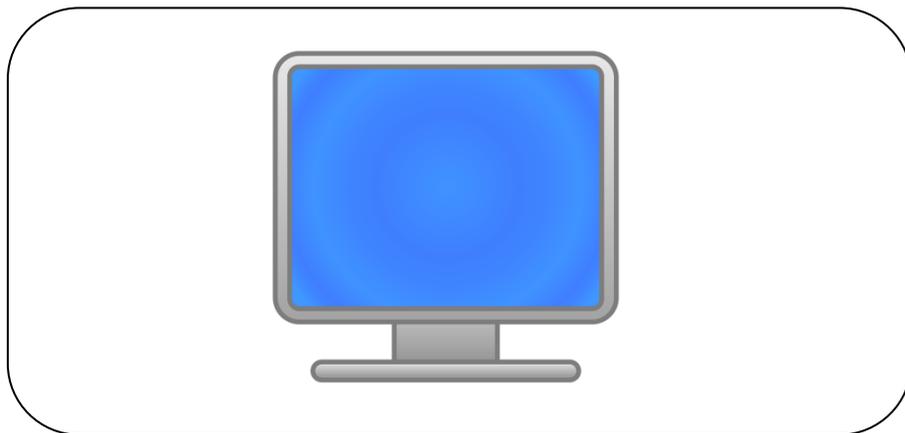


**Métiers de l'informatique, vocabulaires et thèmes pour les métiers du numérique orientés (ingénieur en informatique)**

**Tome 2 : Lexiques**

Par Dimitri PIANETA



## Introductions

Ce tome 2 de « **Métiers de l'informatique, vocabulaires et thèmes pour les métiers du numérique orientés (ingénieur en informatique)** » traite par thèmes de l'informatique. Les lexiques sont non exhaustifs. Je me suis concentré sur le vocabulaire essentiel pour mieux comprendre mon propre métier d'être ingénieur développeur en informatique.

J'ai traité des sujets comme les systèmes de l'information, l'intelligence artificielle, la data science, les bases de données et d'autres thèmes que vous allez découvrir.

J'ai voulu que ce tome 2 soit simple par thèmes dans le sommaire et deuxième partie de ce document un glossaire en ordre alphabétique pour trouver les mots.

Ce document a été le fruit de recherches biographiques dans des livres et des sites Internet qui m'ont aidé à le réaliser.

Je vous souhaite une bonne utilisation de ce document.

Bonne Lecture,

Dimitri PIANETA

Juillet 2020

## Table des matières

Introductions .....	2
Partie 1 : Lexique par thème .....	4
ABREVIATIONS ET SIGLES .....	5
Lexique du système d'information.....	7
Lexiques les architectures Orientées Services (SOA) .....	33
Lexiques les services Web .....	33
Lexiques concept de la programmation .....	34
Lexiques des tests logiciels.....	49
Lexiques des tests logiciels du CFTL/ISTQB .....	51
Lexiques de l'informatique.....	72
Lexiques de l'informatique selon le journal officiel Français .....	107
Lexiques Intelligence artificielle .....	117
Lexiques Intelligence artificielle selon le journal officiel Français .....	127
Lexiques Agilité.....	130
Lexiques modélisation UML .....	133
Lexiques du site web cyberDroit .....	144
Lexiques de la base de données.....	154
Lexiques de la Data Science.....	176
Partie 2 : Glossaire.....	195

## **Partie 1 : Lexique par thème**



---

## ABREVIATIONS ET SIGLES

**ADSI** : Active Directory Service Interface

**ADSL** : Asymmetric Digital Subscriber Line

**AGL** : Atelier de génie logiciel assisté par ordinateur

**AGP** : (Ang. Accelerated Graphic Port)

**ASP** : Application Service Provider

**BI** : Business Intelligence

**BO** : Business Object

**BPM** : Business Process Management

**CAO** : Conception Assistée par Ordinateur

**CFAO** : Conception et Fabrication Assistées par Ordinateur

**CRM** : Custom Relationship Management

**DAF** : Directeur administratif et financier

**DAG** : Le Développement d'Application de Gestion

**DG** : Direction Générale

**DOSI** : Directeur de l'Organisation et des Systèmes d'Information

**DSI** : Direction des Systèmes d'Information

**EAI** : Entreprise Resource Planning

**ESN** : Entreprise de Service Numérique

**FTP** : File Transfer Protocol

**GED** : Gestion Electronique de Documents

**GMAO** : Gestion de la Maintenance Assistée par Ordinateur

**GPAO** : Gestion de la Production Assistée par Ordinateur

**IDE** : Integrated development environment

**IP** : Internet Protocol

**LMDS** : Local Multipoint Distribution Service

**MOA** : Maître d'Ouvrage

**MOE** : Maître d'Œuvre

**MVNO** : Mobile Virtual Network Operator

**MVC** : Model Vue Contrôleur

**MVVM** : Model Vue Vue-Model

**NTIC** : Nouvelles Technologies de l'Information et de la Communication

**OSS** : Open Source Software

**PGI** : Progiciel de Gestion Intégrée

**PKI** : Public Key Infrastructure

**RFID** : Radio Frequency Identification

**RUP** : Rational Unified Process

**SAO** : Système d'Application Organisationnel

**SAP** : Systems Applications and Products for data processing

**SCM** : Supply Chain Management

**SDSL** : Symmetric Digital Subscriber Line

**SET** : Secure Electronic Transaction

**SI** : Système d'Information

**SIS** : Systèmes d'information Stratégiques

**SMS** : System Management Server ou Short Message Service

**SSII** : Société de Service en Ingénierie Informatique

**SSO** : Single Sign On

**TMA** : Tierce Maintenance Applicative

**TRA** : Tierce Recette Applicative

**UML** : Unified Modeling Language

**UMTS** : Universal Mobile Telecommunications System

**WiFi** : Wireless Fidelity

**WiMax** : Worldwide Interoperability for Microwave Access



---

## Lexique du système d'information

**Accès à distance** : Accès à un ordinateur et à ses périphériques à partir d'un terminal et placé à distance via un câble, un réseau ou une ligne téléphonique.

**Access Provider (Ang.)** : Fournisseur d'accès à Internet ou F.A.I..

**Administration des données** : Ce terme recouvre deux fonctions :

- tenue à jour de la documentation des données, tant du côté logique (Modèle conceptuel de données) que du côté technique (Métadonnées).
- Désignation du propriétaire de chaque données c'est-à-dire de celui qui donne les autorisations de mise à jour).

**Administrateur système** : parfois appelée SysAdmin, c'est la personne en charge d'assurer le bon fonctionnement des ressources informatiques. De l'installation au support utilisateur, en passant par le paramétrage et la maintenance des systèmes, ses missions l'amènent à jongler avec les besoins des différents services et métiers.

**A.D.S.L. (Ang. Asymmetric (bit rate) Digital Subscriber Line)** : Raccordement numérique asymétrique (R.N.A) ou liaison numérique à débit asymétrique.

Technologie de connexion Internet à haut débit utilisant le réseau téléphonique. Cette technologie est asymétrique, ce qui signifie que les données ne circulent pas dans le même sens à la même vitesse). Une liaison ADSL n'empêche pas les communications vocales et son coût, forfaitaire, est indépendant du temps de connexion. On peut donc sans inconvénient rester connecté en permanence.

**Agenda partagé** (Sharded Schedule) : Agenda électronique accessible aux diverses personnes d'un groupe de travail, et permettant notamment de détecter les créneaux possibles pour les réunions.

Plusieurs scénarios d'utilisation sont possibles, le plus courant étant le suivant :

- Chacun peut consulter l'emploi du temps des autres, et mettre à jour les informations concernant son propre emploi du temps.
- Des personnes autorisées (secrétaires) peuvent mettre à jour les informations de tout le monde.

**A.G.P. (Ang. Accelerated Graphic Port)** : Port graphique accéléré.

Connecteur se trouvant sur la carte mère et qui permet un transfert rapide entre le processeur et la carte graphique, permettant d'accélérer l'affichage des images et de la vidéo.

**Agent intelligent** (Intelligent Agent) : Messages contenant un programme de recherche d'information sur des bases de données internes ou externes, ainsi que des règles prescrivant leurs parcours et les conditions de la consultation de ces bases. En termes simples : un agent intelligent est un chien de chasse que l'on envoie sur le réseau et qui ramène l'information.

**Agilité** : dans une entreprise, l'agilité est la capacité à s'adapter rapidement au changement. Un management dit agile est généralement caractérisé par la mise en action de l'intelligence collective et une certaine auto-organisation des équipes. Des principes qu'on retrouve dans le monde du développement logiciel, où les méthodes agiles ont pris une place centrale : Scrum, eXtreme

Programming ou encore Behaviour-driven development (BDD) sont des pratiques agiles très populaires.

**Algorithme** : c'est une suite finie d'opérations permettant de résoudre un problème. En informatique, les algorithmes sont les instructions données à la machine dans le but de réaliser une certaine tâche.

**ALT et ALT GR (Ang. Alternate)** : Touches du clavier permettant d'effectuer des fonctions autres que celles initialement prévues pour chacune des touches du clavier.

**Analytics** : un terme un peu valise qui regroupe toutes les techniques d'analyse de données numériques. Dans un contexte Big Data, on parle de Data Analytics pour désigner les activités qui permettent aux organisations d'étudier de grands ensembles de données.

**Analyse des données (Data Analysis)** : Ensemble de méthode de statique descriptive facilitant l'interprétation de gros corpus de données et très utilisées dans le Datamining.

Il comporte :

- des techniques d'analyse factorielle (analyse en composantes principales, analyse des correspondances, analyse discriminante etc.)
- des méthodes de classification automatique (segmentation etc.).

**Analogique** : Qui est représenté par la variation continue d'une valeur physique. S'oppose à numérique.

**Analyse heuristique** : Méthode d'analyse utilisée par les antivirus pour détecter d'éventuels virus non répertoriés qui ne figurent pas dans la base de données du logiciel. Notez que les méthodes heuristiques ne donnent pas de résultats assurés et peuvent parfois générer des fausses alertes.

**Annuaire d'entreprise (Personnel directory)** : Référentiel consacré à l'identification des agents de l'entreprise, et comportant les attributs permettant d'évaluer le profil de chaque agent et de définir ses habilitation.

**API** : c'est l'abréviation d'« Application Programming Interface », ou interface de programmation en français. Les APIs font partie du quotidien des développeurs : elles leur permettent d'accéder facilement aux services d'un logiciel préexistant. Grâce aux APIs, ils peuvent aisément intégrer ces services dans leur propre logiciel ou application. (A ne pas confondre avec l'Allocation Parent Isolé ou l'Alcoolisation Ponctuelle Importante !)

**Applet** : petite application indépendante du matériel et du logiciel utilisés, qui est téléchargée depuis un serveur de la toile mondiale et qui est exécutée localement au sein d'un logiciel de navigation.

**Application** : dans le langage courant, on utilise généralement ce terme pour parler des applications mobiles, pour smartphones. Toutefois, une application désigne tout programme informatique qui sert à réaliser une tâche spécifique à la demande d'un utilisateur. Un éditeur de texte, un navigateur ou un jeu vidéo par exemple sont des applications. A noter que le terme *logiciel* a un champ sémantique plus large et désigne tout ce qui n'est pas matériel (les parties tangibles d'un ordinateur).

**Applications mobiles (apps) :** Les applications mobiles ont été développées spécialement pour les appareils mobiles que sont les smartphones et les tablettes tactiles. Elles sont disponibles pour les environnements Android, Apple, Windows mobile.

**Application Programming Interface (API) :** Protocole standard de communication entre deux ou plusieurs langages de programmation différents, on entre des systèmes d'exploitation différents.

**Architecture à trois niveaux (Three Tiers Architecture) :** Génération récente de l'architecture Client Serveur qui comporte trois composants (un pour les données, un pour les traitements, un pour la présentation) pouvant résider dans des endroits différents. Typiquement : les données sur un mainframe central, les traitements sur un serveur local, la présentation sur le PC (attention : l'expression "trois tiers" est un faux ami).

**Architecture Client Serveur (Client-Server Architecture) :** Architecture informatique qui vise à utiliser au mieux les ressources en puissance de traitement et en mémoire pour l'exécution des applications, en tirant parti des moyens disponibles sur les PC. Typiquement, les données sont concentrées sur le mainframe, la présentation est faite par le PC, les traitements sont répartis entre les deux. Cf. architecture à trois niveaux.

**Assistant à maîtrise d'ouvrage :** Un « assistant à maîtrise d'ouvrage » (AMO) est un consultant externe ou une personne que la direction informatique a mises à la disposition d'une maîtrise d'ouvrage. Il assiste un maître d'ouvrage délégué ou un maître d'ouvrage opérationnel lors des périodes de pointe (par exemple lorsqu'il faut spécifier, suivre ou recetter un gros projet ; concevoir ou mettre en place des méthodes nouvelles ; mettre en œuvre une expertise que l'entité ne possède pas. Il rédige les spécifications ou documents méthodologiques, tient les tableaux de bord à jour etc.

**Automatisation :** élément central de la transformation numérique, l'automatisation consiste à remplacer les tâches réalisées par des humains grâce à des logiciels informatiques. C'est une des clés de la mise en place d'une démarche DevOps.

**Back-end :** Le back-end est la partie en arrière-plan d'un site internet, d'un système informatique ou encore d'un logiciel où les données sont stockées et traitées. Il est en opposition avec le front-end.

**Back-office :** Partie d'un site web, souvent dotée d'une interface, permettant de gérer ses fonctionnalités, ses contenus, etc. Le back-office n'est pas accessible au visiteur du site.

**Backlink :** Un backlink est un lien pointant vers une page web mais inséré sur une autre page. Si sur une page A, on insère un lien vers une page B, alors il s'agit d'un backlink de la page B.

**Balise :** Une balise est une série de caractères permettant de déclencher l'exécution d'une action par un programme informatique. Les plus courantes sont les balises HTML, destinées à la création de pages web (structure et hiérarchisation du contenu).

**Base de données (Database) :** Logiciel qui permet de stocker, classer, retrouver des données, et de réaliser des calculs sur ces données.

**Base documentaire (Documentary Database) :** Logiciel qui permet de stocker, classer, retrouver des documents.

**Bêta :** La version bêta, ou bêta, désigne la version test d'un produit informatique avant sa commercialisation ou distribution. Elle succède à la version alpha.

**Big Data** : Le Big Data désigne des ensembles de données si volumineux qu'ils dépassent les capacités d'analyse humaines ou d'outils informatiques classiques de gestion de base de données ou de l'information.

**Blockchain** : La blockchain est une technologie de stockage et de transmission d'information sans organe de contrôle, supposée plus sécurisée. Elle fut à la base inventée pour le « bitcoin ».

**Bot** : Un bot est un programme informatique opérant de manière autonome et automatique. Il est souvent utilisé pour discuter avec des humains et effectuer des tâches de manière répétée. « Bot » vient du mot « robot ».

**Browser** : Logiciel offrant une interface d'accès aux documents HTML disponibles sur le World Wide Web d'Internet. Netscape, Microsoft Explorer sont des Browsers. Ils permettent de " surfer sur le Web ".

**Bureautique (Office Automation)** : Ce terme a d'abord désigné l'équipement électronique de bureau (photocopieuses, machines à calculer). Puis il s'est appliqué au micro-ordinateur et à ses premières applications (traitement de texte, tableur, grapheur) ainsi qu'aux imprimantes. Avec la mise en réseau des PC, on est passé à la fin des années 80 à la " bureautique communicante " dont les premières applications ont été la messagerie et l'agenda partagé. La bureautique communicante a été vers 1992 englobée sous le concept de groupware.

**Bus** : Dans un ordinateur ou un réseau, le *bus* est un support de transmission sur lequel les signaux sont émis ou reçus par chacun des éléments raccordés au réseau. Seuls ceux auxquels un signal est adressé le traitent effectivement ; les autres le négligent. Le mot " bus " vient de la ressemblance avec les autobus que des passagers peuvent prendre ou quitter à chaque arrêt. Sur un réseau, le bus relie des ordinateurs ; dans un ordinateur, le bus relie la carte mère avec les cartes insérées dans les slots d'extension (lecteurs de disques durs ou de CD-Roms, adaptateurs graphiques, cartes son etc.)

**Business Process Reengineering (BPR)** : Analyse des processus de l'entreprise visant à les rendre plus efficaces en termes de délais, de coût et de qualité, et procédant par remise en question systématique des règles et procédures de l'entreprise.

**Cache** : Ce terme désigne le fait de stocker du contenu relatif aux recherches effectuées sur les moteurs de recherches dans les bases de données des navigateurs web. Ces données en cache peuvent être utilisées par l'utilisateur afin de rapidement trouver le résultat de sa recherche. Le cache d'un navigateur (ou d'un logiciel) est remis à jour régulièrement.

**CAO** : un « logiciel de CAO » consiste en un outil informatique de conception permettant de réaliser des objets complexes comme de décrire des objets simples avec une grande précision. Le logiciel en question vise à simuler la forme, le fonctionnement ou le comportement de l'objet dans ses trois dimensions.

**Chatbot** : Un chatbot est un robot conversationnel capable de répondre à un humain par l'intermédiaire d'une messagerie instantanée, fournissant une réponse immédiate à la demande d'un utilisateur. Il comprend le langage humain, à l'appui de l'intelligence artificielle. On peut distinguer deux types de bots : ceux qui répondent de manière basique à une question posées (FAQ), et ceux qui font appel à une application tierce pour proposer un service (exemple : suivi d'une commande).

**Chef de projet maîtrise d'œuvre** : Le chef de projet maîtrise d'œuvre (CP MOE) assure la gestion de la maîtrise d'œuvre d'un projet particulier. C'est un chef de chantier avec tout ce que cela demande

de compétence, de savoir-faire et d'énergie : il coordonne les entreprises, vérifie la qualité des fournitures, veille au respect de la cohérence entre les diverses livraisons, etc. Son interlocuteur naturel est le maître d'ouvrage opérationnel.

**Circuit intégré (Integrated Circuit) :** Ensemble de composants électroniques fixés, ainsi que leurs interconnexions, sur une " puce " (le plus souvent en silicium).

**Client :** Dans le langage de l'informatique, un " client ", ce n'est pas une personne, mais le PC de l'utilisateur. C'est ainsi qu'il faut comprendre l'expression " client-serveur ".

**Cloud computing : en Français « informatique en nuage »,** correspond à l'accès à des services informatiques (tels que des serveurs, le stockage, la mise en réseau, les logiciels) via Internet (« le cloud ») à partir d'un fournisseur. Le cloud computing est un modèle permettant d'avoir recours à des ressources informatiques et des services à distance à la demande par Internet sans les posséder dans des datacenters). Au sein du cloud computing, les ressources et services sont partagés, selon un certain quota d'utilisation mesurable, plutôt que via le disque dur de son ordinateur

Les principaux services proposés en *cloud computing* sont le SaaS (Software as a Service), le PaaS (Platform as a Service) et le IaaS (Infrastructure as a Service) ou le MBaaS (Mobile backend as a service - lien en anglais). Généralement sur trois niveaux, le Cloud Public - accessible par Internet, le Cloud d'entreprise ou Privé - Accessible que sur un réseau privé, le Cloud intermédiaire ou hybride qui est un mix entre le Cloud public et le Cloud Privé.

**CMS :** CMS est l'acronyme de content management system, soit, en français, « système de gestion de contenu ». Il s'agit d'un programme informatique s'appuyant sur une base de données et permettant de gérer les éléments relatifs à l'apparence et le contenu d'un site web. Il permet, à des individus, de mettre à jour le contenu d'un site web à partir d'un panneau d'administration. Cette solution donne la possibilité de créer des sites web complexes à coût réduit, grâce aux nombreux CMS Open Source existant sur le marché. On peut citer par exemple WordPress, Drupal ou Joomla! comme CMS OpenSource. Des CMS e-commerce existent également comme Prestashop, qui permet de créer des boutiques en ligne.

**Commerce électronique (Electronic Commerce) :** Réalisation du processus de la relation commerciale par voie électronique (réseau, ordinateurs) : présentation des produits, prise de commande, paiement, gestion de la logistique de livraison. Dans le cas des documents ou des logiciels, la livraison elle-même peut se faire par voie électronique.

**Compilateur (Compiler) :** Logiciel qui traduit le programme, écrit par le développeur dans le " langage source ", en une série d'instructions interprétables par l'ordinateur. Un programme compilé est généralement incompréhensible pour un être humain.

**Cookie :** Ce terme se réfère à l'équivalent d'un fichier texte de petite taille, stocké sur le terminal de l'internaute. Les cookies permettent aux développeurs de sites web de conserver des données utilisateur afin de faciliter la navigation et de permettre certaines fonctionnalités.

**CRM :** Le CRM (Customer Relationship Management) ou gestion de la relation client, désigne une stratégie de gestion des relations et interactions d'une entreprise avec ses clients ou clients potentiels. Un système CRM permet aux entreprises d'interagir en continu avec les clients, à rationaliser leurs processus et à améliorer leur rentabilité. L'objectif d'un système CRM est d'améliorer les relations commerciales.

**Cybersécurité** : La cybersécurité consiste en un ensemble de technologies, de processus et de pratiques qui ont pour but de protéger les réseaux, ordinateurs et données contre les attaques, dommages et accès non-autorisés. En informatique, le terme « sécurité » regroupe la cybersécurité et la sécurité physique. Une des problématiques de la cybersécurité tient à la constante et rapide évolution des risques pour la sécurité.

**Datacenter** : littéralement *centre de données*, le datacenter est un lieu regroupant des équipements informatiques (ordinateurs, serveurs, équipements réseaux, baies de stockage...) d'une ou plusieurs organisations. Typiquement, les data centers sont utilisés par les entreprises pour stocker, traiter et organiser de grandes quantités de données. Ils doivent donc répondre à un certain nombre d'exigences en matière de sécurité, surveillance, disponibilité ou encore consommation d'énergie.

**Data Mart** : " Magasin de données " contenant une partie des données de l'entreprise (par produit, département ou activité), organisées sous *une* base de données logique et éventuellement *plusieurs* bases de données physiques.

**Datamining** : Méthodes reposant sur les techniques d'analyse des données et permettant, à partir de phénomènes constatés sur des données agrégées, d'identifier les causes au niveau des variables détaillées. Exemple : repérer l'agence de voyage dont le comportement explique la part de marché d'Air France sur une zone donnée.

**Data scientist** : c'est le scientifique de la donnée. Comme son nom l'indique, il ou elle est chargé-e d'exploiter des données numériques, via des méthodes scientifiques, en vue d'en tirer une certaine valeur. L'idée étant généralement de prendre des décisions stratégiques éclairées, grâce à une analyse poussée de données brutes et parfois très hétérogènes. Les récents progrès en matière de stockage, traitement ou encore visualisation de l'information ont poussé le métier de data scientist sur le devant de la scène : la *Harvard Business Review* l'intronise même « job le plus sexy du XXIème siècle ».

**Datawarehouse** : Méthodes d'agrégation et consultation permettant de calculer et mettre à la disposition de l'utilisateur, sous une forme commode, les séries chronologiques dont il a besoin ainsi que les outils d'analyse de ces séries. Le data warehouse, littéralement *entrepôt de données* est une base de données relationnelle (c-à-d où l'information est organisée en tableaux à deux dimensions, généralement lignes et colonnes, et où les relations entre les données jouent un rôle déterminant en vue de leur exploitation). Les data warehouse sont largement utilisés par les entreprises pour faciliter l'analyse de données et les activités de Business Intelligence.

**Déploiement** : dans un contexte de projet informatique, le déploiement désigne toutes les activités destinées à rendre un logiciel disponible à l'utilisation. Puisque chaque logiciel a ses spécificités, les processus exacts de déploiement varient d'un projet à l'autre. Grâce aux technologies Cloud, un logiciel peut désormais être déployé chez un grand nombre de clients en seulement quelques minutes. Cette nouveauté a fait émerger une nouvelle forme de déploiement, dit en continu : les cycles de développement sont raccourcis et les tests automatisés, pour que chaque modification du logiciel soit rapidement propulsée vers ses utilisateurs finaux.

**Digital** : digital learning, marketing digital, économie ou même culture digitale, les exemples d'utilisation du terme ne manquent pas. Elles sont pourtant... toutes fausses ! Du moins si l'on s'en tient à la définition donnée par l'Académie française : l'adjectif digital veut simplement dire « qui appartient, se rapporte aux doigts ». La confusion vient de l'anglais *digit*, qui signifie « chiffre », et *digital*, « qui utilise des nombres ». En bon français, il faudrait donc dire « numérique », même si le sujet fait encore débat.

**Disque dur (Hard Disk) :** Mémoire de masse d'un ordinateur, à accès lent, résidant sur un support magnétique. Contrairement à la RAM, le disque dur ne s'efface pas lorsque l'ordinateur s'éteint.

**DNS :** cet acronyme désigne le Domain Name Service, ou système de noms de domaine. Sans système DNS, il n'y aurait pas de Web tel qu'on le connaît aujourd'hui. En effet, le DNS permet d'associer un nom de domaine (l'adresse d'un site Internet, en .com, .fr, .net...) à une adresse IP. Pour consulter un site, le nom de domaine ne suffit pas : tout matériel réseau connecté à Internet est accessible via une certaine adresse IP. Le DNS est donc une sorte d'annuaire géant associant chaque nom de domaine à une adresse IP, et permettant ainsi la connexion et l'échange de données entre deux ordinateurs ou serveurs.

**Documentation partagée (Shared Documentation) :** Documentation située sur un serveur et accessible aux utilisateurs autorisés. Pour trouver le document qui l'intéresse, l'utilisateur dispose de divers outils de recherche: index, " vues " présentant des listes de documents selon un classement convenu, recherche sur des mots contenus dans les textes. La documentation partagée est le meilleur moyen pour diffuser les notes administratives, tenir à jour les documentations techniques, et mettre les supports de formation à la disposition des anciens stagiaires.

**Donnée (Data) :** Une donnée, c'est un couple logique formé par (a) une définition, (b) la spécification d'une méthode de mesure, d'observation ou de calcul (" Métadonnée "). Réaliser la mesure (ou l'observation pour les données qualitatives, ou le calcul pour les données agrégées) permet de connaître la valeur prise par la donnée dans un contexte particulier (lieu, date).

**DSI :** le Directeur, ou la Directrice d'ailleurs, du Système d'Information. Au sein d'une organisation, c'est la personne responsable de l'ensemble des ressources informatiques, au sens matériel et logiciel. Postes de travail, serveurs, infrastructures de stockage, de sauvegarde, d'impression, mais aussi applications, données, équipements réseaux ou télécommunications, les DSI gèrent tous les systèmes informatiques nécessaires au bon déroulement des activités de l'organisation. Occupant au départ un rôle essentiellement technique, le DSI est aujourd'hui un acteur clé de la stratégie d'une entreprise et de ce fait souvent subordonné à la direction générale.

**Dynamic Data Exchange (DDE) :** Technique permettant d'insérer dans un document des données provenant d'autres applications.

**E-commerce :** Le e-commerce correspond à de la vente en ligne réalisée par le biais d'un site web marchand.

**Econométrie (Econometrics) :** Méthode permettant de tester des relations de causalité en se fondant sur l'examen des corrélations entre données. Certaines des relations ainsi obtenues peuvent être utilisées à des fins de prévision ou de simulation. L'économétrie est le complément naturel de l'analyse des données.

**EDI (Échange de données informatisé) (Electronic Data Interchange) :** Communication entre applications informatiques d'entreprises différentes ou non, par le moyen de messages dont le format et le codage auront été fixé par un accord d'interchange. La norme Edifact définit le format général des messages. L'EDI facilite les échanges répétitifs d'information (commande, facture, etc.) entre un fournisseur et un client assidu. L' " EDI-ouvert " procure une norme pour décrire les rôles de diverses entreprises dans les montages complexes d'ingénierie d'affaire, et garantit que leurs applications sont capables de communiquer par messages EDI.

**Éditeur logiciel** : Un éditeur logiciel est une entreprise qui conçoit, développe et commercialise un produit logiciel.

**ECMA International** : ECMA International (European association for standardizing information and communication systems) est une organisation de standardisation dans les systèmes d'information et de communication (ex : langages de programmation, technologies de communication).

**ECMAScript** : ECMAScript est un ensemble de normes pour les langages de programmation de type script (ex : JavaScript), standardisée par ECMA International.

**Enterprise Application Integration (EAI)** : L'expression EAI désigne les projets, méthodes et outils qui visent à coordonner et consolider les applications informatiques d'une entreprise. Une entreprise qui a par exemple des applications et bases de données anciennes, et qui veut continuer à les utiliser tout en mettant en oeuvre de nouvelles applications qui utilisent l'Internet (e-commerce, Extranet).

Toutes les grandes SSII ont une offre d'EAI. Pour les mettre en oeuvre, elles utilisent des techniques récentes : langages orientés objet, produits de middleware, brokers de messages, interface avec gestion de files d'attente, accès aux données avec le langage XML etc.

**Enterprise Resource Planning (ERP)** : L'expression ERP désigne un logiciel comprenant divers modules qui permettent à une entreprise de gérer d'importantes parties de ses affaires : planification de la production, achats de produits intermédiaires, gestion du stock de pièces de rechange, relations avec les fournisseurs, services aux clients, suivi de l'exécution des commandes, et aussi comptabilité et gestion des ressources humaines. Un ERP est fondé sur l'utilisation d'une base de données relationnelle. La mise en place d'un ERP peut nécessiter un important travail d'analyse et transformation des processus de travail, ainsi que de formation des agents. Les grands fournisseurs d'ERP sont SAP, Peoplesoft, J. D. Edwards, Oracle etc.

**Ergonomie web** : L'ergonomie web est la discipline s'intéressant au confort, à la sécurité et à l'efficacité sur un site web. On l'utilise pour optimiser le parcours de l'utilisateur sur le site en s'intéressant, par exemple, à la facilité d'utilisation, la rapidité d'exécution ou encore la lisibilité.

**ESN** : Une entreprise de services numérique (ESN) est une société de services experte en nouvelles technologies et en informatique.

**Espace virtuel** (Virtual Space) : Espace à trois dimensions reconstitué par ordinateur, et dans lequel l'utilisateur peut se déplacer en utilisant des outils de pilotage (manche à balai ou " joy-stick ", lunettes pour la vision en relief, gants pour la sensation tactile), voire participer à des réunions avec d'autres utilisateurs représentés par leurs images virtuelles.

**Ethernet** : Protocole de transmission de données sur réseau local, procédant par envoi de trames diffusées sur le réseau et dont l'adresse est reconnue par le récepteur.

**ETL** : De l'anglais Extract, Transform, Loading soit Extraire, Transformer, Charger. Il s'agit d'un système de chargement de données. Il commence par extraire les données, puis les dénormalise, les nettoie, les contextualise pour enfin les charger. Le but de ce système est d'éviter les erreurs d'informations.

**Executive Information System (EIS)** : Système présentant, sous une forme graphique claire et commentée, des indicateurs sélectionnés " sur mesures " pour un ou des responsables de l'entreprise.

**Extension de fichier** : L'extension permet d'identifier le type de fichier. Il s'agit de la dernière partie du nom d'un fichier, précédée par un point (ex : .doc).

**Extranet** : Mise en réseau de plusieurs entreprises qui connectent leurs Intranets. L'Extranet est le moyen idéal pour les relations avec les partenaires, fournisseurs et clients.

**File Transfer Protocol (FTP)** : Protocole de transfert de fichiers utilisé sur l'Internet.

**Firewall** : Ce terme désigne un logiciel qu'on installe sur un ordinateur afin de le protéger de toute menace. Il peut inclure ou non un antivirus selon les modèles. Exemples de firewall : Arkoon Network Security, Nortel etc.

**Flat webdesign** : Le flat webdesign (ou design plat), est un style graphique minimaliste utilisé dans la conception d'interfaces utilisateurs (UI). Le flat webdesign est très présent dans les interfaces web et mobile et renvoie à certains termes qui reviennent souvent : simple, épuré, moderne, etc.

**Flux RSS** : Les flux RSS (également appelé fil RSS) permettent de suivre l'actualité d'un site sans être obligé d'aller le visiter. Le fichier RSS a la possibilité d'être lu par la plupart des navigateurs web récents, par des applications web spécialisés ou encore par des logiciels installés sur votre ordinateur. Certains sites peuvent reprendre de façon automatique des flux RSS d'autres sites.

**Font stack** : Le terme font stack (traduction : pile de polices) est relatif à l'ordre d'affichage des polices de caractères. Autrement dit, il s'agit d'une liste ordonnée de plusieurs polices. Sur un site web, si un visiteur ne possède pas la première police de la pile, alors ce sera la deuxième qui sera utilisée. et ainsi de suite.

**Forum (News)** : Messagerie dans laquelle l'accès à la boîte aux lettres est ouvert à plusieurs personnes autorisées. Questions et réponses sont présentées à la suite. Les forums sont un bon moyen pour les relations entre utilisateurs " de base " et experts. Un forum doit être suivi par un animateur, qui le purge des messages sans intérêt, et veille à ce que les questions reçoivent assez rapidement une réponse. Le forum est un complément naturel de la documentation partagée pour le suivi des formations.

**Framework** : En programmation informatique, il s'agit d'un ensemble d'outils et de composants logiciels à la base d'un logiciel ou d'une application. C'est le framework qui établit les fondations d'un logiciel ou son squelette applicatif. Les développeurs utilisant le framework peuvent l'enrichir pour une meilleure amélioration. Il a pour objectif de simplifier et d'uniformiser le travail des développeurs et son utilisation nécessite d'avoir un profil expérimenté. On peut citer les frameworks Java et les frameworks Web.

**Front end** : Ce terme correspond aux éléments d'un site que l'on voit à l'écran et avec lesquels on peut interagir depuis un navigateur. Les éléments visibles sur un site web sont une combinaison de HTML, CSS et JavaScript. Ces langages de programmation utilisés par le développeur Front End sont interprétés par le navigateur de votre ordinateur pour afficher un résultat visuel. Il s'agit entre autres de polices, de menus déroulants, de boutons, etc.

**FTP** : Cela désigne un protocole ayant pour objectif le transfert de fichiers présents sur ordinateur vers un serveur, ou l'inverse. Le FTP permet d'envoyer de lourds fichiers incompressibles et qui ne peuvent être envoyés d'une autre façon.

**Full stack** : Le full stack correspond à l'ensemble des niveaux techniques de la pile des différentes couches qui constituent une application informatique. De façon plus précise, il peut s'agir de la configuration de l'infrastructure matérielle et du système d'exploitation, de la conception et la gestion des bases de données, de la conception du code « back-end » et « front-end » à l'appui de langages tels que Python ou Javascript.

**Geek** : Prononcer "guic". Le geek est un programmeur passionné, expérimenté, qui connaît bien les possibilités et pièges des divers langages, et l'art de s'en servir pour répondre aux besoins des utilisateurs - ou pour son propre plaisir. Il ne compte pas les heures qu'il passe devant l'ordinateur, consomme beaucoup de café, a le plus souvent des lunettes, porte volontiers la barbe, préfère se vêtir d'un jean et d'un T-shirt que d'un costume-cravate. La version péjorative du geek, c'est le "nerd".

**Génie logiciel** : c'est la discipline qui applique au développement de logiciels les principes structurés et hiérarchisés de l'ingénierie. De la conception à la production, en passant par les phases de développement, test et déploiement, le génie logiciel regroupe l'ensemble des activités qui visent à rationaliser la production de logiciels. On parle généralement de génie logiciel dans le cas de projets complexes et de grande envergure (et non pour des programmes ou applications simples).

**GNU** : un système d'exploitation de type Unix, entièrement libre et dont le développement a démarré en 1984 sous le nom de projet GNU. Généralement, GNU est utilisé avec un noyau Linux, qui alloue les ressources de la machine et communique avec le matériel : cette association constitue le système d'exploitation GNU/Linux, utilisé par des millions de personnes mais souvent appelé simplement « Linux ».

**GPS ("Global Positioning System")** : Le GPS est un système de navigation utilisant un réseau de satellites militaires américains dont l'accès est autorisé aux civils. Il donne la position et l'altitude chaque seconde, 24 heures sur 24, n'importe où dans le monde.

**Gouvernance informatique** : c'est l'ensemble des méthodes, outils et bonnes pratiques mis en œuvre par la direction d'un système d'information (DSI) afin d'atteindre les objectifs liés à la stratégie de l'organisation. La gouvernance définit ainsi la manière dont le SI participe à la création de valeur et délimite le rôle des différentes parties prenantes.

**Graphheur** (Computer Graphics) : Logiciel qui permet de faire des dessins avec un micro-ordinateur.

**Green IT** : en français « informatique durable » ou encore « numérique responsable », le Green IT désigne le concept visant à réduire l'impact environnemental, économique et social des technologies de l'information. C'est le développement durable, appliqué à l'IT.

**Groupware** : Ce terme (en français "collectique") désigne l'ensemble des logiciels qui facilitent le travail de groupe, quelle que soit la localisation géographique des participants : messagerie, documentation partagée, agenda partagé, rédaction coopérative, workflow, etc. La plupart des applications de Groupware se retrouvent maintenant sur l'Intranet.

Ainsi le terme groupware renvoie à des applications diverses et variées concourant à un même but : permettre à des utilisateurs géographiquement éloignés de travailler en équipe. Le travail en équipe peut se concrétiser par le partage d'informations, ou bien la création et l'échange de données informatisées. Il s'agit pour la plupart du temps d'outils de messagerie (instantanée ou non), ainsi que d'applications diverses telles que :

- Agenda partagé,
- Espace de documents partagés,
- Outils d'échange d'information (forums électroniques),
- Outil de gestion de contacts,
- Conférence électronique (vidéoconférence, chat,...)

**GUI** : l'acronyme de Graphical User Interface, ou « interface graphique » en français. En informatique, une interface graphique désigne le dispositif de dialogue homme-machine, soit l'ensemble des éléments qui s'affichent à l'écran et permettent à un utilisateur d'interagir avec un logiciel (menus, boutons, fonctionnalités...).

**Hackathon** : Le nom « hackathon » provient de la contraction entre « hack » et « marathon ». Il s'agit d'un événement où se rassemblent des personnes afin de collaborer, de rencontrer d'autres profils professionnels, et de produire un prototype souvent numérique.

**Header** : le terme « header » peut être traduit par « en-tête » en français. Dans un site web, il s'agit de la partie située en haut de la page, qui contient en général le menu et le nom/logo du site.

**Hyper Text Markup Language (HTML)** : Langage de balisage de documents permettant d'y insérer des liens hypertextes.

**Hyper Text Transfer Protocol (HTTP)** : Protocole de transmission des documents HTML sur l'Internet.

**Hypertexte** (Hyper Text) : Création dans un document de liens, signalés par des icônes, des textes soulignés ou en couleur, contenant l'adresse d'un autre document qui est appelé et chargé lorsque l'on clique sur le lien. L'hypertexte est à l'origine de la création et du succès du WWW.

**Indicateur** (Indicator) : Donnée ou agrégat de données sélectionné pour son intérêt économique, mis sous forme de série chronologique, corrigé des variations saisonnières après interprétation des incidents, ayant fait l'objet d'une modélisation économétrique permettant de l'interpréter et de fournir des prévisions tendancielle (ou extrapolations). La production d'un indicateur est une opération relativement coûteuse : seules certaines données choisies méritent d'être élevées au statut d'indicateur.

**Infocentre** : Outil d'interrogation de bases de données et de reporting. L'Infocentre est plus pauvre que le Datawarehouse : les données ne sont pas mises sous forme de séries chronologiques, et il n'est pas aussi bien outillé pour produire les agrégats souhaités.

**Information** : Une information, c'est une donnée observée par un acteur que cette donnée intéresse. L'observation par un acteur implique la comparaison au moins implicite à d'autres données, car sans comparaison il n'y a pas d'interprétation possible. Passer du rang de *donnée* à celui d'*information* suppose que la connaissance de la donnée contribue à l'action de celui qui l'observe : la notion d'information recèle donc un côté subjectif et un côté objectif.

Dans la " théorie de l'information ", le mot information est pris dans une autre acception que celle-là.

**Informatique** : Le mot " informatique " a été créé en 1962 par Philippe Dreyfus, ingénieur chez Bull, pour traduire l'anglais " computer science ". Il est construit par contraction des mots " information " et " automatique ".

**Intelligence artificielle (IA)** (Artificial Intelligence (AI)) : Méthode de conception de programmes (" systèmes experts ", " systèmes interactifs d'aide à la décision (SIAD) ") automatisant une partie d'un processus de décision. Plusieurs procédés sont utilisables, dont l'un consiste à analyser les procédés des meilleurs experts et à en reproduire le cheminement dans le programme. Le plus souvent, il n'est pas possible d'automatiser complètement le processus de décision, et le système expert offre à l'utilisateur une *assistance* en automatisant les phases répétitives de sa tâche (on dit que le travail est " assisté par ordinateur "). Ainsi la traduction automatique d'un texte littéraire est hors de la portée d'un ordinateur, alors que la traduction assistée est très efficace.

**Interface** : Mise en forme des données permettant leur passage d'une étape à l'autre du traitement. Equipement assurant la transcription des données d'un langage dans un autre. L'interface homme-machine assure la communication entre l'homme et l'ordinateur grâce à des supports (écran, clavier, haut parleurs) accessibles aux sens de l'être humain. L'interface graphique (Graphical User Interface, ou GUI), qui permet d'afficher et de créer des images, fait partie de l'ergonomie standard en 1997.

**Internet** : Réseau d'interconnexion d'ordinateurs utilisant le protocole de transmission de données TCP/IP (qui permet aussi le transfert des images, fixes ou animées, et du son). Efficace, robuste et peu coûteux grâce aux qualités de TCP/IP, l'Internet a été d'abord utilisé par des chercheurs, puis a servi de support à des services devenus populaires (messagerie, forums, Web, commerce électronique, téléchargement de logiciels) qui ont fait de lui un phénomène de société. C'est désormais *le* réseau mondial de communication électronique.

**Interpréteur** (Interpreter) : Logiciel qui transcrit un programme d'un langage dans un autre en procédant instruction par instruction et en les faisant exécuter immédiatement (le traducteur, par contre, transcrit l'intégralité du programme avant l'exécution).

**Intranet** : Utilisation de l'Internet à des fins internes à une entreprise. L'Intranet permet à l'entreprise de bénéficier de l'économie d'échelle acquise par les logiciels sur l'Internet, et d'outils de développement orientés-objet comme Java, C#. On peut réaliser maintenant sur l'Intranet la totalité des applications de groupware. L'Intranet nécessite toutefois une administration soigneuse des droits d'accès, et la mise en place de " Firewalls " pour protéger les données de l'entreprise.

**ISO** : L'International Standard Organisation (ISO), est un organisme de normalisation internationale.

**Firewall** : Logiciel (pare-feu) qui sert à protéger un réseau contre les tentatives d'intrusion.

**Java** : Langage de programmation orienté objet développé par Sun (maintenant maintenu et développé par la société Oracle) et bien adapté à l'échange de petites applications (Applets) sur l'Internet. Java est d'abord une version simplifiée et sécurisée de C++. Le PC de l'utilisateur doit être équipé d'une couche logicielle dite " machine virtuelle ", qui permet notamment de le protéger des virus. L'utilisateur appelle l'applet en cliquant sur un lien hypertexte, et celle-ci est lancée immédiatement après son téléchargement.

**KPI** : Les KPI (Key Performance Indicators) ou indicateurs clés de suivi de la performance, consistent en des regroupements stratégiques de données issues du suivi des interactions et comportement des utilisateurs (tracking) à des fins d'évaluation. Les Kpis se mesurent en taux. Il peut s'agir par exemple du ratio d'utilisateurs par type de terminal (mobile/fixe desktop), du taux d'engagement, du taux de rebond sur une page, du coût au clic, du taux d'ouverture, du taux de réactivité.

**Landing page** : La landing page, ou page d'atterrissage, est une page destinée à recevoir des visiteurs et les faire s'engager à quelque chose. Cette page est optimisée en SEO dans le cadre d'une requête

particulière, avec un travail sémantique spécifique et un design épuré et rassurant au niveau des éléments. Elle contient des liens internes sur une thématique similaire. Une landing page peut être utilisée dans un contexte de campagne publicitaire et doit inciter à l'action.

**Langage de programmation (Programming Language) :** L'ordinateur ne comprend que des instructions simples : chercher des données dont il connaît l'adresse en mémoire, faire une opération arithmétique sur ces données, stocker le résultat de cette opération en mémoire, etc. Le langage qui permet d'écrire ces instructions est le langage machine.

Programmer en langage machine serait une tâche fastidieuse pour un être humain. Les langages de programmation offrent des instructions plus synthétiques et commodes. En fait, l'exécution d'un programme écrit en " langage de haut niveau " (Visual Basic, Java, C++ etc.) nécessite une cascade de traducteurs et interpréteurs pour aboutir à des instructions exécutables par la machine.

Les langages de programmation sont ainsi plus proches du langage " naturel " (ou de sa représentation graphique). Ils restent cependant très conventionnels, et leur utilisation experte suppose une formation approfondie.

**Langage de quatrième génération (L4G) (Fourth Generation Language ) :** Langage de programmation facilitant la conception des applications de type client-serveur, et facilitant les tâches de gestion documentaire et de gestion de configurations.

**Liaison louée (Leased Line) :** Support de transmission loué par un opérateur télécoms pour relier deux établissements distants. On dit parfois encore LS (liaison spécialisée). Le catalogue comporte une offre variée en termes de débit. Le prix dépend du débit et de la distance.

**Lien hypertexte :** un lien hypertexte (ou hyperlien) correspond à un élément placé dans le contenu d'une page Web, et permettant, en cliquant dessus, d'accéder à un autre contenu sur le même site Web (lien interne) ou à un site Web différent (lien externe). Ce lien hypertexte prend généralement la forme d'un texte souligné et d'une autre couleur, et peut également être placé sur une image, un bouton, une photo, etc.

**Learning management system (LMS) :** cela désigne un logiciel visant à accompagner et gérer un processus d'apprentissage ou un parcours pédagogique. En français, on parle notamment de « plateforme d'apprentissage », ou « système de gestion de l'apprentissage ». Ce type de système informatique propose un espace numérique de travail (ENT) comprenant des tests d'évaluation qui sont soit soumis à validation par l'enseignant, soit proposés comme activités de régulation en auto-évaluation.

**Logiciel (ou application) :** cela consiste en un programme apportant à l'ordinateur un ensemble de fonctionnalités supplémentaires, qui ne sont pas forcément présentes à l'origine. Un logiciel peut s'installer sur l'ordinateur via un disque (CD, DVD) ou directement en téléchargement sur Internet. Il existe des logiciels gratuits et d'autres payants. Dans l'informatique, on trouve des logiciels dans tous les domaines, (exemples : comptabilité, gestion de projet, enquêtes, etc).

**Logiciel libre :** il s'agit d'un logiciel qui est distribué selon une licence libre. Un logiciel libre est un logiciel qui peut être utilisé, modifié et redistribué sans restriction par la personne à qui il a été distribué. Les logiciels libres se caractérisent par une certaine fiabilité et réactivité, étant donné qu'ils sont soumis à étude, critique et correction. Mozilla Thunderbird, OpenOffice.org et VLC sont des exemples de logiciels libres.

**Machine Learning** : un domaine de l'informatique dérivé de l'intelligence artificielle et qui permet aux systèmes d'« apprendre », sans avoir été programmés à cet effet. Les technologies de Machine Learning, ou apprentissage automatique, absorbent un grand volume de données, afin d'en extraire des modèles de prédiction et de prise de décision.

**Maître d'œuvre** (Prime Contractor) : Personne morale (entreprise, direction etc.) garant de la bonne réalisation technique des solutions. Il a un devoir de conseil vis-à-vis du Maître d'ouvrage. En cas de pluralité de fournitures, il veille à leur cohérence et à la qualité des interfaces. Il coordonne l'action des fournisseurs en optimisant la qualité technique, en minimisant les risques et en assurant le respect des délais fixés par le Maître d'ouvrage. Il valide la recette technique des solutions.

**Maître d'ouvrage** (Customer) : Personne morale (entreprise, direction etc.) chargée d'une mission.

Le maître d'ouvrage est responsable de l'efficacité de son organisation, de ses méthodes de travail et de son système d'information. Il fait appel à des maîtres d'œuvre pour obtenir les solutions qui lui permettent de réaliser sa mission. Il leur fournit les spécifications fonctionnelles, et valide la recette fonctionnelle de ces solutions.

Les fonctions du maître d'ouvrage sont remplies par des personnes ou des équipes : le "maître d'ouvrage stratégique" est le patron de la personne morale considérée ; le "maître d'ouvrage délégué" est l'expert qui assiste le maître d'ouvrage stratégique pour lui permettre d'exercer ses responsabilités ; le "maître d'ouvrage opérationnel" est un expert spécialisé dans l'automatisation d'un processus (voir ci-dessous les définitions).

**Maître d'ouvrage délégué** (Business Technologist) : Le « maître d'ouvrage délégué » (MOAD) est une personne, soit seule, soit à la tête d'une équipe, qui veille à la qualité du SI de l'entité tant du point de vue de la conception que de la façon dont il est utilisé. Il assiste le maître d'ouvrage stratégique en lui fournissant les éléments nécessaires à la décision et à l'« alignement stratégique » du SI.

Ses interlocuteurs naturels au sein de l'entité sont les chefs de service et les maîtres d'ouvrage opérationnels. Son interlocuteur naturel au sein de la direction informatique est le responsable de domaine.

**Maître d'ouvrage opérationnel** : Le « maître d'ouvrage opérationnel » (MOAO) est, dans l'entité, un expert qui connaît à fond l'un des grands processus du métier. Il recueille les demandes des utilisateurs et établit ou supervise les spécifications générales. Il a pour interlocuteur naturel à la direction informatique le « chef de projet maîtrise d'œuvre ».

**Maître d'ouvrage stratégique** : Le « maître d'ouvrage stratégique » (MOAS) du SI d'une entité est le dirigeant de cette entité : PDG ou DG pour l'entreprise, DGA ou directeur pour une direction de l'entreprise, chef de service pour les unités qui constituent la direction etc. Le SI étant à la fois la concrétisation de la stratégie et la condition de sa mise en œuvre, le MOAS prend les décisions essentielles concernant la maîtrise d'ouvrage (notamment pour le lancement des grands projets). Il arbitre les différends entre ses collaborateurs et signe le contrat avec la maîtrise d'œuvre.

Le MOAS n'est pas, en général, un expert en matière de SI. Il se fait assister par un maître d'ouvrage délégué (MOAD) auquel il *délègue* (c'est le sens même de l'expression) l'expertise en matière de SI.

**Matériel informatique** : cela désigne tous les éléments tangibles d'un ordinateur ou d'un réseau. Disque dur, carte mère, mémoire RAM, imprimante, écran ou même câble, le matériel (*hardware* en anglais) se distingue du logiciel (*software*).

**Mémoire** (Memory) : Support magnétique ou électronique comportant des zones dotées d'adresses, et où un ordinateur peut stocker données et programmes. Une mémoire est d'autant plus chère que son accès est plus rapide (d'où l'utilisation d'un disque dur à accès lent pour la mémoire de masse, et d'une RAM à accès rapide pour les travaux en cours).

**Messagerie électronique** (E-mail) : Système d'adressage et de stockage permettant à des utilisateurs d'échanger des messages en mode asynchrone : le message est stocké dans la " boîte aux lettres " (BAL) de l'utilisateur en attente de sa consultation. L'introduction d'une messagerie a un effet puissant sur l'organisation d'une entreprise.

**Métadonnée** (Metadata) : Spécification d'une méthode de mesure, d'observation ou de calcul d'une donnée. La documentation des métadonnées est l'un des éléments importants de la construction d'un Datawarehouse.

**Modulation par impulsion et codage (MIC)** : Codage numérique du signal sonore (64 kbit/s) qui a permis d'utiliser la transmission numérique d'abord sur les liaisons de transit du réseau téléphonique, puis sur la ligne d'abonné avec le RNIS.

**Micro-ordinateur** (Micro Computer) : Ordinateur dont l'unité centrale est constituée d'un microprocesseur. Le micro-ordinateur contient d'autres circuits intégrés réalisant des fonctions de mémoire et d'interface.

**Microprocesseur** (Micro Processor) : Circuit intégré contenant les circuits arithmétiques, logiques et de contrôle ainsi que les registres nécessaires pour réaliser les fonctions de l'unité centrale d'un ordinateur.

**Migration** : c'est la transition d'un pan du système d'information ou d'une application vers un nouvel état, une nouvelle version. On parle ainsi de migration lorsque l'on change de logiciel, de technologie ou encore de machine, ce qui implique des problématiques de gestion des données existantes.

**Minitel** : Nom donné par France Télécom au terminal de son service vidéotex, permettant d'accéder à des serveurs de données ou de messagerie à travers le réseau téléphonique et Transpac. Le Minitel a été un grand succès commercial en France, parce que France Télécom l'a distribué gratuitement et a ainsi créé un marché pour une offre de services diversifiée et intéressante.

**MOA** : l'abréviation de maîtrise (ou maître) d'ouvrage. Dans un projet informatique, la MOA est une personne ou un groupe qui représente le client, le propriétaire du produit fini. Il définit des objectifs, délais et budgets. Ses responsabilités comprennent ainsi la rédaction du cahier des charges fonctionnel, le choix de la MOE, le pilotage du projet ou encore la validation de la recette fonctionnelle.

**Modèle** : Représentation schématique, à la fois conceptuelle et causale, d'un objet dont on veut pouvoir simuler mentalement le fonctionnement (en se faisant assister éventuellement par l'ordinateur). Un modèle économique simule le fonctionnement d'une économie, un modèle d'entreprise celui d'une entreprise, un modèle mécanique celui d'une machine etc.

**Modèle conceptuel de données (MCD) :** Catalogue des définitions mettant en évidence les liens logiques entre les diverses données de l'entreprise : découpage des nomenclatures sans double compte ni omissions, documentation des agrégats, ratios et autres données dérivées par le calcul des données d'observation directe. La rédaction du MCD est une étape importante de l'administration des données.

**Modèle OSI (OSI Model) :** Le modèle OSI (Open System Interconnection) a été mis au point par l'ISO (International Standard Organization) pour définir un modèle en couches pour la transmission de données sur les réseaux de télécommunications. Il comporte sept couches (physique, liaison, réseau, transport, session, présentation, application).

**Modélisation :** Activité dont le but est de produire un *modèle*.

**Modem :** Appareil (modulateur démodulateur) qui traduit les signaux électriques d'un ordinateur en modulations susceptibles d'être transportées par le réseau téléphonique analogique, et qui effectue la traduction inverse à l'arrivée.

**MOE :** par opposition à la MOA, la MOE (maîtrise d'œuvre) désigne les personnes qui réalisent les projets, qui mettent au point les logiciels, services ou solutions. La MOA est donc chargée de définir ce qu'il faut faire, alors que la MOE définit comment le faire et en assure la réalisation. Ces deux termes ont été empruntés au vocabulaire du bâtiment et des travaux publics.

**Multimédia (Multimedia) :** Ce terme s'applique aux interfaces de communication qui utilisent (presque) toutes les possibilités sensorielles : caractères d'imprimerie, images fixes ou animées, son, voire sensations tactiles et vue en trois dimensions dans les espaces virtuels. Il s'applique aussi aux logiciels qui utilisent ces interfaces, et aux PC disposant de lecteurs de CD-Rom, de cartes graphiques, de cartes son, de micros, de hauts parleurs.

**Multiplateforme :** se dit d'un logiciel ou d'une application pouvant être exécuté sur différentes plateformes (une plateforme désignant ici le couple machine-système d'exploitation). Dans le monde du développement mobile par exemple, une application tournant à la fois sur Android et sur iOS est dite multiplateforme.

**Multiplexeur (Multiplexer) :** Appareil qui permet d'entrelacer à l'émission plusieurs signaux sur un même canal de transmission, et de les séparer à la réception (démultiplexeur). Les multiplexeurs numériques assurent aussi des fonctions de compression du signal, de reroutage en cas d'encombrement du canal, etc. Ils sont utilisés par les entreprises pour optimiser l'utilisation des liaisons louées.

**MVC (Modèle-Vue-Contrôleur) :** c'est un motif d'architecture logicielle très utilisé pour la conception d'applications web. Autrement dit, le pattern MVC désigne une façon bien connue et éprouvée d'organiser son application selon trois types de modules : un modèle (qui contient les données à afficher), une vue (ce que l'on voit à l'écran) et un contrôleur (qui fait le lien entre modèle et vue).

**Navigateur web :** Un navigateur web est un logiciel permettant de parcourir le web et d'afficher les ressources qui s'y trouvent.

**Nerd :** (prononcer "neurd"). Le "nerd", c'est le programmeur sans bon sens, le "Gaston la Gaffe" de l'informatique. Il travaille autant ou plus que le geek, mais n'en a pas l'efficacité. La satisfaction des utilisateurs ne le concerne pas. Son langage est rempli de termes techniques qu'il cite avec componction et toujours hors de propos. Les programmes qu'il écrit sont désespérément bogués. Il

ne tient jamais ses engagements. Il a souvent l'allure d'un adolescent attardé : lunettes sales, acné, gestes mal coordonnés.

**Notes :** Logiciel de groupware portant la marque de Lotus. Lotus Notes est à la fois une interface utilisateur permettant un accès commode à toutes les applications du Groupware ainsi qu'aux données du système d'information, et une plate-forme de développement permettant de réaliser des applications communicantes pour un faible coût. IBM veut faire de Notes l' " interface universelle d'accès au système d'information ". Le serveur de Notes (Domino) comporte une interface commode avec l'Internet. Les produits rivaux ne sont pas au même niveau (Exchange de Microsoft), mais la concurrence est féroce et l'innovation rapide.

**Numéris :** Nom du réseau RNIS commercialisé par France Télécom.

**Object Linking and Embedding (OLE) :** Technique qui permet de faire apparaître une application comme un objet à l'intérieur d'une autre application (exemple : un tableur à l'intérieur d'un traitement de texte), et d'étendre ainsi les fonctionnalités de l'application hôte.

**Object-Oriented Creation and Learning (OOCL) :** Méthode de modélisation des processus mise au point par Agilis Corporation.

**Œuvre :** Produit fourni par un travail et propre à un usage. La façon dont le travail est mené (ouvrage) a une influence sur l'œuvre (délais, qualité, satisfaction des clients etc.).

**Open source :** On dit open source les logiciels dont le code source est distribué sous une licence permettant de consulter, modifier et redistribuer le logiciel.

**Open data :** Une donnée ouverte ou open data, est une information brute et publique pouvant être librement consultée et réutilisée.

**Ordinateur (Computer) :** Appareil électronique capable de recevoir des données et d'exécuter sur ces données des instructions programmées à l'avance.

**Orienté-objet (langage) (Object-Oriented Language) :** La tradition du développement voulait que l'on sépare les données et leur traitement, réalisé par les applications. Puis il est apparu judicieux de rapprocher dans un même petit programme certaines données et certains traitements qui leur sont souvent associés. Ce petit programme, appelé " objet ", est par la suite plus facile à réutiliser que les lignes de code des anciennes applications. A la limite, le travail fait autrefois par une application est fait désormais par des objets qui communiquent en s'échangeant des messages.

L'offre d'outils de développement orientés-objet n'est pas stabilisée, mais ce type de langage semble devoir s'imposer pour des raisons économiques et logiques.

**Orienté-objet (modèle) (Object-Oriented Model) :**

La définition des données (modèle conceptuel de données) est complétée, dans les modèles orientés-objet, par celle des activités qui utilisent les données (" use case ") et par celle des objets que ces activités manipulent. Il sera utile de pousser le modèle conceptuel de données jusqu'à ce degré de précision, même si ensuite les contraintes techniques ne permettent pas de tout réaliser avec un langage orienté-objet.

**Ouvrage** : Définition, financement, organisation de l'action ou du travail en vue de la réalisation d'une mission. L'ensemble composé du produit fini (œuvre) et des conditions de sa réalisation contribue à l'accomplissement de la mission.

**Paquet** (Packet) : Un message qui doit être transmis à travers un réseau commuté est découpé en " paquets " qui comportent l'adresse du destinataire, un numéro d'ordre, et des indications permettant de vérifier l'intégrité de la transmission. Un commutateur qui reçoit un paquet le route vers le commutateur suivant.

**Pare-feu** : Voir « firewall ».

**Peer-to-peer** : Le peer-to-peer (P2P, ou pair-à-pair en français) est un système de connexion directe entre postes clients qui ne passent pas par le serveur. Ces réseaux en peer-to-peer permettent notamment le partage de fichiers.

**Permalien** : Le terme « permalien » provient de la contraction entre « permanent » et « lien ». Il s'agit d'un lien souvent utilisé dans le domaine des blogs et qui renvoie vers l'archive d'un article.

**Personal Computer (PC)** : Nom du micro-ordinateur lancé par IBM en 1981 pour faire face à la concurrence des micro-ordinateurs d'Apple. Le PC est devenu un standard, et on utilise souvent cet acronyme comme synonyme de micro-ordinateur (nous le faisons dans certaines de nos définitions).

**Plug-in** : Un plug-in est un paquet permettant de compléter l'offre de base d'un logiciel en termes de fonctionnalités. Celui-ci ne peut généralement pas fonctionner seul puisqu'il vient étendre un logiciel déjà existant. Un plugin peut également s'appeler différemment selon les cas : plugin, extension, add-in, add-on, ...

**Pont** (Bridge) : Appareil (souvent un micro-ordinateur) qui permet à deux réseaux locaux de communiquer. Un pont filtrant ne laisse sortir d'un réseau que les trames destinées à l'autre. Cf. routeur.

**Prévision** (Forecast) : Une prévision qui se réaliserait exactement relèverait de la magie, non du système d'information. Celui-ci peut seulement aider à établir des projections, ou des extrapolations, en partant des informations disponibles : c'est comme les phares d'une voiture, qui éclairent toujours droit devant, même sur une route sinueuse. Cela ne diminue en rien leur intérêt. La qualité des prévisions peut être améliorée par un recours à l'économétrie.

**Processeur** : Le processeur est un composant qui interprète et exécute les instructions machine. C'est lui qui réalise les calculs permettant une interaction avec l'ordinateur.

**Processus** (Process) : Succession ordonnée des opérations nécessaires à l'exécution d'une tâche.

**Progiciel** : Le terme « progiciel » vient de la contraction entre « produit » et « logiciel ». Il s'agit d'un logiciel ou groupement de logiciels/programmes conçu pour un usage très spécifique, pour un domaine d'activité particulier, par exemple.

**Programmation** : La programmation est le fait d'écrire un programme, dans un langage spécifique, pouvant être interprété puis exécuté par une machine.

**Protocole** : Un protocole (informatique) est l'ensemble des spécifications qui décrivent le déroulement et les règles à suivre lors d'un échange de données. Exemples de protocole : HTTP, FTP, IP, POP.

**Protocole de communication** (Communication Protocol) : Norme définissant les conditions du dialogue entre deux équipements qui doivent communiquer.

**Proxy** : Le proxy est un programme servant d'intermédiaire entre une machine et un autre réseau (internet principalement). On appelle également proxy le matériel utilisé pour le fonctionnement de ce service.

**Puce** : Une puce, en informatique, est une petite plaque de silicium sur laquelle est gravé un circuit électronique. Intégrant un microprocesseur, elle peut réaliser une ou plusieurs fonctions plus ou moins complexes.

**Random Access Memory (RAM)** : C'est dans l'ordinateur la mémoire sur laquelle sont stockés les logiciels en cours d'exécution et les données en cours de traitement. La RAM, ou " mémoire vive ", réside sur un circuit intégré d'accès rapide et de coût élevé.

**Read Only Memory (ROM)** : Mémoire dans laquelle le constructeur d'un ordinateur grave des instructions qui ne pourront pas être modifiées par la suite.

**Réalité augmentée** : La réalité augmentée consiste à afficher des informations virtuelles dans le monde réel par l'intermédiaire d'un smartphone ou de lunettes par exemple. Elle désigne une interface virtuelle, en 2D ou en 3D, qui vient enrichir la réalité en y apportant et superposant des informations complémentaires. Cette technologie fonctionne par l'intermédiaire d'un terminal qui filme le monde réel et y incruste en direct des objets virtuels, animations, textes, données, sons que l'utilisateur visionne à partir de l'écran (exemples : smartphone, tablette tactile, paire de lunettes, casque, etc).

**Réalité virtuelle** : La réalité virtuelle (en anglais, *virtual reality* ou VR) est une technologie permettant de plonger un individu dans un monde artificiel créé numériquement. Il peut s'agir d'une reproduction du monde réel ou bien d'un univers totalement imaginaire. L'expérience est à la fois visuelle, auditive et parfois haptique avec la production d'un retour d'effets. La réalité virtuelle est à distinguer du concept de réalité augmentée.

**Rédaction coopérative** (Cooperative Writing) : Logiciel qui permet à plusieurs auteurs travaillant sur des PC différents de coopérer pour écrire un même texte. Il faut que le texte soit conservé dans la mémoire d'un serveur, et soit protégé contre les risques de corrections qui s'effaceraient ou se superposeraient (" concurrence ").

**Référencement** : Le terme « référencement », lorsque l'on parle d'un site internet, a plusieurs sens. Il peut s'agir de l'inscription du site web dans les moteurs de recherche et les annuaires, qui vont donc le « référencer » dans leurs pages de résultats. Ce terme correspond aussi à l'ensemble des actions et techniques visant à améliorer la position du site internet dans ces résultats de recherche et à en optimiser la visibilité. La notion de référencement naturel fait référence à toutes les techniques utilisées pour améliorer le positionnement du site dans les moteurs de recherche et accroître sa popularité. Le référencement payant correspond à l'utilisation de publicités comme les campagnes adwords, par exemple.

**Référentiel** : Documentation et base de données indiquant, pour chacune des populations considérées par le SI de l'entreprise, les règles d'identification des individus qui la composent, les nomenclatures utilisées pour la classer. La gestion des référentiels relève de l'administration des données.

**Réplication** (Replication) : Procédé qui permet à deux serveurs d'échanger, selon des horaires convenus, les mises à jour des bases de données ou des bases documentaires.

**Reporting** : Document périodique d'une forme convenue à l'avance, par lequel une entité de l'entreprise rend compte de son activité à une autre entité.

**Réseau informatique** : il s'agit d'un ensemble d'éléments informatiques (ordinateur, imprimante, hub, modem..) connectés les uns aux autres. La connexion entre ces éléments peut se faire : de manière physique à travers une connexion filaire avec prise Ethernet ou via des ondes radio avec une connexion sans fil. Un réseau informatique permet de partager des données, des documents, des applications et des imprimantes.

**Réseau local de PC (RLPC) (Local Area Network (LAN))** : Un réseau local, ou RLPC, permet à plusieurs ordinateurs de communiquer entre eux dans le même bâtiment ou le même campus.

**Réseau Numérique à Intégration de Services (RNIS) (Integrated Services Digital Network (ISDN))** : Norme internationale de communication numérique entre les installations des utilisateurs finals. Le RNIS permet à ses utilisateurs de disposer de débits plus élevés que ceux permis par un accès analogique au réseau.

**Résolution** : ce terme, souvent confondu avec la « définition », détermine le nombre de points par unité de surface, exprimé en points par pouce (en français PPP, en anglais DPI pour Dots Per Inch). A noter qu'un pouce représente 2.54 cm. La résolution permet ainsi d'établir le rapport entre le nombre de pixels d'une image et la taille réelle de sa représentation sur un support physique.

**Responsable de domaine** : A l'intérieur de la direction informatique, le responsable de domaine est le correspondant privilégié d'un maître d'ouvrage délégué. Il est responsable des travaux concernant le SI d'un des métiers de l'entreprise. Il supervise le travail des chefs de projet maîtrise d'œuvre.

**Responsive design** : le responsive design englobe les techniques de conception de contenus Internet, proposant des contenus auto-adaptables en fonction des interfaces de consultation utilisées par le visiteur. Dans le cadre du responsive web design, une page web ou une image peut ainsi se redimensionner en fonction de la taille d'écran du terminal utilisé (ordinateur, tablette, smartphone, etc).

**RGPD** : il s'agit de l'acronyme de « Règlement Général sur la Protection des Données » (en anglais : « GDPR » pour « General Data Protection Regulation »). Il s'agit du Règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données, Le texte du RGPD peut être consulté sur le site EUR-Lex. Le RGPD impose de nombreuses obligations nouvelles comme des clauses contractuelles détaillées en matière de sous-traitance dans les contrats informatiques.

**Robotique** : cette notion se réfère à l'ensemble des domaines scientifiques et industriels en rapport avec la conception et la réalisation de robots. Dans le domaine industriel, la robotique produit des automates réalisant des fonctions précises sur des chaînes de montage et concerne également

aujourd'hui la recherche scientifique, l'exploration spatiale et les activités de défense militaire ou de maintien de l'ordre. Elle intéresse également le secteur médical, pour les prothèses, les assistantes aux chirurgiens ou aux infirmiers, par exemple.

**Routeur (Router) :** Appareil (souvent un micro-ordinateur) qui permet à plusieurs réseaux locaux de communiquer, ou plus généralement qui assure une fonction de commutation sur un réseau de transmission de données. Le routeur dispose de tables d'adressage qui lui permettent d'envoyer une trame ou un paquet vers le réseau sur lequel se trouve le destinataire. Cf. pont.

**Salle de marché :** Salle où des opérateurs (synonyme de trader ou de cambiste) interviennent sur le marché d'une part pour coter, d'autre part pour acheter et vendre.

**Script :** ce terme désigne, en informatique, un programme en charge de l'exécution d'une action prédéfinie quand un utilisateur réalise une action ou qu'une page web est en cours d'affichage sur un écran donné. Un script est une suite de commandes simples et souvent peu structurées, dans une logique d'automatisation de certaines tâches successives dans un ordre précis.

**Secure Socket Layer (SSL) :** ce concept désigne le protocole de sécurité le plus répandu qui crée un canal sécurisé entre deux machines communiquant sur Internet ou un réseau interne. De manière générale, le SSL est utilisé lorsqu'un navigateur doit se connecter de manière sécurisée à un serveur web. D'un point de vue technique, le SSL est un protocole transparent qui nécessite peu d'interaction de la part de l'utilisateur final.

**Sécurité informatique ;** elle consiste à protéger les ordinateurs, les serveurs, les appareils mobiles, les systèmes électroniques, les réseaux et les données contre les attaques malveillantes. Avec l'ouverture des systèmes informatiques sur l'extérieur et le rôle de support stratégique tenu désormais par ces derniers, la sécurité informatique est un enjeu majeur aujourd'hui. Avec la multiplication des menaces dont la cybercriminalité, le responsable sécurité informatique doit agir pour prévenir les attaques du type espionnage, vol de données, sabotage, et assurer une protection efficace des biens numériques.

**Segmentation :** Utilisé souvent comme synonyme de classification : identification de classes entre lesquelles l'on peut répartir une population. Une segmentation se construit selon des critères répondant à la fin visée. En marketing, il s'agit de classer les clients selon la ressemblance de leurs comportements en regard de l'offre commerciale qui leur est faite.

**Sémantique :** Ce qui est relatif au sens (d'un mot, d'un texte) ou à une intention (d'une action, d'une organisation).

**Série chronologique (Time Series) :** Présentation des données selon l'ordre des périodes sur lesquelles a porté l'observation. Cette présentation est souvent utile pour l'interprétation des données (mesure du taux de croissance, estimation de la tendance, etc.).

**Serveur (Server) :** Ordinateur connecté à un réseau et qui met à la disposition d'autres ordinateurs ses ressources de mémoire et de puissance, ou qui sert de support au déroulement d'un programme. Exemple : serveurs de messagerie, de base de données, de calcul etc.

**Service Oriented Architecture (SOA) :** en français, Architecture Orientée Services, le SOA formalise le concept d'échange et de partage inter-application. Le SOA propose de définir les échanges en termes de services. Les principes de base du SOA sont notamment une interface standard, un vocabulaire

partagé, ainsi que la mise en relation de composants logiciels clients consommateurs de services avec des composants fournisseurs des susdits services.

**Shadow IT** : cela se rapporte, dans une entreprise, aux matériels ou logiciels qui ne sont pas pris en charge au sein du service informatique central. L'expression peut revêtir une connotation négative car elle implique que le service informatique n'a pas approuvé la technologie ou qu'il n'est même pas informé de son utilisation par les employés de l'entreprise. Le Shadow IT peut présenter des risques pour la sécurité à partir du moment où les équipements et logiciels non pris en charge ne sont pas soumis aux mêmes mesures de sécurité que les autres technologies de l'entreprise.

**SIAD (Système informatique d'aide à la décision)** : Outil d'observation et de description qui, en fournissant une présentation synthétique de l'ensemble des données produites ou acquises par l'entreprise, donne aux managers les moyens de recevoir des alertes, de suivre l'évolution de l'activité et de faire des investigations sur des sujets ou phénomènes particuliers. Il ne fournit pas les explications ni les commentaires qui relèvent d'un travail postérieure à l'observation.

**SIG comptable : système d'information géographique** ou SIG. Est un système d'information conçu pour recueillir, stocker, traiter, analyser, gérer et présenter tous les types de données spatiales et géographiques.

**Software as a Service (SaaS)** : logiciel en tant que Service en Français, est un modèle de distribution de logiciel au sein duquel un fournisseur tiers héberge les applications et les rend disponibles pour ses clients grâce à internet. Le SaaS fait partie des quatre catégories principales de Cloud Computing, tout comme l'Infrastructure en tant que service (IaaS), la Plateforme en tant que Service (PaaS), et le Desktop en tant que Service (DaaS).

**Spam** : Message parasite diffusé sur l'Internet dans le but d'encombrer les boîtes aux lettres électroniques et le réseau lui-même.

**Standard Query Language (SQL)** : Format des questions adressées à une base de données (" requêtes SQL ").

**Stratégie** : Ensemble des questions relevant de l'art du "stratège", du dirigeant de l'entreprise.

**Système d'exploitation (Operating System (OS))** : Logiciel qui exécute les tâches relatives à la création, l'ouverture, la fermeture, la copie, la destruction de fichiers contenant des programmes ou des données, ainsi qu'au lancement des programmes et à leur exécution, à la gestion des interruptions etc.

**Système d'information (Information System (IS))** : Ensemble constitué par la définition des processus des métiers et par celle des stocks et flux d'information éclairant ces processus.

**Système informatique (Information Technology (IT))** : Ensemble de moyens matériels et logiciels assurant le stockage, le traitement et le transport des données sous forme électronique.

**Tableur (Spreadsheet)** : Logiciel comportant des fonctions qui facilitent le calcul sur des données quantitatives rangées dans un tableau. C'est le premier tableur, VisiCalc, qui a fait décoller le marché des micro-ordinateurs après 1978.

**Théorie de l'information (Information Theory)** : Théorie mise au point par Shannon pour quantifier l'information apportée par un message, une distribution statistique etc. et traiter des questions de

dimensionnement des réseaux. La théorie de l'information est indifférente au sens des données considérées. Elle est donc étrangère au sens usuel du mot " information ", mais ses applications en analyse des données montrent qu'elle peut aider à dégager le sens des données.

**TIC** : Le terme TIC (Techniques de l'Information et de la Communication) couvre un large éventail de services, applications, technologies, équipements et logiciels. Il peut s'agir des outils comme la téléphonie et l'Internet, l'e-learning, les télévisions, les ordinateurs, les réseaux et les logiciels indispensables pour employer ces technologies. Les TICE représentent les Technologies de l'Information et de la Communication pour l'Enseignement. Elles recouvrent les outils et produits numériques pouvant être utilisés dans le cadre de l'éducation et de l'enseignement.

**Token Ring** : Protocole de transmission de données sur réseau local, selon lequel le droit d'émettre une trame est donné à l'ordinateur qui détient un jeton (token) qui passe d'un ordinateur à l'autre en suivant leur ordre sur l'anneau (ring).

**Tracking** : ce terme est relatif à l'action consistant à « pister » un visiteur sur un site web. Il est ainsi possible de connaître l'itinéraire de visite ainsi que les temps et dates de visites. Le tracking peut également concerner l'observation des réactions et actions d'un internaute consécutives à la réception d'un e-mail, a posteriori ou en temps réel.

**Traducteur** (Translator) : Logiciel qui transcrit l'intégralité d'un programme d'un langage dans un autre avant de lancer son exécution (l'interpréteur, par contre, procède instruction par instruction et en les faisant exécuter immédiatement).

**Traitement de texte** (Word Processing Application) : Logiciel qui permet à l'utilisateur de mettre en forme un texte selon une présentation dont la qualité se rapproche de celle de la typographie (sans l'égaliser toutefois). Le traitement de texte est, après le tableur et avant le grapheur, l'une des applications qui ont assuré le succès des micro-ordinateurs.

**Trame** (Frame) : Un message qui doit être transmis sur une liaison entre deux points est découpé en trames comportant un numéro d'ordre et les indications nécessaires à la vérification de l'intégrité de la transmission. Un paquet transmis entre deux commutateurs est ainsi découpé en trames. La technique de transmission par " relais de trame " (" frame relay ") évite la reconstitution et le découpage des paquets dans chaque commutateur.

**Transistor** : Equipement de petite taille constitué d'un empilage de semi-conducteurs obtenus en ajoutant des impuretés au silicium, et qui a la propriété d'amplifier, contrôler et produire des signaux électriques. Inventés en 1947, les transistors ont supplanté les tubes à vide vers la fin des années 50, et permis vers la fin des années 60 la production des circuits intégrés.

**Transmission Control Protocol / Internet Protocol (TCP/IP)** : Protocole de transmission de données en mode paquet. Mis au point pour répondre à une demande de l'armée américaine, il est construit pour être indestructible en cas d'attaque nucléaire : pas d'administration centralisée, propagation des tables d'adressage entre serveurs par réplique de proche en proche, routage très simple exigeant le minimum de puissance des ordinateurs. C'est ce protocole qui a permis de construire le réseau Internet. TCP/IP s'impose comme le standard universel de la transmission de données.

**Transpac** : Réseau public français de transmission de données en mode paquet selon le protocole X25. Le prix de la communication est indépendant de la distance (cf. liaison louée). " Transpac " est aussi le nom de l'opérateur qui exploite ce réseau.

**Tutoriel** : il s'agit d'un guide d'apprentissage permettant à un utilisateur débutant de se former au fonctionnement d'un logiciel, d'un outil, d'un langage informatique, entre autres.

**UI** : l'interface utilisateur (UI) est relative aux éléments d'un appareil informatique avec lesquels un être humain peut entrer en interaction : écran, clavier, souris, caractères lumineux, messages d'aide, etc. L'UI renvoie aussi à la façon pour une application ou un site Web de solliciter les interactions avec l'utilisateur et d'y répondre.

**Unicode** : Unicode désigne un système de codage utilisé dans le cadre d'échange de contenus à l'échelle internationale, dans différentes langues.

L'unicode est développé par le Consortium Unicode, une organisation privée sans but lucratif. L'Unicode fait référence à un codage de caractères sur 16 bits, visant à coder l'ensemble des caractères couramment utilisés dans les différentes langues de la planète en spécifiant un nombre unique pour chacun de ces caractères. Ainsi, l'Unicode regroupe plus de 120 000 caractères.

**Unified Modelling Language (UML)** : Langage qui vise à rassembler les meilleurs procédés (" best practices ") dans les modèles conceptuels de données orientés-objet, et qui ambitionne de devenir un standard en unifiant les autres langages de modélisation.

**Urbanisme du système d'information** : Par analogie avec l'urbanisme d'une ville, cette expression désigne (1) la modélisation "à grosses mailles" du SI, représentant l'ensemble du SI d'une entreprise avec ses divers domaines, leurs processus et les relations qu'ils entretiennent ; (2) les règles qu'un SI doit respecter pour être cohérent et utiliser de façon efficace les ressources partagées par les divers processus.

**URL** : de l'anglais Uniform Resource Locator, il correspond à l'adresse unique d'un fichier accessible sur Internet. L'URL contient le nom du protocole permettant d'accéder à la ressource, un nom de domaine identifiant un ordinateur sur Internet et un chemin d'accès indiquant l'emplacement du fichier dans cet ordinateur.

**Use case** : Dans UML, un " use case " est la description formelle d'une activité dans un processus : le " use case " décrit les données consultées ou saisies par l'acteur, ainsi que ses appels de service pour le traitement des données.

**UX** : User eXperience, en anglais, définit la qualité de l'expérience vécue par l'utilisateur dans toute situation d'interaction. L'UX qualifie l'expérience globale ressentie par l'utilisateur lors de l'utilisation d'une interface ou plus largement en interaction avec tout dispositif ou service. L'UX s'appuie aussi bien sur l'utilisabilité d'une interface que sur l'impact émotionnel ressenti.

**Virus informatique** : c'est un programme de taille réduite, doté des caractéristiques suivantes : infection, multiplication, possession d'une fonction nocive. La fonction d'infection permet au virus de s'introduire dans des fichiers de programme, dans des fichiers de données utilisant un langage de script, ou dans une partie de la disquette ou du disque dur contenant un petit programme.

**Virtuel** : Attention : ce mot est un faux ami. En français, une chose est virtuelle quand elle n'existe qu'à l'état de simple potentialité. En anglais, " the virtual leader ", c'est le vrai patron d'une équipe (mais qui n'en a pas le titre). Une entreprise virtuelle, c'est une entreprise *de fait* malgré les apparences, etc. Ce faux ami entraîne de nombreux contresens. Cf. les définitions de "virtuel" et de "virtual" données par de bons dictionnaires.

**VPN** : (Virtual Private Network) est un type de réseau informatique qui permet la création de liens directs entre des ordinateurs distants. Le VPN s'appuie sur la création d'un tunnel via un protocole permettant à deux ordinateurs distants de se connecter à un même réseau local. Le VPN gratuit ou payant permet à certains utilisateurs d'accéder à un réseau interne, malgré la distance.

**VPS** : VPS pour Virtual Private Server est un serveur virtuel que l'utilisateur perçoit comme un serveur dédié, qui est installé sur une machine alimentant plusieurs sites Web. Une même machine peut héberger plusieurs serveurs privés virtuels ayant leur propre système d'exploitation, exécutant le logiciel d'hébergement d'un utilisateur.

**WAI** : (Web Accessibility Initiative en anglais ou Initiative sur l'Accessibilité du web) a été lancée en 1997 par le World Wide Web Consortium (W3C). Une des priorités de la WAI est relative à la proposition de solutions techniques pour rendre le World Wide Web accessible aux personnes handicapées et plus largement à toute personne sans nécessiter de pré-requis particulier. Les actions de la WAI concernent différents domaines : les technologies du Web, le développement de recommandations, le développement d'outils, l'information et la formation, la recherche et le développement.

**Web** : le Web est le terme communément employé pour parler du World Wide Web, ou WWW. Cela fait référence au système hypertexte fonctionnant sur le réseau informatique mondial Internet. Par abus de langage, le Web désigne plus largement tout élément se rapprochant de l'univers Internet.

**Wide Area Network (WAN)** : Un WAN permet à plusieurs ordinateurs de communiquer entre eux dans une même zone géographique en interconnectant des RLPC.

**Wireframe** : cela désigne, dans le domaine du webdesign, le schéma d'une page web ou d'un site web. Ce schéma est utilisé lors des étapes préliminaires à la conception d'une interface web et facilite la communication sur le début du design. Il est simplifié et n'utilise pas de couleurs pour que l'équipe en charge de la conception ne soit pas distraite par les effets visuels. Le but du wireframe est donc d'ajouter, modifier ou supprimer des éléments d'un site web.

**World Wide Web (WWW)** : Le "Web" est constitué par l'ensemble des documents HTML consultables sur l'Internet. C'est l'une des applications les plus intéressantes de l'Internet.

**Workflow** : Logiciel qui permet d'organiser, faire fonctionner et contrôler un processus. Le workflow comporte la définition des masques des documents échangés entre les acteurs du processus et les règles qui codifient son fonctionnement : programmation des routages, délais, alarmes, compteurs de délais et de volumes, édition de comptes rendus automatiques.

**X25** : Norme de transmission de données par paquets utilisant la technique du "circuit virtuel" (le premier paquet établit entre les routeurs un chemin qui sera suivi ensuite par les autres).

**X400** : Norme de messagerie (adressage et stockage) sur réseau public, permettant à des utilisateurs appartenant à des entreprises différentes d'échanger des messages. L'offre de messagerie X400 est vigoureusement concurrencée par la messagerie sur l'Internet.

**Zéro-day** : une vulnérabilité zero-day est relative à une faille de sécurité informatique dont l'éditeur du logiciel ou le fournisseur de service n'a pas encore connaissance, ou qui n'a pas encore reçu de correctif. Par extension, on parle d'exploitation zero-day dans le cas où ce type de faille est utilisé par des cyberdélinquants pour lancer des attaques contre des installations vulnérables.

**ZIP** : ZIP correspond à un format de compression de fichiers très utilisé. Ce format réduit le poids des fichiers sur les unités de stockage (disque dur, clé USB, etc.) ainsi que leur temps de téléchargement. Les fichiers compressés doivent être décompressés afin de retrouver leur forme d'origine. Sur Internet, un grand nombre de programmes sont téléchargeables et permettent de compresser et décompresser des fichiers au format ZIP.

**Zoning** : le zoning a pour but de définir quels hôtes ont le droit de communiquer avec les nœuds de stockage. Chaque hôte peut être membre de plusieurs zones dans un Storage Area Network (SAN), qui est un réseau spécialisé permettant de mutualiser des ressources de stockage. Le zoning permet à un administrateur de contrôler les divers éléments dans un SAN.



## Lexiques les architectures Orientées Services (SOA)

**ERP** : *Enterprise Resource Planning*, ou PGI pour Progiciel de gestion intégré en français. Logiciel permettant de gérer l'ensemble des processus d'une entreprise. Les données sont standardisées sur une seule base de données commune à toutes les applications de l'entreprise.

**http** : Hypertext Transfert Protocol est le standard sur Internet pour les liaisons entre les différentes ressources à partager.

**Service Web** : Composant applicatif présentant une interface standardisée (au format WSDL), accessible au travers d'Internet grâce au protocole SOAP.

**SOA** : Service Oriented Architecture, architecture orientée service, préconisant un système d'information bâti à partir de processus aisément configurable reposant sur des services applicatifs autonomes.

**Workflow** : La modélisation et la gestion informatique de l'ensemble des tâches à accomplir et les communications entre les différents acteurs impliqués pour la réalisation d'un processus métier.

**WSOA** : *Web Service Oriented Architecture*, SOA mise en oeuvre grâce à l'utilisation de services Web et de leur annuaire de référencement UDDI.



## Lexiques les services Web

**CORBA** : Dans l'architecture Corba, ce rôle est joué par l'ORB (*Object Request Broker*), qui transporte les requêtes entre le programme client et les objets sur le serveur.

Depuis la version 2 de la norme Corba, le protocole de communication de l'ORB, appelé IIOP (*Internet InterOrb Protocol*), s'appuie sur TCP/IP et peut donc être employé sur Internet.

**DCOM** : Dans l'architecture Microsoft, ce rôle est précisément assuré par le protocole DCOM (*Distributed Component Object Model*) qui étend le modèle monomachine COM aux réseaux de PC sous Windows.

**EJB** : Dans l'architecture J2EE (*Java 2 Enterprise Edition*), le protocole RMI (*Remote Method Invocation*) permet à un programme Java d'en appeler un autre qui est exécuté sur une machine virtuelle Java distante.

**Requête** : Un langage de **requête** est un langage **informatique** utilisé pour accéder aux données d'une base de données ou d'autres systèmes d'information. Il permet d'obtenir les données vérifiant certaines conditions (on parle de critères de sélection), comme toutes les personnes qui habitent une ville donnée.



---

## Lexiques concept de la programmation

**Abstraction** : Faculté des humains de se concentrer sur l'essentiel et d'oublier les détails. Parfois employé comme synonyme de classe.

**Abstraite** : Se dit d'une classe qui ne peut être instanciée directement.

**Accessibilité** : La conception de produits, d'appareils, de service ou d'environnements prenant en compte la capacité d'accès de tous les utilisateurs, y compris les personnes souffrant de handicaps ou celles qui sont limitées par une technologie plus ancienne ou plus lente.

**Acteur** : 1) Classe de personnes ou de systèmes qui interagissent avec un système.  
2) Objet toujours à l'origine d'une interaction

**Action** : Opération exécutée instantanément lors d'une transition d'un état vers un autre état. Une action ne peut être interrompue.

**Activité** : Opération exécutée au sein d'un état. Une activité est interrompible.

**Agent** : Objet qui peut être origine et destination d'une interaction.

**Agrégation** : Forme d'association non symétrique qui exprime un couplage fort et une relation de subordination.

**Algorithme** : 1. Enchaînement des actions nécessaires à une tâche. 2. Un algorithme est un ensemble de processus ou de règles à suivre dans les calculs ou d'autres opérations de résolution de problèmes. Les algorithmes sont à la base de la plupart des programmations informatiques.

**Analyse** : Détermination du *quoi* et du *quoi faire* d'une application.

**Analyse des besoins** : Détermination des besoins des utilisateurs.

**Analyse du domaine** : Partie de l'analyse qui se concentre sur l'environnement de l'application.

**Ancêtre** : Synonyme de super-classe.

**Angular.JS** : Angular.js est un framework d'application web open-source maintenue par Google.

**Apache** : Apache est un serveur Web open source basé sur Unix. Il a été créé par la Fondation Apache Software.

**API** : Une API est une interface de programmation d'application. C'est un ensemble de routines, de protocoles et d'outils pour créer des applications logicielles. Les composants logiciels express de l'API en termes d'opérations, d'entrées, de sorties et de types sous-jacents.

**Application** : Système logiciel élaboré dans un but précis.

**Architecte** : Spécialiste de l'architecture.

**Architecture** : 1) Art de construire ; en informatique, art de construire les logiciels.  
2) Structure d'un logiciel.

**Array** (Tableau) : Un type de variable qui peut stocker plusieurs bits de données. Un tableau est généralement défini en utilisant des crochets avec le nombre d'éléments dans le tableau passé dans les parenthèses. Chaque élément du tableau peut stocker une donnée comme s'agissait d'une variable. Dans certains langages de programmation, un tableau ne peut alors que dans d'autres langages, les tableaux peuvent contenir des types de données mixtes.

**Artefact** : Élément d'information, produit ou consommé par une des activités lors du développement d'un logiciel.

**ASP** : ASP est une fonctionnalité de Microsoft Internet Information Server (IIS), lequel construit une page HTML. L'on peut créer un fichier ASP en incluant dans un fichier écrit en VBScript ou JScript des instructions de programme ActiveX Data Objects (ADO) dans le fichier HTML.

**Assembleur** : Un assembleur est un programme qui prend des instructions informatiques de base et les convertit en un modèle de bits que le processeur de l'ordinateur peut utiliser pour effectuer ses opérations de base.

**Association** : Relation entre classes qui décrit un ensemble de liens.

**Association dérivée** : Association déduite d'autres associations.

**Asynchrone** : Forme de communication non bloquante et sans accusé de réception.

**Attribut** : Information contenue par un objet.

**Attribut de lien** : Attribut appartenant à un lien entre objets, plutôt qu'aux objets eux-mêmes.

**Attribut dérivé** : Attribut déduit d'autres attributs.

**Automate** : Forme de représentation abstraite du comportement.

**Big bang** : Se dit de la technique d'intégration lorsqu'elle est concentrée en fin de développement.

**Binaire** : Une façon de représenter l'information en utilisant seulement deux options.

**BIT** : Une contraction de « Binary Digit ». Un bit est la seule unité d'information dans un ordinateur, généralement représentée par 0 ou 1.

**Block-Based Programming** (Langage de programmation basé sur des blocs) : Tout langage de programmation qui permet aux utilisateurs de créer des programmes en manipulant des « blocs » ou des éléments de programmation graphique, plutôt que d'écrire du code en utilisant du texte. Les exemples incluent Code Studio, Scratch, Blockly, et Swift.

**Bug** : Une erreur dans un programme qui empêche le programme de s'exécuter comme prévu.

**Byte** (octet) : L'unité fondamentale la plus commune des données numériques. Un octet a une valeur de 8 bits de données.

**Call (une variable)** : Utilisez une variable dans un programme.

**Call (une fonction)** : C'est le morceau de code que vous ajoutez à un programme pour indiquer que le programme doit exécuter le code dans une fonction à un certain moment.

**Cardinalité** : Nombre d'éléments dans une collection.

**Cas d'utilisation** : Technique d'élaboration des besoins fonctionnels, selon le point de vue d'une catégorie d'utilisateurs.

**Catégorie** : Sorte de paquetage de la vue logique ; une catégorie encapsule des classes.

**Chatbot** : Un chatbot est un programme informatique qui tente de simuler la conversation d'un être humain via des interactions texte ou voix. Un utilisateur peut poser une commande, et le chatbot répond ou effectue l'action demandée.

**Classe** : 1.Description abstraite d'un ensemble d'objets ; réalisation d'un type. 2.Un ensemble d'objets ayant le même comportement (mais dont l'état est généralement différent) ou un modèle définissant un tel ensemble.

**Classe abstraite** : Classe qui ne s'instancie pas directement et qui sert uniquement de spécification.

**Classe active** : Classe dont les objets sont actifs.

**Classe concrète** : Par opposition aux classes abstraites, classe qui s'instancie pour donner des objets.

**Classe de base** : Classe racine d'une hiérarchie de classes.

**Classe library** (Une bibliothèque de classes) : Une bibliothèque de classes est une collection de classes ou de modèles codés pré-écrits, dont chacun peut être spécifié et utilisé par un programmeur lors du développement d'un programme d'application. Le programmeur spécifie quelles classes sont utilisées et fournit des données qui instancient chaque classe en tant qu'objet pouvant être appelé lors de l'exécution du programme. L'accès et l'utilisation d'une bibliothèque de classes simplifient grandement le travail du programmeur puisque le code prétendu standard est disponible et que le programmeur n'a besoin de l'écrire.

**Classe paramétrable** : Classe modèle pour la construction de classes.

**Classe paramétrée** : Classe résultant de la paramétrisation d'une classe paramétrable.

**Classe racine** : Classe racine d'une hiérarchie de classes.

**Classe utilitaire** : Classe dégradée, réduite au concept de module.

**Classe-association** : Association promue au rang de classe.

**Classification** : Action d'ordonner dans le but de comprendre.

**Classification dynamique** : Forme de classification par laquelle un objet peut changer de type ou de rôle.

**Classification statique** : Forme de classification par laquelle un objet ne peut pas changer de type ou de rôle.

**Clé** : Tuple des attributs qui réalisent la restriction d'une association.

**Clé naturelle** : Clé primaire issue du domaine de l'application.

**Clé primaire** : Information désignant un objet de manière bijective.

**Client** : Objet à l'origine d'une requête.

**Code** : Le langage que les programmeurs créent et utilisent pour dire à un ordinateur quoi faire.

**Code machine** (ou langage machine) : Système d'instructions et de données directement compréhensible par l'unité centrale d'un ordinateur.

**Code objet** : La sortie d'un compilateur ou d'un assembleur, pas nécessairement exécutable directement sans lien avec d'autres modules.

**Code orienté objet** : Utilisation d'entités appelées objets pouvant traiter des données et échanger des messages avec d'autres objets.

**Code source** : Instructions lisibles par l'utilisateur dans un langage de programmation, à transformer en instructions machine par un compilateur, un interpréteur, un assembleur ou un autre système de ce type.

**Collaboration** : 1) Se dit d'une interaction entre objets réalisée dans le but de satisfaire un besoin de l'utilisateur 2) Élément structurant d'UML pour la description du contexte d'une interaction.

**Collection** : Terme générique désignant tous les regroupements d'objets sans préciser la nature du regroupement.

**Commandes** (ou **instructions**) : Une instruction pour l'ordinateur. De nombreuses commandes constituent des algorithmes et des programmes informatiques.

**Composant** : Élément physique constitutif d'une application, représenté principalement dans la vue de réalisation.

**Composition** : 1) Forme d'organisation complémentaire de la classification. 2) Agrégation par valeur.

**Computational thinking** (pensée machine) : Modifier un problème de telle sorte qu'il puisse être modélisé ou résolu à l'aide d'un ordinateur ou d'une machine. Les stratégies comprennent la décomposition, l'appariement de motifs, l'abstraction et les algorithmes.

**Conception** : Détermination du *comment* d'une application.

**Condition** : Expression booléenne qui valide ou non une transition dans un automate d'états finis.

**Constructeur** : Opération de classe qui construit des objets.

**Construction** : Phase de la vue de l'encadrement dans laquelle le logiciel est réalisé et amené à un état de maturité suffisant pour être livré à ses utilisateurs.

**Conteneur** : Structure de données qui contient des objets.

En informatique, un conteneur est utilisé pour bénéficier d'un espace d'exécution dédié à une application logicielle. Il peut se présenter sous la forme d'un type de données abstrait, d'une structure d'informations ou encore d'une classe.

L'usage d'un conteneur permet de tester des applications en développement, ainsi que des logiciels. L'intérêt est de disposer d'un environnement isolé pour réaliser tous les contrôles nécessaires avant le déploiement. Il est donc plus facile de distinguer les dysfonctionnements, les failles de sécurité et d'éventuels problèmes de stabilité.

### **Quelle est la différence entre un conteneur et une machine virtuelle (VM) ?**

Si le principe de virtualisation est présent dans les deux cas de figure, l'usage et le fonctionnement entre une machine virtuelle et un conteneur sont dissemblables. Une machine virtuelle (ou VM) dispose de son propre système d'exploitation, tandis que le conteneur utilise l'OS de l'ordinateur où il est installé. Bien que ce dernier aspect soit pratique, la puissance sollicitée est inférieure, en particulier sur les ressources de la mémoire.

En revanche, la maintenance est nettement moins contraignante pour un conteneur, car la VM impose une mise à jour du système individualisé. De plus, le conteneur est beaucoup plus léger et compact. Ce qui accroît sa réactivité. Autre différence notable, la VM est un outil de virtualisation de l'infrastructure de calcul et non de l'application logicielle en elle-même, comme le permet le conteneur.

Le terme « conteneur » est l'équivalent anglais de « container ». Certains spécialistes privilégient l'emploi de « software container » pour préciser l'appartenance au domaine informatique.

**Contexte** : Ensemble d'éléments de modélisation qui sert de support à une interaction.

**Contrainte** : Expression qui précise le rôle ou la portée d'un ou plusieurs éléments de modélisation.

**Contrat** : Engagement entre classes rédigé selon les termes de la spécification du fournisseur.

**Contrôle** : Synonyme de flot de contrôle

**Couche** : Segmentation horizontale des modèles

**Couplage** : Dépendance entre éléments de modélisation.

**CRC** : Abréviation de *Class Responsibilities Collaborators*.

**CSS** : Une feuille de style en cascade (CSS) est une page Web dérivée de plusieurs sources avec un ordre de priorité défini. Une feuille de style CSS précise les feuilles de style ou les instructions qui déterminent la manière dont un élément donné est présenté dans une page Web.

**Cycle** : Passage complet par les quatre phases de la vue de l'encadrement.

**Cycle de vie** : Etapes du développement et ordonnancement de ces étapes.

**Data (Données)** : Information. Souvent, des quantités, des caractères ou des symboles qui sont les entrées et les sorties du programmes informatiques.

**Data Structures** (Structures de données) : Une structure de données est une méthode d'organisation des données dans un ordinateur afin qu'elle puisse être utilisée efficacement. Les structures de données peuvent implémenter un ou plusieurs types de données abstraits particuliers, qui sont les moyens de spécifier le contrat d'opérations et leur complexité.

**Débogage** : Trouver et résoudre des problèmes dans un algorithme ou un programme.

**Décomposition** : Séparation en éléments plus petits dans le but de réduire la complexité.

**Délégation** : Mécanisme de communication entre objets qui permet à un objet fournisseur de faire réaliser des tâches par un objet sous-traitant, sans le dire à l'objet client.

**Dépendance** : Relation d'obsolescence entre deux éléments de modélisation.

**Déploiement** : Phase de la vue de l'encadrement qui comprend la transition de l'application dans son environnement final.

**Deprecated** (Obsolète) : En programmation informatique, une entité de langage obsolète est une entité tolérée ou prise en charge mais non recommandée. Par exemple, un certain nombre d'éléments et d'attributs sont déconseillés en HTML 4.0, ce qui signifie que d'autres moyens d'accomplir la tâche sont préférés. Si vous rencontrez ce terme au sujet de déconseillé de les utiliser. Reportez-vous plutôt sur celles qui les ont remplacées dans les dernières versions du langage.

**Destructeur** : Opération de classe qui détruit les objets.

**Diagramme** : Représentation graphique d'éléments de modélisation.

**Discret** : Contraire de continu.

**Discriminant** : Synonyme de critère.

**Django** : Un framework d'application web libre et open source écrit en Python qui suit la structure MVC (Model-view-controller).

**DNS** (Domain Name Server, Service de nom de domaine) : Le service qui traduit les URL en adresses IP.

**Documentation** : Description textuelle des modèles.

**Domaine** : Synonyme de champ d'application.

**Elaboration** : Phase de la vue de l'encadrement dans laquelle l'architecture est définie.

**Élément** : Brique de base d'un modèle.

**Élément de modélisation** : Représentation d'une abstraction issue du domaine du problème.

**Élément de visualisation** : Projection graphique d'une collection d'éléments de modélisation.

**Encapsulation** : Occultation des informations contenues par un objet.

**Entité** : Terme emprunté aux méthodes de modélisation par les données ; il est employé comme synonyme d'objet du domaine.

**Enumération** : Liste de valeurs nommées ; domaine de définition d'un type.

**Erreur** : Occurrence anormale.

**Espace de nom** : Partie de modèle dans laquelle un nom peut être défini : au sein d'un espace de nom, un nom ne possède qu'une seule signification.

**Espace des états** : Ensemble des états possibles.

**Etat** : Condition instantanée dans laquelle se trouvent un objet, un système.

**Etend** : Relation d'extension entre cas d'utilisation/

**Etude d'opportunité** : Phase de la vue de l'encadrement dans laquelle la vision du produit est définie.

**Événement (Event)**: 1. Occurrence qui engendre un changement d'état. 2. Action qui provoque quelque chose.

**Event Manager (Gestionnaire d'événements)** : Un moniteur pour un événement ou une action spécifique sur un ordinateur. Lorsque vous écrivez du code pour un gestionnaire d'événements, il est exécuté chaque fois que cet événement ou cette action se produit. De nombreux gestionnaires répondent aux actions humaines telles que les clics de souris.

**Exception** : Condition exceptionnelle qui correspond à une anomalie lors de l'exécution.

**Export** : Partie visible d'un paquetage.

**Expression** : Chaîne dont l'interprétation donne une valeur d'un type donné.

**Extension** : Forme de programmation qui facilite la prise en compte de nouveaux besoins.

**Factorisation** : Identification puis extraction de similitudes entre classes.

**Flot d'exécution** : Description de la répartition de l'activité entre les objets.

**Flot de contrôle** : Synonyme de flot d'exécution.

**Flot de données** : Description des données qui transitent d'un objet vers un autre objet.

**Fonction** : 1. Expression d'un besoin en termes impératifs, sous la forme de tâches à accomplir. 2. Une routine qui reçoit un ou plusieurs arguments, qui les traite et renvoie un résultat.

**Fonction (appel)** : Le morceau de code que vous ajoutez à un programme pour indiquer que le programme doit exécuter le code dans une fonction à un certain moment.

**Fonction (définition)** : Le code à l'intérieur d'une fonction qui indique au programme ce qu'il faut faire lorsque la fonction est appelée.

**Framework** : 1. Macro-architecture générique. 2. Un framework est souvent une structure en couches indiquant quels types de programmes peuvent ou devraient être construits et comme ils interagiraient.

**Full stack** : Les Full stack Developpers sont des ingénieurs logiciels polyvalents qui ont le savoir-faire pour créer de manière autonome des plates-formes entièrement fonctionnelles, du frontal au back-end. Classiquement, le développeur web nécessite plusieurs variantes d'ingénieurs : développeurs frontaux, administrateurs de serveurs web, ingénieurs back-end et ingénieurs de base de données. Cependant, un développeur de pile complète est tout ce qui précède, et que ce soit dans une grande ou petite équipe d'ingénierie, il/elle peut ajouter de la valeur et de la perspicacité à toutes les couches du projet.

**Garde** : Condition booléenne qui valide une transition dans un automate.

**Généralisation** : Point de vue ascendant porté sur une classification ; nom donné par UML à la relation de classification utilisée pour construire des hiérarchies de classes ; souvent synonyme d'héritage, bien que plus abstrait.

**Généralisation multiple** : Forme de généralisation dans laquelle une classe dérive de plusieurs ancêtres. Souvent synonyme d'héritage multiple, bien que plus abstrait.

**Génération** : Dernière version exécutable produite par un cycle de développement.

**Générique** : Solution générale ; se dit aussi de composants modèles.

**Génie logiciel** : Discipline informatique créée afin de rationaliser la production de logiciels par la définition et l'utilisation de méthodologies et d'outils.

**Gestion de versions** : Enregistrement de l'histoire d'un élément.

**Gestion de configurations** : Etablissement de relations entre les constituants d'un système dans le but de maintenir la cohérence.

**Goto** : Une déclaration (dans le code source) qui transfère le contrôle inconditionnellement à une autre partie du programme. L'abus de Goto conduit généralement à du code spaghetti. C'est indigeste.

**HEAP** : Une zone de mémoire réservée aux objets de données alloués dynamiquement. Complète mais diffère de la pile (voir stack).

**Héritage** : Relation entre classes qui permet le partage de propriétés définies dans une classe ; principale technique de réalisation de la généralisation.

**Héritage d'interface** : Héritage de l'interface d'un élément plus général sans sa réalisation.

**Héritage de (pour la) réalisation** : Héritage de l'interface et de la réalisation d'un élément plus général.

**Héritage multiple** : Relation entre classes qui permet le partage de propriétés définies dans plusieurs classes.

**Hiérarchie** : Arborecence de classes ordonnées par une relation de généralisation.

**Ide** : L'IDE est l'abréviation d'Integrated Development Environment (Environnement de développement intégré). Il s'agit d'outils visuels permettant aux programmeurs de développer des programmes plus efficacement. Généralement, un IDE peut avoir un compilateur, un débogueur, un éditeur de texte et d'autres outils intégrés. Smalltalk a été le premier langage de programmation à avoir un véritable IDE.

**Identifiant** : Un nom formel utilisé dans le code source pour désigner une variable, une fonction, une procédure, un paquet, etc.

**Identité** : Caractéristique fondamentale d'un objet qui le distingue de tous les autres objets.

**Idiome** : Construction élégante, propre à un langage donné.

**Import** : Relation de dépendance entre paquetages qui rend visible les exports d'un paquetage au sein d'un autre paquetage.

**Instance** : Synonyme d'objet ; un objet est instance d'une classe.

**Intégration** : Qualité de l'interdépendance entre éléments de modélisation.

**Interaction** : Description d'un comportement dans le contexte d'une collaboration.

**Interface** : Partie visible d'une classe, d'un groupe d'objets ; parfois synonyme de spécification.

**INTEGER (Nombre entier)** : Un type de données pour les valeurs entières.

**Interpréteur** : Un programme qui exécute un autre programme écrit dans un langage de programmation autre que le code machine.

**Interruption** : Déroutement du flot d'exécution normal, consécutif à la détection d'une condition matérielle.

**Itération** : Une action répétitive ou une commande généralement créée avec des boucles de programmation.

**Invariant** : 1) Expression booléenne dont le changement de valeur déclenche une exception. 2) Critère pour la détection des objets ; un objet est un invariant du domaine.

**Invocation** : Mécanisme par lequel un message déclenche une opération.

**Itérateur** : Objet ou mécanisme qui permet de visiter tous les éléments d'une collection, sans en dévoiler la structure interne.

**Itération** : 1) Action de traverser une collection d'objets. 2) Séquence d'activités de la vue technique qui aboutit à la livraison du prototype.

**JQuery** : JQuery est une bibliothèque Javascript multiplateforme conçus pou simplifier l'écriture de scripts HTML côté client. JQuery est la plus courante actuellement utilisée, avec l'installation sur 35% des 100 millions de sites Web ayant le trafic le plus élevé.

**Liaison dynamique** : Association entre un nom d'objet et une classe réalisée à l'exécution.

**Liaison statique** : Association entre un nom d'objet et une classe réalisée à la compilation.

**Lien** : Connexion sémantique entre un tuple d'objets par laquelle un objet peut communiquer avec un autre objet.

**Ligne de vie** : Représentation de l'existence d'un objet dans un diagramme de séquence.

**Livraison** : Résultat exécutable d'une itération ; parfois synonyme de prototype.

**Maintenance** : Phase du cycle de vie du logiciel qui suit le déploiement ; la maintenance regroupe des activités de correction de défauts et de prise en compte des demandes d'évolution.

**Mécanisme** : Synonyme de collaboration entre objets.

**Membre** : Terminologie C++ pour désigner un attribut ou une opération contenue par une classe.

**Message** : Élément de communication entre objets qui déclenche une activité dans l'objet destinataire ; la réception d'un message correspond à un événement.

**Métaclasse** : Classe d'une classe ; elle contient les données et les opérations qui concernent la classe plutôt que les instances de la classe

**Métamodèle** : Modèle qui décrit des éléments de modélisation.

**Métamodélisation** : Modélisation récursive des éléments de modélisation à partir d'eux-mêmes.

**Méthode** : 1) Souvent synonyme d'opération ; quelquefois utilisé pour distinguer la spécification de l'opération des multiples réalisations - les méthodes - implantées dans les sous-classes.

**Méthode d'instance** : Opération qui concerne les objets.

**Méthode de classe** : Opération qui concerne la classe plus que les objets.

**Mode** : Caractérise les paramètres selon la direction du flot de données : en entrée, en entrée et en sortie, en sortie.

**Modèle** : Construction descriptive à partir des éléments de modélisation.

**Modélisation** : Synonyme d'analyse ; par extension, élaboration des modèles, y compris de conception.

**Modificateur** : Opération qui modifie l'état interne d'un objet.

**Modularité** : Qualité d'un environnement de réalisation qui permet la partition.

**Module** : Espace lexical dans lequel d'autres constructions peuvent être déclarées.

**Monomorphisme** : Concept de la théorie des types, selon lequel un nom ne peut référencer que des objets de la même classe.

**Multiplicité** : Désigne le nombre d'objets qui peuvent participer à une relation.

**Navigabilité** : Qualité d'une relation qui permet le passage d'une classe vers une autre classe.

**Niveau de maturité** : Description de la qualité d'un processus de développement.

**Nœud** : Dispositif matériel susceptible d'exécuter un programme.

**Non interprété** : Représentation d'un type non spécifié par UML, sous la forme d'une chaîne.

**Notation** : Ensemble des signes et symboles qui composent un langage. Dans le cas d'UML, ensemble des représentations graphiques et textuelles qui constituent les diagrammes.

**Note** : Information textuelle qui peut être associée à tout élément ou combinaison d'éléments de modélisation ; la note appartient à la vue, elle ne véhicule aucune sémantique.

**Objet** : **Entité** atomique constituée d'un état, d'un comportement et d'une identité.

**Objet actif** : Objet qui possède son propre flot de contrôle ; instance d'une classe active.

**Occultation d'information** : Synonyme d'encapsulation.

**Opération** : **Élément** de comportement des objets, défini de manière globale dans la classe.

**Opération abstraite** : Opération définie dans une classe, mais dont la réalisation est reportée dans une sous-classe.

**Paquetage** : **Élément** d'organisation des modèles.

**Paradigme** : Un style fondamental de programmation informatique auquel la conception d'un langage de programmation doit généralement répondre, comme la programmation impérative, la programmation déclarative ou, à un niveau plus fin, la programmation fonctionnelle, la programmation logique ou la programmation orientée objet.

**Paramètre** : Une donnée transmise à une fonction ou sous-programme lorsque la fonction ou sous-programme est appelé.

**Partie privée** : Partie de la spécification d'une classe qui regroupe des propriétés invisibles de l'extérieur.

**Partie protégée** : Partie de la spécification d'une classe qui regroupe des propriétés invisibles de l'extérieur, sauf des sous-classes.

**Partie publique** : Partie de la spécification d'une classe qui regroupe des propriétés visibles de l'extérieur.

**Partition** : Segmentation verticale des modèles ; par extension, sous-ensemble d'un modèle.

**Pattern** : Micro-architecture ; élément de conception (ou d'analyse) récurrent.

**Persistance** : Qualité d'un objet à transcender le temps ou l'espace.

**Phase** : Ensemble des activités entre deux points de contrôle d'un processus de développement.

**Pixel** : Abréviation de « élément d'image », l'unité fondamentale d'une image numérique, généralement un petit carré ou point qui contient un seul point de couleur d'une image plus grande.

**Polymorphisme** : Concept de la théorie des types, selon lequel un nom peut référencer des objets instances de plusieurs classes regroupées dans une hiérarchie de généralisation.

**Post-condition** : Condition booléenne qui doit être vraie après l'exécution d'une opération.

**Pré-condition** : Condition booléenne qui doit être vraie avant l'exécution d'une opération.

**Procédure** : Un sous-programme ou une fonction codée pour exécuter une tâche spécifique.

**Processus** : 1) Flot d'exécution lourd. 2) Suite d'étapes, plus ou moins ordonnées, dédiées à la satisfaction d'un objectif.

**Projection** : Relation entre un ensemble et un sous-ensemble.

**Programme** : Un algorithme qui a été codé dans un langage qui peut être exécuté par une machine.

**Propriété** : Caractéristique d'une classe.

**Prototype** : Résultat d'une itération ; version partielle d'un système.

**Pseudo-état** : Désigne des états particuliers comme l'état initial, l'état final, l'historique.

**Récursivité** : Application d'une règle à ses propres résultats pour générer une séquence infinie de résultats.

**Réflexive** : Se dit d'une relation dont les rôles concernent la même classe.

**Réification** : Action de chosifier un concept, une fonction.

**Responsabilité** : Obligation d'une classe ; partie de sa raison d'être.

**Restriction** : Se dit d'une relation dont la portée a été réduite par une contrainte ou une clé.

**Rétro-ingénierie** : Reconstruction des artefacts des activités en amont, à partir des artefacts des activités en aval.

**Réutilisation** : Usage poursuivi ou répété d'un artefact du développement.

**Revue** : Révision formelle d'une documentation, d'un modèle.

**Risque** : Élément susceptible de perturber le bon déroulement du développement.

**Rôle** : Extrémité d'une relation ; par extension, manière dont les instances d'une classe voient les instances d'une autre classe au travers d'une relation.

**RUN (exécuter)** : Faire que l'ordinateur exécute les commandes que vous avez écrites dans votre programme.

**Scénario** : Interaction simple entre objets.

**Schème** : Traduction française de *pattern*.

**SEI** : *Software Engineering Institute*

**Sélecteur** : 1) Opération qui renseigne sur l'état interne d'un objet, sans le modifier.  
2) Dans une expression de navigation, association qui partitionne un ensemble d'objets à partir de la valeur d'une clé.

**Serveur** : Objet qui n'est jamais à l'origine d'une interaction.

**Signal** : Événement nommé qui peut être déclenché explicitement.

**Signature** : Identifiant non ambigu d'une opération, construit à partir du nom de l'opération et de ses paramètres.

**Sous-classe** : Classe spécialisée, reliée à une autre classe plus générale par une relation de généralisation.

**Sous-état** : Etat englobé par un super-état.

**Sous-système** : Sorte de paquetage de la vue de réalisation ; un sous-système contient des éléments de réalisation ; il est le pendant de la catégorie définie dans la vue logique.

**Spécialisation** : Point de vue descendant porter sur une classification

**Spécification** : Description exhaustive d'un élément de modélisation.

**STACK** : La partie de la mémoire de l'ordinateur utilisée pour garder une trace des procédures appelées ou des instructions d'appel

**Stéréotype** : Extension de la sémantique d'un élément du métamodèle.

**Structure** : Relations statiques entre éléments de modélisation.

**Structurée** : Décrit une technique de décomposition, basée sur la notion de modules et la description des flots de données entre ces modules.

**Super-classe** : Classe générale reliée à une autre classe plus spécialisée par une relation de généralisation.

**Super-état** : Etat contenant des sous-états.

**Surcharge** : Emploi d'un même nom pour désigner différentes constructions ; la surcharge est résolue statiquement par les compilateurs en fonction du contexte et de la signature des opérations.

**Synchrone** : Forme de communication bloquante, avec accusé de réception implicite.

**Synchronisation** : Expression de la forme de communication entre deux objets.

**Temps réel** : Caractéristique d'un logiciel dont le temps de réponse est compatible avec la dynamique du système.

**Test** : Ensemble des mesures et des activités qui visent à garantir le fonctionnement satisfaisant du logiciel.

**Topologie** : Décrit la répartition des modules, ainsi que des données et des opérations dans ces modules, au sein d'une application.

**Transition** : 1) Connexion entre deux états d'un automate, déclenchée par l'occurrence d'un événement.

2) Phase de la vue de l'encadrement dans laquelle le logiciel est transféré à ses utilisateurs.

**Transition automatique** : Transition déclenchée dès que l'activité en cours dans l'état s'achève.

**Typage** : Manière dont les langages de programmation supportent la notion de type.

**Type** : Description d'un ensemble d'instances qui partagent des opérations, des attributs abstraits, des relations et des contraintes.

**Type de donnée abstrait** : Description d'une donnée en termes opératoires.

**Type primitif** : Type de base prédéfini par UML.

**URL (Localisateur de Ressources Universel)** : Une adresse facile à mémoriser pour appeler une page Web.

**Use case** : Voir cas d'utilisation.

**Variable** : Un emplacement mémoire nommé dans lequel un programme peut stocker des résultats intermédiaires et à partir duquel il peut les lire.

**Variable d'instance** : Attribut d'objet.

**Variable de classe** : Attribut qui concerne la classe plus que les objets.

**Visibilité** : Niveau d'encapsulation des attributs et des opérations dans les classes.

**Vision** : Définition d'un produit et de sa portée.

**Vue** : Manière de regarder des éléments de modélisation, éventuellement issus de modèles différents.

**Workspace** (espace de travail) : Espace dans votre environnement de programmation où vous faites glisser et déposer des commandes pour construire votre programme.





## Lexiques des tests logiciels

**Anomalie** : Dysfonctionnement du logiciel constaté lors du passage du jeu de test en comparant le résultat attendu et le résultat obtenu.

**Benchmark** : Bancs d'essai de logiciel permettant de vérifier si le logiciel du marché répond aux besoins techniques et/ou fonctionnels du client.

**Bilan de test** : Document présentant le bilan quantitatif (nombre de tests rédigés, passés, en erreur, nombre et état des anomalies, ...) et qualitatif (points forts, points à améliorer) de la mission de test, ainsi que le résultat fournis par les indicateurs qualité mis en place. Il fournit enfin des préconisations afin de palier les points faibles sur les futurs projets.

**Campagne de Test** : Activité qui consiste à dérouler un ensemble de scénarios de test. Un dossier de test est produit à l'issue d'une campagne de tests.

**Cas de test (exigence)** : Point " particulier des spécifications du logiciel nécessitant un test (règle de traitement, de calcul, de gestion, ...)

**Couverture fonctionnelle** : Ensemble des cas à tester (des exigences ou cas de test) de façon à s'assurer que le périmètre de la campagne de tests fonctionnels est bien " couvert ".

**Dossier de Test** : Ensemble documentaire qui contient la description des scénarios, des jeux de test et leurs exécutions, des anomalies et leur suivi. Le dossier de test est le reflet d'une campagne de tests.

**Dysfonctionnement** : Différence constatée entre le résultat attendu et le résultat obtenu lors du passage du jeu de test (il reste à vérifier si c'est l'attendu ou l'obtenu qui est en erreur)

**Jeux de Test** : Ensemble de cas de test permettant de vérifier un produit logiciel. L'enchaînement des jeux de test est relatif à une stratégie de test précisée dans le plan de test et les spécifications.

**Matrice de Couverture Fonctionnelle (MCF)** : Document permettant de vérifier la couverture fonctionnelle. Il s'agit de vérifier que chaque cas de test est "couvert" par un ou plusieurs jeux de test et que chaque jeu de test " couvre " au moins une exigence de test.

**Plan de Test** : Document décrivant la démarche générale de test associée au développement d'un logiciel : la stratégie de test, le périmètre (la couverture), les critères d'arrêt, les moyens à mettre en œuvre, la planification. Sa rédaction intervient durant la phase de définition du projet, il est ensuite mis à jour au cours de l'évolution du projet et du développement du produit logiciel.

**Rapport de Test** : Partie du Dossier de Test rapportant le résultat et l'analyse du passage de chaque jeu de test.

**Recette utilisateur** : Vérifier que le logiciel répond aux besoins « métiers » des utilisateurs.

**Scénarios de Test** : Ensemble des jeux de test cohérents permettant de vérifier un objectif particulier (fonctionnel, performance, etc.).

**Tests d'intégration** : Tests réalisés pour vérifier le bon fonctionnement de la structure du produit et des technologies utilisées (unités centrales, réseaux, architectures, etc.).

**Tests de Non Régression** : Vérification, pour une nouvelle version, que les éléments conservés de l'ancienne version du produit logiciel fonctionne sans dégradation (technique, fonctionnelle, performance) par rapport à la version précédente.

**Tests de recette métier** : Vérification de l'adéquation du logiciel par rapport aux besoins des utilisateurs. Cette recette reprend, en les complétant par les besoins " métiers ", les validations précédentes.

**Tests unitaires** : Vérification qu'un composant d'un produit logiciel est conforme à ses spécifications (techniques et fonctionnelles).

**Tests de Validation** : Vérification de la conformité du produit logiciel livré par rapport aux spécifications.

**Workflow d'une anomalie** : Graphe de changement d'état définissant les règles de passage (et les acteurs concernés) d'un état à l'autre d'une anomalie.



## Lexiques des tests logiciels du CFTL/ISTQB

*Ce lexique s'est inspiré du document d'International Software Testing Qualification Board version 3.2F de 2018.*

**Accessibilité : (accessibility)** La mesure de la façon dont un composant ou un système peut être utilisée par des personnes couvrant le plus grand nombre de caractéristiques et de capacités pour atteindre un but précis dans un contexte d'utilisation précis.

**Amélioration de processus : (process improvement)** Un programme d'activités conçues pour améliorer la performance et la maturité des processus de l'organisation, et le résultat d'un tel programme.

**Amélioration du processus de test : (test process improvement)** Programme d'activités visant à améliorer le rendement et la maturité des processus de test de l'organisation et les résultats d'un tel programme.

**Analyse de test : (test analysis)** L'activité qui identifie les conditions de test en analysant les bases de test.

**Analyse de risque : (risk analysis)** Le processus général d'identification et d'évaluation des risques.

**Analyse des causes racines : (root cause analysis)** Une technique d'analyse au but d'identifier les causes premières de défauts. En dirigeant les mesures correctives sur les causes premières, on espère que la probabilité de réapparition des défauts soit minimisée.  
Synonyme : Analyse causale

**Analyse des valeurs limites : (boundary value analysis)** une technique de conception de tests boîte noire dans laquelle les cas de tests sont conçus sur la base des valeurs limites.  
Voir aussi : Valeur limite

**Analyse dynamique : (dynamic analysis)** le processus d'évaluation du comportement (p.ex. performance mémoire, utilisation CPU) d'un composant ou système pendant son exécution

**Analyse d'impact : (impact analysis)** L'identification de tous les produits d'activités touchés par un changement, y incluant une estimation des ressources nécessaires pour accomplir le changement.

**Analyse statique : (static analysis)** Processus d'évaluation d'un composant ou d'un système sans l'exécuter, en fonction de sa forme, de sa structure, de son contenu ou de sa documentation.

**Anomalie : (anomaly)** toute condition qui dévie des attentes basées sur les exigences de spécifications, documents de conception, documents utilisateurs, standards etc, ou des perceptions ou expériences de quelqu'un. Les anomalies peuvent être trouvées pendant, mais pas uniquement, les revues, tests, analyses, compilations ou utilisation des produits logiciels ou de la documentation applicable  
Voir aussi : défauts, déviation, erreur, faute, défaillance, incident, problème

**Approche du test : (test approach)** l'implémentation de la stratégie de test pour un projet spécifique.

**Aptitude fonctionnelle** : La mesure selon laquelle un composant ou un système fournit des fonctions qui répondent aux besoins exprimés et implicites lorsqu'il est utilisé dans des conditions spécifiées.  
Synonymes : fonctionnalité

**Article de test : (test item)** Une partie d'un objet de test utilisé dans le processus de test.  
Voir aussi : Objet de test

**Assurance qualité : (quality assurance)** partie de la gestion de la qualité visant à fournir l'assurance que les exigences qualité seront atteintes.

**Audit : (audit)** examen indépendant d'un produit du travail, d'un processus ou d'un ensemble de processus qui est effectué par un tiers pour évaluer la conformité aux spécifications, aux normes, aux accords contractuels ou à d'autres critères.

**Automatisation des tests : (test automation)** utilisation de logiciels pour exécuter ou supporter des activités de tests, p.ex. gestion des tests, conception des tests, exécution des tests ou vérification des résultats.

**Base de test : (test basis)** L'ensemble des connaissances utilisées comme base pour l'analyse et la conception des tests.

**Bouchon : (stub)** une implémentation spéciale ou squelettique d'un composant logiciel, utilisé pour développer ou tester un composant qui l'appelle ou en est dépendant. Cela remplace un composant appelé.

**Burndown chart** : Un graphique partagé qui représente l'effort réalisé par rapport au temps dans une itération. Il montre l'état et la tendance de la complétion des tâches de l'itération. L'axe des X représente généralement les jours du sprint, tandis que l'axe des Y représente l'effort restant (habituellement soit en heures d'ingénierie optimales, soit en points).

**Calendrier d'exécution des tests : (test execution schedule)** Un ordonnancement de l'exécution des suites de tests au sein d'un cycle de test.

**Cas de test : (test case)** Un ensemble de conditions préalables, de données d'entrée, d'actions (le cas échéant), de résultats attendus et de postconditions, élaboré sur la base des conditions de test.

**Cas de test de bas niveau : (low level test case)** un cas de test avec des valeurs concrètes en entrée et en sortie.

Voir aussi : Cas de test de haut niveau

Synonymes : Cas de test concret

**Cas de test de haut niveau : (high level test case)** Un cas de test sans valeurs concrètes pour les données d'entrée et les résultats attendus.

Voir aussi : Cas de test de bas niveau

Synonymes : cas de test abstrait, cas de test logique

**Cas d'utilisation : (user case)** Séquence de transactions dans un dialogue entre un acteur et un composant ou un système avec un résultat concret. L'acteur peut être un utilisateur ou tout ce qui peut échanger des informations avec le système.

**Caractéristique : (feature)** Un attribut d'un composant ou système, spécifié ou suggéré par la documentation d'exigences (p.ex. contraintes de fiabilité, disponibilité ou de conception).

Synonyme : caractéristique logicielle

**Caractéristique de qualité : (quality characteristic)** Catégorie d'attributs d'un produit qui a une incidence sur la qualité.

Synonymes : caractéristique de qualité, caractéristique du produit logiciel, caractéristique de qualité du logiciel, attribut qualité.

**Cause racine : (root cause)** Une source de défaut telle que si elle est retirée, l'apparition de ce type de défaut est diminuée ou supprimée.

**Charte de test : (test charter)** Documentation des activités de test dans le cadre de tests exploratoires basés sur des sessions.

Voir aussi : Tests exploratoires

Synonyme : Charte

**Chemin : (path)** séquence d'événements, p.ex instructions exécutables, d'un composant ou système d'un point d'entrée jusqu'à un point de sortie.

Synonyme : Chemin de flot de contrôle

**Clôture des tests : (test completion)** L'activité qui rend les actifs de test disponibles pour une utilisation ultérieure, laisse les environnements de test dans un état satisfaisant et communique les résultats des tests aux parties prenantes concernées.

**Code inatteignable : (unreachable code)** Code qui ne peut être atteint et est de ce fait impossible à exécuter.

Synonyme : Code mort

**Compatibilité : (compatibility)** le degré de capacité d'un composant ou d'un système à échanger des informations avec d'autres composants ou systèmes.

**Complexité : (complexity)** le degré par lequel un composant ou système a une conception et/ou une structure interne qui est difficile à comprendre, maintenir et vérifier.

Voir aussi Complexité Cyclomatique.

**Composant : (component)** un élément logiciel minimal qui peut être testé isolément.

Synonymes : module, unité

**Comportement : (behavior)** la réponse d'un composant ou d'un système à un ensemble de valeurs d'entrées et de préconditions.

**Conception des tests : (test design)** Activité consistant à dériver et à spécifier des cas de test à partir des conditions de test.

Voir aussi : Spécification de la conception des tests

**Condition : (condition)** expression logique qui peut être évaluée à Vrai ou Faux, p.ex. A>B. Voir aussi *Condition de tests*.

Synonyme: condition de branchement

**Condition de test : (test condition)** Un aspect des bases de test qui est pertinent pour atteindre des objectifs de test spécifiques.

Synonymes : exigence de test, situation de test

**Configuration : (configuration)** composition d'un composant ou système défini par le nombre, la nature et les interconnexions de ses parties constituantes.

**Conformité : (compliance)** capacité d'un produit logiciel à adhérer à des standards, conventions ou consignes dans des lois ou prescriptions similaires.

**Constat : (finding)** Résultat d'une évaluation qui identifie une question, un problème ou une possibilité importante.

**Contrôle des tests : (test control)** Une tâche de gestion des tests qui traite du développement et de l'application d'un ensemble d'actions correctives pour remettre un projet de test sur les rails quand les métriques de suivi indiquent une déviation par rapport à ce qui a été prévu.

Voir aussi : Gestion des tests

**Contrôle qualité : (quality control)** Les activités et les techniques opérationnelles de la gestion de la qualité axées sur l'atteinte des exigences qualité.

**Coût de la qualité : (cost of quality)** le coût total imputé aux activités et problèmes liés à la qualité, souvent divisé en coûts de prévention, coûts d'estimation, coûts des défaillances internes et coûts des défaillances externes.

**Couverture : (coverage)** le degré, exprimé en pourcentage, selon lequel un élément de couverture spécifié a été exécuté lors d'une suite de test.

Synonyme : couverture de test

**Couverture de code : (code coverage)** une méthode d'analyse qui détermine quelles parties du logiciel ont été exécutées (couvertes) par une suite de tests et quelles parties ne l'ont pas été, p.ex. couverture des instructions, des décisions ou des conditions.

**Couverture des décisions : (decision coverage)** La couverture des résultats des décisions.

**Couverture des instructions : (statement coverage)** le pourcentage des instructions exécutables qui ont été exécutées par une suite de tests.

**Couverture structurelle : (structural coverage)** mesure de couverture basée sur la structure interne du composant.

**Critère d'acceptation : (acceptance criteria)** le critère de sortie que doit satisfaire un composant ou un système de façon à être accepté par un utilisateur, client ou une autre entité autorisée.

**Critères d'entrée : (entry criteria)** L'ensemble des conditions pour le démarrage formel d'une tâche définie.

Synonyme : définition du prêt

**Cycle de test : (test cycle)** Exécution du processus de test sur une version unique et identifiable d'un objet de test.

**Cycle de vie du développement logiciel : (software development lifecycle)** Les activités réalisées à chaque étape du développement d'un logiciel et la façon dont elles sont reliées logiquement et chronologiquement.

**Cycle de vie logiciel : (software lifecycle)** Une période temporelle qui commence lorsqu'un produit logiciel est conçu et se termine lorsque le logiciel n'est plus disponible à l'usage. Le cycle de vie logiciel inclut typiquement une phase de mûrissement, une phase d'exigences, une phase de conception, une phase d'implémentation, une phase de test, une phase d'installation et livraison, une phase d'opération et de maintenance, et parfois une phase de retrait. Note : ces phases peuvent se recouper ou être exécutées de façon itérative.

**Déboguer : (debugging)** le processus de trouver, analyser et éliminer les causes de défaillance dans les logiciels.

**Décision : (decision)** Type d'instruction dans lequel le choix entre deux ou plusieurs résultats possibles contrôle l'ensemble des actions qui en résulte.

**Défaillance : (failure)** Événement dans lequel un composant ou un système n'exécute pas une fonction requise dans les limites spécifiées.

**Défaut : (defect)** Une imperfection ou une déficience d'un produit d'activités lorsqu'il ne répond pas à ses exigences ou à ses spécifications.  
Synonymes : bug, faute

**Densité de défauts : (defect density)** Le nombre de défauts par unité de mesure d'un produit d'activités.  
Synonyme : densité de fautes

**Diagramme d'États : (state diagram)** Un diagramme qui décrit les états que peut avoir un composant ou système, et qui montre les événements ou circonstances qui causent et/ou résultent en un changement d'un état à un autre.  
Synonyme : Diagramme états-transitions

**Disponibilité : (availability)** degré par lequel un composant ou un système est opérationnel et accessible quand il doit être utilisé.

**Données de test : (test data)** Données créées ou sélectionnées pour satisfaire les préconditions d'exécution et les entrées pour exécuter un ou plusieurs cas de test.

**Échec : (fail)** un test est considéré en échec si son résultat obtenu ne correspond pas au résultat attendu.

**Effet de sonde : (probe effect)** L'effet sur le composant ou le système de l'instrument de mesure lorsque le composant ou le système est mesuré, p. ex. par un outil de test de performance ou un moniteur. Par exemple, les performances peuvent être légèrement inférieures lorsque des outils de test de performance sont utilisés.

**Efficacité : (effectiveness)** Mesure selon laquelle les objectifs corrects et complets sont atteints.  
Voir aussi : Efficience

**Efficience : (efficiency)** Ressources utilisées par rapport à la mesure selon laquelle les utilisateurs atteignent des objectifs précis.  
Voir aussi : Efficacité

**Élément de configuration : (configuration item)** un ensemble de matériels, logiciels (ou les deux), qui fait partie de la gestion de configuration et est traité comme une entité unitaire dans le processus de gestion de configuration.

**Élément de couverture : (coverage item)** un attribut ou une combinaison d'attributs qui est dérivé d'une ou de plusieurs conditions de test en utilisant une technique de test qui permet de mesurer la précision de l'exécution du test.

**Entrée : (input)** Données reçues par un composant ou un système à partir d'une source externe.

**Entrées de test : (test input)** Données reçues d'une source externe par l'objet de test pendant son exécution. Les sources externes peuvent être matérielles, logicielles ou humaines.

**Environnement de test : (test environment)** Environnement contenant le matériel, les instruments, les simulateurs, les outils logiciels et les autres éléments de support nécessaires à l'exécution d'un test.

Synonyme : Banc de test

**Environnement opérationnel : (operational environment)** produits matériels et logiciels installés sur un site utilisateur ou client où le composant, ou système, à tester sera utilisé. Le logiciel peut inclure systèmes d'exploitation, des gestionnaires de bases de données et d'autres applications.

**Erreur : (error)** action humaine produisant un résultat incorrect.

Synonyme : confusion

**Estimation d'erreur: (error guessing)** Technique de test qui consiste à dériver les tests sur la base de la connaissance des défaillances passées ou de la connaissance générale des modes de défaillance.

**Estimation des tests : (test estimation)** Approximation calculée d'un résultat lié à divers aspects du test (ex. : effort accompli, date d'achèvement, coûts induits, nombre de cas de test, etc.) qui est utilisable malgré des données d'entrée incomplètes, incertaines ou altérées.

**Exécution des tests :** Processus consistant à exécuter un test sur un composant ou système en test, en produisant le(s) résultat(s) obtenu(s).

**Exercé : (exit criteria)** un élément d'un programme est dit être exercé par un cas de test quand les valeurs d'entrée causent l'exécution de cet élément, tel qu'une instruction, décision ou autre élément de structure.

**Exigence : (requirement)** Une stipulation qui énonce des critères à satisfaire.

**Exigence fonctionnelle : (functional requirement)** une exigence qui spécifie une fonction qu'un composant ou système doit être capable de remplir.

**Exigence non-fonctionnelle : (non-functional requirement)** Exigence qui décrit comment le composant ou le système réalisera ce qu'il est censé faire.

**Extreme programming (XP):** Une méthodologie d'ingénierie de logiciels utilisée dans le développement logiciel Agile. Les pratiques de base sont la programmation par paire, la revue de code approfondie, les tests unitaires de tout le code, la simplicité et la clarté du code.

Voir aussi : Développement de logiciels en Agile

**Exigence testable : (testable requirements)** Exigence exprimée en des termes permettant le démarrage de la conception des tests (et ultérieurement des cas de test) et l'exécution des tests pour déterminer si l'exigence a été satisfaite.

**Facilitateur : (facilitator)** Le leader et la principale personne responsable d'un processus d'inspection ou de revue.

Voir aussi : Modérateur

**Fiabilité : (reliability)** La mesure selon laquelle un composant ou un système exécute des fonctions spécifiées dans des conditions spécifiées pendant une période de temps spécifiée.

**Flot de contrôle : (control flow)** une représentation abstraite de toutes les séquences d'événements (chemins) dans l'exécution d'un composant ou système.

**Flot de données : (data flow)** une représentation abstraite de la séquence et des modifications possibles de l'état des objets de données, où l'état d'un objet est soit création, utilisation ou destruction.

**Fuite mémoire : (memory leak)** une défaillance d'accès à la mémoire causée par un défaut dans la logique d'allocation dynamique de l'espace de stockage d'un programme. Cette défaillance fait que le programme ne libère pas la mémoire quand il a fini de l'utiliser, causant au bout du compte la défaillance de ce programme et/ou d'autres processus concurrents par manque de mémoire.

**Framework de test unitaire : (unit test framework)** Outil fournissant un environnement pour tests unitaires ou de composant dans lequel un composant peut être testé de façon isolée ou avec des bouchons ou pilotes appropriés. Il fournit aussi une aide au développeur, telle que des fonctionnalités de débogage.

**Gestion de configuration : (configuration management)** une discipline appliquant une direction et surveillance technique et administrative pour identifier et documenter les caractéristiques fonctionnelles et physiques d'un élément de configuration, contrôler les modifications de ces caractéristiques, enregistrer et informer des modifications et états d'implémentation, et vérifier la conformité avec des exigences spécifiées.

**Gestion des défauts : (defect management)** Processus qui consiste à identifier et à enregistrer les défauts, à les classer, à les examiner, à prendre des mesures pour les résoudre et à les clôturer lorsqu'ils sont résolus.

Voir aussi : Gestion des incidents

**Gestion des risques : (risk management)** Les activités coordonnées pour piloter et contrôler une organisation en matière de risque.

**Gestion des tests : (test management)** La planification, l'ordonnancement, l'estimation, la suivi, la production de rapports, le contrôle et la clôture des activités de test.

**Gestion de la qualité : (quality management)** activités coordonnées pour diriger et contrôler une organisation en ce qui concerne la qualité. La direction et le contrôle de la qualité inclus généralement la mise en place de recommandations et d'objectifs qualité, de la planification qualité, du contrôle qualité, de l'assurance qualité et de l'amélioration qualité.

**Guide d'installation : (installation guide)** instructions fournies sur un média adéquat, qui guide l'installateur au cours du processus d'installation. Ceci peut être un manuel, des procédures pas à pas, un démon d'installation, ou tout autre description similaire de processus.

**Harnais de test : (test harness)** Environnement comprenant les bouchons et les pilotes nécessaires pour exécuter un test.

**IDEAL** : Un modèle d'amélioration de l'organisation qui sert de feuille de route pour démarrer, planifier et mettre en oeuvre des actions d'amélioration. Le modèle IDEAL est nommé d'après les cinq phases qu'il décrit: le démarrage (Initiating), le diagnostic (Diagnosing), la planification (Establishing), la mise en oeuvre (Acting), et l'apprentissage (Learning).

**IHM : (GUI)** Acronyme d'Interface Homme/Machine

**Implémentation des tests : (test implementation)** L'activité qui prépare le testware nécessaire à l'exécution des tests sur la base de l'analyse et de la conception des tests.

**Indépendance du test : (independence of testing)** Séparation des responsabilités qui favorise l'atteinte des objectifs de test.

**Indicateur de performance : (performance indicator)** un métrique de haut niveau de rentabilité et/ou d'efficacité utilisé pour guider et contrôler le développement progressif, p. ex., le glissement du temps de développement logiciel.

Synonyme : Indicateur clé de performance

**Infrastructure de test : (test infrastructure)** Ressources organisationnelles nécessaires pour réaliser des tests telles que les environnements de test, les outils de test, les procédures et un environnement bureautique.

**Inspection : (inspection)** Un type de revue formelle visant à identifier les problèmes d'un produit d'activités, qui fournit des mesures pour améliorer le processus de revue et le processus de développement de logiciels.

**Instruction : (statement)** une entité dans un langage de programmation, qui est typiquement la plus petite unité indivisible d'exécution.

Synonyme : Instruction dans le code source

**Intégration : (integration)** le processus de combiner des composants ou systèmes en assemblages plus grands.

**Intelligibilité : (understandability)** Capacité d'un produit logiciel de permettre à l'utilisateur de comprendre si le logiciel est approprié, et comment il peut être utilisé pour des tâches particulières et des conditions d'utilisation.

Voir aussi : Utilisabilité

**Interface utilisateur : (user interface)** Tous les composants d'un système qui fournissent des informations et des contrôles permettant à l'utilisateur d'accomplir des tâches spécifiques avec le système.

**Interopérabilité : (interoperability)** La mesure selon laquelle deux ou plusieurs composants ou systèmes peuvent échanger de l'information et utiliser l'information qui a été échangée.

**Instruction exécutable : (executable statement)** une instruction qui, quand elle est compilée, est traduite en code objet, et sera exécutée de façon procédurale quand le programme s'exécute et peut effectuer une action sur des données ou le déroulement du programme.

**Intégration fonctionnelle : (functional integration)** une approche d'intégration qui combine les composants ou systèmes dans le but de permettre rapidement le fonctionnement d'une fonctionnalité de base.

Voir aussi : Tests d'intégration

**Jalon : (milestone)** moment particulier dans un projet auquel des livrables définis (intermédiaires ou non) et des résultats doivent être prêts.

**Lecture basée sur la perspective : (perspective-based reading)** Technique de revue par laquelle les réviseurs évaluent le produit du travail à partir de différents points de vue.

Synonyme : revue basée sur la perspective

**Logiciel : (software)** Programmes informatiques, procédures et, éventuellement, documentation et données connexes se rapportant au fonctionnement d'un système informatique.

**Logiciel commercial sur étagère : (commercial off-the-shelf software (COTS))** Produit logiciel qui est développé pour le marché de façon générale, c'est-à-dire pour un grand nombre de clients, et qui est livré à de nombreux clients dans un format identique.

Synonyme : Logiciel sur étagère

**Maintenabilité : (maintainability)** La mesure selon laquelle un composant ou un système peut être modifié par les personnes chargées de la maintenance.

**Maintenance : (maintenance)** Processus de modification d'un composant ou d'un système après la livraison pour corriger les défauts, améliorer les attributs de qualité ou s'adapter à un environnement modifié.

**Maturité : (maturity)**

- (1) La capacité d'une organisation quant à l'efficacité et à l'efficience de ses processus et de ses pratiques de travail.
- (2) La mesure selon laquelle un composant ou un système répond aux besoins de fiabilité en fonctionnement normal.

**Mesure : (measure)** le nombre ou la catégorie affectée à un attribut d'une entité à la suite d'une action de mesure.

**Métrique : (metric)** une échelle de mesure et la méthode utilisée pour la mesure.

**Modèle de croissance de fiabilité : (reliability growth model)** Un modèle montrant l'augmentation de la fiabilité dans le temps lors d'un test continu d'un composant ou système, suite à la suppression de défauts qui provoquaient des défaillances de fiabilité.

**Modèle de cycle de vie : (lifecycle model)** Une description des processus, des flux de travail et des activités utilisés dans le développement, la livraison, la maintenance et la mise hors service d'un système.

Voir aussi : Cycle de vie des logiciels

**Modèle de développement incrémental : (incremental development model)** Un modèle de cycle de vie du développement dans lequel la portée du projet est généralement déterminée au début du

cycle de vie du projet, mais les estimations de temps et de coûts sont régulièrement modifiées à mesure que l'équipe du projet comprend mieux le produit. Le produit est développé au moyen d'une série de cycles répétés, chacun fournissant un incrément qui ajoute successivement à la fonctionnalité du produit.

Voir aussi : modèle de développement itératif

**Modèle de développement itératif : (iterative development model)** un modèle de cycle de développement où le projet est séparé en un nombre d'itérations (souvent nombreuses). Une itération est une boucle complète de développement résultant en une livraison (interne ou externe) d'un produit exécutable, un sous-ensemble du produit final en développement, qui grandit d'itération en itération pour devenir le produit fini.

**Modèle de développement séquentiel : (sequential development model)** Un type de modèle de cycle de vie de développement selon lequel un système achevé est développé de manière linéaire en plusieurs phases distinctes et successives, sans chevauchement entre elles.

**Modèle en V : (V-model)** Un modèle de cycle de vie séquentiel du développement décrivant une relation un pour un entre les principales phases du développement

**Modérateur : (moderator)** Une personne neutre qui effectue une session de test d'utilisabilité.

Voir aussi : facilitateur

Synonyme : leader de l'inspection

**Niveau de risque : (risk level)** Mesure qualitative ou quantitative d'un risque défini par l'impact et la probabilité.

Synonyme : Exposition au risque

**Niveau de test : (test level)** Une instanciation spécifique d'un processus de test.

Synonymes : Phase de test

**Norme : (standard)** Ensemble d'exigences formelles et certaines fois obligatoires. Cet ensemble est développé et utilisé afin de préconiser des approches à la façon de travail ou pour fournir des directives (par exemple, les normes ISO/IEC, IEEE et les normes organisationnelles).

**Objectif de test : (test objective)** une raison ou but de la conception et l'exécution d'un test.

**Objet de test : (test object)** Composant ou système à tester.

Voir aussi : Article de test

**Oracle de test : (test oracle)** Source utilisée pour déterminer les résultats attendus à comparer avec les résultats obtenus de l'application en cours de test.

**Outil d'analyse dynamique : (dynamic analysis tool)** un outil qui fournit une information sur l'état d'un code logiciel. Ces outils sont habituellement utilisés pour identifier des pointeurs non affectés, vérifier l'arithmétique des pointeurs et pour suivre l'allocation, l'utilisation et la libération de mémoire et détecter les fuites mémoire.

**Outil de conception de tests : (test design tool)** Outil qui de support à l'activité de conception de test en générant les valeurs d'entrées des test à partir de spécifications qui peuvent être présentes dans un répertoire d'outil CASE, (ex. : outil de gestion des exigences) ou de conditions spécifiques de tests contenues dans l'outil lui-même.

**Outil de couverture : (coverage tool)** un outil qui fournit des mesures objectives de quels éléments structurels (p.ex. instructions, branches) ont été employés lors de l'exécution d'une suite de tests.  
Synonyme : outil de mesure de couverture

**Outil d'exécution des tests : (test execution tool)** Un outil de test qui exécute des tests par rapport à un élément de test donné et évalue les résultats par rapport aux résultats attendus et aux postconditions.

**Outil de gestion de configuration: (configuration management tool)** un outil qui permet d'aider à identifier et contrôler la configuration des composants, leur état en regard des changements et versions, ainsi que la production de baselines constituées d'éléments de configuration.

**Outil de gestion des défauts : (defect management tool)** Un outil qui facilite l'enregistrement et le suivi de l'état des défauts.

Voir aussi : outil de gestion des incidents

Synonymes : outil de suivi des bogues, outil de suivi des défauts.

**Outil de gestion des exigences : (requirements management tool)** un outil qui supporte la consignation des exigences, des attributs des exigences (p.ex. priorité, connaissance responsable) et des annotations, et facilite la traçabilité au travers des couches d'exigences et de la gestion des modifications des exigences. Quelques outils de gestion des exigences fournissent aussi des facilités pour l'analyse statique, tel que la vérification de cohérence et la violation de règles pré-définies de spécification des exigences.

**Outil de gestion des tests : (test management tool)** Outil d'assistance à la gestion des tests et de contrôle partiel du processus de test. Il offre souvent de nombreuses fonctionnalités telles que la gestion du testware, la planification des tests, la traçabilité des résultats, le suivi d'avancement, la gestion des incidents et le Reporting.

**Outil de monitoring : (Monitoring tool)** Outil logiciel ou dispositif matériel qui est utilisé en même temps que le composant ou le système testé et qui supervise, enregistre et/ou analyse le comportement du composant ou du système testé.

Voir aussi : Outil d'analyse dynamique

**Outil de préparation des données de tests : (test data preparation tool)** Type d'outil permettant, à partir de bases de données existantes, de sélectionner, créer, générer, manipuler et éditer des données pour utilisation pendant des tests.

Synonymes : générateur de test

**Outil de test : (test tool)** Produit logiciel qui supporte une ou plusieurs activités de test, telles que la planification, le contrôle, la spécification, la conception des fichiers et données initiaux, l'exécution et l'analyse des tests.

Voir aussi : CAST

**Outil de tests de performances : (performance testing tool)** Outil de test qui génère une charge pour un élément de test désigné et qui mesure et enregistre sa performance pendant l'exécution du test.

**Partition d'équivalence : (equivalence partition)** une portion d'un domaine d'entrée ou de sortie pour laquelle le comportement d'un composant ou système est supposé être le même, basé sur ces spécifications.

Synonyme : classe d'équivalence

**Partitionnement d'équivalence : (equivalence partitioning)** Technique de test boîte-noire dans laquelle les cas de test sont conçus pour exercer des partitions d'équivalence en utilisant un élément représentatif de chaque partition.

Synonyme : tests de partitions

**Performance : (performance efficiency)** La mesure selon laquelle une composante ou un système utilise le temps, les ressources et la capacité dans l'accomplissement de ses fonctions assignées.

Synonymes : Comportement dans le temps, Efficience de la performance.

**Pilote : (driver)** un composant logiciel ou outil de tests qui remplace un composant qui contrôle et/ou appelle un composant ou système.

**Pilotage des tests : (test monitoring)** Une activité de gestion des tests qui comprend la vérification de l'état des activités de test, l'identification de tout écart par rapport à l'état planifié ou attendu, et le reporting aux parties prenantes.

Voir aussi : Gestion des tests

**Plan de test : (test plan)** Documentation décrivant les objectifs de test à atteindre et les moyens et le calendrier pour les atteindre, organisée pour coordonner les activités de test.

**Plan de test maître : (master test plan)** Un plan de test qui est utilisé pour coordonner plusieurs niveaux de test ou types de test.

Voir aussi : Plan de test

**Plan de revue : (review plan)** Un document décrivant l'approche, les ressources et le calendrier des activités de la revue envisagée. Il identifie, entre autres, les documents et le code devant être revus, aussi bien que les critères d'entrée et de sortie à appliquer en cas de revue formelle et la logique ayant motivé leur choix. C'est un rapport du processus de planification de revue.

**Planification de test : (test planning)** Activité de définition ou de mise à jour d'un plan de test.

**Planning Poker : (planning poker)** Une technique d'estimation basée sur le consensus, utilisée la plupart du temps pour estimer l'effort ou la taille relative des User Stories en développement logiciel Agile. C'est une variation de la méthode Wideband Delphi utilisant un jeu de cartes avec des valeurs représentant les estimations de charges de l'équipe.

Voir aussi : Développement logiciel Agile, Wideband Delphi

**Planning de test : (test schedule)** Liste d'activités, de tâches ou d'évènements du processus de test mentionnant leurs dates et/ou heures de début et fin et interdépendances.

**Politique de test : (test policy)** Document de haut niveau décrivant les principes, approches et objectifs majeurs de l'organisation concernant l'activité de test.

Synonyme : politique de test organisationnelle

**Portabilité : (portability)** facilité avec laquelle un produit logiciel peut être transféré d'un environnement matériel ou logiciel vers un autre.

**Postcondition : (postcondition)** L'état attendu d'un élément de test et de son environnement à la fin de l'exécution d'un cas de test.

**Précondition : (precondition)** L'état requis d'un élément de test et de son environnement avant l'exécution d'un cas de test.

**Priorité : (priority)** un niveau d'importance (métier) affecté à un élément (p.ex. à un défaut).

**Problème : (problem)** Une cause sous-jacente inconnue d'un ou de plusieurs incidents.

**Procédure de test : (test procedure)** Une séquence de cas de test dans l'ordre d'exécution, et toutes les actions associées qui peuvent être nécessaires pour mettre en place les préconditions initiales et toutes les activités de bouclage après l'exécution.

Voir aussi : Script de test

**Processus : (process)** ensemble d'activités liées qui transforment des entrées en sorties.

**Processus de test : (test process)** Ensemble d'activités interdépendantes comprenant la planification des tests, le suivi et le contrôle des tests, l'analyse de test, la conception des tests, l'implémentation des tests, l'exécution des tests et la clôture des tests.

**Projet : (project)** un projet est un ensemble unique d'activités, contrôlées et coordonnées, avec des dates de début et de fin, effectuées avec pour objectif de conformité à des exigences spécifiques, incluant des contraintes de temps, de coût et de ressources.

**Qualité : (quality)** degré par lequel un composant, système ou processus atteint des exigences spécifiées et/ou des besoins ou attentes des clients ou utilisateurs.

**Qualité logicielle : (software quality)** La totalité des fonctionnalités et caractéristiques d'un produit logiciel qui influent sur sa capacité à satisfaire des besoins déclarés ou implicites.

Voir aussi : Qualité

**Rapport d'avancement de test : (test progress report)** Un rapport de test produit à intervalles réguliers sur l'état d'avancement des activités de test par rapport à une base de référence, aux risques et aux solutions alternatives nécessitant une décision.

Synonyme : Rapport sur le statut des tests

**Rapport de défaut : (defect report)** Documentation de la survenance, de la nature et de l'état d'un défaut.

Voir aussi : Rapport d'incident

Synonymes : rapport de bogue

**Rapport de synthèse de test : (test summary report)** Un rapport de test qui fournit une évaluation des éléments de test en comparaison avec les critères de sortie.

Synonyme : rapport de test

**Réalisation d'une mesure : (measurement)** le processus d'attribution d'une mesure chiffrée ou d'une catégorie à une entité pour décrire un attribut de cette entité.

**Réussite : (pass)** un test est considéré comme réussi si les résultats effectifs correspondent aux résultats attendus.

**Résultat de décision : (decision outcome)** Le résultat d'une décision qui détermine la prochaine instruction à exécuter.

**Rapport d'incident : (incident report)** Documentation de l'événement, de la nature et de l'état d'un incident.

Synonymes : rapport de déviation, rapport d'incident de test logiciel, rapport d'incident de test

**Rapport de test** : Documentation résumant les activités et les résultats des tests.

**Rational Unified Process (RUP)**: Un cadre propriétaire et adaptable décrivant le processus itératif de développement d'un logiciel. Le processus comprend quatre phases du cycle de vie du projet: création, élaboration, construction et transition.

**Réduction des risques : (risk mitigation)** Le processus par lequel les décisions sont prises et les mesures de protection sont mises en œuvre pour réduire ou maintenir les risques à des niveaux spécifiés.

Synonyme : Contrôle des risques

**Régression : (regression)** Dégradation de la qualité d'un composant ou d'un système en raison d'un changement.

**Relecture technique : (walkthrough)** Type de revue dans laquelle un auteur conduit la revue en parcourant avec les participants un produit d'activités. Les participants posent des questions et font des commentaires sur des problèmes possibles.

Synonymes : Relecture technique Structurée

**Reporting de test : (test reporting)** Collecte et analyse des données provenant des activités de test, puis consolidation des données dans un rapport pour informer les parties prenantes.

Voir aussi : Processus de test

**Résultat : (result)** les conséquences / résultats de l'exécution d'un test. Cela inclut les sorties vers des écrans, les modifications de données, rapports et messages d'information envoyés.

Voir aussi : Résultat effectif, Résultat attendu

Synonyme : Résultat du test

**Résultat attendu : (expected result)** Comportement observable prévu d'un composant ou d'un système s'exécutant dans des conditions spécifiées, sur la base de sa spécification ou d'une autre source.

Synonymes : résultat escompté, résultat prévu

**Résultat effectif : (actual result)** comportement produit / observé quand un composant ou un système est testé.

Synonyme : Sortie effective

**Réunion rétrospective : (retrospective meeting)** Une réunion à la fin d'un projet durant laquelle les membres d'une équipe de projet évaluent le projet et tirent des leçons qui peuvent être appliquées au projet suivant.

Synonyme : réunion post-projet

**Réviseur : (reviewer)** Un participant à une revue, qui identifie les problèmes dans le produit du travail.

Synonymes : Vérificateur, Inspecteur

**Revue : (review)** Type de tests statiques au cours desquels un produit d'activités ou un processus de travail est évalué par une ou plusieurs personnes afin de détecter les problèmes et d'apporter des améliorations.

**Revue ad-hoc : (ad hoc reviewing)** technique de revue réalisée par des examinateurs indépendants de manière informelle, sans processus structuré.

**Revue basée sur une checklist : (checklist-based reviewing)** Technique de revue guidée par une liste de questions ou d'attributs demandés.

**Revue basée sur les rôles : (role-based reviewing)** Technique de revue où les réviseurs évaluent un produit d'activités du point de vue des différents rôles des intervenants.

**Revue basée sur des scénarios : (scenario-based reviewing)** Technique de revue où la revue est guidée par la capacité du produit du travail à réaliser des scénarios spécifiques.

**Revue de pairs : (peer review)** Une forme de revue des produits d'activités effectuée par d'autres personnes qualifiées pour effectuer le même travail.

**Revue formelle : (formal review)** Une forme de revue qui suit un processus défini avec une documentation formelle des résultats.

**Revue informelle : (informal review)** Un type de revue qui n'est pas basée sur une procédure formelle (documentée)

**Revue informelle collective : (informal group review)** Une revue informelle effectuée par trois personnes ou plus.

**Revue technique : (technical review)** Type d'examen officiel effectué par une équipe de personnel qualifié sur le plan technique qui examine l'adéquation d'un produit d'activités à l'usage pour l'usage auquel il est destiné et identifie les divergences par rapport aux spécifications et aux normes.

**Risque : (risk)** Un facteur qui pourrait avoir des conséquences négatives à l'avenir.

**Risque produit : (product risk)** Un risque directement lié à l'objet de test.  
Voir aussi : Risque

**Risque projet : (project risk)** Un risque qui a une incidence sur la réussite du projet.  
Voir aussi : Risque

**Risque qualité :** Un risque Produit lié à une caractéristique de qualité.  
Voir aussi : Caractéristique de qualité, Risque produit.

**Robustesse : (robustness)** le degré pour lequel un composant ou système peut fonctionner correctement en présence de données d'entrée invalides ou de conditions environnementales stressantes.  
Voir aussi : Tolérance aux fautes, Tolérance aux pannes.

**Scribe : (scribe)** Une personne qui consigne l'information pendant les réunions de revue.  
Synonyme : Greffier

**Script de test :** Une séquence d'instructions pour l'exécution d'un test.  
Voir aussi : Procédure de test

**Scrum :** un cadre progressif et itératif pour gérer des projets. Généralement utilisé dans le développement logiciel agile.

Voir aussi : Développement logiciel Agile

**Sécurité : (security)** La mesure selon laquelle une composante ou un système protège l'information et les données afin que les personnes ou d'autres composantes ou systèmes aient le degré d'accès approprié à leurs types et niveaux d'autorisation.

**Session de test :** Laps de temps continu passé à exécuter des tests. En test exploratoire, chaque session de test se focalise sur une charte, mais les testeurs peuvent aussi explorer de nouvelles opportunités ou problèmes durant une session. Le testeur crée et exécute des cas de test à la volée et enregistre leurs évolutions.

Voir aussi : Test exploratoire

**Sévérité : (severity)** degré qu'à l'impact d'une anomalie sur le développement ou l'utilisation d'un composant ou système.

**Simulateur : (simulator)** un appareil, programme ou système utilisé pendant les tests, qui se comporte ou fonctionne comme un système donné à la réception d'entrées.

Voir aussi : Emulateur

**Simulation : (simulation)** représentation d'une sélection de caractéristiques de comportement d'un système physique ou abstrait par un autre système.

**Sortie : (output)** Données transmises par un composant ou un système à une destination externe.

**Spécification : (specification)** Un document qui spécifie, idéalement de façon complète, précise et vérifiable, les exigences, conceptions, comportements et autres caractéristiques d'un composant ou système, et souvent, les procédures pour déterminer si ces stipulations ont été satisfaites.

**Spécification de cas de test : (test case specification)** Documentation d'un ensemble constitué d'un ou plusieurs cas de test.

Voir aussi : Spécification de test

**Spécification de composant : (component specification)** une description des fonctions de composant en termes de valeurs de sortie pour des valeurs d'entrée spécifiées sous des conditions spécifiées, ainsi que les comportements non-fonctionnels requis (p.ex; utilisation des ressources)

**Stabilité : (stability)** La mesure selon laquelle un composant ou un système peut être modifié de manière efficace et efficiente sans introduire de défauts ou dégrader la qualité d'un produit existant.

**Stratégie de test :** Documentation qui exprime les exigences génériques pour tester dans le cadre d'un ou de plusieurs projets exécutés au sein d'une organisation, fournissant des détails sur la façon dont les tests doivent être effectués, et qui est alignée sur la politique de test.

Synonyme : Stratégie de test organisationnel

**Suite de tests :** Ensemble de cas de test ou de procédures de test à exécuter dans un cycle de test spécifique.

Synonymes : Suite de cas de test, Ensemble de tests

**Système : (system)** Ensemble d'éléments en interaction organisés pour accomplir une fonction ou un ensemble de fonctions spécifiques.

**Système sous test : (system under test (SUT))** Un type d'objet de test qui est un système.

**Sûreté : (safety)** La capacité qu'un système ne pourra pas, dans des conditions définies, conduire à un état dans lequel la vie humaine, la santé, la sécurité des personnes et des biens ou l'environnement est en danger.

**Table de décision : (decision table)** Un tableau utilisé pour montrer des ensembles de conditions et les actions qui en résultent.

Synonyme : table de décision cause-effet

**Taux de défaillance : (failure rate)** Rapport entre le nombre de défaillances d'une catégorie donnée et une unité de mesure donnée.

**Technique de test : (test technique:)** Procédure utilisée pour dériver et/ou sélectionner les cas de test.

Synonymes : technique de conception de cas de test, technique de spécification de test, technique de conception de test

**Technique de test basée sur l'expérience : (experience-based test technique)** Procédure pour obtenir et/ou sélectionner des cas de test basée sur l'expérience du testeur, sa connaissance et son intuition.

Synonyme : technique basée sur l'expérience

**Technique de test boîte-blanche : (white-box test technique)** Procédure permettant de dériver et/ou de sélectionner des cas de test sur la base d'une analyse de la structure interne d'un composant ou d'un système.

Synonymes : technique de test structurel, technique de test basé sur la structure, technique basée sur la structure, technique boîte-blanche

**Test : (test)** un ensemble d'un ou plusieurs cas de tests.

**Test d'acceptation : (acceptance testing)** test formel en rapport avec les besoins utilisateur, exigences et processus métier, réalisé pour déterminer si un système satisfait ou non aux critères d'acceptation et permettre aux utilisateurs, clients ou autres entités autorisées de déterminer l'acceptation ou non du système.

Voir aussi : Test d'acceptation utilisateur

**Tests d'acceptation contractuelle : (contractual acceptance testing)** tests d'acceptation effectués pour vérifier si un système satisfait à ses exigences contractuelles.

**Test d'accessibilité : (accessibility testing)** test qui détermine la facilité avec laquelle les utilisateurs avec handicaps peuvent utiliser un composant ou un système.

**Tests Alpha : (alpha testing)** test opérationnel réel ou simulé conduit dans l'environnement de test du développement par des rôles en dehors de ceux de l'organisation en charge du développement.

**Test basé sur une checklist : (checklist-based testing)** Une technique de conception de tests basée sur l'expérience selon laquelle le testeur expérimenté utilise une liste de haut niveau d'éléments à noter, à vérifier, ou à se rappeler, ou un ensemble de règles ou de critères en fonction desquels un produit doit être vérifié.

**Test basé sur l'expérience: (experience-based testing)** test basé sur l'expérience, les connaissances et l'intuition du testeur.

**Test basé sur les modèles : (model-based testing (MBT))** Tests basés sur ou impliquant des modèles.

**Test basé sur les risques : (risk-based testing)** Tests pour lesquels la gestion, la sélection, la priorisation et l'utilisation des activités et des ressources de test sont basées sur les types de risque et les niveaux de risque correspondants.

**Test basé sur des sessions : (session-based testing)** une approche du test dans laquelle des activités de test sont planifiées comme les sessions ininterrompues de conception de tests et d'exécution. Elle est souvent utilisée en conjonction avec le test exploratoire.

**Tests Beta: (beta testing)** Tests opérationnels simulés ou réels effectués sur un site externe, par des rôles à l'extérieur de l'organisation de développement.

Synonyme : Tests terrain

**Test boîte-blanche : (white box testin)** Test basé sur une analyse de la structure interne du composant ou système.

Synonymes : Test de boîtes transparentes, Test basé sur le code, Test de boîtes de verre, Test de couverture logique, Test dirigé par la couverture logique, Test structurel, Test basé sur la structure.

**Tests d'acceptation utilisateur : (user acceptance testing)** Tests d'acceptation effectués dans un environnement opérationnel réel ou simulé par les utilisateurs prévus en mettant l'accent sur leurs besoins, leurs exigences et leurs processus métier.

Voir aussi : Tests d'acceptation

**Test de cas d'utilisation : (use case testing)** Technique de test boîte noire selon laquelle les cas de tests sont conçus pour exécuter des scénarios de cas d'utilisation.

Synonymes : tests de scénarios, tests de scénarios utilisateurs

**Tests de charge : (Load testing)** Type de test de performance effectué pour évaluer le comportement d'un composant ou d'un système sous des charges variables, habituellement entre des conditions prévues d'utilisation faible, typique et de pointe.

Voir aussi : Tests de performance, tests de stress

**Test de confirmation : (confirmation testing)** tests dynamiques effectués après la correction de défauts dans le but de confirmer que les défaillances causées par ces défauts ne se produisent plus.

Synonyme : retest

**Test de composant : (component testing)** le test de composants logiciels individuels.

Synonymes : test de module, test unitaire

**Tests d'acceptation opérationnelle : (operational acceptance testing)** test opérationnel en phase de test d'acceptation, généralement effectué dans un environnement opérationnel (simulé) par l'exploitation et/ou le personnel de l'administration des systèmes en se concentrant sur les aspects opérationnels, par exemple la reprise (après incident), le comportement des ressources, la facilité d'installation et la conformité technique.

Voir aussi : Test opérationnel

Synonyme : tests d'acceptation de production

**Test des décisions : (decision testing)** Une technique de conception de tests boîte blanche selon laquelle les cas de tests sont conçus pour exécuter les résultats de décisions.

**Tests des instructions : (statement testing)** une technique de conception de tests boîte-blanche dans laquelle les cas de tests sont conçus pour exécuter des instructions.

**Test de maintenance : (maintenance testing)** test des modifications d'un système opérationnel ou de l'impact d'une modification d'environnement sur un système opérationnel.

**Test de sécurité : (security testing)** tests effectués pour déterminer la sécurité d'un produit logiciel.  
Voir aussi : Test des fonctionnalités

**Test de portabilité : (portability testing)** Tests pour déterminer la portabilité d'un produit logiciel.  
Synonyme : Test de configuration

**Tests dirigés par mots-clé : (keyword-driven testing)** une technique de script utilisant des fichiers de données qui contiennent non seulement des données de test et des résultats attendus, mais aussi des mots clé liés à l'application à tester. Les mots clé sont interprétés par des scripts de support spécifiques, appelés par le script de contrôle du test.

Voir aussi : Tests déterminés par les données.

Synonyme : Tests dirigés par mots d'actions

**Tests d'intégration : (integration testing)** tests effectués pour montrer des défauts dans les interfaces et interactions de composants ou systèmes intégrés.

Voir aussi : Tests d'intégration de composants, tests d'intégration système

**Tests d'intégration système : (system integration testing)** Tester la combinaison et l'interaction des systèmes.

**Test d'interopérabilité : (interoperability testing)** Tester pour déterminer l'interopérabilité d'un produit logiciel.

Voir aussi : Tests fonctionnels

Synonymes : tests de compatibilité

**Tests exploratoires : (exploratory testing)** Une approche des tests par laquelle les testeurs conçoivent et exécutent dynamiquement des tests en fonction de leurs connaissances, de l'exploration de l'élément testé et des résultats des tests précédents.

**Test fonctionnel : (functional testing)** Test réalisé pour évaluer la conformité d'un composant ou d'un système aux exigences fonctionnelles.

Voir aussi : test boîte-noire

**Test d'acceptation réglementaire : (regulatory acceptance testing)** Tests d'acceptation effectués pour vérifier si un système est conforme aux lois, politiques et règlements applicables.

**Test d'intégration de composants : (component integration testing)** test effectué pour découvrir des défauts dans les interfaces et les interactions entre des composants intégrés.

Synonyme : test de liens

**Test d'utilisabilité : (usability testing)** Tests visant à évaluer dans quelle mesure le système peut être utilisé par des utilisateurs spécifiques avec efficacité, efficacité et satisfaction dans un contexte d'utilisation spécifique.

**Test dynamique : (dynamic testing)** test qui nécessite l'exécution du logiciel d'un composant ou système.

**Tests non-fonctionnels : (non-functional testing)** Test effectué pour évaluer la conformité d'un composant ou d'un système avec les exigences non fonctionnelles.

**Test par tables de décisions : (decision table testing)** une technique de conception des tests boîte noire dans laquelle les cas de tests sont conçus pour exécuter les combinaisons d'entrées et/ou de stimuli (causes) présentes dans une table de décision.

Voir aussi : Table de décision

**Test de performance : (performance testing)** Tests pour déterminer la performance d'un produit logiciel.

Voir aussi : Test de rendement

**Test Leader :** Dans le contexte de grands projets, la personne qui relève du Test Manager et qui est responsable de la gestion de projet d'un niveau de test particulier ou d'un ensemble particulier d'activités de test.

Voir aussi : Test Manager

Synonyme : Test Lead

**Test Manager :** Personne responsable de la gestion de projet des activités et des ressources de test et de l'évaluation d'un objet de test. La personne qui dirige, contrôle, gère l'organisation et règle l'évaluation d'un objet de test.

**Test piloté par les données : (data-driven testing)** une technique de script qui sauvegarde les entrées de tests et les résultats attendus dans une table ou un tableur, de façon à ce qu'un simple script de contrôle puisse exécuter tous les tests de la table. Les tests déterminés par les données sont souvent utilisés pour assister l'utilisation d'outils de tests automatisés tels ceux de capture/rejeu.

Voir aussi : Tests dirigés par mots-clé

**Test de régression : (regression testing)** tests d'un programme préalablement testé, après une modification, pour s'assurer que des défauts n'ont pas été introduits ou découverts dans des parties non modifiées du logiciel, comme suites des modifications effectuées.

**Technique de test boîte-noire : (black-box test techniques)** procédure documentée pour élaborer et sélectionner des cas de tests basés sur une analyse des spécifications, soit fonctionnelles soit non-fonctionnelles, d'un composant ou système sans faire référence à ses structures internes.

Synonymes : Technique boîte noire, technique basée sur les spécifications, technique de test basée sur les spécifications

**Test des transitions d'état : (state transition testing)** Une technique de test de boîte-noire utilisant un diagramme états-transitions ou une table d'état pour dériver des cas de test afin d'évaluer si l'élément de test exécute avec succès des transitions valides et bloque les transitions invalides.

Voir aussi : Test N-switch

Synonyme : Test de machines à états finis

**Tests statiques : (static testing)** Tester un produit d'activités sans que le code soit exécuté.

**Tests système : (system testing)** Tester un système intégré pour vérifier qu'il répond aux exigences spécifiées.

**Testeur : (tester)** Un professionnel qualifié qui participe aux tests d'un composant ou d'un système.

**Transition d'état** : Une transition entre deux états d'un composant ou système.

**Test : (testing)** Processus consistant en toutes les activités du cycle de vie, statiques et dynamiques, concernant la planification et l'évaluation de produits logiciels et produits liés pour déterminer s'ils satisfont aux exigences et démontrer qu'ils sont conformes aux objectifs et détecter des anomalies.

**Testabilité : (testability)** Le degré d'efficacité et d'efficience avec lequel les tests peuvent être conçus et exécutés pour un composant ou un système.

**Testware : (Traceability)** Produits d'activités réalisés au cours du processus de test pour la planification, la conception, l'exécution, l'évaluation et l'établissement de rapports sur les tests.

**Traçabilité** : La mesure dans laquelle une relation peut être établie entre deux ou plusieurs produits d'activités.

Voir aussi : Traçabilité horizontale, Traçabilité verticale

**Type de risque : (risk type)** Ensemble de risques regroupés par un ou plusieurs facteurs communs.

Synonymes : Catégorie de risque

**Type de test : (test type)** Groupe d'activités de test basées sur des objectifs de test spécifiques visant des caractéristiques spécifiques d'un composant ou d'un système.

**User Story** : Une exigence d'utilisateur ou métier de haut niveau communément utilisée dans le développement de logiciels Agile, qui consiste généralement en une phrase dans le langage courant ou le langage métier capturant la fonctionnalité dont un utilisateur a besoin et la raison de ce besoin, les critères non fonctionnels, ainsi que les critères d'acceptation.

Voir aussi : Développement de logiciels Agile, Exigence

**Valeur limite : (boundary value)** Une valeur minimale ou maximale d'une partition d'équivalence ordonnée.

**Validation : (validation)** Confirmation par l'examen et la fourniture de preuves objectives que les exigences, pour un usage ou une application voulue, ont été satisfaites.

**Variable : (variable)** Élément de stockage dans un ordinateur qui est accessible par un programme logiciel en se référant à son nom.

**Vérification : (verification)** Confirmation par l'examen et la fourniture de preuves objectives que des exigences spécifiées ont été satisfaites.

**Wideband Delphi** : Technique d'estimation de test basée sur des experts ayant pour objectif de fournir une estimation correcte en utilisant la connaissance collective des membres de l'équipe.



## Lexiques de l'informatique

**802.11b** : Norme du réseau local d'entreprise (RLE) capable de transmettre jusqu'à 11Mo/s dans la bande non réglementée de 2,4 GHz et qui est efficace sur une distance variant entre 30 à 5 m.

**Abus informatique** : Perpétration au moyen d'un ordinateur d'un acte qui, sans être illégal, est contraire à l'éthique.

**Accès à distance** : Accès à un ordinateur et à ses périphériques à partir d'un terminal et placé à distance via un câble, un réseau ou une ligne téléphonique.

**Accès en mode forage** : Voir **Datamining**

**Actifs complémentaires (Complementary assets)** : Actifs additionnels nécessaires pour tirer parti d'un investissement primaire comme la formation dans un projet de nouveau système d'information.

**Adaptateurs de réseau sans fils** : Carte d'extension (carte réseau) munie d'une radio et d'une antenne pour la transmission sans fil.

**ADSL (Asymmetric Digital Subscriber Line)** : en français, réseau de raccordement numérique asymétrique. C'est une technologie de communication permettant de faire passer de l'information (issue d'Internet par exemple) à hautdébit sur les fils de cuivre utilisés habituellement pour les lignes téléphoniques (RTC). Avec l'ADSL, il est possible d'obtenir des débits plus de 20 fois supérieurs à ceux des modems classiques.

**Administration de la base de données** : Aspects plus techniques et opérationnels de la gestion des données, notamment le modèle physique et la maintenance technique de la base de données.

**Administration des données** : Secteur fonctionnel particulier de l'organisation chargé de la gestion des ressources en données de l'organisation. Il est responsable de la politique de l'information, de la planification des données, de la maintenance du dictionnaire des données et du respect des normes de qualités.

**Adresse IP** : Adresse représentée par quatre nombres entiers qui indique l'emplacement unique d'un ordinateur dans Internet.

**Adresse URL** : Adresse d'une ressource spécifique dans Internet.

**Affaire électronique** : Voir E-business

**Agent intelligent** : Programme logiciel qui utilise une base de connaissances intégrées ou acquises pour exécuter des tâches précises, répétitives et prévisibles pour un utilisateur, un processus d'affaires ou une application logicielle.

**AGL (Atelier de génie logiciel assisté par ordinateur)** : Automatisation des étapes dictée par les méthodes de développement de logiciels et de systèmes en vue de réduire la quantité de travail répétitif que doit accomplir le développeur.

**Alignement stratégique** : Démarche visant à faciliter la convergence et la synchronisation des dispositifs internes dans les entreprises sur le rythme et l'orientation des décisions stratégiques prises par les dirigeants.

**ALT et ALT GR (Anglais Alternate)** : Touches du clavier permettant d'effectuer des fonctions autres que celles initialement prévues pour chacune des touches du clavier.

**Analyse de portefeuille** : Evaluation du portefeuille d'investissements possibles d'une entreprise pour déterminer les risques et les bénéfices, et pour choisir parmi différentes solutions de systèmes d'information.

**Analyse** : Spécialiste qui traduit, dans un cahier des charges fonctionnel, les problèmes et les exigences de l'entreprise en besoins d'information et de systèmes. Il joue le rôle d'agent de liaison entre la direction des systèmes d'information et le reste de l'organisation.

**Analogique** : Qui représente par la variation continue d'une valeur physique. S'oppose à numérique.

**Analyse heuristique** : Méthode d'analyse utilisée par les antivirus pour détecter d'éventuels virus non répertoriés qui ne figurent pas dans la base de données du logiciel. Notez que les méthodes heuristiques ne donnent pas de résultats assurés et peuvent parfois générer des fausses alertes.

**Android** : Système d'exploitation pour tablettes, smartphones et Ultrabook issu de l'univers Linux (base Unix).

**Antivirus** : Logiciel spécialisé dans la lutte contre les virus, outre la vérification et le nettoyage des fichiers enregistrés sur le disque dur ; il peut exercer une surveillance permanente de la mémoire de l'ordinateur, indispensable si vous utilisez Internet.

**Aperçu** : Ouverture d'une petite fenêtre qui permet de voir le fichier sélectionné, avant de l'imprimer par exemple.

**Appel d'offres** : Liste détaillée de questions soumises aux prestataires informatiques (vendeurs de logiciels ou d'autres services) en vue de déterminer dans quelle mesure les prestations produites du vendeur peuvent répondre aux besoins spécifiques de l'organisation.

**Apprentissage organisationnel** : Création de nouvelles procédures d'opérations standard et de processus d'affaires qui reflètent l'expérience de l'organisation.

**Approche progressive ou modulaire** : Introduction de nouveau système par étapes plutôt que par fonctions ou par unités organisationnelles. Elle s'oppose à l'approche du basculement en une seule fois, également nommée « approche big bang ».

**Approvisionnement (Procurement)** : Localisation des sources d'approvisionnement pour les biens et le matériel, paiement de ces biens et établissement des accords de livraison.

**Architectures client-serveur multiniveau (N-tiers)** : Réseau où la charge de travail est équilibrée entre différents niveaux de serveurs et les services et applications exécutables sur les postes des utilisateurs nommés les « clients ».

**Architecture de réseau** : ensemble de liaisons permettant à différents ordinateurs de s'interconnecter et de partager ainsi des données et des services.

**Archive** : Fichier contenant lui-même plusieurs fichiers compressés selon un format particulier. Il faut décompresser l'archive avec un utilitaire pour récupérer les fichiers originaux.

**Argent électronique (Digital cash)** : Argent sous forme électronique qui circule à l'extérieur du circuit monétaire traditionnel.

**Ascenseur** : outil en bordure verticale droite ou (et) horizontale inférieure d'une page ou d'un cadre qui permet de faire défiler horizontalement ou verticalement une page à l'écran. On parle également de barre de défilement.

**ASP (Application Service Provider)** : Entreprise qui offre en location à d'autres entreprises l'accès à l'utilisation de logiciels via le Web ou un réseau privé.

**Assistant numérique personnel** : Petit ordinateur de poche muni d'un stylet et qui possède des fonctions intégrées de télécommunications sans fil capable d'effectuer des transmissions entièrement numériques.

**Asymétrie d'information (Information asymmetry)** : Situation dans laquelle le pouvoir de négociation relatif de deux parties est déterminé par l'une des parties qui possède davantage d'information essentielle à la transaction que l'autre partie.

**ATAPI (Ang. Attachment Packet Interface)** : Protocole avec lequel les lecteurs de CDROM communiquent avec l'ordinateur grâce à l'interface IDE.

**Attaque par déni de service** : Attaque qui inonde un serveur de réseau ou un serveur Web fausses demandes et requêtes de services afin de provoquer une panne réseau.

**Attribut** : Information qui décrit une entité particulière, comme le nombre de caractères d'un champ dans un fichier.

**Attribut de fichier** : Paramètre contenu dans les propriétés du fichier.

**ATX** : Format de carte mère spécifiant l'alimentation électrique, la position des connecteurs et leur format. AT était le format de carte mère précédant.

**Audit de qualité des données** : Examen des fichiers qui sert à déterminer l'exactitude et l'intégralité des données dans un système d'information.

**Authentification** : Procédure de vérification de l'identité d'un utilisateur sur un réseau. Sur Internet, les liaisons sécurisées ne peuvent être établies que par des sites authentifiés par un certificat électronique.

**Authentification biométrique** : Technologie utilisée comme système d'authentification des utilisateurs qui compare les caractéristiques uniques d'une personne comme ses empreintes, son visage ou son image rétinienne, avec un profil stocké en mémoire qui présente l'ensemble de ces caractéristiques.

**Automatisation** : Utilisation de l'ordinateur pour accéder l'exécution des tâches courantes.

**AutoTexte** : Fonction permettant de rédiger des lettres ou d'autres documents à partir de blocs de texte à assembler en saisissant simplement les abréviations qui les désignent.

**Avatar** : Personnage virtuel utilisé dans les espaces collaboratifs ou sur des sites Internet pour représenter un utilisateur ou un agent de service tel qu'un conseiller virtuel, par exemple.

**Backbone (Ang.)** :Dorsale.

Partie principale d'un réseau de télécommunication ou de téléinformatique, caractérisée par un débit élevé, qui concentre et transporte les flux de données entre des réseaux affluents.

**Background color (Ang.)**

Couleur d'arrière-plan.

Il s'agit d'une couleur que vous affecterez comme fond de page sur un site web par exemple.

**Backup (Ang.)**

Sauvegarde.

Souvent utilisé en informatique pour garder une copie de vos fichiers à un autre emplacement ou sur un autre support.

**Balise**

Terme désignant un format préparé et appliqué à un un paragraphe. Les balises sont les commandes de base du langage HTML.

**Bandeau publicitaire** : Affichage graphique sur une page web, qui sert à faire de la publicité. Le bandeau est un lien avec le site Web de l'annonceur de manière à ce que lorsqu'une personne clique sur le bandeau, une page Web contenant davantage d'informations sur l'annonceur apparaisse.

**Barre de défilement** : Voir ascenseur.

**Barre de formule** : Située en haut d'une fenêtre, elle affiche une valeur constante ou une formule utilisée dans la cellule active. Vous pouvez y taper directement une formule de calcul en commençant par le signe =.

**Barre de menu** : Barre qui se situe en haut ou en bas de l'écran et qui contient des menus permettent d'accéder aux commandes les plus courantes.

**Barre des tâches** : En bas de l'écran de Windows, la barre des tâches affiche dans sa partie droite, à côté de l'horloge, les utilitaires chargés en mémoire. De petites icônes signalent ainsi l'activité des logiciels de surveillance, comme les antivirus ou les pare-feu Internet.

**Barres d'outils** : Elles contiennent les commandes des logiciels sous forme de boutons. Vous pouvez les personnaliser en ajoutant ou supprimant des boutons.

**Basculement direct** : Méthode de conversion risquée dans laquelle le nouveau système remplace complètement l'ancien au cours d'une journée.

**Base de données** : Ensemble de données organisées de façon à servir les programmes applicatifs de plusieurs applications.

*Pour effectuer, la récupération des données dans la base de données. Nous devons utiliser un langage comme SQL.*

**Base de données hypermédias** : Base de données qui permet d'organiser les données sous forme de réseau de nœuds reliés dans une structure établie par l'utilisateur. Les nœuds peuvent contenir du texte, des graphiques, des sons, des vidéos ou des programmes exécutables.

**Base de données orientés objet** : Méthode de gestion des données qui permet de stocker les données et les procédures qui les régissent sous forme d'objets automatiquement récupérables, et qu'on peut mettre en commun. Les objets peuvent contenir des composants multimédias.

**Bases de données relationnelles** : est une base de données où l'information est organisée dans des tableaux à deux dimensions appelés des *relations* ou *tables*, selon le modèle introduit par Edgar F. Codd en 1970. Selon ce modèle *relationnel*, une base de données consiste en une ou plusieurs relations. Les lignes de ces relations sont appelées des *nuplets* ou enregistrements. Les colonnes sont appelées des *attributs*.

**Base de données répartie** : Base de données stockée dans plusieurs emplacements physiques. Des parties ou des copies de la base de données sont stockées physiquement dans un emplacement, et d'autres ailleurs.

**Base de règles** : Regroupement de connaissances dans un système d'intelligence artificielle qui est représenté sous forme de règles du type « SI ... ALORS ».

**Base de registres** : Fichier essentiel pour Windows, dans lequel sont consignées des informations de configuration du système et des logiciels.

**Bénéfices intangibles** : Bénéfices difficilement mesurables tels que l'amélioration de la qualité du service à la clientèle.

**Bénéfices tangibles** : Bénéfices quantifiables et auxquels on peut attribuer une valeur monétaire. Citons, entre autres, une diminution des frais d'exploitation et une augmentation des flux de trésorerie.

**Big data** : Expression faisant référence aux méthodes, approche et outils visant à exploiter l'immense masse de données numériques non structurées et hétérogènes qui s'accumulent sur le Web, afin d'en effectuer des analyses utiles pour les entreprises.

**B.I.O.S. (Ang. Basic Input Output System)** : Programme de vérification et d'initialisation de l'ordinateur, exécuté dès la mise en route, avant.

**Bit (Ang. Binary DigiT)** : Unité binaire.

Information représentée par un symbole à deux valeurs généralement notées 0 et 1. Huit bits forment un Octet.

**Bitmap (Ang.)** : Caractéristique d'une image en mode point (pixels). S'oppose à vectoriel. Par exemple, les applications suivantes génèrent des fichiers bitmap : PhotoShop, SuperPaint, Painter,...

**Blockchain** : Technologie de stockage et de transmission d'informations, transparente, sécurisée et fonctionnant sans organe central de contrôle. Par extension, une blockchain constitue une base de données qui contient l'historique de tous les échanges effectués entre ses utilisateurs depuis sa création. Cette base de données est sécurisée et distribuée : elle est partagée par ses différents

utilisateurs, sans intermédiaire, ce qui permet à chacun de vérifier la validité de la chaîne. Il existe des blockchains publiques, ouvertes à tous, et des blockchains privées, dont l'accès et l'utilisation sont limitées à un certain nombre d'acteurs.

**Blogue (Blog) :** De l'anglais *weblog*, désigne un site Web informel mais structuré, où des personnes peuvent publier des témoignages, des opinions et placer des hyperliens vers d'autres sites.

**Bluetooth :** Norme pour les réseaux personnels sans fil capables de transmettre jusqu'à 722 kilo-octets par seconde (Ko/s) à l'intérieur d'une zone de 10 m.

**Boîte aux lettres électronique ou B.A.L. :** Espace (virtuel) où vous recevez vos messages électroniques -en réalité un espace de stockage sur le serveur de votre fournisseur d'accès (voir aussi Compte de messagerie).

**Bogue :** Erreur ou code de programme défectueux.

**Bookmark (Ang.) :** Signet.

Moyen d'accéder rapidement à une adresse universelle préalablement stockée en mémoire par l'utilisateur.

**Boot (Ang.) :** Démarrage.

Séquence de démarrage de l'ordinateur. Par extension, « *booter* » signifie démarrer l'ordinateur.

**Bouton :** Représentation graphique qui fonctionne comme un bouton réel. On l'actionne en cliquant dessus.

**Browser (Ang.) :** Logiciel de navigation, navigateur.

Dans un environnement de type internet, logiciel qui permet à l'utilisateur de rechercher et de consulter des documents, et d'exploiter les liens hypertextuels qu'ils comportent.

**Bruit :** Ce n'est point du tintouin mais de petits points aléatoires, disposées à la surface d'une image, ressemblant à l'effet « neige » sur nos vieilles télévisions.

**Buffer (Ang.) :** Mémoire tampon.

**Bug (Ang.) :** Bogue

Erreur de programmation contenue dans une application (elle se dit alors buggée).

**Bureau :** Espace de travail délimité par la surface de l'écran et sur lequel on peut déposer toutes sortes de documents.

**Bus :** Terminologie informatique désignant un ensemble de fils électriques généralement présents sur la carte-mère qui assure la circulation des données entre les différents éléments du PC ( mémoire vive, carte graphique, USB, bus système, etc...). On caractérise un bus par sa fréquence (cadence de transmission des bits) et sa largeur (nombre de bits pouvant être transmis simultanément).

**BYOD (Bring Your Own Device) :** Mouvement apparu depuis la fin des années 2000 qui consiste à intégrer dans le paysage informatique des entreprises des matériels appartenant aux employés et issus du marché grand public tels que les iPhone, iPad et d'autres smartphones et tablettes.

**Byte (Ang.)** : Octet.

**Câble** : Liaison à Internet à haut débit, utilisant les réseaux câblés de télédiffusion. Comme avec l'ADSL une connexion par câble est généralement ouverte en permanence.

**Câble coaxial** : Support de transmission constitué de fils de cuivre fortement isolés et qui permet une transmission rapide de grandes quantités de données.

**Cache** : Zone mémoire ou disque servant à stocker des instructions ou des données temporaires dont l'accès est fréquent, ce qui accélère le transfert des données en diminuant le temps d'accès pour les trouver.

**Cadre (Picture-It)** : Une animation est composée de plusieurs cadres correspondant aux images composant l'animation.

**Cahier des charges** : décrit précisément les besoins auxquels le prestataire doit répondre, et organise la relation entre les différents acteurs tout au long d'un projet.

**Camembert** : Type de graphique proposé par (ou découlant) des tableurs. Le camembert se présente sous forme d'un cercle découpé en portions.

**Capacité** : Quantité d'informations maximum qu'un support numérique (Disque, CD, ...) peut stocker exprimé en multiples d'octets.

**Capteur CCD (Ang. Charged Coupled Device)** : Dispositif électronique qui transforme une image optique en information numérique acceptée par l'ordinateur.

**Capture d'écran** : Ou copie instantanée de l'écran à un moment donné.

**Carte graphique** : Carte électronique spécialisée dans le traitement et l'affichage des images numériques. Elle se connecte sur la carte mère par un port PCI ou AGP et est reliée à l'écran de l'ordinateur par un câble vidéo (VGA ou DVI). Elle comprend en général un chipset spécifique qui améliore la vitesse d'affichage et le réalisme de la 3D.

**Carte mère** : Carte électronique servant de support aux différents composants de l'unité centrale (microprocesseur et RAM).

**Cartes à puce (Smart card)** : Carte de plastique qui a le format d'une carte de crédit sur laquelle de l'information numérique est enregistrée.

**Casse** : En typographie, indique la grosseur des caractères utilisés. Par extension, on parle de casse pour désigner les minuscules et majuscules.

**Cellule** : Dans un tableur, l'intersection d'une ligne et d'une colonne forme une cellule. Son nom est composé de la référence de la ligne ajoutée à celle de la colonne (Exemple: la cellule A1 se trouve à l'intersection de la ligne 1 et de la colonne A).

**Centre d'appels (Call center)** : Service responsable du traitement des questions relatives au service à la clientèle qui utilise le téléphone ou d'autres moyens comme la messagerie électronique.

**CERN (Centre Européen pour la Recherche Nucléaire)** : C'est au sein du CERN qu'ont été pensés et conçus en 1989-1990 les principes et les premiers outils qui ont permis la création du Web (World Wide Web) par deux chercheurs, Tim Bernes-Lee et Robertt Cailliau.

**Certification** : reconnaissance d'un savoir-faire technique sur un produit donné, de la marque d'un éditeur de logiciels (Microsoft, SAP...) ou d'un fabricant de matériel informatique (IBM, HP, Cisco...).

**Certificat électronique** : Code d'identification d'un site Web ou d'un utilisateur de messagerie, délivré par un organisme de certification et destiné à garantir l'identité du titulaire, indispensable pour les connexions Internet sécurisée ou l'échange d'e-mails cryptés.

**Chaînage arrière** : Méthode de recherche dans la base de règles d'un système expert qui agit comme un résolveur de problèmes en commençant par une hypothèse et en recherchant plus d'information jusqu'à ce que l'hypothèse soit confirmée ou réfutée.

**Chaînage avant** : Méthode de recherche dans la base de règles d'un système expert qui commence par l'information entrée par l'utilisateur pour aboutir à une conclusion.

**Chaîne de valeur Web (Value Web)** : Réseau d'entreprises indépendantes axé sur la clientèle et qui utilise la technologie de l'information pour coordonner leurs chaînes de valeur afin de produire collectivement un produit ou un service destiné à un marché donné.

**Chaîne logistique** : Voir Supply Chain

**Champ clé** : Champ d'un enregistrement qui définit de manière unique selon des règles de codification bijective, ce qui permet de le récupérer, de le mettre à jour ou de la trier sans omission ni confusion.

**Champ** : Information élémentaire dans un fichier ou une base de données et caractérisée par des attributs (nombre de caractères, nombre de décimales pour un nombre, etc.).

Ce peut être le nom ou l'année de naissance d'une personne, ou encore le prix hors taxes d'un article.

**Changement de paradigme** : Méthode qui consiste à repenser de façon radicale la nature des affaires et de l'organisation.

**Chemin d'accès** : Description de l'emplacement d'un document sur le disque dur. Il se compose du nom du lecteur, de ceux du dossier et du sous-dossier, et enfin de celui du fichier.

**Cheval de Troie** : Programme inoffensif en apparence mais qui contient une fonction cachée susceptible de causer des dommages.

**Chiffrement** : Opération qui consiste à transformer un message à transmettre, dit « message clair », en un autre message, inintelligible pour un tiers, dit « messages chiffré », en vue d'assurer le secret de sa transmission.

**Clé** : Code numérique utilisé pour le cryptage de données, et indispensable pour leur décryptage. L'usage simultané de plusieurs clés permet de renforcer le niveau de sécurité.

**Click-and-mortar** : Modèle d'entreprise dont le site Web est l'extension d'une entreprise traditionnelle qui dispose d'installations matérielles.

**Client** : Ordinateur ou logiciel qui reçoit un service de la part d'un autre ordinateur ou logiciel, appelé serveur (voir ce terme). Lorsque vous vous connectez à Internet, votre ordinateur joue le rôle de client.

**Client-serveur** : mode de fonctionnement de plusieurs logiciels entre eux sur des ordinateurs différents. Un environnement client serveur est un réseau d'ordinateurs connectés à un serveur et qui s'échangent des logiciels ou des données.

**Cliquer-glisser ou glisser-déposer** : Action qui consiste à cliquer sur un élément avec un bouton gauche de la souris, puis à déplacer la souris en maintenant la pression sur le bouton. Cette action permet par exemple de sélectionner un mot dans une phrase ou de déplacer un objet.

**Cloud Computing** : Infrastructure informatique issue de la mutualisation dynamique et de la virtualisation de serveurs, permettant de proposer des services et des prestations à des clients internes ou externes, facturés (ou décomptés) à la consommation (voir IaaS, PaaS, SaaS). Également appelé informatique en nuage.

**Coffre-fort numérique** : Espace de stockage situé au sein d'une infrastructure de type Cloud et protégé par encryptions et mots de passe.

**Cogniticien** : Spécialiste qui recueille des informations et l'expertise d'autres professionnels, et qui les traduit en un ensemble de règles ou de trames pour un système expert.

**Collecticiel** : Voir Groupware

**Comité d'entreprise** : Instance représentative des salariés dans une entreprise. Selon les secteurs et la taille des organisations, ces comités doivent être associés aux projets de modernisation des entreprises.

**Commerce collaboratif (Collaborative commerce)** : Utilisation des technologies numériques pour permettre à plusieurs organisations de concevoir, de développer, de produire et de gérer les produits en collaboration au cours des différents cycles de vie du produit.

**Commerce électronique (Electronic commerce ou e-commerce)** : Processus qui consiste à acheter et à vendre des produits et des services par voie électronique via Internet, des réseaux et d'autres technologies numériques.

**Commerce électronique de détail (Business-to-Consumer electronic commerce, B to C)** : Vente par Internet de produits ou de services aux consommateurs eux-mêmes.

**Commerce électronique interconsommateurs (Consumer-to-Consumer electronic commerce, C to C)** : Vente par Internet de produits ou de services entre consommateurs.

**Commerce électronique interentreprise (Business-to-Business electronic commerce, B to B)** : Vente par voie électronique (EDI, Internet, etc.) de produits ou de services entre entreprises.

**Commerce électronique mobile ou commerce mobile (Mobile commerce ou m-commerce)** : Utilisation d'appareils sans fil, comme les Smartphones, les téléphones mobiles et autres appareils mobiles de poche, pour effectuer des transactions de commerce électronique de détail et interentreprises sur Internet.

**Commutateur** : Appareil plus sophistiqué qu'un concentrateur qui relie les composants et le réseau, filtre les données et les achemine vers une destination spécifique.

**Commutateur de paquets** : Technologie qui permet de découper des messages numériques en morceaux appelés « paquets » et de les acheminer de la manière la plus économique vers les canaux de communication disponibles.

**Compression** : Technique permettant de réduire la taille des fichiers afin de minimiser la place occupée sur le disque dur ou les temps de téléchargement sur Internet. Il existe plusieurs formats de fichiers compressés.

**Comptabilité par activités (Activity Based Costing, ABC)** : Modèle destiné à identifier toutes les activités de l'entreprise qui entraînent des coûts lors de la production d'un produit ou d'un service précis afin de permettre aux gestionnaires de déterminer les produits et les services qui sont rentables et ceux qui ne le sont pas, puis d'effectuer des changements pour accroître la rentabilité de l'entreprise.

**Concentrateur** : Appareil très simple qui relie les composants d'un réseau, transmettant des paquets de données aux autres appareils.

**Conception assistée par ordinateur (CAO ; Computer-Aided Design, CAD)** : Système d'information qui aide à la conception initiale puis aux reconceptions des produits au moyen d'un logiciels techniques et graphiques perfectionnés.

**Conception d'application en collaboration (Join Application Development, JAD)** : Processus destiné à accélérer la satisfaction des besoins en information en faisant travailler les utilisateurs finaux et les spécialistes ensemble dans des sessions intensives de conception itératives.

**Conception de systèmes** : Description détaillée de la manière dont le système répondra aux besoins en information déterminés par l'analyse des systèmes.

**Conception sociotechnique** : Conception qui vise à produire des systèmes d'information qui combinent l'efficacité technique au souci de répondre aux besoins organisationnels et humains.

**Configurer** : Régler un ensemble de paramètres en fonction de l'environnement utilisé. Opération en général préliminaire à la mise en service.

**Connaissance** : Concepts, expérience et aperçus qui procurent une structure pour créer évaluer et utiliser de l'information.

**Connaissance explicite** : Connaissance qui a été documentée.

**Connaissance semi-structurée** : Information sous forme d'objets moins structurés comme les courriers électroniques, les sessions de dialogue sur Internet, les vidéos, les graphiques, les brochures ou les panneaux d'affichage.

**Connaissance structurée** : Connaissance sous la forme de documents et de rapports structurés.

**Connaissance tacite** : Connaissance et expertise des membres de l'organisation qui n'ont pas encore été documentées en bonne et due forme.

**Connecteur** : Emplacement ou prise de la carte mère de l'ordinateur, destiné à brancher les cartes ou les câbles des différents composants ou périphériques.

**Connectivité** : Capacité pour les ordinateurs et les différents équipements informatisés de communiquer entre eux et de partager de l'information le plus automatiquement possible.

**Contrôle** : Méthodes, politiques et procédures qui garantissent la protection des éléments d'actif d'une organisation, la précision et la fiabilité de ses enregistrements et la conformité de ses normes de gestion.

**Contrôles d'accès** : Politiques et procédures utilisées par une entreprise pour empêcher les personnes (internes ou externes à l'organisation) non autorisées d'accéder à ses systèmes.

**Contrôles à l'entrée** : Procédures qui permettent de vérifier la précision et l'intégralité des données au moment où on les entre dans un système.

**Contrôles à la sortie** : Mesures qui garantissent que les résultats du traitement informatique sont exacts, complets et adéquatement mis à disposition.

**Contrôles administratifs** : Normes, règles et procédures qui assurent l'exécution appropriée et la mise en vigueur des contrôles de l'organisation.

**Contrôles des applications** : Contrôles uniques, qui portent sur chacune des applications informatisées.

**Contrôles généraux** : Contrôles globaux qui établissent un cadre de vérification de la conception, de la sécurité et de l'usage des logiciels dans une organisation.

**Contrôles sur le traitement** : Sous-programmes qui vérifient l'intégralité et l'exactitude des données durant la mise à jour.

**Conversion** : Processus de passage de l'ancien système au nouveau.

**Cookie** : Petit fichier déposé dans le disque dur d'un ordinateur lorsqu'un utilisateur visite certains sites Web. Il est ensuite utilisé pour identifier le visiteur et pour faire le suivi de ses visites sur le site Web.

**Courrier électronique** : Echange de messages entre usagers qui déposent d'un accès informatique.

**Coût total de possession (Total Cost of Ownership, TCO)** : Indique le coût total lié à la possession de ressources informatiques, incluant les coûts initiaux d'acquisition, les coûts de mises à jour matérielles et logicielles, l'entretien, l'assistance technique et la formation.

**Corbeille** : Espace de stockage temporaire dans lequel sont placés les fichiers supprimés. Il faut penser à la vider si vous ne voulez pas qu'on puisse accéder aux documents que vous avez effacés.

**Correctifs** : Fichiers destinés à corriger les défauts ou améliorer les fonctionnalités d'un logiciel. Ils sont en général diffusés par CD ou sur Internet.

**Courrier électronique** : Echange de messages et de fichiers via Internet. Le courrier électronique est un vecteur privilégié pour les virus.

**C.P.U. (Ang. Central Processing Unit) :** Microprocesseur.

**Cracker (Ang.) :** Pirate.

**CRM (Customer Relationship Management) :** le CRM s'appuie sur une démarche et ensemble d'applications informatiques permettant d'analyser les comportements des consommateurs afin d'optimiser les relations entre l'entreprise et ses clients.

**Cryptage :** Procédé consistant à rendre les données totalement incompréhensibles à l'aide d'un algorithme de cryptage et d'une clé numérique (un code). La clé est indispensable pour le décryptage

**CTRL (Ang. Control) :** Touche du clavier permettant la sélection multiple d'objets par glisser-déplacer

**Cybernaut (Ang.) :** Cybernaute ou internaute.

**Cycle d'horloge :** Unité élémentaire de temps d'un ordinateur, définie par sa fréquence d'horloge. Chaque action d'un microprocesseur requiert au moins un cycle pour s'exécuter

**Cybergouvernement(E-gouvernement) :** Utilisation d'Internet et des technologies connexes afin de permettre aux gouvernements et aux organisations du secteur public d'établir des relations dématérialisées avec les citoyens, les entreprises et d'autres organismes ou collectivités publiques.

**Cybervandalisme :** Perturbation, dégradation ou même destruction préméditée d'un site Web ou d'un système informatique.

**Cybercriminalité :** Désigne l'ensemble des infractions pénales commises via les réseaux informatiques, notamment sur le réseau Internet. Couvre l'atteinte aux biens (fraude à la carte bancaire, encaissement d'un paiement sans livraison, etc. ) et aux personnes (diffusion d'images pédophiles, de recettes d'explosifs ; atteinte à la vie privée, etc...).

**Data Analyst :** Personne dont le métier consiste à utiliser des techniques statistiques et des outils informatiques spécialisées afin d'organiser, de synthétiser et de traduire les informations dont les entreprises ont besoin pour faciliter les prises de décision.

**Data Scientist :** Personne dont les missions sont celles d'un Data Analyst et qui consistent aussi à déterminer sous de nombreux angles ce que les données signifient, puis recommander des moyens pour appliquer les enseignements de ces interprétations.

**Datamart (magasin de données) :** version restreinte d'un datawarehouse, concernant une partie de l'entreprise (finance, marketing, ventes, ressources humaines...).

**Datamining :** processus de gestion et d'extraction des données (du datawarehouse) par des outils statistiques, pour en tirer des corrélations et des tendances pertinentes.

**Datawarehouse (entrepôt de données) :** base de données servant de support à des applications d'aide à la décision et de gestion de la relation client.

**Débit :** Exprimé en bits par seconde, il quantifie la vitesse d'émission (on parle de débit montant ou upload) et de réception (on parle de débit descendant ou download) d'un transfert de données.

**Débogage :** Processus de localisation et d'élimination des erreurs et des défauts (nommés « bogues ») dans un programme.

**Décision non structurée** : Décision inhabituelle qui exige que les managers fassent preuve de jugement et d'intuition en ce qui concerne la définition des problèmes. Aucune procédure n'est préétablie pour cette forme de prise de décisions.

**Décision semi-structurée** : Décision où seule une partie du problème trouve une réponse claire fournie par une procédure préétablie.

**Décision structurée** : Décision répétitive, routinière et qui comporte une procédure préétablie de traitement.

**Définition** : La définition correspond au nombre de points d'affichage utilisés pour composer l'image à l'écran par unité de surface. Plus le nombre de points est important, meilleure sera la définition de l'image. Elle est exprimée en nombre de pixels horizontaux et verticaux. Les définitions courantes utilisées par un écran standard sont 640 x 480, 800 x 600 et 1024 x 768.

**Dématérialisation** : la dématérialisation des données consiste à stocker et à faire circuler des informations sans support matériel autre que des équipements informatiques.

**Déni de service** : Attaque très évoluée visant à rendre indisponible une machine en la submergeant de trafic inutile. Lorsque plusieurs machines sont à l'origine d'une telle attaque distribuée. Elle peut anéantir des serveurs, des sous-réseaux, etc. Elle reste très difficile à contrer ou à éviter.

**Désintermédiation** : Elimination d'étapes intermédiaires des processus d'affaires ou de gestion, ce qui permet de réduire le prix de vente d'un produit.

**Développement des systèmes** : Activités qui contribuent à la recherche et à la mise en œuvre d'une solution informatisée pour répondre à un problème auquel l'organisation se heurte ou afin de saisir une opportunité qui se présente à elle.

**Développement orienté objet** : Méthode de développement des systèmes qui utilise l'objet comme unité de base de l'analyse et la conception des systèmes. Le système est modélisé comme un ensemble d'objets et de relations entre eux.

Je site le langage C#, JAVA, PHP et d'autres. Ce sont des langages dits de haut niveau.

**Développement par composants ou briques logicielles** : Le développement par composant ou briques logicielles combine des briques ou composants de logiciels préexistants pour créer des applicatifs fonctionnels complets.

**Développement par l'utilisateur final** : Construction de systèmes d'information par les utilisateurs finaux avec peu d'aide, voire aucune, de la part des informaticiens.

**Développement rapide d'applications (Rapid Application Development, RAD)** : Processus de développement de systèmes en très peu de temps, au moyen du prototypage, des outils de quatrième génération et d'un travail en collaboration entre les utilisateurs et les spécialistes des systèmes.

**Device (Ang.)** : Périphérique.

**DHTML** : Regroupement des améliorations dynamiques apportées par les navigateurs. Le DHTML comprend donc le HTML 4.0, les CSS et le DOM.

**Diagramme de flux de données** : Outil principal de l'analyse structurée qui illustre graphiquement les processus qui sous-tendent le système et flux de données.

**Diagramme entité-relation** : Méthode utilisée pour documenter les bases de données, qui illustre la relation entre les diverses entités qu'elles contiennent.

**Dictionnaire de données** : Outil automatisé ou manuel qui sert à identifier les éléments d'identification technique et logique liés au stockage et à l'organisation des définitions des éléments de données maintenus dans une base de données.

**Directeur des systèmes d'information** : Voir DSI

**Directory (Ang.)** : Répertoire.

**Disque dur** : Composant de l'ordinateur sur lequel sont enregistrés les fichiers du système d'exploitation, les logiciels et toutes les données. Les disques durs sont en général très fiables, mais la moindre panne peut avoir de grandes conséquences.

**Disquette de démarrage** : Disquette comportant un noyau de système d'exploitation, indispensable pour redémarrer l'ordinateur en cas d'impossibilité de démarrer sur le disque dur à la suite d'une panne ou d'une infection par un virus.

**Documentation** : Description du fonctionnement d'un système d'information du point de vue technique et/ou du point de vue de l'utilisateur final.

**Données** : Sources de faits bruts qui représentent des événements qui ont lieu dans les organisations ou dans leur environnement. Ces faits n'ont pas encore été organisés et disposés de façon à ce que les gens puissent les comprendre et les utiliser.

**Dorsale de réseau** : Partie d'un réseau qui traite le trafic important et qui constitue le circuit principal assurant la transmission de données vers d'autres réseaux ou en provenance de ceux-ci.

**D.O.S. (Ang. Disk Operating System)** : C'est le premier système d'exploitation (voir ce terme) pour PC à avoir connu une très vaste diffusion, dans les années 80. Décliné en de multiples versions, il était édité par Microsoft, comme Windows aujourd'hui. Peu encombrant, il est encore utilisé pour les disquettes de démarrage.

**Dossier** : Espace de stockage sur le disque dur, destiné à classer les fichiers. Un dossier peut contenir des sous-dossiers.

**Dossier compressé** : Voir compression de données.

**Dossier de démarrage** : Les raccourcis placés dans le dossier Démarrage du menu Programmes (via le bouton Démarrer correspondent à des programmes lancés automatiquement au démarrage de Windows.

**Download (Ang.)** : Téléchargement

**D.N.S. (Ang. Domain Name System)** : Système d'adressage par domaines ou annuaire des domaines.

Système de bases de données et de serveurs assurant la correspondance entre les noms de domaine ou de site utilisés par les internautes et les adresses numériques utilisables par les ordinateurs.

**Domain (Ang.)** : Domaine.

Ensemble d'adresses faisant l'objet d'une gestion commune.

**D.P.I. (Ang. Dot Per Inch)** : Point par pouce.

**Drag and drop (Ang.)** : Glisser-déposer ou cliquer-glisser.

Action par laquelle l'utilisateur sélectionne un objet à l'écran, le déplace jusqu'à une autre position, puis le lâche pour déclencher une action sur cet objet

**Driver (Ang.)** : Pilote.

**DSI (Directeur des systèmes d'information ; Chief Information Officer, CIO)** : Cadre supérieur responsable de la fonction des systèmes d'information dans l'entreprise.

**D.V.D (Ang. Digital Versatile Disk)** : Disque numérique polyvalent.

Disque numérique optique de grande capacité, à usages divers (audio, vidéo, multimédia, ...).

**D.V.I. (Ang. Digital Visual Interface)** : Norme de transmission numérique des signaux graphiques informatiques (DVI-D). Existe aussi sous une forme permettant la transmission des signaux en numérique et en analogique (DVI-I).

**EAI (Entreprise Application Integration)** : Logiciel d'intégration d'application différentes dans une même entreprise.

**E-business** : Il s'agit du commerce électronique sous Internet. Le e-business est un mode de fonctionnement qui consiste à transformer les principaux processus commerciaux d'une entreprise en y intégrant les technologies Internet et leurs apports en matière de relation avec la clientèle.

**Echange de données informatisé (EDI ; Electronic Data Interchange, EDI)** : Echange direct de transactions standardisées entre les ordinateurs d'au moins deux organisations, par exemple des commandes, des instructions relatives à l'expédition ou des paiements.

**Economie de réseaux (Network Economics)** : Modèle fondé sur la notion de réseau où l'ajout d'un autre membre comporte un coût marginal nul, mais peut créer des gains marginaux beaucoup plus grands. Il sert de modèle aux systèmes stratégiques sur le plan de l'industrie.

**Ecosystème d'affaires (Business ecosystem)** : Réseaux relativement dispersés, mais interdépendants, de fournisseurs, de distributeurs, d'entreprises délocalisées, d'entreprise de transport et de fabricants de technologies.

**Ecran tactile** : Périphérique qui rassemble les fonctionnalités d'affichage d'un écran (moniteur) et celle d'un dispositif de pointage, comme une souris ou un trackball.

**Ecriture miroir** : Serveur auxiliaire qui reproduit tous les processus et toutes les transactions d'un serveur primaire. En cas de panne de celui-ci, les « données miroir » peuvent être exploitées par un serveur de secours pour prévenir toute interruption de service.

**E.I.D.E. (Ang. Enhanced-Integrated Drive Electronics) :** Nouvelle version de l'interface IDE (interface de disque dur). EIDE permet des disques plus gros et plus rapides; la plupart des systèmes vendus de nos jours intègrent l'EIDE.

**EIS (Executive Information System) :** voir SAD

**Élément de données :** Voir Champ

**Emoticon (Ang.) :** Frimousse.

**Enregistrer :** Ecrire des données dans un fichier sur le disque dur ou une disquette, de manière à les conserver durablement.

**Enregistrement :** Groupe de champs connexes.

**Entité :** Personne, lieu, objet ou événement sur lequel on conserve de l'information.

**Entrée :** Saisie ou collecte des données brutes qui proviennent de l'organisation ou de son environnement extérieur et qui seront traitées par un système d'information.

**Entrepôt de données :** Voir Datawarehouse.

**Entreprise numérique (Digital firm) :** Organisation dans laquelle la plupart des processus d'affaires importants et des relations avec les clients, les fournisseurs et les employés reposent sur l'informatique et dont les principaux actifs de la société sont gérés à l'aide de moyens numériques.

**Équilibrage de charge :** Distribution d'un volume important de requêtes entre de multiples serveurs afin d'éviter de surcharge, voire de bloquer l'un d'entre eux.

**Ergonomie :** Interaction des utilisateurs avec les machines dans leur milieu de travail. Ce terme porte sur la description des tâches, les questions de santé et l'interface utilisateur des systèmes d'information.

**ERP (Enterprise Resource Planning) :** les ERP, aussi appelés Progiciels de Gestion Intégrés (PGI), sont des applications dont le but est de coordonner l'ensemble des activités d'une entreprise ( production, approvisionnement, marketing, forces de vente, gestion des ressources humaines, etc.) autour d'un même système d'information.

**Étalonnage :** Etablissement de normes strictes pour produits, les services ou les activités et évaluation de la performance de l'entreprise par rapport à ces normes.

**Étude de faisabilité :** Dans le cadre d'un projet de transformation ou de création d'un système d'information, manière de déterminer si l'on peut mettre en œuvre la solution proposée, compte tenu des ressources et des contraintes de l'organisation et de ses systèmes en place.

**Étude pilote :** Stratégie destinée à introduire le nouveau système dans un secteur limité de l'organisation jusqu'à ce qu'il démontre qu'il est entièrement fonctionnel. A ce moment seulement, la conversion au nouveau système et son déploiement s'effectuent alors dans toute l'organisation.

**EV-DO (Evolution Data Optimized) :** Technologie utilisée pour les réseaux cellulaires afin d'offrir à partir de n'importe quel endroit et à tout moment un accès à large bande aux ordinateurs personnels et à d'autres appareils à une vitesse moyenne qui se situe entre 300 et 500 Kb/s.

**Exclusion (Opt-out)** : Modèle de consentement éclairé qui permet à l'entreprise de collecter des renseignements personnels, à moins que le consommateur ne demande explicitement que ces renseignements ne soient pas recueillis.

**Extension de fichier** : Les trois lettres qui suivent le nom d'un fichier, séparées de ce dernier par un point « . » forment l'extension. Elle indique le format du fichier.

**Externalisation** : le fait de confier la gestion de tout ou partie de son système d'information à un sous-traitant.

**Extranet** : Intranet étendu, accessible uniquement aux personnes externes dûment autorisées.

**FAT/FAT32 (File Allocation Table)** : Voir Système de fichiers

**F.A.Q. (Ang. Frequently Asked Questions)** : Foire aux questions, questions courantes.

**Fenêtre publicitaire (pop-up ads)** : Publicité qui apparaît automatiquement et qui reste à l'écran tant que l'utilisateur ne clique pas dessus.

**Feuille de style** : Fichier spécial attaché à un document contenant la description de la mise en forme des caractères et des paragraphes ainsi que la description de la mise en page.

**Fibre optique** : Moyen de transmission rapide, léger et durable, constitué de minces torons de fibre de verre transparent reliés en un câble. Les données sont transmises sous forme d'impulsions lumineuses.

**Fichier** : Groupe d'enregistrements du même type. Tout enregistrement sur un support de stockage se fait dans un fichier, identifié par un nom. On distingue deux sortes de fichiers. Les fichiers exécutables contiennent des instructions et des commandes destinées à être exécutées par l'ordinateur. Les fichiers de données correspondent à des textes, des images ou encore à des listes de paramètres.. Tous les fichiers que vous enregistrez sont du second type -à moins que vous ne programmez.

**Fichiers compressés** : fichiers qui subissent une Compression de données.

**Fichiers Internet temporaires** : Lorsque vous consultez des sites Internet, votre navigateur stocke les pages sur le disque dur afin de pouvoir les afficher plus rapidement si vous revenez en arrière. Il s'agit d'un stockage temporaire, les pages récentes venant remplacer les plus anciennes.

**Fichier système** : Fichiers appartenant au système d'exploitation. Il peut s'agir de fichiers programmes ou de bibliothèques de fonctions. Il y a aussi des fichiers de données, comme la base de registres, contenant les paramètres du système.

**Fil torsade** : Support de transmission constitué de paires de fils de cuivres torsadés. On l'utilise pour la transmission téléphonique de signes analogiques, mais aussi pour les transmissions de données.

**Filtrage par paquets** : Examen d'une sélection de fichiers dans les en-têtes des paquets de données qui circulent entre un réseau sécurisé et Internet.

**Filtre de courrier** : Fonction permettant de trier automatiquement les messages électroniques en fonction de l'expéditeur, de l'objet ou encore des termes contenus dans le texte.

**Firewall (Ang.)** : Barrière de sécurité, pare-feu.

**Flux RSS** : RSS (Rich Site Summary ou Really Simple Syndication) vise à syndiquer du contenu Web utilisable ensuite dans un autre cadre. La technologie RSS extrait du contenu spécifique sur un site Web et l'envoie automatiquement vers les ordinateurs des utilisateurs, où il peut être stocké pour être consulté plus tard.

**Fondeur** : Fabricant de microprocesseurs (Exemple: AMD, Intel, Motorola, ...).

**Forage de données** : Analyse d'une grande quantité de données pour dégager des tendances et des règles utilisables afin de prendre des décisions et de prédire des comportements futurs.

**Format de fichier** : Pour traduire un texte ou une image sous forme numérique et l'enregistrer dans un fichier, il faut employer un code, un format particulier. Plusieurs formats différents peuvent être utilisés pour un même type de données. Ainsi, il existe de très nombreux formats d'image. Le format d'un fichier est indiqué par son extension.

**Format paysage** : Présentation d'une page en 29,7x21 par opposition au format classique dit format portrait de 21x29,7 qui est celui utilisé par défaut.

**Fournisseur d'accès à Internet ou F.A.I.** : Société disposant l'une liaison spécialisée à très haut débit vers Internet, et qui offre à ses clients un accès au réseau via ses serveurs

**Fournisseur de service Internet (Internet Service Provider, ISP)** : Organisation commerciale connectée en permanence sur Internet qui vend des connexions temporaires à ses abonnés.

**Framework** : désigne un ensemble cohérent de composants logiciels structurels, qui sert à créer les fondations ainsi que les grandes lignes de tout ou d'une partie d'un logiciel (architecture).

**Freeware** : Logiciel totalement gratuit.

**Fréquence d'horloge** : Exprimée en Mégahertz (Mhz), elle est égale au nombre de cycles effectués par le processeur en une seconde.

**Frimousse** : Dans un message, association facétieuse de quelques caractères typographiques qui évoquent un visage expressif (Exemples: « :- ) » ou « :-( »). (aussi appelé émoticône) 😊

**F.T.P. (Ang. File Transfer Protocol)** : Un des protocoles de transfert de fichiers sur Internet, souvent mis en oeuvre lors des téléchargements.

**Gateway (Ang.)** : Passerelle.

**Génie logiciel** : Voir AGL.

**Gestion de cycle de vie du produit (GCVP ; Product Life cycle management, PLM)** : Systèmes fondés sur un référentiel d'entreprise qui organise chaque détail de l'information d'une entreprise rattachée à un projet spécifique.

**Gestion de la chaîne logistique** : Voir Supply Chain Management

**Gestion de la qualité totale (Total quality management)** : Principe selon lequel le contrôle de la qualité devient responsabilité de tous les membres de l'organisation.

**Gestion de la relation client** : voir **CRM**.

**Gestion des connaissances** : Voir **Knowledge Management**.

**Gestion des processus d'affaires (Business Process Management, BPM)** : Méthodologie conçue pour réviser les processus d'affaires d'une organisation afin qu'ils puissent servir de base essentielle aux systèmes d'information de l'entreprise.

**Gestion flux de travail** : voir **Workflow**.

**Gestion électronique documentaire (GED)** : Politiques, procédures et outils informatiques qui servent à gérer la conservation, la destruction et le stockage des documents électroniques.

**Gestion de liste LISTSERV** : Groupes de discussion en ligne qui utilisent le courrier électronique.

**Gestionnaire de réseaux** : Logiciel technique qui achemine et gère la communication dans un réseau et coordonne ses ressources.

**Gif animé** : Fichier composé de plusieurs images en 256 couleurs au format Gif, le tout formant une animation.

**Green Computing** : Ensemble de technologies et de processus visant à réduire l'empreinte écologique, économique, et sociale des technologies de l'information et de la communication.

**Groupware** : Logiciel qui offre des fonctions de soutien à la communication et au travail d'équipe.

**Hacker (Ang.)** : Fouineur, par extension pirate.

Personne qui contourne ou détruit les protections d'un logiciel, d'un ordinateur ou d'un réseau informatique.

**Hacking**: le terme de hacker est souvent utilisé pour désigner un pirate informatique. Les victimes de piratage sur des réseaux informatiques aiment à penser qu'ils ont été attaqués par des informaticiens chevronnés ayant soigneusement étudié leur système et ayant développé des outils spécifiques pour créer une faille dans leur système.

**Hameçonnage** : Voir **Phishing**.

**Hardware** : Equipement physique utilisé pour exécuter les activités d'entrée, de traitement et de sortie d'un système d'information.

**Historique** : Votre navigateur Internet conserve dans l'Historique les adresses des pages Web que vous avez visitées au cours des jours et des semaines passées afin que vous puissiez facilement les retrouver.

**Home page (Ang.)** : Page d'accueil.

**Hotlist (Ang.)** : Liste de signets.

**HTML (HyperText Markup Language)** : Langage de balisage utilisé pour créer des pages Web et d'autres documents hypermédias.

**H.T.T.P. (Ang. HyperText Transfer Protocol)** : Protocole utilisé pour le transfert des pages Web entre le serveur et les utilisateurs. Il remplace parfois également le protocole FTP pour les téléchargements de fichiers.

**Hypertext (Ang.)** : Hypertexte (n.m.), hypertextuel.(adj.)

**IaaS (Infrastructure as a Service)** : Premier niveau de prestation permis par les infrastructures de type Cloud : mise à disposition d'une infrastructure matérielle sur laquelle le client peut installer et/ou développer et/ou faire tourner des applications.

**Icône** : Symbole graphique représentant un logiciel ou un fichier. Un double-clic sur une icône ouvre le fichier ou lance le logiciel.

**I.D.E. (Ang. Integrated Drive Electronics)** : Interface permettant de connecter des disques durs et des CD-ROMs à des ordinateurs par l'intermédiaire de nappes.

**Identifiant** : Nom d'utilisateur nécessaire à l'ouverture d'une session.

**I.M.A.P (Ang. Internet Messaging Api Protocol)** : Protocole utilisé par les logiciels de messagerie électronique pour récupérer votre courrier sur le serveur de votre fournisseur.

**Implémentation** : Traduire un algorithme dans un langage de programmation.

**Inclusion (Opt-in)** : Modèle de consentement éclairé par lequel une entreprise se voit interdire la collecte de renseignements personnels à moins que le consommateur n'indique clairement qu'il approuve la collecte et l'utilisation des renseignements qui le concernent.

**Incohérence des données** : Présence de valeurs différentes pour le même attribut lorsqu'une même donnée est stockée à différents endroits.

**Index de rentabilité** : Méthode qui consiste à utiliser l'index pour comparer la rentabilité des diverses possibilités d'investissement. On calcule l'index en divisant la valeur actualisée des encaissements totaux d'un investissement par son coût initial.

**Indicateur** : Mesure standard de la performance.

**Infogérance** : le fait de confier tout ou partie de son informatique à un prestataire extérieur (SSII, constructeur...) qui, dans le cadre d'un contrat forfaitaire de plusieurs années, en prend la responsabilité opérationnelle et la fait évoluer. Le contrat peut ou non s'accompagner d'un transfert de salariés. L'infogérance est aussi appelée outsourcing ou facilities management.

**Information** : Données présentées sous forme significative et utile pour les individus.

**Informatique à haut disponibilité** : Outils et technologies, incluant les ressources auxiliaires, qui permettent à un système de récupérer rapidement sa capacité opérationnelle suite à une panne immobilisatrice.

**Informatique à la demande** : Utilisation d'un centre informatique externe pour bénéficier d'une puissance de traitement accrue en période de pointe, ce qui permet à une entreprise de limiter son investissement à une infrastructure capable de traiter des quantités moyennes de données. L'entreprise paie uniquement pour la puissance de traitement supplémentaire qu'elle utilise.

**Informatique autonome** : Effort qui vise à mettre au point des systèmes capables de s'autogérer sans intervention humaine.

**Informatique mobile** : Informatique sans fil qui permet aux téléphones mobiles, aux assistants personnels numériques et aux autres appareils sans fil d'accéder à de l'information numérique dans Internet ou à d'autres sources à partir de n'importe quel endroit.

**Informatique nomade** : Informatique sans fil où les utilisateurs vont d'un point d'accès sans fil à un autre pour accéder à un réseau ou à Internet.

**Infrastructure à clé publique** : Système qui sert à créer des clés publiques et privées à l'aide d'un certificat d'autorisation (CA) et d'un certificat numérique pour l'authentification.

**Infrastructure technologique (Information Technology (IT) infrastructure)** : Matériel (d'ordinateur), logiciels, technologies de stockage et de réseaux qui créent une plateforme commune des ressources en technologie de l'information.

**Inspection dynamique** : Détermine si des paquets de données font partie d'un dialogue continu entre un expéditeur et un destinataire, puis crée des tables d'état pour suivre l'information transmise par des paquets multiples.

**Intégration (d'un système)** : il s'agit de l'assemblage progressif des éléments d'un système, logiciels, matériels... en vue de constituer un système global.

**Intelligence artificielle (IA)** : Systèmes informatisés qui ont la capacité d'apprendre des langages, d'accomplir des tâches physiques, d'utiliser des dispositifs de perception et d'émuler certaines des formes d'expertise et de prise de décisions des individus.

**Interdépendance des données et des programmes** : Relation structurelle et technique entre les données stockées dans des fichiers et les programmes nécessaires pour les mettre à jour et les modifier. Toute modification dans l'organisation ou le format des données exige un changement des programmes qui permettent de traiter les fichiers qui les contiennent. Le recours à un SGBD diminue, voire supprime cette indépendance.

**Interface** : jonction entre deux matériels ou logiciels leur permettant d'échanger des informations par l'adoption de règles communes, physiques ou logiques.

**Interface utilisateur** : Partie d'un système d'information qui permet à l'utilisateur final d'interagir avec le système. Il s'agit d'éléments matériels, mais surtout de l'ensemble des commandes et des réponses à l'écran qui permettent à l'utilisateur de travailler avec un système.

**Internet** : Combinaison internationale de réseaux qui relie des centaines de milliers de réseaux informatiques privés et publics selon les protocoles et technologies standard.

**Interopérabilité** : faculté d'utiliser conjointement des fonctionnalités d'applications basées sur des technologies différentes (J2EE, .NET, PHP, C++, etc.).

**Intranet** : Réseau interne, fondé sur la technologie et les normes d'Internet et du Web.

**IPv6 (Internet Protocole version 6)** : Ce nouveau protocole propose des adresses 128 bits (2 puissances 128), soit plus d'un million de milliards d'adresses uniques.

**ISO 17799** : Ensemble de normes internationales de sécurité et de contrôle.

**Jeton d'authentification** : Dispositif conçu pour confirmer l'identité d'un seul utilisateur.

**Juste à temps (Just-In-Time)** : Système de gestion industrielle et logistique qui permet de réduire au maximum les stocks en recevant les composants juste au moment où on en a besoin et en expédiant les produits finis dès qu'ils quittent la chaîne de montage.

**Knowledge Management (KM)** : Ensemble de processus développés dans une organisation pour créer, réunir, stocker, entretenir et diffuser les connaissances de l'entreprise.

**Langage d'interrogation** : Outil logiciel qui procure des réponses en ligne aux demandes d'information qui ne sont pas prédéfinies.

**Langage de balisage extensible** : Voir **XML**

**Langage de balisage hypertexte** : Voir **HTML**

**Langage de définition de données** : Partie d'un système de gestion de base de données qui définit chaque élément de données tel qu'il apparaît dans la base de données.

**Langage de description des services Web** : Cadre commun pour la description des tâches effectuées par un service Web afin que d'autres applications puissent utiliser ce service.

**Langage de manipulation de données** : Langage associé à un système de gestion de base de données dont se servent les utilisateurs finaux et les programmeurs pour manipuler les données de la base.

**Langage de modélisation unifié** : Voir **UML**.

**Langage de quatrième génération** : Langage de programmation que les utilisateurs finaux ou les programmeurs moins qualifiés peuvent utiliser directement pour développer des applications plus rapidement qu'au moyen des langages de programmation traditionnels.

**Langage de requête structuré** : Voir **SQL**.

**Langage orienté objet** : langage de programmation manipulant des objets, c'est-à-dire fonctionnant avec des sortes de modules ou d'éléments prédéfinis. Ainsi, par exemple, si l'élément « fenêtre » est défini, le programmeur devra juste rappeler celui-ci, sans avoir besoin de le décrire complètement. Il devra seulement renseigner certaines caractéristiques prédéfinies comme la taille ou la couleur de cette fenêtre.

**Large bande** : Technologie de transmission à haut débit.

**Licence utilisateur** : Entente commerciale par laquelle le propriétaire d'une propriété intellectuelle ou une autre partie qui détient des droits sur la propriété concède l'ensemble ou une partie de ses droits à une autre partie, selon des conditions et des limites convenues. Pour un logiciel soumis à une licence utilisateur commerciale, cela ne signifie pas que le logiciel vous appartient, mais que son prix d'achat est en fait une taxe correspondant à un droit d'utilisation pour 1 poste utilisateur. Ne pas respecter les conditions du contrat des logiciels correspond à un acte de piratage, répréhensible par la loi.

**Lien hypertexte** : Renvoi à un emplacement représenté par un texte, en couleur ou souligné, ou par un graphique. Un clic sur un lien hypertexte permet d'accéder à un fichier lié.

**Linux** : Système d'exploitation qui est dérivé d'Unix, compatible avec différentes plateformes matérielles, gratuit ou très économique.

**Liste de diffusion** : Liste d'adresses de personnes s'étant volontairement inscrites dans le but de recevoir des informations sur un sujet précis.

**Logiciel** : Ensemble des éléments informatiques qui permettent d'assurer une tâche ou une fonction.

**Logiciel de base** : Instructions préprogrammées et détaillées qui commandent et coordonnent le fonctionnement des composants du matériel d'un système d'information.

**Logiciel espion (Spyware)** : Technologie qui aide la collecte d'information sur une personne ou sur une organisation à son insu.

**Logiciel libre** : Logiciel qui offre un accès libre à son code programme, ce qui permet aux utilisateurs de le modifier pour y apporter des améliorations ou corriger des erreurs.

**Login (Ang.)** : Identifiant. Code d'entrée dans un système multi-utilisateur où chaque utilisateur avant de pouvoir commencer à travailler ou à exploiter le service, doit s'identifier par un nom ou un pseudo.

**Logique floue** : Intelligence artificielle (IA) à la base des règles pour la résolution de problèmes qui tolèrent l'imprécision, en utilisant des termes spécifiques, appelés « fonctions d'appartenance ».

**Macro** : Liste d'instructions (macro-commandes), destinée à une application bureautique comme Word, Excel -payants (ou OpenOffice, Atlantis, AbiWord –gratuits, LibreOffice). Celles-ci peuvent être exportées à la demande de l'utilisateur ou automatiquement au chargement du document dans lequel elles sont enregistrées.

**Macro-virus** : Catégorie de virus spécialement conçus pour s'insérer dans les macros.

**Mail bombing ou Spam (Ang.)** : Pourriel.

Technique de nuisance consistant à envoyer un très grand nombre de messages vers une boîte aux lettres électronique afin de la saturer.

**Mailing list (Ang.)** : Liste de diffusion.

**Mailto** : Dans une page Web.lien hypertexte vers une boîte e-mail. Ces liens figurent en général à la fin des pages Web.

**Maintenance** : Etape pendant laquelle on apporte des changements au matériel, au logiciel, à la documentation ou aux procédés d'un système en production pour corriger les erreurs, répondre aux nouvelles exigences ou en améliorer l'efficacité.

**Mashups** : On appelle « mashups » l'agrégation de logiciels récupérés de différentes sources et assemblés pour créer une application différente de l'addition des logiciels initiaux.

**Mémoire morte** : Composants de mémoire, présents sur la carte mère, qui ne s'effacent pas à l'extinction de l'ordinateur. La mémoire morte contient notamment le Bios (voir ce terme). Elle ne peut être modifiée que par une opération particulière, appelée flashage.

**Mémoire tampon** : Zone de mémoire vive ou de disque dur utilisée pour stocker temporairement des données, notamment entre deux processus ou matériels ne travaillant pas au même rythme. Certains périphériques peuvent posséder leur propre mémoire tampon.

**Mémoire vive** : C'est la mémoire de travail de l'ordinateur. Concrètement, des puces électroniques enfichées sur la carte mère. Un logiciel est chargé en mémoire vive (à partir du disque dur) lorsque vous le lancez, de même qu'un document que vous affichez sur l'écran. La mémoire vive s'efface entièrement quand vous éteignez l'ordinateur.

**Messagerie unifiée** : Système qui combine les messages vocaux, le courrier électronique et la télécopie de telle manière que l'ensemble des messages transitent par un seul système.

**Méthode de la période de récupération** : Méthode qui mesure le temps nécessaire pour rembourser les coûts de l'investissement initial d'un projet.

**Miage Maîtrise de Méthodes Informatiques Appliquées à la Gestion** : formation supérieure intégrée aux IUP (Instituts Universitaires Professionnalisés) effectuée en 3 ans après une année en Deug, DUT, BTS.

**Microblogging** : voir **Twitter**.

**Middleware (Ang.)** : Logiciel médiateur.

**Migration** : transfert d'un système informatique, de ses données et applications sur un système hétérogène.

**Mise à jour** : Actualisation d'un logiciel ou d'une base de données afin de profiter de ses dernières évolutions. Pour un bon niveau de sécurité, une mise à jour régulière de logiciels de surveillance (antivirus, anti-espions) est indispensable.

**MOA (maîtrise d'ouvrage)** : le client (personne physique ou morale) auquel le maître d'oeuvre doit fournir son service.

**Mode de transfert asynchrone (Asynchronous Transfer Method, ATM)** : Technologie de réseau qui divise l'information en cellule de 8 octets, ce qui permet la transmission de données entre les ordinateurs de différents fournisseurs, à n'importe quelle vitesse.

**Modèle client-serveur** : Modèle de répartition du traitement des données entre les clients et le serveur au sein d'un réseau, où chaque tâche est assignée à la machine la plus apte à l'accomplir.

**Modèle économique (Business model)** : Représentation de ce qu'est une entreprise et de la manière dont elle livre un produit ou un service, et qui montre comment l'entreprise devient productive.

**Modèle d'évaluation du prix des options réelles** : Modèle qui sert à évaluer les investissements dans la technologie de l'information qui procurent des rendements incertains au moyen de techniques pour évaluer les options financières.

**Modèle organisationnel du flux poussé (Push model) :** Chaîne logistique régie par des plans de production fondés sur les meilleures prévisions ou approximations concernant la demande de produits. Les produits sont « poussés » vers les clients.

**Modèle organisationnel du flux tiré (Pull model) :**

**Modèle des points de fonction :** Méthode rapide qui permet de choisir parmi différentes solutions de systèmes, en fonction de leur capacité à atteindre les objectifs fixés.

**Modèle de référence (Open System Interconnexion Model, OSI) :** Modèle de connectivité réseau développé par l'Organisation internationale de standardisation.

**Modèle des forces concurrentielles (Competitive force model) :** Modèle qui sert à décrire l'interaction des influences externes, plus précisément des menaces et des opportunités, qui agissent sur la stratégie d'une organisation et sur sa capacité à affronter la concurrence.

**Modem : ((MODulateur-DEModulateur)** Dispositif qui convertit les signaux numériques en signaux analogiques, et vice versa. Le modem sert d'interface entre l'ordinateur et le réseau téléphonique pour l'établissement d'une connexion à distance. C'est le mode d'accès à Internet le plus répandu, encore très loin devant le câble ou l'ADSL.

**MOE (maîtrise d'oeuvre) :** la personne physique ou morale qui doit fournir un service à son client, alors appelé maître d'ouvrage.

**Mois-homme :** Unité de mesure traditionnelle qu'utilisent les concepteurs pour estimer le temps nécessaire à l'aboutissement d'un projet. Cette mesure désigne la quantité de travail qu'une personne devrait effectuer en l'espace d'un mois.

**Mot de passe :** Élément incontournable dès qu'il s'agit de sécurité, le mot de passe sert à authentifier l'utilisateur dans tous les cas où un accès est protégé : ouverture d'une session Windows, établissement d'une connexion Internet, décryptage d'un fichier etc.

**Moteur d'interférence :** Logiciel d'IA fondé sur la méthode utilisée pour chercher dans la base de règles d'un système expert (chaînage avant ou arrière).

**Multidiffusion :** Transmission de données vers un groupe précis de destinataires.

**MMS (Multimedia Message Service) :** Service de messagerie qui permet d'envoyer et de recevoir sur un téléphone mobile des messages contenant du texte, des images, du son et des clips vidéo ou audio. Ce type de service est possible depuis l'apparition des réseaux 3G.

**MVC : Modèle-vue-contrôleur** est un motif d'architecture logicielle destiné aux interfaces graphiques lancé en 1978 et très populaire pour les applications web. Le motif est composé de trois types de modules ayant trois responsabilités différentes : les modèles, les vues et les contrôleurs.

**MVNO (Mobile Virtual Network Operator) :** Opération de téléphonie mobile qui ne possèdent pas de concession de spectre de fréquences ni d'infrastructure de réseau propre, contraste des accords avec les opérateurs mobiles possédant un réseau mobile pour leur acheter un forfait d'utilisation et le revendre sous sa propre marque à ses clients.

**Mystification** : Tentatives qui visent à cacher la véritable identité en utilisant de fausses adresses de courrier électronique.

**Nanotechnologie** : Technologie de construction de structures et de processus fondée sur la manipulation d'atomes et de molécules.

**Nappe** : Câble spécifique permettant de réaliser le chaînage des périphériques internes tels que les disques durs ou les lecteurs de CD/DVD.

**Navigateur Web (Browser)** : Outil logiciel convivial pour l'accès au Web et à Internet.

**Nearshore** : il s'agit de la sous-traitance dans un pays proche, généralement en Europe du Sud et de l'Est, et en Afrique du Nord.

**Nettoyage de données** : Activités de détection et de correction des données dans une base de données ou dans un fichier incorrect, incomplet, mal structuré ou redondant.

**NFC (Near Field Communication)** : Technologie de communication de proximité (quelques centimes) lancés par Sony et Philips, qui permet d'échanger des données entre un lecteur et n'importe quel terminal mobile, ou entre les terminaux eux-mêmes.

**Niveau OSI (interconnexion de systèmes ouverts)** : ensemble de normes et de protocoles écrivant la communication (l'interface) entre des systèmes informatiques hétérogènes (Linux et Windows, par exemple) utilisés aussi dans l'interconnexion des réseaux.

**Nom de domaine** : Nom en langage courant qui correspond à un protocole IP numérique 32 bits pour chaque ordinateur connecté à Internet.

**Norme** : technologie ou procédé formalisé par un organisme habilité et indépendant.

**Norme technologique** : Spécifications qui établissent la comptabilité des produits (matériels et logiciels) et la capacité de communiquer au sein d'un réseau.

**Numérique** : Type de données ne pouvant prendre qu'un nombre limité, discret, de valeurs intermédiaires. S'oppose à analogique.

**Numériser** : Opération qui consiste à transformer les données analogiques d'une image ou d'un son en valeur numérique, interprétable par l'ordinateur.

**Objet** : Composant logiciel élémentaire qui combine les données et les procédures qui agissent sur ces données.

**O.C.R. (Ang. Optical Character Recognition)** : Logiciel utilisé avec un scanner lui permettant de reconnaître automatiquement les caractères d'une page numérisée et ainsi de la transformer en fichier texte au lieu de la conserver sous forme d'image.

**Octet** : Ensemble ordonné de 8 éléments binaires (bits) traités comme un tout. Un octet est souvent utilisé pour représenter un caractère alphanumérique. Lorsqu'une grandeur est exprimée en multiples de l'octet, les symboles Ko, Mo et Go représentent respectivement le kilo-octet, le méga-octet et le giga-octet.

**Offshore** : c'est l'opération de sous-traitance de tout ou partie d'un projet ou service informatique par un prestataire d'un pays étranger.

**Onduleur** : Périphérique que l'on place entre l'alimentation électrique et l'ordinateur. En régulant l'arrivée de l'électricité, il empêche les microcoupures ou les pics de courant qui pourraient provoquer des pertes de données ou endommager l'ordinateur. Lorsqu'une panne survient, il remplace l'alimentation électrique pendant quelques minutes (selon la capacité de l'onduleur), vous laissant ainsi le temps de sauvegarder les données.

**Onglet** : C'est souvent un rectangle en haut d'une fenêtre (un peu comme les intercalaires alphabétiques d'un répertoire d'adresse). En cliquant dessus, on accède à un thème se rapportant à un sujet précis.

**On-line data service (Ang.)** : Serveur

**Organigramme hiérarchique** : Documentation du système qui monte chaque niveau de conception, la relation entre les niveaux et leur position globale dans la structure de conception. Cet organigramme peut documenter un programme, un système ou une partie du programme.

**Organigramme virtuelle (Virtual organization)** : Organisation utilisant des réseaux qui relient des personnes, des actifs et des idées pour créer des produits et des services et pour les distribuer sans être limitée par les frontières organisationnelles traditionnelles ou par l'emplacement physique.

**Organizer (Ang.)** : Organiseur, agenda électronique  
Logiciel pour la gestion de données personnelles telles que rendez-vous, adresses, etc. Par extension, ordinateur de poche remplissant essentiellement cette fonction.

**Open source (en français, logiciel libre)** : c'est l'ensemble des applications fournies avec leurs codes sources.

Ces logiciels peuvent donc être adaptés en fonction de ses besoins propres.

**O.S. (Ang. Operating System)** : Système d'exploitation.

**Outil de contrôle formel** : Outil de gestion de projet qui aide à surveiller les progrès accomplis au moment des jalons (ou points de synchronisation) importants.

**Outil de planification formelle** : Type d'outil de gestion de projet qui structure et ordonne les tâches en évaluant le temps, l'argent et les ressources techniques nécessaires à leur exécution.

**Outils de localisation** : Services sans fil qui combinent la communication sans fil et les capacités de localisation du système mondial de positionnement (Global Positioning System, GPS).

**PaaS (Plateform as a Service)** : Deuxième niveau de prestation permis par les infrastructures de type Cloud ; mise à disposition d'une infrastructure laquelle le client peut développer et/ou faire tourner des applications.

**Page d'accueil** : Affichage de texte et de graphiques qui accueille le visiteur d'un site et lui présente l'organisation qui a créé la page.

**Page Web** : Techniquement, c'est un fichier HTML, contenant du texte, des images, des liens vers d'autres pages et éventuellement de petits programmes (contrôle Active X, applets Java, etc.) Une page Web est identifiée sur Internet par son adresse URL.

**Page Web Dynamique** : Page créée suite à la demande d'un utilisateur, par exemple par la réponse à un questionnaire. Cette page aura une forme fixe mais le contenu variable sera adapté au profil de l'internaute.

**Palette** : C'est une fenêtre qui affiche un jeu de couleurs sous forme de petits carrés sur lesquels on clique pour choisir une couleur.

**P.A.O. (Publication Assistée par Ordinateur)** : Logiciel permettant des mises en page de documents complexes.

**Paramétrage** : Au sein de possibilités disponibles dans un logiciel applicatif, l'entreprise va décider des options qu'elle retient et fixer ses choix dans des tables de paramètres.

**Pare-feu** : Un pare-feu, ou firewall (en anglais), est un logiciel et/ou un matériel permettant de faire respecter la politique de sécurité du réseau, celle-ci définissant quels sont les types de communication autorisées sur ce réseau informatique.

**Partage de dossiers** : Si vous êtes relié à un réseau, certains dossiers dits partagés sont accessibles à tous les utilisateurs.

**Partition** : Subdivision du disque dur, identifiée sous Windows par une lettre de lecteur. Un disque dur peut ne comporter qu'une seule partition.

**Password (Ang.)** : Mot de passe.

**Patch (Ang.)** : Correctif.

**P.C.I. (Ang. Peripheral Component Interconnect)** : Interconnexion de Composants Périphériques  
Type de connecteur logé sur la carte mère de l'ordinateur et correspondant à un bus d'échange de données rapide normalisé. Ils permettent d'étendre les fonctionnalités de la machine en hébergeant des cartes d'extension.

**P.C.I. Express** : Nouvelle norme de bus remplaçant les anciennes normes AGP et PCI.

**Périphérique** : Ensemble des matériels électroniques entourant une unité centrale (Exemple: imprimante).

**Personnalisation** : Modification d'un logiciel par paramétrage pour qu'il réponde aux exigences uniques de l'organisation sans que l'intégrité fonctionnelle globale n'en soit altérée.

**Personnalisation du Web (Web personalization)** : Personnalisation d'un contenu Web pour un utilisateur spécifique.

**Perspective sociotechnique** : Approche des systèmes d'information fondée à la fois sur les approches techniques et les approches comportementales.

**P.G.P. (Ang. Pretty Good Privacy)** : Cet algorithme de cryptage de données a connu une certaine notoriété. Des logiciels permettant de le mettre en œuvre sont disponibles sur Internet.

**Phishing** : Technique utilisée par les escrocs en ligne pour usurper l'identité d'une personne ou d'une entité connue. A pour but de collecter des données confidentielles qui seront utilisées à mauvais escient.

**Pièce jointe** : Fichier attaché à un message. Tout type de fichiers peut être ainsi transmis par Internet : programme, texte, image, etc. Les pièces jointes peuvent être infectées par un virus.

**Pilote** : Programme contrôlant le fonctionnement d'une carte ou de tout périphérique à l'aide de bibliothèques adaptées.

**Pirate informatique (Hacker)** : Individu qui obtient un accès non autorisé à un réseau informatique pour son profit, afin de commettre un méfait ou pour son plaisir personnel.

**Pixel invisible (Web bug)** : Fichier graphique minuscule inséré dans un courriel ou dans une page Web et qui est conçu pour surveiller, à son insu, les habitudes de l'internaute.

**Place de marché électronique (Net marketplace)** : Place de marché numérique exclusive fondée sur la technologie d'Internet et qui relie plusieurs acheteurs à plusieurs vendeurs.

**Place de marché électronique (Exchange)** : Place de marché électronique indépendante principalement axée sur les transactions et qui met en communication plusieurs fournisseurs et vendeurs pour des achats au comptant.

**Plan de tests** : Plan préparé par l'équipe chargée de l'élaboration d'une solution informatique nouvelle en collaboration avec les utilisateurs. Il comprend la préparation d'une série d'essais à exécuter sur le système avant de décréter qu'il est prêt à être utilisé opérationnellement.

**Planification de la continuité des services (Business continuity)** : Planification qui s'intéresse aux moyens que l'entreprise peut utiliser pour continuer à fonctionner malgré une indisponibilité partielle ou totale de son informatique.

**Planification de la reprise après sinistre** : Planification de la restauration de l'informatique et des communications après une interruption de service.

**Plateforme collaborative** : Application logicielle favorisant le travail collaboratif en permettant le partage et la mise en commun de documents dans des espaces virtuels ergonomiques et à l'utilisation intuitive.

**Plateforme de service** : Intégration de diverses applications qui proviennent d'une multitude de fonctions administratives, d'unités fonctionnelles ou de partenaires d'affaires pour procurer une expérience transparente aux clients, aux employés, aux gestionnaires ou aux partenaires d'affaires.

**Plateforme (économie des)** : Le terme d'«économie des plateformes», comme celui d'«économie collaborative» dont l'emploi est très hétérogène, n'est attaché à aucune définition officielle et le champ retenu par les différentes études sur le sujet est plus ou moins large, comme l'atteste l'abondance des dénominations qui circulent (gig economy, sharing economy). L'économie des plateformes peut être vue comme sous-ensemble de l'économie collaborative. Dans son acceptation

la plus large, l'économie collaborative englobe toutes les plateformes digitales, qu'elles permettent d'échanger des informations ou des ressources privées, d'organiser la production et la vente de services, de référencer des petites annonces ou de co-produire des ressources publiques ou encore co-financer des projets.

**Point d'accès sans fil à Internet** : Endroit qui possède un point d'accès public à un réseau WiFi.

**Politique d'autorisation** : Politique qui détermine les différents niveaux d'accès aux éléments d'actif informationnels pour différents niveaux d'utilisateurs au sein d'une organisation.

**Pouce** : Unité de mesure d'origine anglo-saxonne. Dans le domaine informatique, le Pouce est surtout utilisée pour décrire une diagonale d'écran. Table de conversion : 1 pouce(1") = 2,54 cm (un écran 17" à une diagonale de 43,18 cm).

**P.P.P. (Point Par Pouce)** : Unité qui permet de définir la résolution d'une image à l'impression ou à l'écran (Grain). Plus le nombre est élevé, plus la qualité de l'image est grande.

**P.P.P. (Ang. Point to Point Protocol)** : Protocole permettant la connexion entre le modem d'un ordinateur et celui d'un fournisseur d'accès Internet, sur une ligne téléphonique classique.

**Presse-papier** : C'est une zone mémoire par laquelle transite temporairement toute sélection que vous copiez ou collez afin de permettre l'échange de texte ou d'image entre logiciels. Pour récupérer le contenu du presse-papiers, il suffit de le coller.

**Prise de contrôle à distance** : Communication entre deux ordinateurs, via un réseau local ou Internet par exemple, autorisant l'un des utilisateurs à prendre le contrôle de l'autre machine, comme s'il était installé devant, son clavier et sa souris commandant directement l'ordinateur distant.

**Processeur** : Partie d'un ordinateur qui interprète et exécute les instructions. Peut ce faire avec un microprocesseur.

**Procédures stockées** : est un ensemble d'instructions SQL précompilées, stockées dans une base de données et exécutées sur demande par le SGBD qui manipule la base de données.

**Programme** : Ensemble d'instructions destinées à être exécutées par l'ordinateur. Un programme est conservé sur le disque dur sous forme de fichiers exécutables (qui font appel les uns aux autres) et de différents fichiers de données contenant les paramètres. Pour s'exécuter, un programme doit être chargé en mémoire vive. Les termes logiciels ou applications sont utilisés de manière équivalente.

**Protocole** : Ensemble de normes, de conventions et de codes régissant les échanges de données entre ordinateurs. Différents protocoles sont mis en œuvre sur Internet parfois simultanément, pour la transmission des pages Web, le téléchargement des fichiers, l'émission et la réception de courriers électroniques, etc.

**Prototypage** : Processus de construction rapide et peu coûteux d'une version expérimentale d'un système, version que les utilisateurs peuvent évaluer. Leurs remarques serviront à améliorer la qualité du système final.

**Proxy ou proxy server (Ang.)** : Serveur mandataire, mandataire.

Un serveur Proxy est un ordinateur installé chez votre fournisseur d'accès qui stocke les pages Web les plus consultées, de manière à les envoyer plus rapidement lorsqu'elles sont demandées. Il joue en

outre le rôle de tampon entre votre PC et Internet limitant les possibilités d'intrusion ou de transmission de virus.

**Pseudo (pour pseudonyme)** : Identifiant.

**Raccourci** : Icône (comportant une flèche sous Windows) qui correspond à un élément et sur lequel on peut cliquer directement pour l'ouvrir sans avoir à en rechercher l'original.

**R.A.M. (Ang. Random Access Memory)** : Mémoire vive.

**Redémarrage** : Après la modification de paramètres du système ou l'installation d'un nouveau logiciel, les fichiers d'initialisation de l'ordinateur sont souvent légèrement modifiés ou enrichis. Il faut alors redémarrer l'ordinateur pour en tenir compte.

**Recettage unitaire** : est le résultat et bilan du test unitaire.

**Répertoire** : C'est le nom par lequel on désignait les dossiers du disque dur à l'époque du DOS ou des premières versions de Windows.

**Réplication** : Processus de copie d'un ou d'une partie d'un fichier à un autre emplacement.

**Résolution** : Indique le niveau de qualité d'une image. Plus la résolution n'est élevée, meilleure sera la qualité de l'image. La résolution d'une image peut s'exprimer en DPI ou en pixels. Pour un moniteur, on parle également de pitch.

**Restauration du système** : Dispositif permettant de rétablir le système dans un état antérieur stable lorsqu'un dysfonctionnement apparaît

**Restaurer** : Récupérer des données à partir d'une copie de sauvegarde.

**RFID (Radio Fréquence Identification)** : il s'agit d'une identification par fréquences radio, reposant sur l'implantation dans les objets physiques que l'on veut suivre de petites puces actives ou passives qui envoient un signal unique, d'elles-mêmes ou en réponse à un signal de stimulation.

**R.O.M. (Ang. Read Only Memory)** : Mémoire morte.

**Saut de page** : Séparation entre deux pages, qui est insérée automatiquement par le logiciel ou créée manuellement.

**Sauvegarder** : Souvent employé comme synonyme de Enregistrer (voir ce terme), la sauvegarde désigne plus spécifiquement la copie de données sur un support autre que leur support original afin de disposer d'un double par sécurité.

**Scanner (Ang. dérivé)** : Numériser.

**Secteur d'amorçage** : Zone du disque dur ou d'une disquette de démarrage sur laquelle figurent les instructions nécessaires pour le chargement du système d'exploitation. Cette zone est automatiquement chargée en mémoire au démarrage de l'ordinateur.

**Série** : Connecteur permettant d'envoyer des données élémentaires bit à bit. Situé à l'arrière des PC, il est destiné aux modems ou aux souris mais également utilisés pour d'autres périphériques. Il est progressivement remplacé par le port USB.

**Server (Ang.)** : Serveur.

Tout ordinateur délivrant un service à un autre ordinateur : connexion à un réseau local (serveur de réseau) ou à Internet (serveur du fournisseur d'accès), réception et envoi de courriers électroniques (serveur de messagerie), envoi de pages Web (serveur Web, envoi de fichiers (serveur de téléchargement), etc.

**Session** : Lorsque vous utilisez un compte (compte d'utilisateur, compte d'accès à Internet, etc.) pour vous connecter à votre espace de travail, à Internet ou à un réseau, on dit que vous ouvrez une session. L'ouverture d'une session exige que vous déclinez un nom d'utilisateurs et en général, que vous donniez un mot de passe pour votre authentification. Le terme de session est également utilisé pour l'établissement de communications entre un ordinateur client et un ordinateur serveur (voir ces termes).

**Script** : Ensemble d'instructions destinées à être exécutées par un logiciel.

**Shareware** : Logiciel distribué gratuitement, mais que vous devez payer à l'issue d'une période d'essai (en général 30 jours) si vous désirez le conserver. La plupart des logiciels disponibles en téléchargement gratuit sur Internet sont des Sharewares et non des Freewares.

**SHIFT (Ang.)** : Touche du clavier permettant de passer de minuscule en majuscule, d'avoir accès à certains caractères du clavier ou de faire une sélection multiple par glisser-déposer.

**Signature de virus** : Suite caractéristique de codes, présents dans un virus, qui permettent de détecter sa présence à coup sûr. Les antivirus utilisent une base de données de signatures de virus qui doit être régulièrement mise à jour.

**Signature électronique** : Code électronique délivré par un organisme de certification permettant d'authentifier l'identité de l'expéditeur d'un e-mail.

**Signet** : C'est une marque placée dans un document qui sert à repérer un emplacement, un mot ou une valeur. Par extension, il peut s'agir d'un bookmark internet.

**Skin (Ang.)** : Habillage

Petit fichier qui permet de modifier l'aspect d'un logiciel, par exemple les boutons, le fond, des couleurs, ...

**S.M.B. (Ang. Server Message Block)** : Système de fichier réseau.

Composant destiné à gérer le partage des dossiers et des fichiers sur un réseau. Le plus connu est SAMBA.

**Smiley ou emoticon (Ang.)** : Frimousse.

**S.M.T.P. (Ang. Simple Mail Transfer Protocol)** : Protocole utilisé pour le transfert des e-mails à expédier depuis votre PC vers le serveur de votre compte de messagerie, puis pour leur transport jusqu'aux serveurs destinataires.

**Socket (Ang.)** : Connecteur.

**Software (Ang.)** : Logiciel.

**Spam ou Spamming (Ang.)** : Pourriel. Harcèlement publicitaire et commercial par e-mails.

**Spécifications fonctionnelles** : est la description des fonctions d'un logiciel en vue de sa réalisation.

**Spécifications techniques** : d'un cahier des charges sont une documentation des méthodes, procédés, et technologies sélectionnées pour faire face aux contraintes de réalisation du projet.

**Spyware (Ang.)** : Terme anglais pour logiciel espion.

**SQL : Structured Query Langage.** Langage standard de requête et de manipulation des données, destiné aux systèmes de gestion de bases de données relationnelles.

**SSL (Ang. Secured Socket Layer)** : Protocole de communication sécurisé dans lequel les données échangées sont captées, Utilisé pour les achats en ligne par carte bancaire ou la consultation de données confidentielles (compte bancaire).

**Système d'exploitation (operating system)** : ensemble de programmes fournis par les constructeurs de machines et qui en permettent l'exploitation.

**Système d'information** : l'ensemble des moyens (organisation, acteurs, procédures, systèmes informatiques) nécessaires au traitement et à l'exploitation des informations d'une entreprise.

**Système de fichiers** : Norme utilisée pour la gestion des fichiers et des dossiers sur le disque dur Windows 98 et Millennium prennent en charge les systèmes de fichier FAT (File Allocation Table) et FAT32. Windows XP peut également utiliser le système NTFS (NT File System, hérité de Windows NT), qui apporte notamment la possibilité d'interdire l'accès à certains dossiers.

**Table d'allocation de fichier** : Table dans laquelle sont stockées toutes les informations sur la structure des dossiers et la répartition des différents fichiers dans les systèmes de fichier FAT et FAT32. C'est une sorte d'annuaire dont se sert le système pour lire ou écrire des fichiers sur le disque dur. Souvent cible d'attaques virales.

**Tabulation** : Elle correspond à l'endroit où vous souhaitez disposer et aligner le texte sur la ligne. Il existe quatre types de tabulation : gauche, droite, centrée et décimale.

**Tâche de fond** : Les systèmes d'exploitation sont aujourd'hui tous multitâches, c'est-à-dire qu'ils permettent l'exécution simultanée de plusieurs logiciels.

Ainsi, vous pouvez utiliser votre traitement de texte ou votre navigateur Internet pendant qu'un antivirus veille en tâche de fond.

Certains OS font apparaître les logiciels fonctionnant en tâche de fond dans une barre des tâches.

**Tag** : Balise.

**Taux de transfert** : Vitesse de transmission des informations entre 2 périphériques. S'exprime en méga-octets/seconde (Mo/s).

**TCP/IP (Transmission Control Protocol/Internet Protocol)** : Protocole de communication de base sur Internet, qui régit le fonctionnement de tout le réseau et gère l'attribution des adresses IP aux ordinateurs connectés. Les autres protocoles utilisés (FTP, HTTP, POP3, SMTP etc.) viennent comme un: strate supplémentaire sur les bases fondées par TCP/IP.

**Téléchargement** : Transfert d'un fichier entre deux ordinateurs via Internet. En général, il s'agit de récupérer sur votre ordinateur un programme ou un document mis à disposition par un serveur de fichiers.

**Test unitaire** : Un test unitaire est un procédé permettant de s'assurer du bon fonctionnement d'une unité de programme.

**Test d'intégration** : Le test d'intégration est une phase dans les tests, qui est précédée des tests unitaires et est généralement suivie par les tests de validation. Dans le test unitaire, on vérifie le bon fonctionnement d'une partie précise d'un logiciel ou d'une portion d'un programme (appelée « unité » ou « module ») ; dans le test d'intégration, chacun des modules indépendants du logiciel est assemblé et testé dans l'ensemble.

**Thème** : Ensemble de fichiers de configuration permettant de changer l'apparence d'un environnement (bureau, site web, ...).

**Tierce Maintenance Applicative (TMA)** : maintenance des applications spécifiques d'une entreprise par une autre société (SSII...).

**UC (Unité centrale)** : Unité de traitement informatique composé de la carte mère, du microprocesseur et de la mémoire vive.

**UMTS**: technologie permettant la connexion à hauts débits (2 mégabits/seconde contre 9,6 kilobits/seconde pour le GSM).

Cette norme devrait succéder à la norme GSM et sera utilisée par les téléphones et terminaux mobiles de 3e génération.

**Unix** : Système d'exploitation pour tous les types d'ordinateurs, indépendant de la machine et qui supporte le traitement multi-utilisateurs, le multitâche ainsi que la plupart des types de réseaux.

**Upgrader** : mettre à jour un système.

**V.G.A. (Ang Video Graphics Array)** : Standard d'affichage analogique pour ordinateurs. Par extension, norme pour tout matériel utilisant ce standard (Exemple: carte graphique VGA, câble VGA, ...).

**Vectériel** : Se dit d'une image ou d'une police de caractères dont les contours sont décrits par des courbes géométriques et non point par point (bitmap). Un fichier vectoriel est beaucoup moins encombrant qu'un bitmap. Les images simples, comme les logos, sont le plus souvent enregistrés dans ce mode.

**Ver informatique** : Programme indépendant qui se propage lui-même pour perturber le fonctionnement d'un réseau informatique ou détruire des données et d'autres programmes.

**Virus informatique** : Programme pirate qui s'attache à d'autres programmes ou fichiers de données afin de s'exécuter, ce qui cause un mauvais fonctionnement du matériel ou des logiciels.

**WAN (Wide Area Network)** : Réseau de télécommunications qui permet la transmission de messages sur des grandes distances via des réseaux publics et en recourant aux services d'opérateurs de télécommunications via des câbles, des satellites, etc.

**War Driving** : Dispositifs écoute situés à l'extérieur des immeubles qui tentent de capter le trafic d'un réseau sans fil.

**Web (World Wide Web)** : Système qui comporte des normes mondialement reconnues pour le stockage, la récupération, la mise en forme et l'affichage sous de multiples formes (audio, vidéo, etc.) de l'information dans un environnement interconnecté.

**Webconférence** : Conférence téléphonique et vidéo passant par le réseau Internet.

**Webmaster (Ang.)** : Administrateur de site ou de serveur, webmestre. Personne chargée de la gestion du site Web ou que l'on peut déposer sur le Bureau pour ajouter des fonctionnalités.

**Web sémantique** : Effort collectif mené par le consortium W3C pour accroître l'efficacité de la recherche sur le Web en réduisant la nécessité d'une intervention humaine dans la recherche et le traitement de l'information Web.

**Web site ou website (Ang.)** : Site de la toile, site sur la toile.

**WiFi (Wireless Fidelity)** : c'est la dénomination commerciale qui regroupe des produits de réseaux sans fil compatibles entre eux, qu'ils soient ou non certifiés Wifi.

**Wiki** : Type spécifique de site Web qui encourage les contributions.

**W.W.W. (Ang. World Wide Web)** : Toile d'araignée mondiale, toile mondiale ou T.A.M.



---

## Lexiques de l'informatique selon le journal officiel Français

Ce lexique est basé sur le *JOURNAL OFFICIEL DE LA RÉPUBLIQUE FRANÇAISE* de la rubrique « Vocabulaire de l'informatique (liste de termes, expressions et définitions adoptés) » des années 2005 à 2019.

**aide au dépannage** : (nom masculin) Protocole d'assistance proposé à l'utilisateur d'un logiciel ou d'un matériel pour lui permettre de résoudre une difficulté d'emploi ou de faire face à une défaillance technique.

*En anglais* : troubleshooting.

**arrière-guichet** : (nom masculin) Support informatique et logistique d'un ou de plusieurs guichets.

*En anglais* : back office.

**apportez votre équipement personnel de communication : (AVEC)** Se dit de l'utilisation, dans un cadre professionnel, d'un matériel personnel tel qu'un téléphone multifonction ou un ordinateur.

*En anglais* : bring your own device (BYOD).

**atelier numérique ouvert** : (ANO) Lieu physique ou espace en ligne ouvert aux spécialistes et aux passionnés de technologies numériques, notamment de programmation, qui est destiné au partage de savoir-faire et de ressources.

*Note*: L'atelier numérique ouvert favorise notamment l'émergence de nouvelles pratiques ou la création.

*En anglais* : hackerspace, hacklab, hackspace.

**Bannière** : (nom féminin) Zone d'écran plus haute que large, à vocation publicitaire, informative ou décorative, située sur un des côtés de la page d'un site.

*En anglais* : skyscraper.

**bloc-notes** : (nom masculin) Site sur la toile, souvent personnel, présentant en ordre chronologique de courts articles ou notes, généralement accompagnés de liens vers d'autres sites.

*En anglais* : blog, web log, weblog.

**Blogue** : Site, souvent personnel, présentant, du plus récent au plus ancien, de courts articles ouverts aux commentaires des internautes.

*En anglais* : blog, weblog.

**bloqueur de publicités** : Module d'extension d'un navigateur dont la tâche est de repérer les messages publicitaires et d'en empêcher l'affichage.

*Note:* 1. Un bloqueur de publicités fait appel à des moteurs d'inférence et à des techniques d'apprentissage automatique. 2. Un bloqueur de publicités réajuste en permanence ses techniques de détection en fonction des contremesures prises par les régies publicitaires.

*En anglais :* ad blocker, adblocker.

**bombe logique :** équivaut à bombe programmée.

**bombe programmée :** (nom féminin) Logiciel malveillant conçu pour causer des dommages à un système informatique et qui est déclenché lorsque certaines conditions sont réunies.

*En anglais :* logic bomb.

**Brouillage :** (nom masculin) Transformation du code source d'un programme destinée à le rendre incompréhensible.

*Note :* Le brouillage permet notamment de protéger un programme d'une réutilisation abusive.

*En anglais :* obfuscation.

**Canular :** (nom masculin) Information fautive transmise par messagerie électronique et incitant les destinataires abusés à effectuer des opérations ou à prendre des initiatives inutiles, voire dommageables.

*Note :* Il peut s'agir d'une fautive alerte aux virus, de fautes chaînes de solidarité ou pétitions, de promesses de cadeaux. Les canulars encombrant les réseaux par la retransmission en chaîne du message et incitent parfois le destinataire à effacer des fichiers indispensables au bon fonctionnement de son ordinateur ou à télécharger un fichier contenant un virus ou un ver.

*En anglais :* hoax.

**Carrousel :** (nom masculin) Menu graphique, généralement tridimensionnel, qui fait défiler de manière cyclique sur un écran un ensemble d'images donnant chacune accès à un contenu.

*En anglais :* cover flow-like design.

**Cartouche :** (nom masculin) Fenêtre présentant de façon synthétique et structurée des informations textuelles ou graphiques sur un sujet, qui apparaît notamment en marge de la page des résultats d'une recherche sur la toile.

*Note:*

1. À titre d'exemple, le cartouche relatif à une commune peut afficher son nom, sa localisation, sa superficie et sa population.

2. On trouve aussi, dans le langage professionnel, le terme «encadré» et le terme «infoboîte», qui n'est pas recommandé. *Voir aussi:* graphe de connaissances.

*En anglais :* infobox.

**cédérom autonome :** Cédérom comportant un système d'exploitation qui fonctionne sans installation préalable.

*En anglais :* live CD.

**centre de données** : Site physique où sont regroupées des infrastructures informatiques et de télécommunication destinées à stocker, à traiter ou à distribuer des données de façon sécurisée.

*Note:* 1. Une organisation peut posséder son ou ses propres centres de données ou recourir à ceux d'un prestataire de services. 2. L'informatique en nuage recourt aux centres de données. 3. La taille de certains centres de données nécessite leur implantation à proximité d'une source d'énergie importante et d'une dorsale.

*En anglais* : data center (EU), data centre (GB).

**cercle de confiance** : Ensemble des fournisseurs et des utilisateurs d'un espace de confiance.

*En anglais* : circle of trust.

**charge utile** : Partie d'un élément transmis, tel qu'un message électronique, un flux de données ou un programme d'installation, qui correspond au contenu à transmettre et non aux données d'acheminement.

*En anglais* : payload.

**cheval de Troie** : Logiciel apparemment inoffensif, installé ou téléchargé et au sein duquel a été dissimulé un programme malveillant qui peut par exemple permettre la collecte frauduleuse, la falsification ou la destruction de données.

*En anglais* : Trojan

**Collecteur** : (nom masculin) Programme qui parcourt la toile pour en extraire des éléments de repérage de contenus, destinés à être utilisés par un moteur de recherche.

*En anglais* : crawler.

**conception d'interfaces adaptatives** : Élaboration et mise en oeuvre de méthodes et de techniques qui permettent d'adapter automatiquement les interfaces avec l'utilisateur à différents types de terminaux, fixes ou mobiles.

*Note:* La conception d'interfaces adaptatives peut prendre en compte les options ou les configurations choisies par les utilisateurs, notamment en matière d'ergonomie.

*En anglais* : responsive design, responsive web design (RWD).

**conférence informelle** : Rencontre proposant, sur un sujet donné, un ou plusieurs ateliers ouverts à toute personne souhaitant y contribuer activement, pour aboutir à des conclusions collectives.

*En anglais* : barcamp, unconference.

**Cyberconférence** : Téléconférence utilisant l'internet ou des réseaux informatiques privés et pouvant combiner le son, l'image et le texte.

*Note:* Une cyberconférence permet notamment la présentation, l'échange, le partage ou la modification de documents.

*En anglais* : e-conference, web conference, web conferencing.

**Démon** : (nom masculin) Programme actif en permanence dans un système multitâche, qui effectue certaines fonctions sans l'intervention de l'utilisateur.

*En anglais* : daemon, disk and execution monitor.

**démon de messagerie** : Démon qui assure l'acheminement des messages électroniques et avertit l'expéditeur lorsqu'un message n'a pu atteindre son destinataire.

*En anglais* : mailer daemon.

**Dévoisement** : (nom masculin) Technique consistant à détourner subrepticement des communications à destination d'un domaine vers une adresse différente de son adresse légitime.

*En anglais* : pharming.

**directeur, -trice des données** : Personne chargée, au sein d'une organisation, de définir la politique de gestion des données, de déterminer celles qui sont utiles à la prise de décision et de s'assurer de la qualité et de la cohérence de ces dernières.

*En anglais* : chief data officer (CDO).

**données ouvertes** : Données qu'un organisme met à la disposition de tous sous forme de fichiers numériques afin de permettre leur réutilisation.

*Note* :

1. Les données ouvertes n'ont généralement pas de caractère personnel.
2. Elles sont accessibles dans un format favorisant leur réutilisation.
3. La réutilisation des données ouvertes peut être soumise à conditions

*En anglais* : open data.

**encre électronique** : Matériau composé de microcapsules contenant des particules dont le déplacement, sous l'effet d'un champ électrique, permet d'afficher des textes ou des images.

*En anglais* : e-ink, electronic ink.

**enregistreur de frappe** : Dispositif conçu pour enregistrer la succession des frappes effectuées par un utilisateur sur un clavier.

*Note* : L'enregistreur de frappe peut être un programme malveillant, qui opère à l'insu de l'utilisateur et permet, par exemple, de connaître son mot de passe.

*En anglais* : keylogger.

**environnement de développement** : Infrastructure logicielle qui permet la création de programmes.

*En anglais* : application framework, framework, integrated development environment (IDE).

**espace de confiance** : Ensemble de ressources, de services informatiques et de services de communication qui permettent des échanges dans des conditions de sécurité suffisantes et cohérentes.

*En anglais* : trust environment.

**expérience de l'utilisateur** : Ensemble des impressions que l'utilisateur retire de son interaction avec un dispositif numérique et qui tiennent à l'interface et aux fonctionnalités de ce dispositif.

*Note* : L'expérience de l'utilisateur est prise en compte dans la conception et l'évolution des dispositifs numériques.

*En anglais* : user experience (UX).

**expert, -e en mégadonnées** : Spécialiste de l'extraction et de l'exploitation d'informations pertinentes à partir de mégadonnées, qu'il organise, traite et interprète à l'aide d'outils statistiques, mathématiques et informatiques.

*En anglais* : data scientist.

**exploration de modèle** : Méthode de vérification algorithmique, qui permet de déterminer avec efficacité si un système représenté par un modèle satisfait à un ensemble de spécifications formelles et qui, si une des spécifications n'est pas vérifiée, fournit des contre-exemples servant à identifier la source des erreurs.

*Note* : L'exploration de modèle trouve de nombreuses applications dans les industries du logiciel et du matériel, la vérification de puces, les protocoles de communication, les logiciels pilotes de périphériques, les systèmes critiques embarqués et les algorithmes de sécurité.

*En anglais* : model checking.

**fenêtre d'attente** : Fenêtre qui s'affiche provisoirement sur un écran pendant l'installation d'un logiciel.

*Note* : Cette fenêtre incite l'utilisateur à patienter pendant l'exécution de l'opération tout en lui apportant diverses informations.

*En anglais* : splash screen.

**flux de dépêches** : Suite de données générées automatiquement par un site de la toile pour permettre l'accès aux nouveautés de ce site.

*Note* : Certains sites proposent, sur leur page d'accueil, un bouton qui donne la possibilité de s'abonner au flux de dépêches.

*En anglais* : really simple syndication (RSS), RSS feed.

**fureteur, -euse, n.** : Personne qui, dans un espace d'échanges de l'internet tel qu'un forum ou un blogue, suit les discussions ou consulte les articles sans apporter de contribution.

*En anglais* : lurker.

**gestionnaire d'évènement** : Programme en mémoire qui s'active lorsqu'un évènement particulier survient au cours de l'exécution d'un autre programme.

..

*Note* : La notion d'évènement recouvre aussi bien l'action d'un utilisateur que l'arrivée d'un traitement lié à une date et une heure ou toute autre circonstance susceptible d'interrompre le traitement en cours.

*En anglais* : event handler, handler.

**gestion de versions** : Ensemble de techniques qui permettent de conserver les modifications successives d'un logiciel ou d'un document et leur historique ainsi que d'en restituer toute version antérieure.

*Note*: On trouve aussi, dans le langage professionnel, le terme «versionnage», qui n'est pas recommandé.

*En anglais* : versioning.

**graphe de connaissances** : Représentation graphique des relations sémantiques entre des sujets tels que des notions, personnes, organismes ou lieux, qui permet d'exposer de manière synthétique un ensemble de connaissances.

*En anglais* : knowledge graph.

**Guichet** : (nom masculin) Interface permettant d'accéder aux services en ligne proposés par une entreprise ou une organisation.

*En anglais* : front office.

**Imagette** : (nom féminin) Image dont la taille a été réduite par rapport à l'original pour limiter l'encombrement de l'espace d'affichage.

*En anglais* : thumbnail.

**impression en ligne** : Service en ligne qui permet de faire réaliser tout type d'impression et de façonnage.

*En anglais* : e-printing, web2print, web-to-print

**indexation personnelle** : Classification de contenus de l'internet par l'attribution de mots-clés librement choisis par un utilisateur.

*En anglais* : folksonomy.

**informatique en nuage** : Mode de traitement des données d'un client, dont l'exploitation s'effectue par l'internet, sous la forme de services fournis par un prestataire.

*Note* : L'informatique en nuage est une forme particulière de gérance de l'informatique, dans laquelle l'emplacement et le fonctionnement du nuage ne sont pas portés à la connaissance des clients.

*En anglais* : cloud computing.

**infrastructure à la demande** : Prestation de service qui propose à un client l'utilisation à distance d'une infrastructure comprenant du matériel et des logiciels, et dont le coût correspond à leur usage effectif.

*En anglais* : infrastructure as a service (IaaS).

**ingénierie inverse** : équivaut à rétro-ingénierie.

**interface avec l'utilisateur** : abrégé interface utilisateur. Synonyme : interface homme-machine (IHM). Interface informatique qui régit les interactions entre une machine et son utilisateur.

*Note*:

1. Les interfaces avec l'utilisateur peuvent notamment être graphiques, textuelles, tactiles ou vocales.
2. Pluriel: interfaces avec l'utilisateur, interfaces homme-machine, interfaces utilisateur.

*En anglais* : user interface (UI).

**internet clandestin** : Ensemble de réseaux conçus pour assurer l'anonymat des utilisateurs par la mise en œuvre d'une architecture décentralisée ainsi que de logiciels et d'autorisations d'accès spécifiques; par extension, l'ensemble des activités, souvent illicites, qui y sont pratiquées.

*Note*: 1. L'internet clandestin utilise notamment des réseaux privés virtuels ou pair à pair, ainsi que des méthodes de chiffrement et des processus de cooptation. 2. Il convient de distinguer l'«internet clandestin» de la «toile profonde».

*En anglais* : dark net, darknet.

**logiciel antivirus** : Logiciel conçu pour détecter les virus et les vers et, le cas échéant, les éliminer ou suggérer des remèdes.

*En anglais* : antivirus program, antivirus software, vaccine program.

**logiciel à la demande** : Prestation de service proposant à un client l'utilisation à distance d'un logiciel et dont le coût correspond à son usage effectif.

*En anglais* : software as a service (SaaS).

**logiciel malveillant** : Ensemble de programmes conçu par un pirate pour être implanté dans un système afin d'y déclencher une opération non autorisée ou d'en perturber le fonctionnement.

*En anglais* : malicious logic, malicious software, malware.

**logiciel rançonneur** : Logiciel malveillant qui empêche l'accès aux données stockées sur un ordinateur et propose leur récupération contre le paiement d'une rançon.

*Note*: 1. En général, un logiciel rançonneur chiffre les données de l'ordinateur cible en indiquant les instructions de paiement puis, si l'utilisateur y donne suite, fournit leur mode de récupération. 2. On trouve aussi le terme «rançongiciel», qui est déconseillé. *Voir aussi*: cheval de Troie, cyberattaque, logiciel malveillant, ver.

*En anglais* : ransomware.

**matériel à la demande** : Prestation de service proposant à un client l'utilisation à distance d'un matériel et dont le coût correspond à son usage effectif.

*En anglais* : hardware as a service (HaaS).

**marathon de programmation** : Rassemblement de développeurs, organisés par équipes, ayant pour objectif de fournir, en une ou deux journées, un prototype d'application, de service en ligne ou de produit.

*Note*: On trouve aussi, dans le langage professionnel, le terme «programmathon».

*En anglais* : hackathon.

**Masquage** : (nom masculin) synonyme : camouflage

Référencement abusif qui consiste à proposer deux versions différentes d'une même page de la toile, l'une visible par les internautes, l'autre destinée uniquement aux robots d'indexation des moteurs de recherche. *Note*: Du fait du masquage, le même site pourra être indexé plusieurs fois, modifiant ainsi le classement établi par les algorithmes.

*En anglais* : cloaking.

**Microblogage** : (nom masculin) Publication de minimessages sur un microblogue.

*En anglais* : microblogging.

**Microblogue** : (nom masculin) Blogue constitué de minimessages diffusés en temps réel, qui contiennent souvent des mots-dièse et dont l'enchaînement forme des fils de discussion.

*En anglais* : microblog, thread.

**modérateur, -trice**, n. : Personne qui veille au respect de l'objet et du règlement des échanges de messages électroniques effectués dans un cadre organisé.

*En anglais* : moderator.

**option d'adhésion (à)** : (loc.adj.) Se dit d'un fichier de données personnelles dans lequel un internaute ne peut être inscrit que s'il exprime explicitement son consentement.

*En anglais* : opt-in.

**option de retrait (à)** : (loc.adj.) Se dit d'un fichier de données personnelles dans lequel un internaute est inscrit sans son accord et continue de figurer tant qu'il n'a pas exprimé explicitement son refus.

*En anglais* : opt-out.

**ouverture des données** : Politique par laquelle un organisme met à la disposition de tous des données numériques, dans un objectif de transparence ou afin de permettre leur réutilisation, notamment à des fins économiques.

*En anglais* : open data, open data policy.

**papier électronique** : Support souple fonctionnant avec une source d'énergie et un microprocesseur, sur lequel peuvent s'afficher des textes ou des images au moyen d'encre électronique.

*En anglais* : electronic paper, e-paper.

**parcours sur la toile** : (nom masculin) Succession des pages visitées et des actions effectuées par un internaute, dont l'analyse permet de dégager les habitudes, les centres d'intérêt, les goûts.

*En anglais* : clickstream.

**portail de messagerie** : Service en ligne qui, au moyen d'un logiciel de navigation, permet le traitement des courriels et donne accès aux fonctions habituelles d'un logiciel de messagerie.

*Note* : Les fonctions associées les plus courantes sont la gestion d'un carnet d'adresses ainsi que la gestion des dossiers permettant de classer les messages reçus ou envoyés.

*En anglais* : web mail, webmail.

**Provisionnement** : (nom masculin) synonyme : allocation automatique. Dispositif qui permet d'allouer automatiquement des ressources informatiques ou de télécommunication afin de répondre à une variation d'activité d'un utilisateur.

*Note*: Le provisionnement fait généralement l'objet d'une clause contractuelle.

*En anglais* : provisioning.

**référencement abusif** : Ensemble de techniques qui leurrent les robots d'indexation des moteurs de recherche pour assurer à un site de la toile une meilleure position dans les résultats proposés par ces moteurs.

*Note*: Le référencement abusif recourt par exemple au masquage ou à la dissimulation, dans une page, d'éléments tels que des liens hypertextuels ou des mots clés, qui sont invisibles des internautes mais qui sont pris en compte par les robots d'indexation.

*En anglais* : spamdexing.

**rétro-ingénierie** : (nom féminin) Ensemble des opérations d'analyse d'un logiciel ou d'un matériel destinées à retrouver le processus de sa conception et de sa fabrication, ainsi que les modalités de son fonctionnement.

*En anglais* : reverse engineering.

**Riper** : (v.tr.) Extraire directement d'un support enregistré des données numériques en vue de les reporter sur un autre support, sans recourir à la conversion analogique intermédiaire habituellement nécessaire.

*En anglais* : rip (to).

**service de la toile** : Composant applicatif utilisant les protocoles de la toile, doté d'une adresse universelle et communiquant avec d'autres composants.

*En anglais* : Web service.

**signature de virus** : Suite d'éléments binaires commune à chacune des copies d'un virus ou d'un ver particulier, et utilisée par les logiciels antivirus pour détecter leur présence.

*En anglais* : virus signature.

**synthèse vocale de texte** : (SVT) en abrégé synthèse vocale. Conversion automatique d'un texte en un énoncé oral, qui recourt notamment aux techniques de traitement automatique des langues naturelles et de production de la parole.

*Note*: La synthèse vocale de texte facilite par exemple l'accès des personnes malvoyantes aux outils informatiques.

*En anglais* : text-to-speech (TTS).

**système d'engagement** : Ensemble des services mis en ligne par une organisation pour renforcer ses liens avec ses clients, ses partenaires extérieurs et ses employés ou pour créer des relations entre eux, dans le cadre d'une stratégie commerciale ou managériale, ou encore d'une stratégie de communication.

*Note*: 1. Les téléservices, les réseaux sociaux et les forums sont des exemples de services utilisés dans un système d'engagement. 2. Un système d'engagement permet notamment de recueillir des informations sur les utilisateurs, pour étudier par exemple leur comportement ou leurs centres d'intérêt.

*En anglais* : system of engagement.

**test captcha** (langage professionnel) : Test invitant l'utilisateur d'un site ou d'une application informatique à saisir correctement une suite de caractères présentés à l'écran d'une manière déformée, qui permet de s'assurer que la demande d'accès émane bien d'une personne et non d'un automate.

*En anglais* : captcha, captcha test, completely automated public Turing test to tell computers and humans apart.

**test de reconnaissance humaine** : voir test captcha.

**toile** : (nom féminin) Système de publication, de gestion et de consultation de documents et d'autres ressources, qui est réparti géographiquement et structurellement dans l'internet et fait appel aux techniques de l'hypertexte.

*En anglais* : Web, World Wide Web (WWW).

**toile profonde** : synonyme : abysse. Partie de la toile qui n'est pas accessible aux internautes au moyen des moteurs de recherche usuels.

*Note* : 1. La toile profonde comprend notamment le contenu de certaines bases de données ou de sites à accès restreint. 2. Il convient de distinguer la «toile profonde» de l'«internet clandestin».

*En anglais* : deepnet, deep web, hidden web, invisible web.

**toile sémantique** : Partie enrichie de la toile, dans laquelle la recherche de l'information peut être facilitée grâce à une indexation automatisée et structurée du contenu et des liens.

*Note* : L'indexation est obtenue au moyen d'un classement de l'information en fonction du sens des mots et des expressions des différentes langues naturelles.

*En anglais* : semantic web.

**transcription automatique de la parole** : (TAP) Conversion automatique d'un énoncé oral en texte, qui recourt notamment aux techniques de traitement du signal et de traitement automatique des langues naturelles.

*En anglais* : speech-to-text (STT), voice typing.

**usurpation d'adresse électronique** : Pratique consistant à utiliser frauduleusement une adresse de courrier électronique appartenant à une autre personne.

*En anglais* : e-mail spoofing.

**Ver : (nom masculin)** Logiciel malveillant indépendant qui se transmet d'ordinateur à ordinateur par l'internet ou tout autre réseau et perturbe le fonctionnement des systèmes concernés en s'exécutant à l'insu des utilisateurs.

*Note :*

1. Contrairement au virus, le ver ne s'implante pas au sein d'un autre programme.
2. Les vers sont souvent conçus pour saturer les ressources disponibles ou allonger la durée des traitements.

Ils peuvent aussi détruire les données d'un ordinateur, perturber le fonctionnement du réseau ou transférer frauduleusement des informations. Un ver peut produire des effets soit immédiatement soit de manière différée (à une date donnée, lors de la survenue d'un évènement ou par déclenchement d'une bombe programmée).

3. Bien qu'ils s'en distinguent, les vers sont parfois appelés « virus ».

*En anglais :* worm.

**virus**, n.m. : Logiciel malveillant, généralement de petite taille, qui se transmet par les réseaux ou les supports d'information amovibles, s'implante au sein des programmes en les parasitant, se duplique à l'insu des utilisateurs et produit ses effets dommageables quand le programme infecté est exécuté ou quand survient un évènement donné.

*En anglais :* virus.

---

## Lexiques Intelligence artificielle

**Agent :** Il ne s'agit pas ici d'un homme portant une cravate, mais bien d'un programme informatique qui agit de manière autonome. Les *bots* qui interagissent sur Internet en font partie. Certains systèmes reposent sur une multitude d'agents agissant de concert. Une nouvelle catégorie d'agent est capable d'apprentissage, on les désigne souvent sous le nom d'agents intelligents. Les agents plus conventionnels, eux, n'ont pas la capacité d'apprendre.

**Algorithme :** Un algorithme est une procédure ou une formule qui permet de résoudre un problème. Ce mot est dérivé du nom du mathématicien Muhammad ibn Musa al-Kharezmi qui était à la cour royale à Bagdad et a vécu de 780 à 850 environ.

Les travaux de ce mathématicien seraient également à l'origine du mot algèbre.

Les programmes informatiques peuvent être considérés comme des algorithmes sophistiqués. Dans les domaines des mathématiques et de l'informatique, un algorithme signifie généralement une petite procédure qui résout un problème récurrent.

**Algorithme génétique :** Méthode où l'on étudie un ensemble de solutions possibles et où les solutions les moins performantes sont éliminées. Les solutions les plus performantes sont combinées et étudiées successivement, jusqu'à en arriver à une solution optimale.

**Algorithme du simplex :** Méthode ayant été mise au point lors de la Deuxième Guerre mondiale par le mathématicien George Dantzig, qui vise à trouver la solution optimale à un problème, par exemple pour l'optimisation d'une chaîne de production. Cet algorithme peut être calculé manuellement, mais celui-ci et ses dérivés sont aujourd'hui intégrés à plusieurs solutions informatiques de gestion de la production et des approvisionnements.

**Algorithme évolutionniste ou évolutionnaire :** Un algorithme évolutionniste ou évolutionnaire (AE) met en œuvre des mécanismes inspirés de la nature pour résoudre des problèmes comme le feraient des organismes vivants. Il appartient aussi bien à l'informatique bio-inspirée qu'à l'informatique évolutionniste.

Les AE s'appuient sur les concepts de la théorie de l'évolution. Dans ces algorithmes, les solutions jouent le rôle des organismes individuels d'une population. Le panel des solutions possibles à un problème est d'abord une proposition aléatoire. La population est ensuite testée pour son aptitude à résoudre un problème correctement et rapidement. Puis les individus les plus adaptés sont sélectionnés pour la reproduction. Et le cycle recommence : on évalue le degré d'adaptation de la population et on élimine les individus les moins adaptés.

Comme les mécanismes des AE simulent les organismes vivants et leur évolution, ils font appel à des fonctions telles que la sélection, la reproduction, la mutation et la recombinaison. Le processus d'évolution qui consiste à choisir les meilleures solutions possibles à un problème selon un critère d'adaptation simule la sélection naturelle décrite par Darwin. Les solutions algorithmiques qui donnent les meilleurs résultats parmi les options proposées se reproduisent ; les moins adaptées sont éliminées et ne sont pas reproduites. Grâce à des tests mesurant le degré d'adaptation d'après les performances, l'optimisation s'opère sur plusieurs générations avec des fonctions comme la mutation.

Les AE sont un excellent outil d'optimisation de solutions. Toutefois, ces algorithmes ne trouvent pas nécessairement la solution optimale. En effet, ils ne cessent de trouver des solutions dont ils comparent les performances, ce qui peut ou non faire émerger LA meilleure solution. Il est également important de noter que les AE sont plutôt gourmands en ressources de calcul, notamment à cause de la complexité inhérente à la détermination du degré d'adaptation. Il est possible de réduire cette complexité par l'approximation de l'adaptation.

**Analyse en composantes principales :** Une autre méthode d'analyse de données, mise au point par Karl Pearson au début du 20<sup>e</sup> siècle, et dont on peut retracer l'origine dans des travaux remontant au 19<sup>e</sup> siècle<sup>10</sup>, par laquelle on transforme des données comportant un grand nombre de variables en un ensemble portant sur un nombre plus restreint de variables indépendantes les unes des autres, donc plus facile à traiter. Cette méthode est utilisée notamment dans le traitement d'images et dans le traitement de données sociales, pour faire ressortir les éléments les plus importants de ces données.

**Analyse syntaxique (voir aussi TLN/NLP) :** L'analyse syntaxique consiste à analyser des éléments d'un texte de façon ordonnée. En linguistique, l'analyse syntaxique scinde les mots et les expressions en différentes parties afin d'étudier les relations et les significations. Par exemple, on demande parfois aux élèves d'effectuer l'analyse syntaxique d'une phrase en séparant le sujet et le prédicat, puis les expressions dépendantes, les modificateurs et ainsi de suite.

En général, effectuer l'analyse syntaxique d'un écrit ou d'un discours signifie simplement l'interpréter.

En informatique, l'analyse syntaxique consiste à diviser un énoncé en langage informatique en plusieurs parties exploitables par l'ordinateur. Dans un compilateur, un analyseur syntaxique est un programme qui prend chaque déclaration écrite par un développeur et la découpe en morceaux (par exemple, la commande principale, les options, les objets cibles, leurs attributs, etc.). Ces morceaux peuvent ensuite être utilisés pour développer d'autres actions ou créer des instructions formant un programme exécutable.

**Apprentissage non supervisé (voir aussi Machine Learning) :** L'apprentissage non supervisé consiste à apprendre à un algorithme d'intelligence artificielle (IA) des informations qui ne sont ni classées, ni étiquetées, et à permettre à cet algorithme de réagir à ces informations sans supervision.

Dans ce mode d'apprentissage, le système d'IA peut regrouper des informations non triées en fonction de leurs similitudes et de leurs différences, même si aucune catégorie n'est indiquée. Les systèmes d'IA capables d'utiliser l'apprentissage non supervisé sont souvent associés à des modèles d'apprentissage génératifs, mais ils peuvent aussi fonctionner avec une approche basée sur la récupération (souvent associée à l'apprentissage supervisé). Les approches d'apprentissage supervisé ou non supervisé sont utilisées notamment dans les chatbots, les véhicules autonomes, les programmes de reconnaissance faciale, les systèmes experts et les robots.

Dans l'apprentissage non supervisé, les données fournies au système d'IA ne sont ni étiquetées, ni classées, et les algorithmes du système traitent les données sans aucun entraînement préalable. La sortie dépend des algorithmes codés. L'introduction dans un système d'une approche d'apprentissage non supervisé est un moyen d'expérimenter l'intelligence artificielle.

Les algorithmes d'apprentissage non supervisé peuvent exécuter des tâches de traitement plus complexes que les systèmes d'apprentissage supervisé, mais ils peuvent aussi être plus imprévisibles. Même si un système d'IA d'apprentissage non supervisé parvient tout seul, par exemple, à faire le tri entre des chats et des chiens, il peut aussi ajouter des catégories inattendues et non désirées pour y classer des races inhabituelles, créant la confusion au lieu de mettre de l'ordre.

**Apprentissage automatique :** Souvent désigné par l'expression anglaise *machine learning*, il s'agit d'une très vaste famille de méthodes permettant à une machine d'apprendre des comportements à partir des situations qui lui sont soumises, soit de manière supervisée, soit de manière semi ou non supervisée.

**Apprentissage par renforcement :** Cette méthode vise l'apprentissage du système d'intelligence artificielle en le soumettant à un environnement intégrant diverses situations. Le système est alors récompensé pour son comportement cumulatif, ce qui le guide dans son apprentissage<sup>1</sup>. Contrairement à l'apprentissage supervisé, le système ne reçoit pas d'indications directes sur le résultat souhaité pour une situation particulière. Cette méthode permet donc au système d'apprendre à traiter des situations nouvelles pour lesquelles il n'a pas été entraîné au préalable. Il s'agit d'une approche communément utilisée en robotique. C'est ce type de méthode qui a permis au célèbre logiciel *AlphaGo* de battre le champion mondial au jeu de Go<sup>2</sup>, un résultat qui était impensable jusqu'à tout récemment.

**Apprentissage profond :** Réseaux de neurones à plusieurs couches permettant un haut niveau d'abstraction. Ces réseaux sont entraînés par une méthode de rétropropagation<sup>3</sup>. Cette approche a connu un développement très important dans les dernières années et est en grande partie responsable de l'engouement actuel pour l'intelligence artificielle. Les applications en sont multiples, comprenant notamment la reconnaissance visuelle des visages et des objets, le traitement automatique des langues, ainsi que la recherche scientifique dans le domaine médical.

**Apprentissage supervisé (voir aussi Machine Learning) :** L'apprentissage supervisé, dans le contexte de l'intelligence artificielle (IA) et de l'apprentissage automatique, est un système qui fournit à la fois les données en entrée et les données attendues en sortie. Les données en entrée et en sortie sont étiquetées en vue de leur classification, afin d'établir une base d'apprentissage pour le traitement ultérieur des données.

Les systèmes d'apprentissage automatique supervisé alimentent les algorithmes d'apprentissage avec des quantités connues qui étayeront les futures décisions. Supervisé ou non supervisé, l'apprentissage automatique est utilisé notamment pour les chatbots, les véhicules autonomes, les programmes de reconnaissance faciale, les systèmes experts et les robots. Les systèmes

d'apprentissage supervisé sont associés pour la plupart à une intelligence artificielle basée sur la récupération, mais ils peuvent aussi reposer sur un modèle d'apprentissage génératif.

Les données utilisées pour l'apprentissage supervisé sont une série d'exemples comprenant des paires composées de sujets en entrée et de sorties attendues (appelées également signaux de supervision). Prenons l'exemple d'un système d'apprentissage supervisé pour le traitement d'images dans lequel on introduit des photos de véhicules appartenant aux catégories voitures et camions. Après un temps d'observation suffisant, le système doit être capable de faire la distinction entre plusieurs images non étiquetées et de les catégoriser ; une fois cet objectif atteint, l'apprentissage peut être considéré comme terminé.

Les modèles d'apprentissage supervisé présentent certains avantages sur les modèles non supervisés, mais ils ont aussi des limites. Par exemple, ils sont plus susceptibles de prendre des décisions auxquelles les humains peuvent s'identifier parce qu'elles reposent sur des données fournies par les humains. Mais dans le cas d'une méthode basée sur la récupération, les systèmes d'apprentissage supervisé ont des difficultés à traiter les nouvelles informations. Si un système qui connaît les catégories voitures et camions, par exemple, reçoit une image de vélo, il devra la placer dans l'une ou l'autre de ces deux catégories, ce qui sera incorrect. Alors que si le système était génératif, il ne saurait pas forcément reconnaître un vélo, mais il serait capable de l'identifier comme appartenant à une autre catégorie.

**Bayésien** : Réfère aux méthodes statistiques découvertes par Thomas Bayes au 18<sup>e</sup> siècle et à celles qui en découlent. Cette approche s'intéresse particulièrement à la probabilité d'un résultat compte tenu des données observées<sup>9</sup>. Elle est notamment utile quand la quantité de données à analyser est restreinte, par exemple pour des prédictions financières basées sur la probabilité de différents événements, comme la probabilité que la bourse augmente étant donné une hausse du taux directeur.

**CAPTCHA (Completely Automated Public Turing test to tell Computers and Humans Apart)** : Le CAPTCHA est un type de test de Turing appliqué à l'humain ! En effet, vous réussissez ce test lorsque vous recopiez quelques mots à partir d'images déformées ou floues sur Internet pour avoir accès à une page ou un service. Ces tests simples visent à éviter qu'un programme informatique accède à une base de données de manière systématique pour faire de l'exploration de données sans autorisation. Quand vous réussissez un tel test, vous confirmez que vous n'êtes pas un robot, ou une intelligence artificielle<sup>6</sup>. Toutefois, grâce notamment à l'apprentissage profond, des ordinateurs peuvent de plus en plus facilement déjouer de tels tests.

**Chaines de blocs** : Souvent désigné sous son vocable anglais *blockchain*, il s'agit de protocoles permettant d'avoir une base de données sécurisée et distribuée à travers un réseau. De par sa nature distribuée, ce type de base de données peut être très difficilement corrompu ou falsifié. Ce type de protocole est notamment utilisé pour les crypto monnaies, mais aussi pour certains contrats intelligents, qui sont des protocoles informatiques permettant d'automatiser l'exécution de certaines obligations contractuelles.

**Chatbot** : Un chatbot (à prononcer « tchate-botte »), appelé parfois assistant virtuel, est un programme informatique qui simule une conversation (« chat » en anglais) avec une personne, à l'écrit ou à l'oral. Quand un utilisateur lui pose une question ou formule une commande, le chatbot lui répondra ou exécutera l'action demandée.

Ce sont en quelque sorte les porte-parole de l'intelligence artificielle (IA). Cette forme accessible d'IA est souvent mise en œuvre par les entreprises dans leurs services SAV et d'assistance. On les trouve également dans les secteurs des appareils et des applications grand public.

## *Comment fonctionnent les chatbots ?*

Les chatbots comme ELIZA et PARRY figurent parmi les premières tentatives de créer des programmes capables de tromper une personne, ne serait-ce que temporairement, et lui faire croire qu'elle tient une conversation avec quelqu'un. L'efficacité de PARRY a été mesurée au début des années 1970 à l'aide d'une version du test de Turing : l'attribution correcte par les testeurs des interactions à quelqu'un ou à un chatbot correspond aux résultats que donneraient des réponses aléatoires.

Les chatbots se sont beaucoup améliorés depuis. Ils exploitent les technologies d'IA notamment le Deep Learning, le traitement automatique du langage et les algorithmes d'apprentissage automatique (Machine Learning). Les plus généralistes ont besoin d'énormes volumes de données pour apprendre une langue. Plus un utilisateur interagit avec la machine, plus la reconnaissance de la parole ou du texte améliore sa réponse.

Les chatbots peuvent être « stateless » ou « statefull ». Dans le premier cas, le chatbot aborde chaque interaction comme s'il s'agissait d'un nouvel utilisateur. Dans le deuxième, le chatbot est plus sophistiqué : les interactions passées lui servent à contextualiser ses prochaines réponses.

Aujourd'hui, une entreprise qui dote ses services d'un chatbot n'a que peu de code à écrire. En effet, plusieurs prestataires permettent aux développeurs de créer des interfaces de conversation quasiment clés en mains pour tout type d'application.

Les applications de chatbot mobiles comme Siri d'Apple ou Cortana de Microsoft sont plutôt qualifiées d'assistants virtuels.

**Cloud computing** : Le Cloud computing est un concept général qui désigne la mise à disposition de services hébergés sur Internet. Le Cloud computing permet aux entreprises de consommer les ressources informatiques à la demande (comme elle le ferait d'un service public tel que l'électricité), en leur évitant de créer et de gérer des infrastructures en interne.

Plusieurs avantages pour les entreprises et les utilisateurs découlent du Cloud. Trois de ces avantages sont les suivants :

- Mise à disposition en libre-service : Les utilisateurs peuvent se procurer des ressources informatiques à la demande pour presque tous les types de workloads.
- Elasticité : Les entreprises peuvent monter en puissance lorsque leurs besoins augmentent, puis de nouveau diminuer leur capacité de traitement lorsque la demande baisse.
- Paiement à l'usage : Les ressources informatiques sont évaluées de façon très fine, les utilisateurs ne payant que les ressources qu'ils utilisent.

Les services de Cloud computing peuvent être privés, publics ou hybrides.

Les services de Cloud privé sont fournis par le datacenter d'une entreprise à ses utilisateurs internes. Ce modèle offre souplesse et confort d'emploi, tout en garantissant la gestion, le contrôle et la sécurité. Les services peuvent ou non être facturés aux clients internes par le biais d'une refacturation.

Dans le modèle de Cloud public, c'est un prestataire tiers qui fournit le service de Cloud sur Internet. Les services de Cloud public sont facturés à la demande, généralement à la minute ou à l'heure. Les clients ne payent que les cycles de CPU, le stockage ou la bande passante qu'ils consomment. Les principaux fournisseurs de Cloud public sont Amazon Web Services (AWS), Microsoft Azure, IBM et Google.

**Data Scientist** : Le terme Data Scientist est un titre professionnel qui désigne un employé ou un consultant en informatique décisionnelle (BI, Business Intelligence) qui excelle dans l'analyse des données, particulièrement des grandes quantités de données, et ce afin d'aider une entreprise à en tirer un avantage concurrentiel.

Le titre de Data Scientist est parfois déprécié comme n'étant pas assez spécifique. Il peut être perçu comme un synonyme d'analyste des données, un cran plus haut.

Quoi qu'il en soit, le poste est de plus en plus reconnu dans les grandes entreprises qui s'intéressent à l'interprétation du Big Data, défini comme la vaste quantité de données structurées, semi-structurées et non structurées qu'elles produisent.

Un Data Scientist maîtrise un ensemble de compétences incluant l'analytique, l'apprentissage automatique (Machine Learning), l'exploration de données (data mining) et l'analyse statistique. Il dispose parallèlement d'une expérience des algorithmes et du codage.

Mais sa compétence la plus importante n'en demeure pas moins sa capacité à expliquer la signification des données dans un langage facilement compréhensible par tous.

**Deep learning (réseau neuronal profond) :** L'apprentissage profond, ou « deep learning », est un aspect de l'intelligence artificielle (IA) qui imite la méthode d'apprentissage qu'utilisent les êtres humains pour acquérir certains types de connaissances. Sous sa forme la plus simple, le deep learning peut être considéré comme un moyen d'automatiser l'analytique prédictive.

Alors que les algorithmes traditionnels de l'apprentissage automatique sont linéaires, ceux du deep learning sont empilés dans une architecture d'une complexité et d'une abstraction croissantes.

Pour comprendre le deep learning, imaginez un jeune enfant dont le premier mot est « chien ». Il apprend ce qu'est un chien (et ce qui n'est pas un chien) en pointant des objets du doigt et en disant le mot « chien ». Ses parents lui répondent « Oui, c'est un chien » ou « Non, ce n'est pas un chien ».

A mesure que l'enfant continue de montrer des objets, il comprend de mieux en mieux quelles sont les caractéristiques que possèdent tous les chiens. Ce qu'il fait, sans le savoir, c'est clarifier une abstraction complexe (le concept de chien) en construisant une architecture dans laquelle chaque niveau d'abstraction est créé avec les connaissances acquises au niveau précédent.

Les programmes informatiques qui utilisent le deep learning suivent en grande partie le même processus. Chacun des algorithmes de l'architecture applique une transformation non linéaire aux données en entrée et utilise ce qu'il apprend pour créer un modèle statistique en sortie. Les itérations se poursuivent jusqu'à ce que la sortie ait atteint un niveau de précision acceptable. Le qualificatif « deep » (profond) s'explique par le nombre de couches de traitement par lesquelles les données doivent passer.

Dans l'apprentissage automatique classique, le processus d'apprentissage est supervisé et le programmeur doit être très, très explicite lorsqu'il indique à l'ordinateur le type d'éléments qu'il doit rechercher pour déterminer si une image représente un chien ou pas. C'est un processus laborieux appelé extraction de caractéristiques, et le taux de réussite de l'ordinateur dépend entièrement de la capacité du programmeur à définir de manière précise le jeu de caractéristiques d'un « chien ». L'avantage du deep learning est que le programme crée lui-même le jeu de caractéristiques sans supervision. C'est non seulement plus rapide, mais c'est généralement plus précis.

Au départ, le programme informatique peut être alimenté par des données d'apprentissage, par exemple une série d'images à chacune desquelles un humain aura associé la mention « chien » ou « pas chien » à l'aide de métabalises. Il utilise les informations fournies par ces données pour créer le jeu de caractéristiques du chien et élaborer un modèle prédictif. Dans ce cas, le modèle que l'ordinateur crée en premier peut estimer que dans une image, chaque élément doté de quatre pattes et d'une queue doit être étiqueté comme étant un « chien ». Bien sûr, le programme ne connaît pas les étiquettes « quatre pattes » ou « queue », il recherche simplement des motifs de pixels dans les données numériques. A chaque itération, le modèle prédictif que crée l'ordinateur devient plus complexe et précis.

Parce que ce processus imite la pensée humaine, le deep learning est parfois appelé apprentissage neuronal profond ou réseau neuronal profond. A la différence du petit enfant, qui mettra des semaines, voire des mois, à comprendre le concept de « chien », un programme informatique qui utilise des algorithmes de deep learning peut, après avoir reçu un jeu de données d'apprentissage,

trier des millions d'images, en identifiant avec précision et en quelques minutes celles qui représentent des chiens.

Pour que le niveau de précision soit acceptable, les programmes de deep learning doivent avoir accès à des quantités phénoménales de données d'apprentissage et de puissance de traitement, deux conditions qui étaient difficiles à remplir pour les programmeurs avant l'avènement du Big Data et du Cloud computing. Les programmes de deep learning étant capables de créer des modèles statistiques complexes directement depuis leurs propres sorties itératives, ils peuvent élaborer des modèles prédictifs précis à partir de grandes quantités de données non structurées et sans étiquette. Cette capacité est importante dans le contexte de la propagation de l'Internet des objets (IoT, Internet of Things), car la plupart des données que les humains et les machines créent ne sont ni structurées, ni étiquetées. Les cas d'utilisation du deep learning sont tous les types d'applications d'analytique du Big Data, en particulier celles qui sont axées sur le traitement du langage naturel, la traduction, le diagnostic médical, les signaux boursiers, la sécurité des réseaux et l'identification des images.

**Exploration de données :** Également appelée « analyse de données » et souvent désignée en anglais par data mining, l'exploration des données vise à fouiller de grandes quantités de données de manière automatisée pour en extraire des renseignements précis. Ces renseignements sont généralement utiles pour construire des modèles, souvent par des méthodes non supervisées. Il s'agit souvent d'un moyen, au service de l'entreprise, d'extraire l'information pertinente sur la clientèle, le marché et les opérations de celle-ci.

**Heuristiques :** Se dit d'une méthode non fondée sur un modèle formel et qui fournit un résultat rapide bien que non optimal.

**Intelligence artificielle (IA, AI) :** L'intelligence artificielle ou IA (Artificial Intelligence -AI- en anglais) vise à permettre à des machines, et plus particulièrement à des systèmes informatiques, de simuler les processus cognitifs humains.

Ces processus comprennent l'apprentissage (acquisition d'informations et de règles liées à leur utilisation), le raisonnement (application des règles pour parvenir à des conclusions approximatives ou précises) et l'autocorrection. Les applications spécifiques de l'IA sont notamment les systèmes experts, la reconnaissance vocale et la vision artificielle.

On doit la première occurrence du terme d'IA à John McCarthy, chercheur américain en informatique, à la Dartmouth Conference de 1956 qui vit naître la discipline. Aujourd'hui, le terme recouvre aussi bien l'automatisation robotisée des processus (Robotic Process Automation, RPA) que la robotique proprement dite. Il a récemment gagné en visibilité en partie à cause des Big Data, c'est-à-dire de la vitesse, du volume et de la diversité des données collectées par les entreprises. Plus apte que l'homme à faire ressortir des tendances des données, l'IA permet aux entreprises d'exploiter un maximum d'informations.

### ***Types d'intelligence artificielle :***

Il y a plusieurs façons de classer les systèmes d'IA. En voici deux exemples.

Le premier est que l'intelligence artificielle peut être considérée comme faible ou forte.

L'IA faible est un système d'intelligence artificielle conçu pour reproduire une tâche précise à laquelle il est formé. Les assistants personnels virtuels comme Siri d'Apple en sont une forme. L'IA forte, dite aussi intelligence artificielle générale, est un système doté de capacités cognitives humaines générales qui, lorsqu'on lui présente une tâche inhabituelle, est assez intelligent pour trouver une

solution. Bien que controversé, le test de Turing, élaboré par le mathématicien du même nom en 1950, est une méthode qui sert à déterminer si un ordinateur pense comme un humain.

Le deuxième exemple nous vient d'Arend Hintze, professeur en biologie intégrative et ingénierie informatique à la Michigan State University. Il classe l'IA en quatre types allant de celui des systèmes actuels aux systèmes sensibles à venir. Ses catégories sont :

- Type 1 : machines réactives. Chacun se souvient de Deep Blue, le programme d'IBM qui a battu Garry Kasparov aux échecs dans les années 1990. Deep Blue identifie les pièces sur l'échiquier et émet des prédictions, mais il n'a aucune mémoire et n'utilise pas ses expériences passées pour formuler les prédictions futures. Il analyse les coups possibles, les siens et ceux de son adversaire, et choisit le plus stratégique. Deep Blue et AlphaGO de Google ont été conçus à des fins précises : ils ne sont pas facilement transposables à une autre situation.

- Type 2 : machines à mémoire restreinte. Ces systèmes d'IA s'appuient sur leurs expériences passées pour prendre les décisions suivantes. Certaines fonctions décisionnelles des véhicules autonomes suivent ce modèle. Les observations servent à contextualiser les actions du futur proche, par exemple une voiture qui change de file. Ces observations ne sont pas stockées de manière permanente.

- Type 3 : théorie de l'esprit. Il s'agit d'un concept de psychologie qui se rapporte à la compréhension des gens en tant qu'êtres ayant des pensées, désirs et raisons propres qui les poussent à prendre leurs décisions. Ce type d'IA n'existe pas encore.

- Type 4 : conscience de soi. Dans cette catégorie, les systèmes d'IA ont une identité, une conscience. Ces machines douées de conscience connaissent leur état actuel et utilisent ces informations pour inférer ce que les autres ressentent. Ce type d'IA n'existe pas encore.

**IA explicable (XAI) :** L'IA explicable ou XAI (eXplainable Artificial Intelligence) est une forme d'intelligence artificielle prévue pour décrire son objet, sa logique et sa prise de décision de manière intelligible à une personne lambda. Souvent évoquée de pair avec l'apprentissage profond ou « deep learning », elle joue un rôle primordial dans le modèle d'apprentissage automatique non discriminatoire, redevable et transparent dit FAT ML (Fairness, Accountability and Transparency in Machine Learning).

La XAI renseigne globalement sur la prise de décision d'un programme d'intelligence artificielle en dévoilant :

- Les points forts et les faiblesses
- Les critères précis retenus pour arriver à une décision
- Les motifs qui l'ont conduit à telle décision plutôt qu'aux autres
- Le niveau de confiance approprié selon les différents types de décision
- Les types d'erreur qu'il est susceptible de commettre
- La façon de corriger les erreurs

Un des grands objectifs de la XAI est la redevabilité / responsabilité algorithmique. Jusqu'à maintenant, les systèmes d'IA ont été par essence des boîtes noires. Si on en connaît les données en entrée et en sortie, les algorithmes qui mènent à une décision sont généralement propriétaires ou peu intelligibles, même quand les mécanismes de logique internes sont accessibles gratuitement en open source.

Étant donné que l'intelligence artificielle est de plus en plus répandue, il est plus que jamais important de savoir comment traiter les distorsions et la question de la confiance. Notons par exemple qu'une des clauses du règlement général sur la protection des données (RGPD) de l'UE instaure le droit à l'explication.

**Machine Learning** : L'apprentissage statistique (Machine Learning) est un type d'intelligence artificielle qui confère aux ordinateurs la capacité d'apprendre sans être explicitement programmés. Cette technologie s'appuie sur le développement de programmes informatiques capables d'acquérir de nouvelles connaissances afin de s'améliorer et d'évoluer d'eux-mêmes dès qu'ils sont exposés à de nouvelles données.

Le processus d'apprentissage automatique s'apparente à celui de l'exploration de données (data mining). En effet, il s'agit, dans les deux cas, d'analyser les données à la recherche de schémas récurrents. Cependant, au lieu d'extraire les données pour les soumettre à un traitement humain (comme c'est le cas dans les applications de data mining), l'apprentissage automatique utilise ces données pour améliorer la compréhension du programme lui-même.

Les programmes d'apprentissage automatique détectent des schémas dans les données et ajustent leur fonctionnement en conséquence. Par exemple, le fil d'actualité de Facebook change en fonction des interactions personnelles de l'utilisateur avec ses homologues. Ainsi, si un utilisateur identifie fréquemment un ami dans des photos, écrit sur son mur ou « aime » ses liens, le fil d'actualité montrera un plus grand nombre d'activités de cet ami, car il supposera qu'il s'agit d'un ami proche de l'utilisateur.

**Modèle de Markov caché** : Le mathématicien Andreï Markov n'avait rien de particulier à cacher ! Il s'agit en fait d'un modèle statistique qui ne prend en compte que seuls certains résultats observables connus par l'utilisateur, mais les états du processus qui causent ces résultats sont inconnus, donc « cachés ». Ce modèle est très largement utilisé en intelligence artificielle, notamment pour l'apprentissage par renforcement.

**Perceptron** : Développé dans les années 50, il s'agit de la première forme de réseaux de neurones artificiels, et probablement l'une des plus simples. En effet, les réseaux de neurones ne datent pas d'hier ! Le perceptron n'est que le premier d'une longue série de types d'algorithmes utilisés.

**Réseau de neurones artificiels (RNA)** : Dans le domaine des technologies de l'information, un réseau de neurones est un système logiciel et / ou matériel qui imite le fonctionnement des neurones biologiques. Les réseaux neuronaux, aussi appelés réseaux de neurones artificiels (RNA ou ANN en anglais), font partie des technologies d'apprentissage profond (ou « deep learning »), couvertes également par l'intelligence artificielle (IA).

Les applications commerciales sont souvent axées sur la résolution de problèmes complexes de traitement de signaux ou de reconnaissance de modèles. Parmi les exemples les plus connus depuis 2000, on trouve la reconnaissance de l'écriture manuscrite pour le traitement des chèques, la conversion de parole en texte, l'analyse des données d'exploration pétrolière, les prévisions météorologiques et la reconnaissance faciale.

**Rétropropagation** : La « rétropropagation du gradient », de son nom complet, est une méthode qui permet aux réseaux de neurones d'apprendre. Plus spécifiquement, si l'on soumet des données à un réseau de neurones et qu'on lui indique les résultats désirables et indésirables, cette méthode permet d'augmenter ou de réduire le poids relatif à accorder aux neurones correspondant aux résultats désirables. C'est donc une méthode qui permet à un réseau de neurones d'apprendre à distinguer une bonne d'une mauvaise réponse.

**Stochastique** : Réfère à des méthodes mathématiques où certaines variables sont aléatoires. Ces méthodes sont utilisées pour la compréhension de phénomènes naturels, mais aussi de phénomènes socio-économiques tels les marchés boursiers.

**Systèmes experts** : Logiciel qui permet d'appliquer des règles préétablies, notamment pour l'aide à la prise de décision. Souvent, ces règles sont statiques, contrairement aux algorithmes d'apprentissage, et peuvent être schématisées sous forme d'arbres décisionnels, où les réponses à certaines questions amènent le système à poser des questions de plus en plus précises. Les systèmes experts sont assez répandus notamment dans le domaine médical (aide au diagnostic).

**Test de Turing** : Décrit dès 1950 par Alan Turing, grand pionnier de l'informatique, ce type de test vise à déterminer si une machine a un comportement intelligent. Pour réussir ce test, un ordinateur doit réussir à tromper un évaluateur humain de manière à ce que celui-ci ne se rende pas compte qu'il converse avec un ordinateur. Si des ordinateurs commencent aujourd'hui à tromper des humains dans des cadres particuliers, tels les agents conversationnels (*chatbots* ou *chat bots*), ils n'en sont pas encore capables dans un cadre plus général où une conversation peut porter sur divers sujets.

**Traitement du langage naturel (TLN)** : Le traitement du langage naturel (TLN, ou NLP en anglais) est la capacité pour un programme informatique de comprendre le langage humain tel qu'il est parlé. Il fait partie des technologies d'intelligence artificielle.

Le développement d'applications TLN est difficile parce que traditionnellement les ordinateurs sont conçus pour que les humains leur « parlent » dans un langage de programmation précis, sans ambiguïté et extrêmement structuré, ou à l'aide d'un nombre limité de commandes vocales clairement énoncées. Or le discours humain n'est pas toujours précis, il est souvent ambigu et sa structure linguistique peut dépendre d'un grand nombre de variables complexes, notamment l'argot, les dialectes régionaux et le contexte social.



---

## Lexiques Intelligence artificielle selon le journal officiel Français

Ce lexique est basé sur le *JOURNAL OFFICIEL DE LA RÉPUBLIQUE FRANÇAISE* de la rubrique « Vocabulaire de l'intelligence artificielle (liste de termes, expressions et définitions adoptés) » de l'année 2019.

### **Apprentissage automatique :**

Synonyme : apprentissage machine

En anglais : machine learning (ML).

*Définition:* Processus par lequel un algorithme évalue et améliore ses performances sans l'intervention d'un programmeur, en répétant son exécution sur des jeux de données jusqu'à obtenir, de manière régulière, des résultats pertinents.

*Note:* 1. Un algorithme d'apprentissage automatique comporte un modèle dont il modifie les paramètres, de valeur initiale en général aléatoire, en fonction du résultat constaté. 2. L'apprentissage automatique relève de l'intelligence artificielle. 3. L'apprentissage automatique est fréquemment utilisé pour le traitement du langage naturel et la vision par ordinateur, ou pour effectuer des diagnostics et des prévisions. *Voir aussi:* apprentissage non supervisé, apprentissage par renforcement, apprentissage profond, apprentissage supervisé, dialogueur, intelligence artificielle, réseau de neurones artificiels. *Équivalent étranger:* machine learning (ML).

### **Apprentissage non supervisé :**

En anglais : data clustering, unsupervised learning, unsupervised training.

*Définition:* Apprentissage automatique dans lequel l'algorithme utilise un jeu de données brutes et obtient un résultat en se fondant sur la détection de similarités entre certaines de ces données. *Note:* L'apprentissage non supervisé est utilisé, par exemple, pour l'identification de comportements et la recommandation d'achats.

### **Apprentissage par renforcement :**

En anglais : reinforcement learning.

*Définition:* Apprentissage automatique dans lequel un programme extérieur évalue positivement ou négativement les résultats successifs de l'algorithme, l'accumulation des résultats permettant à l'algorithme d'améliorer ses performances jusqu'à ce qu'il atteigne un objectif préalablement fixé.

*Note:* 1. L'apprentissage par renforcement est fréquemment utilisé dans la robotique. 2. L'efficacité de l'apprentissage par renforcement a été attestée dans certains jeux stratégiques comme le jeu de go.

### **Apprentissage profond :**

En anglais : deep learning, deep structured learning, hierarchical learning.

*Définition:* Apprentissage automatique qui utilise un réseau de neurones artificiels composé d'un grand nombre de couches dont chacune correspond à un niveau croissant de complexité dans le traitement et l'interprétation des données.

*Note*: L'apprentissage profond est notamment utilisé dans la détection automatique d'objets au sein d'images et dans la traduction automatique.

### **Apprentissage supervisé :**

En anglais : supervised learning, supervised training.

*Définition* : Apprentissage automatique dans lequel l'algorithme s'entraîne à une tâche déterminée en utilisant un jeu de données assorties chacune d'une annotation indiquant le résultat attendu.

*Note* : 1. L'apprentissage supervisé recourt le plus souvent aux réseaux de neurones artificiels. 2. L'apprentissage supervisé est utilisé, par exemple, pour la reconnaissance d'images et la traduction automatique.

### **Dialogueur :**

*Synonyme* : agent de dialogue

En anglais : chatbot, conversational agent.

*Définition* : Logiciel spécialisé dans le dialogue en langage naturel avec un humain, qui est capable notamment de répondre à des questions ou de déclencher l'exécution de tâches.

*Note* : 1. Un dialogueur peut être intégré à un terminal ou à un objet connecté. 2. Les dialogueurs sont utilisés, par exemple, dans les techniques de vente, les moteurs de recherche et la domotique. 3. On trouve aussi l'expression « agent conversationnel », qui est déconseillée.

### **Intelligence artificielle : IA**

En anglais : artificial intelligence (AI).

*Définition* : Champ interdisciplinaire théorique et pratique qui a pour objet la compréhension de mécanismes de la cognition et de la réflexion, et leur imitation par un dispositif matériel et logiciel, à des fins d'assistance ou de substitution à des activités humaines.

### **Neurone artificiel :**

*Synonyme* : neurone formel.

En anglais : artificial neuron, artificial neurone, formal neuron, formal neurone.

*Définition* : Dispositif à plusieurs entrées et une sortie, qui simule certaines propriétés du neurone biologique.

*Note* : La valeur de sortie du neurone artificiel est une fonction non linéaire, généralement à seuil, d'une combinaison de valeurs d'entrée dont les paramètres sont ajustables.

### **Réseau de neurones artificiels :**

En anglais : artificial neural network, neural network.

*Synonyme* : réseau de neurones formels, réseau de neurones.

*Définition* : Ensemble de neurones artificiels interconnectés qui constitue une architecture de calcul.



---

## Lexiques Agilité

**Backlog (product ou iteration ou sprint) :** Liste d'exigences priorisées et grossièrement estimées au niveau du produit (product backlog) ou sous-ensemble de cette liste pour l'itération en cours (iteration ou sprint backlog).

**Burndown chart :** Représentation graphique (sous forme de courbe) du travail restant à faire, actualisé chaque jour de l'itération.

**Chemin critique :** Sur un diagramme de réseau, c'est le chemin le plus long dans l'enchaînement des activités. Tout retard sur une activité positionnée sur le chemin critique induit automatiquement un retard sur le projet.

**CMMI (Capability Maturity Model Integration) :** Modèle, mis au point par le SEI (Software Engineering Institute), qui mesure la maturité des organisations pour le développement logiciel, matériel ou système ; c'est un recueil de bonnes pratiques, organisées en cinq niveaux de maturité et vingt-deux domaines de processus.

**Cycle de vie :** C'est « un ensemble, généralement séquentiel, de phases de projet, dont le nom et le nombre sont déterminés en fonction des besoins de suivi par l'(es) organisation(s) impliquée(s) dans le projet.

**Diagramme de Gantt :** Représentation graphique permettant de visualiser l'ordonnancement (enchaînement) et la planification (positionnement dans le temps) des activités nécessaires à la réalisation d'un projet.

**Diagramme de réseau :** Représentation graphique des dépendances entre les activités d'un projet. L'objectif est de trouver le meilleur ordonnancement pour que les ressources soient optimisées (délai, disponibilité, compétences, coût, faisabilité...).

**« Done » :** Etat d'une exigence ou d'une fonctionnalité lorsqu'elle est globalement acceptée par toutes les parties prenantes et lorsqu'elle satisfait aux attentes du client : développée, testée, documentée, validée et potentiellement mise en production.

**Feedback :** L'une des valeurs essentielles des méthodes agiles : recueillir, le plus fréquemment possible, les appréciations, les impressions, les remarques, le ressenti, des acteurs du projet sur le produit, sur les conditions de réalisation, sur la méthodologie, sur les relations au sein de l'équipe...

On l'obtient grâce à la collaboration.

**Incrément :** La ou les fonctionnalité(s) supplémentaire(s) développée(s) et livrée(s) par l'équipe à chaque itération. C'est l'apport de valeur pour le client à chaque itération.

**Intégration continue :** Action visant à consolider tous les développements, le plus fréquemment possible (tous les jours, toutes les heures...), afin d'obtenir une application testable dans son ensemble.

**Itération** : Espace de temps répétitif – une boucle –, de quelques semaines, au cours duquel une ou plusieurs fonctionnalité(s) est(sont) développée(s) ; tous les acteurs y sont impliqués. L'utilisateur visualise un résultat tangible. Il restitue son feedback.

**Maître d'œuvre** : Entité qui réalise « l'ouvrage » ; elle détermine, avec la maîtrise d'ouvrage, les conditions de la réalisation, les délais et le budget. Elle est responsable des choix techniques et du bon déroulement du projet.

**Maîtrise d'ouvrage** : Entité qui achète « l'ouvrage » ; elle porte la vision, définit l'objectif et a la propriété de la liste d'exigences et de leur priorisation. Elle définit, avec la maîtrise d'œuvre, le calendrier et le budget consacrés.

**Offshore** : Pratique qui consiste, pour une entreprise, à externaliser, c'est-à-dire confier, tout ou partie de ses développements à une autre entreprise, le plus souvent localisée, dans des pays étrangers. L'idée d'origine est de réduire les coûts de réalisation grâce à une main d'œuvre moins chère.

**Pair-programming** : L'une des pratiques phares de l'eXtreme Programming, qui consiste à faire travailler les développeurs en binôme, sur un même poste de travail.

**Planning game** : Réunion de planifications à laquelle participent le client et l'équipe de réalisation ; le client décrit les fonctionnalités, puis l'équipe estime le coût d'implémentation en listant et en évaluant les activités nécessaires. Cette réunion a lieu avant chaque itération.

**PMI (Project Management Institute)** : Organisation internationale de standardisation des pratiques en gestion de projet.

**Product backlog item** : Eléments qui constituent le product backlog : cas d'utilisation, user stories, exigences....

Ils sont priorisés et estimés.

**Product owner** : Représentation de la maîtrise d'ouvrage, du client et des utilisateurs, qui a la responsabilité du product backlog : recensement des items, priorisation en fonction de l'apport de valeur ajoutée, ajout/suppression ou modification d'items.

**Refactoring** : Opération de maintenance régulière du code, pour en améliorer la qualité, la lisibilité et la simplicité, qui se déroule tout au long du projet.

**Reste à faire** : Estimation de la charge nécessaire restant à consommer pour l'achèvement d'une activité.

**Rétrospective** : Réunion menée à chaque fin d'itération au cours de laquelle un bilan est dressé ; l'équipe s'interroge sur la façon dont elle a travaillé et sur les adaptations éventuellement nécessaires à apporter pour améliorer le processus.

**Roadmap** : Présentation des dates de sortie des versions majeures d'un produit.

**Scrum** : Réunion quotidienne de quinze minutes, la « mêlée », au cours de laquelle l'équipe, le ScrumMaster et le product owner font le point sur ce qui a été fait la veille, ce qui va être fait le jour même et les difficultés rencontrées par chacun.

**ScrumMaster** : Personne chargée de veiller à la bonne application des pratiques de Scrum, de « protéger » l'équipe des éléments perturbateurs externes et de lever tous les obstacles rencontrés dans la réalisation.

**Sprint** : Itération de trente jours dont les résultats est un ensemble de fonctionnalités validées par le client.

**Timeboxing** : Principe consistant à définir une échéance fixe pour développer et livrer un ensemble de fonctionnalités. Si le travail planifié n'est pas achevé, on ne décale pas la date de fin, mais on analyse les raisons expliquant ce retard pour apporter rapidement des adaptations.

**Team building** : Ensemble d'actions visant à renforcer l'esprit d'équipe et à développer la collaboration au sein d'une équipe, et entre celle-ci et ses interlocuteurs extérieurs.

**Vélocité** : Somme des points correspondant aux fonctionnalités réellement développées au terme d'une itération ; elle détermine la capacité de l'équipe. Cette vélocité est utile pour la planification des itérations suivantes, et est ajustée en fonction des résultats constatées.



---

## Lexiques modélisation UML

**Acteur métier (Business Actor) :** Un acteur métier représente un ensemble de personnes jouant un même rôle en dehors de système métier et en interaction avec lui. En d'autres termes, il s'agit d'un rôle externe impliqué dans le métier.

**Action :** Lorsqu'il s'agit de décrire un comportement, les actions sont centrales. Ils représentent une seule étape dans une activité - une activité, à son tour, est un comportement qui se compose d'un ensemble d'actions. Voici quelques exemples d'actions :

- **Ordre de remplissage**
- **Afficher la page d'erreur**
- **Ordre de process**

**Cas d'application :** Le cas d'application (en tant qu'unité linguistique) représente la configuration système requise. L'acteur (une personne ou un système) est un élément qui détermine qui ou quoi doit effectuer une activité particulière à travers le système. Le système peut également être une classe ou un composant et est donc décrit comme un sujet. Le cas d'utilisation (en tant qu'élément de modèle) informe seulement qu'un comportement nommé est attendu qui est visible par le monde extérieur. Il ne montre généralement pas les actions exactes. Dans une description comportementale, la modélisation assigne les exigences détaillées au cas d'application.

**Classe :** En tant qu'unité de langage, les classes sont un aspect central d'UML. Vous définissez ce qu'est une classe et comment les classes interagissent entre elles. Cette unité linguistique comporte quatre niveaux, allant d'éléments simples à des relations plus complexes :

- Core (describes elements from the UML 2.0 infrastructure such as package, namespace, attribute, etc.)
- Association classes
- Interfaces
- Powertypes (classes whose instances are subclasses within this class)

**Comportement :** L'unité de langage "comportement" ou "description du comportement" désigne la modélisation des aspects dynamiques d'un système. Il contient trois spécifications :

- **Activité :** Les actions interagissent à travers les flux de données et de contrôle. Il en résulte un système complexe de comportements - les activités.
- **Interaction :** Ce méta-modèle décrit comment les flux de messages sont échangés entre les objets, quand un message est envoyé et à quel objet, et quels autres éléments sont affectés par celui-ci.
- **Des machines à états :** Dans un diagramme d'états, ce méta-modèle montre les états (situations aux propriétés immuables) et les pseudo-états (états sans affectation de valeurs) ainsi que leurs transitions. Les objets d'un état peuvent être affectés à des actions ou à des activités

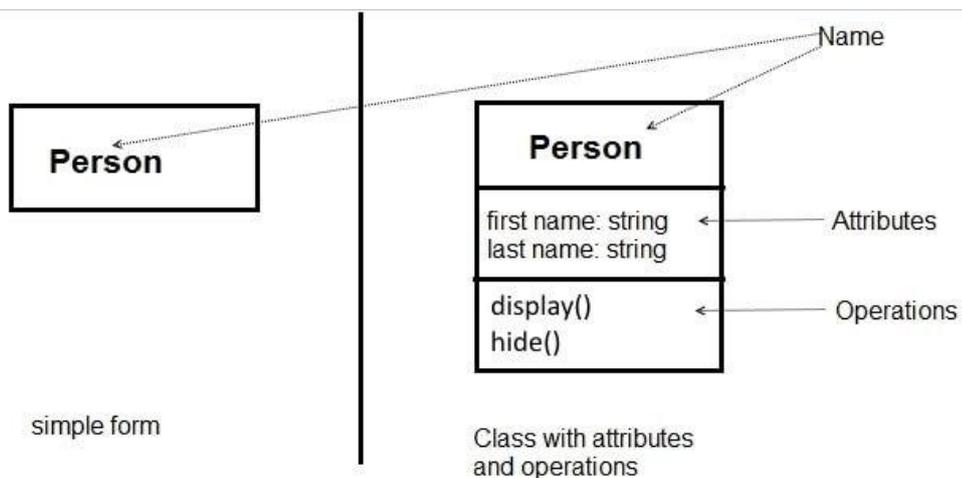
**Composants :** Les composants sont des modules logiciels qui séparent leur contenu du système externe. Il n'y a qu'une connexion vers l'extérieur par l'intermédiaire d'interfaces ou d'un port. Un connecteur de composition établit une connexion avec un autre composant via l'interface. Le

connecteur de délégation relie les éléments internes à une interface à la frontière extérieure. Les composants sont modulaires et interchangeables.

**Diagramme d'activité :** (activity diagram) Les activités consistent en un réseau d'actions reliées par des flux de données et de contrôle. Alors que le diagramme de cas d'utilisation montre la configuration système requise, le diagramme d'activité montre comment ces cas d'utilisation s'exécutent. Dans ce type de diagramme, les jetons jouent un rôle. Dans les processus parallèles, ils sont un marqueur pour lequel les processus sont priorisés et reçoivent des ressources (par exemple, la mémoire de travail).



**Diagramme de classes :** Si les objets ont un comportement commun ou la même structure, vous pouvez les classer ou les affecter à une classe. La classe est donc un élément de simplification et de synthèse (abstraction) pour la représentation visuelle. Les classes et les objets sont reliés entre eux à l'aide d'interfaces. Tous ces composants et leurs relations entre eux peuvent être représentés dans un diagramme de classes. Une classe représente ses diagrammes à l'aide d'un rectangle. Le nom de la classe est en caractères gras, comme indiqué ci-dessous :



**Diagramme de communication** : Semblable à un diagramme de séquence, les diagrammes de communication modélisent un transfert de message à l'aide de lignes de vie. Cependant, ce diagramme UML n'utilise pas de lignes pointillées pour les séquences de temps, mais numérote les séquences avec des chiffres et des lettres. Ces termes de séquence sont situés au-dessus d'une flèche dont l'extrémité pointe dans la direction du récepteur. Les chiffres représentent l'ordre dans lequel les messages sont envoyés, les lettres pour le niveau hiérarchique (comme indiqué dans la figure ci-dessous).

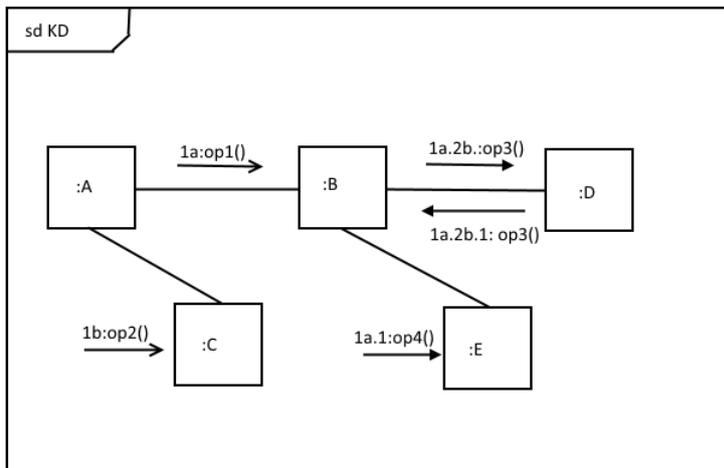
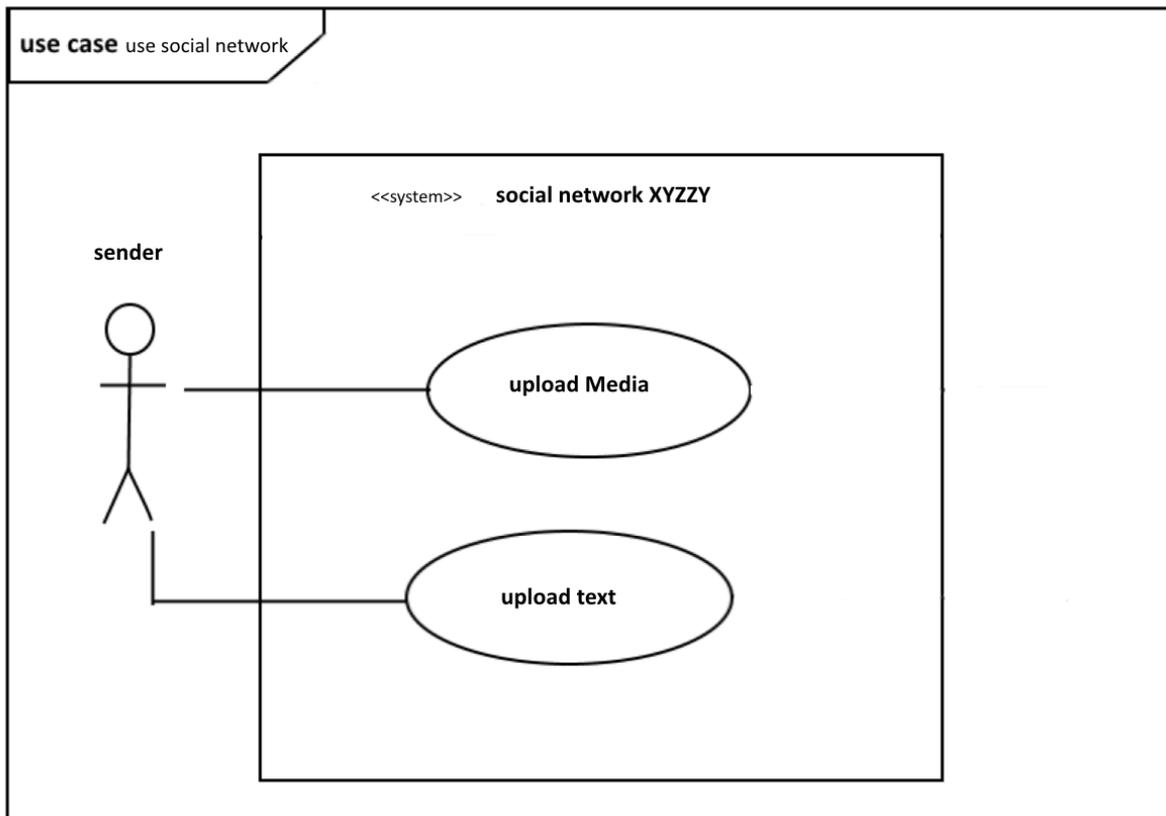


Figure 1 : Les messages asynchrones 1a et 1b sont envoyés en parallèle. Le message avec l'étiquette 1a.1 envoie l'objet :B avant le message 1a.2b. Objet :D envoie le message 1a.2b.1 seulement après réception du message 1a.2b

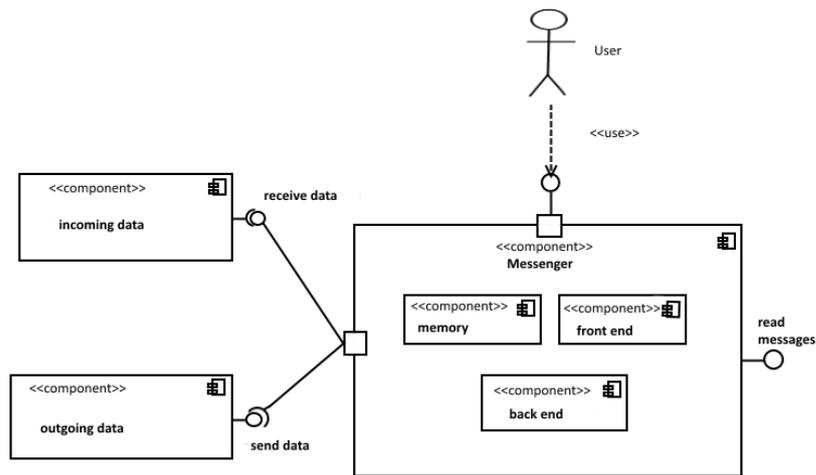
**Diagramme de cas d'utilisation** : Les diagrammes de cas d'utilisation montrent le comportement attendu d'un système ultérieurement. Cette modélisation convient non seulement pour les systèmes logiciels, mais aussi pour prédire les procédures dans l'entreprise, par exemple. Le cas d'utilisation implique un acteur (humain ou système) avec un but. Le diagramme porte généralement le nom de la cible. Les différents cas d'utilisation au sein du système répondent à l'objectif de l'acteur.



**Diagrammes de comportement :** Les diagrammes de comportement couvrent les autres spécifications sous UML. Contrairement aux diagrammes de structure, ils ne sont pas statiques, mais représentent des processus. Les diagrammes de comportement comprennent également des diagrammes d'interaction :

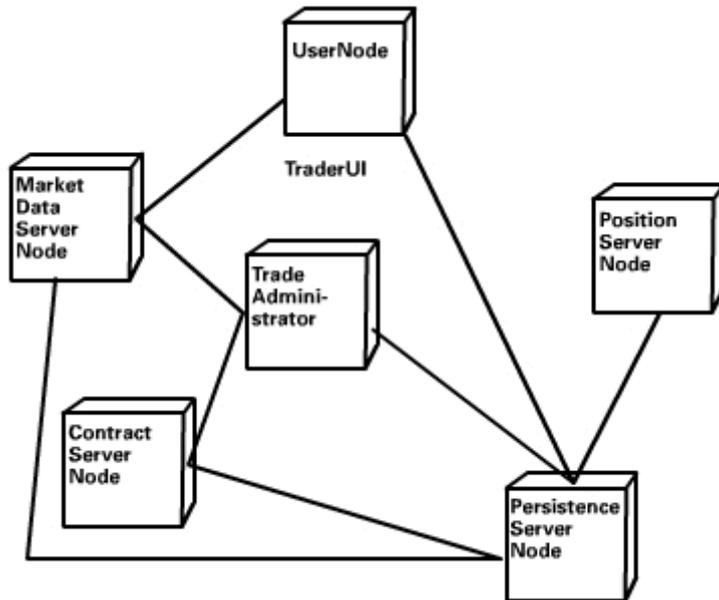
- Diagramme de cas d'utilisation
- Diagramme d'activité
- Diagramme d'état transition

**Diagramme des composants :** Un composant est un module isolé du système externe et interagit avec d'autres composants via des interfaces définies. C'est un sous-type de la classe. Par conséquent, les caractéristiques structurelles telles que les opérations et les attributs définissent plus clairement le composant. Il existe deux options d'affichage pour la modélisation, selon les besoins : la vue boîte noire (le contenu est masqué) et la vue boîte blanche (le contenu est visible).



**Diagramme d’aperçu des interactions :** Le diagramme de vue d’ensemble des interactions récemment ajouté dans UML 2.0 permet d’afficher un système très complexe dans un schéma approximatif quand un diagramme d’interaction normal serait trop confus. Un diagramme de séquence convient pour un affichage détaillé. Le diagramme UML est similaire au diagramme d’activités avec nœuds. Il représente les flux de contrôle entre les interactions. Ceci est différent d’un diagramme d’activités, car un diagramme d’interaction entier peut être imbriqué dans des nœuds qui représentent des activités. Ces imbrications peuvent être affichées directement dans le diagramme (en ligne) ou se référer au modèle (mot clé : réf) qui est montré en détail ailleurs.

**Diagramme de déploiement :** Un diagramme de déploiement modélise la distribution physique des artefacts sur les nœuds. Les nœuds sont soit du matériel (nœuds de périphérique) qui peuvent fournir de la mémoire, soit du logiciel (nœuds d’environnement d’exécution) qui fournit un environnement pour exécuter des processus. Ils sont représentés sous forme de cuboïdes tridimensionnels. Ceux-ci contiennent le nom du fichier. Pour les distinguer d’une classe, les cuboïdes ont des stéréotypes tels que <<artéfact>>. Le diagramme convient à l’affichage des dépendances entre les nœuds et les artefacts, appelées **relations de distribution**.



Les diagrammes de déploiement montrent la disposition physique des matériels qui composent le système et la répartition des composants sur ces matériels. Les ressources matérielles sont représentées sous forme de nœuds. Les nœuds sont connectés entre eux, à l'aide d'un support de communication. La nature des lignes de communication et leurs caractéristiques peuvent être précisées. Les diagrammes de déploiement peuvent montrer des instances de nœuds (un matériel précis), ou des classes de nœuds. Les diagrammes de déploiement correspondent à la vue de déploiement d'une architecture logicielle

**Diagramme d'objet :** Un diagramme d'objets a une structure similaire à celle d'un diagramme de classes. Au lieu du nom tel qu'il apparaît dans le diagramme de classes le **diagramme d'objets** met le nom avec le nom du classificateur/catégorie.

**Diagramme de profil :** Les diagrammes de profil sont utilisés au niveau du méta-modèle. Ils sont utilisés pour assigner un stéréotype à des classes, ou un profil à des packages. Au niveau méta, il est possible d'adapter le modèle à une autre plate-forme ou domaine. Par exemple, si vous limitez la sémantique UML dans un profil, elle transmet les spécifications aux classes subordonnées.

**Diagramme d'ensemble :** Un package combine des éléments tels que des interfaces ou des classes dans un espace de noms (voir note ci-dessous). Les paquets peuvent également fusionner avec d'autres paquets (fusion de paquets), les importer (importation de paquets), ou contenir d'autres paquets (sous-paquets). Les diagrammes de structure de paquet lient les contenus hiérarchiquement, comme dans un diagramme en arbre. Le diagramme de package est utilisé, par exemple, dans le méta-modèle d'UML 2, et est modulaire dans les systèmes logiciels. Strictement spécifié, un paquet se compose d'une tête et d'une zone de contenu.

**Diagramme d'état-transition :** Une machine à états, aussi appelée automate fini, représente un ensemble fini d'états dans un système. Si une condition fixe est remplie dans le système (c'est-à-dire qu'un déclencheur se déclenche), une réaction correspondante se produit. Il peut s'agir d'activités ou d'interactions. Sous UML 2.0, un état représente cette situation. Les états sont considérés comme

des sommets et sont affichés sous forme de rectangles aux coins arrondis. En outre, le diagramme machine d'états modélise les transitions d'un état (nœud source) à l'autre (nœud cible).

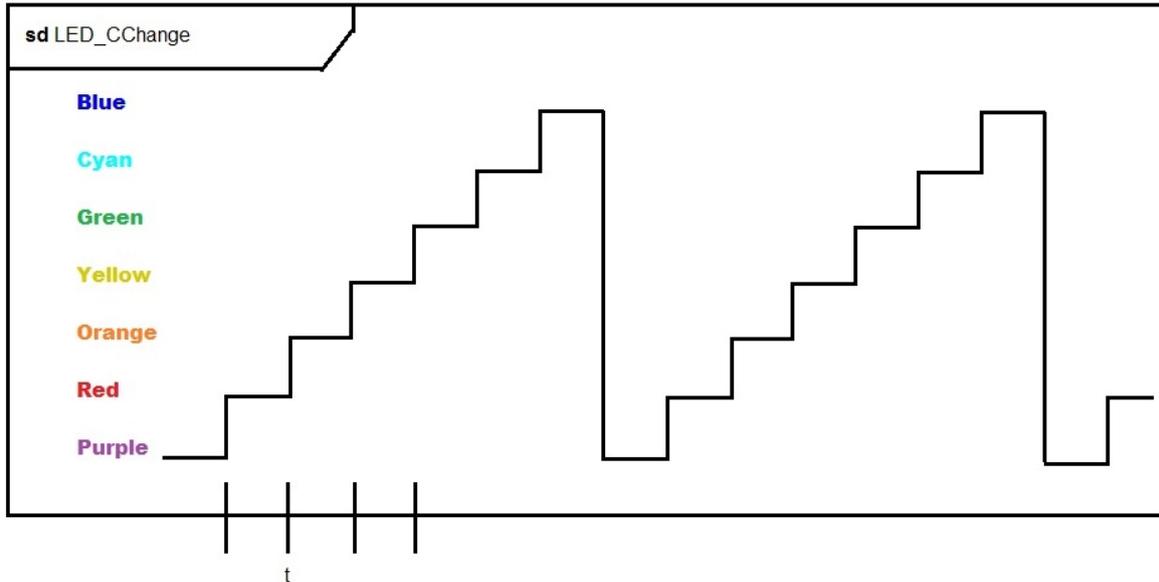
**Diagrammes d'interaction :** Les diagrammes d'interaction sont un **sous-type des diagrammes de comportement**. Ils décrivent également les processus. Ils sont particulièrement adaptés à la modélisation des comportements dans lesquels les éléments échangent des informations. Les diagrammes définissent le rôle des objets impliqués. Ils nomment et classent par ordre de priorité les messages qui sont envoyés d'un objet à l'autre. Les diagrammes d'interaction montrent également comment ces messages affectent les éléments comportementaux, tels que le démarrage ou l'arrêt des activités :

- Diagrammes de séquence
- Diagramme de communication
- Diagramme de temps
- Diagramme d'aperçu des interactions

**Diagrammes de séquence :** En tant que diagramme d'interaction, le diagramme de séquence représente l'échange de messages entre objets. Le type de diagramme UML modélise ces objets comme une ligne de vie. En ce sens, il est similaire à d'autres diagrammes comportementaux tels que le diagramme d'activité. Contrairement à ces derniers, cependant, un **diagramme de séquence n'est pas utilisé pour obtenir une vue d'ensemble** du comportement d'un système, mais pour présenter en détail un comportement possible parmi tant d'autres. Il prescrit une chronologie, et une ligne pointillée représente le cours du temps. UML 2.0 affiche des messages synchrones (une flèche à tête remplie) et asynchrones (une flèche à tête ouverte). Les messages synchrones sont des messages qui bloquent un canal jusqu'à ce qu'ils reçoivent une réponse de l'objet cible. Ils déterminent les caractéristiques comportementales sous forme d'**opérations synchrones**. Les messages asynchrones contrôlent l'objet source appelant. Il s'agit aussi bien d'opérations asynchrones que de signaux (paquets de données envoyés entre les actions).

**Diagramme de structure composite :** Les objets appartiennent à des classes qui, à leur tour, peuvent également être classifiées. Ces soi-disant méta-classes sont appelées **classificateurs** en UML. Le diagramme de structure composite représente les différents éléments et connecteurs d'un classificateur. Les pièces font toujours partie de l'ensemble, même si elles ne sont pas nécessairement nécessaires pour compléter le classificateur. Les connecteurs sont les liens entre les pièces. Les fonctions ou les services qui nécessitent des composants en dehors du classificateur envoient des pièces via une interface.

**Diagramme de temps :** Un diagramme de temps permet de montrer en détail le comportement des systèmes sous l'aspect du séquençage temporel. Les systèmes en temps réel, par exemple, doivent achever certains processus dans un certain laps de temps. Pour mieux représenter un plan de temps, UML 2.0 modélise les diagrammes de temps en diagrammes bidimensionnels, avec un axe des x et un axe des y. Dans cette sous-catégorie du diagramme de séquence, les états des objets se trouvent sur l'axe des x, et les séquences temporelles qui leur sont affectées se déroulent sur l'axe des y.



**Intervenant métier (Business Actor) :** Un intervenant métier représente un ensemble de personnes jouant un même rôle à l'intérieur du système métier et en interaction avec d'autres intervenants et des entités métier. En d'autres termes, il s'agit d'un rôle interne au métier.

**Types de Diagrammes= diagrams types :**

Catégorie	Type de diagramme	Utilité
Diagramme de structure	Diagramme de classes	Représente les classes
	Diagramme d'objet	Affiche l'état d'un système à un moment donné
	Diagramme des composants	Affiche les dépendances et les composants de structure
	diagramme de structure composite	Divise les modules ou les classes en leurs composantes et clarifie leurs relations.
	Diagramme de paquetage	Regroupe les classes en paquets, représente la hiérarchie et la structure des paquets
	Diagramme de déploiement	Affiche la distribution des composants aux nœuds de l'ordinateur
	Diagramme de profil	Illustre les relations d'usage au moyen de stéréotypes, de conditions aux limites, etc.
Diagrammes de comportement	Diagramme de cas d'utilisation	Représente divers usages
	Diagramme d'activité	Décrit le comportement de différents processus (parallèles) dans un système.
	Diagramme d'état-transition	Documente la façon dont un objet

		passer d'un état à un autre par le biais d'un événement.
Diagrammes de comportement - diagrammes d'interaction	Diagramme de séquence	Représente le moment des interactions entre les objets.
	Diagramme de communication	Affiche la répartition des rôles des objets au sein d'une interaction.
	Diagramme de temps	Démontre la limitation temporelle pour les événements qui mènent à un changement d'état.
	Diagramme d'aperçu d'interaction	Montre comment les séquences et les activités interagissent.

**Distribution** : Un réseau est constitué d'objets qui sont reliés les uns aux autres par des mailles. Un cas particulier d'application existe si ces éléments représentent des logiciels exécutables ou des artefacts. Ces artefacts s'exécutent sur des environnements d'exécution ou des périphériques que UML 2.0 classe comme nœuds. L'artefact dépend donc du nœud. La distribution représente cette relation de dépendance qui survient pendant l'installation.

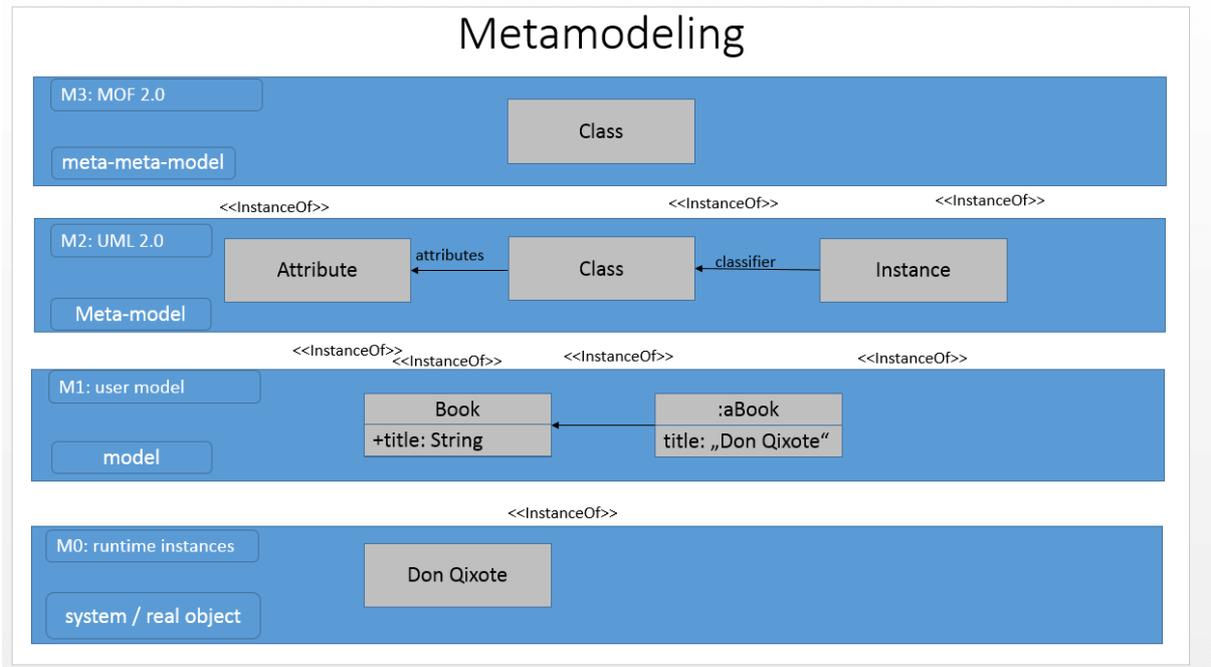
**Encapsulation** : Consiste à masquer les détails d'implémentation d'un objet, en définissant une interface. L'interface est la vue externe d'un objet, elle définit les services accessibles (offerts) aux utilisateurs de l'objet. L'encapsulation facilite l'évolution d'une application car elle stabilise l'utilisation des objets : on peut modifier l'implémentation des attributs d'un objet sans modifier son interface. L'encapsulation garantit l'intégrité des données, car elle permet d'interdire l'accès direct aux attributs des objets (utilisation d'accessieurs).

**Flux d'information** : Cette unité de langage UML décrit l'unité d'information sur les éléments et le flux d'information. Ces éléments de modèle définissent des techniques de description du comportement qui peuvent être très détaillées, comme les activités ou les interactions. Cette représentation simplifiée permet l'utilisation universelle de ces éléments de modélisation dans tous les types de diagrammes UML.

**Héritage et polymorphisme** : L'héritage est un mécanisme de transmission des propriétés d'une classe (ses attributs et méthodes) vers une sous-classe. Une classe peut être spécialisée en d'autres classes, afin d'y ajouter des caractéristiques spécifiques ou d'en adapter certaines. Plusieurs classes peuvent être généralisées en une classe qui les factorise, afin de regrouper les caractéristiques communes d'un ensemble de classes. La spécialisation et la généralisation permettent de construire des hiérarchies de classes. L'héritage peut être simple ou multiple. L'héritage évite la duplication et encourage la réutilisation. Le polymorphisme représente la faculté d'une méthode à pouvoir s'appliquer à des objets de classes différentes. Le polymorphisme augmente la généralité du code.

**Métamodélisation** : UML 2.0 définit des unités de langage qui fonctionnent à différents niveaux. Vous les utilisez pour exprimer la structure et le comportement d'un système. La méta-modélisation inclut tous les éléments d'UML, y compris ceux qui décrivent UML lui-même. Il utilise quatre niveaux hiérarchisés (M0 à M3).

Le niveau méta-méta M3 spécifie les métadonnées du langage de modélisation et ses relations à l'aide de la fonction métaobjet (MOF). Il définit le méta-modèle et permet également le transfert des métadonnées. Le format XML défini par les membres du groupe de gestion d'objets est un outil pratique pour partager des données orientées objet au niveau méta-méta entre les outils de développement. Le langage de contrainte d'objet (OCL), un langage de programmation déclaratif, complète UML et régule les conditions limites de la modélisation. En tant que langage texte, cependant, il n'est qu'un support et ne peut pas être utilisé pour la modélisation.



**Modèle :** Le modèle comprend tous les éléments nécessaires pour représenter une vue spécifique de la structure ou du comportement d'un système. Cela inclut également les influences externes telles que les acteurs.

**OCL:** Object Constraint Language

OCL est une contribution d'IBM à UML 1.1. Ce langage formel est volontairement simple d'accès et possède une grammaire élémentaire (OCL peut être interprété par des outils). Il représente un juste milieu, entre langage naturel et langage mathématique. OCL permet ainsi de limiter les ambiguïtés, tout en restant accessible.

OCL permet de décrire des invariants dans un modèle, sous forme de pseudo-code : pré et post-conditions pour une opération, expressions de navigation, expressions booléennes, etc... OCL est largement utilisé dans la définition dans l'UML.

**Profil :** Un profil configure UML 2.0 pour des besoins spécifiques. Des termes abstraits tels que activité ou objet doivent être spécifiés pour certains projets afin d'améliorer la compréhension. Vous pouvez utiliser un profil pour ajuster la sémantique et les notations qui sont définies de manière approximative.

**Stéréotype :** Les stéréotypes permettent d'étendre la sémantique des éléments de modélisation : il s'agit d'un mécanisme d'extensibilité du métamodèle d'UML.

Les stéréotypes permettent de définir de nouvelles classes d'éléments de modélisation, en plus du noyau prédéfini par UML.

Utilisez les stéréotypes avec modération et de manière concertée (notez aussi qu'UML propose de nombreux stéréotypes standards).

**Structure composite** : La structure composite des unités de langage décrit les éléments qui sont protégés vers l'intérieur et vers l'extérieur, comme les composants. Seuls les ports connectent le contenu au système externe. Les classificateurs dits encapsulés sont composés d'éléments appelés parties. Les pièces communiquent par l'intermédiaire de connecteurs.

**UML** : L'UML est un outil qui donne une dimension méthodologique à l'approche objet et qui permet de mieux maîtriser sa richesse. C'est un outil qui permet de décrire les besoins du client, et l'architecture une solution.

**Advanced Encryption Standard** : Standard de chiffrement choisi par le *National Institute of Standards and Technology* (NIST) destiné à remplacer DES et 3 DES pour lequel des failles étaient apparues. Voir *Data Encryption Standard*.

**Adware** : Logiciel qui conduit automatiquement l'internaute sur des sites où il ne voulait pas aller ou qui l'oblige à visualiser des publicités, en échange de sa gratuité.

**AES** : Voir *Advanced Encryption Standard*.

**Alerte éthique** : Voir *Alerte professionnelle*.

**Alerte professionnelle** : Dispositif permettant aux employés de dénoncer les délits financiers dont ils ont connaissance.

**Bannière** : Annonce publicitaire comportant une image ou un message, affichée en général sur la page d'accueil d'un site et sur laquelle l'internaute est invité à cliquer.

**Bartering** : Opération qui consiste pour une entreprise à coproduire, à travers une marque, le contenu proposé au public.

**Bios** : Abréviation de *Basic input output system*. Programme contenu dans la carte mère qui s'exécute en premier lors de l'allumage de l'ordinateur. Il commande les opérations élémentaires nécessaires au démarrage de l'ordinateur.

**Bittorrent** : Concept imaginé par Bram Cohen qui adopte un principe simple : l'utilisateur ne télécharge plus un fichier unique se trouvant sur l'ordinateur d'un autre utilisateur (principe retenu par le *peer-to-peer*), mais sur celui d'une multitude d'utilisateurs simultanément, tout en envoyant lui-même aux autres des morceaux du fichier qu'il est en train de télécharger. Les utilisateurs sont coordonnés grâce aux informations contenues dans le fichier *Torrent* qu'ils sont en train de télécharger.

**Blog** : Contraction de *web* et *log* (journal), désigne un site personnel, le plus souvent non professionnel, dont l'objet est de présenter un ensemble d'informations, d'images, de liens hypertextes, ou tout autre contenu, régulièrement mis à jour et présenté de manière antéchronologique, sur lequel les internautes peuvent, en général, intervenir par le biais de commentaires, modérés a priori ou non.

**Bombe logique** : Petit programme informatique qui provoque un bug en cas de tentative de lecture sur l'ordinateur.

**Botnet** : Voir *Ordinateur-zombie*.

**Bureau d'enregistrement** : Voir Registrar.

**Caching** : Forme de stockage « intermédiaire » et « temporaire » d'une information qui permet de ne pas encombrer les réseaux et favorise la fluidité des connexions. Le principe consiste pour les fournisseurs d'accès à stocker sur leurs serveurs les pages *web* les plus souvent consultées ou les

contenus les plus souvent téléchargés, ce qui permet de réduire au maximum le volume du trafic et d'accélérer les temps de connexion à un site *web*.

**Centre Français d'exploitation du droit de copie (CFC)** : Société de perception et de répartition des droits de propriété littéraire et artistique, agréé par le ministère de la Culture. C'est le gestionnaire du droit de reproduction par reprographie de la presse et du livre en France. Il représente tous les auteurs et les éditeurs de livres et de presse et peut délivrer des autorisations de reproduction par reprographie dans ces domaines.

**Chat room** : Salon de discussion virtuel synchrone.

**Cheval de Troie** : Désigne un programme introduit par l'utilisateur lui-même, sans le savoir, généralement via un jeu vidéo ou un petit utilitaire. Une fois dans un ordinateur, les chevaux de Troie installent à distance des fonctionnalités non demandées et servent à commettre des actes de cybercriminalité à l'insu du propriétaire de la machine. C'est notamment le cas du « crack » appelé « Windows Vista All Versions Activation 21.11.06 » : alors qu'il est censé activer toutes les versions disponibles de Vista, il permet à ses concepteur d'espionner les saisies clavier.

**Chiffrement** : Méthode de cryptage modifiant les données pour les rendre illisibles sans la clé correspondante.

**Contributory infringement** : Contrefaçon par « contribution ». Cette notion américaine trouve son équivalent, en droit français, dans le délit de complicité de contrefaçon.

**Cookie** : Désigne un fichier témoin, qui peut s'installer sur le disque dur de l'internaute, à l'occasion de la consultation de certains sites. Un *cookie* permet à son expéditeur de collecter des données comportementales de l'internaute et de lui restituer ses habitudes de navigation à son prochain passage sur le site, lui évitant, par exemple, d'entrer de nouveau ses identifiants de connexion.

**Copyleft** : Le terme *copyleft* fait référence au *copyright* dont il est le strict opposé. Le *copyleft* repose sur une licence d'utilisation permettant la réservation privative de l'œuvre à l'auteur mais interdisant son appropriation par autrui afin d'en assurer le partage et l'évolution. Voir *Logiciels libres*.

**Creative commons** : *Œuvre de partage ou œuvre en usage partagé*, expressions recommandées par la Commission générale de terminologie et de néologie (avis sur Legifrance).

**Cybersquatting** : Enregistrement intentionnel d'un nom de domaine, utilisé par un tiers comme nom commercial ou marque, dans le but d'empêcher le propriétaire de la marque d'établir un site *web* identifié par le nom de domaine (traduction proposée par la revue *DIT* 04/1997, p. 37, d'une citation de David Pauker, « Introduction to domain name disputes »).

**Cybersurveillance** : Proposition de définition : « Surveillance des salariés au moyen des matériels issus des NTIC dont les ordinateurs (PC-Terminaux), les logiciels (dont les systèmes de badges), les messageries, les connexions et modes d'accessions aux différents réseaux intra/extra/Internet, voire téléphoniques en cas de liaison téléphonie-informatique. Cette surveillance est rendue possible, de par le matériel précité, mais également par la circulation des données numérisées, circulation qui laisse systématiquement une trace de son passage. Car la cybersurveillance, c'est aussi la "traçabilité" : chaque action du salarié (ouverture, fermeture de session, utilisation de la messagerie,

saisie, traitement de données, connexion et circulation sur un réseau, etc.) laisse l'indication de sa date et de sa nature » (Les cahiers du DRH no 89, 10 oct. 2003, définitions, J.-L. Denier, journaliste de la rédaction).

**Data Encryption Standard** : Algorithme de chiffrement à clé secrète mis au point dans les premières années de 1970. La petite taille de clé utilisée initialement et l'arrivée des microprocesseurs ont rapidement rendu ce système de chiffrement vulnérable. DES a été remplacé par Triple DES ou 3 DES dont le principe consiste en l'enchaînement de 3 applications successives de l'algorithme DES puis par AES. Voir *Advanced Encryption System*.

**DCSSI** : Voir *Direction centrale de la sécurité des systèmes d'information*.

**Deep linking** : Voir *Lien profond*.

**DES** : Voir *Data Encryption Standard*.

**Digg-Like** : Site communautaire diffusant des articles soumis par des internautes et hiérarchisés selon les notes attribuées par les utilisateurs. Ce terme tient son nom du premier site à l'origine de ce concept dénommé Digg qui fait partie des 50 sites Internet les plus fréquentés au monde.

**Direction centrale de la sécurité des systèmes d'Information** : Service du Premier ministre qui assure la fonction d'autorité nationale de régulation qui évalue les menaces pesant sur les systèmes d'information en donnant l'alerte et en développant les capacités à les contrer et à les prévenir.

**Domain tasting** : Ce mécanisme consiste à autoriser les *registrars* à procéder gratuitement, pendant un délai de cinq jours, au dépôt de noms de domaine pour les tester afin de ne conserver que ceux qui génèrent du trafic.

**Domaine public** : Voir *Œuvre du domaine public*.

**Données de trafic** : « On entend par données relatives au trafic toutes les données traitées en vue de l'acheminement d'une communication par un réseau de communication électronique ou en vue de sa facturation » (CPCE, art. L. 32-18 – v. aussi dir. no 2002/58/CE, 12 juill. 2002, concernant le traitement des données à caractère personnel et la protection de la vie privée dans le secteur des communications électroniques, dite directive « vie privée et communications électroniques », art. 2 b.).

**Download** : Voir *Téléchargement descendant*.

**EDI** : Échange par voie électronique de données structurées, organisées en messages normalisés ; l'EDI est utilisé dans les secteurs industriels et commerciaux depuis de nombreuses années.

**Fair use** : Le concept américain de *fair use*, prévu par le paragraphe 107 de la section 17 du US Code, permet de considérer que l'utilisation d'une œuvre protégée n'est pas contrefaisante si elle est « loyale ». L'appréciation relève du pouvoir du juge, auquel il appartient d'examiner plusieurs critères cumulatifs : (i) la finalité de l'utilisation de l'œuvre par le copiste, (ii) la nature de l'œuvre copiée, (iii) l'importance des extraits de l'œuvre copiée, (iv) les conséquences de l'acte de reproduction de l'œuvre protégée sur le marché potentiel ou sur la valeur de la création protégée.

**Firewalls** : Outils de filtrage des flux de données qui transitent sur un réseau.

**Flaming** : Proviens de l'anglais *flame*, la flamme. Le flaming consiste à se livrer à des attaques via l'Internet en ayant la volonté de perturber le système d'information de son interlocuteur et en suscitant un encombrement important de la capacité mémoire.

**Flash** : Voir *Pop-up*.

**Flux RSS (really simple syndication)** : Ce système permet à l'internaute d'avoir accès à l'information in extenso par un lien hypertexte qui renvoie sur le site source de l'information. Les contenus s'affichent automatiquement et sont régulièrement mis à jour.

**Forum de discussion** : Lieu virtuel de discussion autour d'un thème d'intérêt commun ; chaque membre du forum peut ainsi émettre des messages accessibles aux autres participants, lesquels peuvent bien entendu lui répondre et argumenter ; selon la Commission générale de terminologie et de néologie, il s'agit d'un « service permettant l'échange et la discussion sur un thème donné : chaque utilisateur peut lire à tout moment les interventions de tous les autres et apporter sa propre contribution sous forme d'article ».

**Framing** : C'est le terme retenu lorsque le lien hypertexte pointe vers des éléments placés dans une fenêtre distincte de celle du site source. Voir également *Inline linking*.

**Free software** : Voir *Logiciel libre*.

**Freeware** : Voir *Logiciel freeware*.

**Gang name** : Voir *Grabbing*.

**Grabbing** : Cette pratique consiste à enregistrer des noms de domaine correspondant à des marques de grande renommée en vue de les revendre.

**Hacker** : Désigne un pirate informatique et électronique ; du verbe *to hack* : hacher, tailler en pièces. Toutefois, un *hacker* se limite à visiter des sites et à les démonter en procédant à des modifications alors qu'un *cracker* s'emploie à s'introduire dans les sites pour les piller.

**Happy slapping** : Acte d'agression délibérée à l'initiative d'un groupe ou d'une personne dans le but de réaliser une vidéo et de diffuser celle-ci. À titre d'illustration, le meurtre par un groupe d'adolescents d'un jeune garçon à Londres, gare de Waterloo, en décembre 2005 ou encore les images de l'agression d'une enseignante d'un lycée français, dans les Yvelines, ont été diffusées entre les élèves de l'établissement.

**Html** : voir *Hyper Text Mark-up Language (html)*.

**HTTPS** : Voir *IPsec, SSL/TLS, HTTPS*.

**Hyper Text Mark-up Language** : Désigne le langage de transcription des documents accessibles via le *web*.

**Hyperlien** : voir lien Hypertexte.

**IDDN** : Voir *Inter Deposit Digital Number*.

**Inline linking** : C'est le terme retenu lorsque les éléments sur lesquels pointe un lien hypertexte profond ne permettent pas à l'internaute de faire la différence entre les éléments d'origine et les éléments importés. Voir également *Framing*.

**Inter Deposit Digital Number** : Numéro international d'identification d'une œuvre, mis au point par l'agence pour la protection des programmes (APP). Ce scellé virtuel constitue la carte d'identité de l'œuvre, contenant toutes les données se rapportant à l'œuvre. L'utilisateur peut accéder à ces données sur le site créé à cet effet et prendre ainsi connaissance des revendications et des conditions particulières de l'auteur.

**Interactivite** : Grâce à un logiciel de navigation, l'utilisateur peut intervenir dans le déroulement de l'œuvre qui n'est donc plus nécessairement linéaire.

**Internet Protocol Adress** : Grâce à un logiciel de navigation, l'utilisateur peut intervenir dans le déroulement de l'œuvre qui n'est donc plus nécessairement linéaire.

**IP** : Voir *Internet Protocol Adress*.

**IPSEC, SSL/TLS, HTTPS** : Protocoles réseaux intégrant directement le chiffrement des données transmises.

**Licence CeCILL** : Première licence française de logiciel libre mise au point par le CEA, le Cnrs et l'Inria, organismes chargés d'un service public, pour pouvoir diffuser leurs logiciels libres en conformité avec les exigences de la loi française. Cecill est un acronyme constitué de Ce pour « CEA », C pour « Cnrs », I pour « Inria », LL pour « Logiciel Libre ».

**Licence creative commons** : Elle définit les conditions de mise à disposition des œuvres en ligne. Elles se déclinent sous la forme de contrats-type et de leurs résumés explicatifs.

**Licence open source** : Le concept a été mis au point par la *Free Software Foundation* en 1989, sous l'impulsion de Richard Stallman qui, ayant développé un système d'exploitation libre, a décidé, pour favoriser l'évolution du logiciel, de le rendre compatible avec UNIX, le dénommant GNU. Le dispositif contractuel organise l'enrichissement systématique du logiciel, sous le contrôle du donneur de la licence initiale.

**Licence Shrink-Wrap** : L'accord est donné en déchirant l'emballage et le consentement de l'utilisateur étant réputé acquis dès qu'il utilise le logiciel.

**Lien commercial** : voir *Lien sponsorisé*.

**Lien Hypermedia** : Prend la forme d'images, parfois sur fonds sonores, ou encore d'un interstitiel qui occupe l'écran ou une fenêtre pendant quelques secondes.

**Lien Hypertexte** : Procédé qui permet, en cliquant sur un mot affiché à l'écran, d'accéder aux informations auxquelles renvoie ce mot.

**Lien profond** : Le lien hypertexte renvoie vers une page secondaire du site web pointé, permettant ainsi l'« importation » sur un site web, de pages web, d'images, de textes, etc.

**Lien sponsorisé** : Technique publicitaire pratiquée sur l'internet. Elle permet de faire connaître un produit ou un service en ligne par des liens « promotionnels » qui apparaissent en fonction des mots-clés choisis pour la recherche.

**Logiciel freeware** : Dans la traduction littérale du terme *free*, il y a en fait deux notions : « gratuit » et « libre ». Si les logiciels libres sont souvent gratuits, leur caractéristique principale réside avant tout dans une liberté « encadrée » et dans leur caractère évolutif et dynamique. Chaque utilisateur doit disposer d'un droit de libre intervention sur le code source du logiciel libre à la condition que la rediffusion du logiciel, adaptée ou non, soit également libre (mais pas nécessairement gratuite). La licence open source constitue le modèle de référence pour les logiciels libres. Voir *Licence open source*.

**Logiciel shareware** : On dit d'un logiciel qu'il est en shareware lorsqu'il peut être utilisé gratuitement pendant une certaine période ou pour un certain nombre d'utilisations et à l'issue desquels une contribution financière – souvent modeste, visant le plus souvent à couvrir seulement les frais de distribution et le cas échéant, de support technique – est demandée.

**Login** : Permet l'identification des utilisateurs sur un serveur.

**Mailbombing** : Envoi sur une adresse email d'un nombre important de messages, visant à entraver par saturation le fonctionnement d'un système de traitement automatisé de données.

**Meta-tags** : Il s'agit d'éléments — par exemple le nom d'une marque — qui sont insérés dans le code source html, de manière invisible pour l'internaute, afin de permettre le référencement automatique du site au travers de moteurs de recherche.

**MP3, MP4, MPEG** : Le format MP3 a été mis au point par le chercheur Karl-Heinz Brandenburg. Il permet la compression des enregistrements sonores tout en conservant une qualité de restitution du son proche de celle du CD. Plus généralement, ces formats consistent à convertir des données analogiques en langage binaire, c'est-à-dire à décomposer les œuvres en unités d'informations.

**Multimedia** : Voir *Oeuvre multimedia*.

**Netiquette** : Règles de comportement non écrites et adoptées par tous les utilisateurs de l'internet dans le monde.

**Newsgroup** : Forum de discussion accessible sur le réseau Usenet. Voir *Forum*.

**Nom de domaine** : Désigne l'adresse permettant de mémoriser et d'identifier le site sur le réseau internet afin de pouvoir y accéder. Le nom de domaine peut contenir jusqu'à 63 caractères. Un nom de domaine est toujours associé à une adresse numérique fixe, l'adresse IP. Voir *Internet Protocol Adress*.

**Œuvre du domaine public** : Une œuvre protégée par le droit d'auteur est dite tombée dans le domaine public lorsque le droit patrimonial des ayants droit de l'auteur est venu à son terme, c'est-à-dire 70 ans après la mort de l'auteur, les prorogations de guerre (soit 6 ans et 153 jours pour la 1<sup>re</sup> Guerre mondiale et 8 ans et 120 jours pour la Seconde Guerre) et les 30 années supplémentaires accordées en faveur des auteurs morts pour la France, s'ajoutant à cette durée légale. À noter que l'organisation américaine *Creative Commons*, a mis au point une application *web* pour identifier, par

les moteurs de recherche, les créations relevant du domaine public et pour, ainsi, informer les internautes sur les droits des auteurs concernés.

**Œuvre en partage :** Œuvre soumise aux conditions contractuelles de *Creative Commons*, une organisation américaine non gouvernementale, à but non lucratif, fondée en 2002 qui propose d'appliquer les règles de l'open source aux créations numériques (vidéo, texte, photo, musique, etc.). L'objectif est de partager ces créations et d'enrichir le patrimoine commun, les Commons. Voir *Creative Commons*.

**Œuvre en usage partagé :** Voir *œuvre en partage*.

**Œuvre libre de droit :** Œuvre dont le titulaire des droits d'auteur autorise la libre circulation en imposant sa plus large diffusion. Voir *Copyleft*.

**Œuvre multimedia :** Œuvre intégrant des textes, des sons et des images, avec, pour caractéristique essentielle, l'interactivité qui permet à l'utilisateur, au moyen de liens hypertextes et hypermedia d'intervenir dans le déroulement de l'œuvre.

**Opt-in :** Le destinataire d'e-mails se voit offrir la possibilité de s'opposer (droit d'opposition), de manière effective, à l'utilisation de ses coordonnées (règle de l'*opt-in*). Voir *Opt-out*.

**Opt-out :** Il appartient à l'internaute d'effectuer lui-même la démarche pour s'opposer à l'envoi d'*e-mails* publicitaires non sollicités en s'inscrivant sur des registres créés spécialement à cet effet et que les prestataires devront consulter avant tout envoi. À noter que la directive européenne 2000/31/CE du 8 juin 2000 sur le commerce électronique préconise la solution de l'*opt-out* mais la directive du 12 juillet 2002 a retenu le système de l'*opt-in*. C'est ce système qui a été retenu par la législation française avec l'adoption de la LCEN du 21 juin 2004.

**Ordinateur-zombie :** Ordinateur piloté à distance, à l'insu de son propriétaire.

**Pare-feu :** Voir *Firewalls*.

**Passenger name record :** Informations personnelles concernant les passagers en partance pour les Etats-Unis qui sont stockées dans les bases de données des systèmes de réservation renseignés par les agences de voyages et les compagnies aériennes elles-mêmes. Ces données indiquent l'itinéraire suivi par le passager, le nom des personnes avec lesquelles il a effectué une réservation, les coordonnées de celles chez qui il séjournera aux Etats-Unis, les réservations d'hôtels éventuelles ou encore les services demandés à bord (repas, prestations de santé, etc.).

**Peer-to-peer :** Technologie qui permet d'échanger via l'internet des fichiers numériques (audio, vidéo, etc.) entre particuliers (de « pair à pair » ou de « poste à poste ») et de mettre ainsi à la disposition de millions d'internautes des milliers d'œuvres protégées. On distingue les réseaux de type décentralisés (ex. : eDonkey, BitTorrent) des réseaux centralisés (ex. : Napster). Les premiers permettent de mettre en relation directement les internautes, ces derniers étant en même temps émetteur et récepteur de données. Les seconds obligent à passer par un serveur unique qui centralise la liste des fichiers offerts en partage et réoriente l'internaute vers le fichier qui l'intéresse. À signaler également : le logiciel Winny présente quant à lui l'avantage d'offrir l'anonymat à ses

utilisateurs : disponible sur internet depuis mai 2002, il s'agit du programme *peer-to-peer* le plus utilisé au Japon, avec près d'un million d'adeptes.

**Phishing** : Constitue un mode d'usurpation d'identité numérique. Le terme, aujourd'hui traduit par « filoutage », provient d'un mélange des mots *fishing* (pêcher) et *phreaking* (utilisation frauduleuse des lignes téléphoniques). Il s'agit plus précisément d'une technique visant à adresser un mail à un internaute l'invitant à se connecter à un site, copie parfaite d'un site connu de l'internaute (généralement sa banque) et à lui réclamer des informations confidentielles (ses coordonnées bancaires).

**Player** : Nom générique des logiciels de lecture vidéo ou audio.

**PNR** : Voir *Passenger name record*.

**Pop-up** : « Une *pop-up* ou fenêtre surgissante est une fenêtre informatique qui s'affiche au-dessus de la fenêtre de navigation normale lorsqu'on navigue sur internet souvent pour afficher un message publicitaire ».

**Profil d'habilitation** : Définition des droits d'un groupe d'utilisateurs concernant leur accès aux données ou aux applications.

**Protocole IP** : Voir *Internet Protocol Adress*.

**Radio frequency identification** : Technologie qui permet une identification à distance au moyen d'ondes électromagnétiques. Elle est notamment utilisée dans les moyens de transport et dans la logistique, et de plus en plus au sein des hôpitaux. En ce qu'il suppose un traitement de données à caractère personnel, voire des données biométriques, ce dispositif d'identification à distance doit être mis en œuvre conformément aux dispositions de la loi Informatique et Libertés.

**Registrar** : Société ou association permettant le dépôt des noms de domaines, dans les TLD.

**Réseau privé virtuel** : Voir *Virtual Private Network*.

**RFDI** : Voir *Radio Frequency Identification*.

**Rootkit** : Programme permettant de rendre un autre programme invisible à un outil de sécurité.

**Routeur** : Equipement permettant la répartition de données entre différents réseaux ou machines.

**RSS** : Voir *Logiciel shareware*.

**Safe harbor** : Sphère de sécurité entre l'union Européenne et les États-Unis.

**Sampling** : Mixage.

**Scoring** : Technique utilisée par les établissements de crédit pour décider ou non de l'attribution de crédits à la consommation, à l'habitat ou pour les besoins des professionnels.

**Screenner** : Vidéo « amateur » réalisée par la captation à l'aide d'une caméra d'une œuvre audiovisuelle projetée sur un écran dans une salle de cinéma.

**Shareware** : Voir *Logiciel shareware*.

**Short message service** : Il s'agit d'un message court qui s'affiche à l'écran du terminal mobile du destinataire. Ces messages sont également soumis au régime de l'*opt-in*.

**Slamming** : Cette pratique consiste à inciter les titulaires de noms de domaine, arrivés en fin de cycle, à renouveler l'enregistrement pour éviter le risque prétendu de la suppression de celui-ci.

**SMS** : Voir *Short Message Service (SMS)*.

**Snapback** : Désigne un service de surveillance pour avertir le titulaire d'une marque au cas où le nom de domaine, sélectionné au préalable sur <.com>, <.net>, <.org>, deviendrait disponible.

**Space-shifting** : Technique qui consiste à transformer un contenu numérique pour faciliter son utilisation ou sa transmission sur d'autres supports ou médias, par exemple pour passer d'un format CD à un format MP3.

**Spam** : Voir *Spamming*.

**Spamdexing** : Le mot est le résultat de la contraction des mots spam et indexing ; il s'agit des formules mathématiques contenues dans les logiciels (robots) qui indexent et classent les sites ; il permet ainsi de mieux positionner un site dans la page de résultat des moteurs de recherche, en « trompant » les algorithmes des moteurs de recherche.

**Spamming** : Envoi massif, et parfois répété, de courriers électroniques non sollicités, à des personnes avec lesquelles l'expéditeur n'a jamais eu de contact et dont il a capté l'adresse électronique de façon irrégulière. Dans tous les cas, le spamming est caractérisé par trois éléments : (i) « un envoi massif » de messages, (ii) à de très nombreux destinataires, (iii) à l'initiative d'un « expéditeur ».

**Sponsoring** : Opérations visant à associer une marque à un programme en relation avec son activité ou sa cible potentielle. C'est également le principe du *bartering*.

**Spyware** : Logiciel espion qui analyse l'activité d'un ordinateur à l'insu de son propriétaire et qui peut y récupérer des informations confidentielles, généralement des informations personnelles et financières.

**SSL/TLS** : Voir *IPsec, SSL/TLS, HTTPS*.

**STAD** : Voir *Système de traitement automatisé de données*.

**Streaming** : Technique qui offre à l'internaute la faculté d'accéder aux fichiers de données numérisées avant la fin de l'opération de leur téléchargement. Cette technique permet ainsi de recevoir les émissions de radios et de télévision sur l'internet et s'inscrit dans une logique d'interactivité qui pourrait remettre en cause les frontières communément admises entre l'œuvre audiovisuelle et l'œuvre multimédia.

**Suck** : En matière de noms de domaine, cette pratique consiste à enregistrer un nom connu en y ajoutant le suffixe « suck ». Elle a pour finalité de critiquer le titulaire du nom.

**Système de traitement automatisé de données** : « Ensemble composé d'une ou plusieurs unités de traitement automatisé de mémoire, de logiciel, de données, et d'organes d'entrée sortie et de liaisons qui concourent à un résultat déterminé, cet ensemble étant protégé par un dispositif de sécurité » (définition proposée par le Sénat lors de l'élaboration du nouveau Code pénal, v. rapp. J. Thyraud, Doc. Sénat, 1987-88, no 3, p. 52).

**Téléchargement ascendant** : Il s'agit d'adresser un fichier vers un ordinateur distant pour le mettre à disposition des autres internautes (opération de *upload* ou téléchargement « ascendant »).

**Téléchargement descendant** : L'opération consiste pour un internaute à reproduire sur son disque dur un fichier stocké sur un ordinateur distant (opération de *download* ou *téléchargement* « descendant »).

**Torrent** : Concept imaginé par Bram Cohen, un jeune informaticien américain. Il s'agit de « coupler » une description structurée des documents audio-visuels grâce aux *Torrents* et une méthode de distribution via un réseau *peer-to-peer*. Mais contrairement à Kazaa où un même morceau de musique peut parfois exister sous des centaines de noms différents, il n'existe en général qu'un seul Torrent pour des milliers d'utilisateurs ce qui accélère les téléchargements.

**Tunneling** : Encapsulation de données d'un protocole réseau dans un autre.

**Virtual Private Network** : Désigne un réseau privé virtuel sécurisé par l'encapsulation et le chiffrement des données transmises par internet entre ses utilisateurs.

**Virus informatique** : Désigne un programme informatique qui possède la particularité de se reproduire et de se propager en utilisant les capacités d'un système informatique qui l'héberge.

**VPN** : Voir *Virtual Private Network*.

**Web 2.0** : Concept employé pour la première fois par le journaliste Dale Dougherty en 2003. Il a été inventé suite à l'évolution des relations « internaute-site Internet ». Par opposition au « web 1.0 » qui définit des sites dont le contenu est statique et administré par une unique personne, ce concept regroupe l'ensemble des technologies Internet permettant à n'importe quel internaute d'interagir avec soit le contenu d'un site (contenu alors dynamique), soit avec d'autres internautes (aspect communautaire).

**Webforum** : Forum de discussion accessible sur le *web*. Voir *Forum*.

**Weblog** : Voir *Blog*.



---

## Lexiques de la base de données

**Algèbre relationnel** : L'algèbre relationnelle est la théorie mathématique inventée par Edgar Frank Codd sur laquelle repose les bases de données relationnelles. Dérivée de la théorie des ensembles, cette algèbre utilise des opérateurs ensemblistes et des opérateurs relationnels qui permettent de définir comment les informations sont liées entre elles et comment on peut y accéder. À partir d'un nombre basique d'opérations, il est ainsi possible d'accéder des données qui font sens entre elles (relations).

### Opérateurs ensemblistes de l'algèbre relationnelle

- **Union** : sélectionne les données de tous les ensembles
- **Intersection** : sélectionne uniquement les données communes
- **Différence** : sélectionne uniquement les données faisant partie d'un ensemble mais pas de l'autre
- **Produit cartésien** : sélectionne tous les couples possibles entre les éléments des différents ensembles

### Opérateurs relationnels

- Sélection
- Projection
- Jointure
- Division

**Association** : Les bases de données sont constituées de différentes données homogènes qui discutent entre elles (les entités). Lorsque l'on représente les liaisons et les interactions entre ces différentes données cohérentes, on parle d'association.

En conception de base de données, on dit qu'une association précise et explicite le lien entre plusieurs entités. Une association se fait entre 2 entités minimum mais il n'y a pas de maximum.

**Attribut** : Quand on découvre la conception de bases de données, on utilise beaucoup de termes sur mesure. Pas de chance, certains sont synonymes. C'est le cas d'Attribut qui peut être remplacé par Propriété (voir ce lien, la définition y est complète).

Un attribut est un identificateur (idéalement explicite – un nom parlant) et correspond à une donnée que l'on va informatiser en base et qui se caractérise par un nom, un emplacement, un type de données ainsi que d'autres caractéristiques (identifiant par exemple...).

**Banque de données** : Lorsque plusieurs bases de données fonctionnent au sein d'un système homogène (entreprise, recherche médicale par exemple), on parle de banque de données ou data bank en anglais.

Les informations peuvent alors être physiquement être présentes au même endroit ou pas (lieu physique ou site internet regroupant différentes sources). L'idée principale est de faciliter la recherche en proposant un lieu commun.

**Base de données objet** : Dans la grande famille des bases de données existe le vilain petit canard : les bases de données objet. Apparues en même tant que les langages de programmation orientés objet

dans les années 1990, les bases de données objet se distinguent des bases de données classiques (les bases de données relationnelles).

En effet, une base de données objet manipule des objets. En informatique, un objet est une notion théorique qui représente une entité complexe et évolutive. Plutôt que de traiter tous les aspects en direct d'un objet, on y accède via un mécanisme simplifié qui masque la complexité interne. On a donc pas accès à une grande matrice de données (colonnes et lignes comme dans une base de données relationnelles) mais plutôt des portes d'entrées spécifiques que l'on peut ouvrir (ou pas) afin d'y ajouter, modifier et supprimer des contenus particuliers. Classes, héritage, encapsulation, polymorphisme... tout ce qui existe en POO s'applique aux bases de données Objet.

Les bases de données objet prennent tout leur sens lorsqu'il s'agit de modéliser des données très riches et complexes avec de nombreuses variations : documents multimédias, données géographiques, conception par ordinateur...

De par sa conception initiale, les bases de données sont rapides. Rarement besoin de jointures. De plus, les SGBDO permettent de réaliser des opérations sur mesure pour chaque objet au delà de Ajout/Modification/Suppression... Cerise sur le gâteau, une base de données objet et un logiciel réalisé en programmation orientée objet ne nécessitent qu'une seule et même modélisation informatique. Les objets utilisés en programmation se retrouvent directement en base de données.

**Base de données en colonnes :** Aujourd'hui, les bases de données rangent généralement les données en ligne : visuellement, on dispose de vues avec des colonnes qui contiennent le nom des champs et en ligne les valeurs.

Pour les bases de données en colonne, c'est différent car les données sont stockées en colonne. Pour les utilisateurs de haut niveau, ce changement ne se ressent pas du tout. Mais pour l'organisation et la manipulation des données par les outils informatiques, tout change. Ainsi, au lieu de manipuler les données ligne par ligne une fois sérialisée (c'est à dire mises bout à bout pour être traitées plus rapidement), les données sont utilisées colonne par colonne.

Les données en ligne :

1,Jean,DUPONT,7400;2,Julie,VAUQUIER,01100;3,Estelle,MALIA,75400;4,Théo,CONTIN,89130;

Puis en colonne :

1,2,3,4;Jean,Julie,Estelle,Théo;DUPONT,VAUQUIER,MALIA,CONTIN;74000,01100,75400,89130;

L'idée de travailler avec les colonnes n'est pas nouvelle (prémices en 1969) mais ce n'est que depuis 2005 avec l'apparition de tels outils distribués sous licence Open Source que les bases de données en colonnes ont vraiment décollé.

En fonction des volumes de données à traiter et des usages, les bases de données sont plutôt orientées lignes ou colonnes. Les bases de données OLTP sont généralement en lignes et les bases de données OLAP sont plutôt en colonnes.

**Base de données relationnelle :** **Relationnelle** se dit d'une base de données qui manipule des informations hiérarchisées qui sont liées entre elle par des contraintes de relations. À la différence d'autres types de bases de données, les bases de données relationnelles fonctionnent avec des tableaux à deux dimensions nommés tables qui sont liés à d'autres tables par un mécanisme de

relations reposant sur l'algèbre relationnelle. Ces relations définissent et précisent les liens entre les tables.

Conceptualisé par Edgar Frank Codd en 1970, la base de données relationnelle est aujourd'hui la famille la plus répandue des bases de données.

**Big data** : Existant depuis les années 2000, c'est le mot à la mode depuis 2010. Tout le monde parle BIG DATA (on voit plus rarement le terme recommandé mégadonnées).

Avec l'augmentation des capacités de stockage et la réduction des coûts liés, les entreprises collectent de plus en plus d'informations. Ces informations structurées sont rangées dans des bases de données (parmi d'autres outils) et sont parfois tellement volumineuses que leur exploitation nécessite des outils sur mesure. On parle alors de big data.

Ce qui est nouveau et très intéressant c'est que l'on peut exploiter ces données qui, prises une à une, n'ont pas vraiment d'intérêt mais qui, cumulées, recèlent une vraie valeur :

- Prévisions, tendances, prospective, recherche sur les attentes, les usages, les besoins des utilisateurs ;
- Usage en intelligence artificielle ;
- Exploitation statistiques, visualisation de données.

En big data, ce sont les données qui permettent de trouver des modèles. C'est l'inverse en informatique décisionnelle classique. Et ça c'est tout à fait nouveau : nous sommes ici dans un champ de recherche inédit en informatique avec des efforts de R&D importants.

Les secteurs d'activités qui sont déjà très consommateurs d'informations voient ici de nouvelles opportunités, c'est le cas en météorologie, en médecine, en gestion des risques... avec l'apparition de stratégies d'entreprises pilotées par les données (data-centric).

Cependant, le challenge avec les big data est multiple :

- S'assurer que les données collectées sont pertinentes : le plus n'est pas le mieux et le volume ne sert à rien si les données sont fausses.
- Savoir exploiter efficacement les données : vaut-il mieux des statistiques précises ou des statistiques justes ?
- Gérer les aspects liés à l'éthique, à la sécurité de l'information collectée et à la vie privée : qui peut collecter ? sous quelles conditions ? que deviennent les données ?

**Caractéristique** : En conception de base de données avec Merise, une Caractéristique est un synonyme de Propriété. Très simplement, c'est une donnée élémentaire (on ne peut pas la scinder / décomposer / deviner / calculer). Les caractéristiques vont se retrouver un peu partout en base mais toujours bien rangées dans des entités ou au sein d'associations.

**Cardinalités** : Lorsque l'on conçoit une base de données avec le MCD de Merise, on obtient un schéma avec des entités et des associations. Pour préciser au mieux les associations, on utilise des cardinalités. Les cardinalités sont des caractères (0,1, n) qui fonctionnent par couple et qui sont présents de chaque côté d'une association (sur chaque « patte »). Ils donnent des indications très intéressantes et permettent par la suite de construire la base de données :

- avec la création de clés étrangères dans le cas d'une CIF
- avec la création d'une table intermédiaire dans le cas d'une CIM

Les cardinalités possibles sont :

- 0,1 : au minimum 0, au maximum 1 seule valeur (CIF) ;
- 1,1 : au minimum 1, au maximum 1 seule valeur (CIF) ;
- 0,n : au minimum 0, au maximum plusieurs valeurs ;
- 1,n : au minimum 1, au maximum plusieurs valeurs.

Les cardinalités maximales (le côté droit ou le dernier caractère) définissent si la relation est une CIM ou une CIF.

Dans l'image, on a « 1,1 » en cardinalité du côté gauche de l'association et une cardinalité « 0,n » du côté droit. Concrètement, il faut lire les cardinalités ainsi :

- Une structure se situe dans 1 et 1 seule ville ;
- Dans une ville peut se situer 0 ou plusieurs (n) structures.

Les cardinalités sont importantes car elles traduisent les règles de gestion, elles doivent être validées avec l'utilisateur final.

**Champ** : En base de données, un champ peut signifier plusieurs choses différentes et pourtant proches suivant le degré d'informatisation et d'abstraction :

- À l'étape du MCD (Merise), le champ est la valeur contenue à l'intérieur d'une propriété au sein d'une entité ou d'une relation. Si on exportait le contenu d'une entité dans un tableur, on peut imaginer qu'un champ est l'intersection d'une ligne (ensemble cohérent de données) et d'une colonne (propriété).
- Aux étapes ultérieures (MPD, MLD, base de données en production), le champ est simplement équivalent à la propriété ou l'attribut d'une table.
- Aux étapes de conception d'interface, un champ est une zone dans laquelle l'utilisateur pourra lire ou renseigner une valeur (zone de texte, bouton radio, zone de liste...).

Il y a pas mal de confusions avec ce terme. Dans tous les cas, il est préférable lorsqu'il est évoqué, de bien préciser ce qui est sous-entendu.

**CIF** : Une **CIF** ou **Contrainte d'Intégrité Fonctionnelle** est un type d'association entre 2 entités.

Elle se caractérise par un 1 en cardinalité supérieure (0,1 ou 1,1) sur une des pattes de la relation. On dit alors que la relation est **porteuse d'une dépendance fonctionnelle**. Une CIF indique donc une dépendance. Une des entités de l'association est déterminée par la connaissance d'une (ou plusieurs) autre entité présente dans l'association. Cette association est forte et hiérarchique. Sans entité parent, il ne peut pas y avoir d'entité enfant. À ne pas confondre avec une CIM.

Une CIF se transforme ensuite (lorsque l'on passe au MPD / MLD) en une clé étrangère. Cette clé étrangère est un champ ajouté à la table située du côté du 1 en cardinalité maximale qui reprend le champ clé primaire de la table située de l'autre côté de la relation.

**Clé primaire** : Voilà un terme qui revient très souvent et qui est souvent confondu avec le terme identifiant qui est très proche.

Lorsque l'on conçoit une base de données, on utilise souvent une méthode afin de ne rien oublier et de structurer son travail. Merise en est une. Dans la méthode Merise, le concepteur réalise d'abord un MCD qu'il va transformer en MPD et en MLD.

À l'étape du MCD, on a des identifiants dans les entités. À l'étape du MPD et MLD on a des **clés primaires** dans les tables. Identifiants et clés primaires sont la même chose mais à des stades différents de la conception d'une base de données.

**CIM** : Une **CIM** ou **Contrainte d'Intégrité Multiple** est un type d'association entre 2 entités minimum.

Sur un MCD, elle se caractérise par l'absence de 1 en cardinalité maximale (sur aucune « patte » de la relation). Par déduction, on ne peut avoir que 0,n ou 1,n en cardinalité. À ne pas confondre avec la CIF.

Concrètement, une CIM représente un couple unique auquel on peut affecter des propriétés particulières. On dit alors que la CIM est porteuse de propriétés (ou de données). Non hiérarchique, la CIM est construite à partir des identifiants des entités qui lui sont liées.

**Clé étrangère** : Une **clé étrangère** en bases de données est un mécanisme simple qui permet de garantir l'intégrité référentielle entre les données de différentes tables. Concrètement, la clé étrangère oblige une table à être liée aux données d'une autre table. Ainsi dans l'exemple suivant, une structure est obligatoirement liée à une ville par la clé de la ville.

Ces 2 clés sont la traduction concrète de la relation entre les tables « structure » et « ville ».

Par convention, les clés étrangères sont représentées préfixées du symbole dièse #.

Dans une table (au niveau du MPD et du MLD), une clé étrangère est une clé particulière qui n'est pas la clé principale de la table. La clé principale est appelée clé primaire. Lorsqu'une ou plusieurs autres clés apparaissent dans une table, ce sont donc des clés étrangères. À la différence des clés primaires, on peut retrouver plusieurs fois la même valeur dans les clés étrangères d'une table (et c'est logique, dans l'exemple ci-dessus, il est normal que la même ville puisse être utilisée par plusieurs structures).

Les clés étrangères sont la conséquence directe de l'application de règles qui permettent le passage du MCD au MPD / MLD. Les clés sont les identifiants du MCD. Elles deviennent clés primaires ou clés étrangères suivant les cas de figures :

- Dans une CIF (une relation avec une cardinalité supérieure à 1), on obtient une clé étrangère. L'identifiant de l'entité qui n'est pas du côté du 1 est ajouté en tant que clé étrangère à la table du côté du 1. Au passage, cette clé récupère le #.
- Dans une CIM (une relation sans cardinalité supérieure à 1), il n'y a pas de clé étrangère mais une seule clé primaire composée des identifiants de chaque table liée.

**CRUD** : Abréviation en anglais de **Create, Read, Update, Delete**. En français : Écrire, Lire, Mettre à jour, Supprimer.

Il s'agit des opérations de base attendues par un SGBD (mais pas uniquement, on s'en sert aussi dans d'autres types de logiciels). Le terme CRUD date de 1983, c'est James Martin qui a popularisé l'expression dans son livre *Managing the Database Environment*.

On rencontre parfois le terme SCRUD ou le S initial signifie Search (Rechercher) ou encore SCRUDL avec le L final signifiant Large pour les bases de données volumineuses. Les abbréviations DRULAB (Delete, Read, Update, Lock, Add, Browse) et BREAD (Browse, Read, Edit, Add, Delete) sont très similaires.

**Concurrence** : Dans l'univers des **bases de données**, la notion de **concurrence** a une signification particulière bien différente du mot concurrent que nous connaissons tous et qui s'applique à l'économie.

En base de données, la concurrence signifie « **accès simultanés** ». Il est très fréquent que plusieurs personnes (ou programmes) souhaitent accéder « en même temps » aux mêmes données. C'est le cas lorsque l'on consulte un site web. C'est aussi le cas lorsque plusieurs utilisateurs souhaitent accéder à la même fiche client dans un logiciel de gestion commerciale.

Le cas est plus problématique lorsque plusieurs utilisateurs souhaitent modifier « en même temps » les mêmes données ou qu'une donnée consultée par une personne est parallèlement modifiée par une autre.

Pour gérer ces situations problématiques d'accès concurrentiels aux données, les bases de données utilise des règles de gestion et d'intégrité des données nommées ACID. Il en découle des traitements par lots (les transactions) et la mise en place de mécanismes d'isolation (pose de verrous pour interdire l'accès à certaines données pendant la durée d'un traitement). Les verrous posent aussi d'autres problèmes et il faut parfois décaler dans le temps les opérations conflictuelles. Ce qui entraîne encore d'autres situations problématiques (données non à jour, doublons, synchronisation...).

Derrière le terme de concurrence d'accès aux données, il y a aussi un enjeu de performance : en fonction du volume de données à traiter, tous les SGBD ne se valent pas. Certains peuvent absorber des dizaines de demandes simultanées alors que d'autres sont capables d'en traiter beaucoup plus. Enfin, l'aspect matériel et l'architecture technique déployée devient souvent un goulot d'étranglement lorsque les problématiques de volume arrivent.

**Contrainte d'intégrité** : Toutes les contraintes que doivent respecter les données sont appelées contraintes d'intégrité référentiel, ou plus simplement contrainte d'intégrité.

**Data mining** : Le data mining ou exploration de données en français est l'ensemble des techniques qui permet d'extraire des informations au sein de données et plus précisément de découvrir des structures (modèles et comportements) au sein de vastes ensemble de données. Le data-mining utilise les apports des statistiques et de l'intelligence artificielle pour isoler l'information valorisable et la présenter. Les informations extraites doivent être :

- justes, précises, fiables ;
- sourcées (on doit pouvoir savoir d'ou elles viennent précisément) ;
- intéressantes voire nouvelles.

En big-data et informatique décisionnelle, extraire de l'information cachée au sein de masses d'informations est fondamental. Le data mining est ainsi cette tâche clé qui consiste à fouiller, prospecter, forer dans des bases de données afin d'**en faire ressortir une connaissance** (information qui a de la valeur et qui est exploitable).

Cette idée d'exploiter l'information déjà présente (analyse secondaire – on ne cherche pas à collecter l'information mais simplement à l'analyser) est vieille de plusieurs siècles et très tôt en informatique (dès les années 60 avec les statistiques exploratoires), des outils ont été produits dans ce but. Les données textuelles sont les plus simples à traiter mais les éléments multimédias (sons, images) sont aujourd'hui bien supportés par les outils disponibles.

Les outils logiciels les plus utilisés pour le data mining sont :

- Knime, R, Weka, Orange, RapidMiner pour les logiciels libres ;
- KXEN, Excel et SQL Server (avec les modules supplémentaires Business intelligence de Microsoft), Oracle DM, SAS Entreprise Miner, STATISTICA, IBM SPSS, BrainCube, CORICO pour les logiciels commerciaux.

On peut aujourd'hui retenir trois méthodes majeures pour fouiller les données :

- Méthode **SEMMA** : créée par l'éditeur de logiciels SAS Institute, cette méthode s'applique à tous les logiciels (même ceux qui ne sont pas édités par l'entreprise SAS Institute). Elle décrit un ensemble d'outils et de méthodes logiques à mettre en place afin de fouiller au mieux une grande masse d'informations. C'est une approche technique.
- Méthode **CRISP-DM** : il s'agit de 6 étapes qui permettent d'industrialiser la fouille de données (compréhension métier, recherche de données, préparation des données, exploration et modélisation de l'information, segmentation (cœur de la méthode), évaluation et livraison/restitution).
- Méthode **DMAIC Six Sigma** : empruntée à l'industrie, la méthode met l'accent sur la recherche de défauts et l'amélioration de la qualité. Cela s'applique très bien à la fouille de données.

Allant plus loin que les simples calculs statistiques (intelligence artificielle, réseaux de neurones, calculs bayesiens, machine learning), l'exploration de données se base sur deux grandes familles d'algorithmes :

- **Méthodes descriptives** : cette façon de faire consiste à analyser les données disponibles et à les classer sans a-priori de façon à mettre en évidence des normes de comportement et à repérer des aberrations.
- **Méthodes prédictives** : la tâche consiste ici à regarder les données actuelles et à prévoir le futur. En se basant sur les éléments connus, on peut imaginer une certaine reproductibilité.

Les écueils les plus fréquents à l'usage du data-mining sont essentiellement liés à l'humain :

- Poser les bonnes questions ;
- Savoir écarter les données fausses, pouvoir constituer un ensemble de données qui a du sens et qui est vrai ;
- Faire la différence entre corrélation et causalité ;
- Ne pas confondre justesse et précision ;
- Savoir exploiter les résultats et savoir les réfuter parfois ;
- Résoudre les problèmes d'éthique et les respecter les données privées.

**Déclencheur** : Les **déclencheurs** ou **triggers** (en anglais) sont une fonctionnalité avancée très pratique des bases de données. Ils permettent de **déclencher une action lorsque certains paramètres sont réunis** dans une table.

Un déclencheur peut se produire lorsque :

- des données sont insérées (INSERT) ;
- des données sont modifiées (UPDATE) ;
- des données sont supprimées (DELETE) ;
- Juste avant (BEFORE) ou juste après (AFTER) l'action sur les données.

On se sert des triggers pour :

- **mettre à jour** d'autres données en déplaçant une partie de la logique applicative dans la base de données plutôt que dans l'outil utilisé (site web) ;
- **vérifier** les données à rentrer dans la base (contrôle avant ou après la requête) ;
- **remplir** un fichier de log afin de savoir qui a fait quoi ;
- **pallier à un défaut** d'intégrité référentielle (sur certaines bases de données qui ne le gèrent pas ou lorsque le concepteur ne l'a pas mis en place)

### Spécificités des déclencheurs

- Le déclencheur s'exécute pour toutes les lignes. Si 10 lignes doivent être affectées par un traitement, le déclencheur se lancera 10 fois.
- Il ne peut y avoir qu'un seul déclencheur par action et moment. On ne peut pas empiler les triggers. On peut par contre créer des instructions qui se suivent dans le même trigger pour réaliser des opérations multiples (Avant insertion / Après modification...).
- Les déclencheurs ne peuvent pas lancer toutes les sortes d'actions (transaction, sélection de données). Les déclencheurs ne peuvent pas non plus agir sur la table qui l'a déclenché.
- Les triggers peuvent avoir des erreurs. En cas d'erreur et lorsque la base de données le permet, les actions sont déconstruites et la requête n'est pas réalisée (principe du rollback).

**Dépendance fonctionnelle** : Toute ligne d'une table est identifiable à la valeur de sa clef primaire. La valeur des autres champs de la ligne est en quelque sorte rattachée à la valeur de la clef primaire. On dit que ces autres champs sont en dépendance fonctionnelle avec la clef primaire, c'est-à-dire qu'ils lui sont rattachés.

**Dictionnaire des données** : Le dictionnaire des données est une représentation tabulaire de la structure d'une base de données. Il consiste à énumérer tous les champs de toutes les tables d'une base de données. Il prend par exemple la forme suivante :

Champ	Type	Longueur	Vide ?	Par défaut	Description
<b>uneChaine</b>	Chaîne	10	Oui		Chaîne de car. A 10 car. maximum
<b>unEntier</b>	Entier	1	non	0	Entier compris entre -127 et 128
<b>UnDecimal</b>	décimal	1,2	non	0.00	Décimal à 5 chiffres, 2 après la virgule
...	...	...	...	...	...

*Champ* : nom du champ, à savoir dénomination de la donnée à stocker

*Type* : type de donnée du champ (chaîne de caractères, entier, décimal, date, heure, etc) ;

*Longueur* : pour une chaîne de caractères, longueur maximale de la chaîne ; pour entier, en général, le nombre d'octets sur lequel est codé l'entier ; pour un nombre décimal, le nombre de chiffres total suivi du nombre de chiffres après la virgule. ;

*Vide* : Indique si le champ peut être vide, c'est-à-dire avoir la valeur null ;  
*Par défaut* : valeur que le champ prend par défaut ;  
*Description* : précise la signification du champ à stocker.

**Données relationnelle** : Relationnelle se dit d'une base de données qui manipule des informations hiérarchisées qui sont liées entre elle par des contraintes de relations. À la différence d'autres types de bases de données, les bases de données relationnelles fonctionnent avec des tableaux à deux dimensions nommés tables qui sont liés à d'autres tables par un mécanisme de relations reposant sur l'algèbre relationnelle. Ces relations définissent et précisent les liens entre les tables.

**Entité** : Dans la méthode Merise au niveau du MCD, une entité correspond à un **ensemble homogène d'informations** qui correspondent au même « objet » à informatiser. Cette **entité** a un nom unique afin de la manipuler facilement. Plus tard dans l'analyse, l'entité se transforme en table et devient concrètement une table lors de la réalisation effective de la base de données.

Cet ensemble d'informations, **l'entité**, partage les mêmes caractéristiques et peut être manipulé au sein du système d'information mais aussi en discutant entre informaticiens et personnes du métier. L'entité « *contrat* » et l'entité « *client* » sont typiques des bases de données de gestion. Isoler, comprendre le rôle et les informations qui se trouvent dans ces coquilles est une des premières étapes de l'analyse.

Une **entité** reste cependant peu utile si elle n'est pas remplie avec des informations calibrées et si elle ne discute pas avec les autres entités de la base de données : on parle alors de propriétés et d'associations ce qui a donné le nom de modèle entités-associations que l'on croise parfois.

Dans un MCD, les entités sont représentées par des rectangles et dans le cartouche supérieur de chaque rectangle est écrit au singulier le nom de l'entité.

**Entrepôts de données** : Plus souvent rencontré dans sa version anglais data **warehouse**, les **entrepôts de données** sont des bases de données qui agrègent l'information éparpillée au sein du système d'information, les rend compatibles et exploitables dans le cadre de l'informatique décisionnelle.

Ainsi une entreprise peut avoir des données qui proviennent de tableur Excel, d'une base de données Access pour les commerciaux, de la base de données marketing au format MySQL ou encore de catalogues produits venant directement des fabricants. L'entrepôt de données a pour tâche de fournir une vision homogène et cohérente afin de pouvoir croiser les données et ressortir des informations à plus haute valeur ajoutée.

Les entrepôts de données ont aussi un rôle de mémoire. Ils doivent permettre d'historiser les données et de continuer à les interroger même si les données initiales viennent à disparaître ou être transformées. Les données rentrées dans l'entrepôt de données ne sont ainsi plus modifiables.

Un entrepôt de données réalise donc plusieurs tâches :

- Extraire les données de production
- Transformer, épurer, homogénéiser les données
- Stocker les données
- Permettre l'interrogation

**Groupement** : existe une clause bien pratique en SQL qui permet d'opérer des regroupements. Il s'agit de GROUP BY. On s'en sert pour grouper des éléments identiques.

### Table Ventes

identifiant	Date	Nom Client	Montant Vente
1	2017-01-15	ImmoPlus	1600
2	2017-01-15	CanalImmo	800
3	2017-01-15	RégieRE	750
4	2017-01-16	RégieRE	2050
5	2017-01-16	RégieRE	380
6	2017-01-17	ImmoPlus	520
7	2017-01-18	CanalImmo	990
8	2017-01-18	CanalImmo	1050

Par exemple, pour un même client dont on connaît le détail de chaque vente, on peut réaliser un groupement par client et effectuer un calcul statistique sur les ventes (somme des ventes, moyenne des ventes, nombre cumulé de ventes, vente maximum ou minimum)...

Concrètement, la requête va **regrouper les éléments identiques** et **réaliser une fonction d'agrégation sur les autres champs** choisis.

En réalisant la requête suivante, on affiche la somme des ventes par client :

```
SELECT "Nom Client", SUM("Montant Vente") FROM "Table Ventes"  
GROUP BY ("Nom Client");
```

En réalisant la requête suivante, on affiche la moyenne des ventes par jour :

```
SELECT "Date", AVG("Montant Vente") FROM "Table Ventes"  
GROUP BY ("Date");
```

Très pratique, le GROUP BY fait partie des outils de base lorsque l'on interroge les bases de données SQL et que l'on souhaite faire parler les données. Il évite d'afficher des doublons et permet d'avoir un premier niveau de données statistiques.

**Identifiant** : Avec la méthodologie Merise, on utilise les deux termes « **identifiant** » et « **clé primaire** ». Il s'agit de **la même chose** mais pas utilisé dans le même outil. On utilise le terme « **identifiant** » à l'étape du MCD. On utilise le terme « clé primaire » aux étapes du MPD, MLD et lorsque la base est créée.

Dans chaque entité, il existe des propriétés. La première des propriétés, celle **dont la valeur est unique** est l'**identifiant**. C'est grâce à l'identifiant que l'on peut identifier de manière unique les données de chaque propriété de l'entité sans se tromper. On la représente soulignée. Il n'y a qu'un seul identifiant par entité.

Dans la modélisation Merise, les entités peuvent être reliées entre elles. Les relations et les dépendances fonctionnelles reposent sur les identifiants : ce sont eux qui permettent de relier les bonnes valeurs entre elles.

**Index** : En base de données, un index est un outil de performance et un raccourci qui permet d'accéder plus rapidement aux données.

Les données collectées et stockées en base de données doivent pouvoir être interrogeables. Mais toutes les données n'ont pas le même intérêt. Certaines se retrouvent facilement et les calculs nécessaires à leur remontée sont faciles. D'autres sont plus complexes et nécessitent une puissance de calcul plus importante. Enfin, les données clés (les clés primaires, les clés étrangères ou les champs sur lesquels beaucoup d'interrogations sont réalisées) doivent être très rapidement accessibles. C'est pour ces raisons que les concepteurs de bases de données « trichent » et mettent en place des raccourcis vers ces données.

Les index ne sont pas accessibles directement et ne se voient pas dans la structure de la base de données (ils ne sont pas modélisés dans des méthodes telle que Merise et il n'est pas possible d'y accéder via SQL en lecture).

L'index est donc une construction informatique qui double les données présentent en base et qui accélère les recherches.

- C'est surtout le volume d'information qui rend l'index intéressant. Sur de petites bases de données, les gains de performances sont minimes. À l'inverse, lorsque les bases de données sont importantes, les index montrent toute leur efficacité.
- L'efficacité d'un index décroît avec sa longueur. Plus l'index est composé d'un nombre de colonnes important, plus son intérêt baisse (concrètement, l'index ressemble trop aux tables qu'il est censé représenter et le temps nécessaire à son parcours revient peu ou prou à celui de parcourir la table en direct).

Certains SGBDR permettent la création d'index à partir de données calculées.

**Intégrité référentielle** : Ce mot fait parfois peur aux débutants en base de données et il est fondamental en base de données. Lorsque j'étais étudiant, c'était un incontournable (j'ai eu cette question au BAC notamment).

L'intégrité référentielle est un mécanisme qui protège le contenu et les relations d'une base de données. Correctement construite, une base de données fait appel à des données situées dans différents emplacements (les tables). Pour que les données restent utilisables et cohérentes, il ne faut pas que l'on puisse détruire des données qui dépendent les unes des autres. C'est le rôle de l'intégrité référentielle de protéger ces relations.

Dit autrement, l'intégrité référentielle est une contrainte que le concepteur de bases de données s'impose afin de garantir l'intégrité des données stockées.

Techniquement dans une base de données relationnelle, l'intégrité référentielle vérifie qu'une valeur de clé étrangère existe bien en tant que valeur de clé primaire dans une autre table. Quand une clé étrangère est utilisée, elle doit obligatoirement exister en tant que clé primaire dans la table liée.

**Logiciels de base de données** : Pour concevoir une base de données, les professionnels plusieurs types d'outils logiciels :

- Des SGBD pour construire concrètement les bases de données ;
- Des outils de modélisation pour représenter au mieux le fonctionnement de la base de données ;

- Des outils d'administration destinés à optimiser les performances par exemple ou à réaliser des options de maintenance ;
- Des outils d'extraction et de recherche de données (data-mining).

**MCD** : Dans la méthodologie Merise destinée à créer des bases de données, il y a des outils dédiés aux traitements et aux données. Le **MCD (Modèle Conceptuel des Données)** est un des outils majeurs concernant les données.

Le **MCD** est une représentation graphique de haut niveau qui permet facilement et simplement de comprendre comment les différents éléments sont liés entre eux à l'aide de diagrammes codifiés dont les éléments suivants font partie :

- Les entités (1 rectangle = 1 objet) ;
- Les propriétés (la liste des données de l'entité) ;
- Les relations qui expliquent et précisent comment les entités sont reliées entre elles (les ovales avec leurs « pattes » qui se rattachent aux entités) ;
- Les cardinalités (les petits chiffres au dessus des « pattes »).

Utilisé assez tôt en conception de base de données, le **MCD** évoluera ensuite vers les autres outils de Merise, à savoir le MPD et le MLD.

Cet outil **permet d'échanger entre informaticiens et non-informaticiens** sur l'outil à informatiser. On peut ainsi à partir d'un MCD **valider et préciser des règles** qui s'appliqueront à la future base de données (d'après le MCD en image ci-dessus) :

- Une fiche est remplie par un et un seul chargé d'appui ;
- Un contact peut avoir différentes qualités (dans ce modèle « utilisateur, salarié, porteur de projet ») en fonction des fiches.

**Merise** : Merise est une méthode informatique dédiée à la modélisation qui analyse la structure à informatiser en termes de systèmes. Le gros avantage de cette méthode est qu'elle permet de cadrer le projet informatique et de « discuter » en se comprenant entre utilisateurs et informaticiens.

Créée dans les années 70 sur commande de l'État français et destinée aux gros projets informatiques de l'époque, la méthode a perduré jusqu'à aujourd'hui. Son utilisation très répandue en Europe constitue un socle difficilement contournable lorsque l'on s'attache à la création de bases de données.

Merise est en fait un outil analytique qui facilite la création de base de données et de projets informatique. Le principal auteur de la méthode est Hubert Tardieu qui se basa sur les travaux autour du modèle relationnel de Codd. Concrètement Merise (que l'on prononce Meurise) permet de :

- hiérarchiser les préoccupations du gestionnaire de projet informatique
- décrire le fonctionnement du système à informatiser et notamment :
  - Les données (MCD) : quelles sont les relations et les dépendances entre les différents acteurs (*client – commande – produit – fournisseur par exemple*)
  - Les traitements (MCT) : comment les acteurs travaillent-ils ensemble (*comment se passe une commande concrètement par exemple*)
- proposer une implémentation logique (MLD, MLT) du point précédent
- Proposer une construction concrète et utilisable du point précédent (MPD, MOT)

**MLD** : Le **MLD** ou **Modèle Logique des Données** est simplement la représentation textuelle du MPD. Il s'agit juste de la représentation en ligne du schéma représentant la structure de la base de données. Il n'y a pas de travail poussé à réaliser à cette étape, il s'agit juste d'appliquer quelques règles toutes simples. Cette étape est parfois omise.

On représente ainsi les données issues de la modélisation Merise sous la forme suivante :

- Chaque ligne représente une table ;
- C'est toujours le nom de la table qui est écrit en premier ;
- Les champs sont listés entre parenthèses et séparés par des virgules ;
- Les clés primaires sont soulignées et placées au début de la liste des champs ;
- Les clés étrangères sont préfixées par un dièse.

À noter que le MLD prend parfois un R et devient MLDR : le R signifiant simplement Relationnel.

**Moteur de base de données** : Dans un logiciel de base de données il y a plusieurs composants qui discutent entre eux. Certains composants discutent aussi avec l'extérieur et un des composants les plus importants est le **moteur de base de données**.

Le moteur de base de données a un rôle central. Il est le garant des opérations CRUD et vérifie, autorise et exécute les actions suivantes :

- Création de données ;
- Lecture de données ;
- Mise à jour de données ;
- Suppression de données.

Mais le moteur de base de données a d'autres tâches qui lui sont dévolues :

- La gestion des fichiers et du stockage sous-jacent ;
- L'optimisation des accès et des performances (mémoire cache, antémémoire);
- La bonne application des règles de gestion (certaines découlant de CRUD, d'autres non) comme le suivi des droits d'accès, les mécanismes de verrous, les transactions ou l'intégrité référentielle ;
- La journalisation des opérations (fichier log de traçabilité des opérations) ;
- L'automatisation de certaines opérations traitées ensemble (les procédures stockées).

**MPD** : Dans la méthodologie Merise, le **MPD (Modèle Physique des Données)** fait suite au MCD. Ensuite viendra le MLD.

L'étape de création du MPD est presque une formalité comparée à la création du MCD. En s'appuyant sur des règles simples (et qui fonctionnent à tous les coups), l'analyste fait évoluer sa modélisation de haut niveau pour la transformer en un schéma plus proche des contraintes des logiciels de bases de données. Il s'agit de préparer l'implémentation dans un SGBDR.

Concrètement, cette étape permet de **construire la structure finale de la base de données** avec les différents liens entre les éléments qui la composent. Pour la peine, on change aussi de vocabulaire :

- Les entités se transforment en tables ;
- Les propriétés se transforment en champs (ou attributs) ;

- Les propriétés se trouvant au milieu d'une relation génèrent une nouvelle table ou glissent vers la table adéquate en fonction des cardinalités de la relation ;
- Les identifiants se transforment en clés et se retrouvent soulignés. Chaque table dispose d'au minimum 1 clé dite primaire ;
- Les relations et les cardinalités se transforment en champs parfois soulignés : il s'agit de créer des « clés étrangères » reliées à une « clé primaire » dans une autre table.

C'est aussi à cette étape que les aspects performances sont pris en compte. La construction d'index ou des opérations de « dénormalisations » faciliteront la rapidité de la base de données mais pourront aussi entraîner des incohérences. Ce dernier point est à réserver aux professionnels aguerris donc.

**Modélisation** : En conception de base de données, la **modélisation** est l'étape indispensable (pour le professionnel compétent) qui permet de comprendre les processus métiers et de les transcrire informatiquement en flux de données et de traitements.

Derrière cette définition un peu rude se trouve un besoin simple : comprendre comment font les utilisateurs actuellement (avec ou sans informatique), isoler les suites d'actions, les règles de gestion du métier et traduire le tout dans un ensemble de documents à la fois compréhensible par l'utilisateur et par l'informaticien.

Afin d'être réutilisable et compréhensible par des tiers n'ayant pas participé au projet initial, la modélisation repose sur des méthodes (comme Merise par exemple. À ne pas confondre avec UML).

La modélisation est la brique indispensable qui permet de construire une base de données solide et conforme aux attentes des utilisateurs. Cette étape est préalable à la construction concrète de la base de données et la construction des différents écrans que manipuleront les utilisateurs.

**Occurrence** : En base de données, on distingue le modèle d'une côté (la structure avec les entités et les relations) et d'un autre côté le contenu (il faut bien les remplir avec des valeurs ces beaux modèles).

Et bien, une **occurrence**, c'est tout simplement une « **ligne** » de valeurs. Dans une entité, une occurrence correspond à l'ensemble des valeurs des propriétés rattachées à un seul identifiant. Dans une relation, une occurrence correspond à l'ensemble des valeurs des propriétés de la relation (représenté par les clés de chaque entité liée) : on l'appelle alors une **occurrence de relation**.

cli_numauto	cli_nom	cli_prenom	cli_datenaissance
1	DUPONT	Frédéric	11/02/1973
2	DURAND	Marie	31/09/1989
3	NONYME	Alain	2/11/1977
4	BERTRAND	Paul	4/08/1990

**NoSQL** : **NoSQL** (ou **Not Only SQL** ou encore **Non-relational SQL**) est une façon de concevoir les bases de données qui repose non plus sur la notion relationnelle des SGBD (des tables avec des liens entre elles) mais sur une architecture plus simple à manipuler.

Initialement, le besoin est apparu avec les très très grosses bases de données qui avaient des problèmes de rapidité et de cohérence (voir tests ACID). En effet, lorsqu'il y a des milliards de données à traiter, les performances peuvent se dégrader de manière significative. Et lorsqu'il faut s'assurer que les données sont bien synchronisés en temps réel sur différents lieux aux quatre coins du globe, des problèmes insurmontables apparaissent (connu par exemple sous le nom de théorème CAP).

Le principe du NoSQL est donc de **simplifier au maximum pour améliorer les performances**. Cette simplification entraîne des effets de bords indésirables. Les relations très structurées et permettant de représenter des interactions complexes sont délaissées au profit de modèles plus simples de représentation : les informaticiens nomment cela un **tableau associatif** que l'on peut comparer à un immense tableur composé d'une clé (la première colonne) et de valeurs qui en découlent (la ou les autres colonnes).

NoSQL ne s'oppose donc pas à SQL et il est fréquent de retrouver les deux dans les systèmes informatiques d'une certaine ampleur. On utilise alors SQL et les bases de données relationnelles pour modéliser finement et NoSQL pour des besoins plus simples mais où la rapidité est cruciale.

Effet de mode à causes de leur utilisation par les ténors du web, le NoSQL est encore très jeune et le marché est loin d'être mature : les prémices de NoSQL date de 1998 mais le coup d'envoi de cette technologie ne date réellement que de 2009. Pourtant, on trouve déjà sur le marché différents outils spécifiques ou compatibles NoSQL : Cassandra, Voldemort, MongoDB, SimpleDB et CouchDB notamment.

**OLAP** : OLAP signifie « Online Analytics Processing » que l'on peut traduire en français par « Traitement analytique en ligne ». Mais ce n'est pas plus explicite pour autant.

Concrètement, on se sert des d'applications OLAP lorsque l'on a besoin de croiser des données et d'obtenir des informations synthétiques. Beaucoup utilisé en analyse financière notamment, les applications OLAP se situent en dehors des applications classiques. Les entreprises fonctionnent très bien sans elles. Par contre, lorsque le besoin de s'élever un peu plus ou de sortir la tête du guidon se fait sentir, les applications OLAP permettent de faire parler des données diverses. Ces applications sont alors placées au-dessus des autres logiciels et permettent d'interroger les données issues de ces derniers.

Un cube OLAP (c'est le nom que l'on donne aux applications construites sur bases de données OLAP) permet plusieurs opérations pour aller fouiller dans les données :

- Sélection d'un couple de dimension (rotate)
- Extraction d'un bloc de données (scope)
- Extraction d'une tranche d'information – plus précise que l'action précédente (slice)
- Zoom (drill-down)
- Zoom inverse – on s'éloigne (drill-up)
- Accès aux informations de base en cas de données agrégées (drill-through)

C'est le langage MDX qui est utilisé (et non pas le langage SQL comme pour les bases de données classiques) pour interroger les hypercubes OLAP mais l'EIS et des outils sur mesure existent aussi. À noter qu'Excel via des extensions particulières peut lui aussi accéder à ses données.

C'est un peu l'inverse de l'OLTP qui lit et écrit en temps quasi-réel.

**OLTP** : OLTP est l'acronyme d'OnLine Transaction Processing. Ce terme s'applique opérations logicielles qui lisent, écrivent de nouvelles informations et apportent des modifications en temps réel à une base de données.

Concrètement lorsqu'un internaute commande un billet d'avion, réalise un virement, réserve une place de spectacles par exemple, l'OLTP est indispensable. Cela permet d'être toujours certain d'avoir des données justes et à jour dans le système informatique.

La difficulté est ici triple avec les systèmes OLTP :

- Permettent l'accès simultané par plusieurs centaines de personnes ;
- Minimiser le plus possible les arrêts (il faut bien faire des opérations de maintenance) ;
- S'assurer que les données sont à jour.

Aussi pour pallier à ces défis, l'OLTP est souvent :

- Décentralisé afin d'être accessible plus facilement peu importe la localisation géographique ;
- Distribué sur plusieurs machines fonctionnant ensemble afin de permettre la maintenance sur une sans arrêter l'ensemble du système.

Le temps réel n'existe pas en informatique mais les opérations sont exécutées très très rapidement et des verrous et autres mécanisme de sécurité assurent, si besoin est, la cohérence, l'intégrité et la justesse des données et des traitements.

**OQL** : L'Object Query Language est le SQL adapté aux bases de données objet.

Développé par le groupe ODMG dès 1991, l'OQL n'est pas encore implémenté entièrement dans les bases de données à cause de sa grande complexité.

La grammaire et le vocabulaire ressemble beaucoup au SQL.

Le résultat des requêtes sont des objets et utilisent le polymorphisme de la Programmation Orientée Objet.

**ORM** : Le terme **ORM** vient de l'anglais **Object-Relational Mapping**. En français, on croise parfois (rarement) le terme **mapping objet-relational**.

Il s'agit d'une technique de programmation informatique qui permet de simplifier l'accès à une base de données en proposant à l'informaticien des « objets » plutôt que d'accéder directement à des données relationnelles. Ce niveau d'abstraction supplémentaire fait correspondre le monde objet (programmation orientée objet) et le monde relationnel (les bases de données relationnelles classiques et massivement utilisées aujourd'hui).

L'intérêt principal est d'éviter beaucoup de code très similaire pour l'informaticien ce qui est source d'erreurs et de temps perdu. Ainsi plutôt que de devoir recoder à la main les fonctions de base

(CRUD – création, lecture, modification, suppression), les informaticiens ajoutent une couche logicielle intermédiaire qui se charge de la « traduction ».

La difficulté principale consiste à trouver la « bonne traduction » pour toutes les opérations possibles. Souvent, ce n'est pas possible et il faut faire des concessions.

**Procédure stockée** : Une **procédure stockée** (*stored procedure*) est une suite d'instructions SQL qui s'exécutent les unes à la suite des autres et qui est conservée de façon durable (1 jour, 1 mois, 1 an, il n'y a pas de limite). Pour faire une analogie avec Excel, on pourrait dire qu'il s'agit d'une macro qui répète un ensemble de demandes simples préalablement enregistrées.

Une fois créée, on retrouve donc la procédure stockée parmi les autres objets de la base de données (comme une table) et on peut l'appeler pour l'exécuter.

À l'intérieur d'une procédure stockée, on peut utiliser des conditions et des boucles (IF, WHILE, LOOP, REPEAT, CASE, ITERATE...) ce qui permet d'effectuer des opérations beaucoup plus riches sur les données directement dans la base.

Tous les SGBD ne permettent pas l'utilisation de procédures stockées (MySQL ne le permet que depuis la version 5).

```
1 DELIMITER
2 CREATE PROCEDURE afficher_client_selon_region (IN region_id INT)
3 BEGIN
4     SELECT numauto, nom, region
5     FROM Client
6     WHERE region = region_id;
7 END |
8 DELIMITER ;
```

**Propriété** : En conception de base de données avec Merise, il y a des termes un peu barbares. **Propriété** est un de ces termes.

- Une **propriété** est un « champ » de l'entité. Dans l'exemple ci-dessus, par exemple, on a l'entité « séjour » qui contient des propriétés qui commencent par « sej\_ ». Chaque propriété donne une information précise les données stockées liées à l'entité. Le contenu de chaque propriété varie pour chaque occurrence. Ainsi « sej\_datedebut » et « sej\_datefin » varieront d'un séjour à l'autre ;
- Chaque entité possède au moins une propriété ;
- La première propriété d'une entité est un numéro unique, l'identifiant ;
- Ces propriétés sont dites « **atomiques** » c'est à dire que c'est la donnée la plus fine que l'on trouve dans la base (il n'y a pas de sous-données aux propriétés) ;
- Les propriétés se retrouvent aussi dans certaines associations (les CIM). Ici, c'est le cas avec « passecommande » qui contient 3 propriétés et avec « estreferent » qui contient 1 propriété ;

- On dit que les entités et les CIM sont « **porteuses** » de propriétés.
- On ne doit pas avoir de propriétés déductibles (c'est à dire qui se calculent). L'exemple « `sej_duree` » est un faux ami ici : il ne calcule pas la durée d'un séjour mais correspond à un libellé propre à l'entreprise (en fait un arrondi sur mesure de la durée réel afin d'être simple et parlant) ;
- Par convention et pour simplification, on utilise des préfixes aux propriétés. Ça permet de s'y retrouver.

Une propriété est parfois nommée « attribut » ou « caractéristique ».

**Propriétés ACID** : Abréviation pour Atomicity, Consistency, Isolation et Durability. En français Atomicité, Cohérence, Isolation et Durabilité.

Il s'agit des propriétés nécessaires pour garantir une transaction informatique. Très utiles dans les bases de données lors d'opérations critiques (par exemple lors de transferts d'argent), les **propriétés ACID** se matérialisent par les fonctions de transactions. Toutes les bases de données ne mettent pas en place ces propriétés.

Conceptualisés en 1983 par Jim Gray dans les années 1970, les propriétés ACID doivent leur nom à Reuter et Härder.

- **Atomicité** : Tout ou rien. Soit l'opération se fait en entier, soit elle ne se fait pas du tout. La notion d'atomicité sous-entend la possibilité de défaire une opération avortée.
- **Cohérence** : L'opération doit assurer que la base de données sera dans un état valide après l'opération.
- **Isolation** : L'opération doit se faire en toute autonomie sans dépendance à une autre opération.
- **Durabilité** : En cas de problème important (coupure d'électricité), les modifications apportées sont bien enregistrées.

**Requête** : Une **requête** est une opération effectuée sur une base de données. Réalisée en langage SQL, elle permet :

- de rechercher une information au sein d'une base de données en imposant des conditions et en effectuant des filtres et des tris. Le résultat est alors présenté de façon tabulaire.
- d'agir sur la structure de la base de données en y ajoutant / modifiant / supprimant des éléments et en y réalisant des opérations de maintenance.
- d'agir sur les données (ajout / modification / suppression).

**Schéma relationnel** : Le schéma relationnel est représentation qui permet de mettre en évidence les relations entre les données. Ce schéma a une syntaxe, un formalisme, qui lui est propre. Et vous devez connaître. Dans un schéma relationnel, les tables sont appelées « relations » et les champs appelés « attributs ».

Exemple :

Facture(numFacture, #client, dateFacture, montantHT, montantTVA,...)

On note que la clé primaire est numFacture doit toujours être souligné et la clef étrangère est #client toujours précédés d'un #. Facture est le nom de la table.

Les relations entre tables sont dictées par les clefs primaires et clefs étrangères.

- a) **Relation « un à plusieurs »** : entre tables, il peut exister un ou plusieurs liens dits « de un à plusieurs » (one to many). Dans l'exemple ci-après, une facture est par exemple associée à un client. A une facture correspond par conséquent un et un seul client. Inversement, à un client peuvent correspondre aucune à plusieurs factures.

Table « Facture »					Table « Client »		
numFacture	client	dateFacture	montantHT	...	numClient	nomClient	...
1	1	05/01/2020	500,00		1	Nestle	
2	2	05/01/2020	750,00		2	Vinci	
3	2	06/01/2020	25000,00		...	...	...
4	1	06/01/2020	600,00				
...	...	...	...				

On comprend que le « 2 » figurant le champ « client (clef étrangère) » de la table champ renvoie vers le client « 2 » de la table « Client », c'est-à-dire Vinci.

Une clef étrangère fait toujours référence à une clef primaire.

- b) **Relation « plusieurs à plusieurs »** : Entre tables, il peut également exister un ou plusieurs liens dits « de plusieurs à plusieurs » (many to many). Par exemple, un salarié peut travailler sur plusieurs projets. Inversement, plusieurs salariés peuvent travailler sur le même projet. On obtiendra un schéma relationnel de la forme suivante :

Salarie(numSalarie, prenomSalarie, nomSalarie,...)

*Clef primaire : numSalarie*

Equipe(#salarie, #projet)

*Clef primaire : salarie, projet*

*Clefs étrangères :*

- *Salarie référence au champ numSalarie de la relation Salarie*
- *Projet en référence au champ numProjet de la relation Prijet*

Projet(numProjet, dureeProjet, prixProjet,..)

*Clef primaire : numProjet*

**SGBD** : Abréviation de **Système de Gestion de Base de Données**

Noté aussi SGBDD ou DBMS (DataBase Management System en anglais), il s'agit du logiciel qui permet de créer une base de données. Non utilisable directement par l'utilisateur final, le SGBD est un logiciel intermédiaire employé par le concepteur pour se soulager de toutes les tâches indispensables inhérentes au fonctionnement de sa base de données.

Ainsi, le SGBD stocke et partage les données que saisisent les utilisateurs. En arrière plan, le SGBD s'assure que les données sont correctement enregistrées, qu'elles sont diffusées aux bons

utilisateurs, qu'elles sont présentées de façon adéquate... Le SGBD se charge aussi des tâches d'enregistrement, d'accès simultanés et de sauvegardes.

Le concepteur de la base de données choisit son SGBD en fonction de ses caractéristiques, des usages imaginés et des volumes de données à traiter.

**SQL** : Abréviation de **Structured Query Langage**. En français, langage de requête structuré.

Le **SQL** est le langage informatique qui permet de définir et de manipuler les bases de données. En SQL, on peut par exemple interroger une base de données, ajouter, modifier et supprimer des données. Il est aussi possible de réaliser des opérations lourdes (gestion de tables) ou de maintenance / performance (gestion des index...).

Le SQL a été conçu pour manipuler les données de bases de données relationnelles et se base sur l'algèbre relationnelle. Inventé en 1974 d'après les recherches de Edgar Frank Codd qui permirent la naissance des SGBD relationnels modernes, le SQL est aujourd'hui incontournable dans le monde des bases de données.

Standardisé depuis 1986, le SQL existe en différentes versions. Chaque système de gestion de base de données prend en compte tout ou partie des spécifications standardisées. Certains SGBD ajoutent leurs propres jeux d'instructions.

**Table** : Une table est un des éléments fondamentaux des bases de données. Il s'agit d'un **ensemble homogène de données agencées sous forme de tableau** suivant un format particulier et constitué de lignes et de colonnes. Pour les bases de données relationnelles, une table contient des colonnes correspondant aux différents champs connus à l'avance et dont la première ligne explique le contenu (un peu comme la première ligne d'un tableau qui énonce le contenu à venir en dessous).

**Transactions** : En base de données, une **transaction** est un ensemble d'opération qui doit être effectué en respectant les propriétés ACID d'atomicité, de cohérence, de durabilité et d'isolation.

Par exemple lorsqu'une transaction bancaire a lieu (un virement de 100 euros entre un compte A et un compte B), la transaction assure (et rassure) que le virement s'est bien déroulé correctement. Il n'y a pas eu d'erreur, tout a bien été enregistré, les 2 comptes ont bien pris note des mouvements et n'ont pas engendré d'erreur. Si un problème est détecté durant l'opération, la transaction est déconstruite et les données sont restaurées dans leur état initial.

Les méthodes **Commit** et **Rollback** sont les 2 moyens utilisés pour valider définitivement ou détricoter une transaction.

C'est grâce au mécanisme des transactions que l'on est certain que l'opération s'est bien faite en entier (ou pas du tout). Dans les deux cas, les états sont propres, valides, documentés et cohérents.

**Types de données** : Chaque champ d'une table a un type. Les types de données courants sont :

- alphanumérique (=chaîne de caractères). On distingue souvent les chaînes dont le nombre de caractères est fixé ou variable. On peut communément imposer un nombre maximum de caractères ;
- numérique : On distingue par exemple les sous-types suivants : nombre entier ou nombre décimal ;

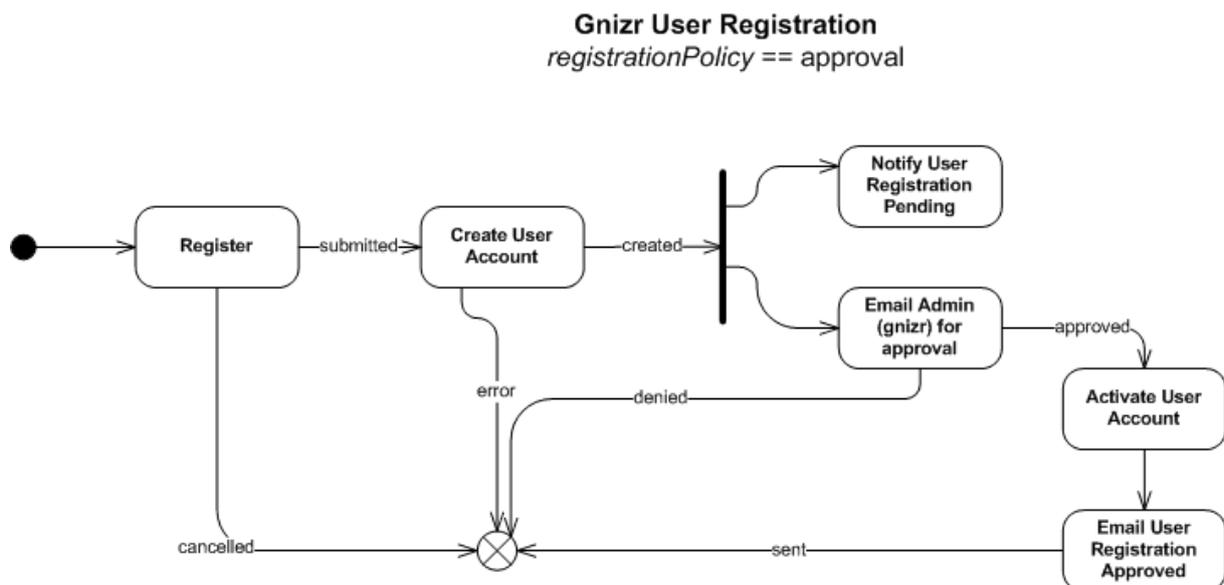
- date et heure : On distingue par les sous-types suivants : date seule, heure seule, date et heure. On notera qu'il existe de nombreuses façons de représenter les dates et/ou les heures.
- Booléen : On rappelle qu'un booléen permet préciser une information qui n'a que deux valeurs possibles : true/false, 0/1 ou vrai/faux en français.

**UML** : Abréviation d'**United Modeling Language**. Langage de Modélisation unifié en français.

**UML** est un outil souvent utilisé en développement informatique. Il permet via des outils graphiques de représenter le fonctionnement d'un outil informatique (c'est particulièrement vrai en Programmation Orientée Objet ou UML permet de modéliser directement les objets qui seront ensuite manipulés par le programme informatique).

Outil plus récent que Merise, UML propose une petite quinzaine de diagrammes qui ne remplacent pas Merise. Les 2 outils ont des objectifs différents :

- Merise est un ensemble cohérent pour construire des bases de données. C'est une méthode idéale de modélisation pour construire une base de données relationnelle.
- UML ne propose pas de démarche ni d'organisation (ce n'est pas une méthode). C'est un outil idéal pour la conception logiciel dans un langage de type objet.
- Les 2 peuvent fonctionner indépendamment ou de concert.



**Verrou SQL** : En SQL, un **verrou** de base de données est un **mécanisme de sécurité**. Il permet d'interdire ou de limiter l'accès à un élément d'une base de données. Ça peut être une ligne d'une table ou une table entière. L'utilité des verrous est simple : conserver l'ACIDité de la base.

#### Verrous SQL : 2 exemples concrets

Imaginons le cas d'un site de vente en ligne de billets d'avions avec un internaute qui souhaite réserver un vol à bord d'un avion. La réservation de la place se fera au dernier moment : c'est à dire lorsque la personne a validé sa commande et est en train de saisir sa CB et d'attendre que la banque veuille bien donner son feu vert. À ce moment, pour protéger l'utilisateur et lui permettre de finaliser

sa transaction, il faut que le site web empêche les autres utilisateurs de réserver le même siège en même temps.

Pour cela, un **verrou en écriture** peut être posé sur le siège sélectionné pendant 3 minutes. Si au bout de 3 minutes, l'internaute n'a pas réalisé l'achat, le verrou peut être libéré et la vente peut être faite avec un autre internaute.

Autre exemple avec des statistiques. Imaginons qu'une base de données soit utilisée en temps réel par des utilisateurs tout autour du globe et qu'en même temps, il faille réaliser des calculs un peu lourd (15 minutes) de statistiques. Si pendant le calcul des utilisateurs viennent ajouter / modifier / supprimer des données, le programme de statistiques se retrouvera avec des données différentes entre le début de son calcul et sa fin.

Dans ce cas, on peut imaginer **verrouiller les données d'une table en écriture et ne conserver que l'accès en lecture**.

### Verrous en base de données

Derrière cette **notion de verrou** se cache une autre notion, importante pour les bases de données : la **concurrence**. Autrement dit, qui et combien de personnes peuvent-elles accéder en même temps aux mêmes données sans compromettre les données de la base.

Les verrous SQL sont utilisés de façons très différentes d'un [SGBD](#) à l'autre et même entre différentes versions du même SGBD. Lorsqu'il y a des [transactions SQL](#) (une suite logique d'opérations qui doivent être réalisées d'une seule traite), les verrous sont nécessairement utilisés. On utilise alors les instructions COMMIT et ROLLBACK pour libérer les verrous après avoir validé/invalidé la suite d'opérations SQL.

- On utilise l'instruction LOCK TABLES en SQL pour verrouiller une ou plusieurs tables ;
- On utilise UNLOCK TABLES pour déverrouiller une ou plusieurs tables ;
- On peut réaliser des verrous exclusifs (EXCLUSIVE), partagés (SHARE) que ce soit sur une table ou sur une ligne.

**Vue** : Une base de données n'a pas besoin d'interface graphique pour fonctionner. Et c'est très perturbant lorsque l'on débute avec ces outils. Alors il existe des outils variés qui permettent de représenter la structure d'une base de données (le MCD de Merise) et les données qui se trouvent à l'intérieur (PhpMyAdmin pour MySQL par exemple). Mais cela ne suffit pas et ne doit pas être confondu avec les vues.

Pour des raisons de rapidité d'accès et de simplification, il existe les **vues**. Les vues sont des requêtes préparées à l'avance dont les paramètres (mais pas les résultats) sont stockés dans la base de données. Les interrogations complexes sont ainsi optimisées : en enregistrant un modèle plutôt qu'en le préparant de A à Z à chaque fois, l'utilisateur va **plus vite** et fait **moins d'erreurs**.

Concrètement, la base de données stocke la requête qui peut ensuite être appelée (une vue a un nom) et exécutée. Grâce à ce nom, on peut ensuite les utiliser tels des tables dans les requêtes. C'est pratique, rapide et génère des requêtes plus lisibles tout en masquant la complexité sous-jacente.

Les vues ont d'autres caractéristiques :

- Une vue peut masquer certaines données. Étant une requête, la vue n'affiche que les informations souhaitées. Elle permet ainsi d'extraire des données choisies et de masquer des données non souhaitées.
- Certains SGBD permettent de stocker les résultats des vues et peuvent faire office de mémoire tampon (cache).
- Les vues ne prennent pas de place. C'est la requête qui est enregistrée (pas le résultat).
- On peut réaliser des regroupements de données (somme, moyenne...) qui peuvent être réutilisés très facilement.
- Les vues sont habituellement en lecture seule (d'où le terme de « vue ») mais dans certains systèmes et sous certaines conditions, on peut manipuler les données d'une vue

### Création d'une vue SQL

```
CREATE VIEW ToutesLesFactures AS
SELECT fac_id, fac_ref, cli_nom
FROM facture, client
WHERE facture.cli_id = client.cli_id ;
```

### Utilisation d'une vue SQL

Plutôt que de retaper la requête ci-dessus avec sa jointure, on peut directement appeler la vue :

```
SELECT * FROM ToutesLesFactures ;
```




---

## Lexiques de la Data Science

**3V** : Volume, Vitesse, Variété : ce sont les trois dimensions retenues, à l'origine par IBM, pour décrire un environnement Big Data. Le volume représente la quantité d'information stockée et/ou traitée par le système, la vitesse rend compte des besoins d'une réponse rapide et la variété souligne les nombreux types de données que l'on peut avoir en entrée.

**4V** : Pareil que 3V avec en plus Vérité : les données utilisées ne sont pas nécessairement à jour, ni même correctes (il y a beaucoup de désinformation sur Internet). La problématique est donc à considérer pour chaque système Big Data étudié.

**Algorithme (Algorithm)** : Un algorithme est un ensemble d'instructions que nous donnons à un ordinateur pour qu'il puisse prendre des valeurs et les manipuler sous une forme utilisable. Cela peut être aussi simple que de rechercher et de supprimer toutes les virgules d'un paragraphe ou aussi complexe que de résoudre une équation aux dérivées partielles.

**Algorithme adaptatif (Adaptive algorithm)** : C'est un algorithme dont le comportement change en fonction de paramètres variés (données entrantes, paramètres d'exécution, propriétés du serveur ou de la mémoire, etc...).

**Algorithme génétique (Genetic algorithm)** : Inspiré des principes de l'évolution génétique, des générations de population sont utilisées pour confronter l'algorithme à un certain environnement

dans lequel un « optimum de survie » finira par apparaître. *Cette famille de techniques s'inspire de la théorie Darwinienne de l'évolution pour résoudre des problèmes d'optimisation.*

La théorie de l'évolution présentée par Charles Darwin en 1859 dans son livre "l'Origine des espèces", repose sur trois principes clés : l'hérédité, la variation et la sélection.

Ainsi, lorsqu'un organisme vivant se reproduit, il transmet ses caractéristiques à ses descendants, à travers ses gènes. Cependant l'hérédité des caractères n'est pas parfaite, les gènes peuvent subir des mutations et les descendants peuvent donc présenter des variations de caractères. Par conséquent, les différents individus appartenant à une population d'organismes vivants ne sont pas tous identiques, et certains d'entre eux peuvent avoir des variations qui leur permettent de mieux survivre et de se reproduire davantage dans un certain environnement. Ces individus ont donc un avantage sélectif, plus de chances de transmettre leurs gènes à la génération future. À travers ce processus, les organismes s'adaptent à leur environnement, au cours des générations.

Les algorithmes génétiques, et les algorithmes évolutionnaires en général, reposent sur ces principes. Comment fonctionnent-ils ? Tout d'abord, on cherche à représenter les solutions possibles du problème d'optimisation sous la forme d'un génome, puis on génère une population de solutions potentielles au problème donné, on sélectionne ensuite les solutions les plus performantes vis à vis de la tâche à optimiser, on crée alors une nouvelle population en copiant à l'identique les solutions sélectionnées, enfin on applique des opérateurs de variation aux génomes des individus de la nouvelle population, afin de créer des solutions différentes. La population d'enfants devient à son tour la population parentale, et on itère la même procédure jusqu'à ce qu'une solution satisfaisante soit trouvée.

En pratique il existe un grand nombre de façons de représenter les solutions potentielles à un problème, les codages les plus courants sont par exemple les vecteurs de nombres (binaires, entiers ou réels) et les graphes (par exemple des structures analogues aux arbres de décision). De même, il existe différentes méthodes de sélection (tournoi, par rang, uniforme...) et de variation (mutations ponctuelles, cross-over), et le bon choix de ces différents aspects est crucial pour obtenir des résultats pertinents.

Ces algorithmes permettent d'explorer l'espace de solutions possibles de manière non exhaustive, afin d'obtenir une solution satisfaisante. Par conséquent, ils sont particulièrement utiles dans le cas d'espaces de très grande taille, présentant des optima locaux, difficiles à explorer avec des algorithmes déterministes d'optimisation. De plus, ces algorithmes s'adaptent facilement à des espaces de données qui changent dans le temps. A contrario, leur principal inconvénient est que leurs résultats dépendent fortement du choix des différents éléments qui les constituent ainsi que des paramètres associés.

**Analyse de données :** *Cette famille de méthodes d'apprentissage permet de dégager les aspects les plus intéressants dans la nature et la structure d'un nombre très important de données. Certaines méthodes permettent de faire ressortir des relations entre les données et de décrire de façon plus succincte les principales informations qu'elles contiennent.*

D'autres techniques permettent de regrouper les données de façon à faire apparaître clairement ce qui les rend homogènes. Toutes ces méthodes ont pour principales caractéristiques d'être multidimensionnelles et descriptives. On distingue les méthodes d'apprentissage non supervisé (segmentation, détection d'anomalie, clustering, etc.) des méthodes d'apprentissage supervisé (régression, classification...). Les premières fournissent des informations sur la structure ou la distribution des données. Quant aux méthodes supervisées, elles s'intéressent aux relations et à la

dépendance des variables entre elles, mettant en évidence le lien entre la variable d'intérêt et les autres variables observées

**API (Application Programming Interface)** : Appelée également Interface de programmation, elle permet simplement à une application d'accéder à une autre application pour des données, des fonctionnalités.

**Apprentissage non supervisé** : *A la différence de l'apprentissage supervisé, le contexte non supervisé est celui où l'algorithme doit opérer à partir d'exemples non annotés.*

Il doit faire émerger automatiquement les catégories à associer aux données qu'on lui soumet pour reconnaître qu'un chat est un chat, une voiture, une voiture comme sont capables de le faire les animaux et les humains. Le problème d'apprentissage non supervisé le plus fréquent est la segmentation (ou clustering) où l'on essaie de séparer les données en groupes (catégorie, classe, cluster...) : regrouper des images de voitures, de chats, etc. Beaucoup d'espoirs sont portés sur la détection d'anomalies pour la maintenance prédictive, la cybersécurité, mais aussi le dépistage précoce de maladies, etc.

De manière générale, l'algorithme cherche à maximiser d'une part l'homogénéité des données au sein des groupes de données et à former des groupes aussi distincts que possible : selon le contexte, on choisit d'utiliser tel ou tel algorithme pour classer les données par exemple selon leur densité ou leur gradient de densité. Dans le cas de la détection d'anomalies, c'est plutôt le caractère extrême ou atypique des valeurs ou d'un pattern dans les données qui est recherché. La métrique sous-jacente joue un rôle clé pour déterminer ce qui est la norme et ce qui s'en éloigne.

**Apprentissage par imitation** : Apprentissage par imitation (ou Apprentissage par renforcement inverse) : *Cette forme d'apprentissage automatique utilise l'expérience des experts pour apprendre, par imitation.*

En apprentissage par renforcement (*Reinforcement Learning* ou RL), l'algorithme essaie de trouver la meilleure stratégie pour atteindre un objectif en interagissant avec l'environnement et en obtenant des « récompenses », façon de qualifier la qualité des actions prises par l'algorithme. En apprentissage par renforcement inverse (*Inverse Reinforcement Learning* ou IRL), l'algorithme observe un « expert » résoudre le problème et essaie d'apprendre à faire aussi bien que lui (voire mieux). Cet expert peut être une personne ou un algorithme qui sait, a priori, résoudre le problème et peut multiplier les exemples. Cet apprentissage aussi qualifié d'apprentissage par imitation (*Apprenticeship Learning*) a l'avantage de ne pas nécessiter de définir de récompenses, problème compliqué et crucial en apprentissage par renforcement.

Prenons un exemple bien connu de problème résolu en utilisant l'IRL : les véhicules autoguidés, ces robots qui se déplacent de façon autonome sans intervention humaine comme on en trouve beaucoup en manutention dans l'industrie ou pour la logistique. Il est facile de faire piloter ces robots par un homme (l'expert) alors qu'il est très difficile de définir des récompenses, autrement dit de qualifier la qualité de la conduite, la vitesse, le positionnement idéal, etc. qui dépendent de multiples paramètres.

En revanche, l'IRL suppose d'avoir accès et d'exploiter de nombreux exemples fournis par l'expert, ce qui peut s'avérer coûteux, en temps, en argent ou en puissance de calcul. Aujourd'hui, les chercheurs ont résolu la plupart des problèmes « d'imitation », grâce à des algorithmes efficaces. Ils travaillent désormais sur des algorithmes qui combinent RL et IRL pour tenter d'obtenir le meilleur des deux modèles d'apprentissage.

**Apprentissage par renforcement :** *L'apprentissage par renforcement (RL pour Reinforcement Learning) fait référence à une classe de problèmes d'apprentissage automatique, dont le but est d'apprendre, à partir d'expériences successives, ce qu'il convient de faire de façon à trouver la meilleure solution.*

Dans un tel problème, on dit qu'un « agent » (l'algorithme, au sens du code et des variables qu'il utilise) interagit avec « l'environnement » pour trouver la solution optimale. L'apprentissage par renforcement diffère fondamentalement des problèmes supervisés et non supervisés par ce côté interactif et itératif : l'agent essaie plusieurs solutions (on parle « d'exploration »), observe la réaction de l'environnement et adapte son comportement (les variables) pour trouver la meilleure stratégie (il « exploite » le résultat de ses explorations). Un des concepts clés de ce type de problèmes est l'équilibre entre ces phases d'exploration et d'exploitation. Cette méthode est particulièrement adaptée aux problèmes nécessitant un compromis entre la quête de récompenses à court terme et celle de récompenses à long terme. Parmi les exemples de problèmes traités de cette façon, on peut évoquer : apprendre à un robot à marcher en terrain difficile, à conduire (cas de la voiture autonome) ou à accomplir une tâche spécifique (comme jouer au jeu de go), piloter un agent à travers un labyrinthe, etc. Les principales familles de problèmes d'apprentissage par renforcement sont les algorithmes de bandits, les problèmes de décisions (partiellement) markovien et les arbres de jeu.

**Apprentissage supervisé :** *Ce cadre de machine learning part du fait que les données historiques (ou exemples) sont annotées.*

Prenons le cas de la reconnaissance d'objets : un problème supervisé correspond au cas où le label « voiture » est bien associé, en base, à des photos de voitures, le label « chat » à des photos de chat, etc. L'algorithme apprend ainsi à partir de milliers ou de millions d'exemples étiquetés : il cherche la relation qui permet de relier les images aux labels. Après avoir classifié correctement les exemples, il peut ensuite généraliser ce classement à de nouvelles données : classifier correctement des images de voiture ou de chat qu'il n'a jamais vues durant la phase d'apprentissage. C'est ce qu'on appelle la capacité de généralisation. Dans un cadre business, on parle souvent d'analyse prédictive. Parmi les exemples d'applications, citons la classification d'email en spam ou non selon le contenu du message, son expéditeur, son sujet..., le diagnostic médical selon les symptômes, etc.

**Approche probabiliste fréquentiste :** *Ce terme générique englobe tous les algorithmes de machine learning basés sur l'approche fréquentiste (par opposition à l'approche bayésienne). Ces méthodes (comme le Kernel Ridge Regression ou le k- Nearest Neighbour,) reposent généralement sur la « loi des grands nombres » et la théorie des « inégalités de concentration ».*

Considérons un jeu de données et un ensemble possible de distributions au sens probabiliste, à savoir les valeurs que peuvent prendre les variables aléatoires et à quelle fréquence. Deux approches probabilistes permettent de classer les variables : l'approche bayésienne et l'approche fréquentiste. La première, que l'on qualifie parfois de théorique ou déductive, combine l'information apportée par les données avec les connaissances a priori provenant soit d'études antérieures soit d'avis d'experts, dans le but d'obtenir une information a posteriori. L'approche fréquentiste, quant à elle, repose sur les observations et consiste à trouver la distribution la plus probable au vue des données, et éventuellement son intervalle de confiance correspondant (c'est à dire l'ensemble des distributions qui ont une chance significative d'être la vraie distribution).

Prenons l'exemple d'une pièce de monnaie et de la probabilité qu'elle tombe sur pile ou face. L'approche fréquentiste se basera sur l'expérience présente (par exemple, la pièce, lancée 10 fois, est tombée 6 fois sur pile) pour établir que la « vraie » probabilité d'obtenir pile. Autrement dit 6/10

soit 0,6 dans cet exemple. Conformément à la « loi des grands nombres », en lançant la pièce un nombre important de fois, cette méthode convergera, au sens mathématique du terme, vers 0,5. Concrètement, on utilise des « inégalités de concentration » pour quantifier l'incertitude du résultat autour de bornes de probabilités. Ainsi, en utilisant l'inégalité d'Azuma-Hoeffding, on sait que si on lance une pièce équilibrée 10 000 fois, la proportion de pile sera comprise entre 0,483 et 0,517 avec une probabilité supérieure à 99 %.

Généralement, du point de vue théorique, l'approche fréquentiste ne nécessite que des hypothèses assez faibles : bien que dans la plupart des cas, on travaille sous l'hypothèse de « variable aléatoire sous-gaussienne », de nombreux travaux ont montré que l'approche fréquentiste peut fonctionner avec des prérequis très faible, comme l'existence d'un « moment d'ordre 2 » [1]).

Enfin, l'une des difficultés majeures des approches fréquentistes est le problème du surapprentissage, qui est généralement contourné en utilisant une régularisation, un outil à la fois riche, flexible mais potentiellement complexe à utiliser. À noter que le débat sur les mérites relatifs entre approche fréquentiste versus bayésienne, qui date des débuts de l'apprentissage automatique, reste toujours d'actualité dans la communauté.

**Arbre de décision :** *Cet outil d'aide à la décision ou d'exploration de données permet de représenter un ensemble de choix sous la forme graphique d'un arbre. C'est une des méthodes d'apprentissage supervisé les plus populaires pour les problèmes de classification de données.*

Concrètement, un arbre de décision modélise une hiérarchie de tests pour prédire un résultat. Il existe deux principaux types d'arbre de décision :

- Les arbres de régression (Regression Tree) permettent de prédire une quantité réelle, une valeur numérique (par exemple, le prix d'une maison ou la durée de séjour d'un patient dans un hôpital) ;
- Les arbres de classification (Classification Tree) permettent de prédire à quelle classe la variable de sortie appartient (cela permet par exemple de répartir une population d'individus, comme des clients d'une entreprise en différents types de profils).

Les décisions possibles sont situées aux extrémités des branches (les « feuilles » de l'arbre) et sont atteintes en fonction de décisions prises à chaque étape. Un arbre de décision fonctionne en appliquant de manière itérative des règles logiques très simples (typiquement des séparations de données par « hyperplan », généralisation d'un plan à plus de 2 dimensions), chaque règle étant choisie en fonction du résultat de la règle précédente. Les arbres de décision ont pour avantage d'être simple à interpréter, très rapide à entraîner, d'être non paramétrique, et de nécessiter très peu de prétraitement des données. Ils peuvent être calculés automatiquement par des algorithmes d'apprentissage supervisé capables de sélectionner automatiquement les variables discriminantes au sein de données non-structurées et potentiellement volumineuses. Ces algorithmes permettent aussi d'extraire des règles logiques qui n'apparaissaient pas dans les données brutes. Un autre usage en machine learning consiste à construire non pas un arbre mais une forêt d'arbres de décision. Une décision est alors prise en faisant « voter » l'ensemble des arbres et en choisissant la réponse majoritaire (pour un choix discret) ou la moyenne des réponses (pour une variable continue). Les résultats ainsi obtenus sont remarquables notamment lorsque les arbres de décision sont utilisés en forêts aléatoires

**Auto-Encoder :** *Les auto-encodeurs sont des algorithmes d'apprentissage non supervisé à base de réseaux de neurones artificiels, qui permettent de construire une nouvelle représentation d'un jeu de données. Généralement, celle-ci est plus compacte, et présente moins de descripteurs, ce qui permet*

de réduire la dimensionnalité du jeu de données. L'architecture d'un auto-encodeur est constituée de deux parties : l'**encodeur** et le **décodeur**.

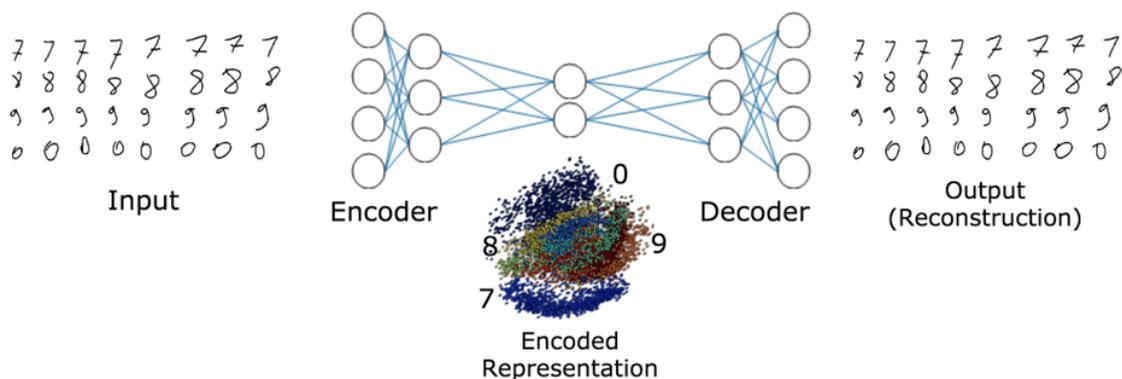
L'encodeur est constitué par un ensemble de couches de neurones, qui traitent les données afin de construire de nouvelles représentations dites "encodées". À leur tour, les couches de neurones du décodeur, reçoivent ces représentations et les traitent afin d'essayer de reconstruire les données de départ. Les différences entre les données reconstruites et les données initiales permettent de mesurer l'erreur commise par l'auto-encodeur. L'entraînement consiste à modifier les paramètres de l'auto-encodeur afin de réduire l'erreur de reconstruction mesurée sur les différents exemples du jeu de données.

La plupart du temps, on ne s'intéresse pas à la dernière couche du décodeur, qui contient uniquement la reconstruction des données initiales, mais plutôt à la nouvelle représentation créée par l'encodeur.

L'architecture la plus simple d'un auto-encodeur est semblable à un perceptron multicouche. Cependant, en fonction des données traitées, on peut utiliser différentes topologies de réseaux de neurones. Par exemple, des couches convolutives afin d'analyser des images ou des couches de neurones récurrentes pour traiter des séries temporelles ou des séquences.

À noter qu'à la différence d'un grand nombre de réseaux de neurones, les auto-encodeurs peuvent être entraînés de manière non-supervisée, ce qui permet d'appliquer ces méthodes à des jeux de données non annotés.

La figure suivante schématise un auto-encodeur simple, dont l'encodeur (encoder) traite des images (inputs), afin de les représenter comme des points dans un espace à deux dimensions (encoded representation), puis décode cette représentation (decoder), afin de retrouver les données de départ (output).



**Bagging** : Le mot *Bagging* est une contraction de *Bootstrap Aggregation*. Le *bagging* est une technique utilisée pour améliorer la classification notamment celle des arbres de décision, considérés comme des « classifieurs faibles », c'est-à-dire à peine plus efficaces qu'une classification aléatoire.

En général, le bagging a pour but de réduire la variance de l'estimateur, en d'autres termes de corriger l'instabilité des arbres de décision (le fait que de petites modifications dans l'ensemble d'apprentissage entraînent des arbres très différents). Pour ce faire, le principe du bootstrap est de créer de « nouveaux échantillons » par tirage au hasard dans l'ancien échantillon, avec remise.

L'algorithme, par exemple l'arbre de décision, est entraîné sur ces sous-ensembles de données. Les estimateurs ainsi obtenus sont moyennés (lorsque les données sont quantitatives, cas d'un arbre de régression) ou utilisés pour un « vote » à la majorité (pour des données qualitatives, cas d'un arbre de classification). C'est la combinaison de ces multiples estimateurs « indépendants » qui permet de réduire la variance. Toutefois, chaque estimateur est entraîné avec moins de données. En pratique, la méthode de bagging donne d'excellents résultats (notamment sur les arbres de décision utilisés en « forêts aléatoires »).

**Biais** : *Le biais est une des deux erreurs utilisée pour définir la qualité d'un algorithme d'apprentissage (l'autre étant la variance).*

Les algorithmes d'apprentissage tentent d'approcher la relation exacte entre des variables d'entrée et de sortie d'un problème, le vrai modèle en quelques sortes. Le modèle utilisé par l'algorithme est plus simple que le problème que l'on cherche à apprendre, il ne permet donc pas de rendre compte de toute sa complexité. On qualifie cette erreur faite dans les hypothèses du modèle de « biais ». Le biais sera d'autant plus faible que le modèle approchera la complexité du problème. Inversement, si le modèle est trop simple, le biais sera très élevé. Par exemple le perceptron est un modèle de classification linéaire trop simple pour des problèmes complexes de classification d'images comme CIFAR<sup>1</sup> : il produira un biais très élevé.

La nature de l'erreur dépend du type de problème considéré. Par exemple, dans un problème de classification d'images, l'erreur pourra être « le % de fois où le modèle se trompe en choisissant les classes » ; dans le cadre d'un problème de régression, le biais pourrait être une erreur des moindres carrés...

Quoi qu'il en soit, l'erreur totale n'est jamais nulle, ne serait qu'à cause du bruit. Cependant, elle peut être très faible. Ainsi, les derniers algorithmes de deep learning atteignent une erreur de 0,01 % sur des problèmes simples comme MNIST<sup>2</sup>. On définit aussi parfois le biais comme la « distance » entre le meilleur modèle pouvant être appris par l'algorithme et le vrai modèle. En machine learning, on cherche, en général, un équilibre entre biais et variance, de telle sorte que ces deux erreurs soient à peu près égales. La « régularisation » est une des nombreuses techniques utilisée pour y parvenir.

**Big data** : *Ensembles de données tellement volumineux qu'ils en deviennent difficiles à traiter avec des outils classiques de gestion de base de données ou de gestion de l'information.*

On parle parfois de mégadonnées ou de données massives. Le big data s'accompagne du développement d'applications analytiques (on parle de Big Analytics ou « Broyage de données ») qui traitent les données pour en tirer du sens. En 2001, un rapport de recherche du META Group (devenu Gartner) définit les enjeux inhérents à la croissance des données comme étant tridimensionnels. Les analyses complexes répondent en effet à la règle dite des « 3V » :  
– volume : données scientifiques mais aussi celles de notre vie de tous les jours comme les données de géolocalisation, issues des réseaux sociaux, les cookies, cartes de fidélité, etc. ;  
– vélocité : mises à jour fréquentes, données arrivant en flux, obsolescence rapide de certaines données, nécessité d'analyses en quasi temps réel ;  
-variété : sources, formes et formats très différents, données structurées ou non, complexes (texte en provenance du web, images, liste d'achats, données de géolocalisation, etc.).

**Bruit** : *Le bruit altère les données collectées et risque de rendre difficile l'apprentissage de la relation que l'on cherche à prédire, voire de rendre la modélisation impossible.*

Dans l'immense majorité des problèmes réels, la collecte des données est confrontée à une voire plusieurs formes de bruits. Ces bruits peuvent être d'origine « matérielle » (bruit blanc dans les signaux électriques, bruit au sens acoustique, etc.), d'imprécision ou de défaut dans les mesures issues de capteurs, d'approximation dans les données ou encore de données manquantes car non mesurées.

De fait, la plupart des modèles et des algorithmes sont conçus pour fonctionner malgré la présence de bruit. Cependant, si le bruit est trop « élevé », aucun algorithme ne peut marcher : le problème devient tout simplement mathématiquement impossible à résoudre.

Prenons l'exemple d'un problème de régression scalaire classique. Le but est de prédire une variable aléatoire  $Y$  à partir d'une variable aléatoire  $X$ . Un des modèles les plus simples consiste à supposer qu'il existe une fonction  $f$ , que l'on cherche à déterminer, telle que dans les données observées,

$$Y = f(X) + \epsilon$$

Où  $\epsilon$  est une variable aléatoire indépendante de  $X$  qui représente le bruit. Si le bruit a ce qu'on appelle de « bonnes propriétés » (il est borné, sous gaussienne, etc.), il existe plusieurs algorithmes pour résoudre le problème et trouver la fonction  $f$  (comme Kernel Ridge Regression). À l'inverse, si n'est pas de carré intégrable, ce qui revient à dire qu'il n'a pas de bonnes propriétés, le problème est beaucoup plus compliqué à résoudre. Finalement si le bruit n'est pas absolument intégrable, il n'y a généralement pas de solution au problème.

**Classification :** *Cette méthode d'analyse de données regroupe des algorithmes d'apprentissage supervisé adapté aux données qualitatives. L'objectif est d'apprendre (autrement dit de trouver) la relation qui lie une variable d'intérêt, de type qualitative, aux autres variables observées, éventuellement dans un but de prédiction. On utilise la classification lorsque la variable d'intérêt est qualitative, c'est à dire qu'elle prend ses valeurs dans un espace qui ne possède pas de métrique naturelle. Par exemple on peut essayer de prédire le genre littéraire d'un livre ; cette variable est discrète (genre « policier », genre « science-fiction », etc.) et il n'y a aucune relation entre les genres, il est difficile de définir une distance entre eux. Les algorithmes de classification les plus simples sont la régression logistique, le k-nearest neighbour (méthode des k plus proches voisins)... ; les plus complexes sont les réseaux de neurones, les support vector machine, les mixture model (modèles de mélange), le Bayesian classifier (classifieur Bayésien), etc.*

**Classifieur Bayésien :** Classifieur bayésien (ou estimateur bayésien) : *Cette méthode de classification réunit une famille d'algorithmes fondés sur le Théorème de Bayes. Leur particularité est de prédire la valeur des paramètres du modèle en termes de probabilité.*

Considérons un jeu de données et un ensemble possible de distributions, au sens probabiliste, à savoir les valeurs que peuvent prendre les variables aléatoires et à quelle fréquence. Supposons que cette famille de distribution puisse être caractérisée par un paramètre (en général baptisé  $\theta$ ). L'objectif de la classification paramétrique est de déterminer ce paramètre pour être capable de prédire la valeur des variables.

Il existe deux approches probabilistes de classification : l'approche fréquentiste et l'approche bayésienne. La première consiste à trouver le paramètre le plus probable au vue des données. Un classifieur *bayésien* essaiera, quant à lui, d'apprendre la distribution du paramètre, en d'autres terme d'estimer la pertinence de chaque valeur possible du paramètre.

Pour ce faire, il faut commencer par lui fournir une distribution « a priori » sur les paramètres ( i.e. reflétant les hypothèses que l'on a, a priori, sur le modèle ; cette distribution est appelée **prior**). L'algorithme apprend la distribution (appelée **posterior**) via la règle de Bayes, en utilisant les données.

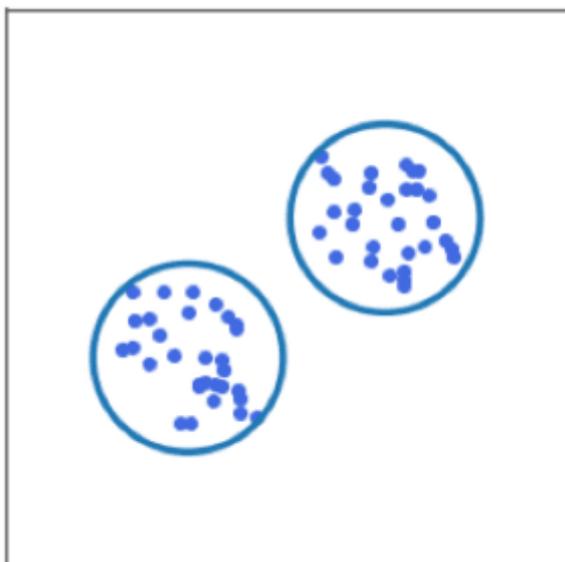
Les classifieurs bayésiens bénéficient de nombreux avantages. Notamment, ils ne présentent pas de risque de surapprentissage, et fournissent beaucoup plus d'informations que les classifieurs fréquentistes (puisque en plus de la valeur, ils en donnent la probabilité). Cependant ils nécessitent de faire des hypothèses fortes à l'avance sur les distributions des données (via le prior et la famille de distribution fournie à l'algorithme), et sont difficilement applicables aux problèmes non paramétriques.

En pratique, la méthode fonctionne relativement bien alors que les hypothèses sur lesquelles elle repose sont généralement fausses dans les problèmes réels – par exemple, les distributions fournies à l'algorithme au début ne sont généralement pas les vraies distributions des variables aléatoires sous-jacentes au problème. Ainsi, le débat sur les mérites relatifs entre classifieurs fréquentistes versus bayésiens, qui date des débuts de l'apprentissage automatique, reste toujours d'actualité dans la communauté.

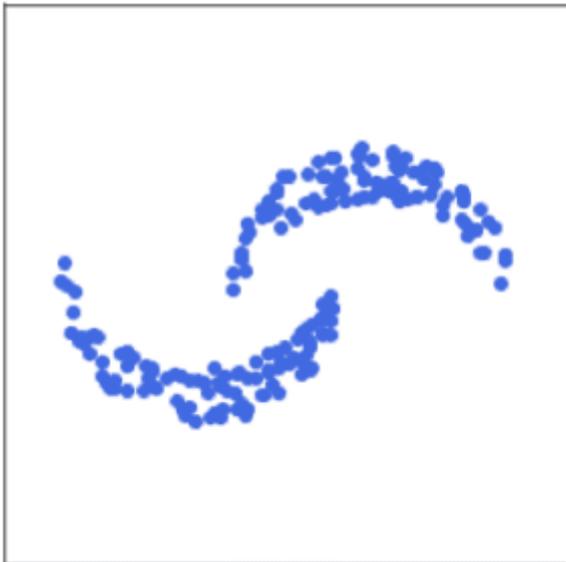
**Clustering** : (ou partitionnement des données) : *Cette méthode de classification non supervisée rassemble un ensemble d'algorithmes d'apprentissage dont le but est de regrouper entre elles des données non étiquetées présentant des propriétés similaires. Isoler ainsi des schémas ou des familles permet aussi de préparer le terrain pour l'application ultérieure d'algorithmes d'apprentissage supervisé (comme le KNN).*

Le clustering est utilisé notamment lorsqu'il est coûteux d'étiqueter le données. C'est néanmoins un problème mal défini mathématiquement : différentes métriques et/ou différentes représentations des données aboutiront à différents regroupements sans qu'aucun ne soit nécessairement meilleur qu'un autre. Ainsi la méthode de clustering doit être choisie avec soin en fonction du résultat attendu et de l'utilisation prévue des données.

Les algorithmes de clustering les plus courants sont le K-Means, les algorithmes de maximisation de l'espérance (de type EM, comme les mixtures gaussiennes) et les partitions de graphes. Voyons comment cela se traduit dans deux exemples :



Ici, les données peuvent être aisément regroupées en deux groupes (dans 2 cercles). L'algorithme K-Means (ou partitionnement en k-moyennes) qui consiste à diviser les points en k groupes appelés clusters, permet d'obtenir ce résultat rapidement et efficacement.



Dans cet autre exemple (le cas des « 2 lunes »), un clustering « naturel » serait de regrouper les données en 2 lunes. L'algorithme K-means ne permet pas de produire ce regroupement. Par contre, les algorithmes de partition de graphes peuvent y parvenir

**Data mining** : *Le data mining couvre l'ensemble des outils et méthodes qui permettent d'extraire des connaissances à partir de grandes bases de données. On parle aussi de fouille, forage ou prospection de données, d'extraction de connaissances à partir de données.*

C'est une analyse préliminaire où l'on explore, on cherche à confirmer des intuitions, à faire émerger des concepts (insights). C'est une façon de produire de la connaissance mais cette étape ne s'automatise pas. Certains y intègrent aussi la transformation des données en informations utiles, en établissant des relations entre les données, des corrélations (on parle aussi de patterns, de motifs, de critères) pour les catégoriser. Le data mining est une extension de l'analyse de données et des statistiques exploratoires pratiquées depuis plus de 30 ans. Il intègre (ou il est le prélude à) des techniques d'analyse issues de l'apprentissage automatique (comme le machine learning), de la reconnaissance de formes et des bases de données de diverses natures dont celles issues d'entrepôts de données (ou Data Warehouse).

Le Data Mining, également surnommé Knowledge Discovery in Data (découverte de savoir dans les données), repose sur des algorithmes complexes et sophistiqués permettant de segmenter les données et d'évaluer les probabilités futures, comme les tendances d'un marché.

**Data science** : *Le but de cet ensemble de disciplines (la « science des données » en français) est d'extraire de la connaissance à partir des données grâce à des techniques et des théories issues des mathématiques appliquées, de la statistique et de l'informatique.*

L'objectif est de produire des méthodes automatisées de tri, d'analyse, de classification de grandes quantités de données et de sources de données, plus ou moins complexes : méthodes statistiques,

traitement du signal, méthodes de référencement, apprentissage automatique, visualisation de données.

**Deep learning :** *Le deep learning (ou apprentissage profond) est un ensemble de méthodes d'apprentissage automatique conçues sur la base de réseaux de neurones profonds, visant à mimer la « profondeur » des couches d'un cerveau :*

le cerveau humain est « profond », dans le sens où chaque action est le résultat d'une longue chaîne de communications synaptiques avec de nombreuses couches de traitement. Le deep learning réunit une classe d'algorithmes d'apprentissage correspondants à ces architectures profondes. Il est souvent utilisé pour un apprentissage « de bout en bout », c'est à dire l'apprentissage simultané des caractéristiques utiles des données, et de la meilleure façon de les utiliser.

Par exemple, pour distinguer une voiture d'une moto, l'algorithme peut apprendre à reconnaître les roues (caractéristique utile), puis à utiliser le nombre de roues pour les distinguer (utilisation de cette caractéristique). Cela fait appel à la fois à des connaissances en neurosciences, en mathématiques et aux progrès technologiques. Le système est constitué d'une série de modules (des couches de neurones), chacun représentant une étape de traitement. Chaque module est entraînable, comportant des paramètres ajustables (similaires aux poids des classifieurs linéaires). A chaque exemple, tous les paramètres de tous les modules sont ajustés de manière à rapprocher la sortie produite par le système de la sortie désirée. Le qualificatif « profond » vient de l'arrangement de ces modules en couches successives. Les réseaux convolutifs et les réseaux récurrents sont deux des architectures les plus en vogue actuellement pour le deep learning. Les applications sont multiples.

Parmi les exemples récents les plus remarquables, citons le programme de reconnaissance des visages de Facebook, la victoire d'AlphaGo ou encore les outils d'aide à la conduite assistée et autonome (ADAS), la santé avec la recherche de cellules cancéreuses par la start-up DreamQuark, ou la reconnaissance de parole.

**Descripteur (feature) :** *Un descripteur est une quantité mesurable ou calculable qui permet de décrire en partie un objet, un signal, une donnée...*

Le choix et la conception de descripteurs est généralement une étape préliminaire cruciale de la fouille de données, nécessaire avant toute utilisation d'algorithmes. Les descripteurs sont généralement numériques (ce sont des nombres). Ils peuvent aussi prendre des formes plus complexes. Par exemple, un descripteur de couleur peut être un nombre : 1-bleu ; 2-noir, etc. Cela peut aussi être une description RGB de la couleur : trois chiffres compris entre 0 et 255, chaque chiffre représentant respectivement le dosage du rouge, du vert et du bleu. Concrètement, pour décrire une voiture, on peut imaginer différents descripteurs tels que les dimensions du véhicule, son poids, sa couleur, sa marque, l'année de fabrication, etc. Comme on peut le voir dans cet exemple, un descripteur seul ne suffit généralement pas à caractériser totalement un objet, mais une collection de descripteurs peut s'en approcher. Il n'existe pas d'algorithme pour créer les descripteurs. Par contre on peut utiliser des descripteurs « génériques » puis utiliser des algorithmes pour « sélectionner » les meilleurs. Mais, la meilleure solution reste de créer des descripteurs spécifiques.

**Détection d'anomalie :** *Le but de la détection d'anomalie est de repérer des données qui ne sont pas conformes à ce à quoi l'on peut s'attendre par rapport aux autres données.*

Il s'agit, par exemple, de données qui ne suivent pas le même schéma ou qui sont atypiques pour la distribution de probabilité observée. La difficulté du problème provient du fait qu'on ne connaît pas

au préalable la distribution sous-jacente de l'ensemble des données. C'est à l'algorithme d'apprendre une métrique appropriée pour détecter les anomalies. Parmi les exemples d'applications courantes, citons les transactions bancaires (où une anomalie sera vue comme une fraude potentielle), la surveillance des données physiologiques d'un malade (l'anomalie est un problème de santé possible), ou encore la détection de défauts dans des chaînes de production. La détection d'anomalie est souvent un problème d'apprentissage de type non supervisé. Les algorithmes typiques de détection d'anomalie sont les one-class SVM, les méthodes d'apprentissage de distribution bayésienne et les random forests.

**Données :** *C'est la matière première de tout algorithme. Les données peuvent être de diverses natures (signaux, vidéos, séquences...) et plus ou moins structurées.*

Tout élément numérisé et stocké sur un serveur peut être appréhendé comme une donnée : les mesures physiques (acoustiques, issues de capteurs comme des accéléromètres, etc.), le texte, les séquences d'ADN, les images, la vidéo, etc. Certaines données sont binaires comme les images, les sons ou les vidéos, numérisées grâce à des approximations (en découpant les images en pixels monochromes, chaque couleur étant une suite de 24 bits), les sons en une suite de nombres... Les données structurées proviennent de bases de données relationnelles ; les données semi-structurées comportent les CSV (données de tableaux structurées pour de l'analyse quantitative), les logs (traces laissées sur les serveurs), les XML et les JSON (deux formats de stockage de données hétérogènes très utiles dans les applications). Quant aux données non structurées, on y trouve les emails, les documents et les PDF

**Hackathon :** Ce mot désigne à la fois le principe, le moment et le lieu d'un événement où un groupe de développeurs volontaires se réunit pour collaborer sur des sujets de programmation informatique pointus et innovants, sur une période généralement courte (une journée, une nuit, un week-end).

Le terme vient de hack (littéralement pirater, autrement dit trouver quelque chose de malin) et marathon (en référence au travail en général sans interruption). Organisés en équipe, les développeurs ont un objectif commun : tester des idées et produire un prototype, un algorithme dans le cas de la science des données. C'est un moyen de se tester dans un contexte d'émulation, avec parfois des récompenses à la clé. Un jury détermine les vainqueurs. Les investisseurs et les entreprises peuvent aussi dénicher des idées et des talents à ces occasions. Les hackathons peuvent également être organisés en interne comme le pratiquent Facebook, Yahoo, Google ou LinkedIn.

Malgré un délai de réalisation extrêmement court et même si la manifestation est plutôt festive et encourage l'entraide, les enjeux sont sérieux et la méthode est rigoureuse.

**Inégalité de concentration :** *Les inégalités de concentration fournissent des bornes sur la probabilité qu'une statistique basée sur des tirages successifs d'une variable aléatoire dévie d'une certaine valeur. Parmi les plus connus, citons Azuma- Hoeffding (pour les variables aléatoires bornées) et McDiarmid (pour les martingales à accroissements finis).*

Contrairement à la loi des grands nombres, les inégalités de concentration sont utiles à la fois en théorie et en pratique. Elles permettent de contrôler l'incertitude liée aux quantités statistiques comme la moyenne empirique. Il existe de nombreuses méthodes de calculs d'inégalités.

Exemple : en utilisant l'inégalité d'Azuma-Hoeffding, on sait que si on lance 10 000 fois une pièce équilibrée, la proportion de tirages côté pile sera comprise entre 0,483 et 0,517 avec une probabilité supérieure à 99 %. Ces résultats permettent de calculer des intervalles de confiance autour des

prédictions des modèles de machine learning. Ils sont aussi fondamentaux dans plusieurs domaines de recherches comme l'apprentissage par renforcement et les problèmes d'algorithmes de bandits.

**Intelligence artificielle** : *L'IA est un ensemble de techniques permettant à des machines d'accomplir des tâches et de résoudre des problèmes normalement réservés aux humains et à certains animaux (selon Y. LeCun, présentation au Collège de France).*

Ces tâches sont parfois très simples pour les humains, moins pour les machines comme reconnaître et localiser les objets dans une image, planifier les mouvements d'un robot pour attraper un objet ou conduire une voiture. Elles requièrent parfois de la planification complexe, comme par exemple pour jouer aux échecs ou au jeu de go. Les tâches les plus compliquées nécessitent beaucoup de connaissances et de sens commun, par exemple pour traduire un texte ou conduire un dialogue.

Depuis quelques années, on associe presque toujours l'intelligence aux capacités d'apprentissage. C'est grâce à l'apprentissage qu'un système intelligent capable d'exécuter une tâche peut améliorer ses performances avec l'expérience. C'est grâce à l'apprentissage qu'il pourra apprendre à exécuter de nouvelles tâches et acquérir de nouvelles compétences. Le domaine de l'IA n'a pas toujours considéré l'apprentissage comme essentiel à l'intelligence. Dans le passé, construire un système intelligent consistait à écrire un programme "à la main" pour jouer aux échecs (par recherche arborescente), reconnaître des caractères imprimés (par comparaison avec des images prototypes), ou faire un diagnostic médical à partir des symptômes (par déduction logique à partir de règles écrites par des experts). Mais cette approche "manuelle" a ses limites.

**Isolation forest** : *Cet [algorithme](#) non supervisé de machine learning permet de détecter des anomalies dans un jeu de données. Il isole les données atypiques, autrement dit celles qui sont trop différentes de la plupart des autres données.*

Cet algorithme calcule, pour chaque donnée du jeu, un score d'anomalie, c'est à dire une mesure qui reflète à quel point la donnée en question est atypique. Afin de calculer ce score, l'algorithme isole la donnée en question de manière récursive : il choisit un descripteur et un "seuil de coupure" au hasard, puis il évalue si cela permet d'isoler la donnée en question ; si tel est le cas, l'algorithme s'arrête, sinon il choisit un autre descripteur et un autre point de coupure au hasard, et ainsi de suite jusqu'à ce que la donnée soit isolée du reste.

Le partitionnement récursif des données peut-être représenté comme un arbre de décision et le nombre de coupures nécessaires pour isoler une donnée correspond tout simplement au chemin parcouru dans l'arbre depuis la racine jusqu'à la feuille, représentant la donnée isolée. La longueur du chemin définit le score l'anomalie : les données ayant un chemin très court, c'est à dire les données faciles à isoler, ont également de grandes chances d'être des anomalies, puisqu'elles sont très loin des autres données du jeu.

Comme pour les forêts aléatoires, il est possible d'exécuter cette démarche indépendamment en utilisant plusieurs arbres, afin de combiner leurs résultats pour gagner en performance. Dans ce cas là, le score d'anomalie correspond à la moyenne des longueurs des chemins sur les différents arbres. Cet algorithme s'avère particulièrement utile car il est très rapide et qu'il ne nécessite pas de paramétrage compliqué.

Dans l'exemple suivant, on a appliqué l'algorithme Isolation forest avec 50 arbres à un jeu de données gaussien avec deux descripteurs (2 axes) comportant quelques anomalies (les 5 points les plus extrêmes en rouge sur la figure). Une fois les scores d'anomalies calculés par l'algorithme, on constate que ce sont bien ces 5 points extrêmes qui ont le score le plus élevé.



**K-Means (ou K-moyennes) :** C'est l'un des algorithmes de clustering les plus répandus. Il permet d'analyser un jeu de données caractérisées par un ensemble de descripteurs, afin de regrouper les données "similaires" en groupes (ou clusters).

La similarité entre deux données peut être inférée grâce à la "distance" séparant leurs descripteurs ; ainsi deux données très similaires sont deux données dont les descripteurs sont très proches. Cette définition permet de formuler le problème de partitionnement des données comme la recherche de  $K$  "données prototypes", autour desquelles peuvent être regroupées les autres données.

Ces données prototypes sont appelés *centroïdes* ; en pratique l'algorithme associe chaque donnée à son centroïde le plus proche, afin de créer des *clusters*. D'autre part, les moyennes des descripteurs des données d'un *cluster*, définissent la position de leur centroïde dans l'espace des descripteurs : ceci est à l'origine du nom de cet algorithme (K-moyennes ou *K-means* en anglais).

Après avoir initialisé ses centroïdes en prenant des données au hasard dans le jeu de données, *K-means* alterne plusieurs fois ces deux étapes pour optimiser les centroïdes et leurs groupes :

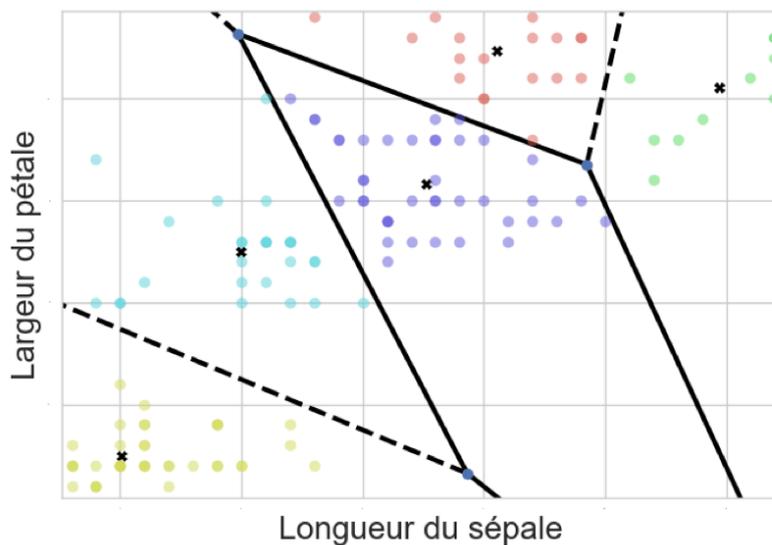
1. Regrouper chaque objet autour du centroïde le plus proche.
2. Replacer chaque centroïde selon la moyenne des descripteurs de son groupe.

Après quelques itérations, l'algorithme trouve un découpage stable du jeu de données : on dit que l'algorithme a convergé.

Comme tout algorithme, *K-means* présente des avantages et des inconvénients : il est simple, rapide et facile à comprendre ; cependant il ne permet pas de trouver des groupes ayant des formes complexes.

L'exemple suivant s'appuie sur le célèbre jeu de données "Iris" qui décrit des fleurs par l'intermédiaire des longueurs et largeurs de leurs pétales et sépales. Les descripteurs considérés ici sont la longueur des sépales et la largeur des pétales. Chaque point correspond à une fleur et la

couleur associée au point reflète son appartenance à un groupe. Les centroïdes de chaque groupe sont représentés par des croix, les frontières par des traits.



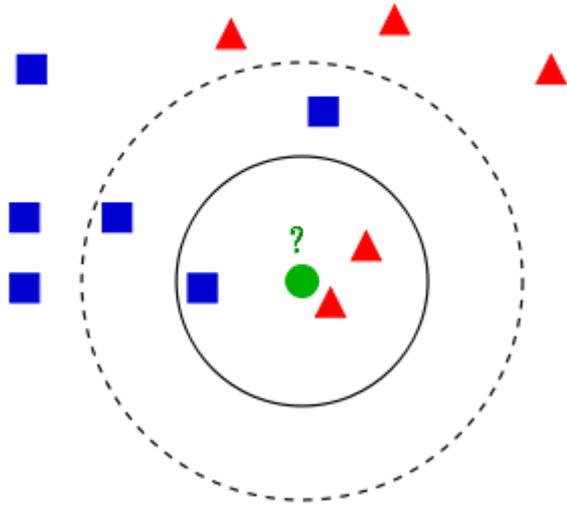
**k-Nearest Neighbours** : (k-NN voire KNN ou méthode des k plus proches voisins) : *k-NN est un algorithme standard de classification qui repose exclusivement sur le choix de la métrique de classification. Il est "non paramétrique" (seul k doit être fixé) et se base uniquement sur les données d'entraînement.*

L'idée est la suivante : à partir d'une base de données étiquetées, on peut estimer la classe d'une nouvelle donnée en regardant quelle est la classe majoritaire des k données voisines les plus proches (d'où le nom de l'algorithme). Le seul paramètre à fixer est k, le nombre de voisins à considérer (voir figure).

Les métriques les plus souvent choisies sont la distance usuelle dite euclidienne (comme dans la figure) et la distance de Mahalanobis (qui tient compte de la variance – du point de vue statistique – et de la corrélation entre les données). Bien que l'algorithme puisse fonctionner avec ces métriques par défaut, il est généralement bien meilleur quand il est utilisé avec une métrique adaptée aux données, métrique qui peut être calculée à partir d'heuristiques connues liées au problème (par exemple la distance euclidienne pondérée).

Les points faibles de cet algorithme sont : d'une part, son coût en puissance de calcul (pour prédire l'image d'un nouveau point, on doit calculer sa distance à tous les autres), d'autre part le fait de devoir conserver toutes les données d'entraînement en mémoire (k-NN convient donc plutôt aux problèmes d'assez petite taille). Il est également important de noter que cet algorithme est vulnérable à la « *curse of dimensionality* » : le nombre de données nécessaires pour avoir un bon estimateur croît potentiellement de manière exponentielle avec la dimension, autrement dit avec la complexité de la représentation des données. Pour ces raisons, k-NN est assez peu utilisé dans sa forme première mais plutôt avec des versions améliorées qui limitent partiellement ces défauts .

A remarquer que l'algorithme peut aussi être adapté pour la régression.



*Quelle est la classe du point vert ? Celle des triangles rouges (délimitée par le cercle continu) ou celle des carrés bleus (cercle tracé en pointillés) ? Si le nombre de plus proches voisins,  $k$ , est fixé à 3, la classe du point vert est celle des triangles rouges, car ces derniers sont au nombre de 2 contre un seul carré bleu. Si  $k$  vaut 5, la classe du point vert est celle des carrés bleus, au nombre de 3 contre 2 triangles rouges. (Source de l'image : Wikipédia CC BY-SA 3.0)*

**Loi des grands nombres :** *La loi des grands nombres est un théorème mathématique fondamental des probabilités et statistiques.*

Cette loi exprime le fait que les caractéristiques d'un échantillon aléatoire se rapprochent des caractéristiques statistiques de la population (ensemble d'individus ou d'éléments) lorsque la taille de l'échantillon augmente à l'infini. En d'autres termes, cela garantit que, lorsque le nombre de tirages effectués selon une loi de probabilité (comme les tirages successifs d'une pièce sur le côté pile ou face) tend vers l'infini, la *moyenne empirique* (moyenne calculée à partir des observations) converge vers la moyenne réelle d'une variable aléatoire suivant cette loi. Cela sous des hypothèses très faibles.

C'est un des premiers résultats qui lie les observations d'un événement – par exemple les tirages d'une pièce – avec sa variable aléatoire – ici une distribution de type Bernoulli (distribution discrète de probabilité qui prend la valeur 1 avec la probabilité  $p$  et la valeur 0 avec la probabilité  $q = 1 - p$ ). C'est sur cette loi que reposent la plupart des sondages (ils interrogent un nombre suffisamment important de personnes pour connaître l'opinion ou les comportements de la population entière) ou l'assurance (en déterminant les probabilités que les sinistres garantis se réalisent ou non).

Il est à noter que la loi des grands nombres n'offre pas d'utilisation propice en pratique, car le nombre de tirages nécessaires pour approcher suffisamment la moyenne réelle est inconnu. Cette loi n'a en fait qu'une valeur asymptotique. Lorsqu'on exécute un nombre fini d'expériences, il y a des écarts par rapport au comportement moyen attendu. Ainsi, après 10 000 tirages d'une pièce équilibrée, on n'a pas la garantie d'observer 5 000 fois « pile » ni même plus de 4 000. Dans la pratique, on utilise d'autres résultats tels que les Inégalités de concentration ou le Théorème central limite.

**Loi gaussienne :** *La loi gaussienne (ou normale) est une des lois de probabilité les plus utilisées dans les sciences appliquées du fait de ses propriétés théoriques remarquables.*

La loi gaussienne est une loi de probabilité paramétrique. Elle est caractérisée par sa moyenne  $\mu$  et sa variance  $\sigma^2$ .

On la note  $\mathcal{N}(\mu, \sigma^2)$

Une variable aléatoire  $X$  suivant une loi  $\mathcal{N}(\mu, \sigma^2)$  a pour densité de probabilité :

$$\forall x \in \mathbb{R}, p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Cette fonction de densité, appelée courbe de Gauss, possède une forme caractéristique rappelant celle d'une cloche. Elle se répartit de manière symétrique autour de la moyenne, point où elle atteint son maximum. Elle décroît ensuite à mesure que les valeurs sont éloignées de la moyenne. Cette concentration de la densité autour de la moyenne est une caractéristique importante de la loi gaussienne. Elle se traduit notamment par le fait qu'une majorité des valeurs observées est contenue dans un intervalle restreint autour de la moyenne. Par exemple, la probabilité d'avoir une valeur comprise dans l'intervalle  $[\mu - 1.96 \sigma, \mu + 1.96 \sigma]$  est égale à 95 % et la probabilité d'avoir une valeur comprise dans l'intervalle  $[\mu - 3.29 \sigma, \mu + 3.29 \sigma]$  est de 99.9 %.

**Machine learning :** *Plutôt que d'expliquer à un ordinateur avec précision comment résoudre un problème, le Machine Learning (ou apprentissage automatique) permet de lui apprendre à apprendre à résoudre un problème par lui-même. Ce champ d'étude comporte des dizaines d'algorithmes.*

On parle aussi de systèmes entraînés car ces algorithmes sont capables de faire émerger des règles mathématiques dans les données en s'entraînant sur la base d'exemples, puis d'appliquer ces règles à de nouvelles données en s'améliorant sans cesse avec l'expérience. Parmi les algorithmes les plus courants, on trouve les SVM (Support Vector Machine), le boosting, les random forests, les réseaux de neurones, les réseaux bayésiens, etc. Ils opèrent dans des contextes variés : supervisé, semi-supervisé ou non-supervisé, en mode séquentiel ou batch (par lot), par renforcement, etc. Ce sont des systèmes « entrée-sortie » avec une donnée en entrée (image, son, texte) et une en sortie (telle que la catégorie de l'objet dans l'image, le mot prononcé, le sujet dont parle le texte).

Toutes les tâches nécessitant d'entrer des données et de les classifier peuvent ainsi être automatisées : cela permet de doter des ordinateurs ou des machines de systèmes de perception de leur environnement comme la vision, la reconnaissance d'objets (visages, schémas, langages naturels, écriture, formes syntaxiques...), de la parole ; sur Internet, cela permet de filtrer des contenus indésirables (spam), d'ordonner des réponses à une recherche, de faire des recommandations ou de sélectionner les informations intéressantes pour chaque utilisateur (moteurs de recherche) ; de concevoir des systèmes d'aide aux diagnostics, médical notamment, des programmes de jeu, des interfaces cerveau-machine, des systèmes de détection de fraudes à la carte de crédit, d'analyse financière, de classification des séquences d'ADN, d'analyse prédictive en matière juridique et judiciaire...

**Modèles de « bandits » :** Cette famille d'algorithmes propose des stratégies optimales pour maximiser l'espérance d'un gain lors d'une succession de choix entre plusieurs actions aux récompenses inconnues (on parle aussi de maximiser le rendement et de minimiser le regret).

**Normalisation :** La normalisation est une méthode de prétraitement des données qui permet de réduire la complexité des modèles. C'est également un préalable à l'application de certains algorithmes.

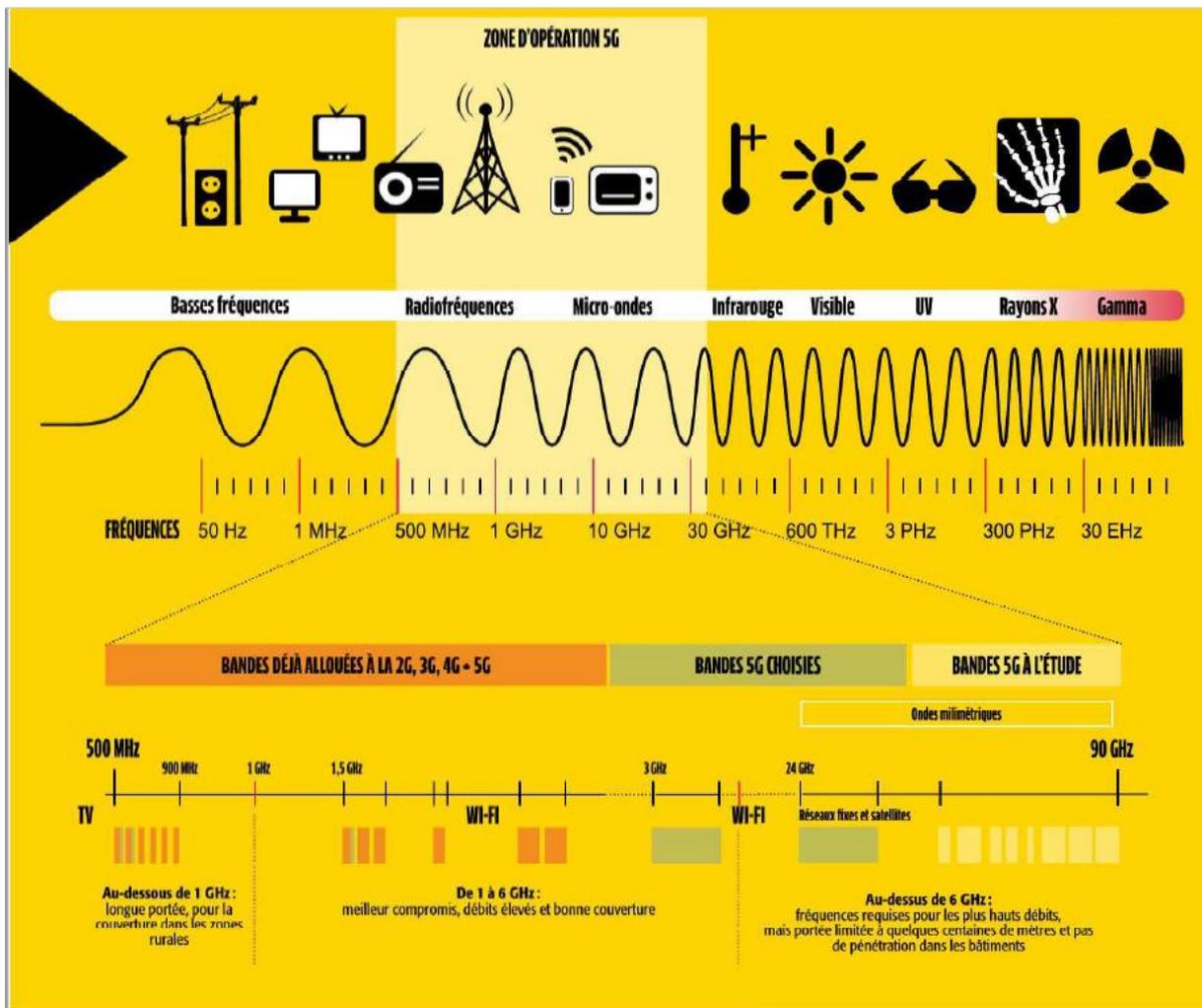
**Overfitting :** Ce phénomène de « surapprentissage » dégrade la performance des algorithmes de machine learning.

Dans quels cas se produit l'overfitting ? Un algorithme d'apprentissage, par exemple supervisé, cherche le modèle qui exprime le mieux la relation entre des données. L'overfitting intervient lorsque l'algorithme sur-apprend (*overfit*), autrement dit, lorsqu'il apprend à partir des données mais aussi à partir de patterns (schémas, structures) qui ne sont pas liés au problème, comme du bruit. Ainsi, l'overfitting est caractérisé par une erreur de type variance très élevée. Concrètement, on observe généralement de l'overfitting lorsqu'on utilise des modèles très complexes sur des problèmes simples mais bruités : par exemple, lors de l'utilisation de Support Vector Machine (SVM ou machine à vecteur de support) avec des noyaux polynomiaux de très haut degré dans le cadre de l'apprentissage d'un problème linéaire (c'est à dire d'un polynôme de degré 1). En d'autres termes, ce type de modèle conduit à de mauvaises performances car, trop complexe, il manque de capacité de généralisation. La technique la plus courante pour limiter le phénomène est la régularisation qui permet de réduire l'erreur de type variance.

**Régression :** Cette méthode d'analyse de données regroupe des algorithmes d'apprentissage supervisé adaptés aux données quantitatives. L'objectif est d'apprendre (autrement dit de trouver) la relation qui lie une variable d'intérêt, de type quantitative, aux autres variables observées, éventuellement dans un but de prédiction. On utilise la régression lorsque la variable d'intérêt est quantitative, c'est à dire « à valeur » dans un espace métrique – la métrique est une notion de distance définie dans l'espace – et souvent « à valeur continue ». Par exemple on peut essayer de prédire l'âge d'un utilisateur en fonction de son comportement ; l'âge est une donnée continue avec la métrique usuelle des nombres réels (23 ans et 22 ans sont distants de 1 an). Les algorithmes de régression les plus simples sont de type régression linéaire, les plus compliqués de type régression à noyau des moindres carrés, réseau de neurones, support vector machine, etc.

**Rétropropagation :** La rétropropagation du gradient de l'erreur (ou *backpropagation*) est un algorithme d'optimisation permettant d'ajuster les paramètres d'un réseau de neurones multicouches pour mettre en correspondance des entrées et des sorties référencées dans une base d'apprentissage.

Pour pouvoir entraîner ces systèmes, il faut savoir comment ajuster les paramètres de chaque couche de neurones. La rétropropagation permet de calculer le gradient de l'erreur pour chaque neurone, de la dernière couche vers la première. Le calcul de ce gradient se fait par la méthode de rétropropagation, pratiquée depuis le milieu des années 80. Cela permet de corriger les erreurs selon l'importance des éléments qui ont justement participé à la réalisation de ces erreurs. Ainsi, les poids synaptiques qui contribuent à engendrer une erreur importante se verront modifiés de manière plus significative que les poids qui ont engendré une erreur marginale. Moyennant quelques précautions lors de l'apprentissage, les procédures d'optimisation finissent par aboutir à une configuration stable, généralement un extremum local, au sein du réseau de neurone.



## Partie 2 : Glossaire

**3V**, 185  
**4V**, 185  
**802.11b**, 76  
**A.D.S.L**, 7  
**A.G.P.**, 7  
**Abstraction**, 35  
**Abstraite**, 35  
**Abus informatique**, 76  
**Accès à distance**, 7, 76  
**Accès en mode forage**, 76  
**Access Provider**, 7  
**Accessibilité**, 35, 53  
**Acteur**, 35  
**Acteur métier**, 140  
**Actifs complémentaires**, 76  
**Action**, 35, 140  
**Activité**, 35  
**Adaptateurs de réseau sans fils**, 76  
**Adaptive algorithm**, 185  
**Administrateur système**, 7  
**Administration de la base de données**, 76  
**Administration des données**, 7, 76  
**Adresse IP**, 76  
**Adresse URL**, 76  
**Advanced Encryption Standard**, 151  
**Adware**, 151  
**AES**, 151  
**Affaire électronique**, 76  
**Agenda partagé**, 7  
**Agent**, 35, 124  
**Agent intelligent**, 7  
**Agilité**, 7  
**AGL**, 77  
**Agrégation**, 35  
**aide au dépannage**, 113  
**Alerte éthique**, 151  
**Alerte professionnelle**, 151  
**Algèbre relationnel**, 162  
**Algorithm**, 185  
**Algorithme**, 8, 35, 124, 185  
**Algorithme adaptatif**, 185  
**Algorithme du simplex**, 124  
**Algorithme évolutionniste ou évolutionnaire**, 124  
**Algorithme génétique**, 185  
**Algorithme génétique**, 124  
**Alignement stratégique**, 77  
**ALT et ALT GR**, 8, 77  
**Amélioration de processus**, 53  
**Amélioration du processus de test**, 53  
**Analogique**, 8  
**Analyse**, 8, 35, 77  
**Analyse de données**, 186  
**Analyse de portefeuille**, 77  
**Analyse de risque**, 53  
**Analyse de test**, 53  
**Analyse des besoins**, 35  
**Analyse des causes racines**, 53  
**Analyse des valeurs limites**, 53  
**Analyse d'impact**, 53  
**Analyse du domaine**, 35  
**Analyse dynamique**, 53  
**Analyse en composantes principales**, 125  
**Analyse heuristique**, 8, 77  
**Analyse statique**, 53  
**Analyse syntaxique**, 125  
**Ancêtre**, 35  
**Android**, 77  
**Angular.JS**, 35  
**Annuaire d'entreprise**, 8  
**Anomalie**, 51, 53  
**Antivirus**, 77  
**Apache**, 35  
**Aperçu**, 77  
**API**, 8, 35, 186  
**Appel d'offres**, 77  
**Applet**, 8  
**Application**, 8, 36  
**Application Programming Interface**, 9  
**Applications mobiles**, 9  
**apportez votre équipement personnel de communication**, 113  
**Apprentissage automatique**, 126, 134  
**Apprentissage non supervisé**, 125, 134, 186  
**Apprentissage organisationnel**, 77  
**Apprentissage par imitation**, 187  
**Apprentissage par renforcement**, 126, 134, 187  
**Apprentissage profond**, 126, 134  
**Apprentissage supervisé**, 126, 135, 188  
**Approche du test**, 54  
**Approche probabiliste fréquentiste**, 188  
**Approche progressive ou modulaire**, 77  
**Approvisionnement**, 77

**Aptitude fonctionnelle**, 54  
**Arbre de décision**, 189  
**Architecte**, 36  
**Architecture**, 36  
**Architecture à trois niveaux**, 9  
**Architecture Client Serveur**, 9  
**Architecture de réseau**, 78  
**Architectures client-serveur multiniveau**, 78  
**Archive**, 78  
**Argent électronique**, 78  
**Array**, 36  
**arrière-guichet**, 113  
**Artefact**, 36  
**Article de test**, 54  
**Ascenseur**, 78  
**ASP**, 36, 78  
**Assembleur**, 36  
**Assistant à maîtrise d'ouvrage**, 9  
**Assistant numérique personnel**, 78  
**Association**, 36, 162  
**Association dérivée**, 36  
**Assurance qualité**, 54  
**Asymétrie d'information**, 78  
**Asynchrone**, 36  
**ATAPI**, 78  
**atelier numérique ouvert**, 113  
**Attaque par déni de service**, 78  
**Attribut**, 36, 78, 162  
**Attribut de fichier**, 78  
**Attribut de lien**, 36  
**Attribut dérivé**, 36  
**ATX**, 78  
**Audit**, 54  
**Audit de qualité des données**, 78  
**Authentification**, 78  
**Authentification biométrique**, 79  
**Auto-Encoder**, 189  
**Automate**, 36  
**Automatisation**, 9, 79  
**Automatisation des tests**, 54  
**AutoTexte**, 79  
**Avatar**, 79  
**B.I.O.S.**, 80  
**Backbone**, 79  
**Back-end**, 9  
**Background color**, 79  
**Backlink**, 9  
**Backlog**, 137  
**Back-office**, 9  
**Backup**, 79  
**Bagging**, 190  
**Balise**, 9, 79  
**Bandeau publicitaire**, 79  
**Bannière**, 113, 151  
**Banque de données**, 162  
**Barre de défilement**, 79  
**Barre de formule**, 79  
**Barre de menu**, 79  
**Barre des tâches**, 79  
**Barres d'outils**, 79  
**Bartering**, 151  
**Basculement direct**, 80  
**Base de données**, 10, 80  
**Base de données hypermédias**, 80  
**Base de données objet**, 162  
**Base de données orientés objet**, 80  
**Base de données relationnelle**, 163  
**Base de données répartie**, 80  
**Base de registres**, 80  
**Base de règles**, 80  
**Base de test**, 54  
**Base documentaire**, 10  
**Bases de données relationnelles**, 80  
**Bayésien**, 127  
**Benchmark**, 51  
**Bénéfices intangibles**, 80  
**Bénéfices tangibles**, 80  
**Bêta**, 10  
**Biais**, 190  
**Big bang**, 36  
**Big data**, 80, 164, 191  
**Big Data :**, 10  
**Bilan de test**, 51  
**Binaire**, 36  
**Bios**, 151  
**Bit**, 81  
**BIT**, 36  
**Bitmap**, 81  
**Bittorrent**, 151  
**Block-Based Programming**, 36  
**Blockchain**, 10, 81  
**bloc-notes**, 113  
**Blog**, 151  
**Blogue**, 81, 113  
**bloqueur de publicités**, 114  
**Bluetooth**, 81  
**Bogue**, 81  
**Boîte aux lettres électronique**, 81  
**bombe logique**, 114  
**Bombe logique**, 151  
**bombe programmée**, 114  
**Bookmark**, 81  
**Boot**, 81  
**Bot**, 10

**Botnet**, 151  
**Bouchon**, 54  
**Bouton**, 81  
**Brouillage**, 114  
**Browser**, 10, 81  
**Bruit**, 81, 191  
**Buffer**, 81  
**Bug**, 37, 81  
**Bureau**, 82  
**Bureau d'enregistrement**, 151  
**Bureautique**, 10  
**Burndown chart**, 54, 137  
**Bus**, 10, 82  
**Business Process Reengineering**, 10  
**BYOD**, 82  
**Byte**, 37, 82  
**C.P.U.**, 87  
**Câble**, 82  
**Câble coaxial**, 82  
**Cache**, 10, 82  
**Caching**, 151  
**Cadre**, 82  
**Cahier des charges**, 82  
**Calendrier d'exécution des tests**, 54  
**Call (une fonction)**, 37  
**Call (une variable)**, 37  
**Camembert**, 82  
**Campagne de Test**, 51  
**Canular**, 114  
**CAO**, 10, 85  
**Capacité**, 82  
**CAPTCHA**, 127  
**Capteur CCD**, 82  
**Capture d'écran**, 82  
**Caractéristique**, 55, 164  
**Caractéristique de qualité**, 55  
**Cardinalité**, 37  
**Cardinalités**, 164  
**Carrousel**, 114  
**Carte graphique**, 82  
**Carte mère**, 82  
**Cartes à puce (Smart card)**, 83  
**Cartouche**, 114  
**Cas d'application**, 140  
**Cas d'utilisation**, 55  
**Cas de test**, 54  
**Cas de test (exigence)**, 51  
**Cas de test de bas niveau**, 54  
**Cas de test de haut niveau**, 55  
**Cas d'utilisation**, 37  
**Casse**, 83  
**Catégorie**, 37  
**Cause racine**, 55  
**cédérom autonome**, 115  
**Cellule**, 83  
**Centre d'appels (Call center)**, 83  
**centre de données**, 115  
**Centre Français d'exploitation du droit de copie**, 152  
**cercle de confiance**, 115  
**CERN**, 83  
**Certificat électronique**, 83  
**Certification**, 83  
**Chaînage arrière**, 83  
**Chaînage avant**, 83  
**Chaîne de valeur Web**, 83  
**Chaîne logistique**, 83  
**Chaines de blocs**, 127  
**Champ**, 83, 165  
**Champ clé**, 83  
**Changement de paradigme**, 83  
**charge utile**, 115  
**Charte de test**, 55  
**Chat room**, 152  
**Chatbot**, 11, 37, 127  
**Chef de projet maîtrise d'œuvre**, 11  
**Chemin**, 55  
**Chemin critique**, 137  
**Chemin d'accès**, 84  
**cheval de Troie**, 115  
**Cheval de Troie**, 84, 152  
**Chiffrement**, 84, 152  
**CIF**, 165  
**CIM**, 166  
**Circuit intégré**, 11  
**Classe**, 37, 140  
**Classe abstraite**, 37  
**Classe active**, 37  
**Classe concrète**, 37  
**Classe de base**, 37  
**Classe library**, 37  
**Classe paramétrable**, 37  
**Classe paramétrée**, 38  
**Classe racine**, 38  
**Classe utilitaire**, 38  
**Classe-association**, 38  
**Classification**, 38, 192  
**Classification dynamique**, 38  
**Classification statique**, 38  
**Classifieur Bayésien**, 192  
**Clé**, 38, 84  
**Clé étrangère**, 166  
**Clé naturelle**, 38  
**Clé primaire**, 38, 166

**Click-and-mortar**, 84  
**Client**, 11, 38, 84  
**Client-serveur**, 84  
**Cliquer-glisser ou glisser-déposer**, 84  
**Clôture des tests**, 55  
**Cloud computing**, 11, 128  
**Cloud Computing**, 84  
**CMMI**, 137  
**CMS**, 11  
**Code**, 38  
**Code machine**, 38  
**Code inatteignable**, 55  
**Code objet**, 38  
**Code orienté objet**, 38  
**Code source**, 38  
**Coffre-fort numérique**, 84  
**Cogniticien**, 84  
**Collaboration**, 38  
**Collecteur**, 115  
**Collecticiel**, 84  
**Collection**, 38  
**Comité d'entreprise**, 84  
**Commandes**, 38  
**Commerce collaboratif**, 84  
**Commerce électronique**, 11, 85  
**Commerce électronique de détail**, 85  
**Commerce électronique interconsommateurs**, 85  
**Commerce électronique mobile**, 85  
**Commutateur**, 85  
**Commutateur de paquets**, 85  
**Compatibilité**, 55  
**Compilateur**, 12  
**Complexité**, 55  
**Comportement**, 56, 140  
**Composant**, 34, 39, 55, 63, 103, 109, 122  
**Composants**, 141  
**Composition**, 39  
**Compression**, 85  
**Comptabilité par activités**, 85  
**Concentrateur**, 85  
**Conception**, 39  
**Conception assistée par ordinateur**, 85  
**Conception d'application en collaboration**, 85  
**conception d'interfaces adaptatives**, 115  
**Conception de systèmes**, 85  
**Conception des tests**, 56  
**Conception sociotechnique**, 86  
**Concurrence**, 167  
**Condition**, 39, 56  
**Condition de test**, 56  
**conférence informelle**, 115  
**Configuration**, 56  
**Configurer**, 86  
**Conformité**, 56  
**Connaissance**, 86  
**Connaissance explicite**, 86  
**Connaissance semi-structurée**, 86  
**Connaissance structurée**, 86  
**Connaissance tacite**, 86  
**Connecteur**, 86  
**Connectivité**, 86  
**Constat**, 56  
**Constructeur**, 39  
**Construction**, 39  
**Conteneur**, 39  
**Contexte**, 39  
**Contrainte**, 40  
**Contrainte d'intégrité**, 167  
**Contrat**, 40  
**Contributory infringement**, 152  
**Contrôle**, 40, 67, 86  
**Contrôle des tests**, 56  
**Contrôle qualité**, 56  
**Contrôles à l'entrée**, 86  
**Contrôles à la sortie**, 86  
**Contrôles administratifs**, 86  
**Contrôles d'accès**, 86  
**Contrôles des applications**, 86  
**Contrôles généraux**, 86  
**Contrôles sur le traitement**, 87  
**Conversion**, 87  
**Cookie**, 12, 87, 152  
**Copyleft**, 152  
**CORBA**, 34  
**Corbeille**, 87  
**Correctifs**, 87  
**Couche**, 40  
**Couplage**, 40  
**Courrier électronique**, 87  
**Coût de la qualité**, 56  
**Coût total de possession**, 87  
**Couverture**, 56  
**Couverture de code**, 56  
**Couverture des décisions**, 56  
**Couverture des instructions**, 56  
**Couverture fonctionnelle**, 51  
**Couverture structurelle**, 57  
**Cracker**, 87  
**CRC**, 40  
**Creative commons**, 152  
**Critère d'acceptation**, 57  
**Critères d'entrée**, 57  
**CRM**, 12, 87

**CRUD**, 166  
**Cryptage**, 87  
**CSS**, 40  
**CTRL**, 87  
**Cyberconférence**, 116  
**Cybercriminalité**, 88  
**Cybergouvernement**, 87  
**Cybernaut**, 87  
**Cybersécurité**, 12  
**Cybersquatting**, 152  
**Cybersurveillance**, 152  
**Cybervandalisme**, 88  
**Cycle**, 40, 62  
**Cycle de test**, 57  
**Cycle de v**, 40  
**Cycle de vie**, 137  
**Cycle de vie du développement logiciel**, 57  
**Cycle de vie logiciel**, 57  
**Cycle d'horloge**, 87  
**D.N.S.**, 90  
**D.O.S.**, 90  
**D.P.I.**, 90  
**D.V.D**, 91  
**D.V.I.**, 91  
**Data**, 40  
**Data Analyst**, 88  
**Data Encryption Standard**, 153  
**Data Mart**, 12  
**Data mining**, 167, 194  
**Data science**, 194  
**Data scientist**, 12  
**Data Scientist**, 88  
**Data Structures**, 40  
**Datacenter**, 12  
**Datamart**, 88  
**Datamining**, 12, 88  
**Datawarehouse**, 12, 88  
**DCOM**, 34  
**DCSSI**, 153  
**Débit**, 88  
**Débogage**, 40, 88  
**Déboguer**, 57  
**Décision**, 57  
**Décision non structurée**, 88  
**Décision semi-structurée**, 88  
**Décision structurée**, 88  
**Déclencheur**, 169  
**Décomposition**, 40  
**Deeep learning**, 128  
**Deep learning**, 195  
**Deep linking**, 153  
**Défaillance**, 57  
**Défaut**, 57  
**Définition**, 88  
**Délégation**, 40  
**Dématérialisation**, 88  
**démon de messagerie**, 116  
**Déni de service**, 89  
**Densité de défauts**, 57  
**Dépendance**, 40  
**Dépendance fonctionnelle**, 169  
**Déploiement**, 13, 40  
**Deprecated**, 40  
**DES**, 153  
**Descripteur**, 195  
**Désintermédiation**, 89  
**Destructeur**, 40  
**Détection d'anomalie**, 195  
**Développement des systèmes**, 89  
**Développement orienté objet**, 89  
**Développement par composants**, 89  
**Développement par l'utilisateur final**, 89  
**Développement rapide d'applications**, 89  
**Device**, 89  
**Dévoisement**, 116  
**DHTML**, 89  
**Diagramme**, 41, 57  
**Diagramme activité**, 141  
**Diagramme d'aperçu des interactions**, 144  
**Diagramme d'ensemble**, 145  
**Diagramme d'États**, 57  
**Diagramme d'état-transition**, 145  
**Diagramme d'objet**, 145  
**Diagramme de cas d'utilisation**, 142  
**Diagramme de classes**, 141  
**Diagramme de communication**, 142  
**Diagramme de déploiement**, 144  
**Diagramme de flux de données**, 89  
**Diagramme de Gantt**, 137  
**Diagramme de profil**, 145  
**Diagramme de réseau**, 137  
**Diagramme de structure composite**, 146  
**Diagramme de temps**, 146  
**Diagramme des composants**, 143  
**Diagramme entité-relation**, 89  
**Diagrammes d'interaction**, 146  
**Diagrammes de séquence**, 146  
**diagrams types**, 147  
**Dialogueur**, 135  
**Dictionnaire de données**, 89  
**Dictionnaire des données**, 169  
**Digg-Like**, 153  
**Digital**, 13  
**Directeur des systèmes d'information**, 89

**directeur, -trice des données**, 116  
**Direction centrale de la sécurité des systèmes d'Information**, 153  
**Directory**, 89  
**Discret**, 41  
**Discriminant**, 41  
**Disponibilité**, 57  
**Disque dur**, 13, 90  
**Disquette de démarrage**, 90  
**Distribution**, 148  
**Django**, 41  
**DNS**, 13, 41  
**Documentation**, 13, 41, 55, 65, 66, 67, 70, 90, 104  
**Domain**, 90  
**Domain tasting**, 153  
**Domaine**, 41, 90  
**Domaine public**, 153  
**Done**, 137  
**Donnée**, 13  
**Données**, 90, 196  
**Données de test**, 58  
**Données de trafic**, 153  
**données ouvertes**, 116  
**Données relationnelle**, 170  
**Dorsale de réseau**, 90  
**Dossier**, 90  
**Dossier compressé**, 90  
**Dossier de démarrage**, 90  
**Dossier de Test**, 51  
**Download**, 90, 153  
**Drag and drop**, 90  
**Driver**, 90  
**DSI**, 13, 91  
**Dynamic Data Exchange**, 14  
**Dysfonctionnement**, 51  
**E.I.D.E.**, 91  
**EAI**, 91  
**E-business**, 91  
**Echange de données informatisé**, 91  
**Échec**, 58  
**ECMA International**, 14  
**ECMAScript**, 14  
**E-commerce**, 14  
**Econométrie**, 14  
**Economie de réseaux**, 91  
**Ecosystème d'affaires**, 91  
**Ecran tactile**, 91  
**Ecriture miroir**, 91  
**EDI**, 14, 153  
**Éditeur logiciel**, 14  
**Effet de sonde**, 58  
**Efficacité**, 58  
**Efficiency**, 58  
**EIS**, 91  
**EJB**, 34  
**Elaboration**, 41  
**Élément**, 36, 41, 45, 48, 75  
**Élément de configuration**, 58  
**Élément de couverture**, 58  
**Élément de données**, 91  
**Élément de modélisation**, 41  
**Élément de visualisation**, 41  
**emoticon**, 109  
**Emoticon (**, 91  
**Encapsulation**, 41, 148  
**encre électronique**, 116  
**Enregistrement**, 92  
**Enregistrer**, 91  
**enregistreur de frappe**, 116  
**Enterprise Application Integration**, 14  
**Enterprise Resource Planning (ERP)**, 14  
**Entité**, 41, 92, 170  
**Entrée**, 58  
**Entrée :**, 92  
**Entrées de test**, 58  
**Entrepôt de données**, 92  
**Entrepôts de données**, 170  
**Entreprise numérique**, 92  
**Enumération**, 41  
**environnement de développement**, 117  
**Environnement de test**, 58  
**Environnement opérationnel**, 58  
**Équilibrage de charge**, 92  
**Ergonomie**, 92  
**Ergonomie web**, 15  
**ERP**, 34, 92  
**Erreur**, 41, 58, 81  
**ESN**, 15  
**espace de confiance**, 117  
**Espace de nom**, 41  
**Espace des états**, 41  
**Espace virtuel**, 15  
**Estimation d'erreur**, 58  
**Estimation des tests**, 59  
**Étalonnage**, 92  
**Etat**, 41, 48  
**Etend**, 41  
**Ethernet**, 15  
**ETL**, 15  
**Etude de faisabilité**, 92  
**Etude d'opportunité**, 41  
**Etude pilote**, 92  
**EV-DO**, 92

**Événement**, 41  
**Event Manager**, 42  
**Exception**, 42  
**Exclusion**, 92  
**Exécution des tests**, 59  
**Executive Information System**, 15  
**Exercé**, 59  
**Exigence**, 59  
**Exigence fonctionnelle**, 59  
**Exigence non-fonctionnelle**, 59  
**Exigence testable**, 59  
**expérience de l'utilisateur**, 117  
**expert, -e en mégadonnées**, 117  
**Exploration de données**, 130  
**exploration de modèle**, 117  
**Export**, 42  
**Expression**, 40, 42, 80  
**Extension**, 15, 42  
**Extension de fichier**, 15, 92  
**Externalisation**, 93  
**Extranet**, 15, 93  
**Extreme programming**, 59  
**F.A.Q.**, 93  
**F.T.P.**, 94  
**Facilitateur**, 59  
**Factorisation**, 42  
**Fair use**, 154  
**FAT/FAT32**, 93  
**Feedback**, 137  
**fenêtre d'attente**, 117  
**Fenêtre publicitaire**, 93  
**Feuille de style**, 93  
**Fiabilité**, 59  
**Fibre optique**, 93  
**Fichier**, 93  
**Fichier système**, 93  
**Fichiers compressés**, 93  
**Fichiers Internet temporaires**, 93  
**Fil torsade**, 93  
**File Transfer Protocol**, 15  
**Filtrage par paquets**, 93  
**Filtre de courrier**, 93  
**Firewall**, 15, 19, 93  
**Firewalls**, 154  
**Flaming**, 154  
**Flash**, 154  
**Flat webdesign**, 15  
**Flot de contrôle**, 42, 59  
**Flot de données**, 42, 59  
**Flot d'exécution**, 42  
**Flux d'information**, 148  
**flux de dépêches**, 117  
**Flux RSS**, 15, 94, 154  
**Fonction**, 42  
**Fonction (appel)**, 42  
**Fonction(définition)**, 42  
**Fondeur**, 94  
**Font stack**, 16  
**Forage de données**, 94  
**Format de fichier**, 94  
**Format paysage**, 94  
**Forum**, 16  
**Forum de discussion**, 154  
**Fournisseur d'accès à Internet**, 94  
**Fournisseur de service Internet**, 94  
**Framework**, 16, 42, 94  
**Framework de test unitaire**, 60  
**Framing**, 154  
**Free software**, 154  
**Freeware**, 94, 154  
**Fréquence d'horloge**, 94  
**Frimousse**, 94  
**Front end**, 16  
**FTP**, 15, 16  
**Fuite mémoire**, 59  
**Full stack**, 16, 42  
**fureteur,,** 118  
**Gang name**, 154  
**Garde**, 42  
**Gateway**, 94  
**Geek**, 16  
**Généralisation**, 42  
**Généralisation multiple**, 43  
**Génération**, 43  
**Générique**, 43  
**Genetic algorithm**, 185  
**Génie logiciel**, 16, 43, 94  
**Gestion de configuration**, 60  
**Gestion de configurations**, 43  
**Gestion de cycle de vie du produit**, 94  
**Gestion de la chaîne logistique**, 95  
**Gestion de la qualité**, 60  
**Gestion de la qualité totale**, 95  
**Gestion de la relation client**, 95  
**Gestion de liste LISTSERV**, 95  
**gestion de versions**, 118  
**Gestion de versions**, 43  
**Gestion des connaissances**, 95  
**Gestion des défauts**, 60  
**Gestion des processus d'affaires**, 95  
**Gestion des risques**, 60  
**Gestion des tests**, 60  
**Gestion électronique documentaire**, 95  
**Gestion flux de travail**, 95

**gestionnaire d'évènement**, 118  
**Gestionnaire de réseaux**, 95  
**Gif animé**, 95  
**Global Positioning System**, 17  
**GNU**, 16  
**Goto**, 43  
**Gouvernance informatique**, 17  
**GPS**, 17  
**Grabbing**, 154  
**graphe de connaissances**, 118  
**Grapheur**, 17  
**Green Computing**, 95  
**Green IT**, 17  
**Groupement**, 171  
**Groupware**, 17, 95  
**GUI**, 17  
**Guichet**, 118  
**Guide d'installation**, 60  
**H.T.T.P.**, 96  
**Hackathon**, 17, 196  
**Hacker**, 95, 154  
**Hacking**, 95  
**Hameçonnage**, 95  
**Happy slapping**, 154  
**Hardware**, 95  
**Harnais de test**, 60  
**Header**, 17  
**HEAP**, 43  
**Héritage**, 43  
**Héritage de (pour la) réalisation**, 43  
**Héritage d'interface**, 43  
**Héritage et polymorphisme**, 148  
**Héritage multiple**, 43  
**Heuristiques**, 130  
**Hiérarchie**, 43  
**Historique**, 96  
**Home page**, 96  
**Hotlist**, 96  
**Html**, 154  
**HTML**, 18, 96  
**http**, 34  
**HTTP**, 18  
**HTTPS**, 155  
**Hyper Text Markup Language**, 18  
**Hyper Text Mark-up Language**, 155  
**Hyper Text Transfer Protocol**, 18  
**Hyperlien**, 155  
**Hypertext**, 96  
**Hypertexte**, 18  
**I.D.E.**, 96  
**I.M.A.P.**, 96  
**IA**, 18, 130, 135  
**IA explicable**, 131  
**laaS**, 96  
**Icône**, 96  
**IDDN**, 155  
**Ide**, 43  
**IDEAL**, 60  
**Identifiant**, 43, 96, 171  
**Identité**, 43  
**Idiome**, 43  
**IHM**, 60  
**Imagette**, 118  
**Implémentation**, 96  
**Implémentation des tests**, 60  
**Import**, 44  
**impression en ligne**, 118  
**Inclusion**, 96  
**Incohérence des données**, 96  
**Incrément**, 137  
**Indépendance du test**, 60  
**Index**, 172  
**Index de rentabilité**, 96  
**indexation personnelle**, 118  
**Indicateur**, 18, 96  
**Indicateur de performance**, 61  
**Inégalité de concentration**, 196  
**Infocentre**, 18  
**Infogérance**, 96  
**Information**, 18, 97  
**Informatique**, 18  
**Informatique à haut disponibilité**, 97  
**Informatique à la demande**, 97  
**Informatique autonome**, 97  
**informatique en nuage**, 119  
**Informatique mobile**, 97  
**Informatique nomade**, 97  
**Infrastructure à clé publique**, 97  
**infrastructure à la demande**, 119  
**Infrastructure de test**, 61  
**Infrastructure technologique**, 97  
**ingénierie inverse**, 119  
**Inline linking**, 155  
**Inspection**, 61  
**Inspection dynamique**, 97  
**Instance**, 44  
**Instruction**, 61  
**Instruction exécutable**, 61  
**instructions**, 38  
**INTEGER**, 44  
**Intégration**, 44, 61, 97  
**Intégration continue**, 138  
**Intégration fonctionnelle**, 61  
**Intégrité référentielle**, 172

**Intelligence artificielle**, 18, 97, 130, 135, 197  
**Inter Deposit Digital Number**, 155  
**Interaction**, 44  
**Interactivite**, 155  
**Interdépendance des données et des programmes**, 97  
**Interface**, 18, 44, 98  
**interface avec l'utilisateur**, 119  
**Interface utilisateur**, 61, 98  
**Internet**, 19, 98  
**internet clandestin**, 119  
**Internet Protocol**, 31  
**Internet Protocol Adress**, 155  
**Interopérabilité**, 61, 98  
**Interpréteur**, 19, 44  
**Interruption**, 44  
**Intervenant métier**, 147  
**Intranet**, 19, 98  
**Invariant**, 44  
**Invocation**, 44  
**IPSEC**, 155  
**IPv6**, 98  
**ISO** :, 19  
**ISO 17799**, 98  
**Isolation forest**, 197  
**IT**, 30  
**Itérateur**, 44  
**Itération**, 44, 138  
**Jalon**, 61  
**Java**, 19  
**Jeton d'authentification**, 98  
**Jeux de Test**, 51  
**JQuery**, 44  
**Juste à temps**, 98  
**k-Nearest Neighbours**, 199  
**Knowledge Management**, 98  
**KPI**, 19  
**L4G**, 20  
**Landing page**, 19  
**Langage d'interrogation**, 98  
**Langage de balisage extensible**, 98  
**Langage de balisage hypertexte**, 98  
**Langage de description des services Web**, 98  
**Langage de manipulation de données**, 98  
**Langage de modélisation unifié**, 98  
**Langage de programmation**, 19  
**Langage de quatrième génération**, 20, 99  
**Langage de requête**, 99  
**Langage orienté objet**, 99  
**Large bande**, 99  
**Learning management system**, 20  
**Lecture basée sur la perspective**, 61  
**Liaison dynamique**, 44  
**Liaison louée**, 20  
**Liaison statique**, 44  
**Licence CeCILL**, 155  
**Licence creative commons**, 155  
**Licence open source**, 155  
**Licence Shrink-Wrap**, 155  
**Licence utilisateur**, 99  
**Lien**, 44  
**Lien commercial**, 156  
**Lien Hypermedia**, 156  
**Lien hypertexte**, 20, 99  
**Lien Hypertexte**, 156  
**Lien profond**, 156  
**Lien sponsorisé**, 156  
**Ligne de vie**, 44  
**Linux**, 99  
**Liste de diffusion**, 99  
**Livraison**, 45  
**Logiciel**, 20, 62, 99  
**logiciel à la demande**, 119  
**logiciel antivirus**, 119  
**Logiciel commercial sur étagère**, 62  
**Logiciel de base**, 99  
**Logiciel espion**, 99  
**Logiciel freeware**, 156  
**Logiciel libre**, 20, 99  
**logiciel malveillant**, 119  
**logiciel rançonneur**, 120  
**Logiciel shareware**, 156  
**Logiciels de base de données**, 172  
**Login**, 99, 156  
**Logique floue**, 100  
**Loi des grands nombres**, 200  
**Loi gaussienne**, 201  
**Machine learning**, 201  
**Machine Learning**, 20, 125, 126, 127, 128, 131  
**Macro**, 100  
**Macro-virus** :, 100  
**Mail bombing ou Spam**, 100  
**Mailbombing**, 156  
**Mailing list**, 100  
**Mailto**, 100  
**Maintenabilité**, 62  
**Maintenance**, 45, 62, 100  
**Maître d'œuvre**, 20, 138  
**Maître d'ouvrage**, 20, 21  
**Maître d'ouvrage délégué**, 21  
**Maître d'ouvrage opérationnel**, 21  
**Maître d'ouvrage stratégique**, 21  
**Maîtrise d'ouvrage**, 138  
**marathon de programmation**, 120

**Mashups**, 100  
**Masquage**, 120  
**matériel à la demande**, 120  
**Matériel informatique**, 21  
**Matrice de Couverture Fonctionnelle**, 51  
**Maturité**, 62  
**MCD**, 22, 173  
**Mécanisme**, 45  
**Membre**, 45  
**Mémoire morte**, 100  
**Mémoire tampon**, 100  
**Mémoire vive**, 100  
**Merise**, 173  
**Message**, 45  
**Messagerie électronique**, 22  
**Messagerie unifiée**, 100  
**Mesure**, 62  
**Métaclasse**, 45  
**Métadonnée**, 22  
**Métamodèle**, 45  
**Métamodélisation**, 45, 148  
**Meta-tags**, 156  
**Méthode**, 45  
**Méthode de classe**, 45  
**Méthode de la période de récupération**, 100  
**Méthode d'instance**, 45  
**Métrie**, 62  
**Miage Maîtrise de Méthodes Informatiques**  
    **Appliquées à la Gestion**, 100  
**MIC**, 22  
**Microblogage**, 120  
**Microblogging**, 101  
**Microblogue**, 120  
**Micro-ordinateur**, 22  
**Microprocesseur**, 22  
**Middleware**, 101  
**Migration**, 22, 101  
**Minitel**, 22  
**Mise à jour**, 101  
**MLD**, 174  
**MMS**, 102  
**MOA**, 22, 101  
**Mode**, 45  
**Mode de transfert asynchrone**, 101  
**Modèle**, 22, 45, 149  
**Modèle client-serveur**, 101  
**Modèle conceptuel de données**, 22  
**Modèle d'évaluation du prix des options**  
    **réelles**, 101  
**Modèle de croissance de fiabilité**, 62  
**Modèle de cycle de vie**, 62  
**Modèle de développement incrémental**, 62  
**Modèle de développement itératif**, 62  
**Modèle de développement séquentiel**, 62  
**Modèle de Markov caché**, 132  
**Modèle de référence**, 101  
**Modèle des forces concurrentielles**, 101  
**Modèle des points de fonction**, 101  
**Modèle économique**, 101  
**Modèle en V**, 63  
**Modèle organisationnel du flux poussé**, 101  
**Modèle organisationnel du flux tiré**, 101  
**Modèle OSI**, 23  
**Modèles de « bandits »**, 202  
**Modélisation**, 23, 45, 175  
**Modem**, 23, 101  
**modérateur**, 120  
**Modérateur**, 63  
**Modificateur**, 45  
**Modularité**, 45  
**Modulation par impulsion et codage**, 22  
**Module**, 45  
**MOE**, 23, 102  
**Mois-homme**, 102  
**Monomorphisme**, 45  
**Mot de passe**, 102  
**Moteur d'interférence**, 102  
**Moteur de base de données**, 174  
**MP3, MP4, MPEG**, 156  
**MPD**, 174  
**Multidiffusion**, 102  
**Multimedia**, 156  
**Multimédia**, 23  
**Multiplateforme**, 23  
**Multiplexeur**, 23  
**Multiplicité**, 45  
**MVC**, 23, 102  
**MVNO**, 102  
**Mystification**, 102  
**Nanotechnologie**, 102  
**Nappe**, 102  
**Navigabilité**, 45  
**Navigateur web**, 23  
**Navigateur Web**, 102  
**Nearshore**, 102  
**Nerd**, 23  
**Netiquette**, 156  
**Nettoyage de données**, 103  
**Neurone artificiel**, 135  
**Newsgroup**, 156  
**NFC**, 103  
**Niveau de maturité**, 46  
**Niveau de risque**, 63  
**Niveau de test**, 63

**Niveau OSI**, 103  
**Nœud**, 46  
**Nom de domaine**, 103, 157  
**Non interprété**, 46  
**Normalisation**, 202  
**Norme**, 63, 103  
**Norme technologique**, 103  
**NoSQL**, 176  
**Notation**, 46  
**Note**, 46  
**Notes**, 24  
**ntelligibilité**, 61  
**N-tiers**, 78  
**Numérique**, 103  
**Numéris**, 24  
**Numériser**, 103  
**O.C.R.**, 103  
**O.S.**, 104  
**Object Linking and Embedding**, 24  
**Objectif de test**, 63  
**Object-Oriented Creation and Learning**, 24  
**Objet**, 103  
**Objet actif**, 46  
**Objet de test**, 63  
**Objet Entité**, 46  
**Occultation d'information**, 46  
**Occurrence**, 175  
**OCL**, 149  
**octet**, 37  
**Octet**, 103  
**Œuvre**, 24  
**Œuvre du domaine public**, 157  
**Œuvre en partage**, 157  
**Œuvre en usage partagé**, 157  
**Œuvre libre de droit**, 157  
**Œuvre multimedia**, 157  
**Offshore**, 103, 138  
**OLAP**, 176  
**OLE**, 24  
**OLTP**, 177  
**Onduleur**, 103  
**Onglet**, 103  
**On-line data service**, 103  
**OOCL**, 24  
**Open data**, 24  
**Open source**, 24, 104  
**Opération Élément**, 46  
**Opération abstraite**, 46  
**Opt-in**, 157  
**option d'adhésion**, 120  
**option de retrait (à)**, 121  
**Opt-out**, 157  
**OQL**, 177  
**Oracle de test**, 63  
**Ordinateur**, 24  
**Ordinateur-zombie**, 157  
**Organigramme hiérarchique**, 104  
**Organigramme virtuelle**, 104  
**Organizer**, 104  
**Orienté-objet (langage)**, 24  
**Orienté-objet (modèle)**, 24  
**ORM**, 178  
**OS**, 29  
**Outil d'analyse dynamique**, 63  
**Outil de conception de tests**, 63  
**Outil de contrôle formel**, 104  
**Outil de couverture**, 63  
**Outil de gestion de configuration**, 64  
**Outil de gestion des défauts**, 64  
**Outil de gestion des exigences**, 64  
**Outil de gestion des tests**, 64  
**Outil de monitoring**, 64  
**Outil de planification formelle**, 104  
**Outil de préparation des données de tests**, 64  
**Outil de test**, 64  
**Outil de tests de performances**, 64  
**Outil d'exécution des tests**, 63  
**Outils de localisation**, 104  
**ouverture des données**, 121  
**Ouvrage**, 25  
**Overfitting**, 202  
**P.A.O.**, 105  
**P.C.I.**, 105  
**P.C.I. Express**, 105  
**P.G.P.**, 105  
**P.P.P.**, 107  
**PaaS**, 104  
**Page d'accueil**, 104  
**Page Web**, 104  
**Page Web Dynamique**, 104  
**Pair-programming**, 138  
**Palette**, 104  
**papier électronique**, 121  
**Paquet**, 25  
**Paquetage Élément**, 46  
**Paradigme**, 46  
**Paramétrage**, 105  
**Paramètre**, 46  
**parcours sur la toile**, 121  
**Pare-feu**, 25, 105, 157  
**Partage de dossiers**, 105  
**Partie privée**, 46  
**Partie protégée**, 46  
**Partie publique**, 46

**Partition**, 46, 105  
**Partition d'équivalence**, 64  
**Partitionnement d'équivalence**, 64  
**Passenger name record**, 157  
**Password**, 105  
**Patch**, 105  
**Pattern**, 46  
**PC**, 25  
 Peer-to-peer, 25, 158  
**Perceptron**, 132  
**Performance**, 65  
**Périphérique**, 105  
 Permalien, 25  
**Persistence**, 47  
**Personal Computer**, 25  
**Personnalisation**, 105  
**Personnalisation du Web**, 105  
**Perspective sociotechnique**, 105  
 PGI, 34  
**Phase**, 47  
**Phishing**, 105, 158  
**Pièce jointe**, 106  
**Pilotage des tests**, 65  
**Pilote**, 65, 106  
**Pirate informatique**, 106  
**Pixel**, 47  
**Pixel invisible**, 106  
**Place de marché électronique**, 106  
**Plan de revue**, 65  
**Plan de test**, 65  
**Plan de Test**, 51  
**Plan de test maître**, 65  
**Plan de tests**, 106  
**Planification de la continuité des services**, 106  
**Planification de la reprise après sinistre**, 106  
**Planification de test**, 65  
**Planning de test**, 65  
 Planning game, 138  
**Planning Poker**, 65  
**Plateforme (économie des)**, 106  
**Plateforme collaborative**, 106  
**Plateforme de service**, 106  
**Player**, 158  
 Plug-in, 25  
 PMI, 138  
**PNR**, 158  
**Point d'accès sans fil à Internet**, 107  
**Point Par Pouce**, 107  
**Point to Point Protocol**, 107  
**Politique d'autorisation**, 107  
**Politique de test**, 65  
**Polymorphisme**, 47  
**Pont**, 25  
**Pop-up**, 158  
**Portabilité**, 65  
**portail de messagerie**, 121  
**Post-condition**, 47  
**Postcondition**, 65  
**Pouce**, 107  
**Pré-condition**, 47  
**Précondition**, 66  
**Presse-papier**, 107  
**Prévision**, 25  
**Priorité**, 66  
**Prise de contrôle à distance**, 107  
**Problème**, 66  
**Procédure**, 47  
**Procédure de test**, 66  
**Procédure stockée**, 178  
**Procédures stockées**, 107  
 Processeur, 25, 107  
**Processus**, 25, 47, 66  
**Processus de test**, 66  
**Product backlog item**, 138  
**Product owner**, 138  
**Profil d'habilitation**, 158  
 Progiciel, 25  
 Programmation, 26  
**Programme**, 47, 107  
**Projection**, 47  
**Projet**, 66  
**Propriété**, 47, 178  
**Propriétés ACID**, 179  
 Protocole, 26, 107  
**Protocole de communication**, 26  
**Protocole IP**, 158  
**Prototypage**, 107  
**Prototype**, 47  
**Provisionnement**, 121  
 Proxy, 26  
**Proxy**, 108  
**Pseudo**, 108  
**Pseudo-état**, 47  
 Puce, 26  
**Qualité**, 66  
**Qualité logicielle**, 66  
**R.A.M.**, 108  
**R.O.M.**, 108  
**Raccourci**, 108  
**Radio frequency identification**, 158  
**RAM**, 26  
**Random Access Memory**, 26  
**Rapport d'avancement de test**, 66

**Rapport de défaut**, 66  
**Rapport de synthèse de test**, 66  
**Rapport de test**, 67  
**Rapport de Test**, 52  
**Rapport d'incident**, 67  
**Rational Unified Process**, 67  
**Read Only Memory**, 26  
**Réalisation d'une mesure**, 66  
 Réalité augmentée, 26  
 Réalité virtuelle, 26  
**Recettage unitaire**, 108  
**Recette utilisateur**, 52  
**Récurtivité**, 47  
**Rédaction coopérative**, 26  
**Redémarrage**, 108  
**Réduction des risques**, 67  
**Refactoring**, 138  
 Référencement, 26  
**référencement abusif**, 121  
**Référentiel**, 27  
**Réflexive**, 47  
**Registrar**, 158  
**Régression**, 67, 202  
**Réification**, 47  
**Relecture technique**, 67  
**Répertoire**, 108  
**Réplication**, 27, 108  
**Reporting**, 27  
**Reporting de test**, 67  
**Requête**, 34, 179  
**Réseau de neurones artificiels**, 132, 135  
 Réseau informatique, 27  
**Réseau local de PC**, 27  
**réseau neuronal profond**, 128  
**Réseau Numérique à Intégration de Services**,  
 27  
**Réseau privé virtuel**, 158  
 Résolution, 27, 108  
**Responsabilité**, 47  
**Responsable de domaine**, 27  
 Responsive design, 27  
**Restauration du système**, 108  
**Restaurer**, 108  
**Reste à faire**, 138  
**Restriction**, 47  
**Résultat**, 67  
**Résultat attendu**, 67  
**Résultat de décision**, 67  
**Résultat effectif**, 67  
 rétro-ingénierie, 121  
**Rétro-ingénierie**, 47  
**Rétropropagation**, 132, 202  
**Rétrospective**, 139  
**Réunion rétrospective**, 68  
**Réussite**, 67  
**Réutilisation**, 48  
**Réviseur**, 68  
**Revue**, 48, 68  
**Revue ad-hoc**, 68  
**Revue basée sur des scénarios**, 68  
**Revue basée sur les rôles**, 68  
**Revue basée sur une checklist**, 68  
**Revue de pairs**, 68  
**Revue formelle**, 68  
**Revue informelle**, 68  
**Revue informelle collective**, 68  
**Revue technique**, 68  
**RFDI**, 158  
**RFID**, 108  
 RGPD, 27  
**Riper**, 122  
**Risque**, 48, 68  
**Risque produit**, 68  
**Risque projet**, 69  
**Risque qualité**, 69  
**RLPC**, 27  
**RNA**, 132  
**RNIS**, 27  
**Roadmap**, 139  
 Robotique, 28  
**Robustesse**, 69  
**Rôle**, 48  
**ROM**), 26  
**Rootkit**, 159  
**Routeur**, 28, 159  
**RSS**, 159  
**RUN (exécuter)**, 48  
**RUP**, 67  
**S.M.B.**, 109  
**S.M.T.P.**, 110  
 SaaS, 29  
**Safe harbor**, 159  
**Salle de marché**, 28  
**Sampling**, 159  
**Saut de page**, 108  
**Sauvegarder**, 108  
**Scanner**, 108  
**Scénario**, 48  
**Scénarios de Test**, 52  
**Schéma relationnel**, 180  
**Schème**, 48  
**Scoring**, 159  
**Screeener**, 159  
**Scribe**, 69

Script, 28, 109  
**Script de test**, 69  
**Scrum**, 69, 139  
**ScrumMaster**, 139  
**Secteur d'amorçage**, 109  
 Secure Socket Layer, 28  
**Sécurité**, 69  
 Sécurité informatique, 28  
**Segmentation**, 28  
**SEI**, 48  
**Sélecteur**, 48  
**Sémantique**, 28  
**Série**, 109  
**Série chronologique**, 28  
**Server**, 109  
**Serveur**, 29, 48  
**service de la toile**, 122  
 Service Oriented Architecture, 29  
**Service Web**, 34  
**Session**, 109  
**Session de test**, 69  
**Sévérité**, 69  
**SGBD**, 181  
 Shadow IT, 29  
**Shareware**, 109, 159  
**SHIFT**, 109  
**Short message service**, 159  
**SIAD**, 29  
**SIG**, 29  
**Signal**, 48  
**Signature**, 48  
**signature de virus**, 122  
**Signature de virus**, 109  
**Signature électronique**, 109  
**Signet**, 109  
**Simulateur**, 69  
**Simulation**, 69  
**Skin**, 109  
**Slamming**, 159  
**Smiley**, 109  
**SMS**, 159  
**Snapback**, 159  
 SOA, 29, 34  
**Socket**, 110  
**Software**, 110  
 Software as a Service, 29  
**Sortie**, 69  
**Sous-classe**, 48  
**Sous-état**, 48  
**Sous-système**, 48  
**Space-shifting**, 159  
**Spam**, 29, 110, 159  
**Spamdexing**, 159  
**Spamming**, 159  
**Spécialisation**, 48  
**Spécification**, 48, 69  
**Spécification de cas de test**, 70  
**Spécification de composant**, 70  
**Spécifications fonctionnelles**, 110  
**Spécifications techniques**, 110  
**Sponsoring**, 160  
**Sprint**, 139  
**Spyware**, 110, 160  
**SQL**, 29, 110, 181  
 SSL, 28, 110  
**SSL/TLS**, 160  
**Stabilité**, 70  
**STACK**, 48  
**STAD**, 160  
**Standard Query Language**, 29  
**Stéréotype**, 48, 150  
**Stochastique**, 132  
**Stratégie**, 29  
**Stratégie de test**, 70  
**Streaming**, 160  
**Structure**, 48  
**Structure composite**, 150  
**Structurée**, 48  
**Suck**, 160  
**Suite de tests**, 70  
**Super-classe**, 49  
**Super-état**, 49  
**Surcharge**, 49  
**Sûreté**, 70  
**Synchrone**, 49  
**Synchronisation**, 49  
**synthèse vocale de texte**, 122  
**Système**, 70  
**système d'engagement**, 122  
**Système d'exploitation**, 29, 110  
**Système d'information**, 30, 110  
**Système de fichiers**, 110  
**Système de traitement automatisé de données**, 160  
**système d'information géographique**, 29  
**Système informatique**, 30  
**Système informatique d'aide à la décision**, 29  
**Système sous test**, 70  
**Systèmes experts**, 132  
**Table**, 181  
**Table d'allocation de fichier**, 110  
**Table de décision**, 70  
**Tableur**, 30  
**Tabulation**, 110

**Tâche de fond**, 110  
**Tag**, 111  
**Taux de défaillance**, 70  
**Taux de transfert**, 111  
**TCP/IP**, 31, 111  
**Team building**, 139  
**Technique de test**, 70  
**Technique de test basée sur l'expérience**, 70  
**Technique de test boîte-blanche**, 71  
**Technique de test boîte-noire**, 74  
**Téléchargement**, 111  
**Téléchargement ascendant**, 160  
**Téléchargement descendant**, 160  
**Temps réel**, 49  
**Test**, 49, 71, 75  
**Test basé sur des sessions**, 71  
**Test basé sur les modèles**, 71  
**Test basé sur les risques**, 71  
**Test basé sur l'expérience**, 71  
**Test basé sur une checklist**, 71  
**Test boîte-blanche**, 71  
**test captcha**, 122  
**Test d'acceptation**, 71  
**Test d'intégration**, 111  
**Test d'intégration de composants**, 73  
**Test d'interopérabilité**, 73  
**Test d'utilisabilité**, 73  
**Test d'acceptation réglementaire**, 73  
**Test d'accessibilité**, 71  
**Test de cas d'utilisation**, 72  
**Test de composant**, 72  
**Test de confirmation**, 72  
**Test de maintenance**, 72  
**Test de performance**, 74  
**Test de portabilité : (portability testing)**, 73  
**test de reconnaissance humaine**, 122  
**Test de régression**, 74  
**Test de sécurité : (security testing)**, 72  
**Test de Turing**, 132  
**Test des décisions**, 72  
**Test des transitions d'état**, 74  
**Test dynamique**, 73  
**Test fonctionnel**, 73  
**Test Leader**, 74  
**Test Manager**, 74  
**Test par tables de décisions**, 74  
**Test piloté par les données**, 74  
**Test unitaire**, 111  
**Testeur**, 75  
**Tests Alpha**, 71  
**Tests Beta: (beta testing)**, 71  
**Tests d'acceptation opérationnelle**, 72  
**Tests d'intégration**, 73  
**Tests d'intégration système**, 73  
**Tests d'acceptation contractuelle**, 71  
**Tests d'acceptation utilisateur**, 72  
**Tests de charge**, 72  
**Tests de Non Régression**, 52  
**Tests de recette métier**, 52  
**Tests de Validation**, 52  
**Tests des instructions**, 72  
**Tests d'intégration**, 52  
**Tests dirigés par mots-clé**, 73  
**Tests exploratoires**, 73  
**Tests non-fonctionnels**, 73  
**Tests unitaires**, 52  
**Testware**, 75  
**Thème**, 111  
**Théorie de l'information**, 30  
**TIC**, 30  
**Tierce Maintenance Applicative**, 111  
**Timeboxing**, 139  
**TLN**, 133  
**toile**, 122  
**toile profonde**, 123  
**toile sémantique**, 123  
**Token Ring**, 30  
**Topologie**, 49  
**Torrent**, 160  
**Traçabilité**, 75  
**Tracking**, 30  
**Traducteur**, 30  
**Traitement de texte**, 30  
**Traitement du langage naturel**, 133  
**Trame**, 30  
**Transactions**, 181  
**transcription automatique de la parole**, 123  
**Transistor**, 31  
**Transition**, 49  
**Transition automatique**, 49  
**Transition d'état**, 75  
**Transmission Control Protocol**, 31  
**Transpac**, 31  
**Tunneling**, 160  
**Tutoriel**, 31  
**Typage**, 49  
**Type**, 49  
**Type de donnée abstrait**, 49  
**Type de risque**, 75  
**Type de test**, 75  
**Type primitif**, 49  
**Types de Diagrammes**, 147  
**Types de données**, 182  
**UC**, 111

UI, 31  
**UML**, 31, 150, 182  
**UMTS**, 111  
Unicode, 31  
**Unified Modelling Language**, 31  
**Unix**, 111  
**Upgrader**, 111  
**Urbanisme du système d'information**, 31  
URL, 31, 49  
**Use case**, 32, 49  
**User Story**, 75  
usurpation d'adresse électronique, 123  
UX, 32  
**V.G.A.**, 111  
**Valeur limite**, 75  
**Validation**, 75  
**Variable**, 49, 75  
**Variable de classe**, 50  
**Variable d'instance**, 49  
**Vectorel**, 111  
Vélocité, 139  
Ver, 123  
**Ver informatique**, 112  
**Vérification**, 75  
**Verrou SQL**, 183  
**Virtual Private Network**, 160  
**Virtuel**, 32  
virus, 123  
Virus informatique, 32, 112, 161  
**Visibilité**, 50  
**Vision**, 50  
VPN, 32, 161  
VPS, 32  
**Vue**, 50, 184  
**W.W.W.**, 112  
WAI, 32  
**WAN**, 32, 112  
**War Driving**, 112  
Web, 32, 112  
**Web 2.0**, 161  
**Web sémantique**, 112  
**Web site**, 112  
**Webconférence**, 112  
**Webforum**, 161  
**Weblog**, 161  
**Webmaster**, 112  
**Wide Area Network**, 32  
**Wideband Delphi**, 75  
WiFi, 112  
**Wiki**, 112  
Wirefarme, 32  
**Workflow**, 33, 34  
**Workflow d'une anomalie**, 52  
**Workspace**, 50  
**World Wide Web**, 33  
**WSOA**, 34  
**X25**, 33  
**X400**, 33  
**XAI**), 131  
**XP**, 59  
Zéro-day, 33  
ZIP, 33  
Zoning, 33