MAIN PAPER

The five tests: designing and evaluating AI according to indigenous Māori principles

Luke Munn[1]

Abstract

As AI technologies are increasingly deployed in work, welfare, healthcare, and other domains, there is a growing realization not only of their power but of their problems. AI has the capacity to reinforce historical injustice, to amplify labor precarity, and to cement forms of racial and gendered inequality. An alternate set of values, paradigms, and priorities are urgently needed. How might we design and evaluate AI from an indigenous perspective? This article draws upon the five Tests developed by Māori scholar Sir Hirini Moko Mead. This framework, informed by Māori knowledge and concepts, provides a method for assessing contentious issues and developing a Māori position. This paper takes up these tests, considers how each test might be applied to data-driven systems, and provides a number of concrete examples. This intervention challenges the priorities that currently underpin contemporary AI technologies but also offers a rubric for designing and evaluating AI according to an indigenous knowledge system.

Keywords AI · Artificial intelligence · Machine learning · Indigenous · Decolonial · Maori · Social justice

1 Introduction: the case against current AI

Artificial intelligence technologies (AI) are being rapidly deployed across an array of high-stakes areas, from welfare to law enforcement, healthcare, and recruitment. For technology pundits, this transformation is a positive one, accelerating innovation and ushering in progress and prosperity (Brynjolfsson and McAfee 2011, 2014). But more critical research has highlighted the social and environmental fallout of AI-driven shifts, its ability to extract capital in novel ways while increasing precarity and inequality. Digital platforms allow homework under a piecework model, a highly exploitative form of labor (Dubal 2020). AI systems can perpetuate gendered stereotypes and contribute to racial injustice (Buolamwini and Gebru 2018; Benjamin 2019). AI systems meticulously track workers, rewarding and punishing individuals based on their performance (Munn 2017). And if workers have suffered, so too has the environment, as high-carbon, high-energy technologies consume natural resources and enact a heavy toll on a warming planet (Munn 2022b).

So while AI is novel, it often seems to continue long-standing paradigms of technology in the service of capital, reducing agency and autonomy (Marx 1977), increasing the precarity of labor (Berardi 2009), undermining the well-being of workers (Huws 2014), and amplifying forms of racialized and gendered inequality (Noble 2018; Beller 2018). This scholarship suggests that the human harms documented in recent AI-driven initiatives are not merely "teething problems," but part of a broader paradigm of capitalist and colonialist values at the core of our current economic and technological systems.

Where do these values come from? AI technologies, like any technology, are embedded with certain values, norms, and priorities drawn from a particular (colonial) history and a particular development environment (white, male, patriarchal, heteronormative). Beller (2018) has shown how the historical development of computation, broadly understood, was intimately connected with capital and its drive to instrumentalize racialized and gendered difference. This means that the values at the heart of contemporary AI systems are not neutral or disinterested, but rather particular and purposeful. Far from being universal, artificial intelligence can be better understood as "artificial Western

✉ Luke Munn
  l.munn@uq.edu.au

1 Digital Cultures & Societies, University of Queensland, Mianjin/Brisbane, Australia

ethno-intelligence" (Williams and Shipley 2020). For this reason, MacQuillan (2019) suggests that whenever AI is adopted without constraints, it will amplify the injustice of the status quo.

Because current AI technologies and their values compound inequality and injustice, an alternate set of values, paradigms, and priorities are urgently needed. AI developers cannot carry out the deep transformations needed to support inclusivity and sustainability while continuing to draw on the same hegemonic epistemological and ethical systems. As AI systems grow in power and permeate further into high-stakes domains of political and social life, the question we are faced with becomes more stark. Will AI technologies continue to extract personal data, to exacerbate inequalities of wealth and power, and to render life more precarious for some of the most marginal and vulnerable (Mejias and Couldry 2019; Ciston 2019; Checketts 2022)? Or can we welcome new knowledge paradigms, establish alternate priorities, and progress towards technologies that underpin care for each other and for the earth in crisis?

2 Towards indigenous AI

Where can we draw an alternate set of AI priorities and principles from? Indigenous cosmologies, epistemologies, and ways of being and doing provide one promising approach. Williams and Shipley (2020) suggest that AI applications might benefit from indigenous wisdom, augmenting the often narrow Western focus on utility and efficiency with concepts such as harmony with others, deeper ecological understanding, and close kinship networks. Similarly, Irwin and White (2019, 1) assert that "indigenous philosophy has a lot to offer the world, as we face the necessary shift from an exploitative, extractive economy, to a more sustainable one."

The aim here is not to "diversify" (in a superficial sense) AI technologies, nor to "solve" or streamline existing AI processes, but to instead radically challenge the foundational assumptions of these technologies. At the same time, this article seeks to do more than critique or debunk. Indeed one of the motivations for this research is that the 5 tests are actionable or operationable, suggesting concrete ways they might be employed in design and evaluation.

This article builds on very recent work exploring indigenous AI alternative approaches to AI technologies. Lewis et al. (2020) carried out workshops and interviews with a number of indigenous groups across Aotearoa, Australia, North America, and the Pacific to develop a rich position paper concerning indigenous protocols and artificial intelligence. While promising, the authors acknowledge that this research is very much in progress. The paper's diverse

mixture of technology descriptions, design guidelines, artworks, and poetry reflect this nascent quality.

Some of this research has started to consider how AI technologies might be conceived and constructed. After working closely with Aboriginal technologists, Abdilla et al. (2021) have shared their insights about how an indigenous-centered AI might be developed. The authors argue that indigenous AI should be regional in its conception and development, be guided by local indigenous laws, and be designed with future cultural and technical interrelationships in mind.

Similarly, but in a Māori context, Shedlock and Hudson (2022) have offered a kaupapa Māori model for the creation of IT artifacts, including AI applications. The duo argue that current approaches to AI reproduce colonial paradigms and historical inequalities. What is needed is a solution developed by Māori and for Māori. The model has three core components: adequately framing the purpose and aim of the artifact according to Māori knowledge systems; meaningfully engaging with end-users during the design process; and maintaining accountability and rapport with communities over the lifetime of the project. These interventions begin to move from theory to practice, exploring how indigenous values might be designed into products and services.

3 The five tests

How might we begin to design and evaluate AI from an indigenous perspective? This article offers one approach by taking up the work of Māori scholar Sir Hirini Moko Mead. Mead is a highly regarded anthropologist, historian, and prominent Māori leader who has founded an indigenous tertiary institution and also represented several iwi in disputes. Mead (2016) unpacked key Māori concepts, practices, and paradigms in his groundbreaking *Tikanga Māori* book, which seeks to provide guidance about tikanga Māori or the correct way of doing something.

After stepping through each of these concepts, Mead closes the book with what he calls the five tests. Mead recognises that there are new global issues that will constantly emerge, from surrogate motherhood to same-sex marriage and genetic engineering. These novel issues have not been encountered before and are not explicitly dealt with by Tikanga Māori or Mātauranga Māori. An existing Māori position cannot simply be plucked from history or tradition. It must be discovered.

To aid in this discovery, Mead (2016, 336) offers the five tests as a framework of assessment, "a method or methods for assessing a situation or event that challenges our thinking and our values." Mead (337) stresses that this process results in *a* position, not *the* definitive position. Different people and communities will make the assessment in different ways. Despite this disclaimer, a tikanga Māori framework can be

immensely helpful in considering a controversial new issue from a Māori (vs a Western or non-indigenous) perspective, working through the benefits and risks, and arriving at a viewpoint. The next sections step through each test, explain key concepts, and discuss how they could be applied to AI technologies.

3.1 Test 1: Tapu

Tapu is frequently translated as sacred and refers to people or places that are special. Mitira (1990) suggests that tapu is better understood and translated as "prohibited" as the rules of tapu are rules of negation or prohibition. Other words associated with tapu in English would be restricted, set apart, or forbidden. *Noa* is often understood to be a contrasting concept to Tapu. Noa designates something which is common, ordinary, or everyday. However, Mead (2016, 33) stresses that noa is not the complete absence of tapu, but rather the idea that a safe balance has been reached: tapu has dropped to a normal level.

When objects or people become tapu, that sacred state must be carefully maintained. This maintenance is accomplished by adhering to a set of strict codes and practices. For example, Mitira (1990) recounts that a priest under heavy tapu could not go near a cooking house or touch food with his hands; this person could not even be approached by someone who was non-tapu. Historically, tapu thus set up a set of binding laws that extended throughout the Māori social space. These tapu laws were taken very seriously and punishment for breaking them or disregarding them could be severe. Tapu thus functioned as a strong form of social and behavioral discipline and maintained a sense of order within a particular tribe or group.

How might we apply this test to AI technologies? I suspect that Mead places this test first because it is foundational. A technology that flagrantly violates tapu is a non-starter. For instance, in Māori cultures, the deceased have a high tapu status and are set apart from others. Yet as Taiuru (2020) notes, in facial recognition databases, images of the living and the dead are stored together. Indeed, Taiuru (2020) compares the government's present-day surveillance and collection of headshots as akin to the colonial practice of collecting Māori heads or *mokomokai*. By failing this first test, such technologies may not be worthy of further debate. Of course, there may be exceptional circumstances where a breach of tapu can be rationalized based on other benefits, and in this case further tests are required.

Fundamentally, this test is about recognising and maintaining tapu principles instead of breaching them. Such an awareness, divorced from the typical Western concerns of optimisation and efficiency, poses a set of novel and non-trivial challenges. The first is the ability to understand the persons, places, and things that may be tapu. This state may be permanent, as in the case of a *rangitira* (chief); daily interactions with this leader are subject to a set of protocols. Alternatively, this state might be temporary: a place where someone has drowned is tapu until a ceremony is conducted to lift that state.

Operationalizing this knowledge might take the form of a database or set of metadata that contains core concepts or conditions regarding what is tapu. In other words, the information ontology of the model (Guarino 1998), which sets out a kind of world-view of objects, relationships, and events, would need to "know" about tapu and what triggers this condition. This foundational knowledge exists outside or beyond the Western canon, and consists not only of scholarly literature but also of insights emerging from oral histories, life experience, and other indigenous knowledge practices. Codifying this knowledge, then, would mean engaging meaningfully with Māori practitioners and experts to develop a socially nuanced yet operationalizable understanding of tapu.

Outside the immediate context of Aotearoa, this principle suggests that AI technologies must be culturally aware and culturally sensitive. Some of the most popular and pervasive technologies over the last two decades have emerged from Silicon Valley. This is a culture historically dominated by wealthy white male engineers, who have embedded their worldviews, norms, and values at the heart of our information technologies. The result is that this Silicon Valley doctrine (Jiménez 2020) becomes a dominant perspective that is then universalized as platforms, software, and services are taken up across the globe. Tapu provides a concrete antidote to this universalizing tendency. AI technologies must be designed with particular people and places in mind. It is both arrogant and insufficient to assume that one model is sufficient for a global audience. Instead, AI developers must be attuned to the needs of a specific community. This means engaging with that community, understanding key practices and concepts, and co-designing solutions that benefit a particular set of stakeholders.

3.2 Test 2: Mauri

Mauri is "the spark of life, the active component that indicates the person is alive" (Mead 2016, 53). Other definitions echo this concise shorthand, with Te Aka Māori Dictionary defining mauri as the "life principle, life force, vital essence, special nature…the essential quality and vitality of a being or entity." This suggests a deep connection or even equivalence between mauri and the self. Once a person or other living thing dies, their mauri is lost.

For Mead (2016, 338), the "mauri test is essentially a test of the risks to the life of the subjects." Any intervention, technological or otherwise, must consider whether the mauri of an object or thing will be enhanced or damaged.

As an illustration, Mead (338) discusses the case of a heart transplant from a pig. This is an intervention that typically helps a person in sustaining or even save their life, contributing to their mauri. However, this requires the sacrifice of the pig, which must be weighed in the balance. In addition, the heart is still living, meaning that some mauri is still retained within it. The high stakes of life and death add another layer of difficulty. Such issues complicate the discussion: there are no clear-cut answers.

How might AI technologies retain the mauri of a person, place, or thing? Or put negatively, how might AI models refrain from compromising or corrupting the innate life force of people and things? Here we are fundamentally talking about protecting communities and sustaining environments. So while this goal of upholding mauri may sound vague from a "rational" Western perspective, with its focus on metrics and quantitative measurements, there are some pragmatic ways of evaluating technologies according to this criteria.

One relevant tool is the Mauri Model, a framework originally developed for water quality assessment (Morgan and Brian 2006). The Mauri Model consists of four interrelated spheres of life which become progressively more expansive. The innermost ring is whānau, loosely correlating with family. The second ring is a community, gesturing to societal impact. The third level is hapū, a subtribe or basic political unit within a Māori mode of governance. And the fourth and largest sphere is the ecosystem, indicating the broader ecologies of air, water, and earth. Each of these spheres is weighted according to what the community and participants decide. In the case considered by Morgan and Brian (2006), the weightings were 40% ecosystem, 30% hapū, 20% community, and 10% whānau. The intervention in question (e.g. a dam, a road, a platform, and so on), is then rated according to its projected impact on mauri within each sphere. Interventions expected to "destroy" mauri score −2 points, while on the other end of the scale, those that "enhance" it score +2 points.

When considering AI's impact on mauri from an ecological perspective, an end-to-end approach is needed. The carbon footprint of AI technology is not just the daily use of the final product, but must include the computation needed for training and inference (Wu et al. 2022). In recent years, the processing needed to carry out many generations of training has surged significantly. Sevilla et al (2022) describe the last few years as the third era of machine learning, with large-scale models demanding a 10 to 100-fold increase in computing power. Google's recently released Palm model, to take just one example, has 540 billion parameters. Such hefty computation requirements threaten to concentrate AI power in the hands of a few major tech companies. But the major point here is that such computation carries an enormous environmental fallout, consuming water and electricity and producing carbon emissions (Hogan 2018). There has been increased attention to this ecological impact in recent years, leading to calls for sustainable AI (van Wynsberghe 2021).

The concept of mauri echoes this call, while stressing the dense connections between care for the earth, care for community, and care for family. Translated into a Western context, it brings together aspects of individual well-being, social support, good governance, and environmental sustainability. Mauri, and the broader Māori world view, recognise in fact that these aspects are often deeply interrelated. Caring for a particular community, for example, means caring for the forest or lake that sustains their lives and livelihoods. In showing the tight connections between ecological and social spheres, this indigenous knowledge system anticipates later concepts like environmental racism and environmental justice (Lazarus 2000; Holifield 2001; Cole and Foster 2001). Mauri is powerful in highlighting a holistic understanding of care for life.

3.3 Test 3: Take-utu-ea

Take-utu-ea refers to an issue that requires resolution. Once an issue or conflict has been identified, the utu refers to a mutually agreed upon cost or action that must be undertaken to resolve the issue. For Mead (2016, 27), an incorrect action is considered to be a breach (a *take*) which then requires some kind of responding action (*utu*) to reach a resolution (*ea*). Take-utu-ea is fundamentally about restoring balance, about making things right through an exchange of some kind. Lévi-Strauss (1996) describes this as the principle of reciprocity.

In the context of the 5 tests, this Take-utu-ea or TUE test is activated based on the results of Tests 1 and 2. In other words, if a breach of tapu or mauri is suspected, then the TUE test is applied. Mead (2016, 342) gives the example of an experimental drug test. The drug had known side effects that the pharmaceutical company failed to warn patients about (*take*); the company takes responsibility and agrees to financially compensate patients (*utu*); a state of satisfaction is then reached (*ea*). However, such a clear-cut example may increasingly be difficult to find. Given the complexity of our technical and political systems, with their chains of events and layers of decision-making, fixing responsibility on any single actor is challenging and often highly contested (Falconer 2002). Deciding what constitutes a transgression and who is to blame is often a fraught exercise.

This problem seems particularly pervasive in our contemporary informational systems, whether framed as AI, automated, or algorithmic. The decision-making process of these systems is often opaque; which factors are considered and how exactly they impact the outcome are obscured within a black box (Pasquale 2015). Because of this, a number of

scholars (Pasquale 2011; Diakopoulos 2016; O'Neil 2017) warned the public early on about the lack of transparency in algorithmic decision-making and called for additional oversight. In the subsequent years, AI technologies have only gained in reach and several prominent examples of bias and discrimination have emerged. In response, some researchers have developed tools to technically audit AI, testing the system and producing reports to reveal risks and threats (Raji et al. 2020). Others have developed methods to support explainability and add social transparency to AI systems (Ehsan et al. 2021).

If transparency aids in revealing the problems with AI systems, it must be accompanied by accountability. Once a breach or act of bias (*take)* is shown, there must be ways to enforce some kind of response (*utu*) to remedy it. Users cannot expect the industry to voluntarily regulate themselves and their technologies. Indeed, regulation is often regarded in the tech sector as being something that stifles "innovation" (Lev-Aretz and Strandburg 2020). Fuzzy claims of human principles and best practices, applied when AI companies find them desirable, has proven to be wholly inadequate (Munn 2022a). There must be a shift from soft regulation to hard law (Floridi 2021). To pursue this goal, scholars have focused on constructing the underlying frameworks (technical, legal, institutional frameworks) for establishing accountability and liability in AI (Smith 2021). These moves seek to produce regulation "with teeth," to couple transparency with accountability. Such laws would force companies to accept responsibility when ethical breaches and issues of discrimination are discovered in their products. Making restitution in this instance may mean engaging more deeply with a community, adding or removing training data, rewriting pieces of code, or even paying compensation to the parties that have been adversely impacted. These insights suggest that applying the Take-utu-ea test to AI tools requires both a serious commitment and an appropriate set of technical and legislative tools.

3.4 Test 4: Precedent

The fourth test is precedent. The aim here is to find examples from the past that might help establish a correct viewpoint and guide actions in the present. "Is there some event in our traditions that might help us understand the issue and help frame a response to it?" asks Mead (2016, 343). An event or issue may seem entirely novel, introducing new technologies, new capabilities, or new controversies. And yet this issue does not emerge from a vacuum, but is instead cumulative, building on historical knowledge, established institutions, and prior techniques. For this reason, Mead suggests looking at indigenous stories, older traditions, and historical examples as a way to develop an appropriate response.

One avenue for guidance is examining pūrākau, a particular form of traditional Māori narrative. In a modern or western context, these are often denigrated as myths, folk tales that are both irrelevant and unscientific. However, these stories distill diverse forms of knowledge (spiritual, empirical, moral) into a memorable and understandable package. Lee (2009) contends these stories contain philosophical thought, epistemological constructs, cultural codes, and world views. Ruth Irwin and Te Haumoana White (2019) note that mythical terms "bridge Māori and contemporary thought." These accounts ground an indigenous cosmology and philosophy in powerful narratives—yet also provide flexibility for this knowledge to be adapted and re-applied to new challenges such as artificial intelligence.

One pūrākau states that there were three baskets of knowledge in the heavens that contain all of humanity's knowledge. Tāne was sent to retrieve these baskets (*kete*), battling his older brother Whiro and overcoming obstacles to ascend through layers of heaven and retrieve the prized possessions. The *kete-aronui* contained knowledge that could help humans; the *kete-tuauri* housed the knowledge of ritual, memory and prayer; and the *kete-tuatea* held knowledge of evil which was harmful to humans. Karaitiana Taiuru (2018) argues that data is today's knowledge basket, a container housing a rich treasure of information regarding all of life. As in the narrative, this information is powerful, granting those who possess its particular advantages. Data is a resource, a treasure for the twenty-first century, but like other resources throughout history, it is one that is often dominated, controlled, or co-opted by colonizers. So, just like the story, this data should not be left to others but should be grasped or at least contested.

As AI applications, algorithmic systems, and automated decision-making encroach on everyday life, indigenous organizations have increasingly recognized the importance of such data. Several years ago, Te Hiku Media, a small Māori non-profit, began assembling a set of annotated audio recordings, a foundational set of data for doing automated language recognition. But all too quickly, an American technology company contacted the organization with an offer to purchase the data. Te Hiku rejected the offer and published a statement explaining that it wanted to retain indigenous knowledge and help revitalize the Māori language (Lucas-Jones 2018). Efforts to maintain control and ownership over data can also be found in the Māori Data Sovereignty Network (Raraunga 2022). This collection of practitioners and scholars conducts research about new technologies, collaborates with government and university partners, and provides guidance on policy initiatives. They recognize that maintaining control over these knowledge-baskets—what information they contain or neglect, where they can circulate, and who has authority over them—is key for autonomy in the times ahead. As their slogan suggests: Our Data, Our

Sovereignty, Our Future. Such work provides a strong example of a famous Māori proverb. *Kia whakatōmuri te haere whakamua:* I walk backwards into the future with my eyes fixed on my past.

3.5 Test 5: Principles

The final test encompasses a range of additional principles that may be drawn upon to evaluate an issue when needed. Mead (2016, 344) is pragmatic here, admitting that in some cases "the first four tests may not be helpful at all and so one may have to consider the principles test." This Principles test thus acts as a fallback for those attempting to develop an indigenous perspective. In some respects it is a catch-all category, containing a number of supplementary creeds and values which may provide novel insights or additional guidance. This section briefly steps through each principle and shows how it might be usefully applied to design or evaluate AI technology.

Test 5.1 is *whanaungatanga*. If a person is a relative, then kin are obliged to assist them or support them as needed. This principle can also be extended to non-kin: classmates, workmates, or the larger iwi that one has membership in. One question this test might ask: does this app or platform support indigenous people in connecting with their extended family, colleagues, or peers? In doing so, it may facilitate a form of *whanaungatanga*, passing this particular test.

Test 5.2 is *manaakitanga*. The concept here is to rise above personal grievances and politics, acknowledging the mana of others and showing care and hospitality to them. This test might question the purpose of AI-driven technology. Is it divisive, fostering polarization and forms of antagonism along racial, cultural, or gendered lines? Or is it inclusive, respecting and potentially even uniting diverse peoples together? The latter case embodies *manaakitanga*, passing that particular test.

Test 5.3 is *mana*. A new event or technology should not damage the mana of the subject or user, nor damage those who are involved with it. In contrast, it should aid in maintaining or even improving this mana. This test might question the impact that a particular technology might have on its participants or end-users. Is it purely driven by short-term interests and business values, undermining their social and mental well-being? Or does it seek to empower individuals, to build their *mana* and help them to cultivate and actualize their potential?

Test 5.4 is *noa*. A novel paradigm or condition may need to be introduced in an appropriate way to those who use it so that it can become commonplace and no longer controversial. This test is fundamentally about reducing the doubt or shock concerning a novel intervention and making it more everyday, more *noa*. In the context of AI, this might be an education campaign that listens to the concerns that a community has and sensitively addresses them.

Test 5.5 is *tika*. When these more targeted principles fail to yield sufficient guidance, then an evaluation may need to fall back to tika: considering "whether something is ethically, culturally, spiritually, and medically right" (Mead 2016, 347). This test opens the framework up to a variety of broader evaluations about the correctness (*tika*) of a technology. This provides a degree of flexibility to the evaluation. An AI developer may want to rely predominantly on indigenous criteria, for instance, yet also splice in some western ethical norms to round off the evaluation.

These additional principles and concepts can be used whenever the first four tests, for whatever reason, fail to yield the kind of insights or guidance that are desired about a particular AI technology.

4 Conclusion: designing and decolonizing

As the problems with current AI paradigms become increasingly clear, the need for alternative visions, values, and frameworks grows urgent. Indigenous concepts from Aotearoa provide a distinctly different set of principles and priorities—a way of knowing and being at once ancient and fresh that productively challenges Western technocratic norms. Importantly, these are not simply high-minded ideals but pragmatic principles that can be applied to novel circumstances and new technical conditions. The Five Tests provide a means for carrying out this evaluation, allowing us to weigh up the potentials and problems of AI according to a rubric that centers on human dignity, communal integrity, and ecological sustainability.

I see two distinct pathways that the work here opens up. The first can be labeled *designing*, a pragmatic application in the present. What would it look like to take these principles and apply them to an AI product that is currently in development? Every day, new technologies are released in the world, rolling out into high-stakes areas such as healthcare, welfare, immigration, human resources, and finance. And so, even though this research is nascent, I want to offer something practical if modest: a starting point for designing or redesigning AI differently.

Designing focuses on the form and function of technology, its appearance and its operations. This is not just the user interface, but more fundamentally the set of decisions that lead to this particular product, this particular technology. In the context of AI, this enfolds aspects like the production of training data, the definition of "ground truth," the architectures used (i.e. transformer vs convolutional), the optimization function, the deployment of the model, and the feedback mechanisms it surfaces to users. These aspects, from classification to explainability and outputs,

can be understood as the design space of a model (Morris et al. 2022a, b). The Five Tests can be applied to these decisions-in-progress, guiding the development of an AI-based intervention. How might this design respect the sacred (tapu)? How might it preserve or enhance the life force of the people and environments it impacts (mauri)? If this design has negative impacts, how might those be reconciled in an acceptable way (take-utu-ea)? These kinds of questions can be asked again and again throughout the design process, from conception through to development and launch. They provide a very different set of criteria from typical business maxims. Engaging genuinely with these questions and "resolving" them through code, architectures, interfaces, and affordances would result in a distinctive technology, something more considered, more inclusive, and more ecologically attuned.

Design also implies redesign, the ability to assess how something is performing and alter particular aspects to improve it. This would apply to AI products and services that have already been developed and deployed. How exactly might these tests be taken up by companies or organizations in this case? This process might begin with a workshop that provided a foundational understanding of indigenous values. This workshop would also enable different stakeholders to come together around the evaluation process. This kind of cross-organizational collaboration is crucial considering the diverse forms of data and feedback that are required. Holistic evaluation could certainly include an engineering team and their quantitative metrics—but might also encompass an environmental report and a series of interviews with community leaders. For some of the tests, evaluative instruments would need to be designed; for others, like Mauri, existing measures might be used. Each of the Five Tests could be weighted according to their importance for the AI developer and for the end-users involved. Scores for each test would highlight areas that need improvement and further iteration.

Critics might object that these measures are unsatisfactory—somewhat improvised and incomplete—and they would be correct. But this is what product development and policy development look like in practice, a process of attempts and iterations. By muddling through (Lindblom 2018) and comparing against prior versions, a material artifact such as code, infrastructures, platforms, or legislation directed at them is gradually improved. Version 0.2 is slightly more diverse, or equitable, or sustainable than Version 0.1. Perfect is the enemy of good.

The more serious criticism in this context would actually be co-option. Indigenous principles are not a smorgasbord where principles can be chosen as desired. In the same vein, splicing one or two concepts into a broader framework of Western values too often leaves them watered down or tokenistic. It would be easy for governments and corporations to gain social and cultural prestige by superficially parroting some of these values without any significant commitment behind them—indeed, we see such a pattern repeatedly in the past. For this reason, the pragmatic, present-focused response discussed above needs to be accompanied by a slower reflection on the nature of power and one's role within it.

The second pathway might thus be termed *decolonizing*, a deeper and more sustained confrontation with current AI regimes. Band-aid fixes and nods to "diversity" or more "humane" AI have proven to be woefully inadequate (Munn 2022a). The failure of such superficial interventions only serve to demonstrate that current technologies are built on a far deeper strata of capitalism and coloniality, perpetuating systems of inequality and oppression. For all their purported rationality and neutrality, technologies often reinforce this violence, causing collateral damage to the marginal and the vulnerable who are least equipped to deal with it (McQuillan 2022).

The exact nature of the Five Test's provocation to current AI regimes is a matter of interpretation, but a few challenges might briefly be sketched. First, it rejects the generic, universalizing frame often imposed by technology companies and even by AI ethicists. Instead, it is grounded in particular worldviews, it highlights communal needs, and it tends to frame impacts according to local norms. Second, it refuses to neatly compartmentalize human well-being and ecological well-being. Instead, it stresses their connection and interdependence moving into the future: social and environmental justice are inseparable (Rixecker and Tipene-Matua 2003). And third, it does not shirk responsibility for its activities, bracketing off downstream effects as a problem for the state or individuals, as is typical of many technology companies. Instead, it carefully considers potential impacts and develops ways to mitigate them or redress them to satisfy involved parties. Even in this brief sketch, we can recognize a way of intervening that is slower, more considered, and more considerate of life in its various forms—the antithesis of our current AI regimes and their production culture of "moving fast and breaking stuff."

The Five Tests, then, are not just about instrumentalization but about confrontation: they pose a more fundamental challenge to current AI paradigms and practices. They raise a series of key questions: what should this data-driven technology be doing? How might we design these technologies in ways that are more inclusive, communal, and sustainable? And what values and norms are we using to judge the success of a particular technology? These are epistemological questions, concerning the knowledge systems that we use to understand the world. They are cultural and historical questions, concerning the violent domination of some peoples by others. And they are social questions, concerning a particular understanding of a healthy society and the good life. These questions resonate with recent calls to decolonize

AI (Mohamed et al 2020; Hanna 2022). Understanding and undoing systems of inequality that have been formalized and fossilized over time is a massive undertaking. Such a programme is daunting when one considers the full range of institutions, relationships, and practices that would need to be rethought (Adams 2021). And yet—as digital technologies and technical systems become increasingly rolled out in high-stakes areas and permeate into our life world— this seems to be the long-term project that social justice demands.

These two pathways are complementary, each augmenting the other. Technical design and codesign initiatives from developers and end-users must be coupled with a radical rethinking of dominant philosophies and paradigms: the bottom-up must be combined with the top-down (Cruz 2021). A genuine engagement with this work might see a company initiate a rapid redesign and evaluation, something material, messy, and ad-hoc that nevertheless generates insights and iterative versions of software—but it would also see that company put in place long-term measures to reflect on their priorities as an organization, to understand the power asymmetries at work in the world, and to restructure their ways of being-and-doing as this journey unfolds. The radical and the actionable must be intertwined.

Data availability There is no dataset associated with this article.

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

Abdilla A, Kelleher M, Shaw R and Yunkaporta T (2021) Out of the black box: indigenous protocols for AI. UNESCO, Paris

Adams R (2021) Can artificial intelligence be decolonized? Interdisc Sci Rev 46(1–2):176–197

Beller J (2018) The message is murder: substrates of computational Capital. Pluto Press, London

Benjamin R (2019) Race after technology: abolitionist tools for the New Jim Code. Polity, London

Berardi F (2009) The soul at work. MIT Press, Cambridge

Brynjolfsson E, McAfee A (2011) Race against the machine: how the digital revolution is accelerating innovation, driving productivity, and irreversibly transforming employment and the economy. Digital Frontier Press, Boston

Brynjolfsson E, McAfee A (2014) The second machine age: work, progress, and prosperity in a time of brilliant technologies. WW Norton & Company, New York

Buolamwini J and Gebru T (2018) Gender shades: intersectional accuracy disparities in commercial gender classification. In: Conference on fairness, accountability and transparency, pp 77–91. ACM, New York

Checketts L (2022) Artificial intelligence and the marginalization of the poor. J Moral Theol 11(1):87–111

Ciston S (2019) Intersectional AI is essential: polyvocal, multimodal, experimental methods to save artificial intelligence. J Sci Technol Arts 11(2):3–8. https://doi.org/10.7559/citarj.v11i2.665

Cole LW, Foster SR (2001) From the ground up: environmental racism and the rise of the environmental justice movement. NYU Press, New York

Cruz CC (2021) Decolonizing philosophy of technology: learning from bottom-up and top-down approaches to decolonial technical design. Philos Technol 34(4):1847–81. https://doi.org/10.1007/s13347-021-00489-w

Diakopoulos N (2016) Accountability in algorithmic decision making. Commun ACM 59(2):56–62

Dubal V (2020) Digital piecework. Dissent Magazine, Fall 2020. https://www.dissentmagazine.org/article/digital-piecework

Ehsan U, Liao QV, Muller M, Riedl MO and Weisz JD (2021) Expanding explainability: towards social transparency in AI systems. In: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, 1–19. New York: ACM

Falconer J (2002) Accountability in a complex world. Emergence 4(4):25–38

Floridi L (2021) The end of an era: from self-regulation to hard law for the digital industry. Philosophy & Technology 34(4):619–622

Guarino N (1998) Formal ontology and information systems. Proc FOIS 98:81–97

Hanna A (2022) Beyond bias: algorithmic unfairness, infrastructure, and genealogies of data. In: Presented at the AI for good conference, Zurich, April 13. https://www.youtube.com/watch?v=j8hn-EMQFhc

Hogan M (2018) Big data ecologies. Ephemera 18(3):631–657

Holifield R (2001) Defining environmental justice and environmental racism. Urban Geogr 22(1):78–90

Huws U (2014) Labor in the global digital economy: the cybertariat comes of age. NYU Press, New York

Irwin R, White TH (2019) Decolonising technological futures: a dialogical tryptich between Te Haumoana White, Ruth Irwin,

and Tegmark's Artificial Intelligence. Futures 112(September):102431. https://doi.org/10.1016/j.futures.2019.06.003

Jiménez A (2020) The silicon doctrine. TripleC 18(1):322–236. https://doi.org/10.31269/triplec.v18i1.1147

Lazarus RJ (2000) Environmental racism—that's what it is. Univ Ill Law Rev 2000:255

Lee J (2009) Decolonising Māori Narratives: Pūrākau as a Method. MAI Rev 2(3):1–12

Lev-Aretz Y, Strandburg KJ (2020) Regulation and innovation: approaching market failure from both sides. Yale J Regul Bull 38:1

Lévi-Strauss C (1996) The principle of reciprocity. In: Komter A (ed) The gift: an interdisciplinary perspective, pp 18–26. Amsterdam University Press, Amsterdam

Lewis JE, Abdilla A, Arista N, Baker K, Benesiinaabandan S, Brown M, Cheung M et al (2020) Indigenous protocol and artificial intelligence position paper. Monograph. Honolulu, HI: Indigenous Protocol and Artificial Intelligence Working Group and the Canadian Institute for Advanced Research. 2020. https://doi.org/10.11573/spectrum.library.concordia.ca.00986506

Lindblom C (2018) The science of 'muddling through'. In: Stein J (ed) Classic readings in urban planning, pp 31–40. Routledge, London

Lucas-Jones P (2018) Indigenous data theft. Te Hiku Media. https://tehiku.nz/p.CFg

Marx K (1977) Capital: a critique of political economy. Vintage, London (**Translated by Ben Fowkes**)

McQuillan D (2019) The political affinities of AI. In: Sudmann A (ed) The democratization of artificial intelligence. Transcript Verlag, Bielefeld, pp 163–173

McQuillan D (2022) Resisting AI: an anti-fascist approach to artificial intelligence. Policy Press, Bristol

Mead HM (2016) Tikanga Maori: living by Maori values. Huia Publishers, Wellington

Mejias UA, Couldry N (2019) The costs of connection: how data Is colonizing human life and appropriating it for capitalism. Stanford University Press, Stanford

Mitira TH (1990) Takitimu. Southern Reprints, Christchurch

Mohamed S, Png M-T, Isaac W (2020) Decolonial AI: decolonial theory as sociotechnical foresight in artificial intelligence. Philosophy Technol 33(4):659–684. https://doi.org/10.1007/s13347-020-00405-8

Morgan TKKB (2006) Waiora and cultural identity: water quality assessment using the mauri model. AlterNative 3(1):42–67. https://doi.org/10.1177/117718010600300103

Morris MR, Cai CJ, Holbrook J, Kulkarni C and Terry M (2022a) The design space of generative models. In: Virtual workshop on human-centered AI, 6. Stanford: NeurIPS Foundation. https://cs.stanford.edu/~merrie/papers/DesignSpaceOfModels.pdf.

Munn L (2017) I am a driver-partner. Work Organisation, Labour & Globalisation 11(2):7–20. https://doi.org/10.13169/workorgalaboglob.11.2.0007

Munn L (2022a) The uselessness of AI ethics. AI and Ethics. https://doi.org/10.1007/s43681-022-00209-w

Munn L (2022b) Data and the new oil: cloud computing's lubrication of the petrotechnical. J Environ Media 2(2):211–227. https://doi.org/10.1386/jem_00063_1

Noble S (2018) Algorithms of oppression how search engines reinforce Racism. New York University Press, New York

O'Neil C (2017) Weapons of math destruction: how big data increases inequality and threatens democracy. Penguin Books, London

Pasquale F (2011) Restoring transparency to automated authority. J Telecommun High Technol Law 9(February):235–254

Pasquale F (2015) The black box society: the secret Algorithms that control money and information. Harvard University Press, Cambridge, MA

Raji ID, Smart A, White RN, Mitchell M, Gebru T, Hutchinson B, Smith-Loud J, Theron D and Barnes P (2020) Closing the AI accountability gap: defining an end-to-end framework for internal algorithmic auditing. In: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, 33–44. FAT* '20. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3351095.3372873

Raraunga TM (2022) "Te Mana Raraunga". Te Mana Raraunga. https://www.temanararaunga.maori.nz

Rixecker SS and Tipene-Matua B (2003) Maori kaupapa and the inseparability of social and environmental justice: an analysis of bioprospecting and a People's resistance to (Bio) cultural assimilation. In: Bullard R, Agyeman J, Evans B (eds) Just sustainabilities: development in an unequal world, pp 252–268. London, Routledge

Sevilla J, Heim L, Ho A, Besiroglu T, Hobbhahn M, and Villalobos P (2022) Compute trends across three eras of machine learning. arXiv. https://doi.org/10.48550/arXiv.2202.05924

Shedlock K, Hudson P (2022) Kaupapa Māori concept modelling for the creation of Māori IT artefacts. J Roy Soc New Zealand. https://doi.org/10.1080/03036758.2022.2070223

Smith H (2021) Clinical AI: opacity, accountability, responsibility and liability. AI & Soc 36(2):535–545

Taiuru K (2018) Why data is a taonga: a customary Māori perspective. Christchurch: Independent. https://www.taiuru.maori.nz/wp-content/uploads/Data-is-a-taonga.pdf

Taiuru K (2020) "Māori cultural considerations with facial recognition technology in New Zealand." Dr Karaitiana Taiuru PhD, JP. December 8, 2020. https://www.taiuru.maori.nz/maori-cultural-considerations-with-facial-recognition-technology-in-new-zealand/

van Wynsberghe A (2021) Sustainable AI: AI for sustainability and the sustainability of AI. AI and Ethics 1(3):213–218

Williams DH, Shipley GP (2020) Enhancing artificial intelligence with indigenous wisdom. Open J Philos 11(1):43–58. https://doi.org/10.4236/ojpp.2021.111005

Wu C-J, Raghavendra R, Gupta U, Acun B, Ardalani N, Maeng K, Chang G, Aga F, Huang J, Bai C (2022) Sustainable AI: environmental implications, challenges and opportunities. Proc Mach Learn Syst 4:795–813