



Norme
internationale

ISO/IEC 5259-4

**Intelligence artificielle — Qualité
des données pour les analyses
de données et l'apprentissage
automatique (AA) —**

Partie 4:
**Cadre pour le processus de qualité
des données**

*Artificial intelligence — Data quality for analytics and machine
learning (ML) —*

Part 4: Data quality process framework

Première édition
2024-07

iTeh Standards
(<https://standards.iteh.ai>)
Document Preview

[ISO/IEC 5259-4:2024](https://standards.iteh.ai/catalog/standards/iso/43cfc2a4-9441-4d50-9f28-3fe804cf02c9/iso-iec-5259-4-2024)

<https://standards.iteh.ai/catalog/standards/iso/43cfc2a4-9441-4d50-9f28-3fe804cf02c9/iso-iec-5259-4-2024>



DOCUMENT PROTÉGÉ PAR COPYRIGHT

© ISO/IEC 2024

Tous droits réservés. Sauf prescription différente ou nécessité dans le contexte de sa mise en œuvre, aucune partie de cette publication ne peut être reproduite ni utilisée sous quelque forme que ce soit et par aucun procédé, électronique ou mécanique, y compris la photocopie, ou la diffusion sur l'internet ou sur un intranet, sans autorisation écrite préalable. Une autorisation peut être demandée à l'ISO à l'adresse ci-après ou au comité membre de l'ISO dans le pays du demandeur.

ISO copyright office
Case postale 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Genève
Tél.: +41 22 749 01 11
E-mail: copyright@iso.org
Web: www.iso.org

Publié en Suisse

Sommaire

Page

Avant-propos	v
Introduction	vi
1 Domaine d'application	1
2 Références normatives	1
3 Termes et définitions	1
4 Symboles et abréviations	3
5 Principes du processus de qualité des données	3
6 Cadre pour le processus de qualité des données	4
6.1 Généralités	4
6.2 Planification de la qualité des données	5
6.3 Évaluation de la qualité des données	6
6.4 Amélioration de la qualité des données	6
6.5 Validation du processus de qualité des données	6
6.6 Utilisation du CPQD	7
7 Processus de qualité des données pour l'AA	7
7.1 Généralités	7
7.2 Exigences en matière de données	8
7.3 Planification des données	9
7.4 Acquisition des données	9
7.5 Préparation des données	10
7.5.1 Généralités	10
7.5.2 AA supervisé	11
7.5.3 AA non supervisé	11
7.5.4 AA semi-supervisé	11
7.5.5 Composition de l'ensemble de données	11
7.5.6 Étiquetage des données	12
7.5.7 Annotation des données	12
7.5.8 Évaluation de la qualité des données	13
7.5.9 Amélioration de la qualité des données	13
7.5.10 Dé-identification des données	16
7.5.11 Encodage des données	16
7.6 Mise à disposition des données	17
7.6.1 Généralités	17
7.6.2 AA supervisé	17
7.6.3 AA non supervisé	17
7.6.4 AA semi-supervisé	17
7.7 Mise hors service des données	17
8 Méthodes et processus d'étiquetage des données	18
8.1 Généralités	18
8.2 Principes d'étiquetage des données	18
8.3 Méthodes d'étiquetage des données	18
8.4 Processus d'étiquetage des données	19
8.4.1 Généralités	19
8.4.2 Spécifications d'étiquetage	19
8.4.3 Rôles des participants à l'étiquetage	20
8.4.4 Outils ou plateformes d'étiquetage	20
8.4.5 Établissement des tâches d'étiquetage	20
8.4.6 Attribution des tâches d'étiquetage	20
8.4.7 Contrôle du processus d'étiquetage	21
8.4.8 Vérification de la qualité des résultats d'étiquetage	21
8.4.9 Révision des résultats d'étiquetage	22

9	Rôles des participants	22
9.1	Généralités.....	22
9.2	Planificateur de données.....	23
9.3	Créateur de données.....	23
9.4	Collecteur de données.....	23
9.5	Ingénieur données.....	23
9.6	Détenteur de données.....	23
9.7	Utilisateur des données.....	23
10	Processus de qualité des données pour l'AA semi-supervisé	23
10.1	Généralités.....	23
10.2	Exigences en matière de données.....	24
10.3	Planification des données.....	24
10.4	Acquisition des données.....	24
10.5	Préparation des données.....	24
10.6	Mise à disposition des données.....	24
10.7	Mise hors service des données.....	24
11	Processus de qualité des données pour l'apprentissage par renforcement	25
11.1	Généralités.....	25
11.2	Exigences en matière de données.....	25
11.3	Planification des données.....	25
11.4	Acquisition des données.....	25
11.5	Préparation des données.....	25
	11.5.1 Processus général.....	25
	11.5.2 Enregistrement des données.....	26
11.6	Mise à disposition des données.....	26
11.7	Mise hors service des données.....	26
12	Processus de qualité des données pour l'analyse de données	26
12.1	Généralités.....	26
12.2	Exigences en matière de données.....	26
12.3	Planification des données.....	27
12.4	Acquisition des données.....	27
	12.4.1 Généralités.....	27
	12.4.2 Chargement des données.....	27
	12.4.3 Stockage des données.....	27
12.5	Préparation des données.....	27
	12.5.1 Généralités.....	27
	12.5.2 Nettoyage des données.....	27
	12.5.3 Transformations des données.....	28
	12.5.4 Agrégation des données.....	28
	12.5.5 Évaluation de la qualité des données.....	28
	12.5.6 Amélioration de la qualité des données.....	28
12.6	Mise à disposition des données.....	29
12.7	Mise hors service des données.....	29
	Bibliographie	30

Avant-propos

L'ISO (Organisation internationale de normalisation) et l'IEC (Commission électrotechnique internationale) forment le système spécialisé de la normalisation mondiale. Les organismes nationaux membres de l'ISO ou de l'IEC participent au développement de Normes internationales par l'intermédiaire des comités techniques créés par l'organisation concernée afin de s'occuper des domaines particuliers de l'activité technique. Les comités techniques de l'ISO et de l'IEC collaborent dans des domaines d'intérêt commun. D'autres organisations internationales, gouvernementales et non gouvernementales, en liaison avec l'ISO et l'IEC, participent également aux travaux.

Les procédures utilisées pour élaborer le présent document et celles destinées à sa mise à jour sont décrites dans les Directives ISO/IEC, Partie 1. Il convient, en particulier, de prendre note des différents critères d'approbation requis pour les différents types de documents. Le présent document a été rédigé conformément aux règles de rédaction données dans les Directives ISO/IEC, Partie 2 (voir www.iso.org/directives ou www.iec.ch/members_experts/refdocs).

L'ISO et l'IEC attirent l'attention sur le fait que la mise en application du présent document peut entraîner l'utilisation d'un ou de plusieurs brevets. L'ISO et l'IEC ne prennent pas position quant à la preuve, à la validité et à l'applicabilité de tout droit de propriété revendiqué à cet égard. À la date de publication du présent document, l'ISO et l'IEC n'avaient pas reçu notification qu'un ou plusieurs brevets pouvaient être nécessaires à sa mise en application. Toutefois, il y a lieu d'avertir les responsables de la mise en application du présent document que des informations plus récentes sont susceptibles de figurer dans la base de données de brevets, disponible à l'adresse www.iso.org/brevets et <https://patents.iec.ch>. L'ISO et l'IEC ne sauraient être tenues pour responsables de ne pas avoir identifié tout ou partie de tels droits de propriété.

Les appellations commerciales éventuellement mentionnées dans le présent document sont données pour information, par souci de commodité, à l'intention des utilisateurs et ne sauraient constituer un engagement.

Pour une explication de la nature volontaire des normes, la signification des termes et expressions spécifiques de l'ISO liés à l'évaluation de la conformité, ou pour toute information au sujet de l'adhésion de l'ISO aux principes de l'Organisation mondiale du commerce (OMC) concernant les obstacles techniques au commerce (OTC), voir www.iso.org/iso/avant-propos. Pour l'IEC, voir www.iec.ch/understanding-standards.

Le présent document a été élaboré par le comité technique mixte ISO/IEC JTC 1, *Technologies de l'information*, sous-comité SC 42, *Intelligence artificielle*. [ISO/IEC 5259-4:2024](https://standards.iteh.ai/catalog/standards/iso/43cfc2a4-9441-4d50-9f28-3fe804cf02c9/iso-iec-5259-4-2024)
<https://standards.iteh.ai/catalog/standards/iso/43cfc2a4-9441-4d50-9f28-3fe804cf02c9/iso-iec-5259-4-2024>
Une liste de toutes les parties de la série ISO/IEC 5259 se trouve sur les sites Web de l'ISO et de l'IEC.

Il convient que l'utilisateur adresse tout retour d'information ou toute question concernant le présent document à l'organisme national de normalisation de son pays. Une liste exhaustive desdits organismes se trouve aux adresses www.iso.org/members.html et www.iec.ch/national-committees.