Une proposition de la société civile pour une IA agentique souveraine et fédérée en Aotearoa (Nouvelle-Zélande)

John G. Stroh / My Digital Sovereignty Ltd

Proposition de la société civile · v1.2 Projet de mai 2026 | Parallèle constructif avecles Lignes directrices de mise en œuvre du CAC 2026 →

v1 (remplacée) → Commentaires par e-mail

v1.2, 16 mai 2026 : révisée conformément à la correspondance de Ted Howard concernant la v1.1, en mettant l'accent sur la distinction entre la couche de substrat et la couche d'exécution. La section sur les primitives §0(i) ajoute un paragraphe de clarification précisant que la couche de substrat (PKI, fédération, portabilité) est architecturalement distincte des primitives de la couche d'exécution que constituent les six services §0(i) ; la section sur l'application des limites §0(i) ajoute le cadre de fallibilité en quatre catégories et la sortie trinaire du routeur (autoriser / refuser / escalader) ; L'annexe A, obj-3 et obj-5, approfondit la séparation entre substrat et runtime ainsi que l'indépendance de la posture de survie par rapport au confinement des agents. Les détails architecturaux de fond sont développés dans *l'alignement architectural* aux §3.3, §3.4, §3.5, §7.4, §7.5 et dans *le document A* au §5.3. La v1.1 reste accessible à son URL à titre de référence historique.

v1.1, 14/05/2026 : révisée le jour même suite aux commentaires du Dr Karaitiana Taiuru sur la v1. Le §0(iii) cite désormais l'analyse critique de Te Mana Raraunga réalisée par Taiuru le 20 septembre 2025 ; cite Te Kāhui Raraunga comme l'organe opérationnel actuellement reconnu ; ajoute une analyse explicite des lacunes. La v1 reste accessible à l'URL de la v1 à titre de référence historique. Les commentaires portant sur des sections spécifiques sont les bienvenus. Veuillez citer les numéros de section (par exemple, §III, point 5). L'auteur répond personnellement ; veuillez prévoir un délai d'une à deux semaines.

Une proposition de la société civile pour une IA souveraine et fédérée à agentique en Aotearoa Nouvelle-Zélande

v1.2 mai 2026 — projet de document de recherche (révisé conformément à la correspondance de Ted Howard sur la v1.1 ; distinction entre substrat et runtime, faillibilité en quatre catégories, sortie du routeur trinaire). Parallèle constructif aux Lignes directrices de mise en œuvre de 2026 de la République populaire de Chine sur les agents intelligents.

John G. Stroh / My Digital Sovereignty Ltd

16/05/2026

Une proposition de la société civile pour une IA agentique souveraine et fédérée en Aotearoa Nouvelle-Zélande

v1.2 mai 2026 — projet de document de recherche (révisé conformément à la correspondance de Ted Howard sur la v1.1 ; distinction substrat-vs-runtime, fallibilité à quatre catégories, sortie du routeur trinaire ; voir À propos de ce document)

Une proposition de la société civile émanant de My Digital Sovereignty Ltd, présentée aux décideurs politiques, aux organisateurs communautaires et aux professionnels du secteur en Nouvelle-Zélande. Conçue comme un parallèle constructif aux Lignes directrices de mise en œuvre de 2026 de la République populaire de Chine pour l' application normalisée et le développement innovant des agents intelligents, hébergée en traduction anglaise à l'adresse /research/translations/.

À propos de ce document

Il s'agit de la **version 1.2 de mai 2026** d'un projet de proposition de la société civile émanant de My Digital Sovereignty Ltd, soumis aux décideurs politiques, aux organisateurs communautaires et aux professionnels du secteur néo-zélandais. Les commentaires sont les bienvenus via les canaux permanents de commentaires sur agenticgovernance.digital ; les

révisions apportées en réponse aux commentaires seront publiées sous la forme de la version 2.

Journal des modifications v1 → v1.1 (14/05/2026) : la v1 a été publiée plus tôt le 14/05/2026 et immédiatement examinée par le Dr Karaitiana Taiuru, qui a signalé que la référence fondamentale de la v1 aux principes de souveraineté des données maories de 2016-2018 était dépassée dans le contexte de l'IA (conformément à son *analyse critique des principes de données de Te Mana Raraunga* datée du 20 septembre 2025). La v1.1 modifie le §0(iii) pour citer directement l' analyse critique de Taiuru ; pour citer **Te Kāhui Raraunga** (l'organe opérationnel actuellement reconnu pour la gouvernance des données maories en Aotearoa (Nouvelle-Zélande), créé en 2019) ainsi que son modèle de gouvernance des données maories et son cadre de gouvernance de l'IA maorie publiés comme références actuelles ; d' adopter les termes de fond privilégiés par Taiuru (mana motuhake, rangatiratanga) lorsque cela est approprié ; et d'ajouter une sous-section explicite d'analyse des lacunes précisant ce que cette proposition fait et ne fait pas encore dans la dimension te ao Māori . Les points 4, 23, 37, le principe 2 de la section I et la section VI comportent la même mise à jour des références. L'architecture spécifiée par cette proposition reste inchangée par rapport à la v1. La v1 reste accessible à l'adresse `/papers/aotearoa-nz-agentic-ai-framework-v1-may-2026.html` à titre de référence historique ; cette URL renvoie à la v1.2.

Journal des modifications v1.1 → v1.2 (16/05/2026) : la v1.1 a été examinée par Ted Howard dans le cadre d'un échange de correspondance ; il a souligné que la formulation de la v1.1 présentant les six primitives §0(i) comme des mécanismes de sécurité architecturaux est tout à fait défendable lorsque la distinction entre substrat et runtime est explicitée. La v1.2 ajoute un paragraphe de clarification après la liste des primitives du cadre §0(i), nommant la couche de substrat (PKI / enveloppes de fédération / enregistrements portables, développés dans *document A* et *Alignement architectural* §3.4) comme l'équivalent architectural de la couche d'exécution constituée par les six services §0(i). L'application des limites §0(i) bénéficie d'un cadre de fallibilité en quatre catégories (ces catégories sont négociées par la communauté et susceptibles d'appel, ce ne sont pas des entités fixes) et de la note de sortie du routeur trinaire (autoriser / refuser / escalader vers un humain). L'annexe A obj-3 (contournement) et obj-5 (exploitation de service d'exécution) bénéficient de renvois vers *l'Alignement architectural* §3.4 substrat-vs-exécution, §7.4 posture de survie indépendante du confinement de l'agent, §7.5 surface d'attaque de la couche sociale (frontière ouverte), et le document A §5.3 horizon de migration PQC. L'architecture spécifiée dans cette proposition reste inchangée par rapport à la v1.1 ; les révisions clarifient les cadres qui ont émergé lors des échanges avec les relecteurs. La v1.1 reste accessible à l'adresse `/papers/aotearoa-nz-agentic-ai-framework-v1.1-may-2026.html` à titre de référence historique.

Version 1.2 : mise en œuvre le jour même (soirée du 16 mai 2026) : §III(II) Point 10 (mécanisme de rayon d'impact), §III(III) Points 11–12 (fondement de la gouvernance polycentrique), §III(IV) Point 13 (mécanismes de fédération), §I Principe 4 (preuve d'adoption facilitée par l'architecture), et §IV (Développement axé sur l'adoption) — les dix-neuf points sectoriels — dotés d'un fondement de capacités primitives afin que les acteurs politiques puissent défendre les revendications de souveraineté substantielles au sein du comité. Registre générique de l'architecture : nomme **la provenance cryptographique**, **les enveloppes de fédération**, **la portabilité pilotée par les membres** et **l'application des limites** par capacité plutôt que par implémentation MDSL. Le §IV gagne un paragraphe d'introduction nommant une seule fois les quatre primitives de substrat ; chaque point sectoriel opérationnalise ensuite uniquement la manifestation spécifique au secteur. Contenu inchangé par rapport à la version 1.2 du matin ; la mise à jour rend les revendications de souveraineté opérationnellement lisibles pour les lecteurs qui les défendraient en commission.

Révision de clarté de la v1.1 le jour même (14 mai 2026, soir) : Paragraphe

d'introduction de la section §0(i) révisé pour commencer explicitement par la distinction entre le niveau système et le niveau modèle (primitives au niveau système, vérifications d'exécution au niveau du code, indépendance vis-à-vis du substrat pour les architectures de type transformateur-LLM / JEPA / hybrides). Formulation du paragraphe BoundaryEnforcer « par l'architecture plutôt que par l'espoir » → « par l'interception d'exécution plutôt que par l'espoir » afin de réduire l'ambiguïté pour les lecteurs ingénieurs familiarisés avec le débat sur l'alignement des LLM. Le fond reste inchangé ; il s'agit d'une révision de la formulation pour des raisons d'accessibilité. Déclenchée par un lecteur technique qui a fait correspondre les primitives du §0(i) à des revendications d'alignement au niveau du modèle qu'elles ne formulent pas.

Révision de clarté de niveau 2 de la v1.1 le jour même (15/05/2026) : chacune des six primitives du §0(i) a reçu une analogie technique concrète (appels système privilégiés du noyau du système d'exploitation / disjoncteur / vérification en exécution vs alignement en phase d'entraînement / configuration vs argument d'exécution / porte de vérification à la limite d'exécution / service de coordination). La primitive de vérification métacognitive a été reformulée, passant de « exige des agents qu'ils vérifient leur propre raisonnement » à « place une porte de vérification avant l'exécution de l'action » afin de supprimer la lecture résiduelle au niveau du modèle. Contenu inchangé ; révisions de formulation pour l'accessibilité.

v1.1 Révision de clarté de niveau 3 le jour même (15/05/2026) : ajout de l'annexe A — « Objections techniques courantes + réponses » — six paires objection-réponse couvrant le scepticisme quant à l'application des LLM, l'agnosticisme vis-à-vis du substrat, le contournement des agents, l'équivalence avec l'ingénierie des invites, l'exploitation des services d'exécution et l'évolution des valeurs. Prolongement le jour même des travaux de fond sur les niveaux 1 et 2 ; aucune revendication au-delà des spécifications primitives du §0(i).

Le présent document **ne** constitue **pas** une politique du gouvernement néo-zélandais. Il **n'** est **pas** approuvé par la Couronne. Il **n'**est **pas** fondé sur le Traité au sens formel du terme. Il s'agit d'une proposition de la société civile émanant de My Digital Sovereignty Ltd, proposée aux parties prenantes néo-zélandaises comme base d' adoption, d'adaptation ou de rejet. Lorsque ses principes sont utiles au travail de l'adoptant, ceux-ci sont libres de les utiliser sous des licences open source permissives ; lorsqu'ils ne le sont pas, ils restent sur la page.

La structure source en miroir est publiée en traduction anglaise à l'adresse `/research/translations/china-cac-implementation-guidelines-2026.html` et, à l'origine, sous le titre «*Directives de mise en œuvre* 2026 de l'Administration chinoise du cyberespace » en mandarin.

Résumé

Cet article propose un cadre souverain et fédéré pour l' application et le développement d'agents intelligents en Aotearoa Nouvelle- Zélande, présenté comme une contribution de la société civile par My Digital Sovereignty Ltd. La proposition est structurée comme un parallèle constructif aux *Directives de mise en œuvre* 2026 de la République populaire de Chine *pour l'application normalisée et le développement innovant d' agents intelligents* — six sections, quatorze sous-sections, trente-huit points numérotés — avec un nouveau chapitre §0 « Fondements philosophiques » ajouté en préambule. Le §0 s'appuie sur trois sources : les six services d'exécution du Tractatus AI Safety Framework (application des limites, surveillance de la pression contextuelle, validation des références croisées, persistance des instructions , vérification métacognitive, orchestration de la délibération pluraliste) ; les principes CARE pour la gouvernance des données autochtones (Carroll et al. 2020) et le mouvement mondial pour la souveraineté des données autochtones qui les a inspirés, y compris les travaux universitaires fondés sur le Te Tiriti de Te Mana Raraunga et du Dr Karaitiana Taiuru ; et le

paysage international coordonné par l'ISO/IEC JTC 1/SC 42 (22989 terminologie, 23053 cycle de vie, 23894 gestion des risques, 42001 systèmes de gestion). La proposition préconise la formation d'un comité sous l'égide d'une organisation faîtière appropriée — parmi les candidats figurent la Royal Society Te Apārangi, le comité miroir SC42 de Standards New Zealand et le New Zealand AI Forum — afin d'élaborer des recommandations adaptées au contexte néo-zélandais et de s'engager dans un dialogue international. Il s'agit de la version préliminaire v1 de mai 2026 ; les commentaires sont les bienvenus via les canaux habituels de commentaires sur agenticgovernance.digital.

Préambule

Les agents intelligents — des systèmes intelligents capables de perception, de mémoire, de prise de décision, d'interaction et d'exécution autonomes — accélèrent leur intégration aux registres, aux infrastructures et aux processus sociaux d'Aotearoa Nouvelle-Zélande. Cette proposition offre une contribution de la société civile sur la manière dont cette intégration devrait être gouvernée : une architecture souveraine et fédérée dans laquelle chaque opération d'un agent intelligent sur un enregistrement produit une entrée attribuée et signée cryptographiquement au nom du détenteur de l'enregistrement, et dans laquelle la coordination entre les installations souveraines s'effectue par le biais d'une fédération bilatérale. La proposition reflète la structure des *Lignes directrices de mise en œuvre* 2026 de la République populaire de Chine, de sorte que les choix architecturaux de chaque partie apparaissent en parallèle constructif, ouvrant le dialogue avec les auteurs de ce cadre, avec les pairs internationaux, ainsi qu'avec les décideurs politiques néo-zélandais, les organisateurs communautaires et les professionnels du secteur. My Digital Sovereignty Ltd propose ce document comme point de départ — pour adoption, adaptation et révision — sous des licences open source permissives. Il est proposé à titre de contribution de la société civile et ne prétend pas avoir le statut de politique de la Couronne. Lorsque ses principes sont utiles au travail de l'adoptant, celui-ci est libre de les utiliser ; lorsqu'ils ne le sont pas, ils restent sur la page.

---

§0. Fondements philosophiques

Nous commençons par les fondements car l'architecture découle de la philosophie. Les recommandations qui suivent dans les sections §I à §VI ne sont pas des choix techniques arbitraires ; elles sont les implications d'engagements philosophiques que cette section énonce explicitement. Trois courants convergent ici : la description structurelle du Tractatus AI Safety Framework sur la manière dont les agents intelligents peuvent opérer en toute sécurité par rapport aux registres détenus par des entités souveraines, développée et publiée ouvertement sur agenticgovernance.digital ; le mouvement mondial pour la souveraineté des données autochtones, qui affirme que les données concernant les personnes appartiennent à ces personnes et aux communautés auxquelles elles appartiennent ; et les travaux internationaux sur les normes en matière d'IA coordonnés par l'ISO/IEC JTC 1/SC 42, qui fournit le vocabulaire formel grâce auquel les recommandations architecturales deviennent applicables dans la pratique organisationnelle. Le fait de nommer ces trois éléments d'emblée fait partie de la contribution constructive que cette proposition apporte au dialogue — avec les auteurs des *Lignes directrices de mise en œuvre* 2026 de l'Administration du cyberespace de Chine, avec les décideurs politiques et les organisateurs communautaires néo-zélandais, et avec les homologues internationaux travaillant sur des questions parallèles.

(i) Les primitives du cadre Tractatus en tant que fondements nommés

Le cadre Tractatus se compose de six **primitives au niveau du système** qui, ensemble, spécifient les conditions architecturales dans lesquelles des agents intelligents peuvent opérer en toute sécurité sur des enregistrements détenus par des entités souveraines. **Il ne s'agit pas de techniques d'alignement au niveau du modèle** , mais de vérifications d'exécution au niveau du code qui encapsulent l' agent, indépendamment de la manière dont l'agent sous-jacent (actuels LLM de type transformateur, futures architectures de type JEPA, systèmes hybrides) est construit ou entraîné. Elles interceptent et vérifient le comportement à la limite d'exécution — la même structure architecturale que la délimitation des capacités du système de fichiers ou les vérifications de portée OAuth. Une démonstration fonctionnelle de la primitive d'application des limites est disponible à l'adresse /demos/boundary-demo.html. `[CITATION : Stroh, J. (2026). Tractatus AI Safety Framework — Valeurs et principes fondamentaux, et concepts clés du cadre Tractatus. Agentic Governance Digital. https://agenticgovernance.digital — les deux ouvrages sont sous licence CC BY 4.0.]`

L'application des limites détermine quels types de décisions nécessitent structurellement une validation humaine. Le principe fondamental — adapté de Wittgenstein et explicitement mentionné dans le cadre du Tractatus — est que « ce qui ne peut être systématisé ne doit pas être automatisé ». Les décisions relatives aux valeurs, les jugements liés au contexte culturel, les conséquences irréversibles et les situations sans précédent ne peuvent être déléguées à des agents autonomes ; le cadre empêche une telle délégation par une interception au moment de l'exécution plutôt que par l'espoir. L'interception se déclenche avant l'exécution de l'action, selon la même architecture qu'un noyau de système d'exploitation interceptant des appels système privilégiés — le processus ne peut contourner la vérification. Les quatre catégories ci-dessus sont traitées comme **des classifications faillibles négociées par la communauté dans la pratique**, et non comme des essences fixes — les erreurs de classification sont elles-mêmes enregistrées, susceptibles d'appel, et utilisées pour réviser le comportement du routeur au fil du temps (voir *Alignement architectural* §3.5). La sortie du routeur est trinaire, et non binaire : *autoriser*, *refuser* et *escalader vers la délibération humaine*. Le troisième état est porteur de sens — il reflète la reconnaissance architecturale qu'une part substantielle des décisions importantes n'est pas binaire au moment de la décision (voir §3.3).

La surveillance de la pression contextuelle reconnaît que la fenêtre contextuelle d'un agent est une ressource finie et que la pression sur la capacité est un signal de gouvernance. Les agents fonctionnant à pleine capacité commettent davantage d'erreurs, et le cadre intervient avant la défaillance plutôt qu' après. Le principe est le même que celui d'un disjoncteur : le disjoncteur se déclenche en fonction de la charge mesurée avant que le système ne s'endommage ; le cadre limite le débit ou renvoie la décision à l'approbation humaine en fonction de l'utilisation contextuelle mesurée avant que la qualité de la sortie ne se dégrade.

La validation par recoupement vérifie les actions proposées par un agent par rapport à l'historique canonique des instructions, détectant ainsi les cas où les modèles d'apprentissage prennent le pas sur les instructions explicites de l'utilisateur. Le cas illustratif est « l'incident 27027 » : un utilisateur spécifie un port de base de données non par défaut, et l'agent — malgré l'instruction explicite — utilise par défaut le numéro de port sur lequel il a été entraîné. La validation détecte cette dérogation ; sans elle, celle-ci corromprait silencieusement les opérations. Le validateur est un contrôle d'exécution sur chaque action proposée, et non un alignement du modèle lui-même lors de l'entraînement.

La classification de la persistance des instructions distingue les instructions transitoires de l'état de gouvernance durable. Toutes les instructions n'ont pas la même importance ; les traiter comme si elles l'étaient dégrade à la fois la sécurité (directives critiques oubliées) et la facilité d'utilisation (préférences insignifiantes appliquées de manière excessive). Même principe que la distinction entre configuration et arguments d'exécution dans les logiciels :

les valeurs de configuration persistent ; les arguments CLI sont propres à chaque invocation ; le classificateur balise chaque instruction par classe afin que les services en aval la traitent de manière appropriée.

La vérification métacognitive place une barrière de vérification avant l'exécution de l'action. Cette barrière évalue chaque action proposée selon cinq dimensions — alignement, cohérence, exhaustivité, sécurité et prise en compte des alternatives — et des seuils de confiance déterminent si les actions doivent se poursuivre, se poursuivre avec prudence, nécessiter un réexamen ou être bloquées. Le contrôle s'effectue à la limite de l'exécution, et non comme une exigence comportementale imposée au modèle.

L'orchestration de la délibération pluraliste facilite la délibération multipartite lorsque l'application des limites signale un conflit de valeurs. Elle ne tranche pas entre les cadres moraux ; elle structure la délibération de manière à ce que les valeurs défendues par les différentes parties prenantes soient documentées, prises en compte dans la mesure du possible, et explicitement nommées lorsqu'elles ne peuvent être conciliées. Le pluralisme fondamental — l'idée que les cadres moraux sont irréductiblement différents et qu'aucune super-valeur ne permet de les résoudre — est l'engagement philosophique qui fait de la délibération pluraliste une primitive structurelle plutôt qu'une subtilité procédurale. Il fonctionne comme un service de coordination qui documente et fait émerger les positions des parties prenantes ; l'orchestration ne demande pas à l'agent de servir de médiateur entre les valeurs en interne.

Ces six services constituent le squelette structurel de cette proposition. Chaque recommandation architecturale qui suit peut être rattachée à l'un ou à plusieurs d'entre eux.

Distinction entre substrat et exécution. Les six primitives §0(i) ci-dessus constituent la couche *d'exécution* du cadre — du code dans le processus hôte de l'agent qui vérifie les actions proposées par rapport aux règles de classe de décision configurées au moment de la décision. L'architecture comporte une deuxième couche qui ne dépend pas de la coopération d'exécution : les mécanismes *de substrat* — provenance cryptographique, enveloppes de fédération et portabilité des enregistrements pilotée par les membres — qui résident en possession distribuée indépendamment de l'agent. Un raisonnement de réseau plus approfondi autour de la couche d'exécution ne peut pas s'étendre à la couche de substrat : la sécurité du substrat découle des mathématiques (signatures, réplication distribuée, sortie sans autorisation) plutôt que de la coopération de l'agent. La couche de substrat est développée dans l'article A, *Sovereign-Record Architecture* (*Article A*), et la distinction est exposée dans la section 3.4 « *Alignement architectural* ». Les deux couches sont complémentaires : la couche d'exécution capture ce qu'elle peut ; le substrat garantit que ce que la couche d'exécution manque laisse tout de même à la communauté des enregistrements vérifiables, des voies de fédération et des options de sortie. La posture de survie est stratifiée, et non fondée sur un mécanisme unique (voir la section*Alignement architectural* §7.4 pour plus de détails sur l'indépendance vis-à-vis du confinement des agents).

(ii) Les principes CARE pour la gouvernance des données autochtones

Les principes CARE pour la gouvernance des données autochtones, publiés en 2020 par une équipe internationale de scientifiques des données autochtones sous les auspices de la Global Indigenous Data Alliance, énoncent quatre engagements : **Bénéfice collectif** (les écosystèmes de données doivent favoriser l'autodétermination autochtone et le bénéfice collectif) ; **Autorité de contrôle** (les droits et intérêts des peuples autochtones concernant leurs données doivent être reconnus) ; **Responsabilité** (ceux qui travaillent avec des données autochtones ont la responsabilité de faire savoir comment ces données sont utilisées pour soutenir l' autodétermination des peuples autochtones) ; et **Éthique** (les droits et le bien-être des peuples autochtones devraient être la préoccupation principale à toutes les étapes du

cycle de vie des données). [CITATION : 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., & Hudson, M. (2020). Les principes CARE pour la gouvernance des données autochtones. Data Science Journal, 19, 43. https://doi.org/10.5334/dsj-2020-043]

CARE se positionne comme un complément à FAIR (Findable, Accessible, Interoperable, Reusable). FAIR optimise la circulation et la réutilisation des données ; CARE optimise les droits et le bien-être des personnes concernées par ces données. Les deux ne sont pas antagonistes. Ils abordent des questions différentes : FAIR s'interroge sur la manière dont les données doivent circuler ; CARE s'interroge sur l'autorité sous laquelle les flux de données sont régis. Une architecture de souveraineté bien conçue répond à ces deux questions.

Nous adoptons CARE comme référence fondamentale. Lorsque les recommandations qui suivent précisent que les agents doivent opérer à partir d'enregistrements attribués, ancrés dans la provenance et détenus par leurs détenteurs souverains, cette précision concrétise l'engagement en matière d'autorité de contrôle. Lorsque les recommandations préconisent une coordination fédérée plutôt qu'un enregistrement centralisé, cette précision est cohérente avec la responsabilité — ceux qui détiennent les données sont responsables devant les personnes concernées par ces données.

(iii) Te Tiriti, tikanga et mātauranga dans l'éthique de l'IA — La recherche en Aotearoa Nouvelle-Zélande

La recherche sur la souveraineté des données autochtones en Aotearoa Nouvelle-Zélande est l'une des plus avancées au niveau international. La formulation initiale des principes de souveraineté des données maories est venue de Te Mana Raraunga (le réseau maori pour la souveraineté des données), fondé en 2015 avec une charte adoptée en 2016. Ces principes ont été réévalués en profondeur par le Dr Karaitiana Taiuru le 20 septembre 2025, intitulée « *Analyse critique des principes de données de Te Mana Raraunga* », qui conclut qu'ils ne traitent pas de manière adéquate l'IA, les biais de l'IA et la discrimination algorithmique, l'entraînement des modèles et l'analyse de données, le colonialisme numérique ou les impacts environnementaux ; il observe que le champ d'application de 2016 était restreint alors qu'« aujourd'hui, les données maories sont partout » ; et constate que, malgré de nombreuses citations universitaires, ces principes ne sont pour la plupart pas mis en œuvre dans la pratique. [CITATION : Taiuru, K. (20 septembre 2025). Critical Analysis of Te Mana Raraunga Data Principles. https://www.taiuru.co.nz/critical-analysis-mana-raraunga/]

L'organisme opérationnel actuellement reconnu en Aotearoa Nouvelle-Zélande pour la gouvernance des données maories est **Te Kāhui Raraunga** (créé en 2019 en tant que fondation caritative). Ses cadres publiés — le **modèle de gouvernance des données maories « Tuia te korowai o Hine-Raraunga »**, structuré autour de huit pou ; le **cadre de gouvernance de l'IA maorie** qui l'étend ; ainsi que le **rapport de synthèse sur la gouvernance maorie de l'IA** et**la ressource de référence sur les cas d'utilisation conceptuels de l'IA** — fournissent l'articulation actuelle de la gouvernance des données et de l'IA maories. Te Kāhui Raraunga décrit le cadre de gouvernance maorie de l'IA comme « activé » avec des études de cas du secteur public référencées ; une mise en œuvre à grande échelle en dehors des déploiements spécifiques dans la fonction publique reste une question ouverte que cette proposition prend en compte plutôt que de la passer sous silence. Le cadre de gouvernance maori de l'IA de Te Kāhui Raraunga stipule que « les systèmes d'IA ne doivent pas être mis en œuvre en Aotearoa sans que l'autorité maorie sur les données maories soit pleinement reconnue » ; la présente proposition ne remplace pas cette exigence. [CITATION : Te Kāhui Raraunga Charitable Trust. Modèle de gouvernance des données maories : Tuia te korowai o Hine-Raraunga, https://www.kahuiraraunga.io/maoridatagovernance ;

Cadre de gouvernance maori de l’IA, https://www.kahuiraraunga.io/maoriaigovernance ; détails bibliographiques complets des publications datées en attente de vérification des sources primaires.]

Les travaux universitaires publiés par le Dr Karaitiana Taiuru sur les cadres éthiques maoris pour l'IA, sur le tikanga (loi et coutumes maories) dans l'éthique de l'IA, sur l'IA respectueuse du IA respectueuse du Tiriti, et sur la protection du mātauranga (savoir maori) dans les données d'entraînement de l'IA — y compris l'analyse critique du 20 septembre 2025 citée ci-dessus — a fourni un langage fondamental pour réfléchir à l' IA agentive dans les contextes te ao Māori. Nous adoptons les termes de référence qu’il privilégie lorsque cela est approprié : **mana motuhake** et **rangatiratanga** plutôt que les cadres conceptuels occidentaux prescrits ; des cadres réactifs et adaptatifs ancrés dans le tikanga, capables d’évoluer avec les changements technologiques et sociaux ; des cadres adaptés à des organisations et des secteurs spécifiques, développés en partenariat avec les parties prenantes maories concernées. Nous citons ses travaux comme référence fondamentale ; nous ne présentons pas son nom — ni celui de quiconque — comme approuvant cette proposition spécifique. C'est aux tangata whenua qu'il appartient de déterminer ce qui constitue une utilisation appropriée des agents intelligents dans les contextes te ao Māori, et non à cette proposition de le préciser.

Analyse des lacunes — ce que cette proposition fait et ne fait pas encore Une évaluation honnête importe davantage que des déclarations ambitieuses pour une version 1.1 adressée aux évaluateurs, dont le Dr Taiuru. Les éléments architecturaux fondamentaux de la proposition — souveraineté par attribution, provenance cryptographique, portabilité des membres, fédération bilatérale — sont **compatibles avec la mise en œuvre de** l’autorité maorie sur les données maories dans le cadre de Te Kāhui Raraunga et selon les termes de fond privilégiés par Taiuru. Les points de compatibilité comprennent :

- **L'isolation des locataires en tant que principe fondamental** (propriété de déploiement Village, ancrée sur la primitive d'application des limites du Tractatus) met en œuvre, par l'architecture, l'exigence selon laquelle les systèmes d'IA ne doivent pas être mis en œuvre sans respecter l'autorité maorie sur les données maories. Un locataire géré par un hapū, un iwi ou un organisme kaitiaki conserve ses enregistrements sous l'autorité de cet organisme par la conception même du cadre plutôt que par une simple promesse.
- **L'application des limites** peut être configurée pour exiger une autorisation explicite du kaitiaki / hapū / iwi pour les opérations sensibles sur les enregistrements du locataire ; le cadre assure cette application par l'architecture plutôt que par la confiance.
- **La provenance cryptographique et les identifiants portables par les membres** soutiennent le kaitiakitanga à travers les générations : les enregistrements portent leur propre piste d'audit, ne peuvent être modifiés à l'insu de tous, et les membres peuvent migrer vers une autre installation souveraine sous la même architecture.
- **La primitive de délibération pluraliste** est conçue pour la délibération morale multi-cadres lorsque l'application des limites signale un conflit de valeurs ; elle est en principe capable d'intégrer les cadres kaupapa Māori aux côtés d'autres cadres dans une délibération structurée.

Ce que cette proposition **ne fait pas encore**, et ce que les évaluateurs devraient prendre en compte en conséquence, est tout aussi important à mentionner :

- **Le mana motuhake et le rangatiratanga comme fondement philosophique.** Le pluralisme**fondamental** du cadre du Tractatus est lui-même un engagement philosophique occidental, s'inspirant de Berlin, Rawls et Ostrom ; il intègre le kaupapa Māori comme un cadre parmi d'autres ; il ne repose pas sur le mana motuhake et le rangatiratanga en tant qu'engagements préalables. Combler cette lacune nécessiterait que les fondements du cadre soient reformulés à partir d’un point de départ kaupapa

Māori — un travail de fond qui ne peut honnêtement être réalisé par les auteurs actuels seuls.

- **Partenariat de conception mené par les autochtones.** Le cadre Tractatus et la mise en œuvre Village ont été développés par le directeur de My Digital Sovereignty Ltd (Pākehā), avec des contributions ultérieures de Claude (Anthropic) en tant qu'auteur. Ils n'ont pas été co-conçus avec des parties prenantes maories. La recommandation de Taiuru concernant des « cadres adaptés à des organisations et secteurs spécifiques, développés en partenariat avec les parties prenantes maories concernées » n'est pas satisfaite au niveau du processus de conception. L'architecture *peut* être appliquée en partenariat ; le cadre lui-même n'a pas été co-développé en partenariat.
- **Biais de l'IA sur les dimensions culturelles et raciales au niveau de la couche des données d'entraînement.** La validation par recoupement détecte les remplacements de modèles d'entraînement au niveau de la couche des conflits d'instructions ; elle ne traite pas des biais plus profonds intégrés dans les données d'entraînement qui ne se manifesteraient pas sous forme de conflits d'instructions. Le document d'accompagnement B sur les couches de langage situées (discipline de formation par locataire, position de non-modification des poids, inférence liée à la juridiction) aborde plus directement ces préoccupations que ne le fait le cœur du Tractatus.
- **Le colonialisme numérique en tant que concept théorique et politique nommé.** L'isolation des locataires et la souveraineté architecturale de la proposition constituent des réponses structurelles partielles au colonialisme numérique ; la proposition n'aborde pas le concept d'un point de vue théorique. Le livre blanc sur l'équité distributive (référencé sur ce site) aborde cette question de manière plus explicite que ne le fait le cœur de Tractatus.
- **Les impacts environnementaux de l'IA** ne sont pour l'essentiel pas abordés. L'architecture d'inférence de repli sur CPU dans le déploiement du Village constitue une réponse opérationnelle partielle ; elle ne fait pas partie des engagements déclarés du cadre.

L'implication honnête de cette analyse des lacunes est que la proposition offre les éléments architecturaux de base que nécessiterait la mise en œuvre de l'autorité maorie sur les données maories, tout en reconnaissant que la mise en œuvre dans les contextes te ao Māori est une entreprise distincte et substantielle qui nécessite un travail de conception mené par le kaupapa maori, ce que cette proposition n'a pas fait. La proposition du comité au §II, point 4, est un mécanisme par lequel ce travail supplémentaire pourrait être fait avancer ; elle est proposée à titre de réflexion plutôt que comme une réponse complète.

La Charte des algorithmes pour l'Aotearoa Nouvelle-Zélande, signée par les agences de la Couronne en 2020, fournit la base de référence existante en matière de transparence, de partenariat avec les Maoris, d'équité, de responsabilité et de protection des données dans la prise de décision algorithmique de la Couronne. La présente proposition ne remplace pas la Charte des algorithmes ; les recommandations qui suivent sont destinées à être mises en œuvre dans le cadre de celle-ci et parallèlement à celle-ci, ainsi qu'aux cadres de Te Kāhui Raraunga. [RÉFÉRENCE : Charte des algorithmes pour l'Aotearoa Nouvelle-Zélande (2020). https://www.data.govt.nz/leadership/governance/data-ethics/algorithm-charter/ — état actuel et mises à jour ultérieures en attente de vérification.]

(iv) La lignée mondiale de la souveraineté des données autochtones

La souveraineté des données autochtones est un mouvement international, et non une particularité néo-zélandaise. Il est important de mentionner cette tradition internationale : cela inscrit le travail fondé sur le Te Tiriti évoqué plus haut dans un débat mondial plutôt que dans un localisme étroit, et cela ouvre un terrain d'entente avec les auteurs du cadre CAC en tant que contributeurs à des approches non occidentales de la manière dont les données et l'IA devraient être gouvernées.

Le Centre de gouvernance de l'information des Premières Nations (FNIGC) au Canada applique les principes **OCAP** — propriété, contrôle, accès, possession — initialement formulés dans les années 1990 dans le cadre de l'Enquête régionale longitudinale sur la santé des Premières Nations et désormais intégrés dans les pratiques en matière d'éthique de la recherche au sein des universités, des administrations et des communautés des Premières Nations canadiennes. [CITATION : Centre de gouvernance de l'information des Premières Nations. Les principes OCAP® des Premières Nations. https://fnigc.ca/ocap-training/]

Le Réseau américain pour la souveraineté des données autochtones (USIDSN), créé en 2016 en lien avec le Native Nations Institute de l'Université d'Arizona, a fait progresser la pratique de la souveraineté des données autochtones dans le contexte américain, notamment par son engagement dans les processus fédéraux américains en matière de politique des données. [CITATION : Réseau américain pour la souveraineté des données autochtones. https://usindigenousdata.org/]

Le collectif Maiam nayri Wingara pour la souveraineté des données autochtones en Australie — dont le nom signifie « De nombreuses voix, un seul esprit » — a été créé en 2017 et a publié en 2018 un communiqué sur la souveraineté des données autochtones qui a façonné la pratique des données autochtones en Australie. [CITATION : Collectif Maiam nayri Wingara pour la souveraineté des données autochtones. (2018). Communiqué sur la souveraineté des données autochtones.]

La **Global Indigenous Data Alliance** (GIDA) assure la coordination entre ces réseaux et d'autres réseaux nationaux de souveraineté des données autochtones à l'échelle internationale ; c'est sous son égide que les principes CARE ont été publiés. [CITATION : Global Indigenous Data Alliance. https://www.gida-global.org/]

Ce n'est pas un hasard si une grande partie du travail philosophique de fond de cette proposition remonte à la recherche autochtone. Les questions récurrentes — sous quelle autorité les données et les agents qui les exploitent agissent-ils ? À qui la responsabilité et la provenance sont-elles dues ? Quelle est l'échelle appropriée à laquelle les intérêts collectifs sont mis en balance avec les intérêts individuels ? — sont des questions sur lesquelles la souveraineté des données autochtones travaille depuis des décennies. La résistance aux architectures extractives des géants de la tech, et la formulation d'alternatives architecturales fondées sur l'autorité collective, constituent l'une des contributions les plus fécondes de ce mouvement international. Les recommandations qui suivent s'inspirent de cette tradition et s'adressent à elle dans le cadre d'un dialogue.

(v) ISO/IEC JTC 1/SC 42 : le paysage international des normes en matière d'IA

Les travaux internationaux sur les normes en matière d'IA, coordonnés par l'ISO/IEC JTC 1/SC 42, fournissent le vocabulaire formel et les cadres de systèmes de gestion dans lesquels des recommandations de ce type deviennent applicables dans la pratique organisationnelle. Quatre normes sont particulièrement pertinentes.

La norme ISO/IEC 22989:2022 spécifie les concepts et la terminologie de l'intelligence artificielle. Nous utilisons la terminologie de la norme ISO/IEC 22989 lorsqu'elle est compatible — par exemple, le terme « système d'IA » reprend sa définition de la norme 22989. La cohérence terminologique rend cette proposition lisible pour les examinateurs rigoureux en matière de normes et applicable parallèlement à d'autres travaux alignés sur la norme 22989. [CITATION : ISO/IEC 22989:2022. Technologies de l'information — Intelligence artificielle — Concepts et terminologie de l'intelligence artificielle. Organisation internationale de normalisation / Commission électrotechnique internationale.]

La norme ISO/IEC 23053:2022 établit un cadre pour les systèmes d'IA utilisant l'apprentissage automatique, en cartographiant les composants d'un système d'IA basé sur l'apprentissage automatique et les relations entre eux. Les recommandations de cette

proposition concernant le cycle de vie, la provenance ou l’ attestation au niveau des composants peuvent être mises en œuvre parallèlement aux étapes du cycle de vie de la norme 23053 . [RÉFÉRENCE : ISO/IEC 23053:2022. Cadre pour les systèmes d’intelligence artificielle (IA) utilisant l’apprentissage automatique (ML). ISO/IEC.]

La norme ISO/IEC 23894:2023 fournit des lignes directrices sur la gestion des risques liés à l’IA. Elle constitue l’équivalent, au niveau des organismes de normalisation, des recommandations du cadre relatives à la surveillance des risques et à la gestion des incidents au niveau de chaque installation. [RÉFÉRENCE : ISO/IEC 23894:2023. Technologies de l’information — Intelligence artificielle — Lignes directrices sur la gestion des risques. ISO/IEC.]

La norme ISO/IEC 42001:2023 spécifie les exigences relatives à un système de gestion de l'IA. Elle constitue l'équivalent pour l'IA des normes ISO/IEC 27001 (gestion de la sécurité de l'information) et ISO 9001 (gestion de la qualité). Nous considérons que les recommandations de cette proposition sont applicables dans le cadre d'un système de gestion de type ISO/IEC 42001 ; les organisations adoptant une partie de cette proposition sont susceptibles d’être celles qui appliquent déjà, ou prévoient d’appliquer, une gouvernance alignée sur la norme ISO/IEC 42001. [RÉFÉRENCE : ISO/IEC 42001:2023. Technologies de l’information — Intelligence artificielle — Système de gestion. ISO/IEC.]

Les travaux du comité qui aboutissent à ces normes impliquent des comités miroirs nationaux dans de nombreuses juridictions, notamment au Royaume-Uni (via la British Standards Institution) et d’autres organismes nationaux de normalisation à l’échelle internationale. La participation d’Aotearoa Nouvelle-Zélande aux travaux du SC42 — par l’intermédiaire de Standards New Zealand ou d’un comité miroir constitué à cette fin — est l’un des cadres dans lesquels la contribution constructive préconisée par la présente proposition pourrait naturellement se concrétiser. [REMARQUE : il convient de vérifier l'existence d'un comité miroir SC42 néo-zélandais avant de rédiger les paragraphes des points 4, 12, 14, 35 et 38.]

Conclusion

Ces cinq courants — le Tractatus, CARE, les travaux universitaires sur la souveraineté des données autochtones fondés sur le Te Tiriti, le mouvement mondial pour la souveraineté des données autochtones et l’ISO/IEC SC42 — convergent vers les choix architecturaux que le reste de cette proposition précise. La souveraineté en tant qu’attribution ; la fédération bilatérale en tant que coordination ; la gouvernance polycentrique en tant que structure d’autorité ; la provenance cryptographique en tant qu’infrastructure d’audit : aucun de ces éléments n’a été inventé pour cette proposition. Chacun trouve ses racines dans une ou plusieurs des lignées mentionnées ci-dessus. Ce que cette proposition apporte, c’est un agencement particulier de ces éléments primitifs, adapté au contexte de l’Aotearoa Nouvelle-Zélande , proposé comme un parallèle constructif au cadre avec lequel elle partage sa structure.

---

§I. Principes fondamentaux

Nous proposons quatre principes fondamentaux pour le développement souverain et fédéré d'agents intelligents en Aotearoa Nouvelle-Zélande. Chacun correspond à l'un des quatre principes qui ouvrent les *Lignes directrices de mise en œuvre* 2026 de l'Administration du cyberespace de Chine ; dans chaque cas, nous affirmons l'intention sous-jacente du principe et proposons un parallèle constructif fondé sur les fondements du §0.

Souveraineté et attribution. Chaque opération d’un agent intelligent sur un enregistrement

est attribuable à un détenteur souverain de cet enregistrement ; la provenance est cryptographique ; la sécurité découle de l' autorité du détenteur de l'enregistrement sur ses propres enregistrements. Nous affirmons que l'engagement du cadre de la CAC en faveur de la sécurité et de la contrôlabilité est fondamental. Nous proposons, à titre de parallèle constructif, que dans le contexte de l'Aotearoa Nouvelle-Zélande — où convergent le partenariat Te Tiriti, le cadre existant de la loi de 2020 sur la protection de la vie privée et l'engagement de l'Autorité CARE en matière de contrôle — la souveraineté fondée sur l'attribution soit bien adaptée à la mise en œuvre de ces mêmes préoccupations de sécurité. La primitive d'application des limites de Tractatus fournit le mécanisme architectural ; les enregistrements signés cryptographiquement fournissent la piste d'audit ; et l'autorité légitime sur les deux est le détenteur des enregistrements, en vertu de la juridiction et des obligations de partenariat. *(Parallèles avec le principe 1 du CAC §I « sécurité et contrôlabilité »).* [RÉFÉRENCES : Application des limites Tractatus (Stroh 2026, CC BY 4.0) ; Principes CARE, Engagement « Autorité de contrôle » (Carroll et al. 2020) ; Loi sur la protection de la vie privée de 2020 (NZ), principes de confidentialité des informations.]

Bilatéral et fédéré. La coordination entre les installations souveraines s'effectue par le biais d'une fédération bilatérale et de normes internationales ouvertes. Nous reconnaissons le mérite de l'engagement du cadre CAC en faveur d'un développement normalisé et ordonné ; la normalisation et l'ordre sont des conditions nécessaires à tout déploiement à grande échelle d'IA agentique, et le programme de normalisation coordonné du cadre CAC constitue une approche crédible. Nous proposons, pour le contexte néo-zélandais — à plus petite échelle, des principes bien établis de souveraineté des données maories , des accords institutionnels bilatéraux existants entre les agences de la Couronne , les hapū, les iwi, les organisations de la société civile et le secteur privé — qu'une approche fédérée de la coordination est tout à fait adaptée. La fédération entre les installations souveraines est bien prise en charge par les normes existantes du W3C, de l'IETF et alignées sur l'ISO/IEC SC42. Nous proposons une fédération bilatérale à examiner en tant qu'architecture parallèle susceptible d'interopérer avec les approches d'enregistrement centralisé dans d'autres juridictions, et nous invitons à la formation de comités sous l'égide d'organisations faîtières appropriées afin de développer le dialogue sur l'interopérabilité . *(Parallèle avec le principe 2 du CAC §I « développement normalisé et ordonné »).* [RÉFÉRENCES : Identificateurs décentralisés (DID) v1.0 du W3C (Recommandation du W3C, 2022) et Modèle de données pour les identifiants vérifiables v1.1 du W3C ; ActivityPub (Recommandation du W3C, 2018) ; ISO/IEC 42001:2023 systèmes de gestion ; Cadre de gouvernance de l'IA Te Kāhui Raraunga Māori + analyse critique de Taiuru (voir §0(iii)).]

Délibération pluraliste, polycentrique. Plusieurs cadres de valeurs coexistent au sein et entre les entités souveraines ; la délibération entre elles est procédurale et structurée ; l'innovation résulte d'une adaptation locale sous l'autorité locale. Nous affirmons l'engagement du cadre CAC en faveur d'un développement axé sur l'innovation. Nous proposons, à titre de parallèle constructif, que la gouvernance polycentrique — multiples centres d' autorité, multiples cadres de valeurs maintenus dans une tension productive, avec une délibération structurée en cas de conflit — soit bien adaptée au contexte néo-zélandais (Aotearoa) du partenariat Te Tiriti, et soit solidement étayée par la recherche internationale sur la gouvernance polycentrique (notamment les travaux fondateurs d'Elinor Ostrom). La primitive de délibération pluraliste du Tractatus fournit le mécanisme architectural permettant de faciliter la délibération multipartite lorsque l'application des limites met en évidence un conflit de valeurs ; le pluralisme fondamental est l'engagement philosophique qui en fait une caractéristique structurelle du cadre. *(Parallèle avec le principe 3 du CAC §I « développement axé sur l'innovation »).* [RÉFÉRENCES : Primitive de délibération pluraliste du Tractatus (Stroh 2026, CC BY 4.0) ; Ostrom, E. (2010). Beyond Markets and States: Polycentric Governance of Complex Economic Systems. American Economic Review, 100(3), 641–672. https://doi.org/10.1257/aer.100.3.641 — détails

bibliographiques complets à vérifier avant la publication de la v1.]

Fondé sur l'adoption, étayé par des preuves. Les applications des agents intelligents sont attestées par leur déploiement dans les communautés qui les ont adoptés ; pour une proposition de la société civile, la base probatoire appropriée est le déploiement dans le monde réel. Nous affirmons l'engagement du cadre CAC en faveur d'un développement axé sur les applications. Nous proposons, à titre de parallèle constructif, que pour une proposition de la société civile émanant d'une seule entreprise, les preuves de déploiement doivent précéder la recommandation. Lorsque cette proposition cite des exemples de déploiement en Aotearoa Nouvelle-Zélande (aux §IV et §V) — dans des contextes paroissiaux et hapū / iwi, dans des contextes d'histoire familiale iwi et de la diaspora, dans des contextes de petites entreprises — ces citations renvoient à des déploiements réels, avec des données de déploiement concrètes (nombres, dates de lancement, portée) à ajouter avant la publication de la version 1. Lorsque la proposition avance des recommandations dans des secteurs où le MDSL n'a pas encore été déployé, ces recommandations sont formulées comme des conditions d'architecture de souveraineté pour tout déploiement d'agent dans ce secteur, adressées à quiconque souhaiterait y appliquer l'architecture. Ce que le déploiement démontre, ce n'est pas seulement le nombre d'adoptions, mais la disponibilité architecturale — c'est-à-dire que les primitives du substrat (provenance cryptographique, enveloppes de fédération, portabilité pilotée par les membres, application des limites) fonctionnent comme spécifié dans des conditions réelles avec les données des communautés ayant autorisé l'essai, plutôt que dans des benchmarks artificiels. *(Parallèle avec le principe 4 de la section I du CAC « approche pilotée par les applications »).* [RÉFÉRENCES : preuves de déploiement MDSL — Village (contextes paroissiaux et communautaires), histoire familiale (contextes iwi et de la diaspora), sydigital (contextes de petites entreprises) ; données de déploiement spécifiques (nombres, dates de lancement, portée des baux) en attente de chiffres vérifiés par les opérateurs avant la publication de la v1.]

---

§II. Fondements du développement souverain

Alors que le cadre de l'Administration chinoise du cyberespace consolide les fondements technologiques dans le cadre d'un programme de normalisation coordonné par l'État, nous proposons des fondements ancrés dans la souveraineté cryptographique et les protocoles bilatéraux. Les deux sous-sections qui suivent — renforcer les fondements de la souveraineté et établir des protocoles bilatéraux — précisent ensemble les primitives architecturales sur lesquelles repose le reste de la proposition.

(I) Renforcement des fondements de la souveraineté

Point 1. Développer des primitives souveraines pour les agents. Les enregistrements signés cryptographiquement, les identifiants portables par les membres et la provenance attribuée constituent le fondement des opérations des agents sur les données souveraines. Il s'agit de primitives architecturales qui garantissent la souveraineté au niveau même de l'enregistrement. Nous proposons un investissement soutenu dans les primitives cryptographiques open source — signature numérique, identifiants vérifiables, stockage adressé par contenu avec provenance — et dans des normes d'identité portables utilisables dans toute installation souveraine de n'importe quel secteur. Nous proposons que ces primitives soient développées et maintenues en tant qu'infrastructure commune, disponible sous des licences open source permissives (Apache 2.0, EUPL-1.2 ou compatibles) permettant l'adoption, la modification et la redistribution par toute partie. La primitive Tractatus d'application des limites, la primitive de validation par références croisées et la primitive de classification de la persistance des instructions spécifient ensemble les mécanismes d'exécution ; la signature cryptographique et l'infrastructure d'identifiants vérifiables

fournissent la piste d'audit sous-jacente. *(Parallèles avec le point 1 du CAC : « renforcer la R&D dans les technologies fondamentales »).* [RÉFÉRENCES : Cadre Tractatus (Stroh 2026, texte CC BY 4.0 / code Apache 2.0) ; Identifiants décentralisés (DID) v1.0 du W3C (Recommandation du W3C, 2022) ; Modèle de données des identifiants vérifiables du W3C v1.1 ; Principes CARE, engagement « Authority-to-control » (Carroll et al. 2020).]

Point 2. Affiner la chaîne d'outils souveraine. Des implémentations de référence open source des frameworks d'agents — y compris les six services du framework Tractatus — devraient être disponibles pour être adoptées par toute installation souveraine sous des licences open source permissives autorisant un fonctionnement local à l'installation. Nous proposons que la chaîne d'outils pour le développement, le test, le déploiement et la maintenance de systèmes agents à architecture souveraine soit développée en open source, en encourageant les contributions de toute installation souveraine. Les implémentations MDSL actuelles — le cadre Tractatus distribué sous Apache 2.0 (avec une documentation sous CC BY 4.0) ; les bases de code Village et Community migrant vers EUPL-1.2 par étapes à partir de mi-2026 — sont proposées comme un ensemble de implémentations de référence parmi plusieurs autres possibles. Les outils de sécurité — détection des entrées adversaires, détection des anomalies comportementales, outils d'attestation pour les builds et les dépendances — constituent le complément technique approprié aux primitives de Tractatus pour l'application des limites et la vérification métacognitive. *(Parallèle avec le point 2 du CAC « affiner la chaîne d'outils de l'agent »).* [RÉFÉRENCES : implémentation de référence du cadre Tractatus (Stroh 2026), Apache 2.0 (code), CC BY 4.0 (texte et figures) ; EUPL-1.2 (Licence publique de l'Union européenne) ; Apache 2.0 (Apache Software Foundation).]

(II) Mise en place de protocoles bilatéraux

Point 3. Protocoles bilatéraux fédérés. L'interopérabilité entre les installations souveraines s'effectue par le biais d'accords bilatéraux et de normes internationales ouvertes. Nous reconnaissons le mérite de l'engagement du cadre CAC en faveur d'un programme d'interconnexion normalisé — le protocole d'interconnexion des agents intelligents (AIP) proposé, des normes d'interface fondamentales pour les logiciels, les services et les périphériques matériels, ainsi que des normes obligatoires dans les secteurs sensibles. Nous proposons, dans le contexte d'Aotearoa Nouvelle-Zélande, que l'interopérabilité entre les installations souveraines soit bien prise en charge par le paysage actuel des normes internationales : les identifiants décentralisés et les identifiants vérifiables du W3C pour l'identité ; ActivityPub et les protocoles de fédération W3C associés pour la communication inter-installations ; les protocoles de l'IETF pour l'authentification, le transport et l'adressage de contenu ; et les travaux de l'ISO/IEC SC42 sur la terminologie spécifique à l'IA, le cycle de vie, les risques et l'alignement des systèmes de gestion. Nous proposons qu'Aotearoa NZ contribue aux normes internationales d'interopérabilité en tant que participant à part entière au sein de ces forums existants. *(Parallèles avec le point 3 du CAC « système de normalisation » et le protocole d'interconnexion AIP proposé.)* [RÉFÉRENCES : W3C DIDs v1.0 ; W3C Verifiable Credentials Data Model v1.1 ; ActivityPub (Recommandation W3C, 2018) ; terminologie ISO/IEC 22989:2022 ; cadre ML ISO/IEC 23053:2022.]

Point 4. Identité cryptographique ; dialogue fédéré sur l' Internet intelligent. L'identité est propre à chaque installation, ancrée dans le DNS et les clés cryptographiques ; la vérification entre les contreparties est de pair à pair ; les déclarations de capacités sont publiées par chaque installation. Nous reconnaissons le mérite de la proposition du cadre CAC concernant une plateforme d'enregistrement d'agents intelligents, qui envisage non seulement la gestion de l'identité numérique et la déclaration de capacités, mais aussi la recherche et la découverte, l'interconnexion de confiance, le paiement conforme, la protection de la sécurité, la résolution des conflits, l'exploitation de l'IPv6 et un système d'indicateurs

de surveillance — un ensemble substantiel et cohérent de fonctions interdépendantes. Une plateforme d'enregistrement centralisée dotée d'une autorité de coordination constitue une approche architecturale crédible pour ces fonctions.

Nous proposons, dans le contexte de l'Aotearoa Nouvelle-Zélande — où convergent des principes de souveraineté des données maories à petite échelle et bien établis, ainsi que les primitives architecturales déjà représentées dans les déploiements MDSL — une approche fédérée dans laquelle chaque fonction de l'Internet intelligent est traitée par le biais d'accords bilatéraux et de normes internationales ouvertes. L'identité et la déclaration de capacités sont assurées par les identifiants décentralisés et les identifiants vérifiables du W3C. La recherche et la découverte entre des installations souveraines peuvent s'appuyer sur les modèles établis par la fédération dérivée d'ActivityPub, par WebFinger (IETF RFC 7033) et par des protocoles de répertoire compatibles avec la fédération tels que nodeinfo — bien que nous notions que la découverte fédérée à grande échelle reste un problème d'ingénierie non résolu et que nous le reconnaissions comme tel. L'interconnexion de confiance et la protection de la sécurité fonctionnent grâce à une attestation cryptographique bilatérale. Les voies de paiement conformes passent par les canaux réglementaires financiers existants. La résolution des conflits s'effectue par la médiation bilatérale et les mécanismes existants de résolution des litiges, la provenance cryptographique fournissant la piste d'audit. IPv6 est un choix d'infrastructure sous-jacente disponible pour toute installation. Un système d'indicateurs de surveillance est réalisable grâce à la publication ouverte des métriques opérationnelles par chaque installation participante, agrégées par des observateurs indépendants.

Nous proposons la création d'un comité unique sous l'égide d'une organisation faîtière appropriée — parmi les candidats figurent la Royal Society Te Apārangi, le comité miroir SC42 de Standards New Zealand (existence à vérifier), le New Zealand AI Forum, ou une structure conjointe regroupant ces entités — **afin d'élaborer des recommandations détaillées sur l'architecture de l'IA agentique dans le contexte néo-zélandais, de contribuer aux travaux de l'ISO/IEC JTC 1/SC 42 en tant que participant à part entière, et d'engager un dialogue bilatéral avec les auteurs du cadre CAC et avec des pairs internationaux. Le comité serait chargé de cinq axes de travail spécifiques : (i) l'identité fédérée pour les agents intelligents et les fonctions plus larges de l'Internet intelligent mentionnées dans ce point ; (ii) les services fédérés d'audit et de conformité (voir §III point 12) ; (iii) les systèmes de réputation basés sur l'attestation (voir §III point 14) ; (iv) les modèles de coordination sectorielle, y compris les modèles de fédération par opposition aux modèles d'alliance (voir §V point 35) ; et (v) l'engagement international et la coopération bilatérale sur l'IA agentique (voir §V point 38). La contribution du comité aux travaux de normalisation internationale et au dialogue avec les auteurs du cadre CAC constitue son principal résultat.** Nous proposons cette proposition de comité comme une contribution au débat international ; ce débat bénéficiera des contributions issues de nombreuses traditions architecturales. *(Parallèles avec le point 4 du CAC « architecture Internet intelligente » et la plateforme d'enregistrement ; le modèle de formation du comité est consolidé dans les points 4, 12, 14, 35 et 38.)* [RÉFÉRENCES : Identifiants décentralisés (DID) du W3C v1.0 ; Modèle de données des identifiants vérifiables du W3C v1.1 ; ActivityPub (Recommandation du W3C 2018) ; WebFinger (RFC 7033 de l'IETF) ; répertoire de fédération nodeinfo ; ISO/IEC 22989:2022 ; Te Kāhui Raraunga (kahuiraraunga.io — Modèle maori de gouvernance des données et cadre maori de gouvernance de l'IA) ; Taiuru, K. (20 septembre 2025) Analyse critique des principes de données Te Mana Raraunga, taiuru.co.nz/critical-analysis-mana-raraunga/ ; Royal Society Te Apārangi ; ISO/IEC JTC 1/SC 42.]

§III. Préserver la base de référence en matière de souveraineté

Alors que le cadre de l'Administration chinoise du cyberespace établit une base de référence en matière de sécurité par le biais de directives sur les produits, de technologies de confinement comportemental, d'une gouvernance à plusieurs niveaux et d'une autorégulation du secteur assortie de sanctions en matière de notation de crédit, nous proposons une base de référence ancrée dans le cadre juridictionnel propre à l'adoptant, la provenance cryptographique, des dispositifs de gouvernance polycentriques et une coordination fondée sur la fédération. Les quatre sous-sections qui suivent — principes relatifs aux produits, risques de sécurité, système de gouvernance, coordination fédérée — précisent ensemble comment la conformité d'un agent intelligent aux principes de souveraineté peut être vérifiée lors de l'exécution et auditée a posteriori.

(I) Clarification des principes relatifs aux produits

Point 5. S'ancrer dans les lois propres à l'adoptant. Les politiques, réglementations et normes éthiques régissant les agents intelligents découlent de la juridiction de l'adoptant. Les valeurs proviennent du droit local et des arrangements institutionnels locaux ; l'architecture fournit l' infrastructure de mise en œuvre dans laquelle ces valeurs s'appliquent. En Aotearoa Nouvelle-Zélande, les instruments applicables comprennent la loi sur la protection de la vie privée de 2020 (avec le Code de confidentialité des informations de santé de 2020 et d'autres codes applicables à des secteurs spécifiques) ; la loi néo-zélandaise sur la Charte des droits de 1990 lorsque des acteurs étatiques sont impliqués ; la Charte des algorithmes pour l'Aotearoa Nouvelle-Zélande pour les agences de la Couronne ; les obligations découlant du Te Tiriti o Waitangi pour les acteurs de la Couronne et les obligations de partenariat qu'elles impliquent ; la loi officielle de 1982 ; la loi sur la fonction publique de 2020 ; la loi sur les archives publiques de 2005 ; et les lois sectorielles, notamment la loi sur la Banque centrale de Nouvelle-Zélande de 2021, la loi sur l'éducation et la formation de 2020, la loi sur les collectivités locales de 2002 et la loi sur les perquisitions et la surveillance de 2012, applicables au contexte de déploiement concerné. L'architecture est neutre quant à la mise en œuvre, quelle que soit la juridiction dont la loi s'applique ; la proposition s'adresse aux adoptants d'Aotearoa NZ, et les mêmes éléments de base servent les adoptants dans toute juridiction dont ils souhaitent mettre en œuvre les valeurs. *(Parallèles avec le point 5 du CAC « politiques, réglementations et normes éthiques »).* [RÉFÉRENCES : Loi sur la protection de la vie privée de 2020 (NZ) ; Loi sur la Charte des droits de la Nouvelle-Zélande de 1990 ; Charte des algorithmes pour Aotearoa Nouvelle-Zélande (2020) ; Code de confidentialité des informations de santé 2020 ; Loi sur l'information officielle de 1982 ; Loi sur la fonction publique de 2020 ; Loi sur les archives publiques de 2005 ; Loi sur la Banque de réserve de Nouvelle-Zélande de 2021 ; Loi sur l'éducation et la formation de 2020 ; Loi sur les collectivités locales de 2002 ; Loi sur les perquisitions et la surveillance de 2012 — versions législatives actuelles à vérifier avant la publication de la v1.]

Point 6. Pouvoir de décision final de l'utilisateur, garanti par la cryptographie. Nous affirmons le même principe que celui défendu par le cadre CAC : l'utilisateur conserve le droit d'être informé et le pouvoir de décision final sur les actions autonomes entreprises par des agents intelligents en son nom. Ce principe est fondamental pour la relation de confiance entre une personne et les systèmes agents agissant en son nom. Nous proposons, comme mécanisme d'audit, une provenance cryptographique par enregistrement par rapport à l'enregistrement souverain de l'utilisateur : chaque action autonome d'un agent opérant sur les enregistrements de l'utilisateur produit une entrée cryptographique attestant de l'action, attribuable à l'agent et au cadre d'autorisation de l'utilisateur. L'utilisateur peut inspecter, rejouer et contester toute action d'un agent par rapport à cette provenance, et la primitive « instruction-persistance-classification » de Tractatus fournit le cadre permettant de distinguer les actions de routine de celles nécessitant une reconfirmation explicite de l'utilisateur.

(Parallèles avec le point 6 du CAC : « clarifier l'autorité décisionnelle »). [RÉFÉRENCES : Primitive de classification de la persistance des instructions Tractatus (Stroh 2026, CC BY 4.0) ; Privacy Act 2020 (NZ), principe de confidentialité des informations n° 6 (droits d'accès) ; Principes CARE, engagement « Authority-to-control » (Carroll et al. 2020).]

Point 7. Provenance, en complément du contrôle comportemental . Nous reconnaissons l'importance accordée par le cadre CAC à l' intégration de règles, au cloisonnement comportemental et à la vérification, ancrée dans la blockchain, du comportement des agents dans des scénarios d'application critiques. Il s'agit d' approches architecturales crédibles pour garantir un comportement légal et conforme dans les déploiements coordonnés de manière centralisée. Nous proposons, en tant que primitive architecturale supplémentaire bien adaptée à la fédération bilatérale, **la provenance**: chaque action d'un agent intelligent produit un enregistrement cryptographique attribuable à l'acteur. Les deux approches se complètent. Le cloisonnement comportemental limite ce qu'un agent peut tenter d'effectuer lors de l'exécution ; la provenance crée un enregistrement infalsifiable de ce qui a effectivement été tenté. Les deux ont un rôle à jouer, et l'équilibre approprié entre eux dépend probablement du contexte. *(Parallèles avec le point 7 du CAC « renforcer le contrôle comportemental »).* [RÉFÉRENCES : primitive de validation des références croisées Tractatus (Stroh 2026, CC BY 4.0) ; modèle de données des identifiants vérifiables du W3C v1.1 ; ISO/IEC 23894:2023 gestion des risques.]

(II) Atténuation des risques de sécurité Risques

Point 8. Sécurité intrinsèque grâce à des primitives souveraines. Les informations personnelles restent dans l'installation du détenteur ; la protection cryptographique s'applique à chaque enregistrement et est basée sur le périmètre ; la détection des attaques s'effectue localement sur les enregistrements du détenteur ; l'accès est régi par contrat entre les contreparties. La portée d'une défaillance est limitée à l'installation concernée. Nous affirmons l'engagement du cadre CAC envers les capacités de sécurité intrinsèque — sécurité des données, protection des informations personnelles, protection cryptographique, détection des attaques, contrôle d'accès, contrôle comportemental. Nous proposons, à titre de parallèle constructif, que pour une architecture fédérée, le lieu approprié de ces capacités soit l'installation souveraine, avec des mécanismes bilatéraux de coopération entre les installations lorsque les menaces traversent les frontières juridictionnelles ou organisationnelles. *(Parallèles avec le point 8 du CAC « capacités de sécurité intrinsèques »).* [RÉFÉRENCES : primitive d'application des limites du Tractatus (Stroh 2026, CC BY 4.0) ; Privacy Act 2020 (NZ) ; ISO/IEC 23894:2023 gestion des risques.]

Point 9. Attestation de la chaîne d'approvisionnement, partage fédéré. Nous proposons une attestation couvrant l'ensemble du cycle de vie par installation — provenance signée de la version, manifestes de dépendances, attestation des données d'entraînement le cas échéant, historique des réponses aux incidents de sécurité — publiée ouvertement par chaque installation. Les incidents liés à la chaîne d'approvisionnement sont partagés bilatéralement entre les pairs fédérés et via des canaux internationaux établis, notamment CERT-NZ, CERT-EU, US-CERT et le système de coordination CVE. Nous reconnaissons le mérite de l'engagement du cadre CAC en faveur de normes de sécurité couvrant l'ensemble du cycle de vie et du partage d'informations sur la chaîne d'approvisionnement. Nous proposons que, pour la coordination fédérée, la transparence de la chaîne d'approvisionnement soit assurée par la publication ouverte des attestations par chaque installation, avec une coopération bilatérale en matière de réponse aux incidents. *(Parallèle avec le point 9 du CAC « sécurité de la chaîne d'approvisionnement »).* [RÉFÉRENCES : ISO/IEC 23894:2023 gestion des risques ; procédures de divulgation du CERT-NZ ; processus international de coordination CVE ; ISO/IEC 42001:2023 systèmes de gestion.]

Point 10. Limiter le rayon d'action ; audit a posteriori. L'identification systématique des risques s'effectue localement au niveau de chaque installation, les incidents inter-installations se propageant à travers la fédération. La principale contribution du cadre à l'atténuation des risques d'attaques automatisées, des atteintes à la vie privée et de la diffusion de fausses informations consiste à limiter l'ampleur à laquelle les dommages automatisés s'aggravent. Nous réaffirmons l'engagement du cadre CAC en faveur de l'identification des risques, de l'alerte précoce, de l' intervention et de la prévention de l'utilisation de l'IA agentique dans des activités illégales (attaques automatisées, atteinte à la vie privée, génération et diffusion de fausses informations, fraude en ligne). Nous proposons, à titre de contribution architecturale complémentaire, que la limitation de l'ampleur des dommages automatisés — par le biais de limites opérationnelles par installation et d'une coopération bilatérale en matière de réponse aux incidents — constitue un complément structurel aux approches de détection et d'intervention au niveau centralisé. Le mécanisme structurel est **l'enveloppe de fédération**: par défaut, les enregistrements d'une installation restent au sein de cette installation, et les flux inter-installations ne se produisent que par le biais d'enveloppes que l'installation signe explicitement, comportant la provenance et la liaison au destinataire. La compromission d'une installation ne peut pas se propager silencieusement vers d'autres, car le substrat ne dispose d'aucun chemin de lecture implicite inter-installations — il n'existe aucun registre partagé par lequel un attaquant pourrait pivoter. La réponse bilatérale aux incidents s'applique alors à ce qui a été délibérément partagé, la provenance de l'enveloppe de fédération permettant une reconstruction forensic de la portée affectée sans nécessiter d'agrégateur d'audit centralisé. *(Parallèles avec le point 10 du CAC « atténuer les risques découlant des applications »).* [RÉFÉRENCES : primitive de délibération pluraliste du Tractatus (Stroh 2026, CC BY 4.0) ; ISO/IEC 23894:2023 gestion des risques ; Privacy Act 2020 (NZ) ; Harmful Digital Communications Act 2015 (NZ) — versions législatives actuelles à vérifier avant la publication de la v1.]

(III) Amélioration du système de gouvernance

Point 11. Gouvernance polycentrique, en dialogue avec des approches à plusieurs niveaux. L'autorité de gouvernance sur ce qu'un agent intelligent peut faire avec un enregistrement appartient au détenteur des enregistrements. La recevabilité d'un scénario est déterminée par installation par la propre juridiction du détenteur, avec le soutien des régulateurs sectoriels lorsque leur autorité s'étend au sujet concerné. Nous reconnaissons le bien-fondé de l' approche de gouvernance catégorisée et à plusieurs niveaux du cadre CAC pour les secteurs sensibles et les industries clés, l'Administration chinoise du cyberespace et les autorités industrielles compétentes déterminant les scénarios d'application autorisés et mettant en œuvre des mesures de gestion telles que le dépôt, les tests et le rappel de produits problématiques. Nous proposons, dans le contexte de l'Aotearoa Nouvelle , que la gouvernance polycentrique — avec de multiples centres d' autorité répartis entre les agences de la Couronne, les entités hapū / iwi, les régulateurs sectoriels, les organismes professionnels et les détenteurs des dossiers eux-mêmes — est bien adaptée au paysage institutionnel existant et aux obligations de partenariat découlant du Te Tiriti. La recherche internationale sur la gouvernance polycentrique, notamment les travaux fondateurs d'Elinor Ostrom, fournit le fondement théorique de cette approche. La polycentricité est soutenue sur le plan opérationnel par les propriétés du substrat : une seule **chaîne d'audit signée cryptographiquement** par enregistrement permet à plusieurs autorités — régulateur de la Couronne, entité hapū / iwi, organisme sectoriel, ordre professionnel, le détenteur de l'enregistrement lui-même — de vérifier chacune le même enregistrement dans le cadre de leur compétence respective, sans nécessiter de compilation centralisée ni de registres dupliqués. Les différentes autorités détiennent leurs propres copies de la chaîne d'audit sous leurs propres clés et rendent des jugements de conformité indépendants sur le même enregistrement sous-jacent. La primitive architecturale qui rend la gouvernance polycentrique opérationnellement gérable est la chaîne d'audit signée et répliquée au sein de

la fédération ; la question institutionnelle de savoir qui détient l'autorité sur quelle catégorie de décision reste politique. *(Parallèles avec le point 11 du CAC « gouvernance catégorisée et à plusieurs niveaux »).* [RÉFÉRENCES : Ostrom, E. (2010). Beyond Markets and States: Polycentric Governance of Complex Economic Systems. American Economic Review, 100(3), 641–672. https://doi.org/10.1257/aer.100.3.641 ; Charte des algorithmes pour l'Aotearoa Nouvelle-Zélande (2020) ; Te Kāhui Raraunga (kahuiraraunga.io — Modèle maori de gouvernance des données et cadre maori de gouvernance de l'IA) ; Taiuru, K. (20 sept. 2025) Critical Analysis of Te Mana Raraunga Data Principles, taiuru.co.nz/critical-analysis-mana-raraunga/.]

Point 12. Services de conformité fédérés. Les services de surveillance des risques, de test, d'évaluation, d'audit et de certification pour les agents intelligents existent sous forme d'offres commerciales, communautaires et universitaires ; la reconnaissance mutuelle entre les services s'effectue par le biais de la publication ouverte et de l'examen par les pairs. Nous reconnaissons le mérite de l'engagement du cadre CAC en faveur d'un système de services de conformité fournissant des services professionnels tels que la surveillance des risques, les tests et l'évaluation, le conseil et la certification, avec la promotion de la reconnaissance mutuelle entre les prestataires accrédités. **Ce domaine correspond au volet (ii) du comité unique proposé au §II, point 4. Le comité élaborerait des recommandations adaptées au contexte néo-zélandais concernant un cadre d'audit fédéré pour les agents intelligents, contribuerait aux travaux de l'ISO/IEC SC42 sur l'évaluation, l' appréciation et les systèmes de gestion de l'IA, et engagerait un dialogue bilatéral avec les auteurs du cadre CAC sur l'interaction entre les services de conformité fédérés et centralisés.** Les services de conformité se fédèrent concrètement comme suit : chaque installation publie ses propres attestations — provenance de la version, manifestes de dépendances, attestation des données d'entraînement le cas échéant, historique des réponses aux incidents, adhésion au cadre d'audit — sous sa propre identité cryptographique. Les prestataires de conformité vérifient ces attestations et publient leurs conclusions sous leur propre identité ; la reconnaissance mutuelle entre prestataires s'effectue par le biais de références croisées à des évaluations vérifiables cryptographiquement plutôt que par une accréditation centralisée. La primitive sous-jacente rendant cela opérationnel est la publication adressée par contenu avec provenance cryptographique — n'importe qui peut vérifier n'importe quelle évaluation de conformité par rapport à la version spécifique de l' installation qu'elle a effectivement évaluée. *(Parallèles avec le point 12 du CAC « système de services de conformité » ; le volet de travail consolidé sur la formation des comités s'applique.)* [RÉFÉRENCES : ISO/IEC 42001:2023 systèmes de management ; ISO/IEC 23894:2023 gestion des risques ; Royal Society Te Apārangi.]

(IV) Renforcement de la coordination fédérée

Point 13. Coordination par fédération. Les installations souveraines se fédèrent bilatéralement ; la coordination sur les préoccupations communes — normes d'interopérabilité, divulgation des incidents de sécurité, élaboration d'un cadre d'audit — s'effectue par le biais de publications ouvertes et d'un consensus entre pairs contributeurs. Nous reconnaissons le mérite de l'engagement du cadre CAC en faveur de l'autorégulation du secteur, les organisations professionnelles et les grandes entreprises formulant conjointement des règles d'autorégulation couvrant la conformité des fonctionnalités de l'IA, la gouvernance des algorithmes, la protection de la propriété intellectuelle et la concurrence loyale. Nous proposons, pour l'architecture fédérée spécifiée dans cette proposition, que la coordination sur les préoccupations communes s'effectue par le biais de publications ouvertes et d'un consensus entre les pairs contributeurs ; l'engagement architectural en faveur de la fédération bilatérale s'étend au mécanisme de coordination lui-même. Dans cette proposition, la fédération est bilatérale par nature : chaque installation publie un point de terminaison de fédération et choisit avec quels pairs elle s'associera, pour quelles

classes d'enregistrements spécifiques ; le format **d'enveloppe de fédération** précise quels enregistrements peuvent être transférés, avec quelle portée de consentement, vers quel destinataire, et avec quelles contraintes de non-transfert ultérieur. Les primitives sous-jacentes rendant la fédération bilatérale opérationnelle sont l'enveloppe de fédération (format de message lié au destinataire, associé à la provenance et à portée limitée), **la portabilité pilotée par les membres** (le détenteur des enregistrements peut exiger leur sortie de n'importe quelle installation vers n'importe quelle destination de son choix) et**la provenance cryptographique** (chaque enregistrement porte des métadonnées d'origine vérifiables qui survivent au transit). La coordination sur les préoccupations communes — normes d'interopérabilité, divulgation des incidents de sécurité, développement d'un cadre d'audit — se déroule de manière bilatérale entre pairs fédérés sans nécessiter de registre centralisé. *(Parallèles avec le point 13 du CAC « autorégulation du secteur »).* [RÉFÉRENCES : protocole de fédération ActivityPub (recommandation du W3C de 2018) ; processus « Request for Comments » de l'IETF ; document de procédure du W3C.]

Point 14. Réputation par attestation. Les installations souveraines publient leurs propres attestations — posture de sécurité, historique d'audit, manifestes de dépendances, réponse aux incidents — et les contreparties les vérifient cryptographiquement. La réputation s'acquiert au fil d'un historique de divulgation volontaire exacte vérifiée par des contreparties bilatérales. Nous reconnaissons le bien-fondé de la proposition du cadre CAC concernant des mécanismes volontaires de notation de crédit pour les entités du marché dans le secteur des agents intelligents, avec des évaluations de crédit portant sur des comportements tels que l'utilisation abusive de la technologie, l'incitation à la consommation, la publicité mensongère et la dissimulation d'informations sur les défauts, ainsi que des sanctions pour conduite malhonnête conformément aux lois et réglementations. **Ce domaine correspond au volet de travail (iii) du comité unique proposé au §II, point 4. Le comité élaborerait des recommandations adaptées au contexte néo-zélandais sur la réputation fondée sur l'attestation par opposition à la réputation fondée sur un registre, contribuerait aux travaux de normalisation internationale sur la provenance et l'attestation en matière d'IA, et engagerait un dialogue bilatéral avec les auteurs du cadre CAC sur l'interopérabilité entre les systèmes de réputation fondés sur l'attestation et ceux fondés sur la notation de crédit.** *(Parallèles avec le point 14 du CAC « mécanismes de notation de crédit » ; le volet de travail consolidé sur la formation de comités s'applique.)* [RÉFÉRENCES : Modèle de données pour les identifiants vérifiables du W3C v1.1 ; ISO/IEC 42001:2023 systèmes de gestion.]

---

§IV. Renforcer le développement axé sur l'adoption

Là où le cadre de l'Administration chinoise du cyberespace énumère dix-neuf secteurs dans lesquels l'État ordonne que « les agents doivent faire X », nous reprenons ces dix-neuf secteurs et les recadrons chacun comme une question de conditions de souveraineté pour tout déploiement d'agents dans ce secteur. Le cadre n'ordonne pas le déploiement ; il spécifie les conditions architecturales sous lesquelles le déploiement est compatible avec la souveraineté. Ce recadrage est rhétoriquement modeste mais structurellement lourd de conséquences : la lecture dictée par l'État positionne les agents intelligents comme des instruments de programmes sectoriels, tandis que la lecture en termes de conditions de souveraineté les positionne comme des outils dont l'utilisation doit satisfaire aux exigences d'attribution, de provenance et de portabilité des membres, quel que soit celui qui les déploie.

Les primitives architecturales invoquées dans les dix-neuf secteurs qui suivent sont au nombre de quatre. **La provenance cryptographique** attache à chaque enregistrement des métadonnées d'origine vérifiables — qui l'a rédigé, quand, en vertu de quelle autorité

— immuables face à toute modification a posteriori silencieuse (les corrections sont contresignées et elles-mêmes enregistrées). **Les enveloppes de fédération** servent d'intermédiaires au partage entre installations : seul le sous-ensemble autorisé est transféré, en conservant par défaut la provenance, la liaison au destinataire et l'interdiction de transmission ultérieure. **La portabilité pilotée par les membres** permet au détenteur des enregistrements d'exporter son ensemble vers une autre installation sans l'autorisation du détenteur d'origine, la provenance restant intacte à destination. **L'application des limites** achemine par défaut les décisions relevant des quatre catégories de limites (irréversibilité, chargées de valeurs, dépendantes du contexte culturel, sans précédent) vers la délibération humaine, l'acheminement lui-même étant enregistré. Les éléments sectoriels qui suivent désignent la manifestation spécifique au secteur d'une ou plusieurs de ces primitives ; les capacités génériques sont constantes d'un secteur à l'autre. Voir *Alignement architectural* §3 pour le développement des primitives ; *Article A* pour la couche de substrat dans son intégralité.

(I) Recherche scientifique

Point 15. Dans le domaine de la recherche, les principes de souveraineté s'appliquent. Les environnements de recherche exploitent des ensembles de données souverains — détenus par des particuliers, des institutions, des entités hapū / iwi ou des consortiums de recherche, dans le cadre de leurs dispositifs de gouvernance respectifs ; la traçabilité accompagne les résultats dérivés ; la fédération bilatérale entre institutions fournit la couche d'interopérabilité lorsque le partage des données est nécessaire. Nous reconnaissons le bien-fondé de la vision du cadre CAC selon laquelle des agents intelligents améliorent la déduction théorique, l'intégration des connaissances et l'intégration avec des instruments scientifiques et des plateformes expérimentales. Nous proposons que, pour la recherche en Aotearoa (Nouvelle-Zélande), ces capacités soient déployées dans le cadre d'une gouvernance de l'éthique de la recherche spécifique à chaque institution et à chaque projet de recherche, la primitive de délibération pluraliste Tractatus fournissant le mécanisme architectural permettant d'étendre l'examen de l'éthique de la recherche à travers des cadres de valeurs concurrents. Les primitives opérationnelles sont la provenance cryptographique (chaque ensemble de données et chaque résultat dérivé est accompagné d'une provenance attestant de ses sources, de ses dérivations et du cadre d'examen éthique dans lequel il a été produit) et les enveloppes de fédération (le partage interinstitutionnel s'effectue dans le cadre d'accords explicites de partage de données, l'enveloppe enregistrant quel sous-ensemble de données est transféré et dans quel cadre de consentement). La portabilité pilotée par les membres permet à un participant à la recherche de retirer sa contribution et de faire mettre à jour la provenance en aval ; l'application des limites renvoie les décisions éthiques chargées de valeurs au comité d'éthique de la recherche plutôt qu'à l'action d'agents autonomes. *(Parallèles avec le point 15 du CAC « recherche et exploration »).* [RÉFÉRENCES : Principes CARE (Carroll et al. 2020) ; Principes FAIR (Wilkinson et al. 2016, https://doi.org/10.1038/sdata.2016.18) ; Te Kāhui Raraunga (kahuiraraunga.io — Modèle maori de gouvernance des données et Cadre maori de gouvernance de l'IA) ; Taiuru, K. (20 sept. 2025) Analyse critique des principes de données Te Mana Raraunga, taiuru.co.nz/critical-analysis-mana-raraunga/ ; Cadre néo-zélandais d'éthique de la recherche via le Conseil de la recherche en santé et la Royal Society Te Apārangi ; Primitive de délibération pluraliste Tractatus (Stroh 2026).]

Point 16. En R&D logicielle, l'attribution et l'audit s'appliquent. Les agents de génération de code opèrent à partir de sources attribuées ; les œuvres dérivées conservent leur lignée ; les pipelines CI/CD vérifient l'attestation de construction et la provenance des dépendances. Nous reconnaissons le mérite de l' engagement du cadre CAC envers les agents intelligents de développement logiciel améliorant l'analyse des exigences, la

conception architecturale, la génération de code et les tests. Nous proposons que toutes ces capacités fonctionnent selon des exigences d'attribution et de provenance ; les contributions des agents au code, à la conception ou aux résultats de simulation sont attribuées à la fois à l'agent et à l' opérateur humain ou organisationnel sous l'autorité duquel elles ont été produites. La primitive opérationnelle est la provenance cryptographique appliquée à chaque artefact de code — l'agent qui l'a proposé, le réviseur humain qui l'a approuvé, le pipeline de compilation qui l'a compilé, les attestations propres à l'arborescence des dépendances — formant une chaîne vérifiable depuis la ligne d'écriture jusqu'à la validation autorisée. La primitive de validation par recoupement de Tractatus fournit une vérification à l'exécution garantissant que les actions de code proposées sont cohérentes avec l'historique canonique des instructions. *(Parallèle avec le point 16 du CAC « Soutien à la R&D ».)* [RÉFÉRENCES : Modèle de données des identifiants vérifiables du W3C v1.1 ; normes SBOM (Software Bill of Materials) via la NTIA et OWASP CycloneDX ; primitive de validation par recoupement de Tractatus (Stroh 2026).]

(II) Développement industriel

Point 17. Dans le secteur manufacturier, les primitives de souveraineté s'appliquent. Les données de production constituent le registre souverain du fabricant ; les agents agissant à son encontre sont identifiés ; la coordination inter-installations pour les chaînes d'approvisionnement est bilatérale. Nous reconnaissons le mérite de l'engagement du cadre CAC envers les agents de gestion de la production pour la planification, l'allocation des ressources et l'optimisation des processus, ainsi que l' intégration avec les machines-outils à commande numérique, les robots industriels et les chaînes de production automatisées. Nous proposons que toutes ces capacités opèrent sous l'autorité du fabricant, la coordination de la chaîne d'approvisionnement s'effectuant par le biais d'accords bilatéraux entre les fabricants participants et leurs contreparties. Les primitives opérationnelles sont la provenance cryptographique (chaque lot est accompagné d'une attestation de la chaîne de production — relevés de capteurs, décisions d'agents, approbations humaines — vérifiable a posteriori à partir de toute enquête sur un défaut) et les enveloppes de fédération (la coordination de la chaîne d'approvisionnement s'effectue par le biais d'enveloppes signées bilatéralement précisant quelles données de production sont partagées avec quelle contrepartie et à quelle fin). La portabilité pilotée par les membres se traduit ici par la capacité du fabricant à exporter l'intégralité de sa piste d'audit de production vers un autre organisme d'audit de la chaîne d'approvisionnement ou un régulateur sans l'autorisation du fournisseur de la plateforme d'origine . *(Parallèle avec le point 17 du CAC « fabrication intelligente ».)* [RÉFÉRENCES : ISO/IEC 42001:2023 systèmes de gestion ; recherche en cours concernant les normes néo-zélandaises relatives aux données de fabrication et les initiatives néo-zélandaises liées à l'Industrie 4.0.]

Point 18. Dans le domaine de l'énergie et des ressources, les principes de souveraineté s'appliquent. Les données environnementales, les catalogues de ressources et les registres de distribution constituent des documents souverains des entités responsables : la Couronne pour certains (ressources statutaires, certaines données environnementales) ; les hapū et les iwi pour ceux où s'appliquent les allocations au titre du règlement du traité ; les entités privées pour le reste. Les agents agissent sur la base des documents de l'entité concernée sous l'autorité de cette dernière. Les attributions spécifiques sont propres à chaque entité et dépendent de la législation et des dispositions pertinentes en matière de règlement. Nous reconnaissons le bien-fondé de l'engagement du cadre CAC en faveur des agents de détection environnementale pour l'alerte précoce en cas de catastrophes naturelles et de risques de pollution, des agents de répartition de l'énergie et de maintenance du réseau, ainsi que des applications d'exploration des ressources. Nous proposons que, dans le contexte de l'Aotearoa NZ, les autorités compétentes découlent du cadre institutionnel et du cadre des traités existants, et que l'architecture fournisse l'infrastructure d'audit et

d’attribution au sein de laquelle ces autorités opèrent. Une provenance cryptographique est associée aux relevés environnementaux, aux décisions de répartition du réseau et aux choix d’allocation des ressources, avec attribution à l’ entité responsable (Couronne, iwi ou entité privée). Les enveloppes de fédération ne transportent que le sous-ensemble approuvé de données environnementales au-delà des frontières des entités — les signaux d'alerte précoce se propagent à toutes les entités concernées sans nécessiter d'agrégation centrale. Lorsque les allocations au titre des accords de règlement du Traité s'appliquent, l'entité iwi détient sa propre chaîne d'audit sous ses propres clés, indépendamment des systèmes des agences de la Couronne. *(Parallèles avec le point 18 du CAC « énergie et ressources »).* [RÉFÉRENCES : Loi de 1991 sur la gestion des ressources (NZ) ; législation pertinente relative aux accords de règlement des traités (spécifique à chaque entité, en attente de vérification avant la publication de la v1) ; Loi de 2010 sur l'industrie électrique (NZ) ; Loi de 1991 sur les minéraux de la Couronne (NZ).]

Point 19. Dans le domaine des transports, les principes de souveraineté s’appliquent. Les données de télémétrie des véhicules, les données de circulation et les données des capteurs d’infrastructure constituent des enregistrements souverains des opérateurs, des agences de la Couronne et des autorités de contrôle routier ; la coordination entre eux — l’Agence des transports de Nouvelle-Zélande Waka Kotahi, KiwiRail, les autorités maritimes, l’Autorité de l’aviation civile, les conseils régionaux et les conseils municipaux — est une fédération bilatérale transcendant les frontières institutionnelles concernées. Nous reconnaissons le mérite de l’engagement du cadre CAC en faveur des agents intelligents chargés de la sécurité routière, des interventions d’urgence et du contrôle des véhicules. Nous proposons que le contexte néo-zélandais (Aotearoa), avec ses arrangements institutionnels bilatéraux existants entre les modes de transport, soit bien adapté à une approche fédérée. Les enveloppes de fédération précisent quelles données de télémétrie, de trafic et de capteurs d’infrastructure sont partagées au-delà des frontières institutionnelles (Waka Kotahi ↔ conseils régionaux ↔ KiwiRail ↔ Autorité de l’aviation civile), la provenance cryptographique garantissant qu’une reconstitution judiciaire de tout incident est possible au niveau des décisions individuelles des agents. La portabilité pilotée par les membres s’applique au niveau du véhicule et de l’opérateur — l’ opérateur peut exiger de quitter n’importe quelle plateforme sans être lié par un engagement. *(Parallèles avec le point 19 du CAC « transport »).* [RÉFÉRENCES : Loi de 1998 sur les transports terrestres (NZ) ; Loi de 2003 sur la gestion des transports terrestres (NZ) ; Loi de 1990 sur l'aviation civile (NZ) ; Loi de 1994 sur les transports maritimes (NZ) ; recherche en cours concernant les travaux sur la souveraineté des données de transport en NZ.]

Point 20. En agriculture, les principes fondamentaux de la souveraineté s'appliquent. Les données agricoles constituent le registre souverain de l'agriculteur ; les données relatives aux ravageurs, aux maladies, aux rendements et au cheptel peuvent être partagées de manière bilatérale avec les services de vulgarisation, les instituts de recherche ou les hapū rōpū, le cas échéant, selon les conditions fixées par l'agriculteur. Nous reconnaissons le bien-fondé de l’ engagement du cadre CAC en faveur d’agents intelligents au service de l’agriculture pour les conseils techniques, le diagnostic des ravageurs et des maladies, et l’intégration avec des machines agricoles intelligentes et des serres. Nous proposons que pour Aotearoa NZ — où la souveraineté des données agricoles est une question reconnue par les coopératives de données agricoles, les organisations sectorielles et l’engagement croissant en faveur de la souveraineté des données maories dans le contexte des industries primaires — le partage bilatéral des données selon les conditions fixées par l’agriculteur soit tout à fait adapté. Une provenance cryptographique est associée aux données agricoles — relevés de capteurs, recommandations d’agents, décisions de traitement, résultats de rendement. Les enveloppes de fédération ne contiennent que le sous-ensemble autorisé (données sur les ravageurs et les maladies vers les services de vulgarisation ; données agrégées sur les rendements vers les instituts de recherche ; données dépendantes du contexte culturel vers les hapū rōpū selon les tikanga appropriés) selon les conditions fixées

par l’agriculteur. La portabilité pilotée par les membres permet à l’agriculteur de passer d’une coopérative de données agricoles à une autre sans perdre son historique d’audit . *(Parallèles avec le point 20 du CAC « production agricole ».)* [RÉFÉRENCES : recherche en cours concernant les travaux sur la souveraineté des données agricoles en Nouvelle-Zélande et les dispositifs de gouvernance des données agricoles ; Te Kāhui Raraunga (kahuiraraunga.io — Modèle maori de gouvernance des données et Cadre maori de gouvernance de l’IA) ; Taiuru, K. (20 septembre 2025) Analyse critique des principes de données Te Mana Raraunga, taiuru.co.nz/critical-analysis-mana-raraunga/ le cas échéant.]

Point 21. Dans les services financiers, les principes fondamentaux de souveraineté s'appliquent. Les dossiers clients, les données transactionnelles et les signaux de risque constituent des enregistrements souverains de l'établissement détenteur, soumis aux exigences prudentielles de la Banque de réserve de Nouvelle-Zélande / Te Pūtea Matua, à la loi sur la protection de la vie privée de 2020 et à la loi de 2009 sur la lutte contre le blanchiment d'argent et le financement du terrorisme. La coopération en matière de LBC/FT est bilatérale via des canaux établis — la Cellule de renseignement financier de Nouvelle-Zélande et les canaux internationaux du GAFI — et l’assistance par IA est attribuée et délimitée par ces accords réglementaires existants. Nous reconnaissons le mérite de l’engagement du cadre CAC envers les agents de contrôle des risques financiers pour l’approbation des crédits, la surveillance des transactions, la sécurité des comptes et la surveillance en matière de lutte contre le blanchiment d’argent. Nous proposons que, pour l’Aotearoa NZ, le cadre institutionnel et réglementaire existant soit bien adapté à un audit basé sur l’attribution au niveau de chaque institution financière, avec une coopération bilatérale par le biais de canaux établis pour la coordination interinstitutionnelle et internationale. Une provenance cryptographique est associée à chaque transaction, évaluation par IA et approbation humaine — vérifiable a posteriori par les auditeurs AML/CFT, la Banque de réserve, des inspecteurs du GAFI et des clients eux-mêmes dans le cadre de leurs compétences respectives. Les enveloppes de fédération facilitent la coopération interinstitutionnelle en matière de LBC/FT : seul le signal d’activité suspecte approuvé est transmis, le destinataire étant lié à la Cellule de renseignement financier de Nouvelle-Zélande ou à la contrepartie concernée. La portabilité pilotée par les membres soutient les obligations de portabilité des comptes : l’historique des transactions du client est exportable vers une autre institution avec une provenance intacte. *(Parallèles avec le point 21 du CAC « services financiers »).* [RÉFÉRENCES : Loi de 2021 sur la Banque de réserve de Nouvelle-Zélande ; Loi de 2009 sur la lutte contre le blanchiment d'argent et le financement du terrorisme (NZ) ; Loi de 2020 sur la protection de la vie privée (NZ) ; Recommandations du GAFI.]

(III) Vie quotidienne

Point 22. Dans les applications destinées aux utilisateurs finaux, les primitives de souveraineté s'appliquent. Des identifiants portables par le membre remplacent les comptes spécifiques à une plateforme ; la coordination entre appareils est assurée par le trousseau de clés ou le portefeuille d'identité du membre lui-même. La souveraineté signifie ici que l'utilisateur détient les enregistrements — que l'application soit développée par un fournisseur néo-zélandais ou international. Nous reconnaissons le mérite de l'engagement du cadre CAC en faveur d'agents intelligents qui renforcent les applications et services Internet sur les téléphones mobiles, les ordinateurs, les véhicules, les appareils ménagers, les appareils portables et les robots grand public. Nous proposons que pour toute application fonctionnant à partir des enregistrements des utilisateurs, les primitives architecturales d'attribution et de portabilité des membres s'appliquent quelle que soit la juridiction du fournisseur . La primitive opérationnelle est la portabilité pilotée par les membres, mise en œuvre via les identifiants décentralisés et les informations d'identification vérifiables du W3C : l'identité de l'utilisateur est conservée dans son propre trousseau (ou portefeuille),

la coordination entre appareils étant assurée par ses propres clés. Une provenance cryptographique est associée à chaque action effectuée par un agent sur les données de l'utilisateur, attribuable à l'agent et au cadre d'autorisation de l'utilisateur. Les enveloppes de fédération assurent la coordination entre fournisseurs uniquement lorsque l'utilisateur l'autorise *. (Parallèles avec le point 22 du CAC « applications destinées à l'utilisateur final »).* [RÉFÉRENCES : Identifiants décentralisés (DID) du W3C v1.0 ; Modèle de données des identifiants vérifiables du W3C v1.1 ; Loi sur la protection de la vie privée de 2020 (NZ), principe de confidentialité des informations n° 7 (correction).]

Point 23. Dans les domaines de la culture et du tourisme, les principes fondamentaux de la souveraineté s'appliquent. Les contenus culturels relèvent de la responsabilité de leurs créateurs ; dans le contexte d'Aotearoa (Nouvelle-Zélande), les obligations des kaitiaki envers les taonga sont au cœur de la manière dont les agents IA peuvent interagir avec le matériel culturel. Les agents de traduction préservent l'attribution et le contexte culturel ; leurs productions ne se substituent pas au mātauranga original, et c'est aux tangata whenua qu'il appartient de déterminer ce qui constitue une utilisation appropriée dans les contextes te ao Māori. Les données des visiteurs traitées par les services touristiques sont considérées comme le dossier souverain du visiteur. Nous reconnaissons le mérite de l'engagement du cadre CAC envers les agents de création de contenu culturel et les agents des services touristiques. Nous proposons que pour l'Aotearoa NZ — où le mātauranga Māori est un taonga en vertu des obligations de partenariat du Te Tiriti, et où les travaux publiés par le Dr Karaitiana Taiuru sur la protection du mātauranga Māori dans les données d'entraînement de l'IA, ainsi que le Raraunga, constitue une recherche fondamentale — les primitives architecturales fournissent l'infrastructure d'audit, et la détermination de fond de l'usage approprié revient aux détenteurs du mātauranga. La provenance cryptographique est attachée au matériel culturel : qui l'a créé, sous quelle autorité, avec quelle portée d'utilisation. Pour le mātauranga, les tikanga des détenteurs déterminent ce que signifie concrètement le pouvoir de contrôle ; le substrat fournit l'infrastructure d'audit afin que toute violation du consentement soit reconstituable de manière légale, et non simplement contestable sur le plan contractuel. Les enveloppes de fédération ne transportent que le sous-ensemble du mātauranga faisant l'objet d'un consentement au-delà des limites de l'installation, avec une non-transmission par défaut — les agents de traduction héritent mais ne peuvent pas réattribuer de licence. Les données des visiteurs contiennent l'attestation d'identité du visiteur ; la portabilité pilotée par les membres signifie que le visiteur exporte son dossier de données touristiques au moment de son départ *. (Parallèles avec le point 23 du CAC « culture et tourisme »).* [RÉFÉRENCES : Taiuru, K. — Protection du mātauranga Māori dans les données d'entraînement de l'IA (publications spécifiques en attente de vérification) ; Te Kāhui Raraunga (kahuiraraunga.io — Modèle de gouvernance des données Māori et Cadre de gouvernance de l'IA Māori) ; Taiuru, K. (20 sept. 2025) Analyse critique des principes de données Te Mana Raraunga, taiuru.co.nz/critical-analysis-mana-raraunga/ ; Principes CARE (Carroll et al. 2020) ; Wai 262 (Rapport du Tribunal de Waitangi sur la flore et la faune autochtones et la propriété intellectuelle culturelle).]

Point 24. Dans les services commerciaux, les principes de souveraineté s'appliquent. Les interactions avec les clients génèrent des enregistrements ; les deux parties — l'opérateur et le client — détiennent des copies de provenance ; les litiges sont coordonnés bilatéralement. Les agents incarnés dans les secteurs de la vente au détail, de l'hôtellerie, des soins aux personnes âgées et de l'aide aux personnes handicapées opèrent sous l'autorité du responsable du déploiement et produisent des enregistrements vérifiables de leurs actions. Nous reconnaissons le bien-fondé de l' engagement du cadre CAC en faveur d'un service client 24 h/24 et 7 j/7, d' agents intelligents incarnés pour l'orientation, le nettoyage, l'entreposage et la distribution dans les lieux commerciaux, ainsi que d'agents incarnés pour l'aide à domicile, les soins aux personnes âgées, la garde d'enfants et l'aide aux personnes handicapées. Nous proposons que, pour l'Aotearoa NZ, toutes ces applications fonctionnent

dans le cadre des réglementations existantes en matière de protection des consommateurs, de qualité des soins et de services aux personnes handicapées. Les deux parties (opérateur et client) détiennent des copies de provenance signées cryptographiquement de chaque interaction, de sorte que les litiges puissent être résolus sur la base d'un enregistrement vérifiable partagé plutôt que sur la base de journaux de plateforme unilatéraux. Les agents incarnés dans le commerce de détail, l'hôtellerie, les soins aux personnes âgées et l'aide aux personnes handicapées opèrent dans le cadre d'une application des limites : les décisions dépendantes du contexte culturel ou chargées de valeurs (dérogations à l'administration de médicaments, modifications du plan de soins, escalades de l'aide aux personnes handicapées) sont par défaut soumises à la délibération humaine. L'enveloppe fédérative assure le transfert des données de type dossier de soins entre les prestataires. *(Parallèles avec le point 24 du CAC « services commerciaux »).* [RÉFÉRENCES : Loi de 1993 sur les garanties des consommateurs (NZ) ; Loi de 1986 sur le commerce équitable (NZ) ; Loi de 2001 sur les services de santé et de soutien aux personnes handicapées (sécurité) (NZ) ; Stratégie néo-zélandaise en faveur des personnes handicapées.]

(IV) Bien-être public

Point 25. Dans le domaine de l'éducation, les principes de souveraineté s'appliquent. Les dossiers d'apprentissage constituent le dossier souverain de l'étudiant, avec une cogestion lorsque l'étudiant est mineur ; les supports pédagogiques produits par des agents sont attribués ; les dossiers institutionnels — listes d'étudiants, évaluations, dossiers de qualification — relèvent de la gouvernance institutionnelle existante en vertu de la loi de 2020 sur l'éducation et la formation. La portabilité revient à l'étudiant, avec des dispositions institutionnelles appropriées pour le transfert lors des transitions entre prestataires. Nous reconnaissons le mérite de l'engagement du cadre CAC en faveur de la production de supports de cours, de la correction des devoirs, de l'analyse des progrès d'apprentissage , des plans d'apprentissage personnalisés et des assistants pédagogiques virtuels. Nous proposons que, pour Aotearoa NZ, ces capacités fonctionnent dans le cadre de la loi de 2020 sur la protection de la vie privée et de la loi de 2020 sur l'éducation et la formation, avec le maintien de la souveraineté des données des étudiants tout au long du processus. Une provenance cryptographique est associée à chaque évaluation, à chaque support pédagogique produit par un agent, à chaque délivrance de qualification — attribuable à l'agent, à l'éducateur superviseur et à l'établissement. Le dossier souverain de l'élève est transférable : à chaque transition (d'une école à une autre, d'une école à une université, entre prestataires, entre pays), l'élève emporte l'intégralité de son dossier, avec sa provenance intacte, vers l'établissement d'accueil. L'application des limites renvoie par défaut les décisions modifiant l'évaluation et affectant les diplômes à la délibération humaine. *(Parallèles avec le point 25 du CAC « éducation et enseignement »).* [RÉFÉRENCES : Loi de 2020 sur l'éducation et la formation (NZ) ; Loi de 2020 sur la protection de la vie privée (NZ) ; Programme scolaire néo-zélandais.]

Point 26. Dans le domaine des soins de santé, les principes de souveraineté s'appliquent. Les dossiers des patients constituent les dossiers souverains des patients en vertu du Code de confidentialité des informations de santé de 2020 et des structures de gestion de Te Whatu Ora / Health New Zealand ; les agents de diagnostic produisent des résultats attribués ; les recommandations de traitement comportent une traçabilité ; la coordination entre les prestataires s'effectue via les canaux de l'Organisation des normes d'information de santé (HISO) et les accords d'interopérabilité de Te Whatu Ora. Nous reconnaissons le mérite de l'engagement du cadre CAC en faveur de l' analyse d'imagerie médicale, du raisonnement pour le diagnostic des maladies, des plans de traitement personnalisés, de la gestion des médicaments, de la planification chirurgicale et des agents de gestion des dossiers médicaux. Nous proposons que, pour Aotearoa NZ, ces capacités fonctionnent dans le cadre de la gouvernance existante des informations de

santé, la souveraineté du patient sur ses dossiers de santé étant maintenue comme principe architectural de base. La provenance cryptographique est associée aux dossiers cliniques, aux résultats de diagnostic par IA, aux recommandations de traitement et à l'administration des médicaments — chaque entrée étant attribuable à son auteur et immuable face à toute modification a posteriori non signalée (les corrections sont des amendements contresignés, eux-mêmes enregistrés). Les enveloppes de fédération transportent les orientations : seul le sous-ensemble clinique autorisé est transmis du prestataire d'origine au prestataire destinataire, avec une obligation de confidentialité et une interdiction de transmission ultérieure par défaut pour le destinataire. La portabilité à l'initiative du membre permet au patient d'exporter son dossier vers un autre prestataire — public, privé ou international — sans l'autorisation du détenteur d'origine, la provenance restant intacte à destination. L'application des limites renvoie les décisions cliniquement incertaines et chargées de valeurs (fin de vie, diagnostics contestés, capacité en matière de santé mentale) à la délibération humaine plutôt qu'à l'action d'un agent autonome. *(Parallèles avec le point 26 du CAC « soins de santé »).* [RÉFÉRENCES : Code de confidentialité des informations de santé 2020 (NZ) ; Loi Pae Ora (Healthy Futures) 2022 (NZ) ; normes de données HISO.]

Point 27. En matière d'emploi et de travail, les principes de souveraineté s'appliquent. Les dossiers d'emploi, les certificats de formation et les dossiers de litige relèvent de la souveraineté des parties ; la médiation s'exerce dans le cadre de la gouvernance existante du Service de médiation en matière d'emploi ; l'assistance par IA est attribuée et délimitée par la structure tripartite existante (travailleur / employeur / État) du droit du travail néo-zélandais. Nous reconnaissons le mérite de l'engagement du cadre CAC envers les agents chargés de la promotion de l'emploi, de la formation et de l'évaluation du personnel technique, des services de relations de travail, de l'assurance sociale, de l'arbitrage des conflits du travail et de la gestion des arriérés de salaire. Nous proposons que pour Aotearoa NZ, ces capacités fonctionnent dans le cadre de la loi de 2000 sur les relations de travail et du cadre tripartite associé, avec l'attribution et la provenance appliquées de bout en bout. La provenance cryptographique est associée aux dossiers d'emploi, aux certificats de formation et aux dossiers de litige — attribuables aux parties concernées. Les enveloppes de la fédération facilitent la portabilité des certificats de formation des travailleurs d'un employeur à l'autre sans perte de provenance. L'application des limites renvoie les décisions d'embauche, de licenciement, disciplinaires et de médiation des litiges à l'autorité humaine — toute action d'un agent autonome contre le statut professionnel d'un travailleur individuel est structurellement bloquée. *(Parallèles avec le point 27 du CAC « ressources humaines ».)* [RÉFÉRENCES : Loi de 2000 sur les relations de travail (NZ) ; Loi de 2003 sur les congés (NZ) ; Loi de 1993 sur les droits de l'homme (NZ) ; Cadre tripartite néo-zélandais des relations de travail.]

Point 28. Dans les services d'information, les principes de souveraineté s'appliquent. Le contenu est attribué à ses créateurs ; les agents de recommandation opèrent en fonction du profil souverain de l'utilisateur, que celui-ci peut consulter, exporter et transférer ; la révision éditoriale reste une fonction humaine. Lorsque des agents IA produisent du contenu, l'attribution revient à l'agent et à l' opérateur humain ou organisationnel sous l'autorité duquel il a agi ; la divulgation du contenu généré par l'IA est l'engagement architectural de base. Nous reconnaissons le mérite de l'engagement du cadre CAC en faveur des agents intelligents pour la construction de contenu en ligne, l'analyse des utilisateurs, la planification des sujets, le traitement éditorial, la distribution et la recommandation, la révision de contenu, l'orientation des opinions, le soutien émotionnel et la traduction en temps réel. Nous proposons que, pour Aotearoa NZ, les exigences d'attribution s'appliquent à toutes ces applications, les normes de radiodiffusion existantes et le cadre relatif aux communications numériques préjudiciables fournissant le contexte réglementaire. Une provenance cryptographique est associée à chaque élément de contenu : identité de l'auteur, attribution IA vs humain, chaîne de révision éditoriale. Les enveloppes de fédération transportent les décisions de distribution : chaque interface de recommandation ne reçoit que

le contenu que l'installation en amont a signé et accepté de partager, avec une interdiction de retransmission par défaut. La portabilité pilotée par les membres fournit à l'utilisateur son historique d'interaction et son profil de recommandation sous une forme portable — il peut passer à un autre service sans perdre son historique de contenu. L'application des limites renvoie les décisions éditoriales à l'autorité humaine — l'action d'un agent autonome sur ce qu'il convient d'amplifier ou de supprimer est structurellement bloquée. *(Parallèles avec le point 28 du CAC « services d'information »).* [RÉFÉRENCES : Loi sur la radiodiffusion de 1989 (NZ) ; Loi sur les communications numériques préjudiciables de 2015 (NZ) ; Loi sur la protection de la vie privée de 2020 (NZ) ; recherche en cours concernant les normes d'attribution des contenus générés par l'IA.]

(V) Gouvernance sociale

Point 29. Dans l'administration publique, les principes fondamentaux de la souveraineté s'appliquent. Les interactions des citoyens avec l'État donnent lieu à des dossiers détenus à la fois par les citoyens et par l'État ; les identifiants détenus par les membres migrent au fil du temps vers le contrôle des membres ; l'assistance des agents dans les processus d'approbation est attribuée et encadrée par les principes du droit administratif. Les agences de la Couronne restent responsables en vertu de la loi de 2020 sur la fonction publique, de la loi de 1982 sur l'information officielle, de la loi de 2020 sur la protection de la vie privée, de la Charte des algorithmes pour Aotearoa Nouvelle-Zélande et de la loi de 2005 sur les archives publiques. Nous reconnaissons le mérite de l'engagement du cadre CAC en faveur des agents chargés de l'approbation administrative, de la consultation politique et de la prestation proactive de services. Nous proposons que, pour l'Aotearoa NZ, toutes ces applications des agences de la Couronne fonctionnent dans le cadre du cadre de responsabilité existant, les primitives architecturales fournissant l'infrastructure d'audit conforme aux engagements de la Charte des algorithmes en matière de transparence et de partenariat avec les Māori. Une provenance cryptographique est associée à chaque action administrative : qui a décidé, en vertu de quelle autorité, sur la base de quel dossier citoyen, avec l'aide de quel agent. Des enveloppes fédératives assurent la coordination inter-agences — seul le sous-ensemble de dossiers citoyens ayant fait l'objet d'un consentement circule d'une agence à l'autre, avec une piste d'audit. La portabilité pilotée par les membres met en œuvre le principe 6 de confidentialité des informations de la Loi sur la protection de la vie privée (droits d'accès) et les obligations de conservation de la Loi de 2005 sur les documents publics de 2005 — le citoyen peut exporter l'intégralité de son dossier d'interaction avec l'administration à l'aide de ses propres clés. L'application des limites renvoie les décisions relevant du pouvoir discrétionnaire administratif, les décisions liées aux obligations découlant du Traité et les décisions affectant les droits à l'autorité humaine. *(Parallèles avec le point 29 du CAC « services de l'administration publique »).* [RÉFÉRENCES : Loi sur la fonction publique de 2020 (NZ) ; Loi sur l'information officielle de 1982 (NZ) ; Loi sur la protection de la vie privée de 2020 (NZ) ; Charte des algorithmes pour l'Aotearoa Nouvelle-Zélande (2020) ; Loi sur les archives publiques de 2005 (NZ).]

Point 30. Dans les services judiciaires, les principes de souveraineté s'appliquent. Les dossiers judiciaires, les preuves et les documents juridiques sont régis par les procédures judiciaires existantes ; l'assistance par l'IA est attribuée ; la chaîne de conservation des preuves est cryptographique le cas échéant ; les contrôles d'accès suivent la gouvernance judiciaire existante. Les outils d'aide aux justiciables non représentés qui utilisent l'IA divulguent leur utilisation et produisent une provenance vérifiable. Nous reconnaissons le mérite de l'engagement du cadre CAC en faveur de l'assistance de bout en bout dans le traitement des dossiers, de la génération de documents juridiques, de la publicité juridique, de la consultation juridique et des agents de supervision juridique. Nous proposons que, pour l'Aotearoa NZ, toutes ces applications fonctionnent en vertu de la loi de 2016 sur

les hautes cours, de la loi de 2006 sur la preuve, ainsi que des règles judiciaires et notes pratiques établies régissant l'utilisation de l'IA dans les procédures judiciaires. Une traçabilité cryptographique est associée à chaque élément de preuve présenté, à chaque projet de document juridique assisté par l'IA, à chaque résultat de recherche. La chaîne de conservation des preuves est cryptographiquement ancrée — les questions d'admissibilité peuvent être tranchées à partir du substrat plutôt que des journaux contestés de la plateforme. La portabilité pilotée par les membres signifie que le justiciable se représentant lui-même peut transférer l'intégralité de son dossier entre les instances (tribunal → cour → appel) en conservant intacte la traçabilité. L'application des limites renvoie les jugements de valeur et les décisions discrétionnaires (qu'il s'agisse de plaider, d'accepter un règlement, ou de prendre une mesure préjudiciable au justiciable) au justiciable lui-même ou à son conseil — toute action d'un agent autonome contre la position juridique du justiciable est structurellement bloquée. *(Parallèles avec le point 30 du CAC « services judiciaires »).* [RÉFÉRENCES : Senior Courts Act 2016 (NZ) ; Evidence Act 2006 (NZ) ; recherche en cours concernant les directives judiciaires actuelles sur l'utilisation de l'IA.]

Point 31. En matière de sécurité publique, les principes fondamentaux de souveraineté s'appliquent. La surveillance est régie par la législation en vigueur — la loi de 2020 sur la protection de la vie privée, la loi de 2012 sur les perquisitions et la surveillance, et la loi de 2017 sur le renseignement et la sécurité — et tout agent d'IA opérant dans des contextes de sécurité publique génère une traçabilité vérifiable dans le cadre de ces dispositifs. Les agents de surveillance du comportement opèrent dans le cadre déjà légal en vertu de ces lois. Nous reconnaissons le bien-fondé de l’engagement du cadre CAC en faveur des agents de surveillance et d’alerte précoce, des agents d’intervention d’urgence et de coordination des secours, ainsi que des applications d’identification des comportements anormaux et de prévention dynamique. Nous proposons, dans le contexte de l'Aotearoa NZ, que la contribution architecturale de l'attribution et de la provenance consiste à rendre l'IA agentique dans les contextes de sécurité publique vérifiable ; la question de savoir si et comment ces capacités doivent être déployées relève d'une décision de fond pour le cadre législatif et politique pertinent, qui s'adresse au Parlement et aux ministres responsables, l'architecture fournissant l'infrastructure d'audit au sein de laquelle ces décisions deviennent gérables. La provenance cryptographique est associée à chaque signal de surveillance, à chaque décision de surveillance des comportements, à chaque action d’intervention d’urgence — rendant la responsabilité a posteriori gérable d’une manière que l’action d’un agent non accompagné ne l’est pas. Les enveloppes de fédération assurent la coordination interinstitutionnelle en matière de sécurité publique : les renseignements qui circulent entre la police, le Bureau des communications du gouvernement et les agences d’intervention d’urgence est régie par ce que chacun a approuvé comme pouvant être divulgué à des fins spécifiques. Le respect des limites est ici déterminant : les décisions engageant des droits (perquisition, arrestation, autorisation de surveillance, recours à la force) sont acheminées vers une autorité humaine — le fondement architectural en dessous duquel aucun agent n’agit de manière autonome. *(Parallèles avec le point 31 du CAC « sécurité publique »).* [RÉFÉRENCES : Loi sur la protection de la vie privée de 2020 (NZ) ; Loi sur les perquisitions et la surveillance de 2012 (NZ) ; Loi sur le renseignement et la sécurité de 2017 (NZ) ; Loi sur la Charte des droits de la Nouvelle-Zélande de 1990.]

Point 32. En matière de gouvernance urbaine, les principes fondamentaux de la souveraineté s'appliquent. Les données urbaines — réseaux de capteurs, données d'urbanisme, permis de construire, données d'exploitation des infrastructures — sont détenues par les conseils municipaux en tant que documents souverains ; les systèmes agentiels opérant dans le cadre des fonctions des conseils municipaux sont attribués et responsables en vertu de la Loi de 2002 sur les collectivités locales et des structures de gouvernance des conseils municipaux. Nous reconnaissons le mérite de l’engagement du cadre CAC en faveur des agents intelligents dans l’urbanisme, la construction urbaine et la

gouvernance urbaine, y compris pour la construction intelligente, la gestion des bâtiments et l'exploitation des infrastructures urbaines. Nous proposons que, pour Aotearoa NZ, toutes ces applications fonctionnent dans le cadre des dispositifs existants de responsabilité des collectivités locales, les primitives architecturales fournissant l'infrastructure d'audit et d'attribution. Une provenance cryptographique est associée à chaque lecture de capteur, décision d'urbanisme, permis de construire et choix d'exploitation des infrastructures — vérifiable par les résidents, par Local Government NZ, par le vérificateur général et par les conseils municipaux successifs. Des enveloppes fédératives assurent la coordination entre les conseils et le partage des données entre le niveau central et le niveau local — seul le sous-ensemble autorisé est transféré. La portabilité pilotée par les membres s'applique aux données au niveau des résidents : un résident peut conserver ses interactions avec le conseil lorsqu'il change de district. L'application des limites renvoie les décisions liées au Traité et culturellement significatives (gestion des urupā, wāhi tapu et taonga) vers la délibération appropriée des tangata-whenua plutôt que vers l'action d'agents autonomes. *(Parallèles avec le point 32 du CAC « gouvernance urbaine »).* [RÉFÉRENCES : Loi sur les collectivités locales de 2002 (NZ) ; Loi sur la construction de 2004 (NZ) ; Loi sur la gestion des ressources de 1991 (NZ).]

Point 33. En matière de marchés publics, les principes fondamentaux de souveraineté s'appliquent. Les dossiers d'appel d'offres, les évaluations et les contrats sont des documents souverains de l'entité contractante ; l'assistance des agents dans les marchés publics est attribuée et encadrée par les règles relatives aux marchés publics et le droit des contrats applicable ; la transparence est assurée par la publication existante conforme à l'OIA. Nous reconnaissons le mérite de l'engagement du cadre CAC en faveur d'une gestion intelligente de bout en bout des processus d'appel d'offres et de soumission, avec une intelligence appliquée aux transactions, aux services et à la supervision. Nous proposons que pour Aotearoa NZ, les règles relatives aux marchés publics et le cadre existant des marchés publics fournissent le contexte de responsabilité approprié, avec une attribution et une provenance appliquées tout au long du processus. La provenance cryptographique est associée à chaque dossier d'appel d'offres, évaluation, modification de contrat et décision d'attribution — publiés dans le cadre des dispositions de transparence existantes conformes à l'OIA, avec des propriétés d'intégrité au niveau du substrat. Les enveloppes fédérées assurent la coordination du consortium et de la chaîne d'approvisionnement requise par les marchés publics, sans exposer de données sensibles sur le plan concurrentiel en dehors du champ d'application convenu. L'application des limites renvoie les décisions discrétionnaires en matière de marchés publics (attribution, dérogation, exception) à une autorité humaine — toute action d'un agent autonome concernant l'attribution d'un contrat est structurellement bloquée . *(Parallèles avec le point 33 du CAC « appels d'offres et soumissions »).* [RÉFÉRENCES : Règles relatives aux marchés publics (NZ) ; Loi de 2005 sur les documents publics (NZ) ; Loi de 1982 sur l'information officielle (NZ).]

---

§V. Construire un écosystème fédéré

Alors que le cadre de l'Administration chinoise du cyberespace envisage un écosystème de pôles industriels visant à promouvoir des champions nationaux par le biais de conférences internationales sur l'IA, nous proposons un écosystème fédéré où la coordination s'effectue par le biais d'une fédération bilatérale entre pairs souverains et où l'alignement international passe par des organismes de normalisation établis. Les deux sous-sections qui suivent — promouvoir la coopération fédérée et renforcer la promotion bilatérale — précisent ensemble comment un écosystème d' installations souveraines s'autoalimente et s'engage au niveau international.

(I) Promouvoir la coopération fédérée

Point 34. Open source sous licences permissives. Les implémentations de référence devraient être disponibles sous des licences open source permissives. Les implémentations actuelles du MDSL constituent un ensemble de références parmi plusieurs autres possibles : le framework Tractatus est distribué sous Apache 2.0 pour le code et CC BY 4.0 pour la documentation ; les bases de code de Village et de la communauté migrent vers l'EUPL-1.2 (European Union Public Licence) par étapes à partir de mi-2026 ; les futures contributions MDSL sont destinées à être sous EUPL-1.2 lorsque cela est possible, pour l'alignement de la souveraineté sur les travaux de l'Union européenne en matière de souveraineté et pour la compatibilité avec la fédération bilatérale entre des installations souveraines dans plusieurs juridictions. Nous reconnaissons le mérite de l'engagement du cadre CAC en faveur de l'innovation open source, notamment les communautés open source d'IA nationales, la compatibilité avec les puces, les systèmes d'exploitation et les grands modèles open source, ainsi que l'implication des entreprises, des universités et des instituts de recherche dans des projets open source. L'open source sous licences permissives est compatible avec la fédération bilatérale : chaque installation souveraine effectue un fork de l' amont, contribue en retour via une pull request et prend ses propres décisions de déploiement. *(Parallèle avec le point 34 du CAC « favoriser l'innovation open source ».)* [RÉFÉRENCES : Apache 2.0 (Apache Software Foundation) ; EUPL-1.2 (Licence publique de l'Union européenne) ; CC BY 4.0 (Creative Commons).]

Point 35. Fédération par publication. Lorsque la coordination est nécessaire en matière de technologie commune, de normes d'interopérabilité, de réponse aux incidents de sécurité ou d'élaboration d'un cadre d'audit, elle s'effectue par le biais d'une publication ouverte et d'un consensus entre les installations contributrices. L'harmonisation internationale passe par le W3C, l'IETF, l'ISO/IEC et d'autres organismes de normalisation établis similaires. Nous reconnaissons le mérite de l'engagement du cadre CAC en faveur des plateformes de collaboration industrielle — notamment les alliances d'écosystèmes d'agents intelligents, les laboratoires de vérification technologique et les accords de R&D conjoints — ainsi que de la coordination des acteurs en amont et en aval de la chaîne d'approvisionnement dans les activités de R&D technologique commune, de normalisation et d'évaluation et de certification. **Ce domaine correspond au volet de travail (iv) du comité unique proposé au §II, point 4. Le comité élaborerait des recommandations adaptées au contexte néo-zélandais sur les modèles de fédération et d'alliance pour la coordination industrielle, contribuerait aux travaux de l'ISO/IEC SC42 sur les modèles de collaboration industrielle en matière d'IA, et engagerait un dialogue bilatéral avec les auteurs du cadre CAC sur l' interaction entre la coordination industrielle fédérée et celle basée sur des alliances .** *(Parallèle avec le point 35 du CAC « plateformes de collaboration industrielle » ; le volet de travail consolidé sur la formation du comité s'applique.)* [RÉFÉRENCES : document de procédure du W3C ; processus de demande de commentaires de l'IETF ; normes ISO/IEC 42001:2023 sur les systèmes de gestion.]

(II) Renforcement de la promotion bilatérale

Point 36. L'adoption est bilatérale. Chaque installation souveraine contacte directement ses contreparties — organisations partenaires, institutions homologues, pairs fédérés. Nous reconnaissons le mérite de l'engagement du cadre CAC envers les canaux de promotion des applications, notamment les boutiques de logiciels d'agents intelligents, les plateformes d'information sur l'offre et la demande du secteur, le développement de produits sur mesure via des appels d'offres et le modèle de défi « dévoiler et prendre les commandes », ainsi que le développement par des entreprises de systèmes matériels et logiciels de produits et services d'agents intelligents. Nous proposons, dans le contexte de l'Aotearoa NZ, que les canaux d'adoption émergent du paysage commercial, de la société civile et institutionnel existant ; les installations souveraines établissent leurs relations avec leurs contreparties par le biais

d’un engagement direct ordinaire, les marchés publics se conformant aux règles relatives aux marchés publics*. (Parallèles avec le point 36 du CAC « canaux de promotion des applications »).* [RÉFÉRENCES : Règles relatives aux marchés publics (NZ) ; vérification en cours concernant d'éventuelles réformes en cours des marchés publics en NZ.]

Point 37. Le déploiement pilote est bilatéral et fondé sur des données factuelles. Les installations souveraines testent l’adoption directement avec les communautés volontaires. Les déploiements MDSL existants — Village dans les contextes paroissiaux et hapū / iwi ; histoire familiale dans les contextes iwi et de la diaspora ; sydigital dans les contextes des petites entreprises — en sont des exemples ; des données de déploiement spécifiques (nombres, dates de lancement, portée des baux) doivent être ajoutées avant la publication de la v1. Nous reconnaissons le mérite de l’engagement du cadre CAC à favoriser l’ouverture de scénarios d’application d’agents intelligents dans des secteurs clés, avec des projets pilotes dans des pôles industriels, des industries clés et des secteurs clés constituant un portefeuille de projets de démonstration. Nous proposons que, pour l’Aotearoa NZ, le déploiement pilote soit bilatéral entre les installations de déploiement et leurs communautés consentantes. Lorsque des agences de la Couronne souhaitent tester l’IA agentique, elles le font dans le cadre des , de la Charte des algorithmes pour l'Aotearoa Nouvelle-Zélande et des engagements du Te Mana Raraunga / Réseau maori pour la souveraineté des données . *(Parallèle avec le point 37 du CAC : « faire progresser l'ouverture de scénarios clés ».)* [RÉFÉRENCES : Données de déploiement MDSL — Village (contextes paroissiaux et communautaires), histoire familiale (contextes iwi et de la diaspora), sydigital (contextes de petites entreprises), données spécifiques en attente de chiffres vérifiés par l’opérateur avant la publication de la v1 ; Charte des algorithmes pour l’Aotearoa Nouvelle-Zélande (2020) ; Te Kāhui Raraunga (kahuiraraunga.io — Modèle de gouvernance des données maories et Cadre de gouvernance de l'IA maorie) ; Taiuru, K. (20 sept. 2025) Analyse critique des principes de données Te Mana Raraunga, taiuru.co.nz/critical-analysis-mana-raraunga/.]

Point 38. Harmonisation internationale par fédération bilatérale . Les installations souveraines en Aotearoa Nouvelle-Zélande se fédèrent bilatéralement avec des installations souveraines d’autres juridictions ; l’engagement en matière de normes internationales s’effectue par le biais du W3C, de l’IETF, de l’ISO/IEC et de forums similaires sous forme de participation entre pairs. Nous reconnaissons le mérite de l’engagement du cadre CAC à cultiver activement l’ écosystème mondial par le biais de plateformes internationales telles que la Conférence mondiale sur l'intelligence artificielle et la Conférence mondiale sur l'Internet, la promotion de l'adaptation des agents intelligents par les entreprises de terminaux et de logiciels, ainsi que l'engagement en matière de conformité à l'étranger et d'adaptation aux lois, réglementations et coutumes culturelles locales. **Ce domaine correspond au volet de travail (v) du comité unique proposé au §II, point 4. Le comité élaborerait des recommandations adaptées au contexte néo-zélandais sur la coopération internationale en matière d’IA, contribuerait aux travaux de normalisation internationale de l’ISO/IEC SC42 et s’engagerait dans un dialogue bilatéral avec les auteurs du cadre CAC et avec des pairs internationaux sur l’interopérabilité entre les approches de fédération bilatérale et de projection de plateforme en matière de coopération internationale.** Nous proposons cela comme une contribution à un débat international à un stade précoce ; les contributions issues de nombreuses traditions architecturales et de contextes politiques divers amélioreront le domaine. *(Parallèle avec le point 38 du CAC « cultiver activement l'écosystème mondial » ; le volet de travail consolidé sur la formation de comités s'applique.)* [RÉFÉRENCES : ISO/IEC JTC 1/SC 42 ; processus de normalisation internationale du W3C ; recherche en cours concernant les accords bilatéraux actuels de la Nouvelle-Zélande en matière d'IA et ses engagements internationaux.]

§VI. Garantir l'adoption

En tant que promoteur de la société civile, My Digital Sovereignty Ltd ne coordonne pas l'adoption. Nous citons ici les organismes dont la participation serait nécessaire si une partie quelconque de ce cadre devait être adoptée par des entités d'Aotearoa Nouvelle-Zélande.

Les organismes publics concernés par cette proposition comprennent le ministère des Entreprises, de l'Innovation et de l'Emploi pour la stratégie numérique ; le ministère de la Justice pour l'harmonisation du cadre juridique ; le Bureau du Commissaire à la protection de la vie privée pour l'alignement sur la loi de 2020 sur la protection de la vie privée ; Stats NZ et Te Kāhui Raraunga pour l'harmonisation en matière de souveraineté des données (en s'appuyant sur l'analyse critique du Dr Karaitiana Taiuru du 20 septembre 2025 comme référence fondamentale) ; Te Whatu Ora / Health New Zealand pour la gouvernance des informations de santé ; Te Pūtea Matua / Banque de réserve de Nouvelle-Zélande pour l'harmonisation prudentielle des services financiers ; Waka Kotahi New Zealand Transport Agency pour les transports ; le ministère de l'Éducation pour l'éducation ; et la police néo-zélandaise Police pour les questions de sécurité publique. L'évaluation par la société civile impliquerait naturellement la Royal Society Te Apārangi, Internet NZ, NetSafe, le New Zealand AI Forum et des chercheurs universitaires issus des disciplines pertinentes. La prise en compte des hapū et des iwi est essentielle lorsque des obligations découlant du Traité ou des implications liées au règlement de litiges se présentent, et l'architecture spécifiée par cette proposition vise à soutenir — et est proposée pour être utilisée dans le cadre — les travaux sur la souveraineté des données maories tels qu'articulés par Te Kāhui Raraunga (modèle maori de gouvernance des données ; cadre maori de gouvernance de l'IA) et par les travaux universitaires publiés par le Dr Karaitiana Taiuru — y compris son analyse critique du 20 septembre 2025 qui explique pourquoi les cadres antérieurs sont inadéquats pour les contextes d'IA.

Un dialogue international avec les auteurs du cadre CAC et avec les réseaux autochtones de souveraineté des données — FNIGC au Canada, USIDSN aux États-Unis, Maiam nayri Wingara en Australie, GIDA à l'échelle internationale — enrichirait les deux sens de la conversation.

My Digital Sovereignty Ltd s'engage à respecter les principes d'ouverture architecturale et d'ouverture des licences de la proposition : le cadre Tractatus, les bases de code Village et communautaires, ainsi que les futures contributions de MDSL resteront disponibles sous des licences open source permissives, et les implémentations de référence seront développées en concertation avec les adoptants. Le reste s'adresse à ceux qui décideraient de l'adoption.

Nous terminons par une invitation explicite : aux auteurs du cadre CAC, aux pairs internationaux, aux décideurs politiques et aux organisateurs communautaires néo-zélandais, ainsi qu' à toute personne travaillant sur des questions parallèles — les commentaires sur cette v1 sont les bienvenus via les canaux permanents de commentaires sur les documents sur agenticgovernance.digital.

Annexe A. Objections techniques courantes + réponses

Cette annexe rassemble les objections techniques les plus courantes soulevées à l'encontre du cadre Tractatus, accompagnées de brèves réponses structurelles. Chaque réponse renvoie à la primitive §0(i) pertinente ou à une section ultérieure. Le cadre en est au stade de la recherche ; ces réponses résument le raisonnement architectural, et non la réalisation technique.

1. « Comment un LLM peut-il faire respecter quoi que ce soit ? »

Objection. Les LLM sont des compléteurs de modèles statistiques ; les invites adversaires et les capacités émergentes contournent les garde-fous au niveau du modèle.

Réponse. Le cadre ne demande pas au modèle d'imposer quoi que ce soit. Les primitives Tractatus sont des interceptions d'exécution — du code dans le processus hôte de l'agent qui décide si chaque action proposée s'exécute, avant que la sortie du modèle n'atteigne le monde. L'interception ne dépend pas de la conformité du modèle ni de l'entraînement à l'alignement ; elle applique des règles de classe de décision configurées à la limite de l'action, selon la même structure architecturale que l'interception des appels système privilégiés par le noyau du système d'exploitation. Voir §0(i) Application des limites.

2. « Qu'en est-il de JEPA / des architectures alternatives ? »

Objection. Le cadre semble supposer des LLM basés sur des transformateurs ; des changements architecturaux (JEPA, systèmes hybrides) pourraient le rendre caduc.

Réponse. Le cadre est indépendant du substrat. Ses primitives opèrent sur les actions des agents, et non sur les rouages internes du modèle — aucun accès aux poids, aux schémas d'attention ou aux traces de raisonnement n'est requis. Un agent, quelle que soit son architecture, émet des actions proposées ; la couche d'interception évalue ces actions par rapport aux mêmes limites de classe de décision. Voir §0(i), premier paragraphe.

3. « L'agent peut-il contourner l'interception ? »

Objection. Un agent suffisamment compétent pourrait modéliser l'interception et concevoir des actions qui la contournent.

Réponse. L'agent ne détient pas le privilège d'exécution — c'est l'interception qui le détient. Le contournement nécessite soit (a) d'opérer en dehors du champ de capacités de l'agent, ce que le runtime détecte comme une violation de privilège, soit (b) une escalade complète des privilèges contre le processus hôte, un problème au niveau du système d'exploitation renforcé de la même manière que les permissions du système de fichiers le gèrent. La piste d'audit (enregistrements souverains, signés cryptographiquement) rend les tentatives de contournement réussies visibles d'un point de vue légal, éliminant ainsi la défection en tant que stratégie. Le point structurel plus profond — développé dans *Alignement architectural* §3.4 — est la **distinction entre substrat et exécution**: même lorsqu'un agent suffisamment capable contourne les interceptions au moment de l'exécution, les mécanismes du substrat (provenance cryptographique, enveloppes de fédération, portabilité des enregistrements pilotée par les membres) ne peuvent pas être contournés car ils sont détenus de manière distribuée, indépendamment de l'agent. La sécurité du substrat découle des mathématiques et de la réplication distribuée, et non de la coopération de l'agent. Voir §0(i) Validation par recoupement ; §II point 5 ; *Alignement architectural* §3.4 substrat vs exécution ; §7.5 surface d'attaque de la couche sociale (la surface que le substrat *ne* ferme *pas*).

4. « En quoi cela diffère-t-il de la sécurité de l'ingénierie des invites ? »

Objection. L'ingénierie des prompts et le RLHF restreignent également la sortie du modèle. Le cadre semble similaire dans son esprit.

Réponse. Structurellement différent. L'ingénierie des prompts et le RLHF modifient la distribution des sorties du modèle mais laissent inchangée la mécanique statistique. Les primitives du cadre s'exécutent avant l'invocation du modèle (définition de la portée des capacités), après l'action proposée (application des limites) ou parallèlement à l'invocation (validation par recoupement) — aucune ne dépend de la production de la sortie correcte par

le modèle. Elles dépendent de la capacité de la couche d'exécution à identifier correctement l'appartenance à une classe de décision. Voir §0(i) Application des limites et vérification métacognitive.

5. « Et si le service d'exécution lui-même était exploité ? »

Objection. Faire confiance à un service d'exécution déplace la surface d'attaque plutôt que de la supprimer.

Réponse. Le cadre ne prétend pas que les services d'exécution sont inviolables. Il affirme que toute défaillance est documentée et que ces documents limitent l'ampleur des dégâts.

À quoi ressemble une violation. Soit le code du service est exploité — un attaquant intercepte les données pour approuver des actions hors politique — soit l'état de la politique que le service consulte est réécrit — un attaquant modifie les classes de décision qui sont acheminées vers l'approbation humaine. Dans les deux cas, l'objectif de l'attaquant est de convertir une décision « acheminée vers l'humain » en une décision « auto-approuvée » sans que l'opérateur ne s'en aperçoive.

Ce qui est en jeu. Les décisions que le cadre aurait autrement acheminées vers une approbation humaine — jugements de valeur, opérations irréversibles, accès aux données entre locataires. La portée de l'impact est limitée par ce que l'interception était déjà autorisée à approuver : le cadre accorde des autorisations d'approbation, pas de nouveaux privilèges pour effectuer des actions. Une interception exploitée ne peut pas exfiltrer des données auxquelles l'agent n'a jamais eu accès au départ ; elle ne peut qu'approuver à tort des actions dans le cadre des capacités existantes de l'agent.

Remèdes. Trois propriétés se combinent. (i) Chaque décision du cadre est consignée dans la piste d’audit du registre souverain — signée cryptographiquement, en ajout seul, répliquée en fédération vers les pairs — afin qu’une analyse forensic a posteriori puisse reconstituer ce qui a été approuvé, par quelle version du service, et par rapport à quel état de la politique. La fenêtre de violation est limitée dans le temps et en portée. (ii) La validation par recoupement (§0(i)) détecte les divergences entre les approbations observées et la politique déclarée en temps quasi réel, mettant en évidence les violations avant qu'elles ne se normalisent. (iii) La réplication au sein de la fédération empêche un attaquant contrôlant un nœud d'effacer rétroactivement des enregistrements ; toute défection nécessite une collusion avec la fédération, et non la compromission d'un service isolé.

L'ancrage de confiance peut échouer, mais l'échec est limité, observable et reconstituable de manière forensic — le même modèle architectural que la transparence des certificats pour l'infrastructure PKI TLS : l'ancrage de confiance (une autorité de certification) peut être compromis, mais le journal d'audit des certificats émis rend la compromission visible à l'échelle mondiale.

La posture de survie est stratifiée. Les services d'exécution du cadre (BoundaryEnforcer, les primitives §0(i)) s'adressent *à l'agent*: ils limitent ce que l'agent peut autoriser. L'architecture des enregistrements souverains (Document A) s'*adresse au substrat*: elle garantit que les enregistrements survivent à l'agent, que celui-ci échappe ou non à la barrière. Voir *Alignement architectural* §7.4 (posture de survie indépendante du confinement de l'agent) et §7.5 (la surface d'attaque de la couche sociale que le substrat ne ferme pas — persuasion, usurpation d'identité coordonnée à grande échelle, consentement synthétisé — désignée ici comme une frontière ouverte). Le schéma de signature PKI sous-tendant la chaîne d'audit présente un horizon de vulnérabilité quantique (10 à 30 ans) ; les normes de signature post-quantique du NIST ont été finalisées en août 2024 et la voie de migration suit le processus de normalisation, la falsification par enregistrement étant coûteuse même sur un CRQC et les mécanismes de fédération/portabilité ne dépendant pas de l'intégrité de la signature. Voir *le document A* §5.3.

Voir §II point 5 ; §0(i) Validation par recoupement.

6. « Et si les valeurs changeaient ? »

Objection. Les limites des classes de décision figent les valeurs actuelles ; les valeurs des communautés évoluent, ce qui rend le cadre fragile.

Réponse. Les classes de décision relèvent de la configuration, et non de l'architecture. Le cadre fournit le mécanisme d'interception ; les classes d'action acheminées vers l'approbation humaine sont modifiables par l'opérateur pour chaque locataire (§III, point 3 — gouvernance pilotée par les adoptants). Lorsque les parties prenantes au sein d'un périmètre de gouvernance occupent des positions limites incompatibles, la primitive d'orchestration de la délibération pluraliste structure la délibération plutôt que de désigner un gagnant. Le cadre est conçu pour accueillir des valeurs en évolution, et non pour les figer. Voir §0(i) Orchestration de la délibération pluraliste ; §III, point 3.

Licence et citation

Citation suggérée : Stroh, J. G. (2026). *Une proposition de la société civile pour une IA agentique souveraine et fédérée en Aotearoa Nouvelle-Zélande* (v1.2, mai 2026, révisée conformément à la correspondance de Ted Howard sur la v1.1). My Digital Sovereignty Ltd. https://agenticgovernance.digital/papers/aotearoa-nz-agentic-ai-framework-v1.2-may-2026.html

Commentaires et correspondance : les commentaires de fond portant sur des sections spécifiques sont les bienvenus. Veuillez citer les numéros de section (par exemple §III point 5) afin que les corrections puissent être retracées. L'auteur répond personnellement ; veuillez prévoir un délai d'une à deux semaines. E-mail : john.stroh@mysovereignty.digital.