



UNIVERSITÉ
DE REIMS
CHAMPAGNE-ARDENNE

Business Intelligence

Nils SCHAEFER

nils.schaefer@snicw.fr

Introduction à la BI

□ Constatation

- Le chiffre d'affaires d'une société a baissé en avril

□ Pour prendre les bonnes décisions, il faut s'interroger...

- Pourquoi le CA a-t-il baissé ?
- Est-ce une baisse habituelle ?
- Quelles sont les gammes de produits concernées ?
- Quels sont les pays concernés ?
- Quels sont les commerciaux concernés ?
- ...

Business Intelligence

❑ En français : Informatique Décisionnelle

❑ Définition

- Terme générique qui englobe les applications, l'infrastructure, les outils et les meilleures pratiques permettant l'accès et l'analyse de l'information afin d'améliorer et d'optimiser les décisions et les performances

Pourquoi construire un système décisionnel ?

- ❑ Les **données** collectées et générées chaque jour constituent un actif de valeur pour les sociétés
- ❑ La plupart du temps, ces informations sont **inaccessibles** ou **réparties** dans différents systèmes
- ❑ Dans une société, on a besoin...
 - D'avoir une vision globale homogène et cohérente des informations manipulées
 - D'accéder directement à l'information nécessaire
 - De donner un sens unique aux données de gestion
 - Calculs de marge, de CA...

Pourquoi construire un système décisionnel ?

- ❑ Le Système d'Information Décisionnel (SID) résultat d'un processus qui consiste à...
 - **Extraire les données** à partir des systèmes opérationnels et de sources externes de données
 - **Transformer les données** en informations de pilotage
 - **Diffuser les données** pour les rendre accessibles aux utilisateurs
- ❑ La Base Décisionnelle est reconnue comme un **actif stratégique** par de nombreuses sociétés

Métriques d'aide à la décision

- ❑ Pour prendre une décision, il est nécessaire d'avoir une vision précise des choses
- ❑ Différentes métriques d'aide à la décision avec différentes importances sur la prise de décision des différents acteurs
 - Donnée
 - Information
 - Connaissance

❑ Valeur élémentaire à faible valeur ajoutée

- Présente dans une base de données interne
- Collectée auprès d'un système externe
- Générée par une action

❑ Une donnée ne permet pas de prendre de décision sur une action à lancer

❑ Exemple

- Le produit « Tarte aux pommes » est passé en caisse du supermarché MySuperMarket de Reims le 5 avril 2023 à 11:23:15

- ❑ Donnée (ou ensemble de données) à laquelle on a donné un sens et une interprétation
- ❑ Une information permet de prendre une décision
- ❑ Exemple
 - Le produit « Tarte aux pommes » a été jeté à la poubelle, faute de vente, à 20 reprises entre le 29/05/2023 et le 04/06/2023
 - Le responsable décide de réduire la production de ce produit

- ❑ Résultat d'une réflexion sur les informations analysées
 - Information enrichie, analytique et pertinente
- ❑ Se base sur les expériences, les idées et les avis des personnes consultées
- ❑ Exemple
 - Le produit « Tarte aux pommes » est le produit qui présente le plus d'invendus devant être jetés
 - Le responsable décide d'annuler la production de ce produit

Risques d'un projet décisionnel

- ❑ Orientation technologique du projet
- ❑ Mise en cause de la fiabilité et cohérence des informations
- ❑ Alimentations trop longues et irrégulières
- ❑ Outils et architecture technique inadaptés
- ❑ Fraicheur insuffisante des informations
- ❑ Pas d'administrateur du système
- ❑ Surenchère fonctionnelle concernant les analyses et les outils à utiliser par rapport aux réels enjeux métiers

Facteurs clés de succès

□ Pour la conduite de projet

- Adopter une démarche incrémentale avec des objectifs précis et quantifiés à l'avance
- Miser dès le départ sur un socle informatique stable, puis coordonner des itérations courtes avec des résultats tangibles

□ Dans la conception fonctionnelle

- Préférer des états figés mais utiles à des analyses libres mais compliquées
- Raisonner en flux amont-aval et en échanges transverses entre utilisateurs

Facteurs clés de succès

❑ Pour la mise en œuvre technique

- S'assurer de l'évolutivité de la solution technique
- Privilégier les solutions paramétrables par un administrateur fonctionnel
- Insérer dans le recette technique des tests de montée en charge

❑ Dans la prise en compte des impacts organisationnels

- Préparer le changement et l'insérer dans le plan de projet
- Fonder et associer le plus tôt possible le futur administrateur du système

Chaîne décisionnelle

- ❑ Chaîne de traitement de l'information
- ❑ Elle permet de transformer les données collectées en informations pouvant être utilisées à des fins décisionnelles
- ❑ 5 grandes étapes
 - Planification
 - ETL
 - Stockage
 - Analyses
 - Restitution

Planification

- ❑ Pour mettre en place une plateforme décisionnelle d'entreprise intégrée, il faut commencer par planifier le projet
- ❑ Ce type de projet nécessite une administration solide
 - Chef de projet robuste (expérience, poids...)
 - Implication de toutes les parties

- ❑ Extract / Transform / Load
- ❑ **Extraction des données** à partir d'une ou plusieurs sources de données
 - BD, Excel, CSV, API...
- ❑ **Transformation des données**
- ❑ **Chargement des données** dans la banque de données de destination
- ❑ Phase importante car elle conditionne et influence la qualité du projet

Stockage

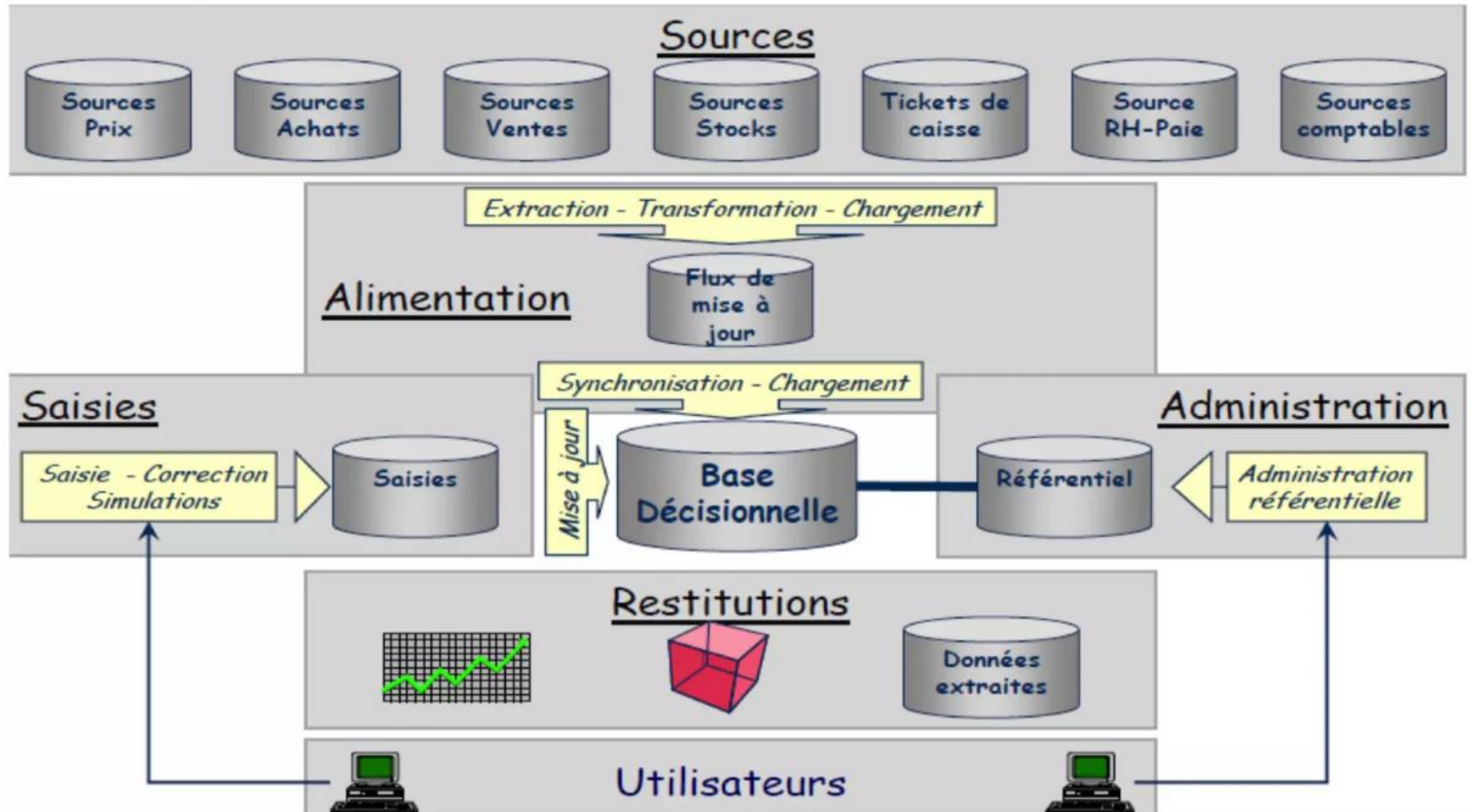
- ❑ Plusieurs manière de stocker des données dans un entrepôt de données
- ❑ Chaque solution ayant ses avantages et ses inconvénients
- ❑ L'administrateur des bases de données décisionnelles pourra notamment choisir entre plusieurs modèles

- ❑ Regroupement de l'ensemble des techniques de statistique, d'économétrie, de Data Mining et de recherche opérationnelle
- ❑ Demande souvent des compétences statistiques avancées
- ❑ Certaines solutions embarquent ces fonctionnalités pré-paramétrées à des cas de figures bien définis, afin d'améliorer la valeur ajoutée de leurs produits

Restitution

- ❑ La génération de tableaux de bord est l'aspect le plus visuel pour les utilisateurs
- ❑ Les utilisateurs disposent d'interface intuitives permettant, en fonction de ses droits, de consulter des rapports, des tableaux de bord mais aussi de les annoter voire de les créer

Architecture d'un système décisionnel



L'alimentation

- ❑ L'alimentation est la partie la plus importante et la plus couteuse
- ❑ Sur le plan des données
 - Définir un langage commun
 - Localiser les données utiles dans les systèmes sources
 - Harmoniser les nomenclatures
- ❑ Sur le plan technique
 - Les règles d'alimentation changent sans cesse
 - Développements modulaires auto-documentés, traçabilité des données
 - Temps de chargement compatibles avec la fenêtre d'exploitation

Web scraping

Besoin en données

- ❑ Nombreux contextes dans lesquels on a besoin de recueillir des données
- ❑ Plusieurs possibilités structurées
 - Base de données externes
 - Flux de données
 - API
 - XML
 - CSV
 - ...
- ❑ Parfois aucune solution n'existe
 - On va donc devoir aller chercher les données

Collecte des données

- ❑ Toutes données visibles peuvent généralement être récupérées « à la main »
 - Copier/coller (Texte)
 - Enregistrement (Documents, vidéos, images...)
- ❑ Traitement manuel long et peu productif
- ❑ Besoin d'automatiser la collecte
 - Surtout lorsque le volume de données est important
 - Surtout lorsque les données doivent être collectées très régulièrement
- ❑ Enchaînement avec d'autres traitements
 - Envois d'emails
 - Génération de documents
 - ...

Web scraper

- ❑ Outil automatisé qui parcourt des listes définies de sites Web pour en extraire le contenu
 - Les sites sont connus et ciblés
 - Le contenu à récupérer est ciblé
- ❑ A ne pas confondre avec un Web crawler qui repère les liens et indexe le contenu
 - Utilisé par les moteurs de recherche
 - Les sites ne sont pas connus
 - Le contenu indexé n'est pas connu

❑ Evolution du site Web

- Web scraper qui ne fonctionne plus
- Mise à jour nécessaire
- Interruption temporaire dans la collecte

❑ Mécanismes de blocage

- CAPTCHA
- Honeypots (faux liens invisibles)
- Mesure de bande passante

❑ Obstacles techniques

- Authentification
- Géolocalisation

Techniques de Web scraping

❑ Une première approche efficace

❑ Analyse du contenu

- Récupération du contenu
 - Code HTML
 - Document Word/Excel/Powerpoint
 - Image
 - Vidéo
 - ...
- Manipulation du contenu
 - Expressions régulières (texte brut, HTML)
 - Parser (HTML, XML, JSON...)
 - Bibliothèques spécifiques (Word/Excel/PowerPoint, Image...)

Difficultés techniques

❑ Nombreuses pages dynamiques

- Structure de la page qui change
- Aucun rechargement global de la page
- Différentes actions nécessaires
 - Scroll, mouvements de la souris, clics...

❑ Besoin d'outils avancés

- Google Chrome (Headless) en mode debug
 - GoogleChrome.exe
 - remote-debugging-address=0.0.0.0
 - remote-debugging-port=9876
- Mink avec PHP

Services permettant de récupérer des données classiques

- Sur des sites connus : Le Bon Coin, Pages Jaunes
- Le service s'occupe de garantir le bon fonctionnement de la collecte de données au cours du temps

Services permettant de construire un scénario spécifique

- Sur n'importe quel site
- L'utilisateur doit lui-même mettre à jour ses processus de collecte des données

SI opérationnel

Processus métier

- ❑ Flux de matériel, d'informations, de connaissances
- ❑ Ensemble d'activités et d'étapes
- ❑ Peut être lié à un domaine fonctionnel ou être interfonctionnel
- ❑ Vision de l'entreprise
 - Ensemble de processus métier

Exemples de processus fonctionnels

Fabrication
et
production

Assembler
un produit

RH

Recruter

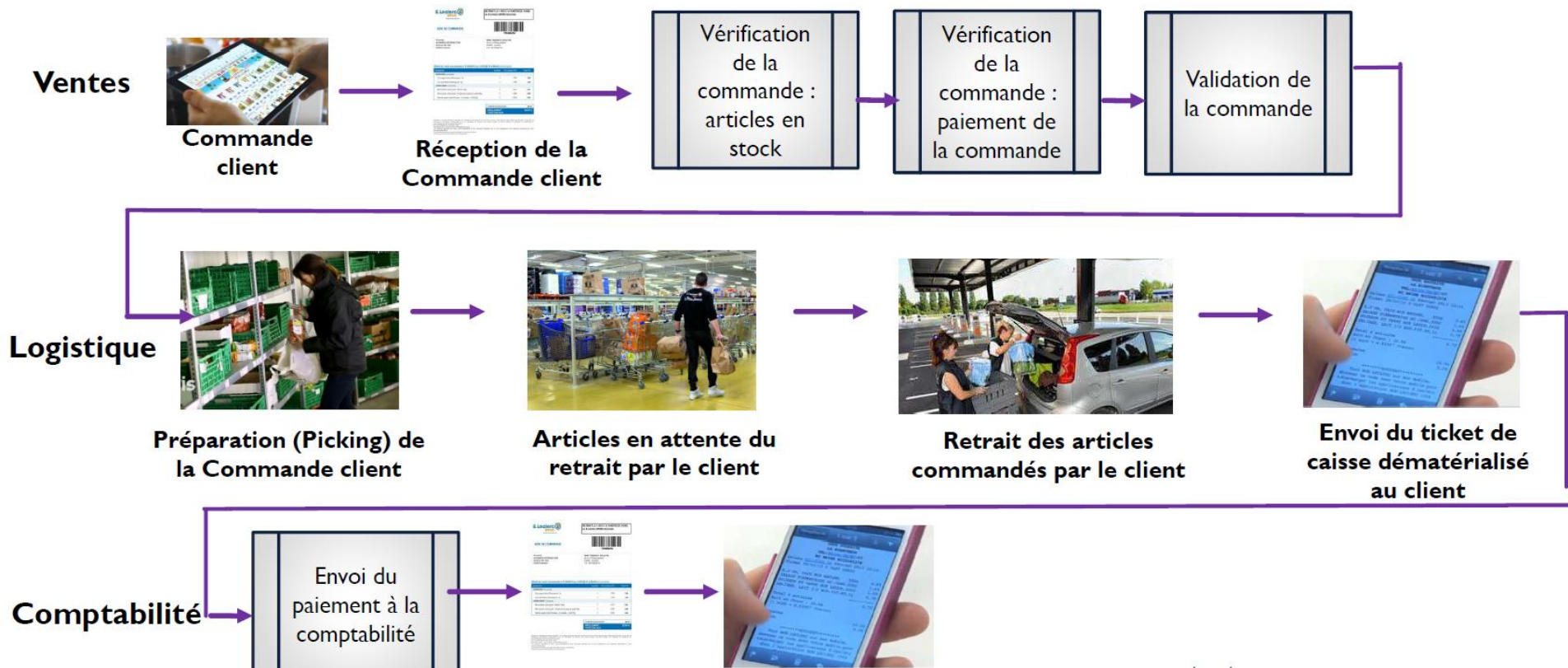
Ventes &
Marketing

Identifier
les clients

Comptabilité
& Finance

Créer les
états
financiers

Exemple d'un processus interfonctionnel



Technologies de l'information et processus

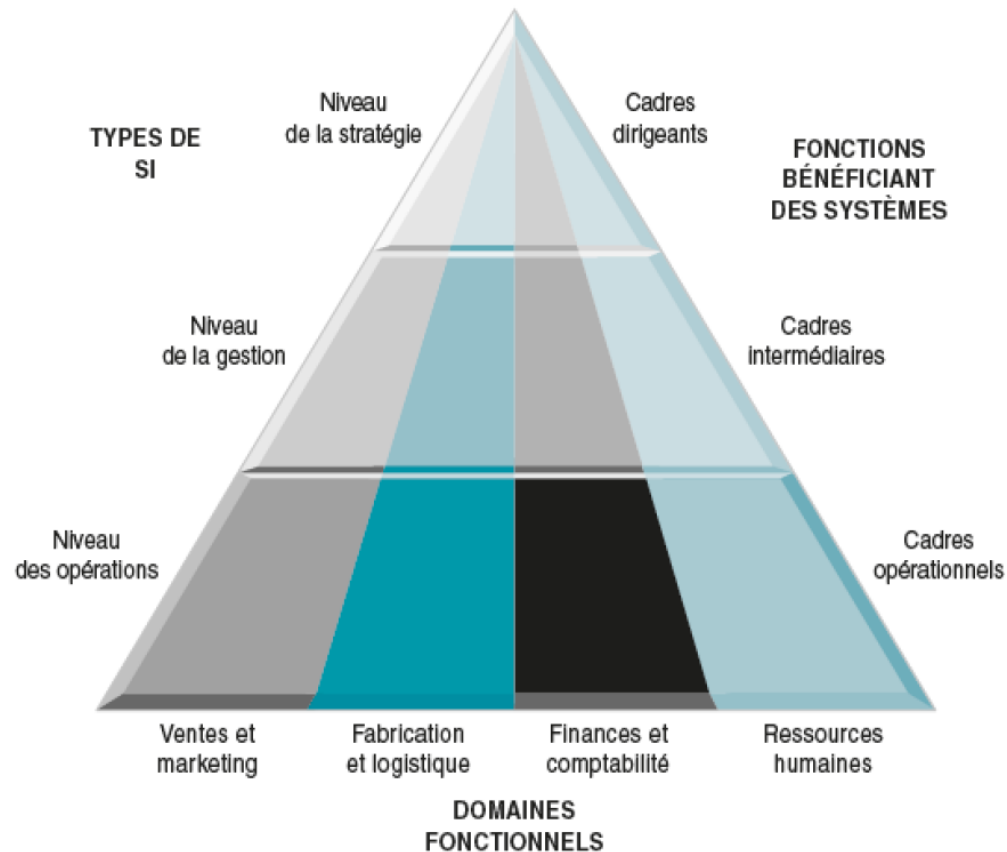
- ❑ Accroître l'efficacité des processus existants
 - Automatisation des étapes manuelles
- ❑ Permettre des processus entièrement nouveaux
 - Modifier le flux d'informations
 - Remplacer les étapes séquentielles par des étapes parallèles
 - Éliminer les délais dans la prise de décision
 - Soutenir les nouveaux modèles d'affaires
- ❑ Les technologies de l'information améliorent les processus métier

SI et processus

- ❑ Les processus métiers sont généralement regroupés en ensembles homogènes appelés fonctions
- ❑ Les SI prenant en charge ces fonctions sont appelés SI fonctionnels
- ❑ Les SI fonctionnels reposent sur des logiciels qui traitent et stockent les transactions quotidiennes de l'entreprise
- ❑ L'intégration des processus est pilotée par l'entreprise
- ❑ Exemples
 - SI financier (logiciels dédiés)
 - SI logistique (logiciels dédiés)
 - SI CRM (logiciels dédiés)
 - ...

Types de SI

- ❑ Interaction avec les SI différentes en fonction du domaine et du niveau hiérarchique



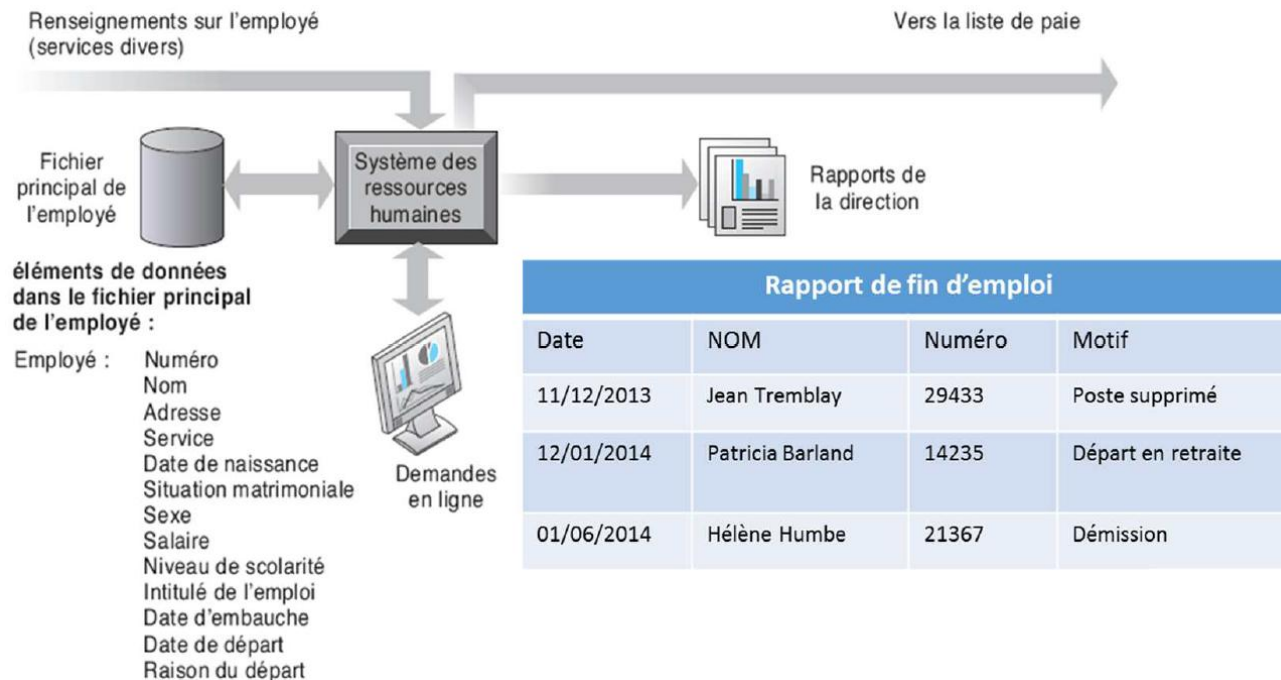
Systemes de Traitement des Transactions (STT)

- ❑ Servir les managers opérationnels
- ❑ Effectuer et enregistrer les transactions de routine quotidienne nécessaires à la conduite des affaires
- ❑ Exemples
 - saisie de commande client, paie, expédition
- ❑ Permettre aux gestionnaires de contrôler la performance des opérations
- ❑ Servir des objectifs et une prise de décision prédéfinis et structurés

Systèmes de Traitement des Transactions (STT)

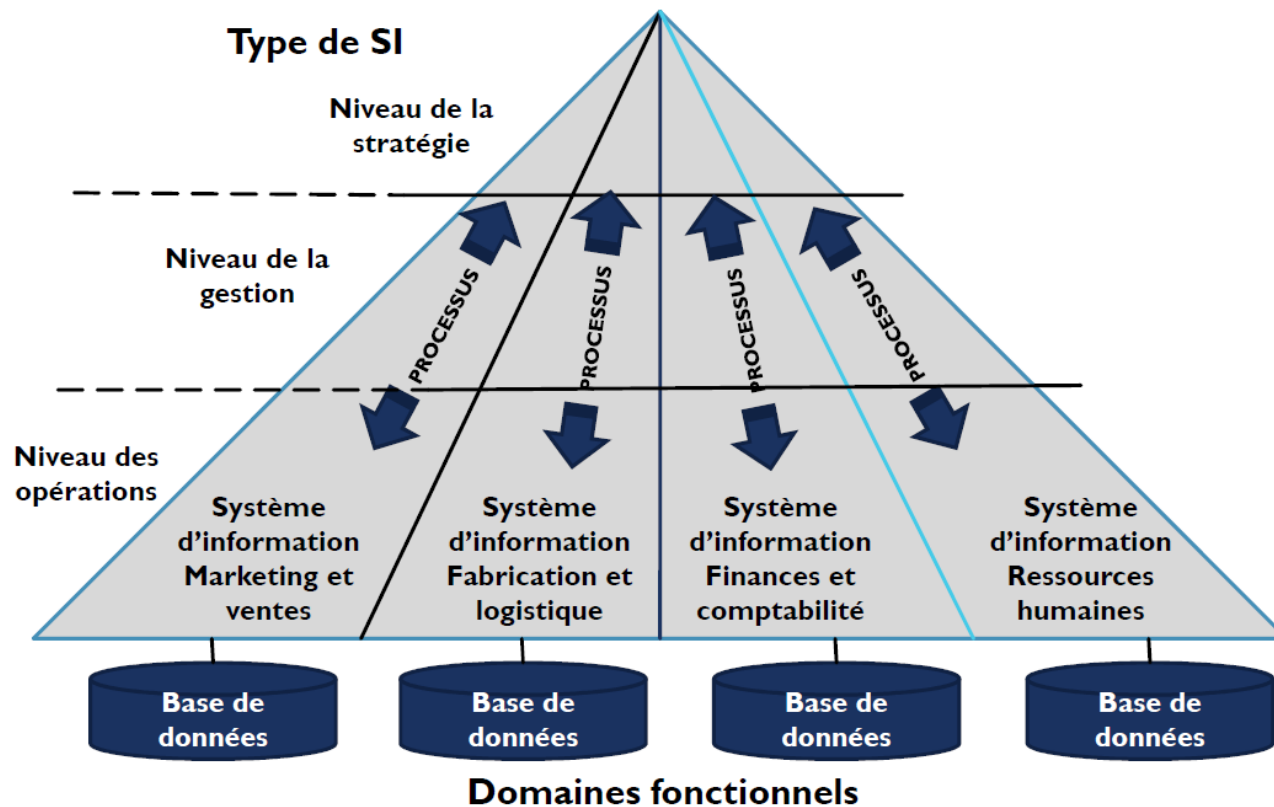
□ Traitement de la paie

- SI au service d'un domaine fonctionnel basé sur des données fiables traitées par des processus métier



SI fonctionnel

- Basé sur des progiciels d'application dédié ou des logiciels d'application spécifiques

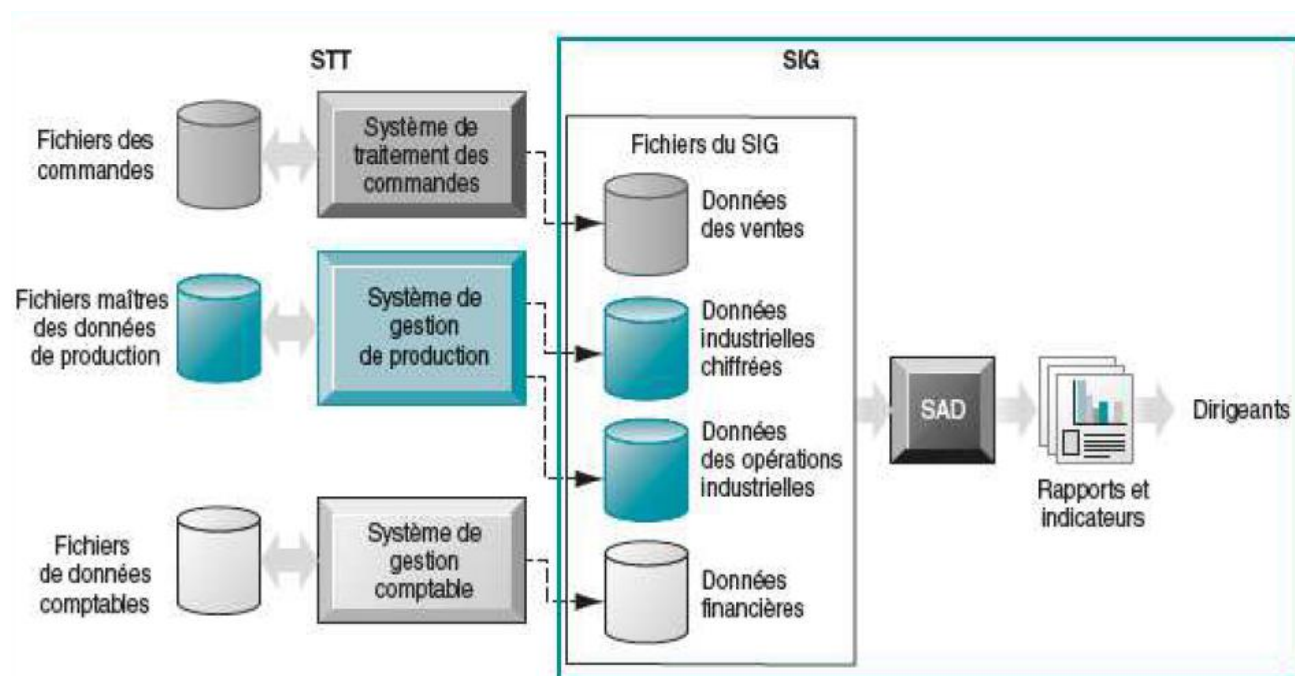


Systèmes d'Information de Gestion (SIG)

- Servir les cadres intermédiaires
- Fournir des rapports routiniers sur la performance sur la base des résumés de transactions
- Fournir des indicateurs à court terme et des réponses à des questions routinières
- Exige peu d'analyse

Systemes d'Information de Gestion (SIG)

- ❑ Les managers ont accès aux données via les rapports fournis par le SIG



Systemes d'Information des Dirigeants (SID)

- Servir les cadres dirigeants
- Aide à la décision non routinière
- Exige des capacités d'analyse et de jugement avancées
- Inclut des informations provenant de l'extérieur de l'entreprise et des données résumées provenant des SIG

Intégration des informations

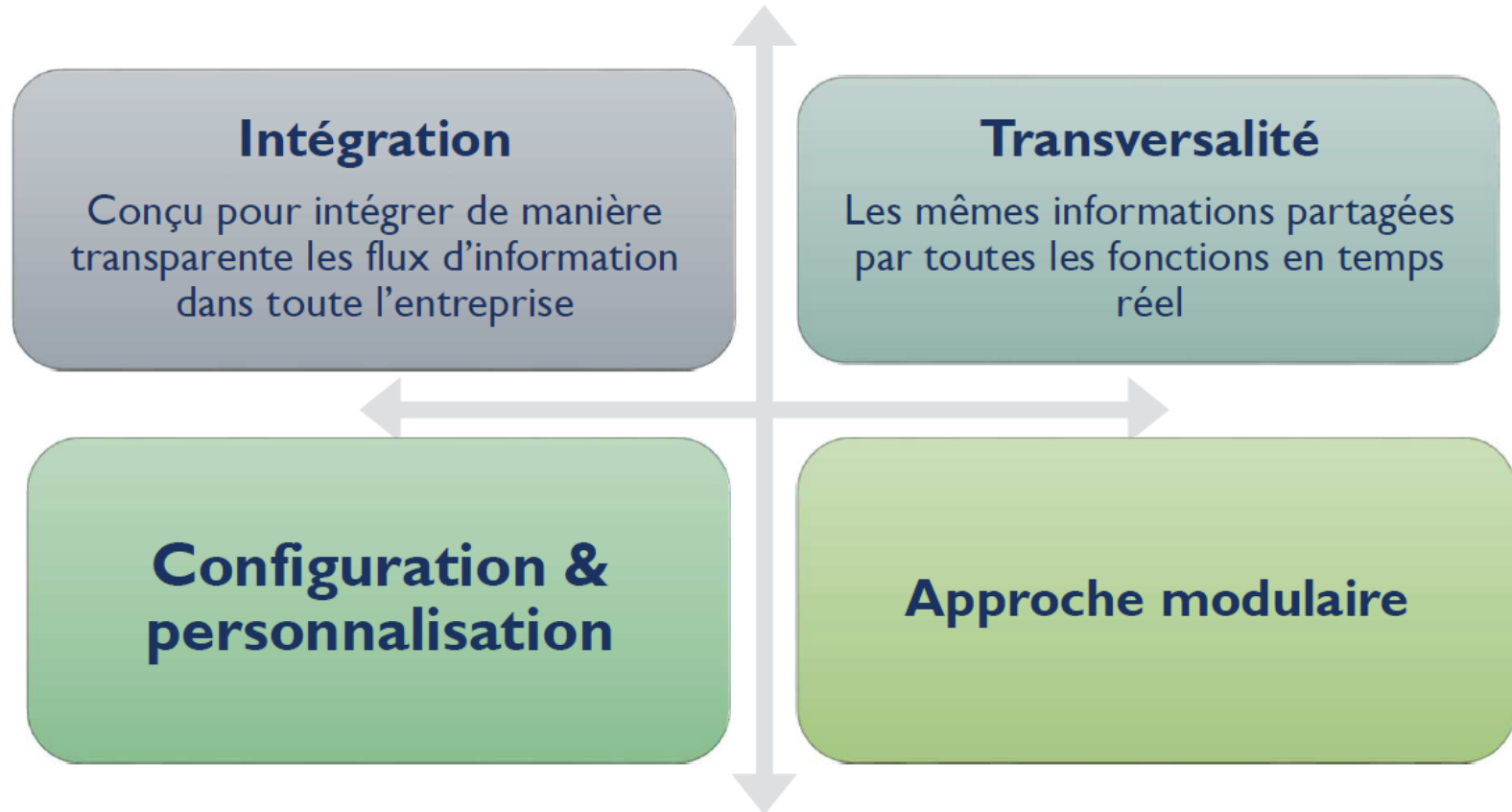
- ❑ Un défi organisationnel
- ❑ Aspect informationnel des processus
 - Une multitude de données
 - De l'information partagée entre différents processus
 - Besoin d'obtenir la bonne information au bon moment et au bon endroit
- ❑ Besoin d'intégration d'information
- ❑ Assurer une transmission fluide des informations entre tous les acteurs et les processus

Entreprise Resource Planning (ERP)

❑ Progiciel de Gestion Intégré (PGI)

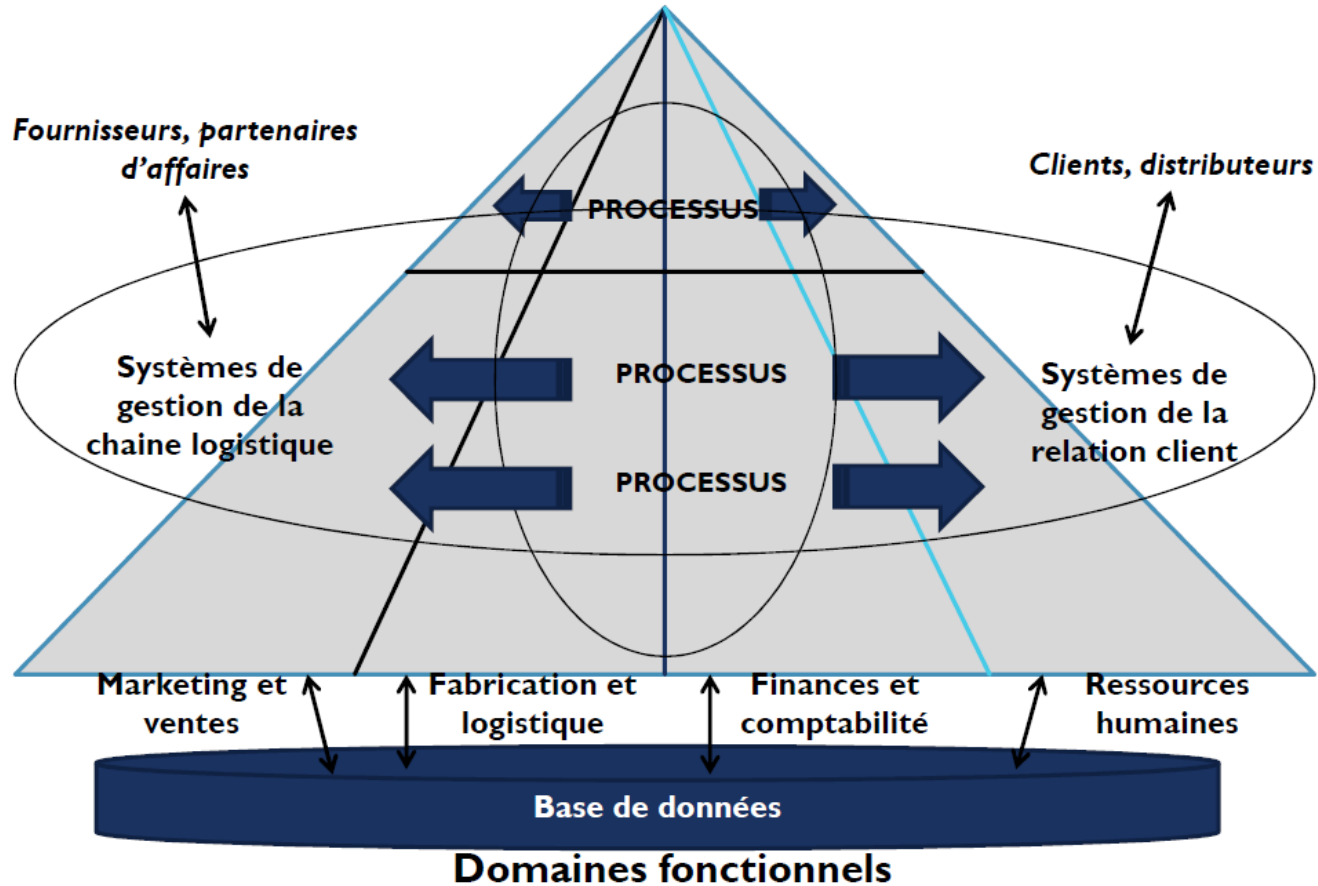
- ❑ Un système ERP est un type de logiciel d'application que les entreprises utilisent pour gérer leurs **activités quotidiennes** telles que la gestion commerciale, la comptabilité, la production, l'approvisionnement, la gestion des ressources humaines, et les opérations de logistique.
- ❑ Les systèmes ERP relient une multitude de processus métier et activent le **flux de données** entre eux. En collectant les données transactionnelles partagées **en temps réel** d'une entreprise à partir de sources multiples, les systèmes ERP éliminent la duplication des données et assurent l'intégrité des données avec une **source unique d'informations**.

Caractéristiques d'un ERP



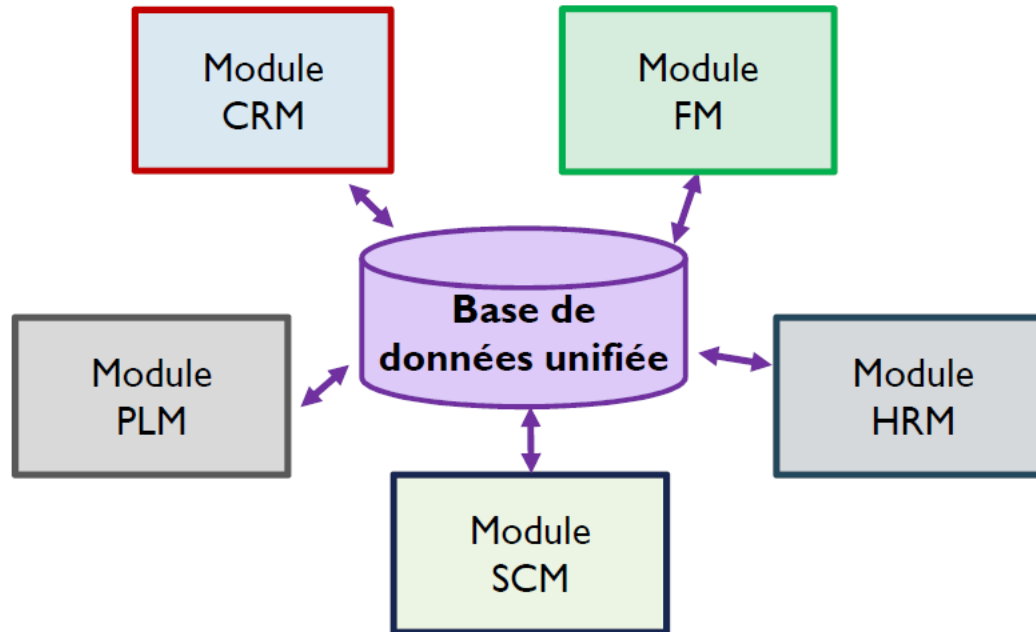
Caractéristiques d'un ERP

□ Vision intégratrice des processus



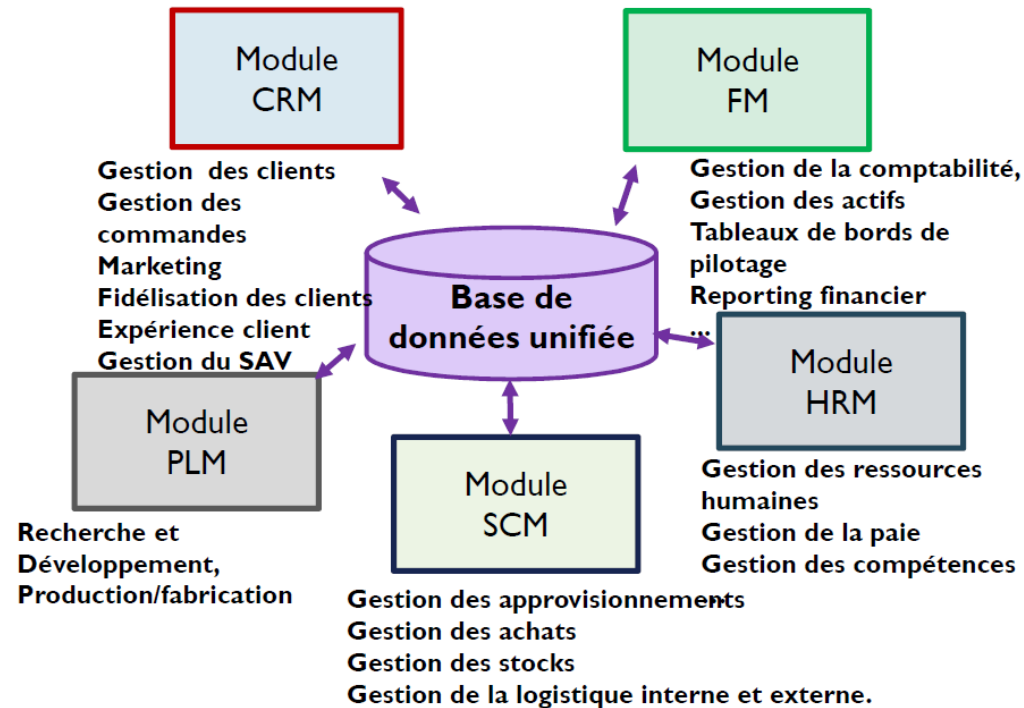
Caractéristiques d'un ERP

□ Une base de données unifiée



Caractéristiques d'un ERP

□ Une structure modulaire



Valeur business des ERP

❑ Création de valeur

- Données homogènes, centralisées et accessibles
- Communication optimisée – meilleure coordination
- Cycles d'opérations et de production plus courts

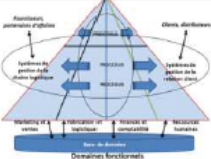


❑ Meilleur management

- Planification des activités
- Meilleure réactivité

❑ Croissance de l'entreprise

- Ajustements de la production et de l'offre
- Améliorer la prise de décision
- Avantage compétitif

Avantages et inconvénients

Progiciel de gestion intégré - ERP	Progiciel d'application dédié	Logiciel d'application spécifique
		
Prêt à implanter	Prêt à utiliser	Sur mesure
<p style="text-align: center;">Avantages</p> <ul style="list-style-type: none"> - Unicité - Paramétrable - Transversalité - Intégration technologique - Couverture fonctionnelle - Contrainte exogène et normes 	<p style="text-align: center;">Avantages</p> <ul style="list-style-type: none"> - Couverture des besoins intra-domaine + ou – adaptée aux besoins de l'entreprise - Coût d'acquisition du progiciel 	<p style="text-align: center;">Avantages</p> <ul style="list-style-type: none"> - Couverture des besoins intra-domaine - Potentiel d'évolution autonome
<p style="text-align: center;">Inconvénients</p> <ul style="list-style-type: none"> - Évolutions disjointes - Contraintes endogènes - Coût d'implantation 	<p style="text-align: center;">Inconvénients</p> <ul style="list-style-type: none"> - Passerelles pr transfert inter-applications : les interfaces 	<p style="text-align: center;">Inconvénients</p> <ul style="list-style-type: none"> - Passerelles pr transfert inter-applications : les interfaces - Coût de conception et de réalisation