Increasing the presence of BIPOC researchers in computational science

Christine Yifeng Chen[1], Alan Christoffels[2,3], Roger Dube[4], Kamuela Enos[5], Juan E. Gilbert[6], Sanmi Koyeji[7], Jason Leigh[8], Carlo Liquido[9], Amy McKee[10,11], Kari Noe[5], Tai-Quan Peng[12], Karaitiana Taiuru[13,14]

[1] Cosmochemical and Isotopic Signatures Group, Nuclear and Chemical Sciences Division. Lawrence Livermore National Laboratory. Livermore, CA 94550, USA
[2] South African MRC (SAMRC) Bioinformatics Unit, South African National Bioinformatics Institute University of the Western Cape, South Africa
[3] Africa Centers for Disease Control and Prevention, African Union, Addis Ababa, Ethiopia
[4] College of Science. Rochester Institute of Technology. Rochester, NY 14623, USA
[5] Office of Indigenous Knowledge and Innovation. University of Hawaiʻi, Honolulu, HI 96822, USA
[6] Computer & Information Science & Engineering Department, University of Florida, Gainesville, FL 32611, USA
[7] Department of Computer Science, Stanford University, Stanford, CA 94305, USA
[8] Department of Information and Computer Sciences, University of Hawaiʻi at Mānoa, Honolulu, HI 96822, USA
[9] Piʻikū Co. Kailua, HI 96734, USA
[10] Department of Educational Foundations, University of Hawaiʻi at Mānoa, Honolulu, HI 96822, USA
[11] ʻĀinaQuest, Honolulu, HI 96822, USA
[12] Department of Communication, Michigan State University, East Lansing, MI 48824, USA
[13] Taiuru & Associates Limited. Auckland, New Zealand
[14] Ngāi Tahu, Ngāti Kahungunu, Ngāti Toa, Auckland, New Zealand

Corresponding authors: chen127@llnl.gov, alan@sanbi.ac.za, rrdube@gmail.com, kenos@hawaii.edu, juan@ufl.edu, sanmi@cs.stanford.edu, leighj@hawaii.edu, carlo.h.liquido@gmail.com, amymckee@hawaii.edu, karinoe@hawaii.edu, pengtaiq@msu.edu, karaitiana@taiuru.maori.nz

Standfirst: *Nature Computational Science* asked a group of scientists to discuss strategies for increasing the presence of Black, Indigenous, People of Color (BIPOC) researchers in computational science, as well as the various considerations to be made for improving education and methods design.

Christine Yifeng Chen- Funding disparities

“It’s an open secret that NSF funds Asians less.” I heard these words on a Zoom call in late 2020, at an informal gathering of Asian American and Pacific Islander researchers in my discipline. We were meeting to support each other amidst the surge of anti-Asian sentiments during the pandemic, sharing stories and exchanging advice. But this revelation about the U.S.National Science Foundation (NSF) was news to me. I was a postdoc at the time, thinking about my future research and future *in* research, so naturally, I was curious and concerned.

Seeking details, my colleague Sara Kahanamoku and I followed up with the professor who made the comment. They pointed us to a little-known corner of NSF’s website that contained nearly 25 years of annual reports on their merit review process and data on proposals and awards. A first look confirmed our fears: for every year between 1999 and 2019, white applicants received funding at above-average rates, while most other racial groups, particularly Asians, were funded at below-average rates.

But that wasn't the whole story. Digging deeper, the funding disparities spanned all scientific fields and were more pronounced for research proposals compared to those for education and training. For instance, in the Directorate for Computer and Information Science and Engineering in 2012–2016, research proposals by white applicants were funded at rates 1.2 and 1.5 times higher than those by Asian and Black applicants. These results offer several insights, including that broadening participation does not automatically lead to equitable outcomes. Asians still face lower funding rates even in disciplines where Asians are well-represented.

While the group with the lowest funding rate varied by context, one pattern remained consistent: white applicants received higher funding rates across all proposal types and disciplines. Because these disparities have persisted year after year, their cumulative impact at NSF amounts to billions of dollars in unbalanced funding. Similar funding advantages for white applicants have also been reported at the U.S. National Institutes of Health (NIH), NASA, philanthropies, and other research funders in the U.K., highlighting a systemic issue underpinning our entire research ecosystem.

With the help of tenured faculty and others, we published our findings in late 2022 [1]. Our publication sparked intense reactions. Of the productive responses, many asked: what's causing this, and how do we fix it? There are no easy answers, but here are some observations. In 2011, the NIH reported that white applicants received funding at a rate 1.7–1.8 times higher than Black applicants in 2000–2006 [2]. Despite the NIH's efforts to address these gaps by amassing data on possible causal mechanisms within their merit review process, the funding advantage for white applicants remained unchanged at 1.7–1.8 times higher for over a decade [3].

The unfortunate reality of these trends is that there is no "one simple fix." The hyperfocus on reducing reviewer bias with double-blind or other tweaks to the review process is a textbook case of costly distraction. The idea that only active animus or unconscious bias can produce racially unequal outcomes in organizations is deceptive. Even if we could magically eliminate all racism and bias in society, these racial funding disparities would persist. Such is the nature of structural racism in contemporary contexts, where policies, processes, and social norms that appear race neutral nevertheless produce and perpetuate racial inequalities.

In many ways, racial disparities in research funding parallel racial wealth gaps in broader American society. Both gaps are widening, contrary to narratives of racial progress [4]. At NSF, the funding advantage for white applicants *increased* from 1999 to 2019. The patterns are self-reinforcing, a "rich-get-richer" effect where past funding success leads to future success [5]. Advantages start early, through prestigious graduate research fellowships [6] and postdoc-to-faculty grants, and compound over time, with downstream impacts on researchers' productivity, recognition, and influence. Under these conditions, it is unsurprising that decades of efforts to diversify university faculty have failed.

How do we break this cycle? Simply put, to eliminate inequalities, we must address the causes of inequality in the first place: unequal access to social prestige, insider knowledge, and most importantly, organizational resources.

Follow the money.

Alan Christoffels- Amplifying computational science research in Africa

Bioinformatics and Computational Biology were formally initiated in Africa with the launch in 1997 of the South African National Bioinformatics Institute (SANBI) at the University of the Western Cape. The following 10 years spawned a range of entities driving bioinformatics activities in Africa, initially focused on human genomics. Sustained funding from the South African Medical Research Council (SAMRC) established the SA bioinformatics unit and the South African Research Chairs Initiative in Bioinformatics and human health. These activities were expanded to genomics of vector-borne diseases through a 10-year investment in the Tropical Diseases Research program, and the human heredity in health (H3Africa).

These regional investments initiated and/or expanded a network of genomics and bioinformatics researchers in Africa that was formalized as the African Society for Bioinformatics and Computational Biology (ASBCB) in 2004. Over a period of 30 years, bioinformatics in Africa has infused the biological science discipline, expanding from algorithmic development in the 1990s in health sciences [7] to a range of other biological sciences from plant genomes to computer aided drug discovery and many others.

As a society, we strive to support this diverse ecosystem and to contribute to shaping computational skills development and fostering collaborative research opportunities both regionally and abroad.

Computational skills development

Working groups within ASBCB affords members an opportunity for peer training within focused community of special interest groups (COSI) including pathogen genomics, meta-omics, agriculture bioinformatics, systems administration and structural biology. These forums allow researchers and students to coalesce around common interests, exchange ideas, and facilitate peer training. These joint initiatives have spawned Codeathons that provide a targeted approach for students to apply their collective programming skills to understand the biology of model organisms. For example using existing bioinformatics tools and writing code to identify SNPs in drug metabolizing enzymes.

Networking opportunity

ASBCB also offers networking opportunities to identify areas of shared interest and collaborations that seek to investigate locally relevant diseases. These linkages translate into joint funding applications that support the growth of the research discipline. Success stories include the Eastern Africa Network for Bioinformatics Training (EANBiT), a collaborative network in Kenya, and Uganda supported by the Fogarty International Center of the National Institutes of Health and similar networks in West Africa.

A regional affiliate of the International Society for Computational Biology (ISCB), the ASBCB has co-hosted biennial ASBCB-ISCB biennial conferences since 2009. These events provide the space to promote locally driven computational science research, amplify the work of young African scientists, and provide a platform for scientists from diverse geographical locations to exchange ideas. These events have translated into tangible student exchange opportunities for exposure to interdisciplinary research environments.

Responding to a dearth of computing infrastructure technical support

The systems administration COSI has endeavored to strengthen the technical support skills needed to meet the growing demands of a data intensive research environment. Over a seven-year period, the Research Software and Systems Engineers (RSSEs) in Africa Forum was launched. This is a discussion and networking forum for professionals involved in building software and computing systems to support research on the African continent.

Equitable access to, and development of, machine learning datasets

While not formally convened as a COSI, there is growing interest among African researchers to use machine learning techniques for applications in biology. The work done in natural language processing provides the impetus for ASBCB to advocate for machine learning datasets extended to biology. For example, despite the advances in natural language procession (NLP), there remains a gap in the development of applications for African languages [8]. Access to, or development of, applications for African languages create opportunities for artificial intelligence (AI) application development in a wide range of domains including healthcare. These efforts have translated into a collection of open-source training datasets, code and results to encourage growth of NLP research in African languages. Similarly, the individual efforts in Africa using AI for biological applications could be strengthened by promoting these efforts via the ASBCB.

Roger Dube- Ongoing initiatives of AISES

The fields of computational science present exciting possibilities for "Indigenous Americans" (that is, American Indian, Alaska Native, Native Hawaiian, Pacific Islander, First Nation). Before North American Indigenous students can benefit from these intellectual and employment opportunities, there are unique challenges and disadvantages some must overcome to improve the likelihood of their success.

These challenges include 1) overcoming inadequate preparation caused by a lack of educational and technological resources (for instance high-speed internet, computers, advanced STEM courses and laboratory equipment) and 2) addressing the cultural and social barriers that lead students consider themselves incapable of a career in science or mathematics. This limiting self-image may come from a lack of role models, knowledge of recent archeological evidence of indigenous scientific and engineering achievements, culturally responsive educational methods, and awareness of opportunities. There are also challenges related to eliminating financial constraints, countering stereotypes and discrimination, which lead to feelings of isolation that impact academic performance and mental health, and the pressure of fulfilling family and community responsibilities, preventing students from attending far-from-home universities.

I've participated in three effective approaches that can address these challenges.

1. Organizations such as AISES (American Indian Science and Engineering Society). This national, nonprofit organization focuses on substantially increasing the representation of "Indigenous Americans" in STEM studies and careers. One approach AISES uses to change the perception that Indigenous Peoples have of potential and achievements in STEM fields is through student poster sessions and competitions, conducted at regional and national AISES meetings. There, awards are presented to reward notable Indigenous accomplishments.

2. Culturally sensitive programs developed within educational institutions, such as University of Manitoba's Wawatay, that supports Indigenous students in their pursuit of STEM degrees, thereby increasing the enrollment and retention of students from these backgrounds, based on a personal communication with Dr. Carrie Selin, Academic Program Lead at Wawatay. AISES is also available to guide schools in creating this type of program.

3. Government-sponsored programs funding university experiences, such as the National Science Foundation's Research Experience for Undergraduates (REU). These programs occur during the summer, providing additional mechanisms reward Indigenous students and strengthen their self-image as STEM students, and eventually professionals.

Although not widely known or discussed, Indigenous Americans have a long history of substantial technological accomplishments, some of which were invented and in use well before corresponding inventions appeared in the rest of the world. Archeology has revealed some of the early inventions in the Americas such as gold plating without the use of batteries , thirty-angstrom scalpels, which promote quick healing of incisions, a compass [9].

Unfortunately, during first contact, Europeans did not recognize the effectiveness and significance of Indigenous inventions and engineering solutions like those listed above. I believe this occurred because of the fundamental difference between how Europeans and Indigenous Americans in acquired and subsequently shared knowledge. Typically, Europeans spread ideas and accomplishments through printed materials, while Indigenous cultures imbedded science in their languages and taught skills and a world view through apprenticeship training. Thus, Europeans dismissed the scientific and engineering capabilities that were already developed and used in the Americas. This gap in the treatment and dissemination of knowledge persists today, which may contribute to the dismissal of the validity of Indigenous discovery and scientific practices.

As more students earn their degrees and assume faculty positions in STEM fields this perspective will hopefully fade. Furthermore, as the number of Indigenous students, professors and professionals in STEM fields grow, the process of new Indigenous STEM faculty attracting more Indigenous students into STEM will become self-sustaining. Eventually there will be a resurgence of the Indigenous "traditional knowledge" and approach that expands and enhances western scientific methods with a more observational and holistic approach to understanding complex systems. In the meantime, initiatives like those discussed here can help to close the gap.

Juan E. Gilbert- Increasing the presence of Black researchers in computational science

According to the most recent Computing Research Association (CRA) Taulbee Survey, there are 236 Black/African-American computer science (CS) PhD students enrolled in PhD programs in the U.S.A., which represents 1.6% of all CS PhD students. The University of Florida's CISE Department has 17 Black CS PhD students. So, roughly 7.2% of the nation's Black CS PhD students are in the CISE Department at the University of Florida. Since 2006, I have produced 26 Black CS PhDs and currently my lab has 7 Black CS PhD students. This means that my lab has produced more Black CS PhDs than most universities in their entire existence. I have produced the most Black CS PhDs for any single faculty member in the country, to our knowledge.

Over the years, increasing the presence of Black researchers in computing science research has been a goal of mine and the data suggests that I have been successful in doing so. Therefore, what are some of the strategic initiatives or approaches I have implemented to achieve this success – I would point to two primary approaches. First, I would say that representation matters. If they see it, they can be it, is something I like to say. Meaning, if you want to reduce the chances of success of Black, or any group of, researchers in AI/computational science, isolate them. There are too many instances where there's only 1 Black student in a program, or there's only 1 Black faculty member in the program. Our PhD program has 17 Black students and we have 5 Black faculty members. The students clearly see themselves represented in the department at all levels. Now, in cases where there are no Black faculty, the students should be given access to meet Black faculty in other departments at the University and at other universities. For example, the NSF Broadening Participation in Computing (BPC) Alliance, Institute for African-American Mentoring in Computing Sciences (iAAMCS) can connect students to a national network of Black computing researchers to reduce, if not eliminate, isolation. iAAMCS has published "The IAAMCS Guidelines for Successfully Mentoring Black/African-American Computing Sciences PhD Students" that departments can use to learn how to best recruit, retain, and graduate Black students in AI/computational science. Second, AI/computational science can be very successful at recruiting Black researchers by enabling them to work on problems related to things they care about. You will find Black researchers in education, health, law, and other areas that can be defined as helping sciences, meaning these areas have clear connections to helping others or people in the society/community. AI/computational sciences are often seen as areas that work with phenomenon and artifacts and not people. If you can show how AI/computational science can be used to help people, you can recruit, retain, and graduate Black researchers. It is often the case, talented Black students pursue these helping disciplines, so if you can connect AI/computational science to problems that relate to their communities, they will come. We have been very successful in doing this by working on projects in voting technologies, law enforcement, and education, to name a few. We often use culturally-relevant approaches to teaching and research. We connect the life experiences of our Black students to research and education in the classroom. We use examples from their communities on problems in educational contexts to engage our students.

These strategies have been instrumental in our efforts to increase the presence of Black researchers in AI/computational science research and I am confident these will work for you too.

Sanmi Koyejo- Insights from Black in AI

I vividly remember the sunny afternoon in Autumn 2015 – I was excited to give my first seminar at a prestigious institution as a new postdoc. I arrived a bit too early and stood in front of the room when a professor who I knew (but did not know me) popped their head in to ask if I would be cleaning the room because there was a seminar starting in a few minutes. Imagine their surprise when they returned later to find that I was their seminar speaker! While I have experienced more or less distressing situations like this throughout my career, my colleagues and friends have experienced far worse. The Black representation gap in computational science is jarring and has remained as far back as I can find. According to the Computing Research Association Taulbee Survey, in 2023, just as in previous years, less than 2% of the PhD's in computing were awarded to individuals who identify as Black. Seeing a Black scientist should not be a surprise, and I believe we can work toward a time when Black computational scientists are not "rare creatures".

Over the years, efforts to increase participation in computing and to normalize the presence of Black scientists and practitioners have taken many forms. In 2016, we launched an initiative to give this emerging community a voice, place, and professional development resources, starting with the Black in AI Workshop. This quickly evolved into the Black in AI nonprofit organization with a mission to shift the power dynamic across the AI ecosystem and help visionaries, creators, thought leaders and builders maximize the multifaceted future of artificial intelligence. Black in AI has granted more than $1 million to support some 400+ AI practitioners to be present at major AI conferences. Black in AI now counts more than five thousand (5000+) members – from beginners to senior professionals and academics, and has helped to prepare 800+ students to apply to graduate school successfully. I am honored to serve as the organization's Board President alongside incredible colleagues and tireless volunteers.

As an organization, we envision a barrier-free field that empowers our community to contribute and accelerate their best, most brilliant work for themselves, fellow practitioners, and their global ecosystems. To this end, Black in AI implements various programs, which include: community engagement, such as conferences & socials, ecosystem support, professional membership, and events; programs to advance educational & career pathways, including research programs, and our Emerging Leaders in AI (ELAI) graduate preparation program; and initiatives to support civil society & policy, research & advocacy, and entrepreneurship innovation, which helps to reduce the financial and social gaps that make it harder for Black Entrepreneurs to be successful – such as our recent whitepaper for the Congressional Black Caucus on exploring the impact of AI on Black Americans.

While the statistics reflect some of Black in AI's success over the years, they miss the human stories of hundreds of people who have contacted me to tell me how Black in AI played a crucial role in their decision to stay in the field at a time where they may have felt alone, or discriminated against, or were facing some financial or knowledge gap that hindered their progress. I believe we can work towards a world where an organization like Black in AI is not needed, and we hope you will join us as members, allies, and partners toward a more equitable future!

Kamuela Enos, Kari Noe, Jason Leigh - Importance of co-designing tech with the community

The first key component for any technology, whether developed or not, to empower Indigenous communities is equity. By equity, we refer not to the DEI (Diversity, Equity, Inclusion) sense of inclusion, but rather the legal sense of co-production and co-ownership. While focusing on the inclusion of Indigenous people due to historic injustices is important, it still often aligns with colonial systems and structures. A more meaningful approach is to recognize the value of

Indigenous practices to contemporary society for their deep, localized insights. Ancestral knowledge represents thousands of years of optimizing natural systems within regenerative frameworks.

The second component is ensuring that Indigenous communities, who hold this knowledge, have the opportunity to own and co-produce technology. The primary goal here is to sustain ancestral practices for current and future generations within their own communities. Innovations from research initiatives should enhance existing systems of ancestral practice rather than disrupt them. While ancestral practices can evolve, they must remain foundational to the development and implementation of new technologies. For example, when a contemporary R1 University adopts that mindset that traditional practices are actually sciences, and that the indigenous communities that hold these practices are ecosystems of innovation, then the whole engagement framework changes. In this ideal mode of engagement, the community is not sharing their knowledge with a researcher to help forward the research interests of the university/academy. Instead, the research unit understands that the community's work of restoration is vital to regional societal well-being, so it orients its research to support community agency and continuity. In this model, the community transitions from a thing to be studied, to becoming a consultant to help society restore biocentric practices whose viability have lasted centuries.

This orientation supports communities and their experts in evaluating contemporary technologies to achieve ancestral outcomes. Adopting this approach can provide solutions to complex societal issues that benefit both Indigenous and broader communities. The ultimate goal is for these communities to be recognized as ecosystems of innovation for societal well-being, provided that appropriate political and economic infrastructures are established to support their practices.

For example, in visualizing climate data in our National Science Foundation funded Change(Hawai‘i) EPSCoR project, our approach goes beyond creating a functional visualization system. We adopt a holistic design framework that incorporates thoughtfulness and equity throughout the process, from creation to implementation and sustainable use [10]. This means not only co-designing the technology with community members but also integrating their practices of biocultural restoration and ancestral sciences. Our goal is to prioritize ancestral observational methodologies alongside other critical mapping components. This approach empowers communities whose ancestral data is often excluded from large-scale decision-making processes or extracted without consent.

The end goal is to re-normalize Indigenous practices, input, and agency in the development and implementation of technology, not merely as a DEI initiative. By doing so, these highly refined projects can be supported with community ownership and provided with the resources needed to sustain their use.

Carlo Liquido, Amy McKee- Towards Kanaka-centric design

As user experience (UX) designers who grew up in Hawai‘i, we have struggled with reconciling the cultural differences between capitalistic design practices in the “real world” and the community-centered values we learned as children. The idea that we should design products and tools meant to extract maximum profit from people and communities felt foreign to us.

While telling a story one day, we began to imagine a different way that we could approach designing digital products and services, and we looked for guidance and inspiration from Indigenous texts and stories (moʻolelo). One term that resonated with us was “Indigenous resurgence,” which offers an affirmative and abundance-based lens of reclaiming ancestral knowledge (ʻike kupuna) and designing solutions. This concept provided us with the theoretical framework to move our design thinking from a capitalist and colonial paradigm towards a kanaka-centered one.

In Hawaiian, kanaka means “human” or “people”. While the word “kanaka” is often used to refer specifically to Native Hawaiians (such as, Kanaka Maoli or Kanaka ʻŌiwi), many old Hawaiian newspapers of the 1800s used the word “kanaka” inclusively to describe people from all around the world (for instance, Kanaka Haole meant white person, Kanaka Kepanī meant Japanese person, etc.). This broader, inclusive interpretation of “kanaka” aligns with the story of the first kanaka, Hāloa. The story of Hāloa illustrates the familial and interdependent relationship between humans and non-humans, where the land, plants, animals, and other phenomena are our elder siblings and we must care for the earth as the younger sibling.

Our goal in creating the term, “Kanaka-Centered Design”, is not to "Hawaiianize" Human-Centered Design, but to respectfully reclaim and assert a design practice rooted in ʻike kupuna (ancestral knowledge). This approach embodies a deep sense of responsibility and ethics rooted in Hawaiian culture and sovereignty, and is inclusive of all beings, including past (kupuna), present (kanaka), and future generations (moʻopuna).

In essence, Kanaka-Centered Design embraces creating impactful solutions, rather than the relentless need to scale. It centers people and the planet, not profit. It designs for the reciprocal and regenerative, not for the extractive. It designs for future generations, not for the fleeting trends of the market.

Organizations like Nalukai and ’Ainaquest embody Kanaka-Centered Design by shifting the focus from a “designer versus user” mentality to one of community empowerment. Nalukai mitigates the issue of savior-driven problem solving by reframing problems as “kuleana,” which focuses on the privilege and responsibility to serve one’s community, thus eliminating the distinction between the designer and the user. ʻĀinaquest is an educational card game that challenges the community to cultivate native and canoe plants in order to deepen their collective pilina (relationships) with the land. This emphasis on reciprocity and shared responsibility resonates deeply with ʻike kupuna, as exemplified in the design of loko iʻa (fishponds), which are complex aquaculture systems. The design of loko iʻa demonstrates a deep understanding of ecological balance and a commitment to long-term sustainability. By working in harmony with natural systems, rather than exploiting them, loko iʻa provided abundance for generations, showcasing the wisdom and ingenuity of ancestral knowledge.

Kanaka-Centered Design also champions excellence, recognizing the inherent gifts within individuals and communities, and fostering their development. In Hawaiʻi, we have experienced and witnessed the detrimental effects of settler colonialism, which often imposes low expectations on our youth and limits their potential. As designers, we echo Queen Kapiʻolani’s call to “Kūlia i ka nuʻu –strive for the highest,” and not only empower individuals to reach their full potential, but also provide the resources to nurture new talent. This is exemplified in the work of Piʻikū Co., which empowers aspiring web designers in Hawaiʻi by providing resources, training, mentorship, and compensation–all the tools necessary for new talent to thrive. This not only nurtures new talent but also creates opportunities for excellence by challenging interns to solve real-world problems that directly affect their community.

In conclusion, Kanaka-Centered Design provides us a way to decolonize our UX methodology and encode ancestral knowledge into our designs. It is our way of recognizing that the future of our planet depends on designers who think holistically, sustainably, and inclusively of human and non-humans, and of past (kupuna), present (kanaka), and future generations (moʻopuna).

Tai-Quan Peng- From promise to practice: Overcoming barriers in computational communication research in the Asia-Pacific region through AI

Computational methods have been heralded as a transformative force in communication research, promising new insights and methodological advancements [11]. However, the expected revolution in the Asia-Pacific region has yet to fully materialize. Computational communication studies are predominantly conducted by scholars from well-represented, often Western, societies. This has resulted in a lack of attention to under-represented communities within the Asia-Pacific region, where unique cultural, linguistic, and social dynamics offer rich research opportunities.

The limited adoption of computational methods in these regions can be attributed to several key challenges. For instance, the technical resources and education for computational research, to provide expertise in programming and data analysis, may be beyond the reach of many scholars in underdeveloped areas. Additionally, the linguistic diversity of the Asia-Pacific region presents a formidable barrier [12]. To date, many natural language processing tools are developed for Western languages, requiring extensive customization to be applicable in local contexts. This adaptation process is both time-consuming and costly, further deterring researchers from utilizing these methods. Moreover, the high costs associated with developing and maintaining the necessary computational infrastructure pose a substantial financial obstacle. This financial barrier exacerbates the digital divide, leaving many researchers unable to contribute to or benefit from advancements in computational methods.

To address these disparities, leveraging current AI technologies, particularly large language models (LLMs), offers a promising solution. LLMs can substantially lower the technical barriers by providing user-friendly interfaces for data analysis and natural language processing tasks, making these tools more accessible to scholars without advanced technical training. Additionally, LLMs are increasingly capable of handling linguistic diversity, allowing for more accurate analysis of texts in various local languages. The decreasing cost of deploying AI technologies also makes it more feasible for researchers to use these powerful tools with less financial burden. By integrating LLMs into their research, scholars in the Asia-Pacific may be able to overcome many of the current obstacles, fostering a more inclusive and comprehensive understanding of communication phenomena in the region.

LLMs can capture the rich tapestry of languages and dialects of the region, enabling researchers to conduct cross-linguistic studies with greater ease. This capability helps break down language barriers, ensuring that research is inclusive and representative of the region’s diverse population.

Moreover, LLMs can also potentially enhance the testing of various interventions. That is, by carefully controlling the inputs through appropriate prompts and training data, LLMs can simulate the impact of different measures aimed at promoting innovative ideas and technologies, altering health behaviors, or implementing new policies. This control allows researchers to test various scenarios and predict outcomes with high accuracy before implementing interventions on a larger scale. In regions like the Asia-Pacific, where socio-political contexts vary widely, such controlled simulations provide critical insights, enabling more informed and tailored decision-making.

Despite its potential, the application of AI in computational communication research in the Asia-Pacific region is not without limitations. One major concern is the bias inherent in AI models, often skewed toward WEIRD (Western, Educated, Industrialized, Rich, and Democratic) populations [13]. This bias arises from inadequate data representing the cultural and linguistic heterogeneity of the Asia-Pacific region, leading to models that may not accurately reflect local nuances. Additionally, the absence of comprehensive "ground truth" datasets complicates the validation of AI performance, risking erroneous conclusions. If AI models continue to be trained predominantly on data from WEIRD populations, they may perpetuate existing biases and fail to address the unique challenges of the Asia-Pacific region. Ensuring that AI tools are developed and validated using diverse datasets is crucial to prevent this outcome. Furthermore, the persistent insufficiency of infrastructure that impedes the development of computational communication research also hampers the effective deployment of AI technologies in the Asia-Pacific region, posing significant challenges to fully leveraging AI's potential.

Karaitiana Taiuru- Treatment of Māori language in language modeling

The Māori language was banned by native schools and other government led assimilation practices in the late 18th century, so effectively, that by 1980 we had less than 20% of native speakers nationwide, and within my own tribe we had 3 native speakers. Tribal dialects were also replaced with one standard version of Māori, influenced by the introduction of the written Bible and by ethnographers who chose to ignore the rich and diverse tribal dialects throughout our country in favor of a standard dialect. Large Language Models (LLMs) have the ability to revitalize dialects, it is not likely a preference for Māori as it is a local distinguishing treasure that is used by tribal members in physically meetings of cultural significance.

Community activism throughout the late 1970's and the establishment of language training for preschoolers and other educational facilities led to the Māori language being recognized as an official language in 1987. The key lessons learned for Māori was that the language had to be normalized in our lives and society, and then to be spoken by at least three generations within one family in order to be preserved. We have now reached that lofty dream, but moving forward we need to address the controversial topic of LLMs that are widely used and that already incorporate our language.

The Māori language has a substantial amount of digitized online resources such as legal records, parliamentary corpus, journals, newspapers, archives and audio-visual materials, and with tens of thousands of new words created to accommodate the translation of products such as Microsoft Office, Windows, and Google. Google had a Māori software engineer contribute to and help develop Google Translate for Māori. This has generated a huge amount of data for artificial intelligence (AI) to incorporate the Māori language and for it to be used relatively well.

For instance, ChatGPT already speaks Māori at a reasonable level of accuracy. I estimate that in less than two years it will be as fluent and accurate as a Māori language expert.

A recent spike with people wanting to learn the Māori language has resulted in a demand issue for teachers that far outweighs the supply. This has led to a large uptake of learners of Māori language using LLM's as both a supplementary tool and an alternative method to learn Māori. Still, many dangers exist. For example, we provide AI technologies with our sacred rituals and esoteric knowledge, our language and history will not be our own and risks being commercialized and changed by international corporations. As Māori, we need to revisit our traditions, go back to our tribal lands and re-engage with elders and tribal members to learn the sacred aspect of our language and ensure that those aspects remain in our human world and in our tribal homes and lands. LLM developers can assist in this area by acknowledging copyright in source data and being transparent about data sources.

Other risks include misogynist ethnographers' historical texts being used to train some LLMs, resulting in incorrect statements about Māori, written in the Māori language. This is predominantly with our creation stories and historical knowledge, where the Māori translated King James Bible is being offered as Māori creation stories and the LLMs are mixing and matching tribal stories to create new ones.

Another emerging risk is phishing attacks or identity theft, as spammers are using LLMs to correspond with Māori and it's becoming more difficult to discern real from fake. Previously, Google Translate often made mistakes when translating, making it easier to identify. LLM developers should consider how best to prevent their tools being used in phishing, bullying and other common forms of online scams and bullying. Moving forward, we hope that LLMs are key to the long-term revitalization and normalization of the Māori language.

Competing Interests

The authors declare no competing interests.

REFERENCES

[1] Chen, C.Y., Kahanamoku, S.S., Tripati, A., Alegado, R.A., Morris, V.R., Andrade, K., & Hosbey, J. Systemic racial disparities in funding rates at the National Science Foundation. *eLife* 11, e83071 (2022). https://doi.org/10.7554/eLife.83071

[2] Ginther, D.K., Schaffer, W.T., Schnell, J., Masimore, B., Liu, F., Haak, L.L., & Kington, R. Race, Ethnicity, and NIH Research Awards. *Science* 333(6045), 1015–1019 (2011). https://doi.org/10.1126/science.1196783

[3] Taffe, M.A. & Gilpin, N.W. Racial inequity in grant funding from the US National Institutes of Health. *eLife* 10, e65697 (2021). https://doi.org/10.7554/eLife.65697

[4] Kraus, M.W., Rucker, J.M., & Richeson, J.A. Americans misperceive racial economic equality. PNAS 114(39), 10324–10331 (2017). https://doi.org/10.1073/pnas.1707719114

[5] Bol, T., de Vaan, M., & van de Rijt, A. The Matthew effect in science funding. PNAS 115(19) 4887–4890 (2018). https://doi.org/10.1073/pnas.171955711

[6] Graddy-Reed, A. & Lanahan L. Prioritizing diversity? The allocation of US federal R&D funding. *Science and Public Policy* 50(1), 104–119 (2023). https://doi.org/10.1093/scipol/scac052

[7] R. T. Miller, A. G. Christoffels, C. Gopalakrishnan, J. A. Burke, A. A. Ptitsyn, T. R. Broveak and W. A. Hide (1999). A comprehensive approach to clustering of expressed human gene structure: The Sequence Tag Alignment and Consensus Knowledgebase. *Genome Research* 9(11): 1143-1155

[8] Baguma, Rehema & Namuwaya, Hajarah & Nakatumba-Nabende, Joyce & Rashid, Qazi. (2024). Examining Potential Harms of Large Language Models (LLMs) in Africa. 10.1007/978-3-031-56396-6_1.

[9] Coe, Michael D., and Richard A. Diehl. In the Land of the Olmecs: The Archaeology of San Lorenzo Tenochtitlan. Vol. 2 Austin: University of Texas Press, 1980.

[10] Kari Noe and Nurit Kirshenbaum. 2024. Where Generalized Equitable Design Practice Meet Specific Indigenous Communities. In Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24). Association for Computing Machinery, New York, NY, USA, Article 616, 1–8. https://doi.org/10.1145/3613904.3642931

[11] van Atteveldt, W., Peng, T.Q., When Communication Meets Computation: Opportunities, Challenges, and Pitfalls in Computational Communication Science. Communication Methods and Measures, 12(2–3), 81–92. https://doi.org/10.1080/19312458.2018.1458084

[12] Peng, T. Q., Liang, H., & Zhu, J. J. H. (2019). Introducing computational social science for Asia-Pacific communication research. Asian Journal of Communication, 29(3), 205–216. https://doi.org/10.1080/01292986.2019.1602911

[13] Santurkar, S., Durmus, E., Ladhak, F., Lee, C., Liang, P., & Hashimoto, T. (2023). Whose Opinions Do Language Models Reflect? (arXiv:2303.17548). arXiv. https://doi.org/10.48550/arXiv.2303.17548

Contributors:

1) Dr. Christine Yifeng Chen is a staff scientist at Lawrence Livermore National Laboratory. As a geologist and geochemist, she researches radiochronometric dating tools for nuclear forensics investigations and supports the International Atomic Energy Agency's nuclear safeguards mission. In late 2020, she co-founded Asian Americans and Pacific Islanders in Geosciences, a grassroots organization building a community that supports Asian American Pacific Islander researchers and students in the broader earth sciences. This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344 (LLNL-JRNL-867596).

2) Dr. Alan Christoffels is the director of the South African National Bioinformatics Institute (SANBI) based at the University of the Western Cape. He is also the director of the SA Medical Research Council Bioinformatics Unit. He is passionate about understanding the genetic drivers of bacterial drug resistance. His research lab focuses on the intersection

of bioinformatics and public health building computational tools to analyze high throughput genomics data. Among his software tools is the Baobab Laboratory Information Management System used by biobanks in Africa. He is leading the Public Health Alliance for Genomic Epidemiology consortium and is an advisor to the Africa CDC for pathogen genomics.

3) Dr. Roger Dube holds a BA in physics from Cornell and a PhD in physics from Princeton. Currently Professor Emeritus at Rochester Institute of Technology and a previous Fulbright scholar at the University of Manitoba, he has held positions at Kitt Peak National Observatory, NASA’s Jet Propulsion Laboratory, the University of Michigan, the University of Arizona and Yale. He is a Board Member of the American Indian Science and Engineering Society (AISES) and chairs the Higher Education Committee. He is cofounder of the Native Alumni of Princeton. He is an enrolled member of the Abenaki Nation and Mohawk Turtle clan and is a Sequoyah Fellow. He was awarded the 2019 Ely Parker Award from AISES.

4) Kamuela Enos is the Director of the Office of Indigenous Knowledge and Innovation and the Vice President for Research and Innovation in the University of Hawaiʻi system. He is also a lecturer at UH Mānoa’s Department of Urban and Regional Planning. He was previously the director of social enterprise at MAʻO Organic Farms, a non-profit that aims to connect Waiʻanae youth to the land.

5) Dr. Juan E. Gilbert is the Andrew Banks Family Preeminence Endowed Professor, UF Distinguished Professor and Chair of the Computer & Information Science & Engineering Department at the University of Florida where he leads the Computing for Social Good Lab. He is a Fellow of the Association of Computing Machinery (ACM), a Fellow of the American Academy of Arts and Sciences, a Fellow of the IEEE, a Fellow of the American Association of the Advancement of Science (AAAS), and a Fellow of the National Academy of Inventors. Dr. Gilbert was named a laureate of the National Medal of Technology and Innovation. Dr. Gilbert received his MS and PhD degrees in computer science from the University of Cincinnati. He also received his BS in Systems Analysis from Miami University in Ohio.

6) Dr. Sanmi Koyejo is an Assistant Professor in the Department of Computer Science at Stanford University. Dr. Koyejo leads the Stanford Trustworthy Artificial Intelligence Research (STAIR) lab, which works to develop the principles and practice of trustworthy machine learning, focusing on healthcare and scientific applications. Dr. Koyejo has received outstanding paper awards, a Skip Ellis Early Career Award, a Sloan Fellowship, an NSF CAREER award, a Kavli Fellowship, and an IJCAI early career spotlight, among others. Dr. Koyejo serves on the Neural Information Processing Systems Foundation Board, the Association for Health Learning and Inference Board, and as the President of the Black in AI organization.

7) Dr. Jason Leigh is the Director of the Laboratory for Advanced Visualization & Applications, Co-director of the Hawaii Data Science Institute at the University of Hawaiʻi at Mānoa, and Director of Create(x) at UH West Oʻahu. He was Director Emeritus of the Electronic Visualization Lab at the University of Illinois at Chicago and held research

appointments at Argonne National Laboratory and the National Center for Supercomputing Applications. His research interests include data visualization, virtual reality, indigenous design, and artificial intelligence.

8) Carlo Liquido, born and raised in Hawaiʻi, is a designer and entrepreneur passionate about leveraging technology for social good. Currently, Carlo is the Executive Director of Piʻikū Co., a workforce development non-profit dedicated to shaping a more equitable future in Hawaiʻi. He also serves as the Director of Entrepreneurship Education at Nalukai, a social and cultural entrepreneurship camp for high school students. He previously led design initiatives for startups in the global healthcare and food security sectors, as well as for large corporations like Autodesk and Shopify. Carlo earned his Economics degree from Brown University and holds a Masters in Information from UC Berkeley, where he served as an adjunct professor teaching full-stack development.

9) Amy McKee is a Kanaka ʻŌiwi (Native Hawaiian) designer, researcher, social entrepreneur, and PhD student in Educational Foundations from Kaimukī, Oʻahu. Amy is passionate about supporting Native Hawaiians in education and tech, developing sustainable communities, and leveraging technology to create positive change. She uses her insights on Indigenous education and research methods to design products that are place-based and community-centered. She is a volunteer at Piʻikū Co. and a co-founder of ʻĀinaQuest, a Hawaiʻi-based game company on a mission to cultivate pilina (relationships) with native plants and foster new generations of ʻāina stewards.

10) Kari Noe is a research assistant at the Laboratory for Advanced Visualization and Applications (LAVA), and a PhD Candidate in Computer Science at the University of Hawaiʻi at Mānoa. She also is the Indigenous Technology Specialist at the University of Hawaiʻi Office of Indigenous Knowledge and Innovation (OIKI). Her research interests are Data Visualization, Immersive Technologies, Equitable Design, and Human Computer Interaction. She co-directs Create(x), an emerging media laboratory for students to develop and showcase their works at the intersection of creative media, computation, culture, and science at the Academy of Creative Media in the University of Hawaiʻi at West Oʻahu.

11) Dr. Tai-Quan Peng is a Professor in the Department of Communication at Michigan State University. His research interests encompass computational social science, health communication, audience analysis, and political communication. He served as the inaugural chair of the Computational Methods Division at the International Communication Association. In 2018 and 2019, he co-edited special issues on computational communication research for *Communication Methods and Measures* and the *Asian Journal of Communication*, respectively. He is currently the Associate Editor of the *Journal of Communication* and the incoming Editor-in-Chief of *Human Communication Research*.

12) Dr. Karaitiana Taiuru is a leading authority and a highly accomplished visionary Māori technology ethicist specializing in Māori rights with AI, Māori Data Sovereignty and Governance with emerging digital technologies and biological sciences. A professional director with membership with Institute of Directors, with roles including membership and Kahui Māori advisor of the New Zealand AI Forum, member and tangata whenua governor of the AI Researchers Association, invited member of the Expert panel on AI and healthcare for the Office of the Prime Minister's Chief Scientific Advisor, a legislated

expert member of the Intellectual Property Office (IPONZ) Trade Marks Advisory Committee.

Stock Image:

https://www.gettyimages.com/detail/illustration/co-working-meeting-royalty-free-illustration/587214234

Credit: Comomolas / Getty Images