façon très remarquable et de nouveaux champs d'investigation se sont ouverts pour la BI.
- Les défis actuels de la BI sont les suivants :

Défis actuels

- **Tâches de la BI** : Il y a aujourd'hui une bonne compréhension de la logique d'entreprise dans pratiquement tous les domaines. Cela conduit à une perspective conceptuelle basée sur les **processus** (on en parlera dans la suite), qui intègre à la BI le **process mining** et **l'organisation du travail**. D'autre part, de nouveaux types d'organisation, par exemple les organisations décentralisées, ont besoin d'outils d'aide à la décision, ce qui fait que des idées liées à **l'intelligence collective** et au **crowd-sourcing** sont aujourd'hui appliquées en BI.

Défis actuels

- **Fondement de la BI** : Au delà des entrepôts de données traditionnels, on doit aujourd'hui prendre en compte **les données sur le Web**. Ces dernière ne sont en général pas totalement structurées, mais seulement semi-structurées, à l'image des données textuelles. Le besoin d'intégrer de façon cohérente différentes sources de données pour l'aide à la décision a conduit à l'émergence, en BI, de nouveaux **modèles d'assimilation des données**. En lien avec ces nouveaux types de données, le champ des outils analytiques utilisés s'est élargi pour inclure le **visual mining**, le **text mining**, l'**opinion mining** et l'**analyse des réseaux sociaux**.

Défis actuels

- **Réalisation de la BI** : Les architectures logicielles contemporaines autorisent de nouvelles réalisations des systèmes de BI. Du point de vue utilisateur, l'approche **SaaS** (Software as a Service - logiciel accessible via Internet contre abonnement) constitue un développement intéressant. Du point de vue computationnel, on doit traiter aujourd'hui des données massives et complexes. De plus, le **cloud** et le **calcul distribué** offrent de nouvelles opportunités à la BI.

Défis actuels

- **Livrables de la BI** : Les appareils mobiles (smartphones, tablettes) offrent de nouvelles opportunités pour la **transmission de l'information en temps réel**. Toutefois, ces développements doivent prendre en compte la **qualité** de l'information ainsi transmise.

Contextualiser la BI

- On a caractérisé la BI et énoncé ses objectifs de façon générale.
- Afin de préciser tout cela, on doit envisager la relation entre entreprise et BI du point de vue du management.
- On entend le terme **entreprise** au **sens large** *i.e.* comme « toute activité d'une organisation visant à fournir des biens ou des services à des clients ».

Contextualiser la BI

- L'organisation peut être active dans différents domaines. Il peut s'agir d'une Entreprise, d'une administration, d'un hôpital, d'une institution éducative comme une Université, etc.
- En plus du domaine d'application, on doit noter que l'aide à la décision est requise par des entreprises de **taille** et de **portée** différentes.
- Par **taille**, on entend une classification des organisation selon des critères tels que le nombre d'employés (PME ou grande Entreprise), implantation (locale ou globale), nombre de clients, chiffre d'affaire.

Contextualiser la BI

- Par **portée** on entend le **nombre d'activités** de l'organisation pour lesquelles une aide à la décision est requise.
- Par exemple, en management, on peut souhaiter disposer d'une aide à la décision au niveau global d'une Entreprise ou à un niveau fonctionnel spécifique (par exemple, Production ou Marketing).
- Dans les applications médicales, on peut mettre l'accent sur l'aide à la décision pour le traitement d'une maladie spécifique ou pour la gestion d'un hôpital.
- Dans un contexte administratif, on peut avoir besoin d'aide à la décision pour organiser efficacement les services ou pour améliorer la satisfaction des clients.

Contextualiser la BI

- Afin de disposer d'un cadre général pour envisager ces différents problèmes, on va organiser les activités de BI selon les principes de la gestion d'Entreprise.
- A cet effet, on distingue le niveau **managérial**, le niveau **organisationnel** et le niveau **fonctionnel**.
- On distingue également différents niveaux **d'organisation** et **d'acquisition des données**.

Contextualiser la BI

- La rôle de la BI en lien avec des **modèles d'entreprise** (**business models**) doit être envisagé.
- Il y a diverses définitions de la notion de **business model** mais pour notre exposé la caractérisation suivante sera suffisante : *un business model reflète la stratégie d'une entreprise pour créer de la valeur.*

Scénarios BI

- On peut alors distinguer **quatre scénarios** qui relient la BI au contexte de l'entreprise, allant d'une simple application de l'aide à la décision à un problème spécifique jusqu'à la situation où la BI est une partie essentielle de la planification stratégique.

Scénarios BI

- **Scénario 1** : *La BI est séparée du management stratégique*. Dans cette situation, la BI s'intéresse principalement à des **objectifs à court terme** dans une **division** d'une organisation ; par exemple, un département d'une Entreprise, une clinique dans un Hôpital. Les résultats de la BI sont typiquement des **rapports** plus ou moins standardisés.

Scénarios BI

- **Scénario 2** : *La BI est utilisée pour de suivre la performance de la stratégie d'entreprise*. Dans ce cas, l'application de la BI est motivée par des objectifs stratégiques globaux et formulée en accord avec ces objectifs. Le suivi de la performance est effectué en définissant des **objectifs mesurables**. Un **entrepôt de données** permettant d'obtenir une vue unifiée de l'entreprise est souvent requis dans ce scénario.

Scénarios BI

- **Scénario 3** : *Il y a rétroaction (feedback) de la BI sur la stratégie d'entreprise*. Dans ce cas, l'application de la BI va au delà du scénario 2 et vise à évaluer la performance à l'aide de méthodes analytiques. Dans le meilleur des cas, elle autorise **l'optimisation** de la stratégie. Un résultat typique dans ce scénario est un **tableau de bord prospectif** (ang. balanced scorecard).

Scénarios BI

- **Scénario 4** : *La BI est une ressource stratégique*. Dans ce cas, l'information générée par la BI est utilisée non seulement pour l'optimisation de la performance, mais également comme intrant essentiel dans la définition de la stratégie d'entreprise au niveau managérial. Des exemples typiques sont le marketing centré sur les clients ou le développement de procédures d'opération standards pour le traitement de patients.

Scénarios BI

- La classification précédente dépend de la **taille** de l'organisation et de la **portée** de l'entreprise considérée.
- Par exemple, une application de la BI dans un département universitaire peut être utilisée pour une rétroaction sur la stratégie du département mais également pour suivre la performance au niveau de l'Université.

Scénarios BI

- A première vue, les scénarios 3 et 4 semblent être les plus favorables, mais il faut en général prendre en compte les spécificités de chaque situation, notamment la **quantité de ressources allouées à la BI** et la **disponibilité de l'information**.
- Pour de grandes Entreprises de production, le scénario 3 est souvent le bon choix et dans les Entreprises de services le scénario 4 est à la base de nombreux succès.
- Toutefois, dans bien des cas, les problèmes de décision à traiter sont spécifiques, difficiles à formaliser et il n'est pas clair si l'implémentation d'une stratégie de haut niveau est rentable à long terme.

Perspectives en BI

- Après avoir examiné la stratégie BI globale, nous allons nous intéresser à la structure des activités de l'entreprise.
- La description de cette structure est souvent faite à l'aide de processus (business processus).
- On utilisera le terme de **processus** pour rendre compte *d'un ensemble d'activités structurées en lien les unes avec les autres, visant à fournir un bien ou un service aux clients, et des activités de réponse des clients.*

Perspectives en BI

- Notons que cette définition est plus générale que celle donnée par la norme **ISO 9000:2015** qui définit un **processus** comme *un ensemble d'activités en relation ou en interaction qui utilisent des inputs pour fournir un résultat attendu (output)*.
- En effet, elle prend en compte également les activités des clients en réponse au produit ou au service fourni, ce qui est plus adapté à la BI.

Perspectives en BI

- On interprètera dans la suite **toute activité** effectuée dans le cadre d'une entreprise comme un **processus**, ce qui signifie que toute activité se déroule au cours d'une période de temps et selon certaines règles.
- Toutefois, le fait de prendre en compte les activités des clients limite la possibilité de considérer les processus comme des ensembles d'activités gouvernées uniquement par des règles.

Perspectives en BI

- Au lieu d'une telle représentation **mécaniste** des processus, la BI s'intéresse à la **réalisation empirique** des processus, incarnée par les **instances** (réalisations) de ceux-ci.
- Afin d'étudier ces **instances**, on introduit trois **perspectives** sur les processus.

Perspectives en BI

- **La perspective de production** : celle-ci envisage l'aide à la décision comme apportant des réponses à des questions telles que :
 1. Quel type de produit offrir aux clients ?
 2. Comment organiser la production ?
- Cette perspective joue un rôle important dans le **développement de produits** et **l'organisation interne** de l'entreprise.

Perspectives en BI

- **La perspective client** : celle-ci met l'accent sur le **comportement des clients** et vise à comprendre comment ils perçoivent les produits ou services offerts et comment il réagissent à cette offre.
- La perspective client joue un rôle essentiel dans les entreprises orientées vers les services.

Perspectives en BI

- **La perspective organisationnelle** : celle-ci s'intéresse aux fondements organisationnels de l'entreprise. Cela peut inclure **l'organisation des opérations** en lien avec la perspective de production ainsi que **l'influence des réseaux sociaux** sur le comportement des clients.
- **La perspective adoptée dépend du domaine d'application, de la taille et de la portée de l'entreprise.**

Perspectives en BI

- Les applications font généralement intervenir **les trois perspectives** simultanément mais, en BI, la typologie précédente est utile pour sélectionner **l'information pertinente** et choisir **un modèle analytique**.
- Elle reflète également, dans une certaine mesure, le développement historique des méthodes et modèles utilisés aujourd'hui en BI.

Perspectives en BI

- Dans la **perspective de production**, la structure organisationnelle de l'entreprise est typiquement spécifiée et maintenue à l'aide de **modèle organisationnels** qui font intervenir différents **rôles**.
- Dans les applications BI, les rôles suivants sont pertinents.

Perspectives en BI

- Le **responsable du processus** (ang. process owner) : *c'est l'entité qui fixe les règles gouvernant le processus.* Traditionnellement, ce rôle est défini dans une perspective de production, mais dans les entreprises orientées vers les services, les responsables du processus peuvent être les clients. C'est le cas, par exemple, des patients qui décident d'un traitement.

Perspectives en BI

- **Les sujets du processus** : *ce sont les entités qui permettent d'identifier les instances du processus*. Dans la plupart des cas, les sujets du processus sont les clients, mais il peut s'agir également de produits ou de réseaux de personnes impliquées dans le processus.

Perspectives en BI

- **Les acteurs** : *ce sont les personnes ou organisations générant des événements affectant le processus*. En gestion, les acteurs font généralement partie de la partie de l'organisation responsable de la production de biens ou de services.

Perspectives en BI

- La **perspective client** requiert souvent une approche plus simple des processus internes à l'entreprise car les clients ne sont en général pas au fait de ceux-ci et ne réagissent que selon leur perception de l'entreprise.
- D'autre part, **comprendre le comportement des clients** requiert une description détaillées de leurs caractéristiques, telles que le sexe, l'âge, le statut social, ainsi que de la nature de leur implication dans le processus.

Perspectives en BI

- Un point important est **l'interaction** entre la perspective de production et la perspective client.
- Cette interaction peut être **simple**, comme dans le cas où un client se contente d'acheter des produits dans un magasin.
- D'autres interactions peuvent être plus **complexes** en faisant intervenir, par exemple, des négociations ou l'utilisation de services multiples.

Perspectives en BI

- Un exemple typique, issu du secteur médical, est le traitement des patients.
- Selon la complexité de l'interaction, une combinaison des différentes perspectives peut être requise pour rendre compte de la situation.

Vues des processus en BI

- L'analyse **structurelle** d'un processus d'entreprise, c'est-à-dire l'analyse **basée sur un modèle** décrivant le processus, est souvent intéressante et utile.
- Toutefois, la BI s'intéresse plutôt à comprendre **le processus réel**.
- Ce processus réel est reflétés par l'exécution, si possible, d'une multitude **d'instances** du processus, qui sont créées, initiées et exécutées selon un modèle rarement complètement explicite.
- **La BI se préoccupe d'exploiter l'information empirique collectée sur les instances d'un processus.**

Vues des processus en BI

- Selon l'effort consacré à la collecte des données, on peut disposer, ou pas, d'un vaste éventail de données sur les instances d'un processus.
- Idéalement, on dispose d'un **journal** (ang. **log**) des **événements** observés lors de l'exécution de chaque instance d'un processus.
- Pour chaque activité, le journal d'un processus commence par enregistrer un **événement de démarrage** (start event), se termine en enregistrant un **événement de fin** (end event) et enregistre également les **événements d'interruption** et de **reprise**.

Vues des processus en BI

- De plus, les **instants d'occurrence** de ces événements sont également enregistrés (usuellement sous la forme d'un « timestamp ») ainsi que des attributs supplémentaires caractérisant l'activité concernée par les événements.
- Il est notamment nécessaire de prévoir un attribut **identifiant** l'activité en question.
- D'autres attributs incluent le résultat de l'activité, les personnes impliquées dans l'activité (essentiellement celles qui travaillent sur celle-ci), le coût de l'activité et les ressources requises pour l'exécution de l'activité.

Vues des processus en BI

- La description de la collecte des données peut laisser penser à un processus complètement automatisé, **ce qui est rarement le cas en pratique**, notamment lorsque les événements sont déclenchés par les clients.
- De plus, **utiliser toute l'information** disponible pour la prise de décision **n'est pas recommandé**, car on peut se perdre dans les détails.
- On a donc besoin en BI de **vues spécifiques** sur les instances des processus étudiés.

Vues des processus en BI

- **La vue par événements** : L'accent est mis sur les **événements** affectant le processus, avec un timestamp de début, un timestamp de fin et, si nécessaire, un timestamp de reprise d'activité après une interruption.
- **La vue par états** : En plus de l'occurrence d'événements, la vue par états prend en compte les **valeurs des attributs**, appelés **variables d'état**, mesurés en lien avec les événements.

Vues des processus en BI

- **La vue transversale** : Dans ce cas, on étudie **l'historique de plusieurs instance du processus** à un instant de référence. Usuellement, cette vue prend en compte l'information sur les événements ainsi que les valeurs des variables d'état et **résume l'information** sur les instances du processus pour la prise de décision.

Vues des processus en BI

- La **vue par événements** met l'accent sur les règles définissant l'ordre partiel des événements affectant le processus dans la perspective de la production.
- L'ordre des événements définit la perspective dite du **flux de contrôle** (control flow) sur le processus.

Vues des processus en BI

- La **vue par états** met l'accent sur **le résultat** des activités impliquées par un processus.
- Cela a pour effet de transporter l'attention sur les **sujets du processus** plutôt que sur le processus lui-même.
- Ce sujets peuvent être des clients, des produits, des services ou un réseau de partenaires.
- La compréhension d'un processus dans cette perspective est basée sur des quantités mesurées, **les variables d'état**.

Vues des processus en BI

- Le processus est traité comme un **système dynamique**.
- Les valeurs des variables d'état changent au cours du temps, soit en raison des activités mêmes, soit en raison de la variabilité inhérente aux sujets du processus.
- Tant dans la **vue par événements** que dans la **vue par états**, la structure de données générée est celle de **série temporelle** (potentiellement **multiple**).

Vues des processus en BI

- La troisième vue, **la vue transversale**, part de l'observation que, dans diverses applications, on s'intéresse à des **quantités agrégées** sur les instances d'un processus à un instant donné.
- Ces agrégations peuvent être obtenues à partir de la vue par événements ou de la vue par états.
- Pour ce qui est de la vue par événements, on peut considérer le **nombre d'instances** des différentes activités ayant lieu au cours d'une période donnée, le **temps d'exécution total** pour ces activités, ou les **ressources consommées** par ces instances.

Vues des processus en BI

- Pour ce qui est de la vue par états, l'agrégation peut consister en une **moyenne des états** des instances sur une période d'observation ou en l'état effectif des instances à un instant donné.
- Les vues par événements, par états et transversale d'un processus peuvent être utilisées en combinaison avec les trois perspectives : client, production et organisation.

Vues des processus en BI

- Dans la perspective **production**, on met généralement l'accent sur la suite d'événements correspondant à la production de biens ou de services.
- Une autre possibilité est de mettre l'accent sur les variables d'état décrivant la production (par exemple, l'utilisation des ressources sur une période de temps donnée), ou sur un résumé des caractéristiques de la production sur une période de temps.

Vues des processus en BI

- Dans la perspective **client**, on adopte en générale les vues par état ou transversale. Les événement jouent un rôle important dans la **classification** des clients.
- Par exemple, un défaut de crédit est un événement qui définit deux classes de clients.
- Dans les applications médicales, la vue par état décrivant le comportement des patients relativement à un paramètre donné est souvent dominante.

Vues des processus en BI

- Dans la perspective **organisation**, la vue transversale est souvent dominante, en particulier pour l'analyse de réseaux de clients, mais les structures organisationnelles générant des événements spécifiques peuvent également être étudiées.

Objectifs de la BI

- A l'origine d'une application BI, il y a en général un **objectif d'analyse**.
- De tels objectifs vont de **l'acquisition d'information** à propos d'un ou plusieurs aspect d'un processus à **l'amélioration du processus** ou à la **compréhension en profondeur** de celui-ci pour atteindre des objectifs stratégiques.
- Les objectifs d'analyse peuvent être formulés de deux manières : à l'aide de **KPI**, ou sous la forme **d'objectifs analytiques**.

Objectifs de la BI

- La première manière est celle basée sur les **KPI** (Key Performance Indicators).
- **Un KPI permet de mesurer la performance d'une entreprise relativement à un de ses objectifs.**
- Dans un contexte économique, un KPI peut référer à l'acquisition nouveaux clients ou à l'amélioration du degré de satisfaction de ces derniers.
- Dans un contexte éducatif, un KPI peut référer au taux d'échec des étudiants, au coût des diplômes ou à la position de l'institution dans un classement.

Objectifs de la BI

- Dans un contexte médical, un KPI peut référer à l'efficacité d'une procédure de traitement ou au bien-être des patients.
- L'identification d'un KPI est basée sur
 1. l'identification d'un processus pré-existant
 2. la définition d'objectifs pour le processus
 3. une **mesure** des résultats de l'entreprise relativement à ses objectifs.
- On verra davantage d'exemples concrets de KPI dans la suite.

KPI comme objectifs

- Un KPI relie les activités d'une entreprise à des objectifs via une quantité mesurable.
- Les KPI peuvent ne prendre en compte que quelques aspects d'un processus comme s'appliquer à l'entreprise dans sa globalité.
- On distingue usuellement entre **indicateurs quantitatifs**, représentés par des nombres, des **indicateurs pratiques**, à l'interface des processus, des **indicateurs directionnels**, indiquant si l'entreprise progresse ou pas, des **indicateurs actionnables**, pour contrôler l'effet du changement et des **indicateurs financiers**.

KPI comme objectifs

- On appelle **facteur influent** tout attribut qui peut avoir une influence sur le comportement d'un KPI dans une perspective BI.
- **Les KPI décrivent une entreprise de façon synthétique.**
- Toutefois, la BI va souvent au delà de la formulation des objectifs et de la mesure de la performance en visant à **comprendre comment la performance de l'entreprise peut être affectée par les facteurs influents**, qui déterminent et expliquent les valeurs des KPI.

KPI comme objectifs

- La relation entre KPI et facteurs influents peut être utilisée pour l'aide à la décision.
- Par exemple, dans le cas de la relation client, on ne s'intéresse pas uniquement au **nombre** de clients qui disparaissent (**churn**) mais on veut comprendre **les raisons** de leur comportement.
- Pour cela, on doit prendre en compte les facteurs influents, dans les trois perspectives de l'entreprise.

KPI comme objectifs

- Dans la **perspective production**, on envisage les facteurs influents en lien avec la production de biens et de services.
- Dans la **perspective client**, les facteurs influents possibles sont souvent issus des caractéristiques socio-démographiques des clients ou des attributs impactant leur perception des produits ou des services offerts.
- Dans la **perspective organisationnelle**, on cherche les facteurs influents en lien avec l'organisation interne de l'entreprise ou l'influence des réseaux sociaux auxquels appartiennent les clients.

KPI comme objectifs

- Pour comprendre la relation entre KPI et facteurs influents, on utilise la **deuxième manière** de formuler les objectifs de la BI.
- Celle-ci est basée sur la notion d'**objectif analytique**.
- Cette formulation correspond à une **typologie** des approches possibles d'une analyse.

KPI comme objectifs

- On distingue usuellement trois types d'objectifs analytiques.
- **Les objectifs descriptifs** : ceux-ci apparaissent dans les trois perspectives BI et peuvent être basés sur les trois approches BI de l'activité de l'entreprise.
- L'objectif descriptif de base est **le reporting**, qui est souvent requis en vue d'autres objectifs analytiques.
- Deux autres objectifs descriptifs sont **la segmentation** et **la détection** de comportements atypiques.

KPI comme objectifs

- Le reporting résume les données de façon à ce qu'on puisse se baser dessus pour prendre des décisions.
- La segmentation (clustering) regroupe les données en classes en se basant sur une mesure de similarité et identifie des individus représentatifs dans les différentes classes.
- La détection de comportement atypique consiste à identifier des événements ayant lieu durant le déroulement d'un processus qui aident à mieux comprendre le processus.

KPI comme objectifs

- Ces deux derniers objectifs sont souvent regroupés sous le terme d'**apprentissage non supervisé**.
- **Les objectifs prédictifs** : ceux-ci sont plus ambitieux et sont particulièrement intéressant dans la perspective client en vue transversale.
- On distingue usuellement deux types de prédiction.

KPI comme objectifs

- **La régression** : trouver une fonction permettant de prévoir une sortie (usuellement un KPI) à partir de variables d'entrée (facteurs influents).
- **La classification** : Etant donnée une partition des données en différentes classes, affecter un nouvel individu à l'une des classes.
- On utilise le terme **apprentissage supervisé** pour désigner ces deux types de prédiction.

KPI comme objectifs

- **Les objectifs de compréhension** : ceux-ci sont encore plus ambitieux et sont souvent liés à la perspective production que ce soit pour la vue événements ou la vue états.
- On distingue usuellement deux types d'objectifs de compréhension.
- **L'identification de processus** : Identifier les règles qui déterminent les relations entre les différents événements d'un processus.

KPI comme objectifs

- **L'analyse des processus** : Etudier la performance des instances en les comparant à un processus type.
- Notons que le choix des objectifs est complémentaire de l'analyse par **cycle de vie** des processus en trois phases : **conception, exécution, évolution.**

KPI comme objectifs

- Lors de la **phase de conception**, le processus est décrit en termes d'un **modèle de processus**, défini en accord avec un KPI, pour répondre à un besoin de l'entreprise.
- Souvent, la phase de conception elle-même repose sur un **Business Plan** et une formulation détaillée du projet requiert la prise en charge de différents objectifs analytiques.
- Par exemple, si on souhaite lancer un nouveau produit, divers objectifs analytiques tels que la prédiction du marché (objectif prédictif) et une description détaillée du produit (objectif descriptif) doivent être atteints.

KPI comme objectifs

- L'analyse d'un processus lors de **la phase d'exécution** est basée sur des instances du processus. Dans ce cas également, on doit formuler précisément les objectifs analytiques.
- De même, l'analyse d'un processus lors de **la phase d'évolution** est souvent basée sur la définition précise d'un KPI.

Activités de la BI

- Atteindre les différents objectifs analytiques de la BI implique un certain nombre **d'activités** et une certaine organisation de ces activités.
- Dans la suite, on va décrire les différentes activités de la BI.

Activités de la BI

- **L'activité data** : elle est à la base de toutes les autres activités BI. L'objectif principal est d'organiser l'information sur l'entreprise et son environnement.
- Typiquement, l'information consiste en des **données sur les propriétés structurelles** de l'entreprise et des clients, des **données transactionnelles** sur des instances de procédés, de données **décrivant les activités de production** et de **traces** de ces activités dans les réseaux sociaux.

Activités de la BI

- Ces données sont collectées selon différentes méthodologies et sont stockées dans différentes **sources de données**. Elles peuvent être **structurées temporellement** ou **spatialement** ou être **semi-structurées** s'il s'agit de données textuelles.
- Le principal défi est **d'organiser les données** de façon à ce qu'elles puissent être utilisées dans diverses activités de BI.

Activités de la BI

- On part souvent d'une organisation pré-existante des données au sein d'un **entrepôt de données** (ang. data warehouse) qui offre des données cohérentes et de qualité, ce qui autorise diverses activités de BI, en particulier le **reporting**.
- Dans cette optique, on va parfois jusqu'à présenter la BI essentiellement comme une activité de **récupération** et de **modélisation** des données.
- Toutefois, en raison de l'évolution de l'entreprise et de son environnement au fil du temps, même l'entrepôt de données le mieux conçu ne peut répondre à toutes les questions.

Activités de la BI

- L'aide à la décision pour de nouveaux problèmes peut requérir une **réorganisation des données** ou de collecter de **nouvelles données** spécifiques.
- Par conséquent, il est nécessaire de maîtriser les méthodes de collection des données et d'augmentation des données existantes pour mener à bien l'activité data.

Activités de la BI

- L'activité data est basé sur des **techniques de modélisation de l'information** qui comprennent divers modèles tels que le modèle **ER (entités-relations)**, **UML (Unified Modeling Langage)** et les modèles semi-structurés.
- Celles-ci incluent également diverses façons d'utiliser ces modèles et prennent en compte l'infrastructure IT qui fournit les données.

Activités de la BI

- **L'activité de compréhension de l'entreprise et des données** : le point de départ de celle-ci est la formulation d'objectifs analytiques, définis à l'aide de KPI dans le meilleur des cas.
- Elle envisage l'entreprise dans la perspective de ces objectifs et identifie les parties de l'entreprise qui ont un impact sur ceux-ci, ainsi que les données qui sont pertinentes pour les réaliser.

Activités de la BI

- Les résultats de cette activité sont une **formulation d'objectifs analytiques, l'identification des activités de l'entreprise pertinentes** pour celui-ci et des **données requises** pour l'atteindre.
- De plus, une première version de l'organisation du travail pour les autres activités de la BI est proposée à l'issue de cette activité.

Activités de la BI

- L'activité de **compréhension de l'entreprise et des données** a lieu selon les axes suivants.
- **L'environnement** : sur cet axe, on s'attache à explorer la taille et la portée des objectifs analytiques au sein de l'entreprise et on identifie le **scénario BI** pertinent. On détermine également les **ressources** et **l'horizon temporel** pour atteindre les objectifs.

Activités de la BI

- **La perspective** : sur cet axe, on envisage les objectifs analytiques selon les différentes perspectives de l'entreprise (production, clients, organisation), ce qui inclut l'identification de responsables de processus, de sujets de processus et d'acteurs.
- **La vue** : on choisit la vue adaptée (par événements, par états ou transversale) et on sélectionne les données pertinentes pour l'analyse.

Activités de la BI

- **Les objectifs analytiques** : on définit de façon précise les **KPI** pertinents et les **facteurs influents**, ainsi que les **types** d'objectifs analytiques (segmentation, régression, classification, etc).
- **L'évaluation des données** : il s'agit de l'examen des données en termes de **caractéristiques des variables** et de **qualité**. Des **transformations des données** peuvent être requises afin de mettre celles-ci sous une forme exploitable par l'analyse.

Activités de la BI

- Le choix de la **vue** dépend non seulement des objectifs à atteindre mais également de la **disponibilité** des données.
- Il est souvent relativement aisé de disposer de données **transversales**, alors même que des données détaillées **par événements** sur les instances de processus ne sont accessibles que **partiellement**.
- Dans la situation où les données **ne sont pas disponibles** pour la vue souhaitée, il faut décider si l'utilisation des données existantes permet d'atteindre l'objectif.

Activités de la BI

- Dans le pire des cas, l'indisponibilité des données est un **facteur limitatif** et il faut envisager **une nouvelle collecte de données**.
- **A l'issue** de l'activité de compréhension de l'entreprise et des données, **on donne une formulation précise des objectifs de l'entreprise en termes d'objectifs analytiques**. Ceci requiert une combinaison de connaissance du domaine, c'est-à-dire de **compréhension de l'entreprise** et de connaissance des propriétés des données, c'est-à-dire de **compréhension des données**.

Activités de la BI

- **L'activité de modélisation** : celle-ci vise à construire **un modèle analytique de l'entreprise**, c'est-à-dire un modèle formel qui permet de répondre de façon précise aux différents objectifs analytiques.
- Selon les perspectives, vues et objectifs retenus, on construit des modèles différents, qui permettent de traduire les objectifs analytiques en **questions formelles** à propos des propriétés de ceux-ci.
- Parfois, le modèle retenu est très simple et ne consiste qu'en un résultat de requête sur les données disponibles.
- D'autres fois, choisir un modèle peut s'avérer complexe.

Activités de la BI

- Considérons par exemple la formulation d'un **modèle d'analyse d'un KPI** en lien avec l'acquisition de clients.
- On peut partir d'une **approche orientée marketing**, adopter la **perspective client** et identifier les facteurs qui attirent de nouveaux clients. Après l'identification de ces facteurs, on peut faire en sorte que les processus internes à l'entreprise répondent aux attentes des clients.

Activités de la BI

- Une autre approche consiste à partir de la perspective production, examiner le processus de production existant et développer des scénarios possibles d'évolution de celui-ci. Ensuite les différents scénarios sont analysés en termes d'appétence des clients.
- Cet exemple montre que **différentes formulations des modèles sont possibles**, selon la perspective et les objectifs analytiques adoptés.

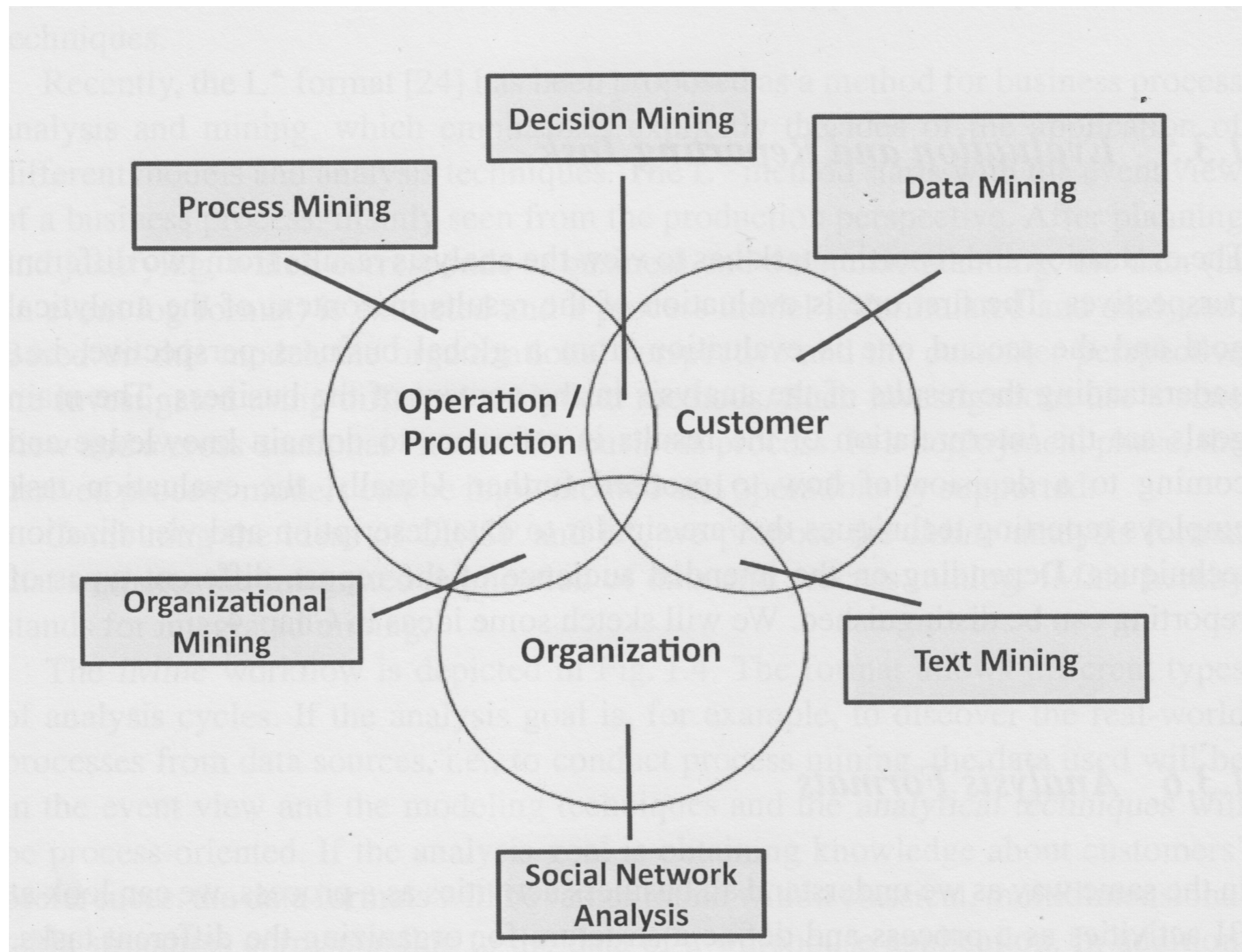
Activités de la BI

- **L'activité d'analyse** : Ayant défini un modèle, on a recours à des **algorithmes** pour identifier une **solution** répondant aux objectifs analytiques posés.
- En BI, ces algorithmes sont généralement qualifiés **d'algorithmes de « mining »**, afin de mettre l'accent sur le fait que la détermination d'une solution est rarement un problème bien défini.

Activités de la BI

- Selon la ou les **perspectives** BI retenues, on aura recours à des **algorithmes différents** pour trouver une solution.
- Le graphique suivant présente la **terminologie** utilisée en BI pour désigner ces algorithmes (en anglais). *Il est issu de la première référence donnée en bibliographie.*

Activités de la BI



Activités de la BI

- **L'activité d'évaluation et de reporting** : Cette activité envisage les résultats de l'activité d'analyse selon deux perspectives différentes. La première consiste à évaluer les résultats dans la perspective des objectifs analytiques. La deuxième consiste à les évaluer dans la perspective de l'entreprise dans sa globalité.
- A cet effet, l'activité d'évaluation fait souvent appel à des techniques de reporting semblables à celles utilisées en description et visualisation des données (tableaux de bord, notamment).

Formats d'Analyse

- Puisqu'on considère toutes les activités d'une entreprise comme faisant partie de processus, **on peut considérer la BI comme un processus à part entière** et définir **une structure pour organiser les différentes activités de la BI.**
- De telles structures sont appelées **formats d'analyse** et sont basées sur des idées issues de la théorie des **cycles de développement** logiciel et de **l'extraction de connaissances** à partir de bases de données.

Formats d'Analyse

- On distingue usuellement entre **format linéaire** et **format itératif**.
- **Les formats itératifs sont les plus utiles en BI** car, en pratique, couvrir les différentes perspectives de la BI requiert souvent une suite de modèles et une évaluation combinée des résultats.
- Un autre argument en faveur de l'approche itérative est qu'une première analyse détecte souvent des caractéristiques nouvelles et inattendues qui requièrent davantage d'analyse.

Formats d'Analyse

- Un format d'analyse très répandu est la **méthode Kimball** (2010). Elle met l'accent sur **l'activité data** et part de la définition des objectifs d'une analyse et de la compréhension du contexte de l'entreprise. Elle se concentre essentiellement sur **l'activité de déploiement**, c'est-à-dire sur l'intégration des résultats des activités BI à un entrepôt de données.

Formats d'Analyse

- Divers autres formats d'analyse ont été proposés pour le processus d'extraction de connaissances à partir des données sous le nom de **formats de data mining** (ang. data mining formats).
- Une approche orientée vers les applications est **CRISP-DM**.
- C'est aujourd'hui un **standard** dans les applications de data mining.
- D'autres formats sont **SEMMA** et **KDD**.

Formats d'Analyse

- **CRISP-DM** se concentre essentiellement sur **la vue transversale** d'un processus, mais il est possible de l'adapter à la vue par états.
- Notre définition des **activités BI** est fortement influencée par **CRISP-DM**.
- Une différence essentielle est que nous faisons un lien plus étroit entre **compréhension des données** et **compréhension de l'entreprise**.
- De plus, notre approche de l'activité de modélisation et d'analyse autorise à prendre en compte **la vue par événements** et les applications de techniques de **process mining**.

Formats d'Analyse

- Récemment, le format **L*** d'analyse et de mining **des processus** a été proposé.
- Il met explicitement l'accent sur l'application de différents modèles et différentes techniques d'analyse.
- La méthode **L*** part de **la vue par événements** d'un processus, essentiellement dans la **perspective production**.
- Après **la planification** et **la justification**, qui correspondent à la compréhension de l'entreprise et des données, les données (sous la forme d'un journal d'événements) sont obtenues et un **modèle de processus** est formulé et analysé.

Formats d'Analyse

- En se basant sur ce modèle, les **perspectives client et organisationnelle** sont envisagées à l'aide de méthodes analytiques diverses.
- Pour ce faire, on peut faire appel à **la vue par états** et à la **vue transversale** du processus.
- Dans la **phase de déploiement**, le modèle de processus obtenu peut être implémenté et utilisé dans la pratique.

Formats d'Analyse

- Le format d'analyse **iMine** (Integrated Mining) combine les idées de **CRISP-DM** et de **L***.
- Il prend en compte simultanément **data mining** et **process mining**.
- Pour des détails sur **CRISP-DM**, on pourra se référer aux pages wikipedia correspondantes.
- Les autres formats d'analyse sont moins répandus et plus spécifiques.