Ein Vorschlag der Zivilgesellschaft für eine souveräne und föderierte agentische KI in Aotearoa Neuseeland

John G. Stroh / My Digital Sovereignty Ltd

Vorschlag der Zivilgesellschaft ·

v1.2 Entwurf vom Mai 2026 |

Konstruktive Parallele zu den CAC 2026-Implementierungsrichtlinien →

v1 (ersetzt) → Feedback per E-Mail

v1.2, 16.05.2026: Überarbeitet gemäß Ted Howards Korrespondenz zu v1.1, mit Schwerpunkt auf der Unterscheidung zwischen Substrat- und Laufzeitebene. Der Abschnitt „§0(i) Primitive“ enthält einen ergänzenden Absatz, der klarstellt, dass sich die Substratschicht (PKI, Föderation, Portabilität) architektonisch von den Primitiven der Laufzeit-Schicht unterscheidet, die die sechs Dienste unter §0(i) bilden; „§0(i) Durchsetzung von Grenzen“ ergänzt den Rahmen der vier Kategorien der Fehlbarkeit und die dreistufige Router-Ausgabe (zulassen / verweigern / eskalieren); Anhang A, obj-3 und obj-5, vertiefen die Trennung zwischen Substrat und Laufzeit sowie die Unabhängigkeit der Überlebensfähigkeit von der Eindämmung durch Agenten. Die wesentlichen architektonischen Details werden in *„Architectural Alignment“* §3.3, §3.4, §3.5, §7.4, §7.5 und in *Paper A* §5.3 ausgearbeitet. v1.1 bleibt unter ihrer URL als historische Referenz zugänglich.

v1.1, 14.05.2026: am selben Tag überarbeitet gemäß dem Feedback von Dr. Karaitiana Taiuru zu v1. §0(iii) zitiert nun Taiurus kritische Analyse von Te Mana Raraunga vom 20. September 2025; nennt Te Kāhui Raraunga als das derzeit anerkannte operative Gremium; fügt eine explizite Lückenanalyse hinzu. v1 bleibt unter der v1-URL als historische Referenz zugänglich. Kommentare zu bestimmten Abschnitten sind willkommen. Bitte geben Sie die Abschnittsnummern an (z. B. §III Punkt 5). Der Autor antwortet persönlich; rechnen Sie bitte mit ein bis zwei Wochen.

Ein Vorschlag der Zivilgesellschaft für eine souveräne und föderierte agentebasierte KI in Aotearoa New Zealand

v1.2 Mai 2026 – Entwurf eines Forschungsartikels (überarbeitet gemäß Ted Howards Korrespondenz zu v1.1; Unterscheidung zwischen Substrat und Laufzeit, vierkategorische Fehlbarkeit, trinarer Router-Ausgang). Konstruktive Parallele zu den Umsetzungsrichtlinien der Volksrepublik China von 2026 für intelligente Agenten.

John G. Stroh / My Digital Sovereignty Ltd

16.05.2026

Ein Vorschlag der Zivilgesellschaft für eine souveräne und föderierte agentische KI in Aotearoa New Zealand

v1.2 Mai 2026 – Entwurf eines Forschungsberichts (überarbeitet gemäß Ted Howards Korrespondenz zu v1.1; Unterscheidung zwischen Substrat und Laufzeit, vier Kategorien der Fehlbarkeit, trinarer Router-Ausgang; siehe „Über diesen Bericht“)

Ein zivilgesellschaftlicher Vorschlag von My Digital Sovereignty Ltd, vorgelegt an neuseeländische politische Entscheidungsträger, Community-Organisatoren und Fachleute der Branche. Konzipiert als konstruktive Parallele zu den „Implementierungsrichtlinien für die standardisierte Anwendung und innovative Entwicklung intelligenter Agenten“

der Volksrepublik China aus dem Jahr 2026, verfügbar in englischer Übersetzung unter /research/translations/.

Über dieses Papier

Dies ist der Entwurf **v1.2 vom Mai 2026** eines zivilgesellschaftlichen Vorschlags von My Digital Sovereignty Ltd, der neuseeländischen Politiker*innen, Community-Organisator*innen und Fachleuten der Branche vorgelegt wurde. Kommentare sind über die ständigen Kanäle für Papierkommentare auf agenticgovernance.digital willkommen; Überarbeitungen als Reaktion auf Kommentare werden als v2 veröffentlicht.

Änderungsprotokoll v1 → v1.1 (14.05.2026): v1 wurde zuvor am 14.05.2026 veröffentlicht und umgehend von Dr. Karaitiana Taiuru geprüft, der darauf hinwies, dass die grundlegende Bezugnahme von v1 Te Mana Raraunga Māori von 2016–2018 im KI-Kontext überholt ist (gemäß seiner *kritischen Analyse der Te Mana Raraunga-Datenprinzipien vom* 20. September 2025). v1.1 überarbeitet §0(iii), um Taiuru kritische Analyse direkt zu zitieren; um **Te Kāhui Raraunga** (das derzeit anerkannte operative Gremium für Māori -Datengouvernanz in Aotearoa, Neuseeland, gegründet 2019) und dessen veröffentlichtes Modell Māori sowie Māori AI Governance Framework als aktuelle Formulierungen zu zitieren; Taiurubevorzugte Grundbegriffe (mana motuhake, rangatira) an geeigneter Stelle zu übernehmen; und einen expliziten Unterabschnitt zur Lückenanalyse hinzuzufügen, in dem aufgeführt wird, was dieser Vorschlag in der Dimension des te ao Māori bereits leistet und was noch nicht. Die Punkte 4, 23, 37, §I Grundsatz 2 und §VI enthalten dieselbe Aktualisierung der Quellenangaben. Die in diesem Vorschlag festgelegte Architektur bleibt gegenüber v1 unverändert. v1 bleibt unter `/papers/aotearoa-nz-agentic-ai-framework-v1-may-2026.html` als historische Referenz zugänglich; diese URL führt zu v1.2.

Änderungsprotokoll v1.1 → v1.2 (16.05.2026): v1.1 wurde von Ted Howard im Schriftwechsel überprüft; er wies darauf hin, dass die Darstellung der sechs §0(i)-Primitiven in v1.1 als architektonische Sicherheitsmechanismen am besten zu rechtfertigen ist, wenn die Unterscheidung zwischen Substrat und Laufzeitumgebung explizit gemacht wird. v1.2 fügt einen klarstellenden Absatz nach der Liste der §0(i)-Framework-Primitiven ein, in dem die Substratschicht (PKI / Föderations-Envelopes / portable Datensätze, entwickelt in *Paper A* und *„Architectural Alignment“* §3.4) als architektonisches Pendant zur Laufzeitschicht, die die sechs §0(i)-Dienste bilden. Die Durchsetzung der §0(i)-Grenzen erhält den Rahmen der vier Kategorien der Fehlbarkeit (die Kategorien sind gemeinschaftlich ausgehandelt und anfechtbar, keine festgelegten Wesensmerkmale) sowie den Hinweis auf die dreistufige Router-Ausgabe (zulassen / verweigern / an Menschen eskalieren). Anhang A: obj-3 (Umgehung) und obj-5 (Ausnutzung von Laufzeitdiensten) erhalten Querverweise auf *„Architectural Alignment“* §3.4 Substrat vs. Laufzeit, §7.4 Überlebensfähigkeit unabhängig von der Eindämmung von Agenten, §7.5 Angriffsfläche der sozialen Ebene (offene Grenze) und Paper A §5.3 PQC-Migrationshorizont. Die in diesem Vorschlag spezifizierte Architektur bleibt gegenüber v1.1 unverändert; die Überarbeitungen

dienen der Klarstellung von Formulierungen, die im Dialog mit den Gutachtern zutage traten. v1.1 bleibt unter `/papers/aotearoa-nz-agentic-ai-framework-v1.1-may-2026.html` als historische Referenz zugänglich.

v1.2 Operationalisierungsdurchlauf am selben Tag (16.05.2026 abends): §III(II) Punkt 10 (Blast-Radius-Mechanismus), §III(III) Punkte 11–12 (polyzentrische Governance-Grundlage), §III(IV) Punkt 13 (Föderationsmechanismen), §I Prinzip 4 (architektonisch ermöglichte Nachweise der Akzeptanz) und §IV (akzeptanzorientierte Entwicklung) – alle neunzehn Sektorpunkte – erhalten eine Verankerung in den Primitivfähigkeiten, damit politische Akteure sich im Ausschuss für die substanziellen Souveränitätsansprüche einsetzen können. Architektur-generisches Register: bezeichnet **kryptografische Herkunft**, **Föderations- Umschläge**, **mitgliederorientierte Portabilität** und **Grenzdurchsetzung** durch Fähigkeiten statt durch MDSL- Implementierung. §IV erhält einen einleitenden Absatz, der die vier Substrat- Primitiven einmalig benennt; jeder Sektorpunkt operationalisiert dann nur die sektorspezifische Ausprägung. Inhaltlich unverändert gegenüber v1.2 am Morgen; die Aktualisierung macht die Souveränitätsansprüche für Leser operativ lesbar, die sich in Ausschusssitzungen dafür einsetzen würden.

v1.1 Klarstellungsüberarbeitung am selben Tag (14.05.2026, abends): §0(i) Einleitungsabsatz überarbeitet, um explizit mit der Unterscheidung zwischen System- und Modellebene zu beginnen (Primitiven auf Systemebene, Laufzeitprüfungen auf Codeebene, substratunabhängig über Transformer-LLMs / JEPA-Stil / hybride Architekturen hinweg). Formulierung des Abschnitts „BoundaryEnforcer“: „durch Architektur statt durch Hoffnung“ → „durch Laufzeit-Interception statt durch Hoffnung“, um Mehrdeutigkeiten für Leser aus dem Ingenieursbereich zu verringern, die mit der LLM-Alignment-Debatte vertraut sind. Inhaltlich unverändert; dies ist eine Formulierungsänderung zur Verbesserung der Verständlichkeit. Ausgelöst durch einen technischen Leser, der die Primitiven in §0(i) mit Alignment-Behauptungen auf Modellebene in Verbindung brachte, die dort nicht aufgestellt werden.

v1.1 Klarstellungsrevision der Stufe 2 am selben Tag (15.05.2026): Jede der sechs §0(i)-Primitiven erhielt eine konkrete technische Analogie (privilegierte Systemaufrufe des Betriebssystemkerns / Circuit Breaker / Laufzeitprüfung vs. Ausrichtung während des Trainings / Konfiguration vs. Laufzeitargument / Verifizierungsgate an der Laufzeitgrenze / Koordinationsdienst). Die metakognitive Verifikationsprimitive wurde von „erfordert, dass Agenten ihre eigene Argumentation überprüfen“ in „setzt ein Verifikationsgatter vor die Aktionsausführung“ umformuliert, um das verbleibende Lesen auf Modellebene zu entfernen. Inhaltlich unverändert; redaktionelle Überarbeitungen zur besseren Verständlichkeit.

v1.1 Klarstellungsrevision auf Ebene 3 am selben Tag (15.05.2026): Anhang A hinzugefügt – „Häufige technische Einwände + Antworten“ – sechs Einwand-Antwort-Paare zu den Themen Skepsis gegenüber der LLM-Durchsetzung, Substrat-Agnostizismus, Umgehung von Agenten, Äquivalenz von Prompt-Engineering, Ausnutzung von Laufzeitdiensten und Werteentwicklung. Erweiterung der inhaltlichen Arbeit zu L1 + L2 am selben Tag; keine Ansprüche über die primitiven Spezifikationen in §0(i) hinaus.

Dieses Papier stellt **keine** Politik der neuseeländischen Regierung dar. Es wird **nicht**

von der Krone unterstützt. Es ist in **keinem**formalen Sinne vertragsbegründet. Es handelt sich um einen Vorschlag der Zivilgesellschaft von My Digital Sovereignty Ltd, der neuseeländischen Interessengruppen als Grundlage für die Übernahme, Anpassung oder Ablehnung angeboten wird. Soweit seine Grundsätze für die eigene Arbeit des Übernehmers nützlich sind, können sie unter freizügigen Open-Source-Lizenzen genutzt werden; wo dies nicht der Fall ist, verbleiben sie auf der Seite.

Die gespiegelte Quellstruktur ist in englischer Übersetzung unter `/research/translations/china-cac-implementation-guidelines-2026.html` und im Original als *„Implementierungsrichtlinien* 2026“ der Cyberspace Administration of China auf Mandarin veröffentlicht.

Zusammenfassung

Dieses Papier schlägt einen souveränen, föderierten Rahmen für die Anwendung und Entwicklung intelligenter Agenten in Aotearoa Neuseeland vor, der als Beitrag der Zivilgesellschaft von My Digital Sovereignty Ltd. angeboten wird. Der Vorschlag ist als konstruktive Parallele zu den *„Implementierungsrichtlinien* 2026“ der Volksrepublik China *für die standardisierte Anwendung und innovative Entwicklung intelligenter Agenten* – sechs Abschnitte, vierzehn Unterabschnitte, achtunddreißig nummerierte Punkte – mit einem neuen vorangestellten Kapitel §0 „Philosophische Grundlagen“. §0 stützt sich auf drei Traditionen: die sechs Laufzeitdienste des Tractatus AI Safety Framework (Grenzdurchsetzung, Kontextdrucküberwachung, Querverweisvalidierung, Anweisungsbeständigkeit , metakognitive Verifizierung, Orchestrierung pluralistischer Deliberation); die CARE-Prinzipien für indigene Daten-Governance (Carroll et al. 2020) und die globale Bewegung für indigene Daten-Souveränität , aus der sie hervorgegangen sind, einschließlich der auf dem Te Tiriti basierenden wissenschaftlichen Arbeiten von Te Mana Raraunga und Dr. Karaitiana Taiuru; sowie die internationale KI-Standards-Landschaft, die durch ISO/IEC JTC 1/SC 42 koordiniert wird (22989 Terminologie, 23053 Lebenszyklus, 23894 Risikomanagement, 42001 Managementsysteme). Der Vorschlag befürwortet die Bildung eines Ausschusses unter einer geeigneten Dachorganisation – Kandidaten sind unter anderem die Royal Society Te Apārangi, das Spiegelkomitee SC42 von Standards New Zealand und das New Zealand AI Forum –, um Empfehlungen für den neuseeländischen Kontext zu entwickeln und einen internationalen Dialog zu führen. Dies ist der Entwurf v1 vom Mai 2026; Kommentare sind über die üblichen Kanäle für Papierkommentare auf agenticgovernance.digital willkommen.

Präambel

Intelligente Agenten – intelligente Systeme, die zu autonomer Wahrnehmung, Gedächtnis, Entscheidungsfindung, Interaktion und Ausführung fähig sind – beschleunigen ihre Integration in die Datensätze, die Infrastruktur und die sozialen Prozesse von Aotearoa New Zealand. Dieser Vorschlag bietet einen Beitrag der Zivilgesellschaft dazu, wie diese Integration geregelt werden sollte: eine souveräne, föderierte Architektur, in der jeder Vorgang eines intelligenten Agenten in Bezug auf

einen Datensatz einen zugeordneten, kryptografisch signierten Eintrag gegenüber dem Inhaber des Datensatzes erzeugt und in der die Koordination zwischen souveränen Einrichtungen durch bilaterale Föderation erfolgt. Der Vorschlag spiegelt die Struktur der *Umsetzungsrichtlinien* 2026 der Volksrepublik China wider, sodass die architektonischen Entscheidungen auf beiden Seiten in konstruktiver Parallele erscheinen und einen Dialog mit den Verfassern dieses Rahmens, mit internationalen Kollegen sowie mit neuseeländischen politischen Entscheidungsträgern, Community-Organisatoren und Fachleuten aus der Branche eröffnen. My Digital Sovereignty Ltd bietet dies als Ausgangspunkt – zur Übernahme, Anpassung und Überarbeitung – unter freizügigen Open-Source-Lizenzen an. Es wird als Beitrag der Zivilgesellschaft angeboten und erhebt keinen Anspruch auf den Status einer staatlichen Politik. Soweit seine Grundsätze für die eigene Arbeit des Anwenders nützlich sind, stehen sie zur freien Verwendung zur Verfügung; wo dies nicht der Fall ist, bleiben sie auf der Seite.

---

§0. Philosophische Grundlagen

Wir beginnen mit den Grundlagen, da Architektur aus der Philosophie folgt. Die Empfehlungen, die in §I–§VI folgen, sind keine willkürlichen technischen Entscheidungen; sie sind Implikationen philosophischer Verpflichtungen, die in diesem Abschnitt ausdrücklich benannt werden. Drei Strömungen laufen hier zusammen: die strukturelle Darstellung desTractatus AI SafetyFramework, wie intelligente Agenten sicher mit Datensätzen souveräner Instanzen umgehen können, die unter agenticgovernance.digital entwickelt und offen veröffentlicht wurde; die globale Bewegung für indigene Datensouveränität, die zum Ausdruck bringt, dass Daten über Menschen diesen Menschen und den Gemeinschaften, denen sie angehören, gehören; und die internationale Arbeit an KI-Standards, die durch ISO/IEC JTC 1/SC 42 koordiniert wird und das formales Vokabular bereitstellt, mit dem architektonische Empfehlungen in der organisatorischen Praxis umsetzbar werden. Die Nennung aller drei zu Beginn ist Teil des konstruktiven Beitrags, den dieser Vorschlag zum Dialog leistet – mit den Verfassern der „2026 *Implementation Guidelines“* der Cyberspace Administration of China, mit neuseeländischen politischen Entscheidungsträgern und Gemeinschaftsorganisatoren sowie mit internationalen Kollegen, die an parallelen Fragestellungen arbeiten.

(i) Tractatus als benannte Grundlagen

Das Tractatus besteht aus sechs **Primitiven auf Systemebene,** die gemeinsam die architektonischen Bedingungen festlegen, unter denen intelligente Agenten sicher mit Datensätzen arbeiten können, die von souveränen Instanzen verwaltet werden. **Es handelt sich dabei nicht um Abgleichtechniken auf Modellebene**, sondern um Laufzeitprüfungen auf Codeebene, die den Agenten umschließen, unabhängig davon, wie der zugrunde liegende Agent (aktuelle Transformer-LLMs, zukünftige Architekturen im JEPA-Stil, hybride Systeme) aufgebaut oder trainiert ist. Sie fangen das Verhalten an der Laufzeitgrenze ab und überprüfen es – dieselbe architektonische Form wie beim Scoping von Dateisystemfunktionen oder bei OAuth-

Bereichsprüfungen. Eine funktionierende Demo der Primitive zur Grenzdurchsetzung finden Sie unter /demos/boundary-demo.html. [ZITAT: Stroh, J. (2026). Tractatus AI Safety Framework — Kernwerte und -prinzipien sowie Kernkonzepte des Tractatus. Agentic Governance Digital. https://agenticgovernance.digital — beide Werke CC BY 4.0.]

Die Durchsetzung von Grenzen legt fest, welche Entscheidungstypen strukturell eine menschliche Genehmigung erfordern. Die grundlegende These – in Anlehnung an Wittgenstein und im Tractatus ausdrücklich benannt – lautet: „Was nicht systematisiert werden kann, darf nicht automatisiert werden." Wertentscheidungen, kulturell-kontextbezogene Urteile, irreversible Konsequenzen und beispiellose Situationen sind nicht an autonome Agenten delegierbar; das Framework verhindert eine solche Delegation durch eine Laufzeit-Interception und nicht durch Hoffnung. Die Interception wird vor der Ausführung der Aktion ausgelöst, in derselben architektonischen Form wie ein OS-Kernel, der privilegierte Systemaufrufe abfängt – der Prozess kann die Prüfung nicht umgehen. Die vier oben genannten Kategorien werden als **fehlbare Klassifizierungen** behandelt, die **in der Praxis von der Community ausgehandelt werden**, nicht als feststehende Essenzen – Klassifizierungsfehler werden selbst protokolliert, sind anfechtbar und werden genutzt, um das Verhalten des Routers im Laufe der Zeit zu überarbeiten (siehe *Architectural Alignment* §3.5). Die Ausgabe des Routers ist trinär, nicht binär: *zulassen*, *verweigern* und *an menschliche Entscheidungsinstanz eskalieren*. Der dritte Zustand ist tragend – er trägt die architektonische Erkenntnis in sich, dass ein erheblicher Teil bedeutender Entscheidungen zum Zeitpunkt der Entscheidung nicht binär ist (siehe §3.3).

Die Überwachung des Kontextdrucks erkennt an, dass das Kontextfenster eines Agenten eine endliche Ressource ist und dass Druck auf die Kapazität ein Governance-Signal darstellt. Agenten, die nahe an ihrer Kapazitätsgrenze arbeiten, machen mehr Fehler, und das Framework greift vor einem Ausfall ein, anstatt danach. Es funktioniert wie ein Leistungsschalter: Der Schalter löst bei gemessener Last aus, bevor das System sich selbst beschädigt; das Framework drosselt die Leistung oder leitet die Entscheidung zur menschlichen Genehmigung weiter, basierend auf der gemessenen Kontextnutzung, bevor die Ausgabequalität abnimmt.

Die Querverweis-Validierung vergleicht die vom Agenten vorgeschlagenen Aktionen mit dem kanonischen Anweisungsverlauf und erkennt Fälle, in denen Muster aus der Trainingsphase explizite Anweisungen des Benutzers außer Kraft setzen. Ein anschauliches Beispiel ist der „Vorfall 27027": Ein Benutzer gibt einen vom Standard abweichenden Datenbankport an, und der Agent verwendet – trotz der expliziten Anweisung – standardmäßig die Portnummer, auf die er trainiert wurde. Die Validierung erkennt die Überschreibung; ohne sie würde die Überschreibung unbemerkt die Abläufe beeinträchtigen. Der Validator ist eine Laufzeitprüfung jeder vorgeschlagenen Aktion, keine Anpassung des Modells selbst während des Trainings.

Die Klassifizierung der Anweisungsbeständigkeit unterscheidet vorübergehende Anweisungen von dauerhaften Governance-Zuständen. Nicht alle Anweisungen sind gleich wichtig; sie so zu behandeln, als wären sie es, beeinträchtigt sowohl die Sicherheit (kritische Anweisungen werden vergessen) als auch die Benutzerfreundlichkeit (triviale Präferenzen werden übermäßig durchgesetzt). Dies entspricht der

Unterscheidung zwischen Konfigurations- und Laufzeitargumenten in der Software: Konfigurationswerte sind persistent; CLI-Argumente gelten pro Aufruf; der Klassifikator kennzeichnet jede Anweisung nach Klasse, damit nachgelagerte Dienste sie entsprechend behandeln.

Die metakognitive Verifizierung setzt eine Verifizierungsschranke vor die Ausführung von Aktionen. Die Schranke bewertet jede vorgeschlagene Aktion anhand von fünf Dimensionen – Ausrichtung, Kohärenz, Vollständigkeit, Sicherheit und Berücksichtigung von Alternativen – und Konfidenzschwellenwerte bestimmen, ob Aktionen fortgesetzt werden, mit Vorsicht fortgesetzt werden, einer Überprüfung bedürfen oder blockiert werden. Die Prüfung erfolgt an der Laufzeitgrenze und ist keine Verhaltensanforderung an das Modell.

Die Orchestrierung pluralistischer Beratungen erleichtert die Beratung unter mehreren Interessengruppen, wenn die Durchsetzung von Grenzen einen Wertekonflikt signalisiert. Sie entscheidet nicht zwischen moralischen Rahmenwerken; sie strukturiert die Beratung so, dass die Werte der verschiedenen Interessengruppen dokumentiert, wo möglich berücksichtigt und ausdrücklich benannt werden, wenn sie nicht in Einklang gebracht werden können. Fundamentaler Pluralismus – die Ansicht, dass moralische Rahmenwerke irreduzibel unterschiedlich sind und dass kein übergeordneter Wert sie auflöst – ist die philosophische Verpflichtung, die pluralistische Beratung zu einer strukturellen Grundkomponente statt zu einer prozeduralen Feinheit macht. Er läuft als Koordinationsdienst, der die Positionen der Stakeholder dokumentiert und sichtbar macht; die Orchestrierung verlangt vom Agenten nicht, Werte intern zu vermitteln.

Diese sechs Dienste bilden das strukturelle Gerüst dieses Vorschlags. Jede der folgenden architektonischen Empfehlungen lässt sich auf einen oder mehrere davon zurückführen.

Unterscheidung zwischen Substrat und Laufzeit. Die sechs oben genannten §0(i)- Primitiven bilden die *Laufzeitschicht* des Frameworks – Code im Host-Prozess des Agenten, der vorgeschlagene Aktionen im Moment der Entscheidung anhand konfigurierter Regeln der Entscheidungsklasse überprüft. Die Architektur verfügt über eine zweite Schicht, die nicht von der Zusammenarbeit zur Laufzeit abhängt: *Substrat*- Mechanismen – kryptografische Provenienz, Föderations-Envelopes und mitgliedsgesteuerte Portabilität von Datensätzen –, die sich in verteiltem Besitz unabhängig vom Agenten befinden. Eine tiefere Netzwerk-Argumentation rund um die Laufzeit- Schicht kann nicht auf die Substrat-Schicht angewendet werden: Die Sicherheit des Substrats ergibt sich aus der Mathematik (Signaturen, verteilte Replikation, Beenden ohne Erlaubnis) und nicht aus der Zusammenarbeit der Agenten. Die Substrat-Schicht wird in Paper A, *Sovereign-Record Architecture* (*Artikel A*), und die Unterscheidung wird in *„Architectural Alignment“* §3.4 erörtert. Die beiden Schichten ergänzen sich: Die Laufzeit erfasst, was sie kann; das Substrat stellt sicher, dass das, was die Laufzeit übersieht, der Gemeinschaft dennoch überprüfbare Datensätze, Verbundwege und Ausstiegsmöglichkeiten belässt. Die Überlebensstrategie ist mehrschichtig und basiert nicht auf einem einzigen Mechanismus (siehe *„Architectural Alignment“* §7.4 zur Erläuterung der Unabhängigkeit von der Agenten-Eindämmung).

(ii) Die CARE-Prinzipien für indigene Datenverwaltung

Die CARE-Prinzipien für indigene Datenverwaltung, die 2020 von einem internationalen Team indigener Datenwissenschaftler unter der Schirmherrschaft der Global Indigenous Data Alliance veröffentlicht wurden, formulieren vier Verpflichtungen: **Kollektiver Nutzen** (Datenökosysteme sollten die indigene Selbstbestimmung und den kollektiven Nutzen fördern); **Kontrollbefugnis** (die Rechte und Interessen indigener Völker an ihren Daten müssen anerkannt werden); **Verantwortung** (diejenigen, die mit indigenen Daten arbeiten, haben die Verantwortung, offenzulegen, wie diese Daten zur Unterstützung der Selbstbestimmung indigener Völker genutzt werden); und **Ethik** (die Rechte und das Wohlergehen indigener Völker sollten in allen Phasen des Datenlebenszyklus im Vordergrund stehen). [ZITAT: 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). Die CARE-Prinzipien für die Datenverwaltung indigener Völker. Data Science Journal, 19, 43. https://doi.org/10.5334/dsj-2020-043]

CARE versteht sich als Ergänzung zu FAIR (Findable, Accessible, Interoperable, Reusable). FAIR optimiert den Datenfluss und die Wiederverwendung; CARE optimiert die Rechte und das Wohlergehen derjenigen, um die es bei den Daten geht. Die beiden stehen nicht im Widerspruch zueinander. Sie befassen sich mit unterschiedlichen Fragen: FAIR fragt, wie Daten fließen sollten; CARE fragt, unter wessen Autorität der Datenfluss geregelt wird. Eine gut konzipierte Souveränitätsarchitektur beantwortet beide Fragen.

Wir übernehmen CARE als grundlegende Referenz. Wo die folgenden Empfehlungen festlegen, dass Akteure auf der Grundlage von zugeordneten, herkunftsverankerten Datensätzen handeln müssen, die von ihren souveränen Inhabern gehalten werden, operationalisiert diese Spezifikation die Verpflichtung zur Kontrollhoheit. Wo die Empfehlungen eine föderierte Koordination anstelle einer zentralen Registrierung vorsehen, steht diese Spezifikation im Einklang mit der Verantwortung – diejenigen, die Daten halten, sind gegenüber denjenigen rechenschaftspflichtig, die von den Daten betroffen sind.

(iii) Te Tiriti, tikanga und mātauranga in der KI-Ethik – Wissenschaft in Aotearoa New Zealand

Die wissenschaftliche Arbeit zu indigener DatenhoheitAotearoa New Zealandgehört zu den international am weitesten entwickelten. Die frühe Formulierung der PrinzipienMāori stammt von Te Mana Raraunga (dem Māori Data Sovereignty Network), das 2015 gegründet wurde und dessen Charta 2016 verabschiedet wurde. Diese Prinzipien wurden von Dr. Karaitiana Taiuru *„Kritischen Analyse der Datenprinzipien von Te Mana Raraunga“* vom 20. September 2025 eingehend überprüft, die feststellt, dass diese KI, KI-Voreingenommenheit und algorithmische Diskriminierung, Modelltraining und -analyse, digitalen Kolonialismus oder Umweltauswirkungen nicht angemessen behandeln; sie stellt fest, dass der Anwendungsbereich von 2016 eng gefasst war, während „heute Māori überall zu finden sind“; und stellt fest, dass die Grundsätze trotz umfangreicher akademischer

Zitierung in der Praxis weitgehend nicht umgesetzt werden. [ZITAT: Taiuru, K. (20. September 2025). Kritische Analyse der Te Mana Raraunga. https://www.taiuru.co.nz/critical-analysis-mana-raraunga/]

Die derzeit anerkannte operative Stelle in Aotearoa New Zealand für die Māori -Datenverwaltung ist **Te Kāhui Raraunga** (gegründet 2019 als gemeinnützige Stiftung). Ihre veröffentlichten Rahmenwerke – das **Māori „Tuia te korowai o Hine-Raraunga“**, das auf acht Säulen basiert; das **Māori KI-Governance-Rahmenwerk**, das dieses erweitert; sowie der begleitende **Māori** sowie**die Referenzressource für konzeptionelle KI-Anwendungsfälle** – bieten die aktuelle Darstellung der Māori und KI-Governance. Te Kāhui Raraunga beschreibt das Māori AI Governance Framework als „aktiviert“ und verweist dabei auf Fallstudien aus dem öffentlichen Sektor; eine breite Operationalisierung außerhalb spezifischer Einsätze im öffentlichen Dienst bleibt jedoch eine offene Frage, die dieser Vorschlag berücksichtigt, anstatt sie zu beschönigen. Das Te Kāhui Raraunga -KI-Governance-Framework Māori besagt, dass „KI-Systeme in Aotearoa nicht implementiert werden dürfen, ohne die Autorität Māori über Māori vollständig zu verwirklichen“; dieser Vorschlag hebt diese Anforderung nicht auf. [ZITAT: Te Kāhui Raraunga Charitable Trust. Māori: Tuia te korowai o Hine-Raraunga,hui; Māori AI Governance Framework,hui; vollständige bibliografische Angaben zu datierten Veröffentlichungen stehen nach Überprüfung der Primärquellen noch aus.]

Dr. Karaitiana Taiuruveröffentlichte wissenschaftliche Arbeiten zu ethischen Rahmenwerken Māori für KI, zu tikanga (Māori Recht und -Bräuche) in der KI-Ethik, zu Tiriti-konforme KI sowie zum Schutz von mātauranga (Māori) in KI-Trainingsdaten – einschließlich der oben zitierten kritischen Analyse vom 20. September 2025 – hat eine grundlegende Sprache für das Nachdenken über agente KI im Kontext von te ao Māori bereitgestellt. Wir übernehmen seine bevorzugten Grundbegriffe, wo dies angemessen ist: **mana motuhake** und **rangatira** anstelle vorgeschriebener westlicher Konzeptrahmen; reaktionsfähige und anpassungsfähige Rahmen, die auf tikanga basieren und sich mit technologischem und sozialem Wandel weiterentwickeln können; Rahmen, die auf bestimmte Organisationen und Branchen zugeschnitten sind und in Partnerschaft mit relevanten Māori entwickelt wurden. Wir zitieren seine Arbeit als grundlegende wissenschaftliche Grundlage; wir stellen weder ihn noch irgendjemanden so dar, als würde er diesen konkreten Vorschlag befürworten. Was als angemessener Einsatz intelligenter Agenten im Kontext von te ao Māori gilt, ist von den Tangata Whenua zu bestimmen und nicht in diesem Vorschlag festzulegen.

Lückenanalyse – was dieser Vorschlag leistet und was noch nicht Eine ehrliche Bewertung ist wichtiger als ambitionierte Behauptungen in einem Entwurf der Version 1.1, der an Gutachter, darunter Dr. Taiuru, gerichtet ist. Die architektonischen Grundelemente des Vorschlags – Souveränität durch Zuordnung, kryptografische Herkunftsnachweise, Mitgliederportabilität, bilaterale Föderation – sind**vereinbar mit der Operationalisierung** Māori über Māori im Rahmen des Te Kāhui Raraunga und Taiurubevorzugten Grundlagen. Zu den Kompatibilitätsaspekten gehören:

- **Die Isolation von Mandanten als grundlegende** (Village Bereitstellungseigenschaft, verankert in der Tractatus zur Grenzdurchsetzung) setzt durch die Architektur

die Anforderung um, dass KI- Systeme nicht implementiert werden dürfen, ohne die Autorität Māori über Māorizu berücksichtigen. Ein Mandant, der von einem hapū, iwi oder kaitiaki betrieben wird, verwahrt seine Datensätze gemäß der Konzeption des Frameworks unter der Autorität dieses Gremiums und nicht nur aufgrund einer Zusage.

- **Die Durchsetzung von Grenzen** kann so konfiguriert werden, dass sie eine ausdrückliche Kaitiaki-/Hapū-/Iwi-Genehmigung für wertkritische Vorgänge an den Datensätzen des Mandanten erfordert; das Framework setzt dies durch Architektur statt durch Vertrauen durch.
- **Kryptografische Provenienz und mitgliederportable Identifikatoren** unterstützen kaitiakiüber Generationen hinweg: Datensätze verfügen über einen eigenen Prüfpfad, können nicht unbemerkt verändert werden, und Mitglieder können zu einer anderen souveränen Installation unter derselben Architektur migrieren.
- **Die Primitive für pluralistische Deliberation** ist für moralische Deliberation über mehrere Frameworks hinweg konzipiert, wenn die Durchsetzung von Grenzen einen Wertekonflikt signalisiert; sie ist grundsätzlich in der Lage, kaupapa Māori Frameworks neben anderen Frameworks in einer strukturierten Deliberation zu vereinen.

Was dieser Vorschlag **noch nicht leistet** und was die Gutachter entsprechend abwägen sollten, ist ebenso wichtig zu benennen:

- **Mana motuhake und rangatiraals grundlegende philosophische Grundlage.** Der grundlegende Pluralismus des Tractatus ist selbst ein westliches philosophisches Bekenntnis, das sich auf Berlin, Rawls und Ostrom stützt; er berücksichtigt kaupapa Māori als einen Framework unter vielen; er gründet nicht auf Mana motuhake und rangatiraals vorrangige Verpflichtungen. Um diese Lücke zu schließen, müssten die Grundlagen des Rahmens von einem kaupapa Māori aus neu formuliert werden – eine umfangreiche Arbeit, die ehrlich gesagt nicht von den vorliegenden Autoren allein geleistet werden kann.
- **Eine von Indigenen geleitete Designpartnerschaft.** Das Tractatus Framework und die Village wurden vom Geschäftsführer (Pākehā) von My Digital Sovereignty Ltd entwickelt, mit anschließenden Beiträgen von Claude (Anthropic) als Autor. Sie wurden nicht gemeinsam mit Māori Interessengruppen entworfen. TaiuruEmpfehlung für „auf bestimmte Organisationen und Branchen zugeschnittene Rahmenwerke, die in Partnerschaft mit relevanten Māori entwickelt wurden“ wird auf der Ebene des Designprozesses nicht erfüllt. Die Architektur steht für die Anwendung in Partnerschaft *zur Verfügung*; das Rahmenwerk selbst wurde nicht in Partnerschaft mitentwickelt.
- **KI-Voreingenommenheit in kultureller und ethnischer Hinsicht auf der Ebene der Trainingsdaten.** Die Validierung durch Querverweise erfasst Überschreibungen von Trainingsmustern auf der Ebene der Anweisungskonflikte; sie behebt jedoch nicht die tiefer liegenden Voreingenommenheiten, die in den Trainingsdaten verankert sind und nicht als Anweisungskonflikte zutage treten würden. Die begleitende Arbeit in Paper B zu Situated Language Layers (trainingsspezifische Disziplin pro Mandant, Haltung ohne Gewichtsmodifikation, jurisdiktionsgebundene Inferenz) setzt sich direkter mit diesen Bedenken auseinander als der Tractatus.

- **Digitaler Kolonialismus als benanntes theoretisches und politisches Konzept. Die** im Vorschlag vorgesehene Mandantenisolierung und architektonische Souveränität sind teilweise strukturelle Antworten auf den digitalen Kolonialismus; der Vorschlag setzt sich nicht theoretisch mit dem Konzept auseinander. Das Whitepaper „Distributive Equity“ (mit Querverweis von dieser Website) behandelt dies expliziter als der Tractatus.
- **Umweltauswirkungen von KI** werden weitgehend nicht behandelt. Die CPU-Fallback-Inferenzarchitektur im Village Einsatz ist eine teilweise operative Antwort; sie ist nicht Teil der erklärten Verpflichtungen des Frameworks.

Die ehrliche Schlussfolgerung aus dieser Lückenanalyse lautet, dass der Vorschlag die architektonischen Grundelemente bietet, die für die Operationalisierung der Maori-Autoritätüber Maori-Daten erforderlich wären, wobei jedoch anerkannt wird, dass die Operationalisierung im Kontext von te ao Māori ein wesentliches, eigenständiges Unterfangen ist, das einevon Kaupapa Māori geleitete Entwurfsarbeit erfordert, die dieser Vorschlag nicht geleistet hat. Der Vorschlag des Ausschusses in §II Punkt 4 ist ein Mechanismus, durch den diese weitere Arbeit vorangetrieben werden könnte; er wird zur Prüfung und nicht als vollständige Antwort angeboten.

Die Algorithmus-Charta für Aotearoa New Zealand, die 2020 von Behörden der Krone unterzeichnet wurde, bildet die bestehende Grundlage für Transparenz, Partnerschaft mit Māori, Fairness, Rechenschaftspflicht und Datenschutz bei algorithmischen Entscheidungsprozessen der Krone. Dieser Vorschlag ersetzt nicht die Algorithmus-Charta; die folgenden Empfehlungen sollen innerhalb und parallel zu dieser sowie zu den Rahmenwerken KāhuiRaraunga umsetzbar sein. [QUELLE: Algorithmus-Charta für Aotearoa New Zealand (2020). https://www.data.govt.nz/leadership/governance/data-ethics/algorithm-charter/ – aktueller Stand und etwaige spätere Aktualisierungen vorbehaltlich der Überprüfung.]

(iv) Die globale Tradition der indigenen Datenhoheit

Indigene Datenhoheit ist eine internationale Bewegung, keine neuseeländische Besonderheit. Die Nennung der internationalen Tradition ist wichtig: Sie stellt die oben beschriebene, auf dem Te Tiriti basierende Arbeit in einen globalen Kontext und nicht als engstirnigen Lokalismus dar, und sie schafft eine gemeinsame Basis mit den Autoren des CAC-Rahmenwerks als Mitwirkende an nicht-westlichen Konzepten zur Regulierung von Daten und KI.

Das **First Nations Information Governance Centre** (FNIGC) in Kanada wendet die **OCAP-Prinzipien** an – Ownership, Control, Access, Possession (Eigentumsrecht, Kontrolle, Zugang, Besitz) –, die ursprünglich in den 1990er Jahren im Rahmen der First Nations Regional Longitudinal Health Survey formuliert wurden und heute in der Forschungsethikpraxis an kanadischen Universitäten, in Regierungsbehörden und in First Nations-Gemeinschaften verankert sind. [ZITAT: First Nations Information Governance Centre. The First Nations Principles of OCAP®. https://fnigc.ca/ocap-training/]

Das **United States Indigenous Data Sovereignty Network** (USIDSN), das 2016 in Verbindung mit dem Native Nations Institute an der University of Arizona gegründet

wurde, hat die Praxis der indigenen Datenhoheit im US-amerikanischen Kontext vorangetrieben, unter anderem durch die Einbindung in datenpolitische Prozesse der US-Bundesregierung. [QUELLE: United States Indigenous Data Sovereignty Network. https://usindigenousdata.org/]

Das **Maiam nayri Wingara Indigenous Data Sovereignty Collective** in Australien – der Name bedeutet „Viele Stimmen, ein Geist“ – wurde 2017 gegründet und veröffentlichte 2018 ein Kommuniqué zur Datenhoheit indigener Völker, das die australische Praxis im Bereich indigener Daten geprägt hat. [QUELLE: Maiam nayri Wingara Indigenous Data Sovereignty Collective. (2018). Indigenous Data Sovereignty Communiqué.]

Die **Global Indigenous Data Alliance** (GIDA) koordiniert diese und andere nationale Netzwerke zur indigenen Datenhoheit auf internationaler Ebene; unter ihrer Schirmherrschaft wurden die CARE-Prinzipien veröffentlicht. [ZITAT: Global Indigenous Data Alliance. https://www.gida-global.org/]

Dass ein Großteil der philosophischen Grundlagenarbeit in diesem Vorschlag auf indigene Wissenschaft zurückgeht, ist kein Zufall. Die immer wiederkehrenden Fragen – unter wessen Autorität agieren Daten und die damit arbeitenden Akteure? Wem sind Rechenschaftspflicht und Herkunft nachzuweisen? Was ist der angemessene Maßstab, an dem kollektive Interessen gegen individuelle abgewogen werden? – sind Fragen, mit denen sich Indigenous Data Sovereignty seit Jahrzehnten beschäftigt. Der Widerstand gegen ausbeuterische Big-Tech-Architekturen und die Formulierung alternativer Architekturkonzepte, die auf kollektiver Autorität beruhen, ist einer der fruchtbarsten Beiträge dieser internationalen Bewegung. Die folgenden Empfehlungen knüpfen an diese Tradition an und richten sich im Dialog an sie.

(v) ISO/IEC JTC 1/SC 42: die internationale Landschaft der KI-Standards

Die internationale KI-Normungsarbeit, koordiniert durch ISO/IEC JTC 1/SC 42, liefert das formale Vokabular und die Rahmenwerke für Managementsysteme, in denen Empfehlungen dieser Art in der organisatorischen Praxis umsetzbar werden. Vier Normen sind dabei besonders relevant.

ISO/IEC 22989:2022 legt die Konzepte und Terminologie der künstlichen Intelligenz fest. Wir verwenden die Terminologie von ISO/IEC 22989 soweit kompatibel – beispielsweise hat der Begriff „KI-System“ die in 22989 festgelegte Definition. Die terminologische Konsistenz macht diesen Vorschlag für normenorientierte Prüfer lesbar und ermöglicht die Umsetzung neben anderen an 22989 ausgerichteten Arbeiten. [ZITAT: ISO/IEC 22989:2022. Informationstechnologie — Künstliche Intelligenz — Konzepte und Terminologie der künstlichen Intelligenz. Internationale Organisation für Normung / Internationale Elektrotechnische Kommission.]

ISO/IEC 23053:2022 legt einen Rahmen für KI-Systeme fest, die maschinelles Lernen nutzen, und bildet die Komponenten eines auf maschinellem Lernen basierenden KI-Systems sowie die Beziehungen zwischen ihnen ab. Empfehlungen in diesem Vorschlag, die den Lebenszyklus, die Herkunft oder die Zertifizierung auf Komponentenebene betreffen, können parallel zu den Lebenszyklusphasen der Norm

23053 umgesetzt werden. [ZITAT: ISO/IEC 23053:2022. Rahmenwerk für KI-Systeme (Artificial Intelligence), die maschinelles Lernen (ML) nutzen. ISO/IEC.]

ISO/IEC 23894:2023 bietet Leitlinien zum KI-Risikomanagement. Sie ist das Pendant der Normungsorganisation zu den Empfehlungen des Rahmens zur risikobezogenen Überwachung und zum Vorfallmanagement pro Installation. [ZITAT: ISO/IEC 23894:2023. Informationstechnologie – Künstliche Intelligenz – Leitlinien zum Risikomanagement. ISO/IEC.]

ISO/IEC 42001:2023 legt Anforderungen an ein KI- Managementsystem fest. Es ist das KI-Pendant zu ISO/IEC 27001 (Informationssicherheitsmanagement) und ISO 9001 (Qualitätsmanagement). Wir positionieren die Empfehlungen in diesem Vorschlag als umsetzbar innerhalb eines Managementsystem nach dem Vorbild von ISO/IEC 42001 umsetzbar; Organisationen, die Teile dieses Vorschlags übernehmen, sind wahrscheinlich solche, die bereits eine an ISO/IEC 42001 ausgerichtete Governance betreiben oder dies planen. [ZITAT: ISO/IEC 42001:2023. Informationstechnologie – Künstliche Intelligenz – Managementsystem. ISO/IEC.]

An der Ausschussarbeit zur Erstellung dieser Normen sind nationale Spiegelausschüsse in zahlreichen Ländern beteiligt, darunter das Vereinigte Königreich (über die British Standards Institution) und andere nationale Normungsgremien weltweit. Die Beteiligung Aotearoa New Zealand an der Arbeit des SC42 – über Standards New Zealand oder einen zu diesem Zweck gebildeten Spiegelausschuss – ist einer der Orte, an denen der in diesem Vorschlag befürwortete konstruktive Beitrag auf natürliche Weise erfolgen würde. [HINWEIS: Die Existenz eines aktuellen neuseeländischen SC42-Spiegelausschusses muss vor der Ausarbeitung der Entwürfe für die Punkte 4, 12, 14, 35 und 38 überprüft werden.]

Schluss

Diese fünf Strömungen – Tractatus, CARE, die auf dem Te Tiriti basierende Forschung zur indigenen Datenhoheit, die globale Bewegung für indigene Datenhoheit und ISO/IEC SC42 – laufen in den architektonischen Entscheidungen zusammen, die im weiteren Verlauf dieses Vorschlags dargelegt werden. Souveränität als Zuschreibung; bilaterale Föderation als Koordination; polyzentrische Governance als Autoritätsstruktur; kryptografische Provenienz als Audit-Infrastruktur: Keines dieser Elemente wurde für diesen Vorschlag neu erfunden. Jedes hat seine Wurzeln in einer oder mehreren der oben genannten Traditionen. Der Beitrag dieses Vorschlags besteht in einer bestimmten Anordnung dieser Grundelemente, angepasst an den Kontext Aotearoa New Zealand und angeboten als konstruktive Parallele zu dem Rahmenwerk, mit dem er seine Struktur teilt.

---

§I. Grundprinzipien

Wir schlagen vier Grundprinzipien für die souveräne und föderierte Entwicklung intelligenter Agenten in Aotearoa New Zealand vor. Jedes entspricht einem der vier Prinzipien, mit denen die Cyberspace Administration of China *ihre*

Umsetzungsrichtlinien für 2026 einleitet; in jedem Fall bekräftigen wir die dem Prinzip zugrunde liegende Absicht und bieten eine konstruktive Parallele, die auf den in §0 dargelegten Grundlagen beruht.

Souveränität und Zuordnung. Jeder Vorgang eines intelligenten Agenten in Bezug auf einen Datensatz ist einem souveränen Inhaber dieses Datensatzes zuzuordnen; die Herkunft ist kryptografisch gesichert; Sicherheit ergibt sich aus der Autorität des Datensatzinhabers über seine eigenen Datensätze. Wir bekräftigen das Bekenntnis des CAC-Rahmenwerks zu Sicherheit und Kontrollierbarkeit als grundlegend. Wir schlagen als konstruktive Parallele vor, dass im Kontext Aotearoa New Zealand – wo die Te Tiriti-Partnerschaft, das bestehende Rahmenwerk des Privacy Act 2020 und die CARE-Verpflichtung zur Kontrollhoheit zusammenlaufen – die auf Zuordnung basierende Souveränität gut geeignet ist, um genau diese Sicherheitsanliegen umzusetzen. Die Tractatus zur Durchsetzung von Grenzen bietet den architektonischen Mechanismus; kryptografisch signierte Datensätze liefern den Prüfpfad; und die legitime Autorität über beides liegt beim Inhaber der Datensätze, aufgrund der Rechtshoheit und der Partnerschaftsverpflichtungen. *(Parallelen zu CAC §I Prinzip 1 „Sicherheit und Kontrollierbarkeit“.)* [QUELLENANGABEN: Tractatus (Stroh 2026, CC BY 4.0); CARE-Prinzipien, Verpflichtung zur Kontrollbefugnis (Carroll et al. 2020); Datenschutzgesetz 2020 (NZ), Grundsätze zum Schutz personenbezogener Daten.]

Bilaterale und föderierte Zusammenarbeit. Die Koordination zwischen souveränen Einrichtungen erfolgt durch bilaterale Föderation und offene internationale Standards. Wir erkennen den Wert des Engagements des CAC-Rahmenwerks für eine standardisierte und geordnete Entwicklung an; Standardisierung und Ordnung sind notwendige Voraussetzungen für jeden groß angelegten Einsatz von agentischer KI, und das koordinierte Standardisierungsprogramm des CAC-Rahmenwerks ist ein glaubwürdiger Ansatz. Wir schlagen für den Aotearoa Neuseeland – kleinerer Maßstab, gut etablierte Prinzipien Māori , bestehende bilaterale institutionelle Vereinbarungen zwischen staatlichen Behörden, hapū, iwi, zivilgesellschaftlichen Organisationen und dem privaten Sektor – einen föderierten Ansatz zur Koordination als gut geeignet an. Die Föderation zwischen souveränen Einrichtungen wird durch bestehende, an W3C, IETF und ISO/IEC SC42 angepasste Standards gut unterstützt. Wir schlagen eine bilaterale Föderation zur Prüfung als parallele Architektur vor, die mit zentralen Registrierungsansätzen in anderen Rechtsräumen interoperabel sein könnte, und laden zur Bildung von Ausschüssen unter geeigneten Dachorganisationen ein, um den Interoperabilitätsdialog voranzutreiben. *(Parallelen zu CAC §I Grundsatz 2 „standardisierte und geordnete Entwicklung“.)* [QUELLENANGABEN: W3C Decentralized Identifiers (DIDs) v1.0 (W3C-Empfehlung, 2022) und W3C Verifiable Credentials Data Model v1.1; ActivityPub (W3C-Empfehlung, 2018); ISO/IEC 42001:2023 Managementsysteme; Te Kāhui Raraunga Māori AI Governance Framework + Taiuru kritische Analyse (siehe §0(iii)).]

Pluralistische Deliberation, polyzentrisch. Mehrere Wertegerüste koexistieren innerhalb und über souveräne Strukturen hinweg; die Deliberation zwischen ihnen ist prozedural und strukturiert; Innovation entsteht aus lokaler Anpassung unter lokaler Autorität. Wir bekräftigen das Bekenntnis des CAC-Rahmenwerks zu einer innovationsgetriebenen Entwicklung. Wir schlagen als konstruktive Parallele vor,

dass polyzentrische Governance – mehrere Autoritätszentren, mehrere Wertesysteme, die in produktiver Spannung zueinander stehen, mit strukturierter Deliberation bei auftretenden Konflikten – gut zum Kontext der Te Tiriti-PartnerschaftAotearoa New Zealand passt und durch die internationale Forschung zu polyzentrischer Governance (insbesondere Elinor Ostroms wegweisende Arbeit) gut gestützt wird. Die Tractatus Primitive der pluralistischen Deliberation bietet den architektonischen Mechanismus zur Ermöglichung einer Multi-Stakeholder-Deliberation, wenn die Durchsetzung von Grenzen einen Wertekonflikt aufzeigt; der fundamentale Pluralismus ist das philosophische Bekenntnis, das dies zu einem strukturellen Merkmal des Rahmenwerks macht. *(Parallelen zu CAC §I Prinzip 3 „innovationsgetriebene Entwicklung“.)* [QUELLENANGABEN: Tractatus Primitiv für pluralistische Deliberation (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 – vollständige bibliografische Angaben müssen vor der Veröffentlichung von v1 überprüft werden.]

Anwendungsorientiert, belegt. Der Einsatz intelligenter Agenten wird durch die Implementierung in Gemeinschaften belegt, die diese übernommen haben; für einen zivilgesellschaftlichen Vorschlag ist die angemessene Beweisgrundlage der Einsatz in der Praxis. Wir bekräftigen das Bekenntnis des CAC-Rahmenwerks zu anwendungsorientierter Entwicklung. Wir schlagen als konstruktive Parallele vor, dass bei einem zivilgesellschaftlichen Vorschlag, der von einem einzelnen Unternehmen stammt, der Nachweis des Einsatzes der Empfehlung vorausgehen muss. Wo dieser Vorschlag Beispiele für den Einsatz in Aotearoa Neuseeland anführt (in §IV und §V) – im Kontext von Gemeinden und Hapū/Iwi, im Kontext von Iwi und Familiengeschichte in der Diaspora, im Kontext kleiner Unternehmen – beziehen sich diese Verweise auf tatsächliche Einsätze, wobei konkrete Einsatzdaten (Anzahlen, Startdaten, Umfang) vor der Veröffentlichung von Version 1 hinzugefügt werden müssen. Wo der Vorschlag Empfehlungen für Sektoren vorbringt, in denen MDSL noch nicht eingesetzt wurde, sind diese Empfehlungen als Bedingungen der Souveränitätsarchitektur für jeden Agenten-Einsatz in diesem Sektor formuliert und richten sich an jeden, der die Architektur dort anwenden möchte. Was der Einsatz belegt, ist nicht nur die Anzahl der Einführungen, sondern die architektonische Verfügbarkeit – dass die Substrat-Primitiven (kryptografische Herkunft, Föderations-Envelopes, mitgliedergeführte Portabilität, Durchsetzung von Grenzen) wie spezifiziert unter realen Bedingungen mit den Daten der Gemeinschaften funktionieren, die den Test autorisiert haben, und nicht in künstlichen Benchmarks. *(Parallelen zu CAC §I Prinzip 4 „anwendungsorientierter Ansatz“.)* [ZITATE: MDSL-Einsatznachweise – Village (Gemeinde- und Gemeinschaftskontexte), Familiengeschichte (iwi- und Diaspora-Kontexte), sydigital (Kontexte kleiner Unternehmen); spezifische Einsatzdaten (Zahlen, Startdaten, Umfang der Nutzung) stehen noch aus, bis die vom Betreiber verifizierten Zahlen vor der Veröffentlichung von v1 vorliegen.]

---

§II. Grundlagen für souveräne Entwicklung

Während das Rahmenwerk der Cyberspace Administration of China technologische Grundlagen unter einem staatlich koordinierten Standardisierungsprogramm konsolidiert, bieten wir Grundlagen, die in kryptografischer Souveränität und bilateralen Protokollen verwurzelt sind. Die beiden folgenden Unterabschnitte – Stärkung der Souveränitätsgrundlage und Etablierung bilateraler Protokolle – spezifizieren gemeinsam die architektonischen Grundelemente, auf denen der Rest des Vorschlags aufbaut.

(I) Stärkung der Souveränitätsgrundlage

Punkt 1. Aufbau souveräner Primitive für Agenten. Kryptografisch signierte Datensätze, mitgliederportable Identifikatoren und attribuierte Provenienz bilden die Grundlage für Agentenoperationen an souveränen Daten. Es handelt sich um architektonische Primitive, die Souveränität auf der Ebene des Datensatzes selbst konstituieren. Wir schlagen nachhaltige Investitionen in Open-Source-Kryptografie-Grundbausteine vor – digitale Signatur, überprüfbare Berechtigungsnachweise, inhaltsadressierter Speicher mit Herkunftsnachweis – sowie in Standards für übertragbare Identitäten, die in jeder souveränen Installation in jedem Sektor einsetzbar sind. Wir schlagen vor, dass diese Primitive als gemeinsame Infrastruktur entwickelt und gewartet werden, die unter freizügigen Open-Source-Lizenzen (Apache 2.0, EUPL-1.2 oder kompatibel) verfügbar ist und die Übernahme, Änderung und Weiterverbreitung durch jede Partei erlaubt. Die Tractatus Primitive zur Durchsetzung von Grenzen, die Primitive zur Validierung von Querverweisen und die Primitive zur Klassifizierung der Persistenz von Anweisungen legen gemeinsam die Laufzeitmechanismen fest; die Infrastruktur für kryptografische Signaturen und verifizierbare Zugangsdaten liefert den zugrunde liegenden Prüfpfad. *(Parallelen zu CAC-Punkt 1 „Stärkung von Forschung und Entwicklung im Bereich grundlegender Technologien“.)* `[QUELLENANGABEN: Tractatus (Stroh 2026, CC BY 4.0 Text / Apache 2.0 Code); W3C Decentralized Identifiers (DIDs) v1.0 (W3C-Empfehlung, 2022); W3C Verifiable Credentials Data Model v1.1; CARE-Prinzipien, Verpflichtung zur Kontrollhoheit (Carroll et al. 2020).]`

Punkt 2. Verfeinerung der souveränen Toolchain. Open-Source- Referenzimplementierungen von Agenten-Frameworks – einschließlich der sechs Dienste des Tractatus Frameworks – sollten für die Übernahme durch jede souveräne Installation unter freizügigen Open-Source-Lizenzen verfügbar sein, die einen installationslokalen Betrieb erlauben. Wir schlagen vor, dass die Toolchain für die Entwicklung, das Testen, die Bereitstellung und die Wartung von souveränitätsorientierten Agentensystemen offen entwickelt wird, wobei Beiträge von jeder souveränen Installation gefördert werden. Die aktuellen MDSL-Implementierungen – das Tractatus, das unter Apache 2.0 vertrieben wird (mit Dokumentation unter CC BY 4.0); die Code-Basen Village und Community, die ab Mitte 2026 schrittweise auf EUPL-1.2 migrieren – werden als eine von potenziell mehreren Referenzimplementierungen angeboten. Sicherheitswerkzeuge – Erkennung feindlicher Eingaben, Erkennung von Verhaltensanomalien, Attestierungswerkzeuge für Builds und Abhängigkeiten – sind die geeignete technische Ergänzung zu den Tractatus zur Grenzdurchsetzung und metakognitiven Verifizierung. *(Parallelen zu CAC-Punkt 2 „Verfeinerung der*

Agent-Toolchain“.) [QUELLENANGABEN: Tractatus (Stroh 2026), Apache 2.0 (Code), CC BY 4.0 (Text und Abbildungen); EUPL-1.2 (European Union Public Licence); Apache 2.0 (Apache Software Foundation).]

(II) Einrichtung bilateraler Protokolle

Punkt 3. Verbundene bilaterale Protokolle. Die Interoperabilität zwischen souveränen Einrichtungen erfolgt durch bilaterale Vereinbarungen und offene internationale Standards. Wir erkennen den Wert des Engagements des CAC-Rahmenwerks für ein standardisiertes Verbindungsprogramm an – das vorgeschlagene Intelligent Agent Interconnection Protocol (AIP), grundlegende Schnittstellenstandards für Software, Dienste und Hardware-Peripheriegeräte sowie verbindliche Standards in sensiblen Sektoren. Wir schlagen für den Kontext Aotearoa New Zealand vor, dass die Interoperabilität zwischen souveränen Einrichtungen durch die bestehende internationale Standardlandschaft gut unterstützt wird: W3C Decentralized Identifiers und Verifiable Credentials für die Identität; ActivityPub und verwandte W3C-Föderationsprotokolle für die Kommunikation zwischen Einrichtungen; IETF-Protokolle für Authentifizierung, Transport und Inhaltsadressierung; sowie die Arbeit der ISO/IEC SC42 zu KI-spezifischer Terminologie, Lebenszyklus, Risiko und der Angleichung von Managementsystemen. Wir schlagen vor, dass Aotearoa NZ als gleichberechtigter Teilnehmer in diesen bestehenden Foren zu internationalen Interoperabilitätsstandards beiträgt. *(Parallelen zu CAC-Punkt 3 „Standardisierungssystem“ und dem vorgeschlagenen AIP-Verbindungsprotokoll.)* [QUELLENANGABEN: W3C DIDs v1.0; W3C Verifiable Credentials Data Model v1.1; ActivityPub (W3C-Empfehlung, 2018); ISO/IEC 22989:2022 Terminologie; ISO/IEC 23053:2022 ML-Framework.]

Punkt 4. Kryptografische Identität; föderierter Dialog über das Intelligente Internet. Die Identität ist installationsspezifisch, verankert in DNS und kryptografischen Schlüsseln; die Verifizierung zwischen den Parteien erfolgt Peer-to-Peer; Fähigkeitserklärungen werden von jeder Installation veröffentlicht. Wir erkennen den Wert des Vorschlags des CAC-Rahmenwerks für eine Registrierungsplattform für intelligente Agenten an, der nicht nur die Verwaltung digitaler Identitäten und die Angabe von Fähigkeiten, sondern auch Suche und Erkennung, vertrauenswürdige Vernetzung, konforme Zahlungen, Sicherheitsschutz, Konfliktlösung, Nutzung von IPv6 und ein Überwachungsindikatorensystem umfasst – eine umfangreiche und kohärente Reihe miteinander verbundener Funktionen. Eine zentralisierte Registrierungsplattform mit einer koordinierenden Behörde ist ein glaubwürdiger architektonischer Ansatz für diese Funktionen.

Wir schlagen für den Kontext Aotearoa New Zealand – wo kleinere, etablierte Prinzipien Māori und die architektonischen Grundelemente, die bereits in MDSL-Implementierungen vertreten sind, zusammenlaufen – einen föderierten Ansatz vor, bei dem jede Intelligent-Internet-Funktion durch bilaterale Vereinbarungen und offene internationale Standards geregelt wird. Identität und Fähigkeitserklärung werden durch dezentrale Identifikatoren und überprüfbare Berechtigungsnachweise des W3C gewährleistet. Die Suche und Erkennung zwischen souveränen Installationen kann auf den Mustern aufbauen, die durch die von ActivityPub abgeleitete Föderation, durch WebFinger (IETF RFC 7033) und durch föderationsfähige Verzeichnisprotokolle

wie nodeinfo etabliert wurden – wobei wir anmerken, dass die föderierte Erkennung in großem Maßstab ein offenes technisches Problem bleibt, und dies als solches anerkennen. Vertrauenswürdige Vernetzung und Sicherheitsschutz erfolgen durch bilaterale kryptografische Beglaubigung. Konforme Zahlungswege über bestehende finanzregulatorische Kanäle. Die Konfliktlösung erfolgt durch bilaterale Mediation und bestehende Streitbeilegungsmechanismen, wobei die kryptografische Herkunft den Prüfpfad liefert. IPv6 ist eine zugrunde liegende Infrastrukturwahl, die jeder Installation zur Verfügung steht. Ein Überwachungsindikatorensystem ist durch die offene Veröffentlichung von Betriebskennzahlen durch jede teilnehmende Installation realisierbar, aggregiert durch unabhängige Beobachter.

Wir schlagen die Bildung eines einzigen Ausschusses unter einer geeigneten Dachorganisation vor – Kandidaten sind unter anderem die Royal Society Te Apārangi, das Spiegelkomitee SC42 von Standards New Zealand (Existenz zu überprüfen), das New Zealand AI Forum oder eine gemeinsame Struktur über diese hinweg – **um detaillierte Empfehlungen zur Architektur agentischer KI im neuseeländischen Kontext zu entwickeln, als gleichberechtigter Teilnehmer zur Arbeit von ISO/IEC JTC 1/SC 42 beizutragen und einen bilateralen Dialog mit den Autoren des CAC-Rahmenwerks sowie mit internationalen Kollegen zu führen. Der Ausschuss würde fünf benannte Arbeitsstränge bearbeiten: (i) föderierte Identität für intelligente Agenten und die in diesem Punkt genannten weiter gefassten Funktionen des intelligenten Internets; (ii) föderierte Audit- und Compliance-Dienste (siehe §III Punkt 12); (iii) bescheinigungsbasierte Reputationssysteme (siehe §III Punkt 14); (iv) Muster der Branchenkoordination, einschließlich Föderations- versus Allianzmodelle (siehe §V Punkt 35); und (v) internationales Engagement und bilaterale Zusammenarbeit im Bereich der agentenbasierten KI (siehe §V Punkt 38). Der Beitrag des Ausschusses zur internationalen Normungsarbeit und zum Dialog mit den Autoren des CAC-Rahmenwerks ist sein Hauptprodukt.** Wir legen diesen Ausschussvorschlag als einen Beitrag zur internationalen Diskussion vor; die Diskussion wird von Beiträgen aus vielen verschiedenen Architekturtraditionen profitieren. *(Parallelen zu CAC Punkt 4 „intelligente Internetarchitektur“ mit Registrierungsplattform; das Muster der Ausschussbildung wird in den Punkten 4, 12, 14, 35 und 38 konsolidiert.)* [ZITATE: W3C Decentralized Identifiers (DIDs) v1.0; W3C Verifiable Credentials Data Model v1.1; ActivityPub (W3C-Empfehlung 2018); WebFinger (IETF RFC 7033); nodeinfo-Föderationsverzeichnis; ISO/IEC 22989:2022; Te Kāhui Raraunga (hui.io — Māori und Māori AI Governance Framework); Taiuru, K. (20. September 2025) Kritische Analyse der Te Mana Raraunga, taiuru.co.nz/critical-analysis-mana-raraunga/; Royal Society Te Apārangi; ISO/IEC JTC 1/SC 42.]

---

§III. Wahrung der Souveränitätsgrundlage

Während das Rahmenwerk der chinesischen Cyberspace-Behörde eine Sicherheitsbasislinie durch Produktrichtlinien, Verhaltensbegrenzungstechnologien, mehrstufige Governance und Selbstregulierung der Industrie mit Bonitätsbewertungssanktionen festlegt, bieten wir eine Basislinie an, die im eigenen Rechtsrahmen des Anwenders, in

der kryptografischen Herkunft, in polyzentrischen Governance-Vereinbarungen und in einer föderationsbasierten Koordination verwurzelt ist. Die vier folgenden Unterabschnitte – Produktgrundsätze, Sicherheitsrisiken, Governance-System, föderierte Koordination – legen gemeinsam fest, wie die Einhaltung der Souveränitätsgrundsätze durch einen intelligenten Agenten zur Laufzeit überprüft und nachträglich auditiert werden kann.

(I) Klärung der Produktgrundsätze

Punkt 5. Verankerung in den eigenen Rechtsvorschriften des Anwenders. Richtlinien, Vorschriften und ethische Standards für intelligente Agenten ergeben sich aus der Rechtsordnung des Anwenders. Die Werte leiten sich aus dem lokalen Recht und lokalen institutionellen Regelungen ab; die Architektur stellt die Umsetzungsinfrastruktur bereit, in der diese Werte zum Tragen kommen. In Aotearoa New Zealand gehören zu den anwendbaren Rechtsinstrumenten das Datenschutzgesetz 2020 (zusammen mit dem Health Information Privacy Code 2020 und anderen Kodizes, soweit diese für bestimmte Sektoren gelten); das New Zealand Bill of Rights Act 1990, sofern staatliche Akteure beteiligt sind; die Algorithm Charter for Aotearoa New Zealand für Behörden der Krone; die Verpflichtungen aus Te Tiriti o Waitangi für Akteure der Krone und die damit verbundenen Partnerschaftsverpflichtungen; das Official Information Act 1982; der Public Service Act 2020; der Public Records Act 2005; sowie sektorale Gesetze, darunter der Reserve Bank of New Zealand Act 2021, der Education and Training Act 2020, der Local Government Act 2002 und der Search and Surveillance Act 2012, die für den jeweiligen Einsatzkontext gelten. Die Architektur ist implementierungsneutral hinsichtlich der Frage, welches Recht gilt; der Vorschlag richtet sich an Anwender Aotearoa Neuseeland, und dieselben Grundelemente dienen Anwendern in jeder Rechtsordnung, deren Werte sie umsetzen möchten. *(Parallelen zu CAC-Punkt 5 „Richtlinien, Vorschriften und ethische Standards“.)* [QUELLENANGABEN: Datenschutzgesetz 2020 (NZ); Neuseeländisches Grundrechtegesetz 1990; Algorithmus-Charta für Aotearoa New Zealand (2020); Health Information Privacy Code 2020; Official Information Act 1982; Public Service Act 2020; Public Records Act 2005; Reserve Bank of New Zealand Act 2021; Education and Training Act 2020; Local Government Act 2002; Search and Surveillance Act 2012 – aktuelle Gesetzesfassungen sind vor der Veröffentlichung von Version 1 zu überprüfen.]

Punkt 6. Endgültige Entscheidungsbefugnis des Nutzers, kryptografisch gesichert. Wir bekräftigen denselben Grundsatz, den das CAC-Rahmenwerk bekräftigt: Der Nutzer behält das Recht, über autonome Handlungen, die von intelligenten Agenten in seinem Namen vorgenommen werden, informiert zu werden und die endgültige Entscheidungsbefugnis darüber zu haben. Dieser Grundsatz ist grundlegend für das Vertrauensverhältnis zwischen einer Person und den in ihrem Namen handelnden agentenbasierten Systemen. Wir schlagen als Prüfmechanismus eine kryptografische Provenienz pro Datensatz gegenüber dem eigenen souveränen Datensatz des Nutzers vor: Jede autonome Handlung eines Agenten, der auf die Datensätze des Nutzers einwirkt, erzeugt einen kryptografischen Eintrag, der die Handlung bestätigt und dem Agenten sowie

dem Autorisierungsrahmen des Nutzers zugeordnet werden kann. Der Nutzer kann jede Agentenhandlung anhand dieser Herkunft prüfen, wiedergeben und anfechten, und die Tractatus „Anweisungs-Persistenz-Klassifizierung“ bietet den Rahmen zur Unterscheidung zwischen Routinehandlungen und solchen, die eine ausdrückliche erneute Bestätigung durch den Nutzer erfordern. *(Parallelen zu CAC-Punkt 6 „Klärung der Entscheidungsbefugnis“ .)* [ZITATE: Tractatus zur Klassifizierung der Befehlsbeständigkeit (Stroh 2026, CC BY 4.0); Privacy Act 2020 (NZ), Datenschutzgrundsatz 6 (Zugriffsrechte); CARE-Prinzipien, Verpflichtung zur Kontrollbefugnis (Carroll et al. 2020).]

Punkt 7. Provenienz als Ergänzung zur Verhaltenskontrolle. Wir würdigen die Betonung des CAC-Rahmenwerks auf Regel- Einbettung, Verhaltensbegrenzung und Blockchain-verankerte Verifizierung des Agentenverhaltens in kritischen Anwendungsszenarien. Dies sind glaubwürdige architektonische Ansätze zur Gewährleistung rechtmäßigen und konformen Verhaltens in zentral koordinierten Bereitstellungen. Wir schlagen als zusätzliche architektonische Grundkomponente, die sich gut für bilaterale Verbünde eignet, **die Provenienz** vor: Jede Aktion eines intelligenten Agenten erzeugt einen kryptografischen Datensatz, der dem Akteur zugeordnet werden kann. Die beiden Ansätze ergänzen sich. Verhaltensbegrenzung schränkt ein, was ein Agent zur Laufzeit versuchen darf; Provenienz erstellt einen fälschungssicheren Datensatz darüber, was tatsächlich versucht wurde. Beide spielen eine Rolle, und das angemessene Gleichgewicht zwischen ihnen ist wahrscheinlich kontextspezifisch. *(Parallelen zu CAC-Punkt 7 „Verhaltenskontrolle stärken“.)* [QUELLENANGABEN: Tractatus zur Validierung von Querverweisen (Stroh 2026, CC BY 4.0); W3C Verifiable Credentials Data Model v1.1; ISO/IEC 23894:2023 Risikomanagement.]

(II) Minderung von Sicherheitsrisiken Punkt

Punkt 8. Intrinsische Sicherheit durch souveräne Primitive. Personenbezogene Daten verbleiben in der Installation des Inhabers; der kryptografische Schutz erfolgt sowohl auf Datensatzebene als auch perimeterbasiert; die Angriffserkennung läuft lokal gegen die Datensätze des Inhabers; der Zugriff ist vertraglich zwischen den Vertragsparteien geregelt. Der Auswirkungsbereich eines Ausfalls ist auf die betroffene Installation beschränkt. Wir bekräftigen das Bekenntnis des CAC-Rahmenwerks zu intrinsischen Sicherheitsfunktionen – Datensicherheit, Schutz personenbezogener Daten, kryptografischer Schutz, Angriffserkennung, Zugriffskontrolle, Verhaltenskontrolle. Wir schlagen als konstruktive Parallele vor, dass für eine föderierte Architektur der geeignete Ort für diese Funktionen die souveräne Installation ist, mit bilateralen Mechanismen für die Zusammenarbeit zwischen Installationen, wenn Bedrohungen jurisdiktionelle oder organisatorische Grenzen überschreiten. *(Parallelen zu CAC Punkt 8 „intrinsische Sicherheitsfunktionen“.)* [QUELLENANGABEN: Tractatus -Primitiv zur Durchsetzung von Grenzen (Stroh 2026, CC BY 4.0); Privacy Act 2020 (NZ); ISO/IEC 23894:2023 Risikomanagement.]

Punkt 9. Lieferketten-Zertifizierung, föderierter Austausch. Wir schlagen eine anlagenbezogene Zertifizierung über den gesamten Lebenszyklus vor – signierte Build-Herkunft, Abhängigkeitsmanifeste, gegebenenfalls Zertifizierung der Trainingsdaten, Verlauf der Reaktionen auf Sicherheitsvorfälle –, die von jeder

Anlage offen veröffentlicht wird. Vorfälle in der Lieferkette werden bilateral zwischen föderierten Partnern und über etablierte internationale Kanäle ausgetauscht, darunter CERT-NZ, CERT-EU, US-CERT und das CVE-Koordinationssystem. Wir erkennen den Wert des Engagements des CAC-Rahmenwerks für Sicherheitsstandards über den gesamten Lebenszyklus und den Austausch von Informationen zur Lieferkette an. Wir schlagen vor, dass für die föderierte Koordination die Transparenz der Lieferkette durch die offene Veröffentlichung von Bescheinigungen durch jede Installation erreicht wird, bei bilateraler Zusammenarbeit bei der Reaktion auf Vorfälle. *(Parallelen zu CAC-Punkt 9 „Sicherheit der Lieferkette“.)* [ZITATE: ISO/IEC 23894:2023 Risikomanagement; CERT-NZ-Offenlegungsverfahren; internationaler CVE-Koordinierungsprozess; ISO/IEC 42001:2023 Managementsysteme.]

Punkt 10. Begrenzung des Schadensradius; nachträgliche Überprüfung. Die routinemäßige Risikoidentifizierung erfolgt lokal für jede Anlage, wobei anlagenübergreifende Vorfälle sich über den Verbund ausbreiten. Der wesentliche Beitrag des Rahmens zur Minderung des Risikos automatisierter Angriffe, von Datenschutzverletzungen und der Verbreitung falscher Informationen besteht darin, das Ausmaß zu begrenzen, in dem sich automatisierter Schaden vervielfacht. Wir bekräftigen das Bekenntnis des CAC- Rahmenwerks zu Risikoidentifizierung, Frühwarnung, Intervention und Prävention des Einsatzes agentischer KI für illegale Aktivitäten (automatisierte Angriffe, Datenschutzverletzungen, Erzeugung und Verbreitung falscher Informationen, Online-Betrug). Wir schlagen als ergänzenden architektonischen Beitrag vor, dass die Begrenzung des Ausmaßes automatisierter Schäden – durch installationsspezifische operative Grenzen und bilaterale Zusammenarbeit bei der Reaktion auf Vorfälle – eine strukturelle Ergänzung zu Erkennungs- und Interventionsansätzen auf zentraler Ebene darstellt. Der strukturelle Mechanismus ist der **Föderations-Envelope**: Standardmäßig bleiben die Datensätze einer Installation innerhalb dieser Installation, und ein installationsübergreifender Datenfluss erfolgt nur über Envelopes, die die Installation explizit signiert und die Herkunfts- und Empfängerinformationen enthalten. Eine Kompromittierung einer Installation kann sich nicht unbemerkt auf andere ausbreiten, da das Substrat keinen impliziten installationsübergreifenden Lesepfad aufweist – es gibt keine gemeinsame Registrierungsstelle, über die sich ein Angreifer verschieben könnte. Die bilaterale Reaktion auf Vorfälle erfolgt dann auf der Grundlage dessen, was bewusst geteilt wurde, wobei die Herkunftsangaben des Föderations-Envelopes eine forensische Rekonstruktion des betroffenen Umfangs ermöglichen, ohne dass ein zentraler Audit-Aggregator erforderlich ist. *(Parallelen zu CAC-Punkt 10 „Risiken aus Anwendungen mindern“.)* [QUELLENANGABEN: Tractatus pluralistische Deliberation (Stroh 2026, CC BY 4.0); ISO/IEC 23894:2023 Risikomanagement; Privacy Act 2020 (NZ); Harmful Digital Communications Act 2015 (NZ) — aktuelle Gesetzesfassungen sind vor der Veröffentlichung von v1 zu überprüfen.]

(III) Verbesserung des Governance-Systems

Punkt 11. Polyzentrische Governance im Dialog mit abgestuften Ansätzen. Die Governance-Befugnis darüber, was ein intelligenter Agent mit einem Datensatz tun darf, liegt beim Inhaber der Datensätze. Die Zulässigkeit von Szenarien wird pro Installation durch die eigene Zuständigkeit des Inhabers bestimmt, unterstützt

durch sektorale Regulierungsbehörden, sofern deren Zuständigkeit sich auf den jeweiligen Sachverhalt erstreckt. Wir erkennen den Wert des kategorisierten und abgestuften Governance-Ansatzes des CAC-Rahmenwerks für sensible Sektoren und Schlüsselindustrien an, wobei die Cyberspace Administration of China und die zuständigen Branchenbehörden zulässige Anwendungsszenarien festlegen und Verwaltungsmaßnahmen wie die Registrierung, Prüfung und den Rückruf problematischer Produkte umsetzen. Wir schlagen für den neuseeländischen Neuseeland eine polyzentrische Governance vor – mit mehreren Zentren der Autorität, verteilt auf staatliche Behörden, Hapū-/Iwi-Einrichtungen, sektorale Regulierungsbehörden, Berufsverbände und die Inhaber der Datensätze selbst –, die gut zur bestehenden institutionellen Landschaft und zu den Partnerschaftsverpflichtungen aus dem Te Tiriti passt. Die internationale Forschung zur polyzentrischen Governance, insbesondere die wegweisenden Arbeiten von Elinor Ostrom, liefert die theoretische Grundlage für diesen Ansatz. Die Polyzentrizität wird operativ durch die Eigenschaften des Substrats unterstützt: Eine einzige **kryptografisch signierte Audit-Kette** pro Datensatz ermöglicht es mehreren Behörden – staatliche Regulierungsbehörde, Hapū-/Iwi-Einheit, sektorale Stelle, Berufsverband, der Datensatzinhaber selbst –, jeweils denselben Datensatz im Rahmen ihrer jeweiligen Zuständigkeit zu verifizieren, ohne dass eine zentralisierte Zusammenführung oder doppelte Register erforderlich sind. Verschiedene Behörden halten ihre eigenen Kopien der Auditkette unter ihren eigenen Schlüsseln und treffen unabhängige Konformitätsbeurteilungen über denselben zugrunde liegenden Datensatz. Die architektonische Grundkomponente, die polyzentrische Governance operativ handhabbar macht, ist die föderativ replizierte, signierte Auditkette; die institutionelle Frage, wer die Autorität über welche Entscheidungsklasse hat, bleibt politisch. *(Parallelen zu CAC-Punkt 11 „kategorisierte und abgestufte Governance“.)* [ZITATE: 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; Algorithm Charter for Aotearoa New Zealand (2020); Te Kāhui Raraunga (kahuiraraunga.io – Māori-Daten -Governance-Modell und Māori-KI-Governance-Rahmenwerk); Taiuru, K. (20. September 2025) Kritische Analyse der Te Mana Raraunga-Datenprinzipien, taiuru.co.nz/critical-analysis-mana-raraunga/.]

Punkt 12. Verbundene Compliance-Dienste. Risikoüberwachung, Tests, Bewertung, Audit und Zertifizierungsdienste für intelligente Agenten werden als kommerzielle, gemeinschaftliche und akademische Angebote bereitgestellt; die gegenseitige Anerkennung zwischen den Diensten erfolgt durch offene Veröffentlichung und Peer-Review. Wir erkennen den Wert des Engagements des CAC-Rahmenwerks für ein Compliance-Dienstleistungssystem an, das professionelle Dienstleistungen wie Risikoüberwachung, Tests und Bewertung, Beratung und Zertifizierung bereitstellt und die gegenseitige Anerkennung zwischen akkreditierten Anbietern fördert. **Dieser Bereich ist der Arbeitsstrang (ii) des in §II Punkt 4 vorgeschlagenen einheitlichen Ausschusses. Der Ausschuss würde Empfehlungen im neuseeländischen Kontext für ein föderiertes Audit-Rahmenwerk für intelligente Agenten ausarbeiten, zur Arbeit des ISO/IEC SC42 an KI-Bewertungs-, Evaluierungs- und Managementsystemen beitragen und einen bilateralen Dialog mit den Autoren des CAC-Rahmenwerks über die Interaktion zwischen föderierten und zentralisierten Compliance-Diensten führen.** Compliance-Dienste werden

konkret wie folgt föderiert: Jede Installation veröffentlicht ihre eigenen Bescheinigungen – Build-Herkunft, Abhängigkeitsmanifeste, gegebenenfalls Bescheinigungen zu Trainingsdaten, Vorgeschichte der Incident-Response, Einhaltung des Audit-Rahmens – unter ihrer eigenen kryptografischen Identität. Compliance-Anbieter überprüfen diese Bescheinigungen und veröffentlichen ihre Ergebnisse unter ihrer eigenen Identität; die gegenseitige Anerkennung zwischen Anbietern erfolgt durch gegenseitige Verweise auf kryptografisch überprüfbare Bewertungen und nicht durch eine zentrale Akkreditierung. Die grundlegende Komponente, die dies ermöglicht, ist die inhaltsadressierte Veröffentlichung mit kryptografischer Herkunftsnachweisbarkeit – jeder kann jede Compliance-Bewertung anhand der spezifischen Version der Installation überprüfen, die tatsächlich bewertet wurde. *(Parallelen zu CAC-Punkt 12 „Compliance-Service-System“; es gilt der konsolidierte Arbeitsstrang zur Ausschussbildung .)* [ZITATE: ISO/IEC 42001:2023 Managementsysteme; ISO/IEC 23894:2023 Risikomanagement; Royal Society Te Apārangi.]

(IV) Stärkung der föderierten Koordination

Punkt 13. Koordinierung durch Verbände. Souveräne Einrichtungen schließen sich bilateral zu Verbänden zusammen; die Koordinierung bei gemeinsamen Anliegen – Interoperabilitätsstandards, Offenlegung von Sicherheitsvorfällen, Entwicklung von Prüfungsrahmen – erfolgt durch offene Veröffentlichung und Konsens unter den beteiligten Akteuren. Wir würdigen den Wert des Engagements des CAC-Rahmenwerks für die Selbstregulierung der Industrie, bei der Branchenverbände und große Unternehmen gemeinsam Selbstregulierungsregeln formulieren, die die Einhaltung von KI-Funktionalitäten, die Algorithmus-Governance, den Schutz geistigen Eigentums und fairen Wettbewerb abdecken. Wir schlagen vor, dass für die in diesem Vorschlag spezifizierte Verbundarchitektur die Koordination bei gemeinsamen Anliegen durch offene Veröffentlichung und Konsens unter den beteiligten Peers erfolgt; das architektonische Bekenntnis zur bilateralen Verbundbildung erstreckt sich auf den Koordinationsmechanismus selbst. Die Föderation ist in diesem Vorschlag von Natur aus bilateral: Jede Installation veröffentlicht einen Föderations- Endpunkt und wählt aus, mit welchen Peers sie sich zu welchen spezifischen Datensatzklassen zusammenschließt; das Format **des Föderations-Envelopes** legt fest, welche Datensätze mit welchem Zustimmungsumfang an welchen Empfänger und mit welchen Einschränkungen hinsichtlich der Weiterleitung übertragen werden dürfen. Die grundlegenden Elemente, die die bilaterale Föderation funktionsfähig machen, sind der Föderations-Envelope (ein empfängergebundenes, mit Herkunftsangaben versehenes, in seinem Umfang begrenztes Nachrichtenformat), **die mitgliederorientierte Portabilität** (der Inhaber von Datensätzen kann die Weiterleitung von jeder Installation an ein Ziel seiner Wahl verlangen) und**die kryptografische Herkunft** (jeder Datensatz enthält überprüfbare Herkunftsmetadaten, die den Transport überstehen). Die Koordination bei gemeinsamen Anliegen – Interoperabilitätsstandards, Offenlegung von Sicherheitsvorfällen, Entwicklung von Audit-Frameworks – erfolgt bilateral zwischen den Verbündeten, ohne dass eine zentrale Registrierungsstelle erforderlich ist. *(Parallelen zu CAC-Punkt 13 „Selbstregulierung der Branche“.)* [QUELLENANGABEN: ActivityPub-Föderationsprotokoll (W3C-Empfehlung 2018); IETF-Request-for-

Comments-Prozess; W3C-Prozessdokument.]

Punkt 14. Reputation durch Beglaubigung. Souveräne Installationen veröffentlichen ihre eigenen Beglaubigungen – Sicherheitsstatus, Audit- Historie, Abhängigkeitsmanifeste, Reaktion auf Vorfälle – und die Gegenparteien überprüfen diese kryptografisch. Die Reputation entsteht durch die Historie genauer Selbstauskünfte, die von bilateralen Gegenparteien verifiziert wurden. Wir erkennen den Wert des Vorschlags des CAC-Rahmenwerks für freiwillige Bonitätsbewertungsmechanismen für Marktteilnehmer im Bereich intelligenter Agenten an, mit Bonitätsbewertungen für Verhaltensweisen wie den Missbrauch von Technologie, die Anstiftung zum Konsum, falsche Werbung und das Verschweigen von Informationen über Mängel sowie Sanktionen für unehrliches Verhalten in Übereinstimmung mit Gesetzen und Vorschriften. **Dieser Bereich ist der Arbeitsstrang (iii) des in §II Punkt 4 vorgeschlagenen einzigen Ausschusses. Der Ausschuss würde Empfehlungen im neuseeländischen Kontext zu attestbasierter Reputation im Vergleich zu registerbasierter Reputation entwickeln, zur internationalen Normungsarbeit zu AI-Provenienz und -Attestierung beitragen und einen bilateralen Dialog mit den Autoren des CAC- Rahmenwerks über die Interoperabilität zwischen attestbasierten und kreditratingsbasierten Reputationssystemen führen.** *(Parallelen zu CAC Punkt 14 „Bonitätsbewertungsmechanismen“; es gilt der konsolidierte Arbeitsstrang zur Ausschussbildung .)* [ZITATE: W3C Verifiable Credentials Data Model v1.1; ISO/IEC 42001:2023 Managementsysteme.]

---

§IV. Stärkung einer anwendungsorientierten Entwicklung

Wo das Rahmenwerk der Cyberspace Administration of China neunzehn Sektoren auflistet, in denen der Staat vorschreibt, dass „Akteure X tun sollen“, spiegeln wir die neunzehn Sektoren wider und formulieren jeden neu als eine Frage der Souveränitätsbedingungen für jeden Einsatz von Akteuren in diesem Sektor. Das Rahmenwerk schreibt den Einsatz nicht vor; es legt die architektonischen Bedingungen fest, unter denen der Einsatz mit der Souveränität vereinbar ist. Die Umformulierung ist rhetorisch bescheiden, aber strukturell folgenreich: Die staatlich gelenkte Lesart positioniert intelligente Akteure als Instrumente sektoraler Programme, während die auf Souveränitätsbedingungen ausgerichtete Lesart sie als Werkzeuge positioniert, deren Einsatz Anforderungen an Zuordnung, Herkunft und Mitgliederportabilität erfüllen muss, unabhängig davon, wer sie einsetzt.

Die architektonischen Grundelemente, die in den neunzehn folgenden Sektoren herangezogen werden, sind vier. **Die kryptografische Herkunftsangabe** fügt jedem Datensatz überprüfbare Metadaten zur Herkunft hinzu – wer ihn verfasst hat, wann und auf Grundlage welcher Befugnis – die gegen nachträgliche stille Bearbeitung unveränderlich sind (Korrekturen werden gegengezeichnet und selbst aufgezeichnet). **Föderations-Envelopes** vermitteln den installationsübergreifenden Austausch: Nur die genehmigte Teilmenge wird übertragen, wobei Herkunft, Empfängerbindung und das Verbot der Weiterleitung standardmäßig mitgeführt werden. **Die mitgliederorientierte Portabilität** ermöglicht es dem Inhaber von Datensätzen, sein Bündel ohne die Erlaubnis des ursprünglichen Inhabers in eine

andere Installation zu exportieren, wobei die Herkunft am Zielort erhalten bleibt. **Die Durchsetzung von Grenzen** leitet Entscheidungen in den vier Grenzkategorien (Unumkehrbarkeit, wertgeladen, kulturkontextabhängig, beispiellos) standardmäßig an menschliche Beratung weiter, wobei die Weiterleitung selbst aufgezeichnet wird. Die folgenden Sektorenelemente benennen die sektorspezifische Ausprägung einer oder mehrerer dieser Primitiven; die generischen Fähigkeiten sind sektorübergreifend konstant. Siehe *„Architektonische Ausrichtung“* §3 für die Entwicklung der Primitiven; *Paper A* für die vollständige Beschreibung der Substratschicht.

(I) Wissenschaftliche Forschung

Punkt 15. In der Forschung gelten Souveränitätsprimitive. Forschungsumgebungen arbeiten mit souveränen Datensätzen – die von teilnehmenden Einzelpersonen, Institutionen, Hapū-/Iwi-Einheiten oder Forschungskonsortien im Rahmen ihrer jeweiligen Governance-Regelungen gehalten werden; die Herkunftsangabe begleitet die abgeleiteten Ergebnisse; eine bilaterale Föderation zwischen Institutionen stellt die Interoperabilitätsschicht bereit, wo Datenaustausch erforderlich ist. Wir erkennen den Wert der Vision des CAC-Rahmenwerks an, wonach intelligente Agenten die theoretische Deduktion, die Wissensintegration sowie die Integration mit wissenschaftlichen Instrumenten und Experimentierplattformen verbessern. Wir schlagen vor, dass für die Forschung in Aotearoa NZ diese Fähigkeiten im Rahmen einer forschungsethischen Governance eingesetzt werden, die spezifisch für jede Institution und jedes Forschungsprojekt ist, wobei die Tractatus-Primitive der pluralistischen Deliberation den architektonischen Mechanismus bereitstellt, um die forschungsethische Überprüfung über konkurrierende Werte-Rahmenwerke hinweg zu skalieren. Die operativen Primitive sind kryptografische Provenienz (jeder Datensatz und jedes abgeleitete Ergebnis trägt eine Provenienz, die seine Quellen, Ableitungen und den ethischen Überprüfungsrahmen, unter dem es erstellt wurde, bescheinigt) und Föderations-Envelopes (der institutionsübergreifende Austausch erfolgt im Rahmen expliziter Datenaustauschvereinbarungen, wobei der Envelope aufzeichnet, welche Teilmenge der Daten übertragen wird und unter welchem Einwilligungsumfang). Mitgliedergeführte Portabilität ermöglicht es einem Forschungsteilnehmer, seinen Beitrag zurückzuziehen und die Provenienz nachgelagert aktualisieren zu lassen; die Durchsetzung von Grenzen leitet die wertgeladenen ethischen Entscheidungen an den Forschungsethikausschuss weiter, anstatt sie autonomen Agenten zu überlassen. *(Parallelen zu CAC-Punkt 15 „Forschung und Erkundung“.)* [ZITATE: CARE-Prinzipien (Carroll et al. 2020); FAIR-Prinzipien (Wilkinson et al. 2016, https://doi.org/10.1038/sdata.2016.18); Te Kāhui Raraunga (kahuiraraunga.io — Māori-Modell zur Datenverwaltung und Māori-Rahmenwerk zur KI-Verwaltung); Taiuru, K. (20. September 2025) Kritische Analyse der Te Mana Raraunga-Datenprinzipien, taiuru.co.nz/critical-analysis-mana-raraunga/; Neuseeländischer Rahmen für Forschungsethik über den Health Research Council und die Royal Society Te Apārangi; Tractatus-Primitiv für pluralistische Deliberation (Stroh 2026).]

Punkt 16. In der Software-F&E gelten Urheberangabe und Audit. Code-Generierungsagenten arbeiten auf der Grundlage zuordnbarer Quellen; abgeleitete Werke tragen ihre Herkunftsangabe; CI/CD-Pipelines überprüfen die Build-

Bescheinigung und die Herkunft der Abhängigkeiten. Wir anerkennend den Wert des Engagements des CAC-Rahmenwerks für intelligente Softwareentwicklungsagenten, die die Anforderungsanalyse, den Architekturentwurf, die Codegenerierung und das Testen verbessern. Wir schlagen vor, dass alle derartigen Fähigkeiten unter Anforderungen an Zuordnung und Herkunft operieren; agentische Beiträge zu Code, Design oder Simulationsergebnissen werden sowohl dem Agenten als auch dem menschlichen oder organisatorischen Betreiber zugeordnet, in dessen Auftrag sie erstellt wurden. Die operative Grundfunktion ist die kryptografische Herkunftsnachweis, die auf jedes Code-Artefakt angewendet wird – den Agenten, der es vorgeschlagen hat, den menschlichen Prüfer, der es genehmigt hat, die Build-Pipeline, die es kompiliert hat, die eigenen Bescheinigungen des Abhängigkeitsbaums – und so eine überprüfbare Kette von der verfassten Zeile zurück zum autorisierten Commit bildet. Die Tractatus für Querverweise bietet eine Laufzeitüberprüfung, dass vorgeschlagene Code-Aktionen mit dem kanonischen Anweisungsverlauf übereinstimmen. *(Parallelen zu CAC-Punkt 16 „F&E-Unterstützung“.)* [ZITATE: W3C Verifiable Credentials Data Model v1.1; SBOM-Standards (Software Bill of Materials) über NTIA und OWASP CycloneDX; Tractatus -Primitiv zur Validierung von Querverweisen (Stroh 2026).]

(II) Industrielle Entwicklung

Punkt 17. In der Fertigung gelten Souveränitätsprimitive. Produktionsdaten sind die souveränen Aufzeichnungen des Herstellers; Agenten, die darauf zugreifen, werden zugeordnet; die installationsübergreifende Koordination für Lieferketten erfolgt bilateral**.** Wir erkennen den Wert des Engagements des CAC-Rahmenwerks für Produktionsmanagement-Agenten in den Bereichen Terminplanung, Ressourcenzuweisung und Prozessoptimierung sowie für die Integration mit CNC-Werkzeugmaschinen, Industrierobotern und automatisierten Produktionslinien an. Wir schlagen vor, dass alle diese Funktionen unter der Aufsicht des Herstellers betrieben werden, wobei die Koordination der Lieferkette durch bilaterale Vereinbarungen zwischen den beteiligten Herstellern und ihren Partnern erfolgt. Die operativen Grundelemente sind kryptografische Herkunftsnachweise (jede Charge trägt eine Produktionslinienbescheinigung – Sensorwerte, Agentenentscheidungen, menschliche Genehmigungen –, die bei jeder Fehleruntersuchung rückverfolgbar ist) und Föderations-Envelopes (die Koordination der Lieferkette erfolgt über bilateral signierte Envelopes, die festlegen, welche Produktionsdaten zu welchem Zweck mit welcher Gegenpartei geteilt werden). Mitgliederorientierte Portabilität bedeutet hier die Fähigkeit des Herstellers, seinen vollständigen Produktions-Audit-Trail an eine andere Lieferketten-Prüfstelle oder Aufsichtsbehörde zu exportieren, ohne die Genehmigung des ursprünglichen Plattformanbieters. *(Parallelen zu CAC-Punkt 17 „intelligente Fertigung“.)* [QUELLENANGABEN: ISO/IEC 42001:2023 Managementsysteme; Recherche zu neuseeländischen Standards für Fertigungsdaten und Initiativen zu Industrie 4.0 in Neuseeland steht noch aus.]

Punkt 18. Im Bereich Energie und Ressourcen gelten Souveränitätsgrundsätze. Umweltdaten, Ressourcenkataloge und Versandprotokolle sind souveräne Aufzeichnungen der zuständigen Stellen: der Krone für einige (gesetzlich festgelegte Ressourcen, bestimmte Umweltdaten); hapū und iwi für diejenigen, für die Zuweisungen

aus dem Vertragsabkommen gelten; privater Stellen für den Rest. Akteure handeln auf der Grundlage der Aufzeichnungen der jeweiligen Stelle unter deren Aufsicht. Die konkreten Zuweisungen sind einrichtungsspezifisch und hängen von den einschlägigen Gesetzgebungen und Vereinbarungen im Rahmen der Vertragsabwicklung ab. Wir erkennen den Wert des Engagements des CAC-Rahmenwerks für Umweltüberwachungsagenten zur Frühwarnung vor Naturkatastrophen und Verschmutzungsrisiken, für Stromverteilungs- und Netzwartungsagenten sowie für Anwendungen zur Rohstoffexploration an. Wir schlagen vor, dass im Kontext Aotearoa NZ die zuständigen Behörden aus dem bestehenden institutionellen und vertraglichen Rahmen hervorgehen und die Architektur die Audit- und Zuordnungsinfrastruktur bereitstellt, innerhalb derer diese Behörden agieren. Kryptografische Provenienz wird an Umweltmesswerte, Entscheidungen zur Netzsteuerung und Entscheidungen zur Ressourcenallokation angehängt, mit Zuordnung zur zuständigen Stelle (Krone, iwi oder privat). Föderations-Envelopes übertragen nur die genehmigte Teilmenge der Umweltdaten über Entitätsgrenzen hinweg — Frühwarnsignale werden an alle relevanten Entitäten weitergeleitet, ohne dass eine zentrale Aggregation erforderlich ist. Wo Zuweisungen aus Vertragsabkommen gelten, verwaltet die iwi ihre eigene Audit-Kette unter ihren eigenen Schlüsseln, unabhängig von den Systemen der Kronbehörden. *(Parallelen zu CAC-Punkt 18 „Energie und Ressourcen“.)* [QUELLENANGABEN: Resource Management Act 1991 (NZ); einschlägige Gesetzgebung zu Vertragsabkommen (einheitsspezifisch, Überprüfung vor Veröffentlichung von Version 1 ausstehend); Electricity Industry Act 2010 (NZ); Crown Minerals Act 1991 (NZ).]

Punkt 19. Im Verkehrsbereich gelten Souveränitätsgrundsätze. Fahrzeugtelemetrie, Verkehrsdaten und Sensordaten der Infrastruktur sind souveräne Aufzeichnungen von Betreibern, staatlichen Behörden und Straßenverkehrsbehörden; die Koordination zwischen ihnen – der neuseeländischen Verkehrsbehörde Waka Kotahi, iwi, den Seeverkehrsbehörden, der Zivilluftfahrtbehörde, den Regionalräten und den Stadträten – erfolgt im Rahmen einer bilateralen Föderation über die jeweiligen institutionellen Grenzen hinweg. Wir erkennen den Wert des CAC-Rahmenwerks hinsichtlich seines Engagements für intelligente Agenten in den Bereichen Verkehrssicherheit, Notfall-Einsatzleitung und Fahrzeugsteuerung an. Wir schlagen vor, dass der Kontext Aotearoa NZ mit seinen bestehenden bilateralen institutionellen Vereinbarungen über Verkehrsträger hinweg gut für einen föderierten Ansatz geeignet ist. Föderationsumschläge legen fest, welche Telemetrie-, Verkehrs- und Infrastruktursensordaten über institutionelle Grenzen hinweg (Waka Kotahi ↔ Regionalräte ↔iwi↔ Zivilluftfahrtbehörde), wobei die kryptografische Herkunftsnachweisbarkeit sicherstellt, dass eine forensische Rekonstruktion jedes Vorfalls auf der Ebene einzelner Agentenentscheidungen möglich ist. Mitgliederorientierte Portabilität gilt auf Fahrzeug- und Betreiberebene – der Betreiber kann den Austritt aus jeder Plattform ohne Bindung verlangen. *(Parallelen zu CAC-Punkt 19 „Verkehr“.)* [QUELLENANGABEN: Land Transport Act 1998 (NZ); Land Transport Management Act 2003 (NZ); Civil Aviation Act 1990 (NZ); Maritime Transport Act 1994 (NZ); Recherche zu Arbeiten zur Datenhoheit im neuseeländischen Verkehrsbereich steht noch aus.]

Punkt 20. In der Landwirtschaft gelten Souveränitätsgrundsätze. Landwirtschaftliche Daten sind die souveränen Aufzeichnungen des Landwirts; Daten zu Schädlingen,

Krankheiten, Erträgen und Bestandsdichte können bilateral mit Beratungsdiensten, Forschungseinrichtungen oder hapū rōpū geteilt werden, sofern dies zutrifft, und zwar zu den Bedingungen des Landwirts. Wir erkennen den Wert des Engagements des CAC-Rahmenwerks für intelligente Agenten im Bereich landwirtschaftlicher Dienstleistungen für technische Beratung, Schädlings- und Krankheitsdiagnose sowie die Integration mit intelligenten landwirtschaftlichen Maschinen und Gewächshäusern an. Wir schlagen vor, dass für Aotearoa Neuseeland – wo die Souveränität über landwirtschaftliche Daten ein anerkanntes Thema in landwirtschaftlichen Datengenossenschaften, Branchenorganisationen und im zunehmenden Engagement für die Datenhoheit Māori im Kontext der Primärindustrie ist – der bilaterale Datenaustausch zu den Bedingungen des Landwirts gut geeignet ist. Kryptografische Provenienz wird an landwirtschaftliche Daten angehängt – Sensorwerte, Empfehlungen von Agenten, Behandlungsentscheidungen, Ertragsergebnisse. Föderations-Envelopes übertragen nur die genehmigte Teilmenge (Schädlings- und Krankheitsdaten an Beratungsdienste, aggregierte Ertragsdaten an Forschungseinrichtungen, kulturkontextabhängige Daten an hapū rōpū unter Einhaltung angemessener tikanga) zu den Bedingungen des Landwirts. Mitgliederorientierte Portabilität ermöglicht es dem Landwirt, zwischen landwirtschaftlichen Datengenossenschaften zu wechseln, ohne seinen historischen Prüfpfad zu verlieren. *(Parallelen zu CAC-Punkt 20 „landwirtschaftliche Produktion“.)* [QUELLENANGABEN: Recherche zu neuseeländischen Arbeiten zur Datenhoheit in der Landwirtschaft und zu Regelungen zur Verwaltung landwirtschaftlicher Daten steht noch aus; Te Kāhui Raraunga (hui.io — Māori zur Datenverwaltung und Māori AI Governance Framework); Taiuru, K. (20. September 2025) Kritische Analyse der Te Mana Raraunga, taiuru.co.nz/critical-analysis-mana-raraunga/, soweit zutreffend.]

Punkt 21. Im Finanzdienstleistungssektor gelten Souveränitätsgrundsätze. Kundendaten, Transaktionsdaten und Risikosignale sind souveräne Aufzeichnungen des verwahrenden Instituts und unterliegen den aufsichtsrechtlichen Anforderungen der Reserve Bank of New Zealand / Te Pūtea Matua, dem Datenschutzgesetz von 2020 sowie dem Gesetz zur Bekämpfung von Geldwäsche und Terrorismusfinanzierung von 2009. Die Zusammenarbeit bei der Bekämpfung von Geldwäsche und Terrorismusfinanzierung erfolgt bilateral über etablierte Kanäle – die neuseeländische Finanzermittlungsstelle und die internationalen FATF-Kanäle – und die KI-Unterstützung wird durch diese bestehenden regulatorischen Vereinbarungen zugewiesen und begrenzt. Wir erkennen den Wert des CAC-Rahmenwerks hinsichtlich der Einbindung von Akteuren zur Finanzrisikokontrolle bei der Kreditgenehmigung, Transaktionsüberwachung, Kontosicherheit und Überwachung zur Bekämpfung von Geldwäsche an. Wir schlagen vor, dass für Aotearoa NZ der bestehende institutionelle und regulatorische Rahmen gut geeignet ist für eine zuordnungsbasierte Prüfung auf der Ebene jedes einzelnen Finanzinstituts, mit bilateraler Zusammenarbeit über etablierte Kanäle für die institutionsübergreifende und internationale Koordination. Jede Transaktion, jede KI-Bewertung und jede menschliche Genehmigung ist mit einer kryptografischen Herkunftsangabe versehen – rückverfolgbar durch AML/CFT-Prüfer, die Reserve Bank, FATF-Inspektoren und den Kunden selbst im Rahmen ihrer jeweiligen Zuständigkeiten rückverfolgt werden kann. Föderations-Envelopes vermitteln die institutionsübergreifende AML/CFT-Zusammenarbeit: Nur das genehmigte Signal für verdächtige Aktivitäten wird weitergeleitet, wobei der Empfänger an

die neuseeländische Financial Intelligence Unit oder die zuständige Gegenstelle gebunden ist. Die mitgliederorientierte Portabilität unterstützt Verpflichtungen zur Kontoportabilität: Die Transaktionshistorie des Kunden ist mit intakter Herkunftsnachweisbarkeit in ein anderes Institut exportierbar. *(Parallelen zu CAC Punkt 21 „Finanzdienstleistungen“.)* [QUELLENANGABEN: Reserve Bank of New Zealand Act 2021; Anti-Money Laundering and Countering Financing of Terrorism Act 2009 (NZ); Privacy Act 2020 (NZ); FATF-Empfehlungen.]

(III) Alltag

Punkt 22. In Endbenutzeranwendungen gelten Souveränitätsprimitive. Mitgliederportable Identifikatoren ersetzen plattformspezifische Konten; die geräteübergreifende Koordination wird durch den eigenen Schlüsselbund oder die Identitäts-Wallet des Mitglieds vermittelt**.** Souveränität bedeutet hier, dass der Nutzer die Datensätze besitzt – unabhängig davon, ob die Anwendung von einem Anbieter Aotearoa NZ oder einem internationalen Anbieter entwickelt wurde. Wir erkennen den Wert des CAC-Rahmenwerks an, das sich für intelligente Agenten einsetzt, die Internetanwendungen und -dienste über Mobiltelefone, Computer, Fahrzeuge, Haushaltsgeräte, Wearables und Verbraucherroboter hinweg befähigen. Wir schlagen vor, dass für jede Anwendung, die mit Benutzerdatensätzen arbeitet, die architektonischen Primitive der Attribution und der Mitgliederportabilität unabhängig von der Zuständigkeit des Anbieters gelten. Das operative Element ist die mitgliederorientierte Portabilität, die über dezentrale Identifikatoren und überprüfbare Berechtigungsnachweise des W3C umgesetzt wird: Die Identität des Nutzers wird in seinem eigenen Schlüsselbund (oder Wallet) gespeichert, wobei die geräteübergreifende Koordination durch seine eigenen Schlüssel vermittelt wird. Eine kryptografische Herkunftsangabe ist mit jeder Agentenaktion verbunden, die in Bezug auf die Datensätze des Nutzers durchgeführt wird, und ist dem Agenten sowie dem Autorisierungsrahmen des Nutzers zuzuordnen. Föderations-Envelopes vermitteln die herstellerübergreifende Koordination nur, wenn der Nutzer dies autorisiert. *(Parallelen zu CAC-Punkt 22 „Endnutzeranwendungen“.)* [QUELLENANGABEN: W3C Decentralized Identifiers (DIDs) v1.0; W3C Verifiable Credentials Data Model v1.1; Privacy Act 2020 (NZ), Datenschutzgrundsatz 7 (Korrektur).]

Punkt 23. In Kultur und Tourismus gelten Souveränitätsgrundsätze. Kulturelle Inhalte unterliegen der Kontrolle ihrer Schöpfer; im Kontextvon Aotearoa Neuseeland sind die Kaitiaki- Verpflichtungen gegenüber Taonga von zentraler Bedeutung dafür, wie KI-Agenten mit kulturellem Material interagieren dürfen. Übersetzungsagenten bewahren die Urheberschaft und den kulturellen Kontext; ihre Ergebnisse ersetzen nicht das ursprüngliche mātauranga, und was im te ao Māori-Kontext als angemessene Nutzung gilt, ist von den tangata whenua zu bestimmen. Besucherdaten, die von Tourismusdiensten verarbeitet werden, werden als souveräne Aufzeichnungen des Besuchers behandelt. Wir erkennen den Wert des Engagements des CAC-Rahmenwerks für Agenten zur Erstellung kultureller Inhalte und Tourismusdienstleister an. Wir schlagen vor, dass für Aotearoa Neuseeland – wo mātauranga Māori gemäß den Partnerschaftsverpflichtungen des Te Tiriti ein taonga ist und wo Dr. Karaitiana Taiurusveröffentlichte Arbeit zum

Schutz von mātauranga Māori in KI-Trainingsdaten sowie das Raraungas Māori-KI-Governance-Rahmenwerk grundlegende wissenschaftliche Arbeit darstellen – die architektonischen Primitive die Audit-Infrastruktur bereitstellen und die inhaltliche Festlegung der angemessenen Nutzung den Inhabern des Mātauranga obliegt. Kryptografische Provenienz wird mit kulturellem Material verknüpft: wer es erstellt hat, unter welcher Autorität, mit welchem Nutzungsumfang. Für mātauranga bestimmtdas tikanga der Inhaber, was Kontrollbefugnis in der Praxis bedeutet; das Substrat stellt die Audit-Infrastruktur bereit, sodass eine Verletzung der Einwilligung forensisch rekonstruierbar ist und nicht nur vertraglich bestritten werden kann. Föderations- Envelopes transportieren nur die genehmigte Teilmenge von mātauranga über Installationsgrenzen hinweg, wobei standardmäßig keine Weiterleitung erfolgt – Übersetzungsagenten erben die Daten, können sie aber nicht neu lizenzieren. Besucherdaten enthalten die Identitätsbescheinigung des Besuchers; mitgliederorientierte Portabilität bedeutet, dass der Besucher bei seiner Abreise seinen Tourismusdatensatz exportiert. *(Parallelen zu CAC-Punkt 23 „Kultur und Tourismus“.)* [QUELLENANGABEN: Taiuru, K. – Schutz von mātauranga Māori in KI-Trainingsdaten (spezifische Veröffentlichungen stehen noch zur Überprüfung aus); Te Kāhui Raraunga (kahuiraraunga.io – Māori-Daten-Governance -Modell und Māori-KI-Governance-Rahmenwerk); Taiuru, K. (20. September 2025) Kritische Analyse der Te Mana Raraunga-Datenprinzipien, taiuru.co.nz/critical-analysis-mana-raraunga/; CARE-Prinzipien (Carroll et al. 2020); Wai 262 (Bericht des Waitangi-Tribunals über indigene Flora und Fauna sowie kulturelles geistiges Eigentum).]

Punkt 24. Bei kommerziellen Dienstleistungen gelten Souveränitätsgrundsätze. Kundeninteraktionen erzeugen Aufzeichnungen; beide Parteien – Betreiber und Kunde – verfügen über Herkunftskopien; Streitigkeiten werden bilateral koordiniert. Verkörperte Agenten im Einzelhandel, im Gastgewerbe, in der Altenpflege und in der Behindertenbetreuung agieren unter der Aufsicht des Einsatzverantwortlichen und erstellen überprüfbare Aufzeichnungen ihrer Handlungen. Wir erkennen den Wert des CAC-Rahmenwerks an, das sich für einen 24/7-Kundenservice, verkörperte intelligente Agenten für Beratung, Reinigung, Lagerung und Vertrieb in gewerblichen Einrichtungen sowie verkörperte Agenten für Haushaltshilfe, Altenpflege, Kinderbetreuung und Behindertenhilfe einsetzt. Wir schlagen vor, dass für Aotearoa Neuseeland alle derartigen Anwendungen im Rahmen bestehender Regulierungsrahmen für Verbraucherschutz, Pflegequalität und Behindertenhilfe betrieben werden. Beide Parteien (Betreiber und Kunde) verfügen über kryptografisch signierte Herkunftskopien jeder Interaktion, sodass Streitigkeiten anhand einer gemeinsamen, überprüfbaren Aufzeichnung und nicht anhand einseitiger Plattformprotokolle beigelegt werden können. Verkörperte Agenten im Einzelhandel, im Gastgewerbe, in der Altenpflege und in der Behindertenbetreuung arbeiten unter der Durchsetzung von Grenzen: Kulturkontextabhängige oder wertgeladene Entscheidungen (Übersteuerungen bei der Medikamentenverabreichung, Änderungen am Pflegeplan, Eskalationen in der Behindertenbetreuung) werden standardmäßig an menschliche Entscheidungsträger weitergeleitet. Der Föderations-Envelope vermittelt die Übergabe von Daten der Klasse „Pflegeaufzeichnungen“ zwischen Anbietern. *(Parallelen zu CAC-Punkt 24 „kommerzielle Dienstleistungen“.)* [QUELLENANGABEN: Consumer Guarantees Act 1993 (NZ); Fair Trading Act 1986 (NZ); Health and Disability Services (Safety)

Act 2001 (NZ); New Zealand Disability Strategy.]

(IV) Soziales

Punkt 25. Im Bildungswesen gelten Souveränitätsgrundsätze. Lernunterlagen sind die souveränen Unterlagen des Schülers bzw. der Schülerin, mit gemeinsamer Verwaltung, sofern der Schüler oder die Schülerin minderjährig ist; von Akteuren erstellte Unterrichtsmaterialien werden zugeordnet; institutionelle Unterlagen – Anwesenheitslisten, Bewertungen, Qualifikationsnachweise – unterliegen der bestehenden institutionellen Governance gemäß dem Education and Training Act 2020. Die Übertragbarkeit gilt für den Schüler, mit entsprechenden institutionellen Vorkehrungen für die Übergabe bei Wechseln zwischen Bildungsanbietern. Wir erkennen den Wert des Engagements des CAC-Rahmenwerks für die Erstellung von Unterrichtsmaterialien, die Benotung von Hausaufgaben, die Analyse des Lernfortschritts, personalisierte Lernpläne und virtuelle Lehrassistenten an. Wir schlagen vor, dass diese Funktionen für Aotearoa Neuseeland im Rahmen des Datenschutzgesetzes von 2020 und des Bildungs- und Ausbildungsgesetzes von 2020 betrieben werden, wobei die Souveränität über die Schülerdaten durchgehend gewahrt bleibt. Jede Bewertung, jedes von einem Akteur erstellte Lehrmaterial und jede Qualifikationsvergabe ist mit einer kryptografischen Herkunftsangabe versehen – zuordenbar an den Akteur, den betreuenden Pädagogen und die Einrichtung. Die souveräne Akte des Schülers ist übertragbar: Bei jedem Übergang (von Schule zu Schule, von Schule zu Universität, zwischen Anbietern, zwischen Ländern) bringt der Schüler sein vollständiges Aktenbündel mit intakter Herkunftsangabe zur aufnehmenden Einrichtung mit. Die Durchsetzung von Grenzen leitet Entscheidungen, die die Bewertung verändern und sich auf Qualifikationen auswirken, standardmäßig an menschliche Entscheidungsinstanzen weiter. *(Parallelen zu CAC-Punkt 25 „Bildung und Unterricht“.)* [QUELLENANGABEN: Bildungs- und Ausbildungsgesetz 2020 (NZ); Datenschutzgesetz 2020 (NZ); Neuseeländischer Lehrplan.]

Punkt 26. Im Gesundheitswesen gelten die Grundprinzipien der Souveränität. Patientenakten sind gemäß dem Health Information Privacy Code 2020 und den Verwaltungsstrukturen von Te Whatu Ora / Health New Zealand die souveränen Aufzeichnungen des Patienten; Diagnoseagenten erzeugen zuordnungsfähige Ergebnisse; Behandlungsempfehlungen sind mit einer Herkunftsangabe versehen; die Koordination zwischen den Leistungserbringern erfolgt über die Kanäle der Health Information Standards Organisation (HISO) und die Interoperabilitätsvereinbarungen von Te Whatu Ora. Wir erkennen den Wert des Engagements des CAC-Rahmenwerks für die Analyse medizinischer Bilddaten, die Begründung von Krankheitsdiagnosen, personalisierte Behandlungspläne, das Medikamentenmanagement, die Operationsplanung und Agenten für die Verwaltung medizinischer Unterlagen an. Wir schlagen vor, dass diese Funktionen für Aotearoa Neuseeland im Rahmen des bestehenden Governance-Rahmens für Gesundheitsinformationen ablaufen, wobei die Souveränität der Patienten über ihre Gesundheitsunterlagen als architektonische Grundvoraussetzung beibehalten wird. Kryptografische Provenienz wird klinischen Unterlagen, KI-Diagnoseergebnissen, Behandlungsempfehlungen und der Medikamentenverabreichung zugeordnet – jeder Eintrag ist dem Verfasser zuordenbar und kann nachträglich nicht

unbemerkt geändert werden (Korrekturen sind gegengezeichnete Änderungen, die selbst aufgezeichnet werden). Föderationsumschläge enthalten Überweisungen: Nur der genehmigte klinische Teil wird vom überweisenden Anbieter an den empfangenden weitergeleitet, wobei der Empfänger standardmäßig an die Daten gebunden ist und diese nicht weiterleiten darf. Mitgliederorientierte Portabilität ermöglicht es dem Patienten, sein Datensatzbündel an einen anderen Anbieter – öffentlich, privat oder international – ohne die Erlaubnis des ursprünglichen Inhabers zu exportieren, wobei die Herkunftsangaben am Zielort erhalten bleiben. Die Durchsetzung von Grenzen leitet klinisch ungewisse und wertgeladene Entscheidungen (Lebensende, umstrittene Diagnosen, geistige Zurechnungsfähigkeit) an menschliche Entscheidungsfindung weiter, anstatt sie autonomen Agenten zu überlassen. *(Parallelen zu CAC-Punkt 26 „Gesundheitswesen“.)* [QUELLENANGABEN: Health Information Privacy Code 2020 (NZ); Pae Ora (Healthy Futures) Act 2022 (NZ); HISO-Datenstandards.]

Punkt 27. Im Bereich Beschäftigung und Arbeit gelten Souveränitätsgrundsätze. Beschäftigungsunterlagen, Ausbildungsnachweise und Streitfallunterlagen unterliegen der Souveränität der Parteien; die Mediation erfolgt im Rahmen der bestehenden Regelungen des Employment Mediation Service; KI-Unterstützung wird der bestehenden dreigliedrigen Struktur (Arbeitnehmer / Arbeitgeber / Staat) des neuseeländischen Arbeitsrechts zugeschrieben und durch diese begrenzt. Wir erkennen den Wert des CAC-Rahmenwerks an, das sich für Akteure in den Bereichen Beschäftigungsförderung, Ausbildung und Bewertung von Fachpersonal, Arbeitsbeziehungsdienste, Sozialversicherung, Schlichtung von Arbeitskonflikten und Verwaltung von Lohnrückständen einsetzt. Wir schlagen vor, dass diese Funktionen für Aotearoa Neuseeland im Rahmen des Employment Relations Act 2000 und des damit verbundenen dreigliedrigen Rahmens operieren, wobei Zurechnung und Herkunft durchgehend angewendet werden. Kryptografische Herkunftsnachweise werden mit Beschäftigungsunterlagen, Ausbildungszertifikaten und Streitfallunterlagen verknüpft – zuordenbar an die beteiligten Parteien. Verbund-Umschläge ermöglichen die Übertragbarkeit von Ausbildungsnachweisen zwischen Arbeitgebern ohne Verlust der Herkunft. Grenzsicherungsmaßnahmen leiten Entscheidungen über Einstellung, Entlassung, Disziplinarmaßnahmen und Streitbeilegung an menschliche Instanzen weiter – autonomes Handeln von Agenten gegen den Beschäftigungsstatus eines einzelnen Arbeitnehmers wird strukturell verhindert. *(Parallelen zu CAC-Punkt 27 „Human Resources“.)* [QUELLENANGABEN: Employment Relations Act 2000 (NZ); Holidays Act 2003 (NZ); Human Rights Act 1993 (NZ); Neuseeländischer dreigliedriger Rahmen für Arbeitsbeziehungen.]

Punkt 28. In Informationsdiensten gelten Souveränitätsgrundsätze. Inhalte werden ihren Urhebern zugeschrieben; Empfehlungsagenten agieren im Rahmen des souveränen Profils des Nutzers, das dieser einsehen, exportieren und übertragen kann; die redaktionelle Überprüfung bleibt eine menschliche Aufgabe. Wenn KI-Agenten Inhalte produzieren, erfolgt die Zuordnung an den Agenten und an den menschlichen oder organisatorischen Betreiber, in dessen Auftrag er gehandelt hat; die Offenlegung KI-generierter Inhalte ist die grundlegende architektonische Verpflichtung. Wir erkennen den Wert der Verpflichtung des CAC-Rahmenwerks an, intelligente Agenten für die Erstellung von Online-Inhalten, Nutzeranalyse, Themenplanung, redaktionelle Bearbeitung, Verbreitung und

Empfehlung, Inhaltsprüfung, Meinungsführung, emotionale Unterstützung und Echtzeit- Übersetzung einzusetzen. Wir schlagen vor, dass für Aotearoa NZ die Anforderungen an die Urheberschaft für alle derartigen Anwendungen gelten, wobei die bestehenden Rundfunkstandards und der Rahmen für schädliche digitale Kommunikation den regulatorischen Kontext bilden. Jeder Inhalt ist mit einer kryptografischen Herkunftsangabe versehen: Identität des Autors, Zuordnung „KI vs. Mensch“, redaktionelle Überprüfungskette. Föderations- Umschläge enthalten Verteilungsentscheidungen: Jede Empfehlungsoberfläche erhält nur die Inhalte, die die vorgelagerte Instanz signiert und deren Weitergabe sie zugestimmt hat, wobei eine Weiterleitung standardmäßig unterbunden ist. Mitgliederorientierte Portabilität stellt dem Nutzer seinen Interaktionsverlauf und sein Empfehlungsprofil in einer portablen Form zur Verfügung – er kann zu einem anderen Dienst wechseln, ohne seinen Inhaltsverlauf zu verlieren. Die Durchsetzung von Grenzen leitet redaktionelle Entscheidungen an menschliche Autoritäten weiter – autonome Handlungen von Agenten darüber, was verstärkt oder unterdrückt werden soll, werden strukturell blockiert. *(Parallelen zu CAC-Punkt 28 „Informationsdienste“.)* [QUELLENANGABEN: Rundfunkgesetz von 1989 (NZ); Gesetz über schädliche digitale Kommunikation von 2015 (NZ); Datenschutzgesetz von 2020 (NZ); Suche nach Standards zur Zuordnung von KI-Inhalten noch ausstehend.]

(V) Soziale Governance

Punkt 29. In der öffentlichen Verwaltung gelten souveränitätsbezogene Grundprinzipien. Interaktionen der Bürger mit dem Staat führen zu Aufzeichnungen**,** die sowohl vom Bürger als auch vom Staat aufbewahrt werden; die vom Mitglied gehaltenen Identitätsnachweise wandern im Laufe der Zeit in die Kontrolle des Mitglieds über; die Unterstützung durch Agenten in Genehmigungsverfahren wird durch verwaltungsrechtliche Grundsätze geregelt und begrenzt. Kronbehörden bleiben rechenschaftspflichtig gemäß dem Public Service Act 2020, dem Official Information Act 1982, dem Privacy Act 2020, der Algorithm Charter for Aotearoa New Zealand und dem Public Records Act 2005. Wir erkennen den Wert des CAC-Rahmenwerks hinsichtlich der Verpflichtung zu Agenten für Verwaltungsgenehmigungen, politische Konsultationen und proaktive Dienstleistungserbringung an. Wir schlagen vor, dass für Aotearoa NZ alle derartigen Anwendungen staatlicher Behörden innerhalb des bestehenden Rechenschaftsrahmens betrieben werden, wobei die architektonischen Grundelemente die Prüfungsinfrastruktur bereitstellen, die mit den Verpflichtungen der Algorithmus-Charta zu Transparenz und Partnerschaft mit Māori im Einklang steht. Jede Verwaltungshandlung ist mit einer kryptografischen Herkunftsangabe versehen: wer entschieden hat, auf welcher Grundlage, anhand welcher Bürgerdaten, mit welcher Unterstützung durch einen Akteur. Föderationsumschläge vermitteln die behördenübergreifende Koordination – nur die genehmigte Teilmenge der Bürgerdaten wird von Behörde zu Behörde weitergeleitet, mit Prüfpfad. Die mitgliederorientierte Portabilität setzt das Datenschutzgesetz (Privacys Act) – Informationsschutzgrundsatz 6 (Zugriffsrechte) und die Aufbewahrungspflichten des Gesetzes über öffentliche Aufzeichnungen von 2005 – der Bürger kann seine vollständigen Interaktionsdaten mit der Regierung unter Verwendung seiner eigenen Schlüssel exportieren. Die Durchsetzung von Grenzen leitet Entscheidungen nach Verwaltungsermessen, Entscheidungen im Zusammenhang mit Verpflichtungen aus

dem Vertrag sowie Entscheidungen, die Rechte betreffen, an menschliche Instanzen weiter. *(Parallelen zu CAC-Punkt 29 „Dienstleistungen der öffentlichen Verwaltung“.)* [QUELLENANGABEN: Public Service Act 2020 (NZ); Official Information Act 1982 (NZ); Privacy Act 2020 (NZ); Algorithm Charter for Aotearoa New Zealand (2020); Public Records Act 2005 (NZ).]

Punkt 30. In der Justiz gelten die Grundprinzipien der Souveränität. Gerichtsakten, Beweismittel und Rechtsdokumente unterliegen den bestehenden gerichtlichen Verfahren; KI-Unterstützung wird ausgewiesen; die Beweiskette ist, soweit anwendbar, kryptografisch gesichert; Zugriffskontrollen folgen den bestehenden gerichtlichen Vorschriften. Tools zur Unterstützung von Prozessparteien, die sich selbst vertreten und KI nutzen, legen deren Einsatz offen und erzeugen eine überprüfbare Herkunftsnachweis. Wir erkennen den Wert des Engagements des CAC-Rahmenwerks für durchgängige Unterstützung bei der Fallbearbeitung, die Erstellung von Rechtsdokumenten, Rechtsberatung, Rechtsberatung und Rechtsaufsicht an. Wir schlagen vor, dass für Aotearoa Neuseeland alle derartigen Anwendungen unter dem Senior Courts Act 2016, dem Evidence Act 2006 sowie den etablierten Gerichtsregeln und Praxisleitfäden zur Regelung des KI-Einsatzes in Gerichtsverfahren betrieben werden. Jedes vorgelegte Beweismittel, jeder KI-gestützte Entwurf eines Rechtsdokuments und jedes Suchergebnis wird mit einer kryptografischen Herkunftsangabe versehen. Die Beweiskette ist kryptografisch verankert – Fragen zur Zulässigkeit lassen sich anhand der zugrunde liegenden Daten beantworten und nicht anhand umstrittener Plattformprotokolle. Mitgliederorientierte Portabilität bedeutet, dass der sich selbst vertretende Prozessbeteiligte seine vollständige Aktenlage zwischen den Instanzen (Gericht → Gericht → Berufung) mit intakter Herkunftsnachweisbarkeit mitnehmen kann. Die Durchsetzung von Grenzen leitet Werturteile und Ermessensentscheidungen (ob eine Klage eingereicht, ein Vergleich angenommen oder eine für den Prozessbeteiligten nachteilige Maßnahme ergriffen werden soll) an den menschlichen Prozessbeteiligten oder dessen Rechtsbeistand weiter – autonome Handlungen des Agenten gegen die Rechtsposition des Prozessbeteiligten sind strukturell blockiert. *(Parallelen zu CAC-Punkt 30 „Justizdienstleistungen“.)* [QUELLENANGABEN: Senior Courts Act 2016 (NZ); Evidence Act 2006 (NZ); aktuelle gerichtliche Leitlinien zum Einsatz von KI werden noch recherchiert.]

Punkt 31. Im Bereich der öffentlichen Sicherheit gelten Souveränitätsgrundsätze. Die Überwachung unterliegt bestehenden Gesetzen – dem Privacy Act 2020, dem Search and Surveillance Act 2012 und dem Intelligence and Security Act 2017 – und alle KI-Agenten, die im Kontext der öffentlichen Sicherheit tätig sind, erzeugen unter diesen Rahmenbedingungen eine überprüfbare Herkunft. Agenten zur Verhaltensüberwachung agieren innerhalb des Rahmens, der gemäß diesen Gesetzen bereits rechtmäßig ist. Wir erkennen den Wert des CAC-Rahmenwerks hinsichtlich seines Engagements für Überwachungs- und Frühwarnagenten, Notfall- und Rettungskoordinierungsagenten sowie Anwendungen zur Identifizierung und dynamischen Prävention von abnormalem Verhalten an. Wir schlagen für den Kontext Aotearoa (Neuseeland) vor, dass der architektonische Beitrag von Attribution und Provenienz darin besteht, agentische KI in Kontexten der öffentlichen Sicherheit überprüfbar zu machen; ob und wie solche Fähigkeiten eingesetzt werden sollten, ist eine Wertentscheidung für den jeweiligen gesetzlichen und

politischen Rahmen, die an das Parlament und die zuständigen Minister gerichtet ist, wobei die Architektur die Prüfungsinfrastruktur bereitstellt, innerhalb derer diese Entscheidungen handhabbar werden. Kryptografische Provenienz ist mit jedem Überwachungssignal, jeder Entscheidung zur Verhaltensüberwachung und jeder Notfallmaßnahme verbunden – wodurch die nachträgliche Rechenschaftspflicht auf eine Weise handhabbar wird, wie es bei eigenständigen Handlungen von Agenten nicht der Fall ist. Föderations-Envelopes vermitteln die behördenübergreifende Koordination der öffentlichen Sicherheit: Welche Informationen zwischen der Polizei, dem Government Communications und den Notfallbehörden, ist an das gebunden, was jede Behörde für welchen spezifischen Zweck als freigebbar genehmigt hat. Die Durchsetzung von Grenzen ist hier tragend: Entscheidungen, die Rechte betreffen (Durchsuchung, Festnahme, Überwachungsgenehmigung, Gewaltanwendung), werden an menschliche Autoritäten weitergeleitet – die architektonische Grundlage, unterhalb derer kein Akteur autonom handelt. *(Parallelen zu CAC-Punkt 31 „öffentliche Sicherheit“.)* [QUELLENANGABEN: Datenschutzgesetz 2020 (NZ); Durchsuchungs- und Überwachungsgesetz 2012 (NZ); Geheimdienst- und Sicherheitsgesetz 2017 (NZ); Neuseeländisches Grundrechtegesetz 1990.]

Punkt 32. In der Stadtverwaltung gelten Souveränitätsgrundsätze. Städtische Daten – Sensornetzwerke, Planungsdaten, Baugenehmigungen, Infrastrukturbetriebsdaten – werden von den Gemeinderäten als souveräne Aufzeichnungen geführt; agentische Systeme, die in kommunalen Funktionen operieren, sind durch das Local Government Act 2002 und die Verwaltungsstrukturen der Gemeinderäte zugeordnet und rechenschaftspflichtig. Wir erkennen den Wert des Engagements des CAC-Rahmenwerks für intelligente Agenten in den Bereichen Stadtplanung, Städtebau und Stadtverwaltung an, einschließlich für intelligentes Bauen, Gebäudemanagement und den Betrieb städtischer Infrastruktur. Wir schlagen vor, dass für Aotearoa NZ alle derartigen Anwendungen im Rahmen bestehender Rechenschaftsregelungen der Kommunalverwaltung betrieben werden, wobei die architektonischen Grundelemente die Infrastruktur für Audit und Zuordnung bereitstellen. Eine kryptografische Herkunftsangabe ist mit jedem Sensorwert, jeder Planungsentscheidung, jeder Baugenehmigung und jeder Entscheidung zum Infrastrukturbetrieb verbunden – überprüfbar durch die Einwohner, durch Local Government NZ, durch den Rechnungshof und durch nachfolgende Gemeinderäte. Föderations-Envelopes vermitteln die Koordination zwischen den Gemeinderäten und den Datenaustausch zwischen Zentral- und Lokalebene – nur die genehmigte Teilmenge wird übertragen. Mitgliederorientierte Portabilität gilt für Daten auf Anwohnebene: Ein Anwohner kann seine Interaktionen mit dem Gemeinderat mitnehmen, wenn er in einen anderen Bezirk umzieht. Die Durchsetzung von Grenzen leitet Entscheidungen, die den Vertrag betreffen und kulturell bedeutsam sind (Urupā, Wāhi Tapu, taonga), an die entsprechenden Tangata-Whenua-Beratungsgremien weiter, anstatt sie autonomen Akteuren zu überlassen. *(Parallelen zu CAC-Punkt 32 „Stadtverwaltung“.)* [QUELLENANGABEN: Local Government Act 2002 (NZ); Building Act 2004 (NZ); Resource Management Act 1991 (NZ).]

Punkt 33. Bei der Beschaffung gelten die Grundprinzipien der Souveränität. Ausschreibungsunterlagen, Bewertungen und Verträge sind souveräne Unterlagen des Auftraggebers; die Unterstützung durch Akteure bei der Beschaffung unterliegt den Regeln für das öffentliche Beschaffungswesen und dem geltenden Vertragsrecht;

Transparenz wird durch bestehende, OIA-konforme Veröffentlichungen gewährleistet. Wir erkennen den Wert des CAC-Rahmenwerks an, das sich für ein durchgängiges intelligentes Management von Ausschreibungs- und Bietprozessen einsetzt, wobei Intelligenz auf Transaktionen, Dienstleistungen und die Überwachung angewendet wird. Wir schlagen vor, dass für Aotearoa Neuseeland die Vorschriften für das öffentliche Beschaffungswesen und der bestehende Rahmen für das öffentliche Beschaffungswesen den geeigneten Rechenschaftskontext bieten, wobei Zuordnung und Herkunft durchgängig angewendet werden. Kryptografische Herkunftsnachweise sind mit jedem Ausschreibungsdatensatz, jeder Bewertung, jeder Vertragsänderung und jeder Vergabeentscheidung verknüpft – veröffentlicht im Rahmen bestehender OIA-konformer Transparenzregelungen mit Integritätseigenschaften auf Substratebene. Föderationsumschläge vermitteln die für die Beschaffung erforderliche Koordination von Konsortien und Lieferketten, ohne wettbewerbsrelevante Daten außerhalb des vereinbarten Umfangs offenzulegen. Die Durchsetzung von Grenzen leitet die diskretionären Beschaffungsentscheidungen (Vergabe, Außerkraftsetzung, Ausnahme) an menschliche Instanzen weiter – autonome Handlungen von Agenten bei einer Auftragsvergabe werden strukturell blockiert. *(Parallelen zu CAC-Punkt 33 „Ausschreibung und Angebotsabgabe“.)* [QUELLENANGABEN: Vorschriften für das öffentliche Beschaffungswesen (NZ); Gesetz über öffentliche Aufzeichnungen von 2005 (NZ); Gesetz über amtliche Informationen von 1982 (NZ).]

---

§V. Aufbau eines föderierten Ökosystems

Während das Rahmenkonzept der chinesischen Cyberspace-Behörde ein Ökosystem aus Industrieclustern vorsieht, das durch internationale KI-Konferenzen eine Vorreiterrolle auf nationaler Ebene anstrebt, bieten wir ein föderiertes Ökosystem an, in dem die Koordination durch bilaterale Zusammenschlüsse zwischen souveränen Partnern erfolgt und die internationale Abstimmung über etablierte Normungsgremien erfolgt. Die beiden folgenden Unterabschnitte – Förderung der föderierten Zusammenarbeit und Stärkung der bilateralen Förderung – legen gemeinsam dar, wie sich ein Ökosystem souveräner Einrichtungen selbst trägt und international agiert.

(I) Förderung der föderierten Zusammenarbeit

Punkt 34. Open Source unter freizügigen Lizenzen. Referenzimplementierungen sollten unter freizügigen Open-Source-Lizenzen verfügbar sein. Die aktuellen MDSL-Implementierungen sind eine von möglicherweise mehreren Referenzen: Das Tractatus wird unter Apache 2.0 für den Code und CC BY 4.0 für die Dokumentation vertrieben; die Codebasen Village und Community werden ab Mitte 2026 schrittweise auf EUPL-1.2 (European Union Public Licence); zukünftige MDSL- Beiträge sollen, soweit praktikabel, unter EUPL-1.2 stehen, um eine souveräne Angleichung an die Souveränitätsbemühungen der Europäischen Union zu gewährleisten und um die Kompatibilität mit bilateralen Verbünden zwischen souveränen Installationen über mehrere Rechtsordnungen hinweg sicherzustellen. Wir würdigen den Verdienst des

CAC-Rahmenwerks bei der Förderung von Open-Source-Innovation, einschließlich inländischer Open-Source-Communities für KI, der Kompatibilität mit Open-Source- Chips, Betriebssystemen und großen Modellen sowie der Einbindung von Unternehmen, Universitäten und Forschungseinrichtungen in Open-Source- Projekte. Open-Source unter freizügigen Lizenzen ist bilateraler-Föderation-freundlich: Jede souveräne Installation erstellt einen Fork des Upstreams, leistet über Pull-Requests einen Beitrag und trifft ihre eigenen Entscheidungen zur Bereitstellung. *(Parallele zu CAC-Punkt 34 „Förderung von Open-Source- Innovation“.)* [ZITATE: Apache 2.0 (Apache Software Foundation); EUPL-1.2 (European Union Public Licence); CC BY 4.0 (Creative Commons).]

Punkt 35. Föderation durch Veröffentlichung. Wo Koordination bei gemeinsamer Technologie, Interoperabilitätsstandards, Reaktion auf Sicherheitsvorfälle oder der Entwicklung von Audit-Rahmenwerken erforderlich ist, erfolgt diese durch offene Veröffentlichung und Konsens unter den beitragenden Installationen. Die internationale Abstimmung erfolgt über W3C, IETF, ISO/IEC und ähnliche etablierte Normungsgremien. Wir würdigen den Wert des Engagements des CAC-Rahmenwerks für Plattformen der Industriekooperation – einschließlich Allianzen im Bereich intelligenter Agenten-Ökosysteme, Technologieprüflabore und gemeinsame F&E-Vereinbarungen – sowie für die Koordination von vor- und nachgelagerten Teilnehmern der Lieferkette bei der gemeinsamen Technologie-F&E, der Festlegung von Standards sowie bei Bewertungs- und Zertifizierungsarbeiten. **Dieser Bereich ist der Arbeitsstrang (iv) des in §II Punkt 4 vorgeschlagenen einheitlichen Ausschusses. Der Ausschuss würde Empfehlungen im neuseeländischen Kontext zu Föderations- und Allianzmodellen für die Industriekoordination entwickeln, zu den Arbeiten des ISO/IEC SC42 an Modellen für die Zusammenarbeit in der KI-Industrie beitragen und einen bilateralen Dialog mit den Autoren des CAC-Rahmenwerks über die Wechselwirkung zwischen föderierter und allianzbasierter Industriekoordination führen.** *(Parallelen zu CAC Punkt 35 „Plattformen für die Zusammenarbeit in der Industrie“; es gilt der konsolidierte Arbeitsstrang zur Ausschussbildung .)* [ZITATE: W3C-Prozessdokument; IETF-Request-for-Comments-Prozess; ISO/IEC 42001:2023 Managementsysteme.]

(II) Stärkung der bilateralen Förderung

Punkt 36. Die Übernahme erfolgt bilateral. Jede souveräne Einrichtung wendet sich direkt an ihre Partner – Partnerorganisationen, Peer-Institutionen, verbundene Peers. Wir erkennen den Wert des Engagements des CAC-Rahmenwerks für Anwendungsförderungskanäle an, darunter Software-Stores für intelligente Agenten, Informationsplattformen für Angebot und Nachfrage in der Industrie, maßgeschneiderte Produktentwicklung über Ausschreibungen und das „Unveil-and-Take-the-Helm“-Challenge-Modell sowie die Entwicklung von Hardware-Systemen und Software für intelligente Agentenprodukte und -dienstleistungen durch Unternehmen. Wir schlagen für den Kontext Aotearoa Neuseeland vor, dass Einführungskanäle aus der bestehenden kommerziellen, zivilgesellschaftlichen und institutionellen Landschaft hervorgehen; staatliche Einrichtungen bauen ihre Geschäftspartnerbeziehungen durch gewöhnliche direkte Zusammenarbeit auf, wobei die öffentliche Beschaffung den Regeln für das öffentliche Beschaffungswesen

folgt. *(Parallelen zu CAC-Punkt 36 „Kanäle zur Anwendungsförderung".)* [QUELLENANGABEN: Vorschriften für das öffentliche Beschaffungswesen (NZ); Überprüfung auf aktuelle Reformen des neuseeländischen Beschaffungswesens steht noch aus.]

Punkt 37. Die Pilot-Einführung erfolgt bilateral und evidenzbasiert. Staatliche Einrichtungen führen Pilotprojekte zur Einführung direkt mit bereitwilligen Gemeinschaften durch. Bestehende MDSL-Implementierungen – Village im Kontext von Gemeinden und Hapū/Iwi; Familiengeschichte im Kontext von Iwi und Diaspora; sydigital im Kontext kleiner Unternehmen – sind Beispiele; spezifische Implementierungsdaten (Anzahlen, Startdaten, Umfang der Nutzung) sollen vor der Veröffentlichung von v1 hinzugefügt werden. Wir würdigen den Wert des CAC-Rahmenwerks, das sich dafür einsetzt, die Erschließung von Anwendungsszenarien für intelligente Agenten in Schlüsselbereichen voranzutreiben, wobei Pilotprojekte in Industrieclustern, Schlüssel branchen und Schlüsselbereichen ein Portfolio von Demonstrationsprojekten aufbauen. Wir schlagen vor, dass für Aotearoa Neuseeland die Pilot-Einführung bilateral zwischen den einführenden Einrichtungen und ihren bereitwilligen Gemeinschaften erfolgt. Wenn Behörden der Krone agentische KI pilotieren möchten, tun sie dies im Rahmen bestehender Verfahren zur Datenschutz-Folgenabschätzung, der Algorithmus-Charta für Aotearoa Neuseeland und den Verpflichtungen des Te Mana Raraunga / Māori Data Sovereignty Network . *(Parallele zu CAC-Punkt 37 „die Erschließung von Schlüssel- Szenarien vorantreiben".)* [QUELLENANGABEN: Belege für den Einsatz von MDSL – Village (Gemeinde- und Gemeinschaftskontexte), Familiengeschichte (Iwi- und Diaspora-Kontexte), Sydigital (Kleinunternehmenskontexte), spezifische Daten stehen noch aus, bis vom Betreiber verifizierte Zahlen vor der Veröffentlichung von v1 vorliegen; Algorithmus-Charta für Aotearoa Neuseeland (2020); Te Kāhui Raraunga (kahuiraraunga.io – Māori-Daten -Governance-Modell und Māori-KI-Governance-Rahmenwerk); Taiuru, K. (20. September 2025) Kritische Analyse der Te Mana Raraunga-Datenprinzipien, taiuru.co.nz/critical-analysis-mana-raraunga/.]

Punkt 38. Internationale Angleichung durch bilaterale Föderation. Souveräne Einrichtungen in Aotearoa New Zealand verbünden sich bilateral mit souveränen Einrichtungen in anderen Rechtsordnungen; die Mitwirkung an internationalen Standards erfolgt über W3C, IETF, ISO/IEC und ähnliche Foren als gleichberechtigte Teilnahme. Wir erkennen den Wert des Engagements des CAC-Rahmenwerks an, das globale Ökosystem durch internationale Plattformen wie die World Artificial Intelligence Conference und der World Internet Conference, die Förderung der Anpassung intelligenter Agenten durch Endgeräte- und Softwareunternehmen sowie das Engagement für die Einhaltung internationaler Vorschriften und die Anpassung an lokale Gesetze, Vorschriften und kulturelle Gepflogenheiten. **Dieser Bereich ist der Arbeitsstrang (v) des in §II Punkt 4 vorgeschlagenen einheitlichen Ausschusses. Der Ausschuss würde Empfehlungen im neuseeländischen Kontext zur internationalen KI-Zusammenarbeit entwickeln, zur internationalen Normungsarbeit von ISO/IEC SC42 beitragen und einen bilateralen Dialog mit den Verfassern des CAC-Rahmenwerks sowie mit internationalen Fachkollegen über die Interoperabilität zwischen bilateralen Föderations- und Plattform-Projektionsansätzen für die internationale Zusammenarbeit führen.** Wir bieten dies als einen Beitrag zu einer internationalen Diskussion im Frühstadium an; Beiträge aus vielen architektonischen Traditionen und politischen

Kontexten werden das Fachgebiet verbessern. *(Parallelen zu CAC-Punkt 38 „das globale Ökosystem aktiv pflegen“; es gilt der konsolidierte Arbeitsstrang zur Ausschussbildung.)* [ZITATE: ISO/IEC JTC 1/SC 42; internationaler Normungsprozess des W3C; Überprüfung der aktuellen bilateralen KI-Abkommen Neuseelands und internationaler Verpflichtungen steht noch aus.]

---

§VI. Sicherstellung der Übernahme

Als zivilgesellschaftlicher Antragsteller koordiniert My Digital Sovereignty Ltd die Umsetzung nicht. Wir nennen hier die Stellen, deren Beteiligung erforderlich wäre, sollte ein Teil dieses Rahmens von Einrichtungen Aotearoa Neuseeland übernommen werden.

Zu den staatlichen Stellen, deren Arbeit von diesem Vorschlag berührt wird, gehören das Ministerium für Wirtschaft, Innovation und Beschäftigung für die digitale Strategie; das Justizministerium für die Angleichung des Rechtsrahmens; das Amt des Datenschutzbeauftragten für die Angleichung an das Datenschutzgesetz 2020; Stats NZ und Te Kāhui Raraunga für die Angleichung der Datenhoheit (mit Dr. Karaitiana Taiuruskritischer Analyse vom 20. September 2025 als grundlegender Referenz); Te Whatu Ora / Health New Zealand für die Steuerung von Gesundheitsinformationen; Te Pūtea Matua / Reserve Bank of New Zealand für die Anpassung der Aufsichtsvorschriften für Finanzdienstleistungen; Waka Kotahi New Zealand Transport Agency für den Verkehr; das Bildungsministerium für den Bildungsbereich; sowie die neuseeländische Police für Fragen der öffentlichen Sicherheit. An der Bewertung durch die Zivilgesellschaft wären natürlich die Royal Society Te Apārangi, Internet NZ, NetSafe, das New Zealand AI Forum sowie akademische Forscher aus den relevanten Disziplinen beteiligt. Die Einbeziehung von Hapū und Iwi ist unerlässlich, wenn Verpflichtungen aus dem Vertrag oder Auswirkungen auf die Siedlungsvereinbarungen auftreten, und die in diesem Vorschlag festgelegte Architektur soll die Arbeit zur Datenhoheit der Māori unterstützen – und wird zur Nutzung im Rahmen — die Arbeit zur Datenhoheit der Māori, wie sie von Te Kāhui Raraunga (Māori Data Governance Model; Māori AI Governance Framework) und in den veröffentlichten wissenschaftlichen ArbeitenvonDr. Karaitiana Taiurudargelegt wird – einschließlich seiner kritischen Analyse vom 20. September 2025, die aufzeigt, warum frühere Rahmenkonzepte für KI-Kontexte unzureichend sind.

Ein internationaler Dialog mit den Autoren des CAC-Rahmenwerks und mit gleichgesinnten Netzwerken für indigene Datenhoheit – FNIGC in Kanada, USIDSN in den Vereinigten Staaten, Maiam nayri Wingara in Australien, GIDA auf internationaler Ebene – würde den Austausch in beide Richtungen bereichern.

My Digital Sovereignty Ltd bekennt sich zu den Elementen der architektonischen Offenheit und Lizenzfreiheit des Vorschlags: Das Tractatus, die Code-BasenVillage und Community sowie zukünftige Beiträge von MDSL werden weiterhin unter freizügigen Open-Source-Lizenzen verfügbar sein, und die Referenzimplementierungen werden im Dialog mit den Anwendern entwickelt. Der Rest richtet sich an diejenigen, die über die Übernahme entscheiden würden.

Wir schließen mit einer ausdrücklichen Einladung: an die Autoren des CAC-Frameworks, an internationale Kollegen, an neuseeländische politische Entscheidungsträger und Community-Organisatoren sowie an alle, die an ähnlichen Fragen arbeiten – Kommentare zu dieser Version 1 sind über die üblichen Kanäle für Papierkommentare auf agenticgovernance.digital willkommen.

Anhang A. Häufige technische Einwände + Antworten

Dieser Anhang sammelt die häufigsten technischen Einwände gegen das Tractatus mit kurzen strukturellen Antworten. Jede Antwort verweist auf die relevante Primitive §0(i) oder einen späteren Abschnitt. Das Framework befindet sich im Forschungsstadium; diese Antworten fassen die architektonische Argumentation zusammen, nicht die technische Fertigstellung.

1. „Wie kann ein LLM irgendetwas durchsetzen?“

Einwand. LLMs sind statistische Mustervervollständiger; gegnerische Eingabeaufforderungen und neu entstehende Fähigkeiten umgehen die Sicherheitsvorkehrungen auf Modellebene.

Antwort: Das Framework verlangt vom Modell nicht, irgendetwas durchzusetzen. Die Tractatus sind Laufzeit-Intercepts – Code im Host-Prozess des Agenten, der entscheidet, ob jede vorgeschlagene Aktion ausgeführt wird, bevor die Ausgabe des Modells die Welt erreicht. Der Intercept hängt nicht vom Compliance- oder Alignment-Training des Modells ab; er wendet konfigurierte Regeln der Entscheidungsklasse an der Aktionsgrenze an, was der gleichen architektonischen Form entspricht wie die Interception privilegierter Systemaufrufe im OS-Kernel. Siehe §0(i) Durchsetzung an der Grenze.

2. „Was ist mit JEPA / alternativen Architekturen?“

Einwand. Das Framework scheint von transformatorbasierten LLMs auszugehen; architektonische Veränderungen (JEPA, Hybridsysteme) könnten es ungültig machen.

Antwort. Das Framework ist substratunabhängig. Seine Primitive wirken auf Agentenaktionen, nicht auf Modellinterna – es ist kein Zugriff auf Gewichte, Aufmerksamkeitsmuster oder Schlussfolgerungsspuren erforderlich. Ein Agent auf einer beliebigen Architektur gibt vorgeschlagene Aktionen aus; die Intercept-Schicht bewertet diese Aktionen anhand derselben Entscheidungsklassen-Grenzen. Siehe §0(i) einleitender Absatz.

3. „Kann der Agent die Intercept-Schicht umgehen?“

Einwand. Ein ausreichend fähiger Agent könnte die Intercept-Schicht modellieren und Aktionen entwickeln, die diese umgehen.

Antwort: Der Agent verfügt nicht über die Berechtigung zur Ausführung – diese liegt bei der Abhörfunktion. Eine Umgehung erfordert entweder (a) einen Vorgang

außerhalb des Zuständigkeitsbereichs des Agenten, was von der Laufzeitumgebung als Berechtigungsverletzung erkannt wird, oder (b) eine vollständige Berechtigungserweiterung gegenüber dem Host-Prozess, was ein Problem auf Betriebssystemebene darstellt, das auf dieselbe Weise abgesichert ist wie Dateisystemberechtigungen. Der Prüfpfad (souveräne, kryptografisch signierte Aufzeichnungen) macht erfolgreiche Umgehungsversuche forensisch sichtbar und verhindert so „Defection-as-Strategy“. Der tiefere strukturelle Punkt – der in *„Architectural Alignment“* §3.4 entwickelt wird – ist die **Unterscheidung zwischen Substrat und Laufzeit**: Selbst wenn ein ausreichend fähiger Agent Laufzeit-Interceptions umgehen kann, lassen sich die Substrat-Mechanismen (kryptografische Herkunft, Föderations-Envelopes, mitgliedergeführte Portabilität von Aufzeichnungen) nicht umgehen, da sie sich in verteiltem Besitz befinden, unabhängig vom Agenten. Die Sicherheit des Substrats ergibt sich aus der Mathematik und der verteilten Replikation, nicht aus der Kooperation des Agenten. Siehe §0(i) Querverweisvalidierung; §II Punkt 5; *Architectural Alignment* §3.4 Substrat vs. Laufzeit; §7.5 Angriffsfläche der sozialen Ebene (die Fläche, die das Substrat *nicht* abdeckt).

4. „Inwiefern unterscheidet sich dies von der Sicherheit durch Prompt-Engineering?“

Einwand. Prompt-Engineering und RLHF schränken ebenfalls die Modellausgabe ein. Das Framework scheint vom Grundgedanken her ähnlich zu sein.

Antwort. Strukturell unterschiedlich. Prompt-Engineering und RLHF verändern die Verteilung der Modellausgaben, lassen aber die statistische Mechanik unverändert. Die Primitive des Frameworks laufen vor dem Aufruf des Modells (Capability-Scoping), nach der vorgeschlagenen Aktion (Boundary Enforcement) oder parallel zum Aufruf (Cross-Reference-Validation) – keine davon hängt davon ab, dass das Modell die richtige Ausgabe erzeugt. Sie hängen davon ab, dass die Laufzeitebene die Zugehörigkeit zur Entscheidungsklasse korrekt identifiziert. Siehe §0(i) Boundary Enforcement und Metakognitive Verifikation.

5. „Was passiert, wenn der Laufzeitdienst selbst ausgenutzt wird?“

Einwand. Das Vertrauen in einen Laufzeitservice verlagert die Angriffsfläche, anstatt sie zu beseitigen.

Antwort. Das Framework behauptet nicht, dass Laufzeitdienste unhackbar sind. Es behauptet, dass ein Missbrauch nachweisbar ist und dass diese Nachweise den Schadensumfang begrenzen.

Wie eine Sicherheitsverletzung aussieht. Entweder wird der Dienstcode ausgenutzt – ein Angreifer erlangt Zugriff auf die Abfrage, um Aktionen außerhalb der Richtlinie zu genehmigen – oder der Richtlinienstatus, den der Dienst abfragt, wird umgeschrieben – ein Angreifer ändert, welche Entscheidungsklassen zur menschlichen Genehmigung weitergeleitet werden. In beiden Fällen ist es das Ziel des Angreifers, eine „Weiterleitung an den Menschen“-Entscheidung in eine „automatische Genehmigung“-Entscheidung umzuwandeln, ohne dass der Betreiber dies bemerkt.

Was ist gefährdet? Entscheidungen, die das Framework andernfalls zur menschlichen Genehmigung weitergeleitet hätte – Werturteile, irreversible Vorgänge, mandantenübergreifender Datenzugriff. Der Schadensumfang ist durch das begrenzt, was der Intercept bereits genehmigen durfte: Das Framework erteilt Genehmigungsberechtigungen, keine neuen Privilegien zur Ausführung von Aktionen. Ein ausgenutzter Intercept kann keine Daten exfiltrieren, auf die der Agent von vornherein keinen Zugriff hatte; er kann nur Aktionen innerhalb des bestehenden Zuständigkeitsbereichs des Agenten fälschlicherweise genehmigen.

Abhilfemaßnahmen. Drei Eigenschaften kommen zusammen. (i) Jede Framework-Entscheidung wird in den Audit-Trail des Sovereign-Records geschrieben – kryptografisch signiert, nur anhängbar, föderativ an Peers repliziert –, sodass eine nachträgliche forensische Analyse rekonstruieren kann, was genehmigt wurde, durch welche Version des Dienstes und unter welchem Richtlinienstatus. Das Zeitfenster für Sicherheitsverletzungen ist zeitlich und vom Umfang her begrenzt. (ii) Die Validierung durch Querverweise (§0(i)) erkennt Abweichungen zwischen beobachteten Genehmigungen und der deklarierten Richtlinie nahezu in Echtzeit und deckt Verstöße auf, bevor sie zur Normalität werden. (iii) Die Replikation innerhalb der Föderation verhindert, dass ein Angreifer, der einen Knoten kontrolliert, Datensätze rückwirkend löscht; ein Überlaufen erfordert die Kollaboration mit der Föderation, nicht die Kompromittierung eines einzelnen Dienstes.

Der Vertrauensanker kann versagen, aber der Ausfall ist begrenzt, beobachtbar und forensisch rekonstruierbar – dasselbe Architekturmuster wie bei der Zertifikatstransparenz für die TLS-PKI: Der Vertrauensanker (eine Zertifizierungsstelle) kann kompromittiert werden, aber das Audit-Protokoll der ausgestellten Zertifikate macht die Kompromittierung global sichtbar.

Die Überlebensstrategie ist mehrschichtig. Die Laufzeitdienste des Frameworks (BoundaryEnforcer, die §0(i)-Primitiven) sind *agentenorientiert*: Sie schränken ein, was der Agent autorisieren kann. Die Sovereign-Records-Architektur (Paper A) ist *substratorientiert*: Sie stellt sicher, dass die Datensätze den Agenten überdauern, unabhängig davon, ob der Agent das Tor umgeht. Siehe *„Architectural Alignment“* §7.4 (Überlebensfähigkeit unabhängig von der Eindämmung des Agenten) und §7.5 (die Angriffsfläche der sozialen Ebene, die das Substrat nicht schließt – Überredung, massenkoordinierte Identitätsfälschung, synthetisierte Zustimmung – hier als offene Grenze bezeichnet). Das der Audit-Kette zugrunde liegende PKI-Signaturschema hat einen Horizont der Quantenanfälligkeit (10–30 Jahre); die NIST-Standards für postquanten-sichere Signaturen wurden im August 2024 fertiggestellt, und der Migrationspfad folgt den Standards, wobei Fälschungen auf Datensatzebene selbst auf einem CRQC kostspielig sind und Föderations-/Portabilitätsmechanismen nicht von der Signaturintegrität abhängen. Siehe *Paper A* §5.3.

Siehe §II Punkt 5; §0(i) Querverweisvalidierung.

6. „Was passiert, wenn sich Werte ändern?“

Einwand. Die Grenzen der Entscheidungsklassen legen aktuelle Werte fest; die Werte von Gemeinschaften entwickeln sich weiter, was das Framework anfällig macht.

Antwort. Entscheidungsklassen sind Konfiguration, keine Architektur. Das Rahmenwerk stellt den Intercept-Mechanismus bereit; welche Aktionsklassen zur menschlichen Genehmigung weitergeleitet werden, ist pro Mandant vom Betreiber editierbar (§III Punkt 3 – anwendergesteuerte Governance). Wenn Stakeholder innerhalb eines Governance-Bereichs unvereinbare Grenzpositionen vertreten, strukturiert die Primitive zur Orchestrierung pluralistischer Beratungen die Beratung, anstatt einen Gewinner zu bestimmen. Das Rahmenwerk ist darauf ausgelegt, sich entwickelnde Werte zu beherbergen, nicht sie einzufrieren. Siehe §0(i) Pluralistische Deliberationsorchestrierung; §III Punkt 3.

Lizenz und Zitierweise

Vorgeschlagene Zitierweise: Stroh, J. G. (2026). *Ein zivilgesellschaftlicher Vorschlag für souveräne und föderierte agentische KI in Aotearoa New Zealand* (v1.2, Mai 2026, überarbeitet gemäß Ted Howards Korrespondenz zu v1.1). My Digital Sovereignty Ltd. https://agenticgovernance.digital/papers/aotearoa-nz-agentic-ai-framework-v1.2-may-2026.html

Kommentare und Korrespondenz: Substanzielles Feedback zu bestimmten Abschnitten ist willkommen. Bitte geben Sie die Abschnittsnummern an (z. B. §III Punkt 5), damit Korrekturen nachvollziehbar sind. Der Autor antwortet persönlich; rechnen Sie bitte mit ein bis zwei Wochen. E-Mail: john.stroh@mysovereignty.digital.