

Recueil du Symposium de 2024 de Statistique Canada : Le futur des statistiques officielles

Sur l'interaction des exigences légales, des aspects de la qualité et des risques éthiques lors de l'utilisation de l'apprentissage automatique dans les statistiques officielles allemandes

par Florian Dumpert

Date de diffusion : le 8 septembre 2025



Statistique
Canada

Statistics
Canada

Canada

Sur l'interaction des exigences légales, des aspects de la qualité et des risques éthiques lors de l'utilisation de l'apprentissage automatique dans les statistiques officielles allemandes

Florian Dumpert¹

Résumé

L'intelligence artificielle (IA), et son sous-domaine qu'est l'apprentissage automatique (AA), ont trouvé leur place dans l'administration en général et dans les statistiques officielles en Allemagne en particulier. Le présent article souligne les enjeux éthiques qui peuvent surgir lors de l'utilisation de l'IA et de l'AA dans les statistiques officielles. Il examine s'il faut créer un cadre éthique distinct pour composer de manière appropriée avec ces enjeux, comme le proposent les établissements institutionnels d'autres pays et les institutions intergouvernementales de la statistique officielle. On présente les résultats de cette étude pour montrer que la mise en œuvre des exigences des cadres de référence existants qui, dans la plupart des cas, ne sont pas propres à l'IA ou à l'AA, comme la loi et la qualité, est déjà suffisante pour tenir compte de manière adéquate des enjeux éthiques se fondant sur des scénarios de risque.

Mots-clés : Intelligence artificielle; apprentissage automatique; loi; qualité; éthique.

1. Introduction

Il y a environ 10 ans, les organismes nationaux de statistique ont commencé à se pencher sur l'apprentissage automatique, un domaine des algorithmes. Certains considèrent que ces domaines sont associés à l'informatique, tandis que d'autres jugent qu'ils sont associés à la statistique. Sans égard à l'identité des développeurs des algorithmes, les organismes de statistique peuvent les utiliser. D'ailleurs, ils le font de plus en plus. Cette situation est favorisée par une communauté active qui met en œuvre et crée efficacement des méthodes d'apprentissage automatique en R et Python. Alors que le caractère expérimental était au premier plan au début, des questions organisationnelles, méthodologiques, juridiques, éthiques et en matière de personnel sont apparues au fil du temps. On n'a pas encore obtenu réponse à toutes ces questions. Il en va de même pour les attentes associées à l'apprentissage automatique dans les bureaux. Même si l'on a observé un certain degré de scepticisme à l'égard de la nouveauté au début, les attentes ont aussi été temporairement exagérées. Dans certains cas, on s'attendait à ce que l'apprentissage automatique (souvent associé à de nouvelles sources de données, mot-clé « mégadonnées ») permette de relever le « défi olympique de la statistique publique » de Holt, à savoir permettre la production de statistiques élargies, plus approfondies, plus rapides, meilleures et moins chères, car « les statisticiens officiels ont tiré parti des percées technologiques pour améliorer le rendement [...] dans le cas des statisticiens, d'énormes avantages ont été tirés des avancées techniques en matière de collecte, de traitement et de présentation des données [...] [et] de meilleures méthodologies et des avancées au chapitre des méthodes et de l'analyse statistique [traduction] » (Holt, 2007, p. 1). Évidemment, cette situation ne s'est pas concrétisée. Cependant, l'utilisation de l'apprentissage automatique pour produire des statistiques officielles a évidemment donné lieu à des réussites, souvent à la phase 5 (« processus ») du Modèle générique du processus de production statistique (CEE-ONU, 2019), c'est-à-dire pour certaines tâches, comme

¹ Florian Dumpert, Office fédéral de la statistique de l'Allemagne, Gustav-Stresemann-Ring 11, 65189 Wiesbaden, Allemagne, florian.dumpert@destatis.de. Cette participation au Recueil du Symposium 2024 de Statistique Canada se fonde directement et, en partie, mot pour mot, sur les articles (publiés en allemand; deepl.com a été utilisé pour de quelques traductions) de Florian Dumpert, de Jannik Reichel, d'Elisa Oertel, d'Holger Leerhoff et de Christian Salwiczek qui ont été publiés dans *WISTA – Wirtschaft und Statistik* en février 2025 (Dumpert et coll., 2025a, 2025b). Cette participation au Recueil ne renferme donc pas de perspectives fondamentalement nouvelles ou de perspectives qui dépassent celles-ci figurant dans ces articles. En revanche, on peut trouver des détails et des références additionnelles sur le sujet de cet article dans Dumpert et coll. (2025a, 2025b).

l'intégration des données, la classification et le codage, la vérification et l'imputation. Habituellement, ces réussites découlent du fait que les étapes de traitement nécessaires sont réalisées de manière plus efficace. Cela a été rendu possible en remplaçant le travail de routine manuel par des ordinateurs entraînés en matière d'apprentissage automatique. Les agents responsables des cas ainsi déchargés d'un travail simple et répétitif ont désormais la possibilité de se consacrer aux cas particulièrement importants ou difficiles lors des étapes de traitement et de prendre en charge le contrôle de la qualité des résultats de l'algorithme. Bien que l'allègement des tâches soit bien accueilli, des inquiétudes ont été et sont encore exprimées. Ces préoccupations portent souvent sur la confiance que l'on peut accorder aux algorithmes, sur l'identité du responsable définitif d'un produit statistique malgré l'automatisation, et sur la manière d'éviter les distorsions. Cela a donné naissance au domaine des considérations éthiques de l'apprentissage automatique dans les statistiques officielles.

Voici une note pour réduire le champ d'application : le présent article ne porte que sur l'utilisation de l'apprentissage automatique pour produire des statistiques (c.-à-d. pas dans le cadre de l'administration générale) et, plus particulièrement, dans le cadre des travaux sur les données. Plus précisément, on n'y discute pas des considérations en matière d'éthique dans les exemples d'application de l'IA dite générative pour les résumés de texte, la création de documents, les traductions et autres.

La brève contribution au Recueil du Symposium 2024 de Statistique Canada est structurée comme suit : tout d'abord, on présente les risques en matière d'éthique que les statistiques officielles allemandes ont décelés lors de l'utilisation de l'apprentissage automatique (section 2). Puis, dans la section 3, on décrit comment les risques en matière d'éthique peuvent déjà être pris en compte en même temps que l'on respecte les exigences établies par la loi allemande et les normes de qualité européennes. À la section 4, on trouve un résumé.

2. Risques en matière d'éthique

2.1 Établir les risques en matière d'éthique

La première étape consiste à déceler les risques en matière d'éthique. L'approche choisie consistait à passer en revue les travaux nationaux et internationaux existants sur les risques en matière d'éthique que pose l'apprentissage automatique en général, mais surtout en ce qui concerne les statistiques officielles, et à établir leurs points clés. Dans ce cas-ci, on a déterminé que le *Cadre pour l'utilisation des processus d'apprentissage automatique de façon responsable* de Statistique Canada (Statistique Canada, 2021), le document *Ethical considerations in the use of Machine Learning for research and statistics* de l'organisme de la statistique du Royaume-Uni (UK Statistics Authority, 2021), et l'*International Framework on Responsible AI for Official Statistics* de la CEE-ONU (CEE-ONU, 2023) sont des sources exceptionnelles. Les objectifs que sont le fait d'éviter les biais, la transparence, l'explicabilité, la responsabilité, la confidentialité, la valeur, la prévention des préjudices, l'équité, la protection des renseignements personnels, la sécurité, la reproductibilité, la qualité, l'inférence valide, la modélisation rigoureuse, le but éthique, la durabilité environnementale, la non-discrimination et l'inclusion, la robustesse, et plusieurs autres objectifs, peuvent être trouvés ici et encore une fois dans une forme semblable. À partir de ces objectifs, il est possible d'extraire et de regrouper les risques les plus importants en matière d'éthique en utilisant le contraire des objectifs. Ainsi, on a utilisé les groupes qui forment un point d'intérêt particulier pour réaliser un examen plus approfondi.

2.2 Risques relevés en matière d'éthique

De cette manière, six risques ont été relevés :

1. Production de statistiques incorrectes
2. Création du citoyen transparent
3. Transfert de responsabilité à un algorithme et manque de transparence
4. Fuite de données délicates
5. Amplification de biais indésirables
6. Manque de durabilité écologique

Ces risques sont décrits plus amplement ci-dessous.

Production de statistiques incorrectes Cela peut se produire lorsque, en raison de l'utilisation de l'IA et de l'AA, les statistiques (p. ex. les indicateurs ou les tableaux) sont produites de manière inexacte. Cela peut causer, de manière indirecte, des erreurs d'affectation ou la prise de décisions erronées par les personnes qui interprètent les statistiques officielles ou les utilisent pour prendre des décisions (par exemple, en politique). D'un point de vue sociétal, c'est préjudiciable et devrait donc être considéré de manière négative.

Création du citoyen transparent L'utilisation de l'IA et de l'AA couple les renseignements au sujet d'une unité statistique (comme une personne naturelle) d'une manière qui fait en sorte que l'unité statistique devient « transparente », ce qui permet le profilage. Des exemples montrent qu'un petit nombre de points de données sont suffisants pour identifier de manière unique les personnes. Le grand volume de données qui peut être couplé par l'IA et l'AA au sein d'un organisme de la statistique propose une belle occasion d'élargir cette possibilité. Cet aspect est également contraire aux idées de la société. Il faut donc le considérer comme étant préjudiciable.

Transfert de responsabilité à un algorithme et manque de transparence Les processus d'IA et d'AA utilisés sont tenus responsables de décisions qui ont été prises pendant la production de statistiques. Parmi les exemples, mentionnons l'affectation d'unités statistiques à certaines catégories (comme les catégories professionnelles ou les activités économiques). Personne n'est responsable d'une affectation. Il n'est pas possible de remettre en question la décision prise automatiquement par l'IA et l'AA. Cela s'explique notamment par le fait qu'il est impossible de retracer les étapes essentielles ou qu'elles n'ont pas été documentées. Dans un état de droit (primauté du droit) (« Rechtsstaat »), on s'attend à ce que les mesures administratives puissent être examinées. Si ce n'est pas le cas, il faut considérer qu'il s'agit d'un aspect négatif.

Fuite de données délicates Il est manifestement préjudiciable et incompatible avec la perception des statistiques officielles qu'a le public que des données qui doivent être protégées (en raison de la loi sur la statistique, de la protection des données, des obligations en matière de confidentialité des employés ou de toute autre norme) sortent de la zone protégée des statistiques officielles. Si les données, qui ont été fournies avec la certitude qu'elles seront utilisées exclusivement à des fins statistiques, sont mises à la disposition de personnes non autorisées, il y aura une perte de confiance majeure et une grave atteinte à la réputation des statistiques officielles (en plus des conséquences juridiques qui pourraient en découler).

Amplification de biais indésirables En principe, il est concevable que l'utilisation de l'IA et de l'AA amplifie les biais indésirables (p. ex. dans les données d'entraînement des algorithmes). Par exemple, l'algorithme classe les catégories professionnelles qui ne sont pas suffisamment représentées dans les données d'entraînement comme étant négligeables. Ces catégories ne sont plus prises en considération dans les classifications subséquentes réalisées par l'algorithme. Cette situation peut être particulièrement critique en cas de classification croisée (p. ex. sexe, catégorie professionnelle et région). Ainsi, les statistiques officielles ne décrivent plus correctement les phénomènes de masse, ce qui pourrait augmenter le risque de produire des statistiques inexactes et d'observer les conséquences correspondantes (p. ex. au chapitre de la politique ou de l'industrie, mais également dans le cadre des décisions privées). Il faut l'évaluer de manière négative du point de vue social et personnel.

Manque de durabilité écologique L'utilisation de l'IA et de l'AA, en particulier de ce que l'on appelle l'apprentissage profond, et des grands modèles de langage appartenant à cette catégorie, est souvent exigeante en termes de calcul et entraîne une consommation d'énergie et d'eau extrêmement élevée, en particulier lors de l'entraînement initial et de tout nouvel entraînement éventuellement nécessaire, mais aussi, bien que souvent dans une moindre mesure, au cours de l'application. Cet aspect est particulièrement négatif si l'utilisation de cette catégorie de méthodes, qui peut ne pas être durable sur le plan écologique, n'est pas nécessaire (p. ex. lorsque des méthodes plus durables offrent des résultats quasiment aussi bons).

3. Lutter contre les risques

3.1 Cadres de référence

Différents cadres de référence sont envisageables pour tenir compte des risques en matière d'éthique mentionnés ci-dessus, c'est-à-dire pour répondre à la question à savoir comment traiter les risques en matière d'éthique. De quels

aspects les organismes des statistiques doivent-ils tenir compte lorsqu'ils produisent des statistiques officielles à l'aide de l'IA et de l'AA?

Deux cadres de référence s'appliquent aux statistiques officielles. Tout d'abord, et avec une primauté par rapport à tous les autres cadres de référence, il y a la loi avec ses normes et ses décisions judiciaires contraignantes qui sont pertinentes pour les statistiques officielles. Ensuite, il y a les engagements volontaires (autorisés par la loi) pris par les autorités chargées des statistiques officielles elles-mêmes. On peut les trouver dans les cadres de la qualité des statistiques officielles en Allemagne et en Europe (tout particulièrement dans le Cadre d'assurance de la qualité, Système statistique européen, 2019). De nombreux autres pays disposent de cadres comparables de la qualité, comme le Cadre d'assurance de la qualité de Statistique Canada (Statistique Canada, 2017). Il existe également un cadre correspondant au niveau des Nations Unies (ONU, 2019).

Loi Les normes fondamentales du cadre juridique de référence pour l'Allemagne sont la Loi fondamentale de la République fédérale d'Allemagne, la loi fédérale sur la statistique, les lois sur les statistiques spécialisées (certaines de ces lois renferment même des précisions très détaillées quant aux variables qui peuvent être recueillies au moyen de statistiques officielles) et l'arrêt de la Cour constitutionnelle fédérale sur le recensement de 1983. Il existe un principe selon lequel (avec très peu d'exceptions) il est interdit de recueillir des statistiques officielles sans loi. Cela permet de s'assurer que le parlement évalue les aspects éthiques dans le cadre du processus législatif. Le Bundestag allemand représente les citoyens et sert d'organe législatif décisif. C'est au sein de celui-ci que la société décide de ce qu'elle souhaite autoriser ou non. Si une loi de niveau supérieur (en particulier la Loi fondamentale) n'est pas respectée ou est délibérément ignorée, il est possible de saisir la Cour constitutionnelle fédérale de la question.

D'autres normes renferment des règlements concernant la sécurité de l'information, la protection des données ou l'efficacité énergétique.

Qualité Les différents cadres de qualité (y compris le Cadre d'assurance de la qualité, le code de bonnes pratiques de la statistique européenne et le Qualitätshandbuch, qui est l'adaptation allemande des anciens cadres) comprennent des exigences pour l'organisme (l'organisme de statistique), ses processus et ses produits. Dans ce cas-ci, les principes comprennent, par exemple, l'indépendance professionnelle, une méthodologie rigoureuse, l'exactitude et la fiabilité, l'actualité et la ponctualité, ou l'accessibilité et la clarté (Système statistique européen, 2019). Un extrait des dimensions de la qualité pour l'IA et l'AA dans les statistiques officielles tiré des cadres généraux de qualité, ainsi qu'une proposition pour les opérationnaliser pour qu'elles prennent la forme de lignes directrices en matière de qualité, se trouve dans Saidani et coll. (2023), un document qui est lui-même fondé sur un document de Yung et coll. (2022), et donc sur des travaux préliminaires conjoints dans le cadre d'un projet de la CEE-ONU. Les dimensions de la qualité tirées de Saidani et coll. (2023), c'est-à-dire l'exactitude, la robustesse, l'explicabilité, la reproductibilité, l'actualité et la ponctualité, et l'efficacité par rapport au coût, ont été depuis classées parmi les principes contraignants des statistiques officielles en Allemagne. En outre, les lignes directrices en matière de qualité ont été définies, sous une forme légèrement modifiée, comme base contraignante pour les projets d'IA et d'AA en ce qui concerne les statistiques officielles allemandes. Ces lignes directrices en matière de qualité visent, par exemple, l'exactitude, « *les méthodes d'apprentissage automatique suivent une méthodologie scientifique normalisée en ce qui concerne les statistiques et l'AA. La méthodologie utilisée a été documentée (y compris les raisons expliquant pourquoi une méthode précise a été retenue) et communiquée au public cible* », et quant à la robustesse, « *l'objet désiré des analyses de la robustesse a été défini : prédictions précises, coefficients du modèle, mesures de l'exactitude ou valeurs agrégées pour la variable cible* ».

La loi représente, forcément, le principal cadre de référence à mettre en application. Cette primauté de la loi est associée, d'une part, au deuxième cadre de référence (qualité) et s'appliquerait également à un cadre éthique hypothétique. Le fait qu'un cadre éthique établi par une autorité, comme un organisme des statistiques, ne peut être supérieur aux lois établies par un parlement, découle de la Loi fondamentale. En outre, un tel cadre éthique devrait aller de pair avec les valeurs de la Loi fondamentale, en particulier les droits fondamentaux, y compris la Charte des droits fondamentaux de l'UE, et les objectifs de l'État.

3.2 Aperçu détaillé des risques

Production de statistiques inexactes Les principes selon lesquels la méthodologie et les procédures utilisées doivent correspondre à l'état actuel de la recherche scientifique et que les résultats d'une statistique doivent être exacts (principes 7, 8 et 12 du Cadre d'assurance de la qualité) supposent l'obligation d'assurer et de vérifier cette exactitude. Bien entendu, de par la nature des statistiques et de toutes les méthodes statistiques (y compris l'IA et l'AA), il arrive souvent que les estimations et les prédictions ne correspondent pas exactement à la « vraie » valeur lorsqu'elles sont appliquées à un cas particulier. C'est le prix à payer, car tous les renseignements nécessaires ne sont pas recueillis directement sans erreurs de mesure et sans difficultés au chapitre de la délimitation et de l'interprétation. Lorsqu'il s'agit de l'application précise de l'IA et de l'AA (et cela s'applique également à l'utilisation d'autres méthodes), l'affectation, par exemple, d'une personne particulière à une classification d'emploi, peut être inexacte. Toutefois, tant que l'on s'assure que les statistiques elles-mêmes correspondent aux exigences en matière d'exactitude, ce point n'est pas critique. Le cadre de référence sur la qualité tient donc compte du risque en matière d'éthique que représentent les statistiques inexactes.

Création du citoyen transparent Le risque en matière d'éthique lié à la création d'un citoyen transparent ne peut pas être présent dans les statistiques officielles allemandes, car il est déjà exclu par le cadre de référence juridique et, ici, en particulier par la décision sur le recensement de 1983. Cette décision indique, de manière explicite, qu'il n'est pas permis de combiner les statistiques et les données personnelles d'une manière qui rend le profilage possible. En outre, en vertu de la loi fédérale sur la statistique, les organismes des statistiques sont tenus de supprimer les caractéristiques auxiliaires (comme le nom et l'adresse) dès que possible. En outre, il est interdit de tenter de procéder à une réidentification. Les données à partir desquelles les statistiques peuvent être fusionnées sont également régies, de manière explicite, par la loi. Dans ce contexte, la question à savoir si l'IA et l'AA peuvent être utilisés n'est pas pertinente. Le cadre de référence juridique tient donc compte de ce risque en matière d'éthique.

Transfert de responsabilité à un algorithme et manque de transparence D'un point de vue juridique, l'algorithme (ou l'IA) n'agit jamais. C'est l'organisme des statistiques qui agit, qui ne fait qu'utiliser l'algorithme. C'est à ce niveau que les responsabilités sont définies. En outre, les lois sur la protection des données donnent aux personnes le droit de ne pas devenir le sujet du processus décisionnel automatisé. En principe, cela peut, évidemment, s'appliquer également aux processus relatifs aux statistiques officielles. En outre, l'utilisation d'algorithmes ne doit pas mener à une situation dans laquelle les étapes essentielles de l'activité ne sont pas compréhensibles ou documentées (non-respect des obligations en matière de transparence et de responsabilité), car cela pourrait signifier que l'activité officielle ne peut plus être contrôlée par les tribunaux (non-respect de la garantie de recours juridique de la Loi fondamentale) et que les instances de contrôle (en particulier les autorités de surveillance, les parlements) ne peuvent plus contrôler les mesures administratives (non-respect du caractère licite des mesures administratives dans la Loi fondamentale). En outre, selon le Cadre d'assurance de la qualité, la direction (c'est-à-dire les personnes) des producteurs de statistiques officielles est seule responsable quand vient le temps de déterminer les méthodes, normes et procédures statistiques (indicateur 1.3). En outre, les indicateurs 6.4 (accessibilité publique des méthodes et des procédures utilisées) et 15.6 (informer les utilisateurs au sujet de la méthodologie des processus statistiques) du Cadre d'assurance de la qualité, et la dimension « explicabilité » de la qualité dans Saidani et coll. (2023, p. 279 et suivantes) fournissent des lignes directrices en ce qui concerne la transparence du point de vue de la qualité. Le cadre de référence sur la qualité et le cadre de référence juridique tiennent tous les deux compte du risque en matière d'éthique associé au transfert de responsabilité à un algorithme et au manque de transparence.

Fuite de données délicates La loi fédérale sur la statistique et le Règlement relatif aux statistiques exigent déjà des organismes des statistiques qu'ils prennent toutes les mesures juridiques, administratives, techniques et organisationnelles requises pour garantir la protection matérielle et logique des données confidentielles. Il existe également des exigences juridiques spéciales en ce qui concerne le traitement des données personnelles. Les normes juridiques pertinentes interdisent donc les mesures préjudiciables. Les mesures qui ne sont pas interdites par les normes juridiques sont acceptables du point de vue du législateur. En outre, le principe 5 du Cadre d'assurance de la qualité du SSE exige de garantir la confidentialité des statistiques.

Amplification des biais indésirables Dans Saidani et coll. (2023), on résume l'équité comme une circonstance dans laquelle les agrégats statistiques pour certains sous-groupes ne sont pas systématiquement surestimés ou sous-estimés (Saidani et coll., 2023, p. 292 et suivantes). Les statistiques officielles doivent tenir compte des conditions actuelles.

Cependant, la manière dont on réagit à ces conditions est une décision politique, économique ou personnelle, qui ne relève pas des statistiques officielles. Pour obtenir une représentation correcte et non biaisée, il est essentiel de veiller à ce que tous les sous-groupes pertinents soient représentés correctement par les méthodes statistiques dans le cadre du processus de production. Les statistiques officielles doivent l'assurer par l'intermédiaire des exigences en matière de qualité. Le cadre de référence sur la qualité tient donc déjà compte du risque associé au fait de renforcer les biais indésirables.

Manque de durabilité écologique En fonction des connaissances de Dumpert et coll. (2025a, 2025b), les cadres de référence (loi et qualité) mentionnés ci-dessus ne tiennent pas entièrement compte de l'aspect de la durabilité (écologique). Les tentatives visant à établir des lignes directrices correspondantes pour la durabilité écologique restent dans un premier temps indirectes, vagues ou non contraignantes. Le fait que ces lignes directrices ne soient pas aussi précises que pour les autres risques en matière d'éthique signifie qu'il faut accorder une attention particulière à ce risque à l'avenir.

4. Résumé et perspectives

La mise en œuvre des exigences des cadres de référence existants qui, dans la plupart des cas, ne sont pas propres à l'IA ou à l'AA, comme la loi et la qualité, est déjà suffisante pour tenir compte de manière adéquate des enjeux éthiques se fondant sur des scénarios de risque. Il n'est donc pas nécessaire de créer un cadre éthique spécial pour les statistiques officielles allemandes. Tant que les cadres de référence juridiques et de qualité existent et sont mis rigoureusement en œuvre, il n'est pas nécessaire, à l'exception du cas de la durabilité écologique, d'agir d'un point de vue éthique lors de l'utilisation de l'IA ou de l'AA pour produire des statistiques. Néanmoins, des changements concevables aux cadres de référence ou à l'élaboration de l'IA et de l'AA exigent de garder constamment à l'esprit les circonstances, et que l'on vérifie régulièrement si les cadres de référence continuent de répondre à toutes les questions d'ordre éthique en ce qui concerne les statistiques officielles.

Bibliographie

- Dumpert, F., Reichel, J., Oertel, E., Leerhoff, H., et Salwiczek, C. (2025a), "Ethische Fragen beim Einsatz von KI/ML in der Produktion amtlicher Statistiken – Teil 1: Identifikation", *WISTA – Wirtschaft und Statistik*, 77(1), p. 15–24.
- Dumpert, F., Reichel, J., Oertel, E., Leerhoff, H., et Salwiczek, C. (2025b), "Ethische Fragen beim Einsatz von KI/ML in der Produktion amtlicher Statistiken – Teil 2: Auseinandersetzung", *WISTA – Wirtschaft und Statistik*, 77(1), p. 25–36.
- European Statistical System (2019), "Quality Assurance Framework of the European Statistical System (Version 2.0)", <https://ec.europa.eu/eurostat/web/quality/european-quality-standards/quality-assurance-framework>.
- Holt, D. (2007), "The Official Statistics Olympic Challenge: Wider, Deeper, Quicker, Better, Cheaper", *The American Statistician*, 61(1), p. 1–8.
- Saidani, Y., Dumpert, F., Borgs, C., Brand, A., Nickl, A., Rittmann, A., Rohde, J., Salwiczek, C., Storfinger, N., et Straub, S. (2023), "Qualitätsdimensionen maschinellen Lernens in der amtlichen Statistik", *AStA Wirtschafts- und Sozialstatistisches Archiv*, 17 (3+4), p. 253–303.
- Statistique Canada (2017), "Le cadre d'assurance de la qualité de Statistique Canada", 3^{ième} édition, accessible en ligne à l'adresse <https://www150.statcan.gc.ca/n1/fr/catalogue/12-586-X>.
- Statistique Canada (2021), "Cadre pour l'utilisation des processus d'apprentissage automatique de façon responsable à Statistique Canada", <https://www150.statcan.gc.ca/n1/fr/catalogue/892000062021001>.

UK Statistics Authority (2021), “Ethical considerations in the use of Machine Learning for research and statistics“, <https://uksa.statisticsauthority.gov.uk/publication/ethical-considerations-in-the-use-of-machine-learning-for-research-and-statistics/pages/1/>.

UNECE (2019), “Generic Statistical Business Process Model (GSBPM)”, https://unece.org/sites/default/files/2023-11/GSBPM%20v5_1.pdf.

UNECE (2023), “International Framework on Responsible AI for Official Statistics”, publication à venir.

United Nations (2019), “United Nations National Quality Assurance Frameworks Manual for Official Statistics”, <https://unstats.un.org/unsd/methodology/dataquality/references/1902216-UNNQAManual-WEB.pdf>.

Yung, W., Tam, S.-M., Buelens, B., Chipman, H., Dumpert, F., Ascari, G., Rocci, F., Burger, J., et Choi, I. (2022), “A quality framework for statistical algorithms”, *Statistical Journal of the IAOS*, 38(1), p. 291–308.