



THE MISSING MILLIONS: BRIDGING DIGITAL DIVIDES

SCIENCE AND TECHNOLOGY SUFFER FROM UNDERREPRESENTATION OF MANY SOCIAL GROUPS, WHICH AFFECTS THE QUALITY AND SOCIETAL IMPACT OF NEW ADVANCEMENTS. THIS IS ESPECIALLY TRUE IN DATA AND COMPUTATION, WHERE THE LACK OF ACCESS AND PARTICIPATION FOR MANY MEANS THAT THEIR VIEWS, VALUES AND SKILLS ARE NOT ACCOUNTED FOR IN NEW DEVELOPMENTS

Alan Blatecky has been leading a pioneering project from the National Science Foundation entitled “The Missing Millions: Democratizing Computation and Data to Bridge Digital Divides and Increase Access to Science for Underrepresented Communities.” From his position at RTI International in North Carolina in the U.S., he worked with other researchers and countless practitioners to build a strong case for the democratisation of computational resources.

The report focuses on the many millions of people who are currently excluded from the world of research. “The ‘missing millions’ are those students, faculty and citizens who are not involved or engaged in STEM activities because they don’t have access to the digital resources or technologies required,” says Alan. “This means they are severely limited in their ability to fully participate in the economic and research engines driving society.”

Given that these people are underrepresented, their knowledge and perspectives are also marginalised within science and technological advancement. “There are countless research questions of importance to these missing millions that are not being pursued because of

the lack of support,” explains Alan. Ensuring underrepresented people have the opportunity to participate in, drive and lead research, specifically through computational and data means, is the focus of the Missing Millions report.

ACCESS AND ABILITY

Providing access to resources is often identified as a straightforward solution, but access alone will not solve this issue. Digital technologies require some level of expertise to be able to use them; any forward-thinking programme needs to facilitate training, as well as supply resources. “Although having access is fundamental, without having a corresponding ability to use and deploy computation and data resources, providing access is a hollow goal,” says Alan. People need the ability to use statistical software, data visualisation tools, data curation and storage infrastructure, data quality assurance and artificial intelligence. They also require equitable funding for training, hiring qualified personnel and maintaining infrastructure.

Despite the assumption that science is objective, scientific and technological progress

is inevitably shaped by the demographics of the researchers behind it. “We know there is the risk of bias and adverse societal impacts of data, so the people who are at risk need to be part of the solution,” says Alan. “Without having the ability to use computation and data resources, the digital divide remains and, worse, will continue to expand.”

THE RESEARCH PROCESS

The Missing Millions project involves engaging with a diverse range of people. “The people we talked to range from cyberinfrastructure experts to computer and data researchers and practitioners, to leaders of underrepresented institutions, as well as organisations focused on broadening data access,” says Alan. “Broad communication was essential in this research as the changes required go far beyond just making targeted investments with technology.”

Beyond discussions, the project also helped set the groundwork for building the networks needed to bring the discussed improvements to fruition. “Changes need to provide opportunities for people to familiarise themselves with data and computer



The opinions, findings, conclusions, and recommendations expressed are those of the co-authors and do not necessarily reflect the views of the National Science Foundation. Award number NSF OAC 2127459.

We acknowledge the traditional owners of the lands where Washington, Maryland and Virginia now stand – the Pamunkey, Chickahominy, Upper Mattaponi, Rappahannock, Monacan and Nansemond. We pay respect to their Elders past, present and emerging.



ALAN BLATECKY

RTI International, Research Triangle Park, North Carolina, USA

FIELD OF RESEARCH

Cyberinfrastructure
(for Research and Education)

RESEARCH PROJECT

'The Missing Millions: Democratizing Computation and Data to Bridge Digital Divides and Increase Access to Science for Underrepresented Communities'

FUNDER

National Science Foundation (NSF)

TOP TIPS FOR YOU AND YOUR STUDENTS

- 01** Find things that you're passionate about and people who are also passionate about them.
- 02** Focus on local social or science issues that are important to your community.
- 03** Work with others to accomplish together what you can't do alone.
- 04** Consider what your role is in contributing to your communities in the future. What is your responsibility today to prepare for tomorrow?

resources, so they can build experience and expertise," says Alan. "This can only be accomplished by broad educational, social and community efforts."

ENTERING THE FIELD

Established entry-level work and training opportunities are often emphasised as gateways into research careers but often fail to meet their intended purpose. "Many internships and apprenticeships are given to those students already 'qualified,' who have already made it into the research community, while others continue to remain outsiders," says Alan. "To address this, these programmes must have a component that enables those with non-traditional backgrounds to participate."

Alan's team engaged closely with educators from establishments such as Historically Black Colleges and Universities (HBCUs), Tribal Colleges and Universities (TCUs), and other Minority-Serving Institutions (MSIs) to understand how these barriers could be overcome. The team found that approaching local issues through research provided a gateway to build communities of students that could tackle issues that matter to them, while

simultaneously building research skills.

THE IMPORTANCE OF COMMUNITY

The project concluded that engaging the missing millions was best achieved at the grassroots level, by supporting communities to come together to address research issues that affect them. "A sense of community in research has always been important," says Alan. "However, to become a member of the research community, you often need to have certain resources or backgrounds. Those individuals who don't have the 'correct' experience or sufficient financial capacity are often left out." Local communities that focus on local issues can help build the skills needed to access the broader research community.

"One of the most effective ways to nurture research communities at the ground level is to enable small groups of students to address and explore issues in their own neighbourhoods," says Alan. "By addressing local issues that matter to them and their local communities, students can begin to create their own research communities. They also gain a sense of belonging and receive validation from their peers and community." Ultimately, these efforts can feed into broader scientific

projects, helping to enrich findings through the inclusion of diversity and new perspectives.

MODERN CHALLENGES

The global issues facing the modern world – climate change, pandemics and conflicts – affect the marginalised in society more severely. To ensure that science and technology address these impacts and meet the needs of the people affected, it is essential that the voices of the missing millions are included. "It isn't just about having large numbers of people engaged but also ensuring that large numbers of people from diverse backgrounds, experiences and perspectives are asking questions and exploring solutions," says Alan. "The complexity of research demands it."

Reiterating the vital combination of resources, training and engagement that meaningful research requires, Alan explains, "We use the expression 'access + ability = capability.' Nurturing this capability to cross the digital divide is essential to allow the missing millions to participate fully in the economic prosperity and opportunities of modern society."

MEET THE TEAM



LAUREN MICHAEL



RESEARCH COMPUTING FACILITATION LEAD, CENTER FOR HIGH THROUGHPUT COMPUTING (CHTC), UNIVERSITY OF WISCONSIN-MADISON, USA

Lauren's role involves engaging with researchers to advise on the most impactful computing approaches, services and relevant learning pathways.

As well as addressing disparities within its domain, computing also has an important role to

play in democratising science across all research domains and underrepresented groups by making research data and methodologies accessible to new communities of researchers. Repeatedly, we see research emphasising the importance of diversity and representation to the quality of the nation's science and engineering endeavours.

We specifically looked for individuals who could provide insights around multiple angles addressed by the study and also asked them for recommendations from their networks. This included people experienced in expanding diversity and representation at different levels and people bringing research, data and computation to wider audiences.

Educators showed us the need to incorporate computational training at earlier stages. However,

they lack access to standardisable datasets, tools and methodologies that they can incorporate into classrooms easily. Additionally, there is a lot of research demonstrating that engagement is best driven by societally relevant research problems and datasets rather than methodologies and tools.

I was led to this role by my passion for communication and teaching, for computation and for impacting and learning about a wide range of research topics. The greatest rewards continue to be the achievements and gratitude of the thousands of researchers who I have been fortunate to work with, where I have seen significant gains in the diversity of research and researchers using large-scale computing over the last decade.

I think the most successful community-building efforts

engage members to 'own' their organisational work by building in processes that achieve ongoing, regular community input and by creating opportunities for members to share the work.

Communication is the most important skill for Research Computing Facilitation and probably for any form of facilitation. You do not have to know all the answers, but you do have to really understand the client's questions, navigate the plethora of people and information sources that have the answers and then translate these for the client's specific context and background.

A diverse and representative research community will ensure that the STEM community is appropriately prioritising research that representatively addresses societal issues, including other social disparities.

PROFESSOR JOEL E. CUTCHER GERSHENFELD



THE HELLER SCHOOL FOR SOCIAL POLICY AND MANAGEMENT, BRANDEIS UNIVERSITY, BOSTON, MASSACHUSETTS, USA

Joel's career spans social impact enterprises, large-scale systems change, cyberinfrastructure and many other areas.

Long-standing disparities in society have become more visible

with the rise of social protest and the COVID-19 pandemic. The Missing Millions project's mission to democratise data and computation is an essential and positive response to these inequalities.

We conducted online interviews with small groups of participants covering the following topics:

- Visions of success: what the next 10 years could look like if efforts to democratise research computing were successful.
- Enabling technologies: how new or emerging technologies, or novel research applications, could facilitate this democratisation.
- Communities of practice: how to encourage diverse and inclusive innovation driven by communities.
- Institutional and societal barriers: entrenched issues, potential risks and avenues for overcoming these barriers.

Discussions were lively and enthusiastic, with many participants offering to take the conversations further.

From speaking to educators, we found that there is great innovation at the "edge" of computation. For instance, electronic lab notebooks are helping make science more accessible. The potential for communities to control their own data is a powerful idea.

Much more data could be open access than they are at present. Fields and disciplines with more openly available data will progress faster. Considerable work is needed to reduce the disincentives to make data available and provide accessible computing infrastructure.

Inclusive and representative research communities need strong leadership, effective collaboration

and appreciation of diverse voices. They also need improved infrastructure and funding and further support in removing barriers to progress.

My career has seen several "eureka" moments that I take pride in. I have helped develop what has become the Ford-UAW Production System and pioneer new strategies for labor-management negotiations. In recent years, I have helped document the rise of the third digital fabrication revolution and how it relates to broad societal challenges.

The Missing Millions project aims to inspire a collaborative mindset, which leads to diverse, inclusive and equitable institutions in the 21st century. It is essential that we learn to accomplish together what we can't do separately.

CHRIS CALDWELL



PRESIDENT OF THE COLLEGE OF MENOMINEE NATION (CMN), KESHENA, WISCONSIN, USA

Chris has held positions as Tribal Resources Director/Compliance and Enforcement Officer for the Menominee Indian Tribe, forest products technician with the USDA Forest Service's Forest Products Laboratory in Madison and Director of Sustainable Development Institute.

Through my work with our Sustainable Development

Institute (SDI) at the College (CMN), I understand how a lack of control over the research process can negatively impact Tribal communities. Being able to control the research process means prioritising topics of inquiry relevant to the community, employing culturally appropriate approaches to addressing the topic and having access to resources and expertise to carry out the work. Tribal research institutions often have limited access to the funding, resources and infrastructure required to carry out the necessary research – the Missing Millions report helps to address this.

I was asked to participate in a focus group of Tribal Colleges and Universities (TCUs) presidents and leaders. This came at the same time the SDI and the University of Wisconsin-Madison were discussing ways to increase capacity for storing, managing

and protecting data derived from research projects at CMN SDI.

I'm a PhD candidate in Environment and Resources, looking at how environmental issues impact human and environmental relationships. This is important to me as a researcher and educator for my Tribe.

The Menominee theoretical model of sustainability was developed by Menominee Tribal leaders and is used to develop a research agenda, processes and projects. SDI is the steward of this model of sustainability on behalf of the community.

As College President, it is my job to communicate with students, Tribal communities and the general public. Past Tribal leaders have been forward-looking, thinking about future generations and what decision-making today means for that future. I see my

role continuing this.

Our focus as a Tribal college is how we can help our Tribal communities. We understand our students need to be prepared to engage the broader world – student researcher and research leadership opportunities are important and can be achieved by providing more resources and funding to institutions like TCUs.

Reaching the missing millions will see Indigenous peoples having more control over research.

The data generated in their communities is important for sovereignty, self-determination and revitalising Indigenous approaches to inquiry. Accepting and valuing other ways of knowing beyond dominant Western-based research paradigms is important to this story as well. This is a grounding point from which STEM can advance beyond its current state.

DR. DEBORAH F. DENT



CHIEF INFORMATION OFFICER AT JACKSON STATE UNIVERSITY, JACKSON, MISSISSIPPI, USA

Deborah oversees the management of the university's cybersecurity and network and communications infrastructure. She is a founding member of the Minority Serving – Cyberinfrastructure Consortium (MS-CC), working to advance cyberinfrastructure capabilities on historically underserved campuses.

I was drawn to the Millions project's goals because I support the desire to expand, diversify and support the development of new communities of researchers. How

else can we truly begin to address pressing research, social and global issues today and in the future? I'm a co-author of the report and participated in meetings, focus groups and webinars.

At the Minority Serving - Cyberinfrastructure Consortium, we envision a transformational partnership to promote advanced cyberinfrastructure capabilities on historically underserved campuses with data, research computing, teaching, curriculum development and implementation, collaboration and capacity-building connections among institutions. We're committed to learning and adjusting our operational model to ensure that we continue to lift all participating institutions by advancing cyberinfrastructure for research and education across diverse fields, disciplines and communities.

We have the data and empirical evidence that show what the immediate cyberinfrastructure needs for Historically Black Colleges and Universities

(HBCUs), Hispanic-Serving Institutions (HSIs), Tribal Colleges and Universities (TCUs), and other Minority-Serving Institutions (MSIs) are and what barriers are preventing campuses from achieving them. A common need is the actual infrastructure – the hardware and connectivity – without which you can't enable data, computing services and resources essential to supporting advanced science, engineering and mathematics.

There has been an increase in pandemic-related government funding, allowing campuses to improve their basic cybersecurity equipment and services. Recently, the MS-CC, in collaboration with Internet2 and American Indian Higher Education Consortium, received funding to provide cyberinfrastructure resources, support and training for HBCUs and TCUs.

To nurture inclusive research communities, we must understand the barriers limiting access to research data and computing, and

that's what the Missing Millions report addresses. We must be intentional in making incremental changes while addressing the broader, systemic changes in the science and engineering enterprise.

Mentorship works! Creating mentorship opportunities for students, such as afterschool programmes and summer camps, can have profound impacts on increasing STEM participation. This is especially needed for students living in rural areas.

I believe inclusive and representative teams across STEM fields will bring different life experiences and add valuable perspectives to research. A more diverse research community will see our societies flourish as we begin to tackle some of the world's most pressing research problems.