VOCABULAIRE DU TRAITEMENT DE DONNÉES

L'extraction de données, leur stockage, leur codage, leur saisie, leur fusion, leur tri… Toutes ces opérations sont effectuées par des humains et des ordinateurs qui transforment cette matière première en ressource exploitable. Le grand domaine de l'informatique qui regroupe ce type d'opérations est le traitement de données.

Plusieurs concepts fondamentaux relèvent de cet univers de chiffres, de caractères, de processus et d'analyses. Il suffit de s'attarder au partitionnement, à la recherche sémantique, à l'intégration de données et au traitement par lots pour s'apercevoir que le champ d'expertise des spécialistes qui y travaillent est vaste.

Pour faire connaître les termes à la base de ce domaine, l'Office québécois de la langue française, en collaboration avec la Faculté des sciences et de génie de l'Université Laval et Druide informatique, propose un vocabulaire de plus de quarante concepts.

Symboles

Termes privilégiés

Termes utilisés dans certains contextes

Termes déconseillés

Ce vocabulaire est accessible en ligne à l'adresse suivante :
oqlf.gouv.qc.ca/ressources/bibliotheque/dictionnaires/vocabulaire-traitement-donnees.aspx.
Version PDF du 30 juin 2025

Avertissement : Lors de la conversion du format HTML au format PDF, il est possible que certains caractères spéciaux ou signes typographiques (comme les espaces insécables) n'aient pas été correctement conservés. En cas de disparité, c'est la version en ligne du vocabulaire qui prévaut.

Index

A

analyse des données, 1
analyse par groupe, 2
anonymisation, 3

B

base de données multidimensionnelle, 4
bloc de données, 5

C

capture de données, 6
centre de données à très grande échelle, 7
centre de traitement de données, 8

D

données brutes, 9
données continues, 10
données de diffusion en continu, 11
données discrètes, 12
données internes, 13
données non structurées, 14
données partenaires, 15
données semi-structurées, 16
données structurées, 17
données tierces, 18

E

entrepôt de données, 19
expert en science des données, 20

F

fusion de données, 21

G

génération de variables explicatives, 22
gestion de données, 23
groupement de données par classe, 24

H

hypercube de données, 25

I

informatique en périphérie, 26
informatique géodistribuée, 27
ingestion de données, 28
intégration de données, 29
intégrité des données, 30

J

jeu de données, 31

L

lac de données, 32

M

métadonnée, 33
minientrepôt de données, 34
modèle de données, 35
modélisation des données, 36

N

nettoyage de données, 37

P

petites données, 38
plateforme de gestion des données, 39
processus d'extraction, de chargement et de transformation des données, 40
processus d'extraction, de transformation et de chargement des données, 41

R

recherche sémantique, 42

S

science des données, 43

T

traitement de données, 44
traitement intégré de données, 45
traitement massivement parallèle, 46
traitement par lots, 47

V

valorisation de données, 48

1. analyse des données

Définition
Technique de recherche consistant à dégager des corrélations au sein d'ensembles de données à l'aide d'outils informatiques spécialisés ou d'algorithmes, et à en tirer des conclusions permettant d'orienter les actions ou d'influencer la prise de décision.

Notes
L'analyse des données est généralement sous-catégorisée en analyse prédictive, prescriptive, descriptive et de diagnostic.

analyse des données n. f.
analytique des données n. f.

anglais
data analytics
DA
analytics
data analysis

En anglais, le terme *data analysis* est parfois employé pour désigner plus précisément l'analyse des données orientée vers la compréhension du passé, et le terme *data analytics* pour désigner l'analyse visant à la prédiction de ce qui est à venir.

---

2. analyse par groupe

Définition
Méthode d'analyse qui consiste à grouper, au moyen de systèmes logiciels fonctionnant librement, des objets ou des éléments qui sont semblables ou différents au sein d'un ensemble, de manière à dégager des tendances ou à tirer des conclusions.

Notes
Le groupement de données par classe est une forme d'analyse par groupe.

La méthode d'analyse par groupe est semblable à celle de l'apprentissage non supervisé de l'intelligence artificielle.

analyse par groupe n. f.
analyse par regroupement n. f.
analyse par segmentation n. f.
analyse typologique n. f.
analyse par grappe n. f.

Il existe dans l'usage en français plusieurs variantes des termes *analyse par groupe* et *analyse par grappe*. Le complément de nom peut être précédé de la préposition *de* (ex. : *analyse de groupe*, *analyse de grappe*) ou il peut prendre la forme du pluriel (ex. : *analyse par groupes*, *analyse par grappes*, *analyse de groupes* ou *analyse de grappes*).

anglais
cluster analysis
segmentation analysis
clustering

---

3. anonymisation

Définition
Procédure qui consiste à remplacer les renseignements personnels contenus dans un document par un code d'identification, avant de les supprimer définitivement, de manière à rendre virtuellement impossible l'identification des individus auprès desquels ils ont été recueillis.

Notes
Les renseignements pouvant mener à l'identification d'un individu sont, notamment, le nom, l'adresse, le numéro de carte de crédit, le numéro d'assurance sociale.

Il ne faut pas confondre l'anonymisation avec la dépersonnalisation, par laquelle les renseignements personnels ne sont pas supprimés définitivement.

anonymisation n. f.
dénominalisation irréversible n. f.

anglais
anonymization
anonymisation

---

4. base de données multidimensionnelle

Définition
Base de données dans laquelle les données sont représentées sous forme d'hypercubes.

Notes
Le stockage des données est réalisé de façon à optimiser l'accès aux informations en cas de requêtes non prévues lors de la création de la base.

base de données multidimensionnelle n. f.
BDM n. f.

anglais
multidimensional database
MDB
MDD
MDDB

---

5. bloc de données

Définition
Ensemble de données de même nature organisé de manière à être traité et distribué comme un seul élément.

Notes
Un bloc de données est généralement contenu dans un seul fichier débutant par un en-tête de bloc qui fournit diverses métadonnées sur le contenu et la nature.

bloc de données n. m.
bloc n. m.

En France, le terme *bloc* a été officialisé par la Commission générale de terminologie et de néologie, en 1998.

anglais
data block
block

---

6. capture de données

Définition
Acquisition, réalisée autrement que par la saisie conventionnelle, de données destinées à être converties en format traitable par ordinateur et à être utilisables immédiatement.

Notes
Une capture de données peut s'effectuer pendant qu'une transaction ou une opération s'exécute (par exemple, l'enregistrement automatique d'un code à barres pour la gestion du stock, au moment même où ce code est lu par un terminal de point de vente pour préparer la facture), ou encore au cours d'un balayage effectué par un scanneur sur un document papier, pour en extraire des données répondant à des paramètres de recherche précis.

capture de données n. f.

anglais
data capture

---

7. centre de données à très grande échelle

Définition
Centre de traitement de données caractérisé par une grande rapidité de traitement de mégadonnées, une capacité à répondre à de très forts volumes de demandes, et dont l'architecture prévoit généralement l'utilisation de l'infonuagique.

centre de données à très grande échelle
n. m.

centre de données hyperscale

Le terme *centre de données hyperscale* est déconseillé parce qu'il est employé depuis peu en français et qu'il ne s'intègre pas au système du français. En anglais, le mot *scale* désigne une échelle de mesure.

anglais
hyperscale data center
hyperscale data centre

---

8. centre de traitement de données

Définition
Lieu où sont regroupées des ressources informatiques et de télécommunication afin d'offrir des services de gestion, de stockage et de traitement des données informatiques.

Notes
Les centres de traitement de données disposent généralement de plusieurs mesures de sécurité (système contre l'incendie ou contre le vol, système d'alimentation d'urgence, etc.). La température et l'humidité y sont contrôlées.

centre de traitement de données n. m.
centre de données n. m.
centre de traitement informatique n. m.
CTI n. m.
centre de traitement de l'information n. m.

Par métonymie, ces termes peuvent également désigner l'organisation qui gère un tel lieu.

Le terme *centre de données*, calqué sur l'anglais *data center*, est acceptable parce qu'il est légitimé en français au Québec et ailleurs en francophonie et qu'il s'intègre au système linguistique du français.

anglais

data center
data centre
data processing center
DPC
data processing centre

---

9. données brutes

Définition

Ensemble de données présentées sous la forme où elles ont été collectées, avant d'avoir subi un quelconque traitement ou une interprétation.

Notes

Des données brutes peuvent être, par exemple, des données météorologiques enregistrées par un ordinateur, des données d'enquête recueillies par une maison de sondage ou des notes prises par un professionnel de la santé lors d'un rendez-vous avec un patient.

données brutes n. f. pl.

anglais

raw data

---

10. données continues

Définition

Données quantitatives représentées par des valeurs numériques d'une séquence continue de valeurs possibles.

Notes

Les données continues, contrairement aux données discrètes, sont mesurables. Elles peuvent être décomposées en fractions ou en décimales. Par exemple, la taille d'une personne et la température sont des données continues.

données continues n. f. pl.

anglais

continuous data

---

11. données de diffusion en continu

Définition

Ensemble des flux de données pouvant provenir de sources multiples, qui sont générés, diffusés et analysés de manière continue et en temps réel dans un système de traitement de données.

données de diffusion en continu n. f. pl.
données diffusées en continu n. f. pl.
données en flux continu n. f. pl.
flux de données en continu n. m. pl.
données en continu n. f. pl.

données de streaming
données en streaming

Les emprunts hybrides *données de streaming* et *données en streaming* ne sont pas acceptables parce que *streaming* a été emprunté à l'anglais depuis peu de temps pour former ces appellations et qu'il ne s'intègre pas au système linguistique du français. De plus, plusieurs termes bien construits sont disponibles en français pour exprimer le présent concept.

anglais

streaming data

---

12. données discrètes

Définition

Données quantitatives ou qualitatives représentées par des valeurs énumérables et identifiables séparément.

Notes

Les données discrètes, contrairement aux données continues, sont comptables. Elles ne peuvent être constituées que de nombres entiers. Par exemple, le nombre d'enfants dans une famille ou l'âge (arrondi) d'une personne sont des données discrètes. Par ailleurs, les données discrètes peuvent comporter des données qualitatives ordinales.

données discrètes n. f. pl.

anglais

discrete data

---

13. données internes

Définition

Ensemble des données que possède une entreprise sur ses clients.

Notes

Les données internes proviennent notamment de la fréquentation du site de l'entreprise, de ses points de vente, des réseaux sociaux.

données internes n. f. pl.
données de première main n. f. pl.

anglais

first-party data

---

14. données non structurées

Définition

Données stockées sans être organisées de manière prédéfinie, ce qui rend leur utilisation plus difficile pour un système d'information.

Notes

On distingue les données non structurées textuelles, comme celles qui se trouvent dans les documents Word, et les données non structurées non textuelles, comme les images ainsi que les fichiers audio et vidéo.

On trouve également des données structurées et des données semi-structurées.

données non structurées n. f. pl.

anglais

unstructured data

---

15. données partenaires

Définition

Données obtenues par une entreprise auprès de partenaires.

Notes

Les données partenaires complètent les données internes, notamment dans une perspective de portée commerciale accrue.

données partenaires n. f. pl.
données secondaires n. f. pl.
données de seconde main n. f. pl.

anglais

second-party data

---

16. données semi-structurées

Définition

Données dont l'organisation ne correspond pas à un modèle prédéfini, mais comportant des informations associées telles que des métadonnées ou des marqueurs sémantiques qui facilitent leur analyse.

Notes
Une page HTML, par exemple, contient des données semi-structurées.

On trouve également des données structurées et des données non structurées.

données semi-structurées n. f. pl.

Semi- se joint par un trait d'union au nom ou à l'adjectif qu'il accompagne.

anglais
partially structured data
semi-structured data

---

17. données structurées

Définition
Données stockées selon un format prédéfini, de façon à permettre leur interprétation par un logiciel.

Notes
Des dates, des numéros de téléphone sont des exemples de données structurées.

L'utilisation de données structurées entraîne un meilleur référencement sur Internet.

On trouve également des données semi-structurées et des données non structurées.

données structurées n. f. pl.

anglais
structured data

---

18. données tierces

Définition
Données achetées par une entreprise auprès de fournisseurs extérieurs.

Notes
Les données clients sont un exemple de données tierces.

données tierces n. f. pl.
données externes n. f. pl.

anglais
third-party data

---

19. entrepôt de données

Définition

Système de stockage dans lequel est centralisé un volume important de données issues des différentes sources internes et externes d'une organisation, et qui est conçu pour offrir un accès rapide à l'information stratégique nécessaire à la prise de décision.

Notes

Dans un entrepôt de données, les données sont sélectionnées et préparées (pour répondre aux questions vitales de l'organisation), intégrées (à partir des différentes sources de renseignements) et datées (elles gardent la trace de leur origine).

entrepôt de données n. m.
dépôt de données n. m.

En France, le terme *entrepôt de données* a été officialisé par la Commission générale de terminologie et de néologie, en 2007.

anglais

data warehouse

---

20. expert en science des données, experte en science des données

Définition

Personne spécialisée dans la collecte, l'exploration, l'analyse et l'interprétation des données, et qui a notamment pour tâche d'orienter les actions et les prises de décisions d'une organisation.

Notes

Il est généralement attendu de l'expert en science des données qu'il ait des connaissances étendues en statistique et en informatique.

L'expert en science des données peut notamment être appelé à concevoir des modèles prédictifs ou encore à coordonner les travaux d'autres spécialistes du domaine, comme les analystes de données.

expert en science des données n. m.
experte en science des données n. f.
expert en mégadonnées n. m.
experte en mégadonnées n. f.

En France, le terme *expert en mégadonnées* et sa variante sont recommandés officiellement par la Commission d'enrichissement de la langue française, depuis 2017.

scientifique des données

Le terme *scientifique des données*, calqué sur l'anglais *data scientist*, est déconseillé parce qu'il ne s'intègre pas au système linguistique du français. En effet, le présent concept désigne une personne ayant des connaissances approfondies sur un sujet particulier, ici les données, plutôt qu'un scientifique ayant pour objectif d'étendre les connaissances dans son domaine d'étude. Le terme *scientifique des données* est par ailleurs peu fréquemment employé en français.

anglais

data scientist

---

21. fusion de données

Définition
Opération qui consiste à combiner des données issues de sources diverses afin de permettre leur exploitation en tant qu'ensemble.

Notes
La fusion de données a pour objectif d'extraire les informations les plus spécifiques et les plus pertinentes sur une entité, une activité, un événement ou un phénomène. Ces informations proviennent de sources telles que des capteurs, des bases de données, des systèmes de navigation ou des connaissances acquises a priori.

fusion de données n. f.

anglais
data fusion

---

22. génération de variables explicatives

Définition
Procédé consistant à extraire des données brutes d'un ensemble, à les traiter et à les recoder afin qu'elles soient exploitées le mieux possible par les algorithmes d'apprentissage.

génération de variables explicatives n. f.
création de variables explicatives n. f.
conception de variables explicatives n. f.

anglais
features engineering
feature engineering

---

23. gestion de données

Définition
Organisation des données dans un système informatique, qui détermine, tout au long de leur cycle de vie, la manière dont elles sont enregistrées, structurées, conservées en mémoire et distribuées en toute sécurité.

gestion de données n. f.
gestion des données n. f.

anglais
data management
DM

---

24. groupement de données par classe

Définition

Méthode d'analyse des données dans laquelle celles-ci sont réparties dans des groupes distincts, sur la base de caractéristiques communes ou de différences, qui sont repérées par des systèmes logiciels fonctionnant librement.

Notes

Le groupement de données par classe est une forme d'analyse par groupe.

Cette méthode d'analyse de données est semblable à celle de l'apprentissage non supervisé de l'intelligence artificielle.

groupement de données par classe n. m.
partitionnement de données n. m.
regroupement de données n. m.
segmentation de données n. f.

Dans l'usage, on trouve également *partitionnement des données*, *regroupement des données*, *groupement des données* et *segmentation des données*.

anglais

data clustering
clustering
binning

---

25. hypercube de données

Définition

Représentation graphique des informations contenues dans une base de données, sous forme de structure multidimensionnelle permettant de prendre en compte plusieurs axes d'analyse.

hypercube de données n. m.
hypercube n. m.
cube multidimensionnel n. m.
cube de données n. m.
cube n. m.

On emploie souvent les synonymes *cube de données* et *cube* pour désigner ce concept, bien qu'un hypercube comporte normalement plus de trois dimensions.

anglais

hypercube
multidimensional cube
data cube
cube

---

26. informatique en périphérie

Définition

Traitement des données qui se fait en périphérie de réseau, à l'échelle des utilisateurs finaux, par le biais des objets connectés.

Notes

En traitant et en stockant les données au moyen d'objets connectés (par exemple les téléphones, les montres, les véhicules automobiles, les terminaux domotiques) et non pas dans un nuage informatique, l'informatique en périphérie permet de réduire le volume de données circulant sur le réseau infonuagique ainsi que le délai de transit.

informatique en périphérie n. f.
informatique en périphérie de réseau n. f.

Le terme *informatique en périphérie*, calqué sur l'anglais *edge computing*, est acceptable parce qu'il est intégrable au système linguistique du français. En effet, le concept de « périphérie » évoque, aussi bien en français qu'en anglais, l'idée d'un traitement de données complètement décentralisé.

anglais

edge computing

---

27. informatique géodistribuée

Définition

Traitement de données dont la structure sert d'intermédiaire entre l'infonuagique et l'informatique en périphérie, de sorte que les données soient distribuées vers l'endroit le plus adéquat.

Notes

L'algorithme qui distribue les données vers les nuages informatiques ou vers les appareils périphériques connectés se base sur plusieurs critères comme la vitesse de récupération, l'accessibilité aux données, le contrôle du trafic sur les réseaux, etc.

informatique géodistribuée n. f.
informatique en brouillard n. f.

Le terme *informatique en brouillard*, calqué sur l'anglais *fog computing*, s'intègre au système linguistique du français, sa formation étant analogue à celle du terme *informatique en nuage*.

anglais

fog computing
fog networking
fogging

---

28. ingestion de données

Définition

Collecte d'un grand volume de données et importation de celles-ci dans une structure de stockage.

Notes

Les données peuvent provenir de sources différentes, comme des fichiers CSV, du contenu en continu, des interfaces de programmation d'applications, des bases de données, etc., et se présenter sous différents formats. Elles devront par la suite être modifiées pour former un ensemble cohérent.

La structure de stockage peut être un entrepôt de données, une base de données, etc.

L'ingestion de données peut se faire en temps réel ou par lots, à intervalles réguliers.

ingestion de données n. f.

anglais

data ingestion

---

29. intégration de données

Définition

Opération qui consiste à combiner des données qui proviennent de sources hétérogènes de manière qu'elles forment un ensemble cohérent et uniforme.

intégration de données n. f.
intégration des données n. f.

anglais

data integration

---

30. intégrité des données

Définition

Propriété des données qui ne subissent aucune altération accidentelle ou non autorisée lors de leur traitement, de leur transmission ou de leur conservation.

intégrité des données n. f.
intégrité n. f.

anglais

data integrity
integrity

---

31. jeu de données

Définition

Groupe de données cohérent et structuré, portant sur un sujet déterminé, accessible pour la consultation ou le téléchargement en un ou plusieurs fichiers.

Notes

Les jeux de données existent sous différents formats, notamment le fichier CSV et le document en langage XML.

Un jeu de données se présente parfois sous la forme d'un tableau dans lequel chaque colonne (ou ligne) correspond à une variable, et chaque ligne (ou colonne), à la valeur de celle-ci.

jeu de données n. m.
ensemble de données n. m.

anglais

data set

---

32. lac de données

Définition

Infrastructure logicielle extensible qui s'appuie sur un service de fichiers répartis et sur la distribution des opérations au sein d'une grappe de serveurs pour stocker et interroger des mégadonnées de tous types.

Notes

Les lacs de données permettent de stocker autant des données structurées que des données non structurées, des fichiers journaux que le contenu généré par la communication entre machines.

Les lacs de données tirent profit du processus d'extraction, de chargement et de transformation des données pour l'interrogation de volumes importants de données disparates.

lac de données n. m.

Le terme *lac de données*, calqué sur l'anglais *data lake*, est acceptable parce qu'il est intégrable au système linguistique du français. En effet, l'analogie faite avec le lac, considéré comme un vaste réservoir alimenté par de multiples sources (rivières, ruisseaux) se transpose bien en français et permet d'exprimer tant la grande quantité de données stockées que la multiplicité de portes d'entrée permettant d'alimenter le lac de données. Une telle analogie avec une étendue d'eau se retrouve notamment dans le terme mer de portes.

anglais

data lake architecture
data lake

---

33. métadonnée

Définition

Donnée qui renseigne sur la nature de certaines autres données dans le but d'en faciliter la compréhension et la gestion.

Notes

Le nom d'un fichier, sa taille, sa date de création ou de modification sont des exemples de métadonnées.

métadonnée n. f.

anglais

metadata

---

34. minientrepôt de données

Définition

Système de stockage calqué sur le modèle d'un entrepôt de données, qui correspond à une base de données spécialisée se rapportant à un secteur d'activité particulier d'une organisation ou à un métier qui y est exercé.

minientrepôt de données n. m.
comptoir de données n. m.

Le minientrepôt peut contenir un volume important de données. Le terme *minientrepôt* ne fait donc pas référence au fait qu'il est petit, mais au fait qu'il constitue une solution moins lourde à gérer qu'un entrepôt de données.

Les mots formés avec le préfixe *mini-* s'écrivent généralement sans trait d'union.

anglais

data mart

---

35. modèle de données

Définition

Représentation schématique d'un ensemble de données et de leurs relations à l'intérieur d'une base de données.

Notes

Le modèle de données permet d'inventorier, de décrire et d'organiser les données pour lesquelles la base de données est créée.

modèle de données n. m.

anglais

data model

---

36. modélisation des données

Définition

Opération par laquelle on définit le modèle de données, selon les besoins des utilisateurs.

Notes

La modélisation des données permet également de conditionner la structure de la base de données et de définir les règles de manipulation des données qui y sont stockées.

modélisation des données n. f.
modélisation de données n. f.

anglais

data modelling
data modeling

---

37. nettoyage de données

Définition

Opération par laquelle sont détectées, corrigées, remplacées ou éliminées les données d'un ensemble stocké, qui ne sont pas conformes à certaines règles, qui présentent des anomalies de forme ou qui sont incomplètes, de manière à assurer leur cohérence.

Notes

Avec l'avènement des techniques spécialisées d'exploration de données et d'extraction de connaissances, cette opération de nettoyage a une importance de plus en plus grande et va bien au-delà de la simple correction d'erreurs d'ordre technique. C'est en effet sur l'exactitude des données premières, à partir desquelles s'effectuent les analyses, que reposent la qualité ainsi que la fiabilité de l'information et des connaissances qui en sont tirées.

nettoyage de données n. m.
nettoyage des données n. m.

anglais

data cleaning
data cleansing
data scrubbing

---

38. petites données

Définition

Ensemble structuré de données rassemblées pour répondre à un besoin précis, d'un volume assez petit pour être contenues dans un seul ordinateur et être gérées et analysées par un humain à l'aide des outils classiques de gestion de bases de données ou de gestion de l'information.

Notes

Les fichiers de photos contenus dans un ordinateur familial, les données de vente d'une petite ou moyenne entreprise ainsi que les résultats des parties jouées dans une ligue sportive constituent des exemples de petites données.

Les petites données peuvent être issues d'une source locale ou être extraites d'un ensemble de mégadonnées.

petites données n. f. pl.
données de petite taille n. f. pl.

anglais

small data
little data

---

39. plateforme de gestion des données

Définition

Plateforme permettant, à l'aide de témoins, de données partenaires ou de données tierces, de collecter et d'assembler des données pour les exploiter à des fins de commercialisation.

Notes

Les réseaux sociaux et les traces laissées par la navigation dans les sites Web constituent des sources de données pour les plateformes de gestion de données.

Les plateformes de gestion de données permettent notamment l'optimisation du site d'une société et le ciblage des clients.

En Europe francophone, on utilise aussi le terme anglais *data management platform* et le sigle anglais *DMP* pour désigner ce concept en français.

plateforme de gestion des données n. f.
PGD n. f.
plateforme de gestion de données n. f.
PGD n. f.

anglais

data management platform
DMP

---

40. processus d'extraction, de chargement et de transformation des données

Définition

Processus de stockage de données pour lequel la transformation ou le traitement de celles-ci sont réalisés après leur déplacement dans la base de données de destination, en tirant profit des ressources locales de cette dernière.

Notes

Le processus d'ECT est au cœur de l'architecture des lacs de données.

Il ne faut pas confondre le processus d'extraction, de chargement et de transformation des données et le processus d'extraction, de transformation et de chargement des données.

processus d'extraction, de chargement et de transformation des données n. m.
ECT n. m.
processus d'ECT n. m.
processus ECT n. m.

anglais

extraction, loading, transformation process
ELT
ELT process
extract, load, transform process

---

41. processus d'extraction, de transformation et de chargement des données

Définition

Processus de stockage de données pour lequel la transformation ou le traitement de celles-ci sont réalisés parallèlement à leur déplacement vers une base de données de destination.

Notes
Le processus d'ETC est au cœur de l'architecture des entrepôts de données.

Il ne faut pas confondre le processus d'extraction, de transformation et de chargement des données et le processus d'extraction, de chargement et de transformation des données.

processus d'extraction, de transformation et de chargement des données n. m.
ETC n. m.
processus d'ETC n. m.
processus ETC n. m.

On trouve parfois le sigle *ETC* dans des collocations qui font l'ellipse du mot *processus*, comme *outil d'ETC* et *méthode ETC*.

Les formes courtes *processus d'ETC*, *processus ETC* et *ETC* se rencontrent plus fréquemment que la forme longue *processus d'extraction, de transformation et de chargement des données*.

anglais
extraction, transformation, loading process
ETL
ETL process
extract, transform, load process

---

42. recherche sémantique

Définition
Technique permettant à un moteur de recherche, grâce à des algorithmes de langage naturel, d'interpréter les requêtes des internautes à partir de la signification contextuelle des mots utilisés dans celles-ci.

Notes
La recherche sémantique permet au moteur de recherche de fournir les résultats les plus pertinents possible.

La recherche sémantique est également mise à profit dans la technologie des assistants vocaux personnels.

recherche sémantique n. f.

anglais
semantic search

---

43. science des données

Définition
Science qui s'intéresse aux méthodes et aux pratiques permettant d'acquérir des connaissances à partir de l'extraction de données et ainsi d'orienter la prise de décision.

Notes
La science des données fait appel à de nombreuses disciplines, notamment les mathématiques, les statistiques et l'informatique.

Le spécialiste en science des données est appelé expert en science des données.

science des données n. f.

anglais

data science

---

44. traitement de données

Définition

Ensemble des opérations que les ordinateurs d'un système informatique peuvent effectuer sur des données dans le but de les transformer en ressources exploitables.

Notes

Les données font l'objet d'opérations telles que leur extraction, leur codage, leur saisie, leur fusion, leur tri, leur conversion, leur transfert, etc.

traitement de données n. m.
traitement des données n. m.
traitement de l'information n. m.

Les termes *traitement de données, traitement des données* et *traitement de l'information* sont également employés pour désigner le déroulement logique des opérations, selon un processus prédéterminé.

Certaines sources établissent une distinction entre les termes *traitement de données* et *traitement de l'information*. Au sens strict, le terme information implique que des données brutes ont acquis une signification particulière dans un contexte donné. Toutefois, plusieurs auteurs ne font pas cette distinction dans le cas des termes composés et, dans l'usage, on les utilise l'un pour l'autre.

traitement électronique de données n. m.
vieilli
traitement automatique de données n. m.
vieilli

anglais

data processing
DP
information processing
computerized data processing vieilli
automatic data processing vieilli
ADP vieilli
electronic data processing vieilli
EDP vieilli

---

45. traitement intégré de données

Définition

Traitement de données dans lequel les opérations effectuées automatiquement sur les données, qui proviennent généralement de sources diverses, sont coordonnées, harmonisées et intégrées au sein même d'un système informatique, de manière à répondre adéquatement aux besoins en ressources informationnelles d'une entreprise ou d'un organisme.

traitement intégré de données n. m.
traitement intégré des données n. m.
traitement intégré de l'information n. m.

anglais

integrated data processing
IDP
integrated information processing

---

46. traitement massivement parallèle

Définition

Mode de traitement de données utilisant un grand nombre de processeurs de façon coordonnée et simultanée.

Notes

Dans le traitement massivement parallèle, chaque processeur possède sa propre mémoire, son propre système d'exploitation et sa propre partie des données, et communique avec les autres pour exécuter un même programme.

Le traitement massivement parallèle est notamment mis à profit dans les systèmes d'aide à la décision et dans les applications embarquées industrielles telles que le véhicule autonome.

traitement massivement parallèle n. m.

anglais

massively parallel processing
MPP

---

47. traitement par lots

Définition

Mode de traitement de données dans lequel celles-ci sont divisées en petits groupes pour former des unités destinées à être analysées séparément en une seule opération et à un moment déterminé.

Notes

Dans le cas des données de diffusion en continu, celles-ci ne sont pas traitées par lots, mais au fur et à mesure de leur transfert.

traitement par lots n. m.

En France, le terme *traitement par lots* est recommandé officiellement par la Commission d'enrichissement de la langue française, depuis 2000.

anglais

batch processing
deferred time processing
off-line processing

---

48. valorisation de données

Définition

Processus de collecte, de traitement et d'analyse de données, permettant l'utilisation optimale de celles-ci dans la poursuite d'un objectif donné.

Notes

La valorisation de données peut avoir comme résultats de faciliter la prise de décisions, de permettre la détection d'anomalies et de trouver la cause de certains problèmes.

La valorisation de données comporte plusieurs étapes, comme la détection, le stockage (au moyen d'un lac de données, par exemple), la manipulation, l'analyse et la restitution.

valorisation de données n. f.

anglais

data valorization

data valorisation

VOCABULAIRE DU TRAITEMENT DE DONNÉES

Pour accéder à l'ensemble des vocabulaires de l'Office québécois de la langue française :
oqlf.gouv.qc.ca/ressources/bibliotheque/dictionnaires/index_lexvoc.html.

Pour connaître les outils et les services linguistiques de l'Office :
vitrinelinguistique.oqlf.gouv.qc.ca/a-propos-de-la-vitrine-linguistique/offre-de-services-linguistiques.

Pour consulter les ressources de la Vitrine linguistique :
vitrinelinguistique.oqlf.gouv.qc.ca.

Pour visiter le site de l'Office :
oqlf.gouv.qc.ca/accueil.aspx.

© Office québécois de la langue française, 2025

Office québécois
de la langue
française
Québec