

Démystifier les Technologies

Décembre 2023

Cette brochure à l'intention des professionnels des ISC vise à fournir des orientations visuelles claires sur certaines technologies pertinentes et susceptibles d'influer sur le travail des auditeurs publics, tout en établissant une terminologie commune.





Quel est l'objectif de ce document ?

Cette brochure à l'intention des professionnels des ISC vise à fournir des orientations visuelles claires sur certaines technologies pertinentes et susceptibles d'influer sur le travail des auditeurs publics, tout en établissant une terminologie commune. L'objectif est essentiellement d'apporter un aperçu de haut-niveau sur les principales technologies et concepts aux professionnels ou auditeurs qui ne sont pas spécialisés dans les technologies.

Quelles sont les technologies traditionnelles et émergentes ?

Les technologies peuvent être classées en deux catégories : les technologies traditionnelles et émergentes.

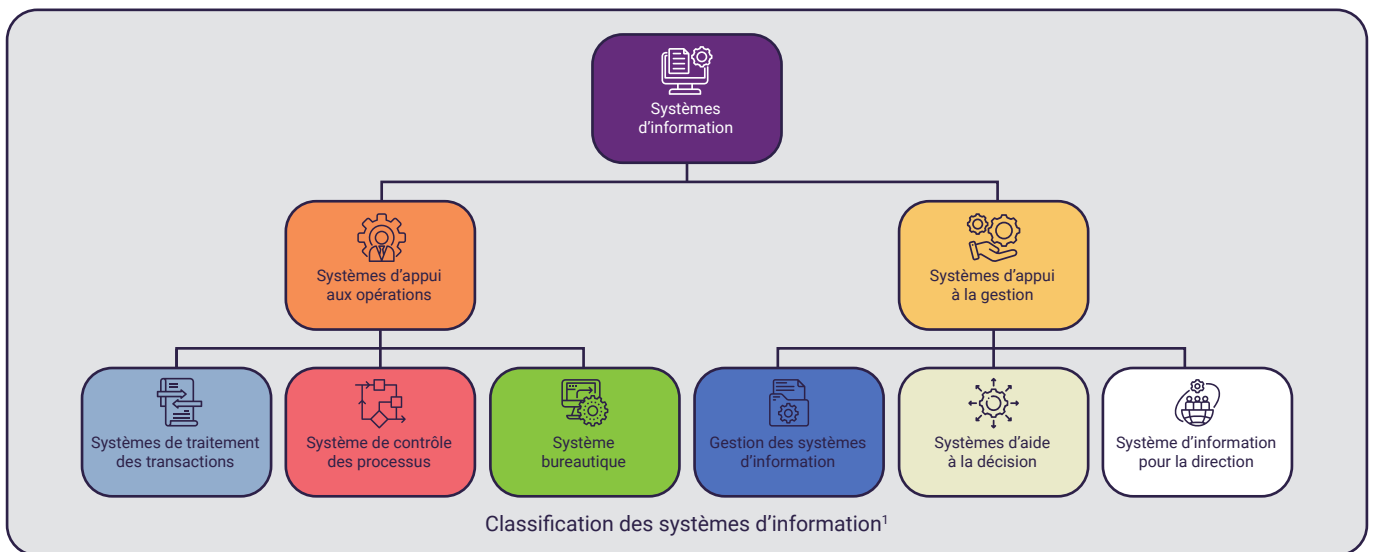
Les **technologies traditionnelles** sont utilisées depuis longtemps mais restent d'actualité. Elles ont déjà fait la preuve de leur efficacité et de leur fiabilité. Ces technologies traditionnelles comprennent notamment les applications de traitement de texte, les logiciels de productivité bureautique, les bases de données, les entrepôts de données et l'intelligence artificielle.

À l'inverse, les **technologies émergentes** sont relativement récentes. On peut par exemple citer l'intelligence artificielle, la blockchain (ou chaîne de blocs), l'automatisation robotisée des processus (RPA) ou le cloud computing (informatique en nuage). Ces technologies sont en constante évolution et présentent le potentiel de révolutionner nos modes de vie et de travail.

Systemes et applications : de quoi parle-t-on ?

Un **système d'information** est une compilation de composantes interdépendantes qui fonctionnent de concert pour recueillir, traiter, stocker et diffuser des informations. La finalité d'un système d'information est de faciliter la prise de décision, la coordination, le contrôle, l'analyse et la visualisation au sein d'une organisation. Les systèmes d'information sont utilisés en appui des opérations, des travaux intellectuels et de la gestion des organisations.

Dans le langage courant, le terme « système d'information » est souvent confondu avec « système informatique ». Il convient néanmoins de noter que les systèmes informatiques ne sont qu'un volet des systèmes d'information. Un système d'information comprend l'ensemble des composantes et processus constitutifs du système, comme les personnes et les processus.



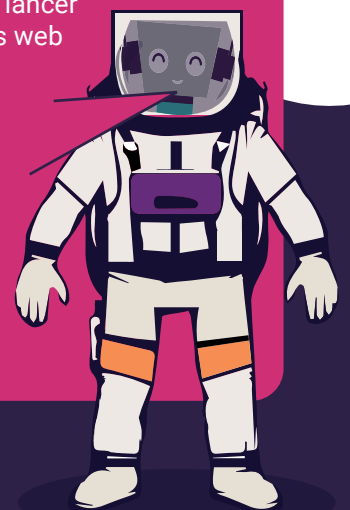
Les **systèmes informatiques** comprennent les éléments logiciels et matériels à même de fournir des services de type automatisation, transactions, traitement de données et intégration.

Les **applications informatiques** sont des logiciels conçus pour être utilisés par des personnes ou des appareils, par exemple des outils de connaissance, des logiciels d'accès à l'information, d'aide à la décision, de services aux entreprises, de commerce électronique, de divertissement, de médias et de jeux.

En fonction de l'architecture, les applications informatiques peuvent être :

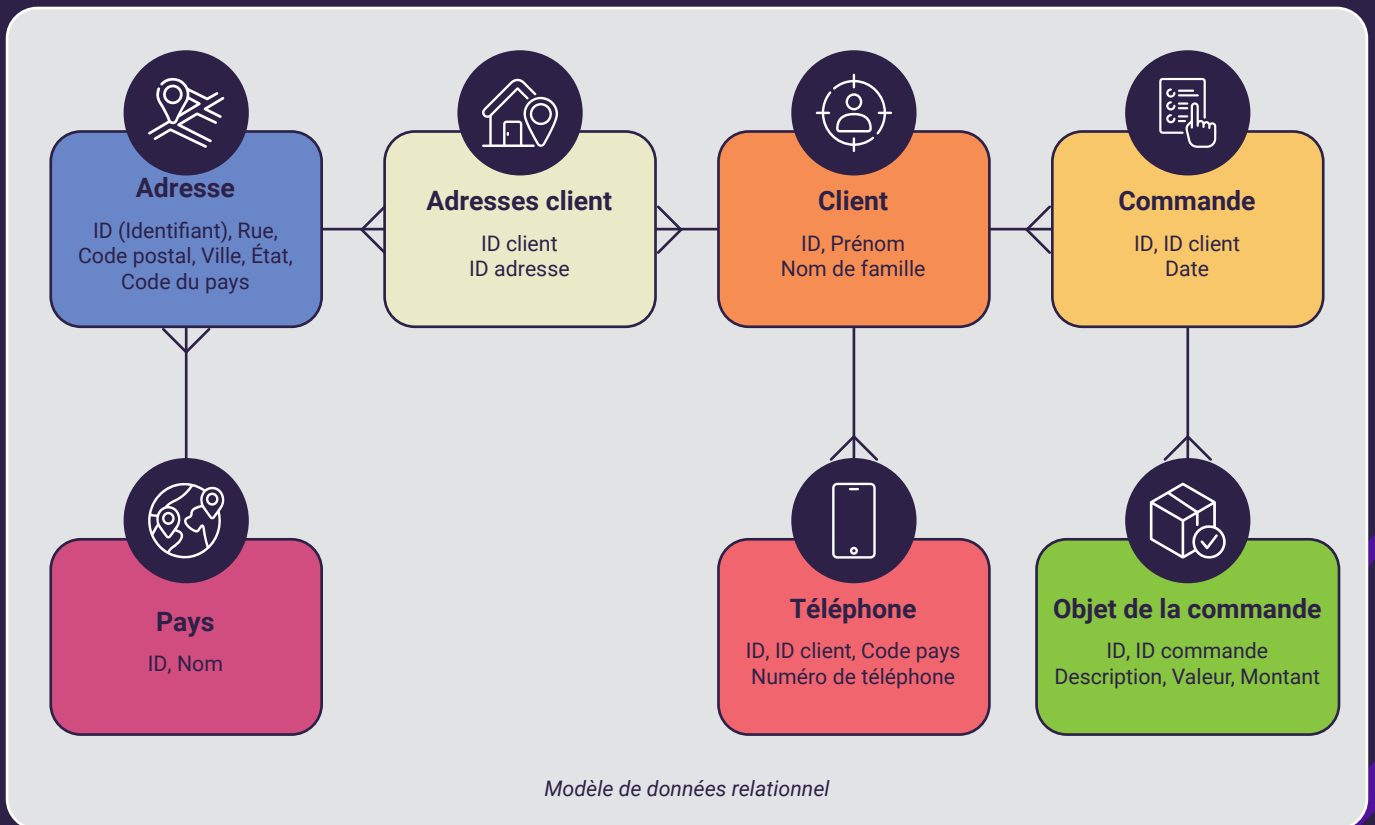
- des **applications de bureau**, c'est à dire un programme logiciel qui fonctionne localement sur un ordinateur et doit être lancé à l'aide du système d'exploitation ;
- des **applications web**, à savoir un logiciel que le client (ou interface utilisateur) doit lancer avec un navigateur web. Contrairement aux applications de bureau, les applications web ont forcément besoin d'un navigateur internet pour être utilisées ;
- des **applications mobiles**, soit un logiciel conçu pour fonctionner sur un appareil mobile comme un smartphone, une tablette ou une montre.

Notez que certaines applications peuvent être proposées dans plusieurs, voire toutes les formes ci-dessus. Par exemple, Microsoft Office est disponible en tant qu'application de bureau installée sur un ordinateur, en tant qu'application web disponible à l'aide d'un navigateur internet, et en tant qu'application mobile fonctionnant sur téléphone ou tablette.

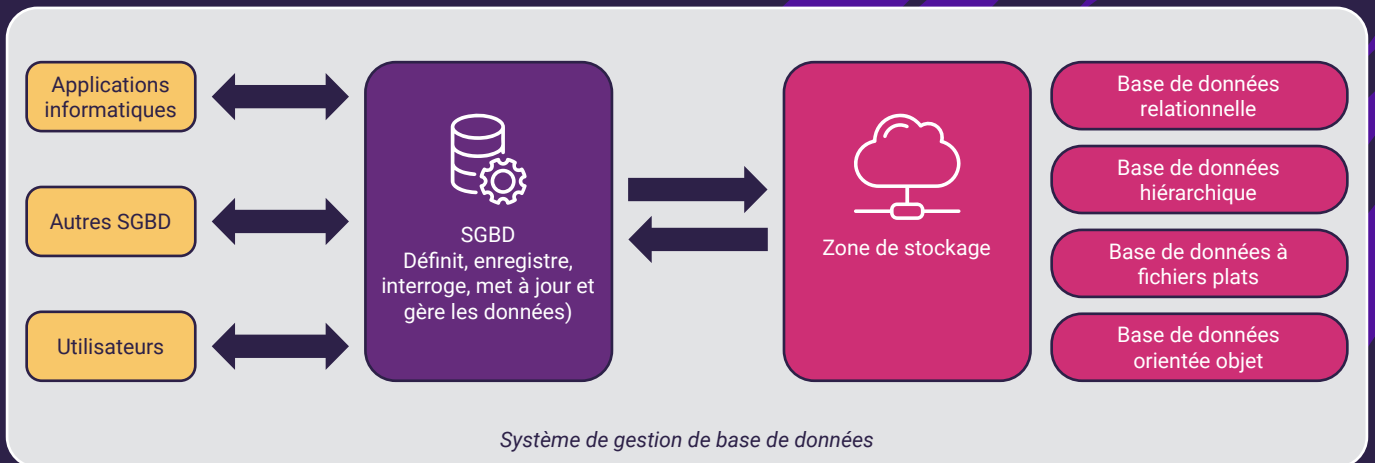


Comment collecter, stocker et diffuser les données ?

Un **modèle de données** est un modèle abstrait qui permet d'organiser des éléments de données et de normaliser leurs relations entre elles, ainsi qu'avec des entités concrètes. Les modèles de données courants sont les modèles à plat, les modèles hiérarchiques, les modèles en réseau, les modèles relationnels et les modèles en étoile.



Une **base de données** est une compilation organisée d'informations structurées (ou de données), le plus souvent stockées sous forme électronique dans un système informatique. Les bases de données sont généralement contrôlées par un **système de gestion de bases de données (SGBD)**.



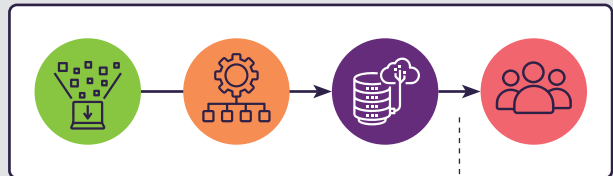
Un **entrepôt de données** (data warehouse) est une réserve centralisée qui permet de stocker toutes les données structurées collectées et utilisées par une organisation. Il est organisé de façon à faciliter l'interrogation et l'analyse des données. En général, les entrepôts de données reposent sur des bases de données spécifiques organisées en étoile (ou en flocon de neige). Les données sont le plus souvent exportées vers l'entrepôt de données à l'aide des processus **Extract, Transform et Load (ETL)**.

Un **datamart** (ou magasin de données) est un entrepôt de données simplifié, axé sur un métier ou un domaine d'activité spécifique comme les ventes, la finance ou le marketing. Compte tenu de leur spécialisation, la source des données des datamarts est plus restreinte que celle des entrepôts de données.

Un **lac de données** (data lake) est une grande réserve centralisée conçue pour stocker des données brutes sous leur format original. Ils peuvent contenir une grande variété de types et de formats de données, et facilitent l'accès et l'analyse de celles-ci.

Entrepôt de données

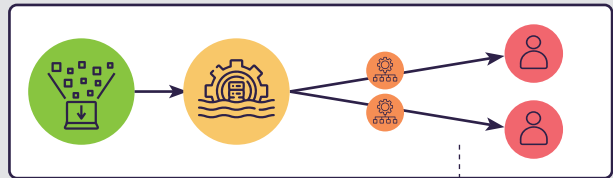
Les données entrantes sont nettoyées et organisées dans un schéma unique et cohérent avant d'être exportées vers l'entrepôt de données



L'analyse est réalisée directement sur les données sélectionnées de l'entrepôt de données

Lac de données

Les données entrantes intègrent le lac de données sous leur forme brute

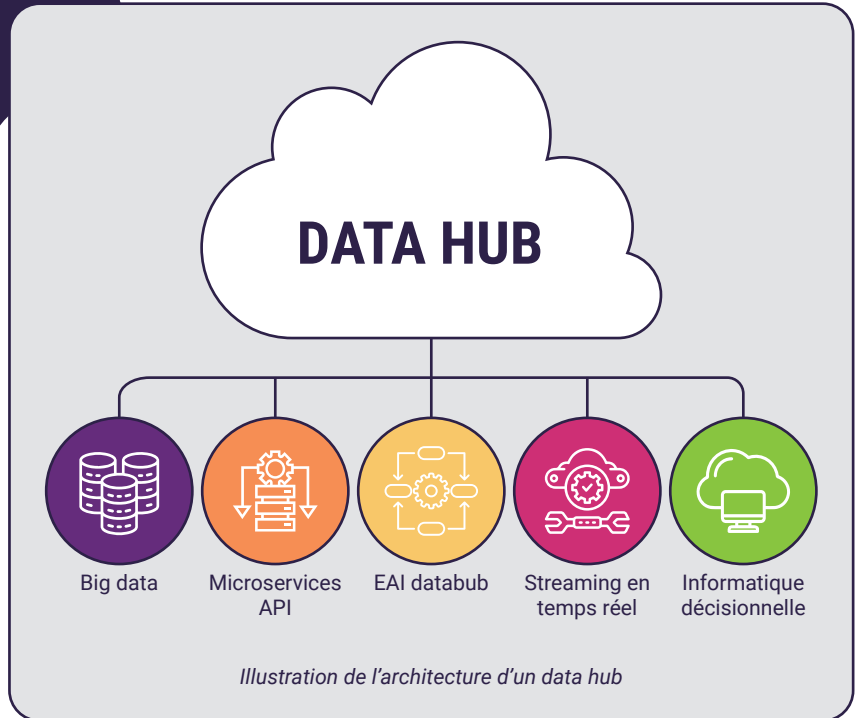


Les données sont sélectionnées et organisées en fonction des besoins individuels

Comparaison entre un entrepôt de données et un lac de données ²

Les **big data** (mégadonnées) sont des ensembles d'informations volumineux et variés qui se développent de façon exponentielle. Ils englobent le *volume* d'informations, la *vitesse* à laquelle elles sont créées et collectées, et la *variété* ou l'étendue des points de données couverts (les 3V du big data). Parmi les exemples de big data, citons les systèmes de traitement des transactions, les bases de données clients, les documents, les courriers électroniques, les dossiers médicaux, les registres de données *clickstream* (flux des clics effectués par les utilisateurs d'un site internet), les applications mobiles et les réseaux sociaux.

Un **data hub** est un centre d'échange de données reposant sur la science des données, l'ingénierie des données et les technologies de stockage des données afin de créer des interactions avec des points de terminaison tels que les applications ou les algorithmes. Un *data hub* est une architecture et une stratégie de gestion des données, et non pas un simple produit. Imaginez une réserve de données centrale équipée de rayons irradiant les systèmes et les clients. Ainsi, l'architecture du *data hub* permet le partage des données en mettant tout simplement en lien producteurs et consommateurs des données.

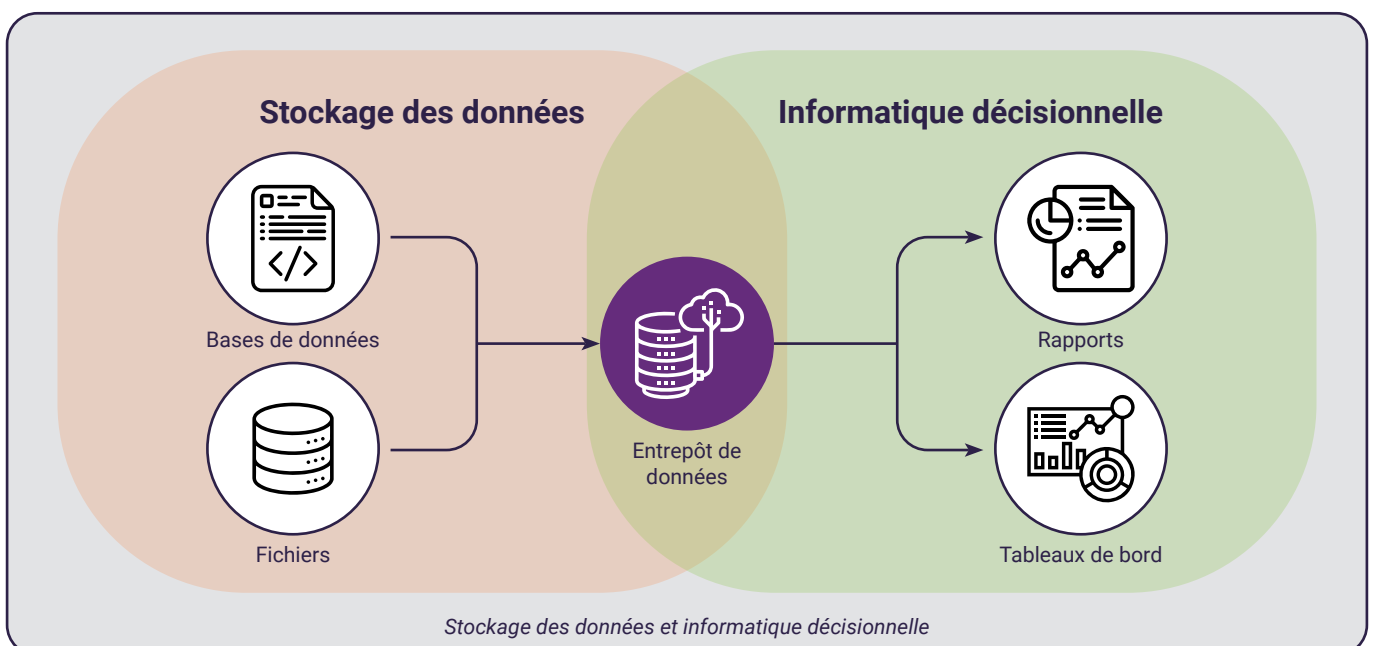


Comment analyse-t-on les données ?

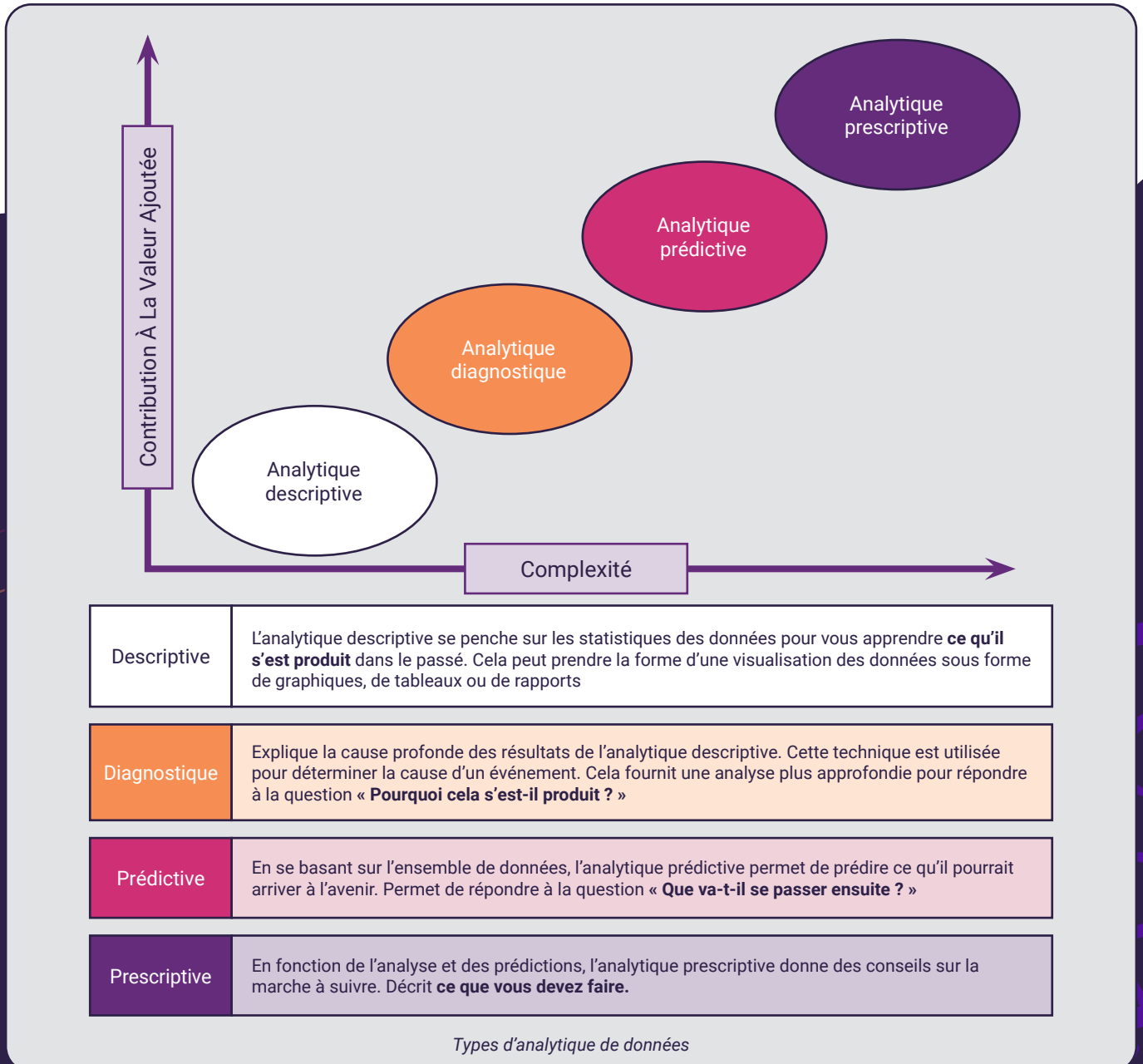
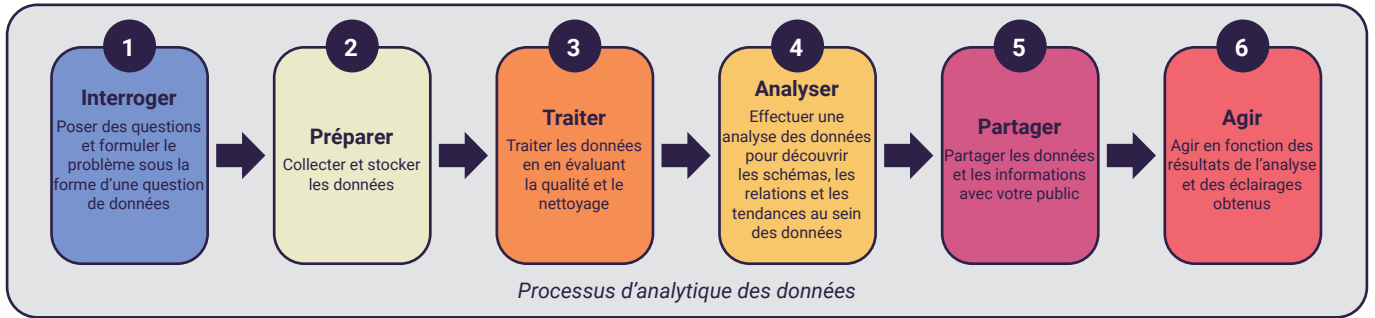
La **communication des données** (data reporting) implique la collecte et le formatage de données brutes afin d'évaluer la performance continue d'une organisation. Les rapports élaborés permettent de répondre à des questions élémentaires sur l'entreprise.

La **visualisation des données** désigne l'utilisation de graphiques courants, comme des diagrammes, des tracés, des infographies, et même des animations, pour représenter les données. Ces représentations visuelles de l'information permettent de retranscrire des relations complexes entre les données et autres informations reposant sur les données d'une manière facile à comprendre.

L'**informatique décisionnelle** désigne les processus qui convertissent des données brutes en rapports significatifs et en permettent la diffusion dans toute l'organisation. Ces connaissances concrètes aident les dirigeants à prendre de meilleures décisions. L'informatique décisionnelle peut être mise en œuvre en complément d'une base et d'un entrepôt de données. Elle implique généralement des tableaux de bords interactifs, des rapports, etc.



L'**analytique de données** (data analytics) est le processus d'examen de données brutes en vue d'en tirer des conclusions et éclairages. Il s'agit d'un domaine très vaste impliquant l'usage de données et d'outils afin de prendre des décisions commerciales éclairées. Le processus d'analytique des données comprend la collecte et l'analyse des données. En revanche, l'**analyse des données** ne représente qu'une composante de l'analytique des données. Celle-ci implique l'observation, la transformation, le nettoyage et la modélisation de données brutes en vue d'obtenir des éclairages concrets utiles pour l'optimisation d'une stratégie ou d'un processus. L'analyse des données est néanmoins limitée à des ensembles de données préalablement préparés.



Quels langages peut-on utiliser pour parler aux ordinateurs ?

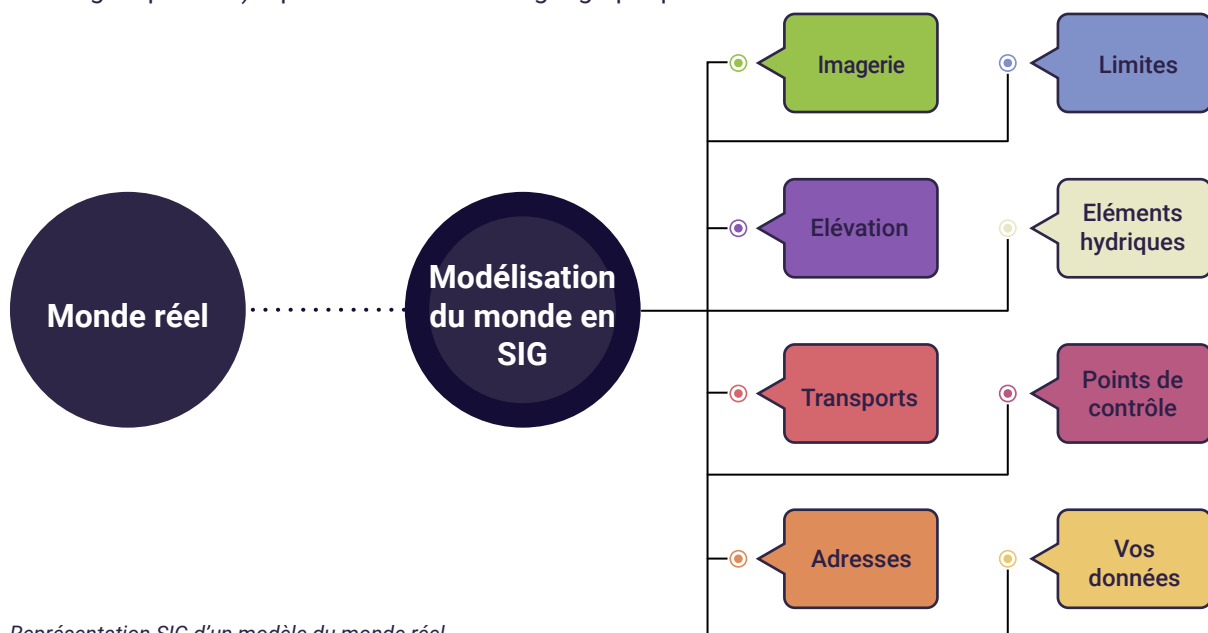
Les langages de programmation sont des notations utilisées pour écrire des programmes informatiques. Les développeurs utilisent de nombreux langages de programmation en fonction de leurs besoins. Ces langages ont été conçus pour différentes plateformes et systèmes d'exploitation. Selon certaines sources, il existerait près de 9000 langages de programmation, mais seule une poignée est encore utilisée aujourd'hui. Les langages les plus courants sont par exemple Python, JavaScript et C++. Vous pouvez commencer à coder après avoir appris un seul de ces langages. Dans notre brochure, nous couvrirons les trois principaux langages utilisés pour le traitement des données.

- **Structured Query Language (SQL)** est un langage dédié utilisé en programmation et conçu pour la gestion des données stockées dans les systèmes de gestion de bases de données relationnelles.
- **R** est un langage et un environnement pour l'informatique statistique et les graphiques.
- **Python** est un langage de programmation de haut-niveau interprété, orienté objet et doté d'une sémantique dynamique.

La **méthodologie de développement de logiciels** désigne le processus ou l'ensemble de processus utilisé pour le développement d'un logiciel. Cela comprend les phases de conception, de développement, et d'autres façons de penser le processus de développement. La méthodologie de développement de logiciels prend la forme d'une approche structurée du développement de logiciels. Parmi les méthodologies courantes de développement de logiciels, on retrouve la méthode agile, la méthodologie de déploiement DevOps, la méthode Waterfall et la méthode de développement rapide d'application (RAD).

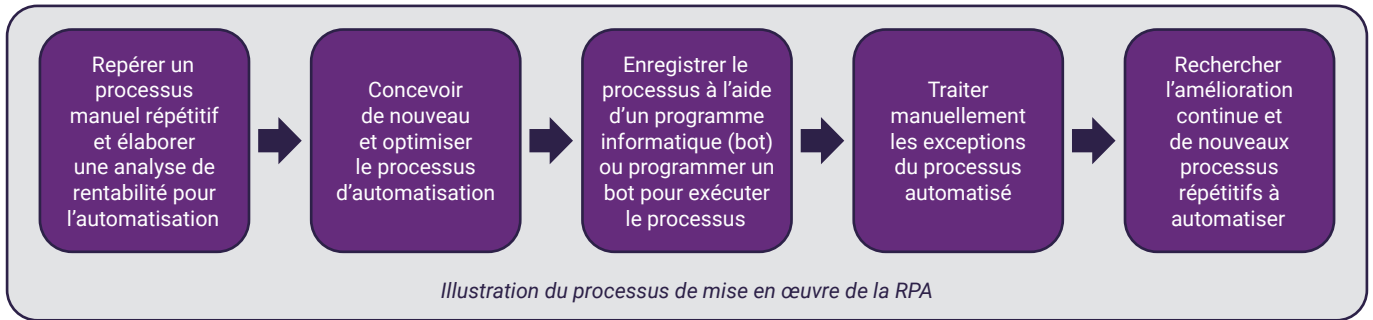
Quelles autres technologies peut-on rencontrer ?

Les **systèmes d'informations géographiques**, ou SIG, sont des outils informatiques utilisés pour stocker, visualiser, analyser et interpréter des données géographiques. Les données géographiques (aussi appelées données spatiales ou géospaciales) repèrent la localisation géographique des éléments.



Représentation SIG d'un modèle du monde réel

L'automatisation robotisée des processus (RPA) est la mise en œuvre de logiciels d'automatisation spécialisés pour exécuter des tâches répétitives, très volumineuses et à forte intensité.



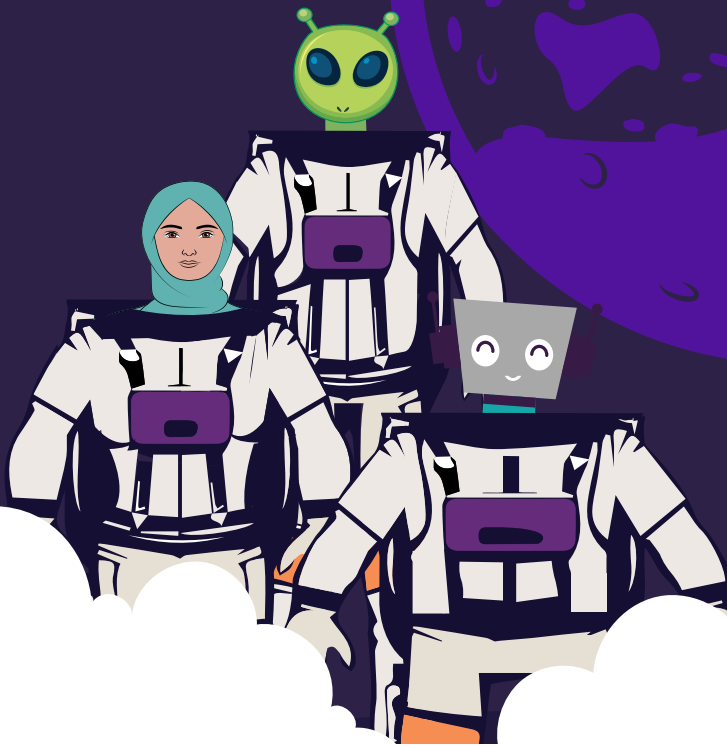
Le cloud computing (informatique en nuage) désigne la disponibilité de ressources de systèmes informatiques à la demande, particulièrement en matière de stockage de données et de performances, sans gestion active directe de la part de l'utilisateur. Les gros clouds disposent souvent de fonctions réparties sur plusieurs sites (chacun étant un centre de données, ou data center). L'essentiel des services de cloud computing se répartissent en trois catégories : l'infrastructure en tant que service (IaaS), la plateforme en tant que service (PaaS) et le logiciel en tant que service (SaaS).

	Sur site	IaaS <i>Infrastructure as a Service</i>	PaaS <i>Platform as a Service</i>	SaaS <i>Software as a Service</i>
Données				
Applications				
Environnement d'exécution				
Système d'exploitation				
Serveurs				
Stockage				
Mise en réseau				

Les cases rouges sont gérées par un prestataire de services - Les cases vertes sont gérées par l'organisation

Classification des plateformes de cloud



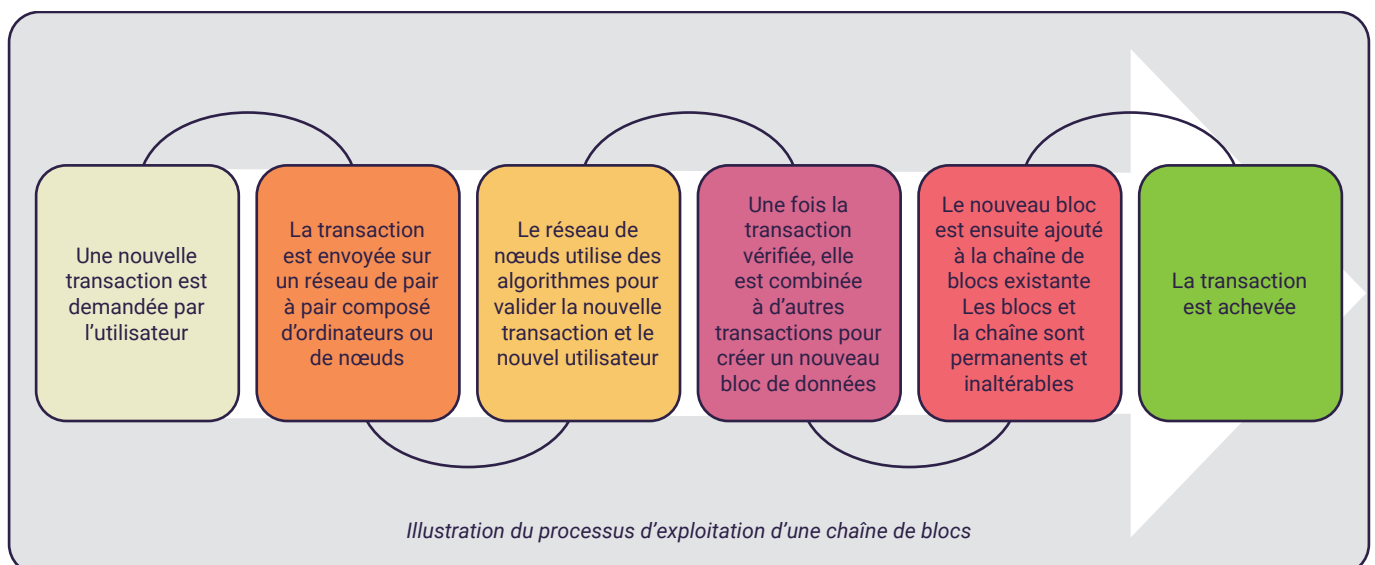


Le cloud computing présente cinq caractéristiques clés :





1. Libre-service à la demande
2. Accès global au réseau
3. Mutualisation des ressources
4. Élasticité rapide
5. Service mesuré

Ces caractéristiques permettent aux consommateurs de se doter facilement et rapidement en capacités informatiques, d'y accéder via le réseau, de mettre en commun des ressources, de fournir des capacités de manière élastique, et de contrôler et d'optimiser l'utilisation de ces ressources.

Une **blockchain (chaîne de blocs)** est un registre distribué disposant de listes croissantes d'enregistrements (les blocs) liées entre elles de façon sécurisée par hachage cryptographique. Grâce au registre distribué, les transactions enregistrées sur la chaîne de blocs sont difficiles à falsifier.

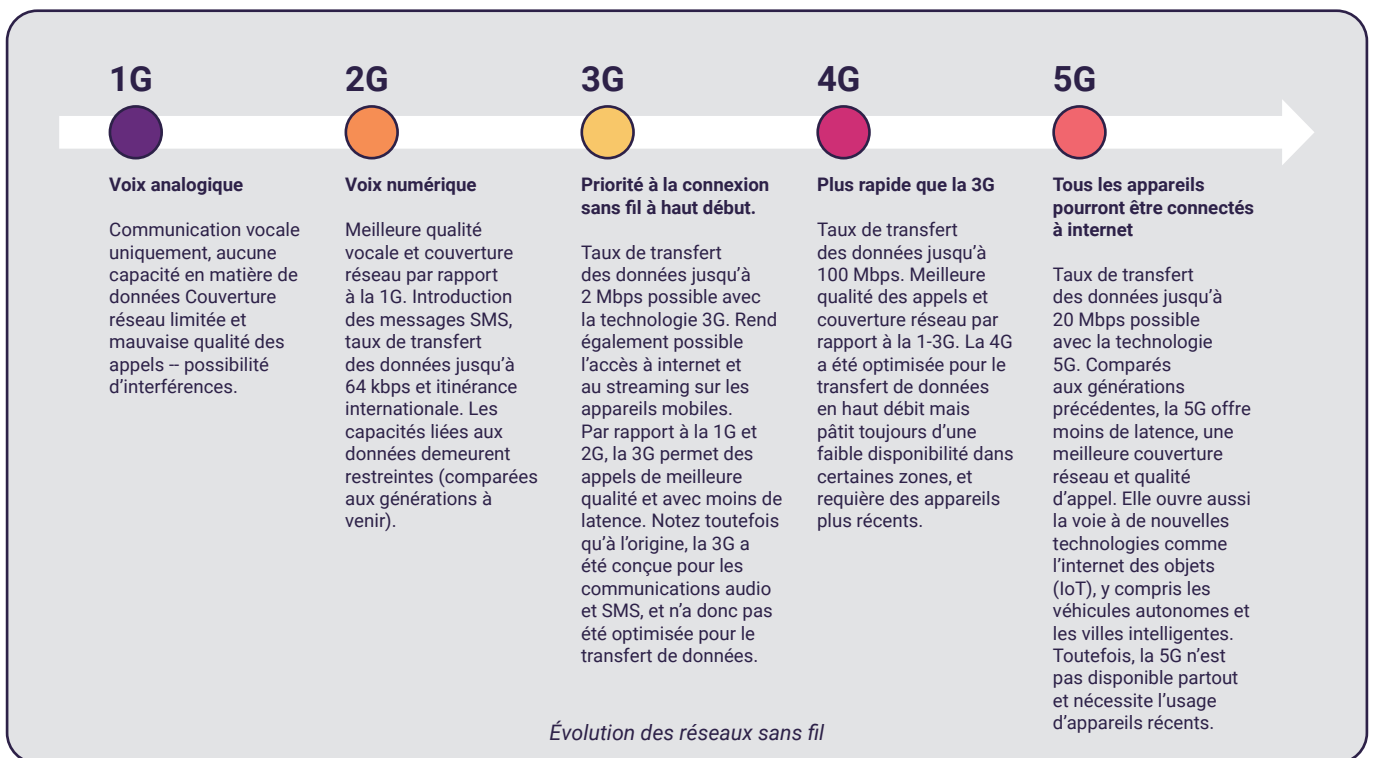


L'intelligence artificielle (IA) désigne la capacité d'un ordinateur numérique ou d'un robot contrôlé par ordinateur d'exécuter des tâches généralement associées à des êtres intelligents. Le terme est souvent associé aux systèmes de développement dotés de processus intellectuels caractéristiques des humains, comme la capacité de raisonner, de découvrir des significations, de faire des généralisations ou d'apprendre de ses expériences.

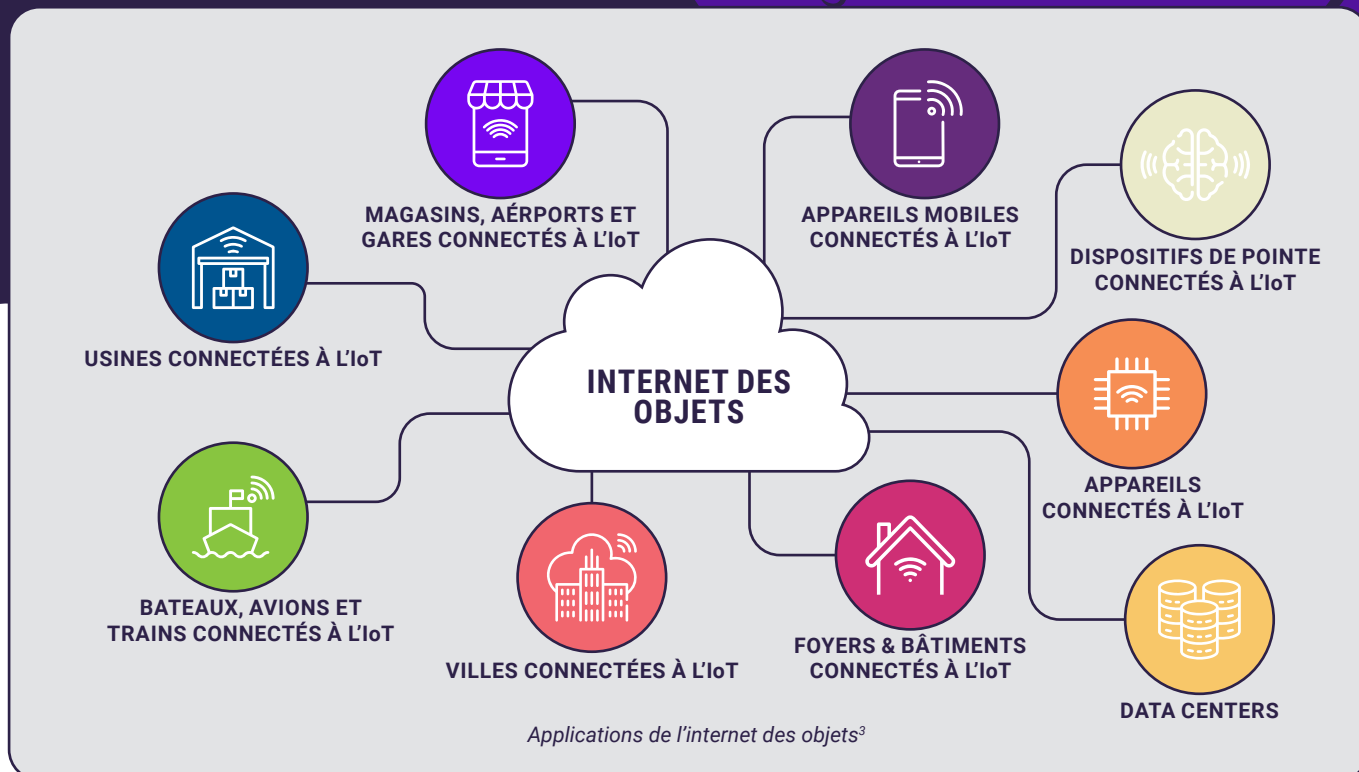
 <p>Apprentissage automatique Utilise des échantillons de données pour entraîner les programmes informatiques à reconnaître des schémas reposant sur des algorithmes</p>	 <p>Réseaux neuronaux Systèmes informatiques conçus pour imiter les neurones d'un cerveau</p>
 <p>Traitement du langage naturel Capacité de comprendre une discussion, mais aussi à comprendre et à analyser des documents</p>	 <p>IA générative Capacité de créer de nouveaux contenus, comme des textes, des images, des documents audio et vidéo</p>

Types d'IA

La **5G** désigne le réseau sans fil de cinquième génération, qui a le potentiel de considérablement améliorer les communications mobiles.



L'internet des choses (IoT) désigne un réseau d'objets physiques, également appelés « choses », équipés de capteurs, de logiciels et d'autres technologies. Ces appareils sont conçus pour être connectés et échanger des données avec d'autres appareils et systèmes à l'aide d'internet ou d'autres réseaux de communication. L'internet des objets couvre les domaines de l'électronique, des communications et de l'ingénierie informatique.



Qu'est-ce tout cela implique pour les auditeurs des ISC ?

Nous vivons dans un monde en constante évolution, et si elles veulent rester à la page, les institutions supérieures de contrôle (ISC) doivent s'adapter à de tels changements. Les avancées technologiques ont révolutionné nos modes de vie, et ce processus ne fera que s'accélérer. Même ceux qui se sont un temps montrés sceptiques face à la numérisation investissent désormais dans les technologies, car elles sont devenues incontournables à la survie. Face à ces progrès, les organisations et les individus doivent traiter d'immenses volumes de données et ce, plus rapidement que jamais auparavant.

Les gouvernements investissent dans ces technologies et mettent en place de nouveaux systèmes. D'un côté, cela leur permet d'optimiser leurs fonctions et d'offrir de meilleurs services à leurs citoyens. D'un autre, ces évolutions imposent aux ISC d'acquérir de nouvelles compétences et connaissances afin de faire face à ces nouveaux sujets d'audits. Le rapport du Comité de renforcement des capacités (CBC) « The Relevant Value-Adding Auditor » et le principe INTOSAI 12 « valeur et avantages des institutions supérieures de contrôle – faire une différence dans la vie des citoyens » mettent tous deux en exergue ce même aspect.

Le rôle de l'auditeur d'ISC devra **évoluer et acquérir de nouvelles capacités** en vue de :

1. Utiliser des solutions technologiques pour conserver leur pertinence et maximiser leur impact. L'utilisation des technologies par les équipes d'audit contribue à la poursuite d'un *développement et d'opérations durables*, en assurant la *pertinence* des contrôles réalisés et de l'ISC, dans l'intérêt des citoyens. Les audits portant sur les technologies permettent de maximiser l'*impact* découlant de l'utilisation et du contrôle des technologies, ainsi que missions d'audit réalisées. Enfin, elles contribuent à maintenir les *capacités* requises pour que les ISC demeurent des organisations pertinentes, professionnelles et à fort impact.

2. Contrôler l'utilisation des technologies par le gouvernement pour renforcer la confiance dans celui-ci. Les audits portant sur les technologies assurent au public davantage de **confiance dans les systèmes technologiques gouvernementaux**. De plus, les audits technologiques favorisent des **sociétés et des vies meilleures** grâce à l'usage des technologies. Ils créent également une culture de **transparence**, de **responsabilité** et de **conformité**, ce qui entraîne un sentiment de confiance dans l'usage des technologies par les pouvoirs publics.

Les audits **technologiques** sont des examens portant sur l'usage des technologies au sein d'une organisation, notamment au niveau de ses infrastructures, politiques, procédures, applications, utilisation des données, et technologies émergentes (au moment de la rédaction de cette brochure, il s'agit par exemple de l'intelligence artificielle, de l'automatisation robotisée des processus, des blockchains ou du cloud computing). L'objectif de ces audits technologiques est de veiller à ce que la mise en œuvre de ces technologies par le gouvernement réponde à ses besoins, sans compromettre la sécurité, le respect de la vie privée, les coûts, l'équité et l'inclusion, ou d'autres éléments essentiels au bon fonctionnement de l'institution. Ces audits impliquent aussi d'examiner si le développement, la mise en œuvre et la maintenance de ces technologies satisfont aux objectifs de l'institution, permettent de protéger ses biens informationnels et de maintenir l'intégrité de ses données. Les audits technologiques permettent de repérer les cas où ces outils s'écartent des critères établis, en fonction du type de mission d'audit (performance, financier ou de conformité).

Les audits technologiques varient en fonction des types d'audits qu'ils ont pour but de contrôler. Par exemple :

- Dans le cadre d'un **audit financier**, un audit technologique pourrait prendre la forme d'un examen des contrôles généraux garantissant le fonctionnement des systèmes d'information sur lesquels reposent les processus financiers de l'entité, tels que décrits dans ses états financiers ;
- Dans le cadre d'un **audit de performance**, un audit technologique pourrait être mené pour déterminer quels avantages ont été produits et quelles économies mesurables ont été réalisées à l'échelle de l'administration grâce à l'adoption de nouvelles technologies ;
- Dans le cadre d'un **audit de conformité**, un audit technologique pourrait consister à vérifier la conformité du système d'information développé et/ou du processus de développement du système d'information par rapport aux normes et pratiques internationales.

