

## La Déclaration de Vancouver relative aux Collections-Données

Padilla, Thomas, Scates Kettler, Hannah, Varner, Stewart, & Shorish, Yasmineen. (2023). La Déclaration de Vancouver relative aux Collections-Données. <https://doi.org/10.5281/zenodo.8341863>.

Dernière mise à jour : 2023

Précédée par : la Déclaration de Santa Barbara relative aux Collections-Données<sup>1</sup>, 2017

Depuis la publication de la Déclaration de Santa Barbara relative aux collections-données<sup>2</sup> (2017), celles-ci ont fait l'objet d'un engagement international croissant. Des institutions, de grande et petite taille, ont consenti à des investissements individuels et collectifs afin de développer et de soutenir l'utilisation informatique responsable des collections en tant que données et de favoriser l'accès à ces dernières.

La Déclaration de Vancouver propose un ensemble de principes visant à examiner en détail les questions que soulève le travail relatif aux collections-données. Ces principes s'inscrivent dans le cadre d'un effort mondial, interprofessionnel et interdisciplinaire ayant pour objectif d'autonomiser les intendants de la mémoire, de la connaissance et des données<sup>3</sup> (notamment les praticiens et les chercheurs) qui aspirent à favoriser le développement et l'utilisation informatique responsables des collections en tant que données. Ce rôle d'intendance ne cesse de gagner en importance étant donné que les applications d'intelligence artificielle, qui disposent d'une énorme quantité de données, y compris les collections en tant que données, sont de plus en plus omniprésentes dans nos vies.

Le concept de collections-données est apparu à un moment précis de l'histoire des galeries, des bibliothèques, des archives et des musées. Pendant des décennies, les praticiens et les chercheurs ont œuvré au développement de collections numériques. Simultanément, un éventail hétéroclite de groupes d'utilisateurs (notamment des chercheurs, des artistes, des décideurs politiques, des entrepreneurs, des journalistes et des praticiens GLAM) ont utilisé les outils informatiques pour dériver des informations à partir de collections à grande échelle. Ce travail s'appuie sur diverses méthodes telles que la fouille de textes, la vision par ordinateur, l'apprentissage automatique, l'intelligence artificielle, la visualisation de données, la représentation graphique, l'analyse d'images, l'analyse audio et l'analyse de réseau ; cette liste n'étant toutefois pas exhaustive. Si les efforts visant à favoriser un développement et une utilisation informatique responsables des collections ont pris une dimension internationale, il reste encore beaucoup de travail à fournir pour initier, favoriser et socialiser des approches qui cadrent avec

---

<sup>1</sup> <https://zenodo.org/record/3066209>

<sup>2</sup> Dans ce document, le terme composé « collections-données » est utilisé pour faire référence au concept, à la communauté et au domaine d'exercice et domaine de pratique lié au soutien d'une utilisation informatique responsable des collections en tant que données.

<sup>3</sup> steward: la personne qui facilite et préconise l'accès et la prise en charge à long terme des collections.

un éventail de ressources locales, régionales, nationales et internationales, ainsi que les priorités stratégiques.

Malgré la diversité des visions concernant les modalités de développement, de mise à disposition et d'utilisation des collections en tant que données, n'importe quel matériel numérique sera potentiellement accessible sous forme de données compatibles avec une utilisation informatique. On peut encourager l'utilisation et la réutilisation des données en travaillant avec un système de licence adéquat et des formats non exclusifs, particulièrement dans le cas où les données sont mises à disposition par le biais de canaux très divers et où l'accès à ces données est conçu pour répondre à des besoins communautaires, avec une attention pour les préoccupations éthiques.

La compréhension et la prise en compte des préoccupations éthiques font partie intégrante des collections-données. Les décisions en matière d'accessibilité des collections-données doivent être sous-tendues par un éventail de priorités communautaires. Les arguments avancés par la Global Indigenous Data Alliance's CARE Principles<sup>4</sup> et le travail sur la Maori Data Sovereignty<sup>5</sup> sont très instructifs. Ceux-ci préconisent le respect des exigences légales et des besoins en matière d'accessibilité, le respect de différentes normes et valeurs culturelles, la prise en considération des risques encourus par certains groupes vulnérables, ainsi que le respect des droits et de l'autorité de ces communautés.

L'ampleur de certaines collections peut masquer l'exhaustivité – ou le manque d'exhaustivité – des histoires qu'elles sont censées représenter. Les praticiens du souvenir, de la connaissance et des données, mais aussi les chercheurs, doivent être attentifs à ces absences ou à ces déformations et œuvrer pour éviter leur récurrence. La documentation doit être éclairée par des principes archivistiques et des pratiques de reproductibilité efficaces afin de veiller à ce que les utilisateurs disposent des informations dont ils ont besoin pour travailler de manière responsable avec les collections.

La création de la Déclaration de Vancouver réside dans la volonté d'interpeller les personnes nouvellement acquises au soutien du développement et de l'utilisation informatique responsables des collections en tant que données, mais également les personnes rompues aux approches très anciennes de la pratique des galeries, des bibliothèques, des archives et des musées.

Les répercussions des principes énoncés ci-dessous ne doivent pas être nécessairement prises en compte toutes à la fois. En effet, certaines peuvent demeurer des aspirations pendant un certain temps étant donné que l'on commence à travailler de concert avec des collègues dont les expériences, les valeurs et les approches du travail s'avèrent différentes. L'essentiel est d'avoir un point de départ, d'évaluer, d'affiner et d'améliorer.

---

<sup>4</sup> Carroll, S.R., Garba, I., Figueroa-Rodríguez, O.L., Holbrook, J., Lovett, R., Materechera, S., Parsons, M., Raseroka, K., Rodriguez-Lonebear, D., Rowe, R., Sara, R., Walker, J.D., Anderson, J. and Hudson, M., 2020. The CARE Principles for Indigenous Data Governance. *Data Science Journal*, 19(1), p.43. DOI: <https://doi.org/10.5334/dsj-2020-043>

<sup>5</sup> <https://www.temanararaunga.maori.nz/>

1. **Le développement des collections-données contribue à l'utilisation informatique responsable des collections numérisées et de celles qui ont été créées numériquement.** Les intendants de la connaissance, des données et du souvenir développent l'ensemble des opportunités d'interaction avec les collections en concevant et en créant des collections en tant que données et en donnant accès à celles-ci.
2. **Les intendants de collections-données s'engagent à combattre les inégalités historiques et contemporaines symbolisées dans l'acquisition, la portée, la description, l'accès et l'utilisation.** Les engagements visant à combattre les inégalités doivent être documentés et accessibles au public, sauf si la publication de tels engagements compromet la sécurité de certaines communautés. Les silences et/ou les distorsions connus qui entachent les collections sont, dans la mesure du possible, repris dans cette documentation. Si les détails de ces engagements varient en fonction des communautés qui ont accès aux collections, ils servent néanmoins une cause commune : la tentative de répondre aux besoins des plus vulnérables. Les intendants peuvent imposer des restrictions mûrement réfléchies à l'utilisation des données – y compris l'éventuelle décision de ne pas donner accès à telle ou telle collection en tant que donnée. Les intendants respectent les droits et besoins des communautés à l'origine des contenus qui constituent des collections, de celles qui sont représentées dans les collections, ainsi que des communautés qui les utilisent.
3. **Les collections en tant que données doivent être largement accessibles, dans les limites des attentes éthiques, légales et communautaires.** Les conditions d'utilisation des collections en tant que données doivent être explicites et éclairées par des pratiques communautaires telles que le RightsStatements.org et les principes CARE, et par des licences standards telles que les licences Creative Commons<sup>6</sup>, Open Data Commons<sup>7</sup>, and Traditional Knowledge<sup>8</sup>. Les matériaux doivent être compatibles avec les outils et logiciels d'accessibilité et disponibles dans plusieurs langues ou dans un format qui se prête à une traduction automatique.
4. **Les collections-données tirent profit de la conception participative.** Au lieu d'émettre des suppositions à propos des besoins ou d'imaginer les communautés d'utilisateurs, les intendants doivent collaborer avec d'autres parties et (le cas échéant), s'efforcer de réduire les obstacles à l'utilisation. Ils doivent également prendre les principes de conception universels en considération et continuer d'évaluer les besoins de l'utilisateur au fil du temps. Lorsque les ressources le permettent, ils favorisent la multiplicité des approches en matière de développement et de mise à disposition des données afin de faciliter l'interopérabilité (tel que décrit au Principe 6). Il convient de motiver et d'encourager les personnes à construire et à

---

<sup>6</sup> <https://creativecommons.org/about/cclicenses/>

<sup>7</sup> <https://opendatacommons.org/>

<sup>8</sup> <https://localcontexts.org/licenses/traditional-knowledge-licenses/>

partager l'aboutissement de ce travail.

5. **La documentation partagée fournit un contexte et aide d'autres parties à trouver une voie vers la réalisation du travail.** Le travail sur les collections-données nécessite des prises de décisions concernant le développement, l'accès et l'utilisation. La documentation doit expliquer clairement la manière dont telle ou telle décision a été prise dans le cadre du travail sur les collections-données, ainsi que le bien-fondé de cette décision. Bien qu'aucune documentation ne soit totalement complète, une documentation incomplète ou en cours de réalisation est préférable à une absence de documentation. La planification stratégique, les descriptions de fonctions, les schémas de métadonnées lisibles par l'homme et par machine, les fiches techniques, les flux de travail, les profils d'application, les actes de donation et les livres-codes – accompagnés de traductions dans plusieurs langues –, constituent notamment des exemples de documentation. Le choix de la langue doit refléter les besoins de l'utilisateur. La documentation doit être accessible au public et résister à l'épreuve du temps, compte tenu de considérations éthiques et légales.
6. **Le développement des collections-données valorise l'interopérabilité sociale et technique.** L'interopérabilité entraîne un alignement sur des modèles (gestion de personnel, équipes matricées, maintenance et stratégies en matière d'innovation), normes et infrastructures communautaires émergents et/ou établis et facilite l'intégration, que ce soit avec une architecture centralisée ou une architecture distribuée. Il faut être attentif aux distorsions présentes dans les normes et infrastructures. L'interopérabilité s'étend au-delà de l'aspect technique pour devenir une interopérabilité sociale – la faculté qu'ont les personnes d'interagir efficacement et de soutenir une infrastructure humaine durable.
7. **Les intendants de collections-données travaillent en toute transparence afin de maintenir l'intégrité des collections en vue de leur accessibilité à long terme.** L'intégrité des collections-données dépend des efforts déployés afin de garantir et de documenter publiquement l'intégrité technique des données, leur origine et les relations qui unissent les intendants de collections et les communautés et/ou les groupes d'utilisateurs. Pour ce faire, les intendants sont tenus de reconnaître les lacunes, les distorsions et les zones d'incertitude. Pour être transparentes, les collections-données doivent comporter des métadonnées ouvertes et robustes (conformément au Principe 3) et être assorties d'un engagement de préservation.
8. **La progression du travail sur les collections-données dépend des engagements organisationnels en termes d'infrastructures et de processus durables.** Le développement continu des collections données peut impacter les modèles de gestion de personnel, les flux de travail et les infrastructures techniques. Les entités et les organisations doivent favoriser le développement de collaborations et d'intégrations productives entre les flux de travail, les postes et les modèles de services existants et nouveaux.

9. **Le développement des collections en tant que données doit établir un équilibre entre les avantages et les préoccupations en matière de climat.** Le travail numérique repose, par nature, sur des centres de données énergivores et utilise notamment des métaux lourds, nocifs pour l'environnement. Il faut concevoir, dans la mesure du possible, des processus exploitant des ressources computationnelles qui font preuve de transparence concernant les manières dont elles impactent, négativement ou non, l'environnement.
10. **Le développement des collections-données doit établir un équilibre entre les avantages et les préoccupations en matière d'exploitation par le travail.** La création de collections en tant que données est ou a éventuellement été facilitée par l'exploitation d'êtres humains par le travail. Les créateurs doivent reconnaître (si possible) les contextes de travail en produisant une documentation appropriée et en utilisant des taxonomies telles que CRediT<sup>9</sup>.
11. **En créant l'accès aux données, les intendants de collections-données doivent accorder une attention adéquate à la consommation des données par l'intelligence artificielle ou d'autres technologies.** Les collections en tant que données peuvent être utilisées dans des ensembles de formation destinés à l'intelligence artificielle et au développement de produits. Les intendants doivent être attentifs aux implications éthiques, y compris la propriété intellectuelle et les revendications en matière de souveraineté des données, et développer des dispositifs afin d'éviter une utilisation abusive, une perte de contexte préjudiciable et une amplification des distorsions par le biais de ces technologies.

---

<sup>9</sup> <https://credit.niso.org/>