

## Exploitez la richesse de vos données!



**Ralph Tonezzer**  
Associé AWK Group

Les données nous fournissent des informations qui, dans notre société de l'information, existent en abondance. On a vite fait de s'y perdre et de laisser de côté les informations importantes permettant de fournir savoir et sagesse.

Dans beaucoup d'entreprises, ce ne sont pas les outils et les méthodes qui manquent pour tirer des données commerciales plus d'informations et une meilleure image de la réalité.

Dans les projets, nous constatons souvent des lacunes dans les méthodes d'analyse de ces données ou que les outils existants ne sont pas exploités de manière optimale.

Cet AWK Focus vise à vous donner un aperçu des méthodes et des outils servant à l'analyse des données et, par des exemples d'utilisation, à vous inciter à exploiter également dans votre entreprise le potentiel existant.

Cordialement vôtre

## Mine your data – Pour une meilleure compréhension de l'entreprise

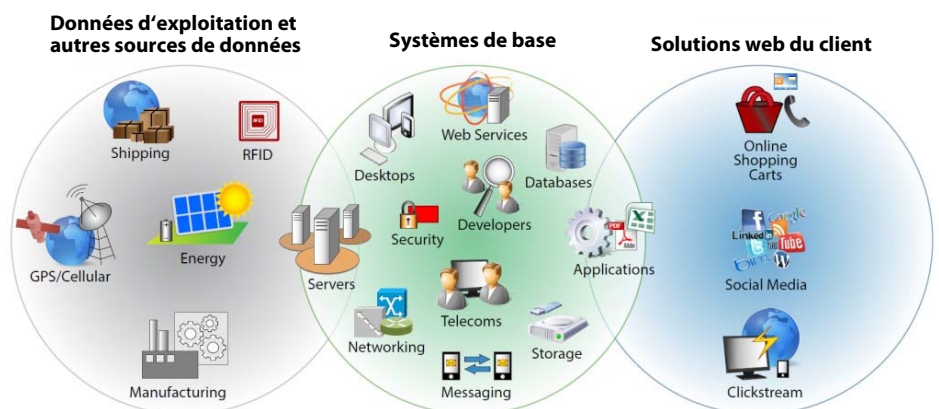
**De nos jours, presque tous les processus importants sont assistés par des systèmes IT et gérés électroniquement – tout au moins au niveau des données. Et pourtant, on recourt trop peu à ces données dans le cadre de décisions stratégiques, d'optimisation de processus ou de la mise en œuvre de nouvelles idées commerciales. Nous voulons exposer ci-après comment les analyses de données peuvent profiter à l'entreprise.**

*Patrik Rüegge, Dr Jonas Dischl, Dr Dimitrios Tombros*

La quantité de données traitées dans les entreprises augmente en moyenne de 40 % par an. En général, ces données ne servent qu'à des applications précises. Outre les données d'applications classiques et les collections de documents non structurées, les données générées automatiquement par des machines et systèmes ont également fortement augmenté. En relation avec les médias sociaux et le commerce électronique, de nouveaux outils sont apparus sur le marché pour analyser et interpréter les données sans recourir à un data-warehouse.

Une meilleure analyse des données et le profit qui en résulte entraînent divers avantages:

- Meilleure compréhension de la chaîne de création de valeur: une fois traitées, les données exposent les processus effectifs et sont la source d'idées d'amélioration et de conception plus efficace des processus.
- Base des prévisions stratégiques: une visualisation des données permet d'exposer aux décideurs des situations complexes sous une forme plus simple, d'où la possibilité de procéder à des adaptations stratégiques plus ciblées et plus efficaces.
- Meilleures bases de décision: les décisions ne doivent pas être prises par intuition, mais sur la base de faits concrets.



*Sources pour l'analyse des données*

## Formats de données et outils d'analyse

En règle générale, les données viennent de diverses sources et ont des structures variées. De nouveaux outils et technologies permettent d'analyser ces données.

Les formats peuvent être les suivants:

- Données structurées en tableaux
- Données semi-structurées, comme logs et fichiers XML
- Données non structurées (documents ou commentaires en texte libre)
- Graphes ou données relationnelles

Pour ces données, il est toujours possible de recourir aux méthodes présentées ici si les outils appropriés sont disponibles et les données correctement préparées.

Un vaste éventail d'outils est disponible pour l'analyse des données – des simples interfaces SQL pour l'analyse manuelle jusqu'aux outils étendus de data mining.







Un choix d'outils différents pour l'analyse des données

## Méthodes d'analyse des données

Il existe des méthodes très différentes pour analyser les données. Qui y voit clair pourra obtenir un résultat optimal en choisissant la méthode appropriée. Les méthodes «manuelles» sont connues, éprouvées et largement répandues: l'utilisateur se sert de l'ordinateur pour afficher ses données sous différentes formes et tire des conclusions de ce qu'il voit.

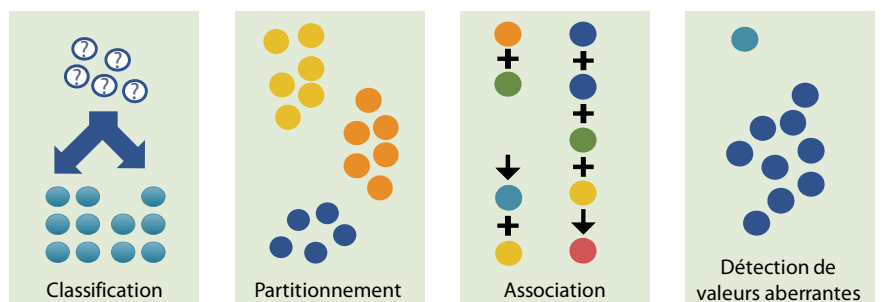
Mais en cas de données avec beaucoup d'attributs et un contexte complexe, les méthodes manuelles atteignent leurs limites. C'est là qu'entre en jeu le «data mining» – l'ordinateur fouille automatiquement les données à la recherche de certains modèles ou contextes.

Analyse manuelle des données	Data mining
 Regroupement, isolation, classement, affichage	 Recherche de modèles et structures Calcul de modèles
 Recherche de modèles et structures Confirmation d'hypothèses	 Commande d'algorithmes de recherche Interprétation des résultats

Deux types d'analyse de données

Depuis de nombreuses années, le data mining est employé avec succès pour un grand nombre de problèmes et prend toujours plus de l'importance avec l'actualité de l'analyse de données en masse. Il vaut donc la peine d'examiner les principales problématiques pouvant être résolues par le data mining.

- Avec la **classification**, les nouvelles données sont automatiquement classées en fonction des expériences passées. C'est par exemple la méthode employée par les banques pour évaluer la solvabilité d'un client.
- Le **partitionnement** sert à regrouper les données analogues. Les entreprises connaissent mieux leurs clients potentiels et peuvent par exemple placer leur publicité de manière plus ciblée.
- Par l'**association**, on voit quelles données apparaissent fréquemment ensemble. Ainsi, on peut par exemple établir quels produits sont achetés souvent ensemble et optimiser le portefeuille de produits ou la présentation des produits en magasin.
- La **détection de valeurs aberrantes** fait ressortir les divergences – même là où on ne les avait pas remarquées par les méthodes manuelles. Ceci est notamment de grande importance pour la détection de fraudes.



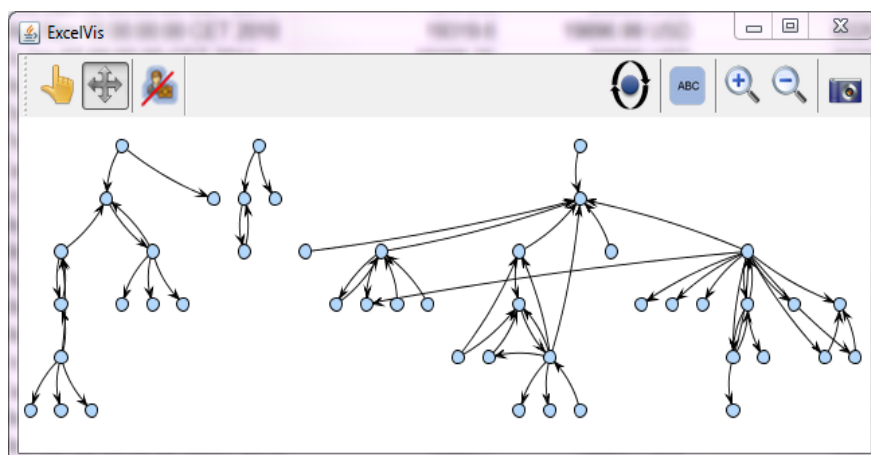
Les principales problématiques dans le data mining

## Utilisations possibles

Les possibilités d'analyse des données sont aussi variées que les données qui sont disponibles dans les entreprises. Les emplois concrets suivants ne couvrent qu'une petite partie de l'éventail:

- Les **analyses de ventes croisées** se justifient dans diverses branches. Ainsi, par exemple, l'information concernant une nouvelle assurance-vie d'une grosse compagnie d'assurance a pu être communiquée à des clients potentiels.
- Une **analyse des médias sociaux** peut par exemple montrer dans quelle mesure les tweets des clients d'un prestataire de services en communication peuvent servir à la détection précoce d'incidents. En outre, les médias sociaux se prêtent à une «analyse des sentiments» pour établir une relation entre les clients et les produits et services.
- L'**analyse des demandes de services** permet d'identifier les produits exigeant un support important et d'optimiser le portefeuille de produits. La fouille de textes, ou text-mining, cherche, dans les textes libres de demandes de services, des modèles qui signalent des problèmes jusqu'alors inconnus.
- Dans la **détection de fraude**, on emploie diverses méthodes d'analyse des données pour identifier des modèles de fraude jusqu'alors inconnus ou classifier des incidents suspects. Ceci permet, par exemple, la reconnaissance automatique de déclarations de sinistre douteuses faites aux compagnies d'assurance ou de certaines corrélations entre des sinistres.
- Les analyses de marché aident par exemple, dans le cadre du développement du «**Fiber-To-The-Home**», à identifier le potentiel de certaines régions et à fixer des priorités en conséquence.

Une application intuitive particulièrement attractive est le **visual data mining**. Cette méthode visualise les données de manière optimale pour permettre au cerveau humain une reconnaissance des modèles. Les graphes dits de transaction aident par exemple à reconnaître en un coup d'œil les constellations douteuses de flux financiers dans une grosse quantité de données – à supposer que l'on possède l'expérience correspondante pour effectuer une interprétation correcte. Excelvis est un outil de visual data mining développé par AWK qui est employé dans les banques à des fins d'analyses diverses.



L'outil de visual data mining «Excelvis» d'AWK

## Exemple: la lutte contre la fraude

Les analyses de données traditionnelles se concentrent sur des objets ou individus examinés isolément et sur leurs caractéristiques (p. ex. les clients et leurs achats ou les filiales et leur chiffre d'affaires).

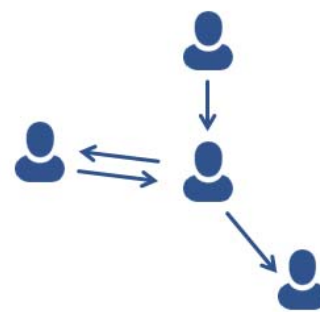
Dans notre monde de réseaux, il est de plus en plus essentiel de tenir également compte, dans l'analyse, des liens entre les objets et les personnes (p. ex. transactions entre divers clients).



Sexe: m  
Âge: 42  
Type de client: X

*Approche traditionnelle: le client et ses caractéristiques*

Notamment dans la détection des fraudes, les analyses par graphes prennent une importance croissante et s'imposent dans les outils du commerce. Ces méthodes, également dénommées analyses de liens ou de réseaux, aident à reconnaître des modèles qui restent invisibles par les méthodes classiques.



*Approche par graphe: les clients avec leurs interactions (p. ex. flux financiers ou entretiens téléphoniques)*

Les modèles d'interaction rendent visible un comportement douteux en révélant par exemple des liens inhabituels (indirects) entre diverses personnes ou organisations par le biais de flux financiers. AWK a déjà mis en œuvre de telles solutions dans l'environnement bancaire.

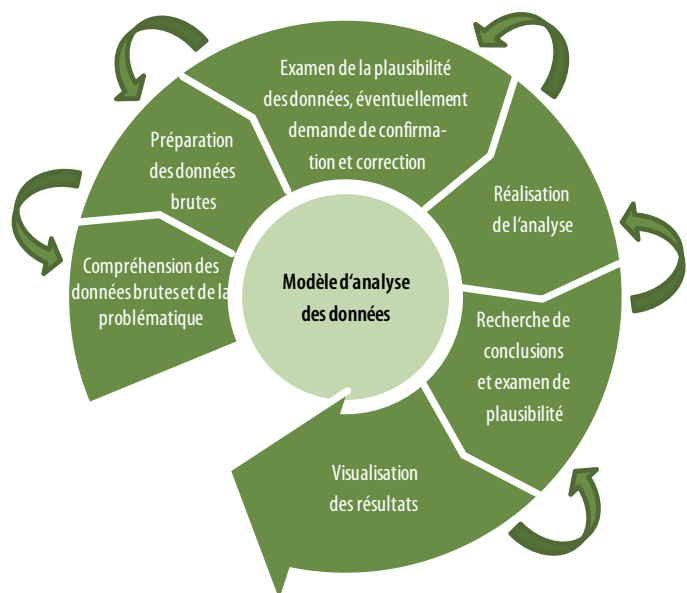
## Comment procéder

Correctement appliquées, les diverses méthodes d'analyse manuelle des données peuvent souvent fournir très rapidement des informations précieuses. Mais selon la situation, les informations vraiment intéressantes ne sont fournies que par un algorithme de data mining approprié.

Manuellement ou à l'aide d'outils de data mining, les analyses de données offrent en tout cas un gros potentiel mais sont également délicates. Les fausses interprétations et d'autres erreurs peuvent mener à de mauvaises décisions. Une solide compréhension des bases théoriques et statistiques des diverses méthodes alliée à une longue expérience pratique et pragmatique permet de découvrir le trésor qui se cache dans les données.

Le modèle d'AWK pour les analyses de données attache une grande importance à la compréhension des données brutes (Business Understanding) car ceci est essentiel pour le concept d'analyse et l'interprétation correcte des résultats.

Les données brutes doivent être préparées pour l'analyse et mises sous une forme appropriée. Selon le format et la qualité des données à analyser, cette opération constitue une grande part du travail global. Dans ce domaine, l'efficacité, grâce aux Best Practices, est payante. Un examen critique constant de la plausibilité après chaque étape assure la qualité et l'absence d'erreurs. Les visualisations adaptées aident à communiquer les résultats essentiels en fonction des groupes cibles respectifs.



AWK effectue des analyses de données et offre son assistance pour l'évaluation et l'application optimale de solutions d'analyse chez le client. Notre expérience montre que le potentiel d'outils, souvent coûteux, n'est exploité que partiellement. Quelle que soit la branche, nous avons constaté que de nombreux utilisateurs ne connaissent pas ou n'utilisent pas les outils modernes et peu coûteux. Dans ce cas, une formation peut fournir le savoir-faire nécessaire pour permettre à l'organisation de retirer de ses données des résultats plus complets sans l'emploi d'outils coûteux.

## Les auteurs



**Patrik Rüegge**  
Chef de secteur  
Dipl. El.-Ing. ETH

*Patrik Rüegge est chez AWK responsable du secteur Banques et assurances*



**Dr. Jonas Dischl**  
Consultant,  
Dr. Inform., Dipl. Inform.

*Jonas Dischl est chez AWK expert en matière d'analyse de données et de détection de fraudes*



**Dr. Dimitrios Tombros**  
Chef de secteur,  
Dr. Inform., MScS,  
lic. oec. publ.

*Dimitrios Tombros est chez AWK responsable du domaine de compétence Management de l'information*

## A propos d'AWK Group

*Avec 130 collaborateurs, AWK est l'une des principales sociétés de consulting suisses pour les technologies de l'information.*

*Nous travaillons dans toute la Suisse à partir de nos sites de Zurich et de Berne.*

*Nos prestations de services comprennent le consulting, l'ingénierie et la gestion de projet.*



**AWK GROUP**  
Consulting | Engineering | Project Management

[www.awk.ch](http://www.awk.ch)