

ISO/IEC 5259-2:2024(fr)

ISO/IEC JTC 1/SC 42

Date: 2025-06-05

Première édition

2024-11

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

Partie 2:

Mesures de la qualité des données

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

Part 2: Data quality measures

ISO/IEC 5259-2:2024

<https://standards.iteh.ai/catalog/standards/iso/a1df5595-1f27-4bcb-835f-3e7f1e5354b8/iso-iec-5259-2-2024>

ISO/IEC 5259-2:2024(fr)

© ISO/IEC ~~2025~~ 2024

Tous droits réservés. Sauf prescription différente ou nécessité dans le contexte de sa mise en ~~œuvre~~ 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~~ CP 401 ~~+~~ - Ch. de Blandonnet 8

CH-1214 Vernier, ~~Genève~~ Geneva

~~Tél.:~~ +Phone: + 41 22 749 01 11

E-mail: copyright@iso.org

~~Web~~ Website: www.iso.org

Publié en Suisse

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

ISO/IEC 5259-2:2024

<https://standards.iteh.ai/catalog/standards/iso/a1df5595-1f27-4bec-835f-3e7f1e5354b8/iso-iec-5259-2-2024>

Sommaire-Page

Avant-propos iv

Introduction v

1 Domaine d'application 1

2 Références normatives 1

3 Termes et définitions 1

4 Symboles et abréviations 5

5 Composantes de la qualité des données et modèles de qualité des données pour l'analyse de données et l'apprentissage automatique 5

5.1 Composantes de la qualité des données dans le cycle de vie des données 5

5.2 Modèle de qualité des données 7

6 Caractéristiques et mesures de la qualité des données 9

6.1 Généralités 9

6.2 Caractéristiques de qualité des données inhérentes 10

6.3 Caractéristiques de qualité des données inhérentes et dépendantes du système 16

6.4 Caractéristiques de qualité des données dépendantes du système 20

6.5 Caractéristiques de qualité des données supplémentaires 21

7 Mise en œuvre d'un modèle de qualité des données et de mesures de la qualité des données pour une tâche d'analyse de données ou d'AA 31

8 Établissement de rapports sur la qualité des données 32

8.1 Cadre pour l'établissement de rapports sur la qualité des données 32

8.2 Informations relatives aux mesures de la qualité des données 32

8.3 Recommandations à l'intention des organismes 32

Annexe A (informative) Conception et documentation d'une fonction de mesure 34

Annexe B (informative) Modèle UML de cadre de mesure de la qualité des données 36

Annexe C (informative) Aperçu des caractéristiques de qualité des données 37

Annexe D (informative) Autres groupes de caractéristiques de qualité des données 39

Annexe E (informative) Comparaison entre les caractéristiques de qualité des données de l'ISO/IEC 25012 et de l'ISO/IEC 5259-2 41

Bibliographie 43

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*.

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.

Introduction

La prise de décision basée sur les données pose de nouveaux défis pour la gestion de la qualité des données dans le cadre de l'analyse de données et de l'intelligence artificielle (IA) basée sur l'apprentissage automatique (AA). Les problèmes de qualité des données, tels que des données incomplètes, fausses ou obsolètes, peuvent nuire aux processus et résultats d'analyse de données et d'AA. Les données provenant de sources diverses, y compris les données structurées (par exemple, bases de données relationnelles) et les données non structurées (par exemple, documents, images, audios), peuvent être directement utilisées dans le cycle de vie des données pour le développement de modèles d'analyse de données et d'AA. Les données sont transformées à chaque étape du cycle de vie des données dans le cadre de l'analyse de données et de l'AA. Une approche holistique normalisée pour contrôler, produire et fournir suffisamment de données de haute qualité est nécessaire pour que les modèles d'analyse de données et d'AA soient sûrs, fiables et interopérables. Pour développer une gestion crédible de la qualité des données pour l'analyse de données et l'AA, des Normes internationales sur la qualité intrinsèque des données, comprenant des concepts et cas d'utilisation, des caractéristiques et mesurages, des exigences de gestion et un cadre de processus, peuvent être envisagées.

Le présent document fait partie de la série ISO/IEC 5259. Il s'appuie sur la série ISO 8000, l'ISO/IEC 25012 et l'ISO/IEC 25024. L'objectif du présent document est de décrire un modèle de qualité des données en définissant des caractéristiques et mesures de la qualité des données basées sur l'ISO/IEC 25012 et l'ISO/IEC 25024. Les modèles de qualité des données peuvent être étendus ou modifiés conformément au présent document.

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

ISO/IEC 5259-2:2024

<https://standards.iteh.ai/catalog/standards/iso/a1df5595-1f27-4bcb-835f-3e7f1e5354b8/iso-iec-5259-2-2024>

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

Partie 2: Mesures de la qualité des données

1 Domaine d'application

Le présent document spécifie un modèle de qualité des données, des mesures de la qualité des données et des recommandations concernant l'établissement de rapports sur la qualité des données dans le contexte de l'analyse de données et de l'apprentissage automatique (AA).

Le présent document s'applique à tous les types d'organismes qui souhaitent atteindre leurs objectifs de qualité des données.

2 Références normatives

Les documents suivants sont cités dans le texte de sorte qu'ils constituent, pour tout ou partie de leur contenu, des exigences du présent document. Pour les références datées, seule l'édition citée s'applique. Pour les références non datées, la dernière édition du document de référence s'applique (y compris les éventuels amendements).

ISO/IEC 5259-1, *Intelligence artificielle — Qualité des données pour les analyses de données et l'apprentissage automatique (AA) — Partie 1: Vue d'ensemble, terminologie et exemples*

ISO/IEC 25024, *Ingénierie des systèmes et du logiciel — Exigences et évaluation de la qualité des systèmes et du logiciel (SQuARE) — Mesurage de la qualité des données*

ISO/IEC 22989, *Technologies de l'information — Intelligence artificielle — Concepts et terminologie relatifs à l'intelligence artificielle*

3 Termes et définitions

Pour les besoins du présent document, les termes et les définitions de l'ISO/IEC 5259-1, l'ISO/IEC 22989 ainsi que les suivants s'appliquent.

L'ISO et l'IEC tiennent à jour des bases de données terminologiques destinées à être utilisées en normalisation, consultables aux adresses suivantes:

— ISO Online browsing platform: disponible à l'adresse <https://www.iso.org/obp>

— IEC Electropedia: disponible à l'adresse <https://www.electropedia.org/>

3.1 3.1 donnée

représentation réinterprétable d'une information sous une forme conventionnelle convenant à la communication, à l'interprétation ou au traitement