12-2-2023

IndigiCloud – Enacting Māori Data Sovereignty

Ian Welch
Victoria University of Wellington, New Zealand, ian@ecs.vuw.ac.nz

Kevin Shedlock
Victoria University of Wellingon, New Zealand, kevin.shedlock@vuw.ac.nz

IndigiCloud – Enacting Māori Data Sovereignty

Research-in-progress

Ian Welch
School of Engineering and Computer Science
Te Herenga Waka Victoria University of Wellington
North Island, New Zealand
Email: ian.welch@vuw.ac.nz

Kevin Shedlock
School of Engineering and Computer Science
Te Herenga Waka Victoria University of Wellington
North Island, New Zealand
Email: kevin.shedlock@vuw.ac.nz

Abstract

The advent of cloud technologies has brought critical questions about data sovereignty to the fore, particularly for Indigenous communities. This research-in-progress paper focuses on the challenges of Māori data sovereignty and introduces IndigiCloud, a prototype for a community cloud infrastructure designed to uphold Māori data sovereignty. Leveraging an open-source software stack, IndigiCloud aims to allow for data management under Māori governance in accordance with the recent Māori Data Governance Model report. We have developed a prototype that the community can use as part of a co-design process to align with collective priorities. This paper contributes to the larger discourse on Indigenous data sovereignty by offering a tangible prototype for the Māori community while acknowledging the prototype as a work in progress, subject to further refinement through community engagement.

Keywords Māori Data Sovereignty, Community Cloud, Data Governance, Indigenous Communities, Open-source Technologies.

1 Introduction

In an increasingly digitised world, the importance of data sovereignty for Indigenous communities, particularly the Māori, cannot be overstated. As more data migrates to cloud-based storage and processing solutions, questions surrounding control, access, and rights over this data become paramount. This paper outlines research-in-progress work exploring the potential benefits and challenges of employing a community cloud approach with an open-source stack to enhance Māori data sovereignty.

2 Literature Review

2.1 Data Sovereignty Concepts

The Native Nations Institute (NNI) are located on Tohono O'odham Nation traditional homelands and was founded in 2001 by The University of Arizona and the Morris K. Udall and Stewart L. Udall Foundation as a self-determination, self-governance, and development resource for Native nations. We use their definitions of "data sovereignty" and "indigenous sovereignty" (Caroll et al. 2017). The NNI define the mainstream usage of sovereignty as "data sovereignty is the concept that information which has been converted and stored in binary digital form is subject to the laws of the country in which it is located". Indigenous data sovereignty (IDSov) is "… the right of Native nations to govern the collection, ownership and application of its data." The key difference is that this recognises rights negotiated between nations and settler-colonial governments. This definition is centred in Native nations located in North America but has been adopted in the Aotearoa context.

Te Mana Raraunga (the Māori Data Sovereignty Network) advocates for the realisation of Māori rights and interests in data and the ethical use of data. Members include Māori and iwi data users, Information Computer Technology (ICT) providers, researchers, policymakers and planners, businesses, service providers and community advocates. It was formed following a workshop on Data Sovereignty for Indigenous Peoples (UNDRIP) that built upon previous workshops organised by the United Nations Permanent Forum on Indigenous Issues (UNPFII) on 'data collection and disaggregation' (in 2004), on 'indicators of wellbeing' (in 2006) and on 'development with culture and identity' (in 2010) (Te Mana Raraunga, 2016).

Māori data is considered a living tāonga (treasure), it is data produced by Māori, or that is about Māori and the environment with which they have relationships. Māori data is protected by both the Treaty /Tiriti of Waitangi (1840 Treaty of Waitangi),and the UN's Declaration on the Rights of Indigenous Peoples (United Nations, 2007), affirming the rights of Māori people over their data. Māori Data Sovereignty (MDSov) asserts that Māori data should be governed by Māori people themselves, in line with their customs and values. This form of sovereignty is instrumental in upholding tribal sovereignty and advancing the aspirations of Māori communities (Te Mana Raraunga, 2016; "Te Mana Raraunga" 2022).

Tino rangatiratanga (absolute chiefly authority) is a concept central to MDSov. As explained by Kukutai and Cormack, tino rangatiratanga exists whether the state recognises it or not (Kukutai and Cormack 2021, p. 23). Additionally, the independent Waitangi Tribunal has confirmed that sovereignty was never ceded in the Māori version of the Treaty (Te Tiriti) and it was the Māori language verson signed by the majority. Kulutai and Cormack also highlight two closely aligned concepts are mana motuhake ("self-determination") and mana whenua ("territorial rights associated with long-term occupation").

2.2 Māori Data Governance

Data governance refers to the processes, policies, roles, responsibilities, and metrics that ensure the effective and efficient use of information in enabling an organization to achieve its goals. It encompasses the people, processes, and technology required to create a consistent and proper handling of an organization's data across the business enterprise (Khatri and Brown 2010, p. 149).

Māori data governance (MDGov) is required for a resilient and trustworthy data system that meets MDSov needs and aspirations. It defines the processes, practices, standards and policies that enable Māori, as collectives and as individuals, to have control over Māori data (Kukutai, Tahu et al. 2023, p. 3). Māori and the Aotearoa public sector have been working on the co-design of a MDGov model since 2021. The MDGov co-design process has been co-led by Data Iwi Leaders Group (Data ILG) and New Zealand's official data agency Stats NZ Taturanga Aotearoa (data.govt.nz 2021). Phase I saw consultation and engagement through Māori data wānanga (seminar or conference) involving te ao Māori and public service agencies. Phase II has produced a The recent Māori Data Governance Model

report (Kukutai, Tahu et al. 2023) describes a Māori Data Governance model that Māori data experts have designed for use across Aotearoa by both Māori and public agencies. Phase III is expected to see te ao Māori (the Māori world) resourced by government to accelerate the critical steps needed for an autonomous Māori data system that supports the concept of mana motuhake.

The model outlines Eight Data Pou or pillars that define critical areas of data governance and specify the actions that should be undertaken to realise six desired outcomes. The pou from the Māori Data Governance report are:

- Pou #1 – Data capacities and workforce development – invest in anti-racist and decolonial ways of working with data as well as resourcing Māori to lead development and training of a sustainable Māori workforce.
- Pou #2 – Data infrastructure – shared decision-making about ongoing investment in data infrastructure, invest in and support the development of mana motuhake data infrastructure, and offer technology options that power choices close to where decisions are made.
- Pou #3 – Data collection – use practices that maintain and ideally strengthen relationships, only collect necessary data have protocols that guide ethical practice.
- Pou #4 – Data protection – the scope of data protection needs to be broader than personal privacy to include collective privacy, data classified as Māori needs to be subject to data security procedures guided by Māori leadership and expertise, and there needs to be a proactive approach to strengthening local infrastructure so that data can be stored locally.
- Pou #5 – Data access, sharing and repatriation – rather than a "one off" event, data access should be treated as a process, barriers to access by Māori to Māori data should be removed, Māori should be in control of the sharing of Māori data, and repatriation of data back to where it belongs should be supported.
- Pou #6 – Data use and reuse – agencies need to address issues of consent, reframe data analysis to meet Māori priorities, and exercise upmost care when developing algorithms and associated decision matrices.
- Pou #7 – Data quality and system integrity – shift from focusing only on data accuracy to fitness of use in terms of user perspectives.
- Pou #8 – Data classification – new ways of making sense of data are needed that both recognise pre-existing structures of thought in te ao Māori but also looking forward to future possibilities.

2.3 Māori Data Applications

What are the types of Māori data applications that have been developed that might be seen a exemplars of the MDGov model outlined above? In this section, we highlight two current projects and briefly explain how they meet the requirements.

The first is Te Whata (a whata is a non-carved storehouse), a storehouse of iwi data developed by Data ILG and Te Kāhui Raraunga Charitable Trust and has been resourced by Stats NZ under the Mana Ōrite relationship agreement between the parties. It is a data platform for iwi members, technicians and leaders. Data sets from StatsNZ, the Ministry of Education and the Ministry of Health have been provided to Te Whata. The fact that this is a Māori led project and provides access to iwi for their own purposes supports pou #2, #4, #5, #7 and #8. The resourcing suggests that #1 was met although the development was by two design agencies and it is unclear if they are Māori or not. It also isn't clear if the data is actually stored locally to the iwi who are using it and future work by our team will be to contact the provider and ascertain this information.

The second is Āhau ("Āhau | I Am" n.d.), a decentralised system for archiving traditional knowledge and storing genealogy and whakapapa developed by Kaye-Maree Dunn (Te Rarawa, Ngā Puhi) and Ben Tairea (Ngāti Nurou, Kuki Airani). The focus has been on co-creation and co-design with incorporation of tikanga (customs and traditional values). Data can be stored on their own device or within their own chosen networks. As part of Āhau is a self-hosted cloud server called Pātaka, and future extensions include considering the application of blockchain technology. This is a Māori led project with Māori developers who have placed tikanga and involvement with users at the forefront of their approach and their data can be stored on a self-hosted local server with data being curated and controlled by Māori. The meet the requirements of pou #1-8.

2.4 Local Data Infrastructure

Both Te Whata and Āhau substantially meet the requirement for a MDGov model. However, the area that they do not address is how to strengthen the infrastructure that they use. This relates substantially to the issue of strengthening local data infrastructure mentioned that is part of pou #4.

It is not clear that Te Whata is hosted locally to the iwi and it does not address that issue directly. As mentioned in pou #4, data should be stored locally rather than offshore because countries such as USA and China assert jurisdiction over stored by companies headquartered in their respective countries. This even applies if the data is on a local server farm as long as the company owning the farm is headquartered in these two countries.

We argue that the data should be stored as locally as possible to the users of the data. This means in the case of whānau or hapū data stored at the relevant marae. This is indeed the case with Āhau who have a Pātaka server that be installed on a personal computer, for example within the home of whānau or at te marae. What is missing is guidance on how best for Māori running their own infrastructure to secure it against data security risks such as ransomware where data can be made unavailable through encryption or misuse of the infrastructure to send spam or carry out crypto mining. These risks are common in Aotearoa as shown in the most recent CERT NZ cyber security insights report (CERT NZ, 2023)

We have been working at the marae level to investigate what is required through a prototyping approach. This involves co-design of a local cloud infrastructure based on readily available open source components that we call IndigiCloud. This has involved working with the marae to initially establish a base infrastructure for future research. Our aim is to provide an infrastructure for securely hosting Māori data applications such as Te Whata and Āhau.

Our approach is a kaupapa Māori methodology modelled upon work by Shedlock & Vos (Shedlock and Vos 2018). A key concept here is “Mahi atu, mahi mai” (give and take), applied to research to achieve mutual agreement and benefit during technology research. We have been meeting over the last two years with marae workers to define initial requirements and to understand their existing infrastructure. We have supplemented fortnightly meetings with visits to the marae and visits from marae workers to our university. We have done work for the marae to establish the relationship and share the resources available from our institution.

3 IndigiCloud prototype

This section outlines the prototype infrastructure called " IndigiCloud " implemented at a marae. This is a fairly technical description, we have included this so it can be replicated elsewhere.

The server uses a Linux Operating System that has been hardened against network security threats by removing unnecessary services that might provide an entry point for attackers. We have also chosen Ubuntu Server 22.04 LTS because OS security updates are guaranteed until April 2027.

We use a content management system (CMS) to access authenticated users via a web interface. We have chosen Silverstripe 4.0, this open-source CMS developed and maintained within Aotearoa and provides an “out of the box” solution that can be customised

The CMS runs within a Virtual Machine, and we are using Oracle Virtual Box 7.0, which is open source again. We use a Virtual Machine to prevent access to the underlying Linux operating system in case an attacker manages to compromise the CMS.

Data is stored on a Network Attached Storage device, a Synology DiskStation 8-bay networked attached storage (NAS) server with a Quad Core Ryzen processor. There are four 12-TB drives configured as a redundant array of independent disks that use disk striping with parity (RAID-5). This configuration allows one failed drive to be removed without any system downtime. This comes at the cost of reducing the total storage to 36 TB but with the benefit of redundancy and reliability.

In this case, a managed Ubiquiti Unifi switch is used to create a private network for the data server and the networked attached storage. We have this configured so that all access to the network-attached storage is mediated by the CMS and Linux operating system. This allows acceptable grained security policies to be imposed and protect against direct access to data from attackers on the Internet.

A router/firewall, in this case, a Spark router that connects to fibre internet. A static IP address is used, and web traffic is forwarded to the CMS. Firewall rules are in place to block all traffic except SSH for remote management of the data server and HTTP/S for web access to the CMS.

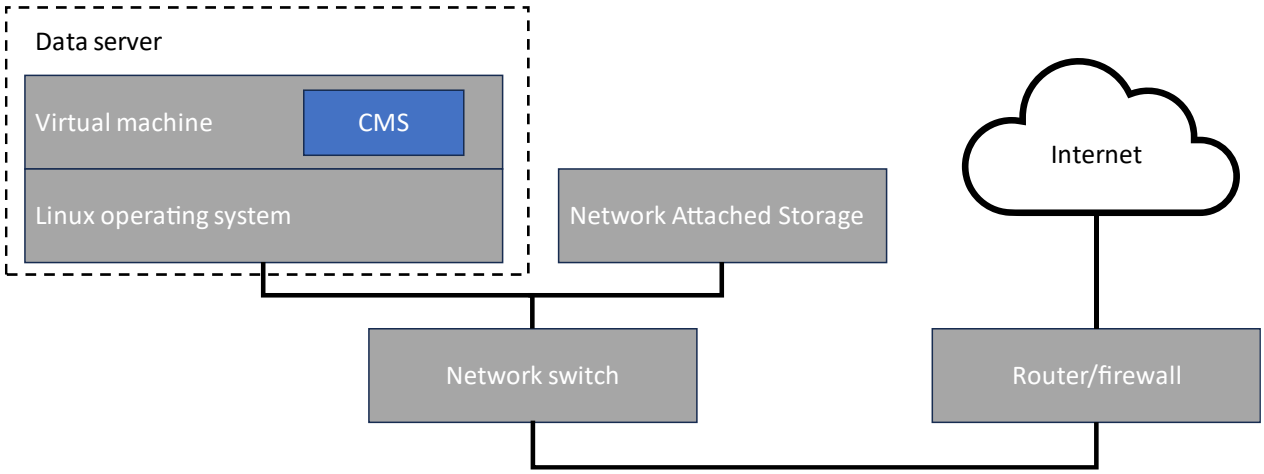

Figure 1: System Architecture for IndigiCloud

4 Open questions

The prototype as designed is a platform for exploring pou #4. It is also a physical artefact that we can use in discussions to refine answers to questions about how best to secure the platform. An important aspect of securing the platform is addressing the issue of resourcing. As mentioned with respect to the Māori Data Governance Model, the expectation is that Māori will be resourced. However, as with the rest of the world there is a lack of cybersecurity professionals available to maintain infrastructure. Worldwide there is a critical need for cybersecurity professionals with a worldwide gap of 3.4 million cybersecurity workers (ISC2 2023).

This need to work with limited resources raises the following general questions:

- How to implement security in such as way that the system requires minimal security management? There is a world-wide shortage of trained security and system administrators so there is an urgent need to automate security management where appropriate. We believe that automation, secure-by-design and secure architecture would be appropriate here.
- How to best define use cases that encapsulate user roles, goals, actions, and specific conditions like data accessibility and logging needs (Goldman and Song 2005). Issues to be considered are parameters for uploading, accessing, and modifying files, including size limitations.
- What does it mean to develop community capability for ongoing maintenance and further development? Training is another focal point and must meet the needs of local administrators and community members.
- Is there a role for community-centric security solutions? These would be designed around their priorities and using appropriate cultural language. Further considerations include physical security, system availability, disaster recovery procedures, and responses to security incidents.
- How do we deal with the obsolescence of hardware? A benefit of a local cloud is that you are not paying to keep data available, unlike a public cloud. However, there does need to be a plan for upgrading and replacing hardware to ensure continued availability.,
- What are economic opportunities for the community? Can this approach stimulate local innovation, fostering job creation in IT sectors, software development, and data management tailored to Māori needs?

5 Conclusion

This paper has outlined a new research project to investigate the methodology for developing an open-source data infrastructure to support Māori data sovereignty. We have outlined three essential key Pou to implement and provided an overview of two related data sovereignty projects. We believe that building a prototype using a “Mahi mai, mahi atu” kaupapa Māori methodology will help us to surface and work towards investigating several research problems related to practical data sovereignty. The

work presented should be considered a work in progress and is likely to evolve from its first iteration in this paper.

6 References

Caroll, S. R., Lonebear, D. R., and Martinez, A. 2017. "Policy Brief: Indigenous Data Sovereignty in the United States | NNI Database," Arizona Native Nations Institute. (https://nnigovernance.arizona.edu/policy-brief-indigenous-data-sovereignty-united-states).

CERT NZ,. 2023. "Quarter One Cyber Security Insights 2023," *Quarter One Cyber Security Insights 2023*. (https://www.cert.govt.nz/about/quarterly-report/quarter-one-cyber-security-insights-2023/, accessed October 23, 2023).

data.govt.nz. 2021. "Co-Designing Māori Data Governance - Data.Govt.Nz," *Co-Designing Māori Data Governance - Data.Govt.Nz*. (https://www.data.govt.nz/toolkit/data-governance/maori/, accessed October 23, 2023).

Goldman, J. L., and Song, I.-Y. 2005. "Organizing and Managing Use Cases," in *Perspectives in Conceptual Modeling*, Lecture Notes in Computer Science, J. Akoka, S. W. Liddle, I.-Y. Song, M. Bertolotto, I. Comyn-Wattiau, W.-J. van den Heuvel, M. Kolp, J. Trujillo, C. Kop, and H. C. Mayr (eds.), Berlin, Heidelberg: Springer, pp. 43–52. (https://doi.org/10.1007/11568346_6).

ISC2. 2023. "Cybersecurity Workforce Study." (https://www.isc2.org/research).

Khatri, V., and Brown, C. V. 2010. "Designing Data Governance," *Communications of the ACM* (53:1), pp. 148–152. (https://doi.org/10.1145/1629175.1629210).

Kukutai, T., and Cormack, D. 2021. "'Pushing the Space' Data Sovereignty and Self-Determination in Aotearoa NZ," in *Indigenous Data Sovereignty and Policy* (1st ed.), Routledge Studies in Indigenous Peoples and Policy, 2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN: Taylor & Francis Group, pp. 21–35. (https://lccn.loc.gov/2020020862).

Kukutai, Tahu, Campbell-Kamariera, K., Mead, A., Mikaere, K., Moses, C., and Whitehead, C. 2023. "Māori Data Governance Model," Te Kāhui Raraunga.

Shedlock, K., and Vos, M. 2018. "A Conceptual Model of Indigenous Knowledge Applied to the Construction of the IT Artefact," in *Proceedings of the 31st Annual CITRENZ Conference*, pp. 1–7.

Te Mana Raraunga,. 2016. "Te Mana Raraunga Charter," Te Mana Raraunga, p. 5. (https://static1.squarespace.com/static/58e9b10f9de4bb8d1fb5ebbc/t/5913020d15cf7dde1df34482/1494417935052/Te+Mana+Raraunga+Charter+%28Final+%26+Approved%29.pdf).

"Te Mana Raraunga." 2022. *Te Mana Raraunga*, , July 11. (https://www.temanararaunga.maori.nz, accessed August 25, 2023).

United Nations,. 2007. *United Nations Declaration on the Rights of Indigenous Peoples*, Declaration. (https://www.un.org/development/desa/indigenouspeoples/declaration-on-the-rights-of-indigenous-peoples.html).

Acknowledgements

This research is supported by the Cyber Security Research Programme|Artificial Intelligence for Automating Response to Threats from the Ministry of Business, Innovation, and Employment (MBIE) of New Zealand as part of the Catalyst Strategy Funds under the grant number MAUX1912.

Copyright