

**Empowering Health Data Science Through The
Philippine Datathon:**

How Artificial Intelligence Can Contribute to Universal Health Coverage

RFS, MD

Datathon 2023, Overall Committee Chair

ABM, MD

Datathon 2023, Committee Co-Chair

BJT, MD

Datathon 2023, Ways and Means Committee Head

PCZ

Datathon 2023, Technical Committee Head

BPS, PhD; MC III, PhD

Datathon 2023, Logistics Committee Heads

RDC, MD; RAC, PhD; ARL, PhD

Datathon 2023, Organizing Committee Members

RBN, RN

Datathon 2023, Registration Desk Lead

JMF

Datathon 2023, Secretariat

Abstract

The Philippine Datathon 2023 aimed to convene advocates of health data governance and management to explore the integration of artificial intelligence (AI) within the health sector's UHC initiative. It was the third in a series following previous conferences such as Big Data Philippines in 2017 and 2018. The event utilized a dynamic approach incorporating lectures, panels, workshops, and a competition (datathon) to foster active participation and collaboration among attendees. Key insights were distilled from presentations used by the speakers.

The datathon successfully cultivated a network of institutions and individuals committed to advancing data governance and management. Critical discussions surfaced recurring challenges and ethical considerations necessary for the responsible use of AI in healthcare, which were meticulously documented for future reference.

In conclusion, Philippine Datathon 2023 underscored the critical role of a united community in driving forward the ethical application of AI in healthcare. It highlighted the necessity for ongoing dialogue to demystify AI's complexities and leverage its potential responsibly. The datathon set the stage for future collaborations aimed at enhancing public health outcomes through innovative data-driven approaches, reinforcing the imperative for ethical consideration and community engagement in the evolving landscape of health data management and AI.

1. Program Overview

The event takes place at the Henry Sy, Sr. Hall, De La Salle University, Taft Avenue, Manila. The schedule for the first day is structured as follows:

Registration: Attendees were welcomed to register, facilitated by the datathon staff. Prayer and National Anthem: Dr. Ruth Angelie Cruz led the reflective prayer and the national anthem to start the proceedings.

Dr. Raymond Francis Sarmiento gives an introduction, setting the stage for the day's agenda.

Dr. Raymond Girard R. Tan, focusing on the event's importance and the collaborative spirit.

Welcome Remarks were given by NIH Executive Director, Dr Eva Maria Cutiongco-dela Paz.

PCHRD Executive Director Dr Jamie C. Montoya provided an important address that contributed to the thematic depth of the event.

Objective: Gather advocates of data governance and management and discuss opportunities and challenges for artificial intelligence in the health sector.

Methods: The event adopted various methods such as lectures, panels, workshops, and a competition. Participants were actively engaged to participate through open fora and collaborative activities. Slides were collected and summarized into their salient points using ChatGPT4. Authors validated the output.

Results: A network of advocates, composed of institutions and individuals, was formed. Recurring issues on data governance and management including ethics were identified and documented in these proceedings.

Conclusion: The Philippine Datathon 2023, the third in a series of activities after Big Data Philippines 2017 and 2018, demonstrated the importance of a community of advocates to stimulate and maintain interest in the ethical management of health data and its use for artificial intelligence. Continued discourse among the advocates is key to clarifying ambiguous concepts of artificial intelligence and in presenting its risks and benefits for advancing public health objectives and for fostering innovation.

2. Introduction

The growing importance of data science in public health is driven by its ability to analyze vast amounts of health data to identify trends, predict outbreaks, and improve healthcare delivery. By leveraging machine learning, artificial intelligence, and statistical tools, data science enables public health professionals to make data-driven decisions, enhancing the effectiveness of health interventions and policies for better population health outcomes.

Since 2011, the University of the Philippines Manila has hosted the MIT Critical Data team for a series of visits and lectures around health information management. In 2017, the first Big Data Philippines was held with co-organization with UP Cebu followed by another one co-hosted by UP Diliman in 2018. For this third in a series, the new organizers, CARE PH, De La Salle University agreed to rename the event to Philippine Datathon 2023. The theme selected was "How can artificial Intelligence contribute to universal health coverage" with the goal of threshing out ideas on the ability of new and emerging technologies in catalyzing healthcare in the country.

3. Methodology

This report contains observations of the authors as well as synthesis of slides and documents provided by resource speakers using ChatGPT4.

4. Results

Setting the context

Prof. Philip Zuñiga focused on the significance of HL7 fast Healthcare Interoperability Resources (FHIR) as a syntactic and semantic standard for health, highlighting its crucial role in data interoperability and efficient decision-making. He listed the challenges of using big data, particularly when the data is invalidated, untimely, and cannot be aggregated, rendering it ineffective for practical use. Prof. Zuñiga elaborated on the benefits of FHIR, including its use of APIs, implementable data model, open-source nature, and maximization of standard web technologies, underscoring its adaptability and efficiency in health data management.

Dr. Romulo de Castro emphasized the importance of preparing data for machine learning (ML) and artificial intelligence (AI) especially in health care. He emphasized the importance of interoperability, ethical use of data, and its governance. He outlined various projects and ongoing initiatives aimed at enhancing health data management and utilization, including efforts to improve data quality, security, and ethical considerations. His main message was the critical role of data governance in achieving effective health data interoperability, supporting patient care, and facilitating research and policy development.

Dr. Jong de Castro and Dr. Trixie Tiangco explored various facets of data science, emphasizing its application in health and health research. They delved into the definition of data science, highlighting its nature as both a science and a multidisciplinary field, and discussed its role in innovative health solutions, including drug discovery and aging research. They each related the state of data science from their perspectives and suggested directions for future development and the importance of collaborative research efforts.

Dr. Leo Celi's presentation on "UHC and the Bias in Health Data" discusses the impact of data bias on artificial intelligence in healthcare, highlighting how such biases in data collection and interpretation can lead to disparities in diagnosis and treatment recommendations. He emphasized the importance of recognizing and addressing these biases to ensure equitable and effective healthcare outcomes. Dr. Celi advocated for relevant education, diversity of expertise, and continuous monitoring in the deployment of AI to mitigate these biases and improve health data's reliability and utility.

Dr. Fresthel Climacosa introduced REDCap (Research Electronic Data Capture) as a secure, web-based application designed for data collection and management in research studies, highlighting its development by Vanderbilt University in 2004. She discussed REDCap's global impact, with a vast network of collaborators and projects worldwide, and its specific application in various COVID-19 related studies. She enumerated REDCap's features, such as high data quality, secured data sharing, timely analysis, and cost-effectiveness, making it a valuable tool for health research data management.

Workshops

In Workshop 1, Matthew McDermott explored the challenge of confounding AI models for medical imaging, particularly focusing on pneumothorax detection. He discussed the importance of identifying and addressing hidden stratifications and shortcut features that could lead to biased or inaccurate model predictions. The session emphasized the value of collaboration in small groups to examine these issues, and in using tools like saliency maps for error auditing and improving model validation.

Workshop 2 with Matthew McDermott discussed the challenge of AI models for medical imaging, particularly focusing on pneumothorax detection. The group detailed the importance of identifying and addressing hidden stratifications and shortcut features that could lead to biased or inaccurate model predictions.

The workshop on "Introduction to Causal Inference" provided foundational knowledge on distinguishing between causation and correlation, emphasizing the importance of understanding causal relationships in healthcare data analysis. The group was introduced to methodologies and statistical techniques for identifying causal effects, including the use of directed acyclic graphs (DAGs) and the do-calculus for controlling confounding variables. The workshop highlighted practical applications of causal inference in health research, illustrating how to apply these concepts to real-world data to inform effective and accurate decision-making.

Giovanni Angelotti facilitated the workshop on the implications of proxies and social determinants of health (SDoH) in healthcare data, with focus on disparities in treatment and outcomes. The group examined the importance of recognizing and addressing proxies of disparity in health data to ensure equitable care and research practices. They studied the critical role of data analysis and management in identifying and mitigating biases related to race, ethnicity, gender, and insurance status, aiming to improve health outcomes and care delivery.

Joe Byers facilitated the workshop that highlighted AI's potential to address healthcare disparities by improving access, diagnosis, and treatment across diverse populations. The group discussed challenges such as data bias and accessibility, and various strategies for mitigating bias in AI systems, including diverse data collection and algorithmic fairness. The presentation underscored the importance of ethical data collection and responsible AI development to ensure equitable and unbiased healthcare solutions.

Dr. Romulo de Castro co-facilitated the workshop that addressed disparities, proxies, and social determinants of health (SDoH) in healthcare. In that group, he highlighted the importance of recognizing and addressing social and environmental factors that impact health outcomes, using real-world examples like medication access issues. The

session emphasized the need for healthcare systems to consider SDoH in patient care and policy making to reduce health disparities and improve overall health equity.

Competition

The "Dengue Dataset MinDSS Against Dengue" team outlined the methodology for collecting and processing dengue incidence data on a weekly basis, aggregating at regional and city/provincial levels, including cases and deaths. They demonstrated data cleaning and preprocessing techniques like outliers and null replacement with 3-day moving averages and uniform date formatting. The presentation showcased the application of LSTM (Long Short-Term Memory) networks for predicting dengue cases, demonstrating improved accuracy over traditional models and highlighting the potential of machine learning in public health decision-making and dengue prevention efforts.

The presentation titled "Scoliosis x BackBend Detectives" showcases a patient-driven registry aiming to support individuals with scoliosis in the Philippines. It highlights the development and launch of a mobile app that facilitates patient registry, community support, and access to resources, with features including QR Code Patient ID and a comprehensive guidebook. The document also discusses the challenges of data collection and proposes solutions like good data governance and AI utilization for diagnostic screening, underlining the project's role in influencing policy and healthcare practices for scoliosis patients.

Ethics of AI

Dr. Marita V. Tolentino-Reyes talked about "ETHICS IN THE USE OF AI" emphasizing the importance of ethical in AI development and use, particularly in research and teaching contexts. She listed key ethical principles such as transparency, respect for persons, beneficence, non-maleficence, and justice, highlighting the need to address biases, ensure data integrity, and promote fairness. She recommended practical strategies for educators in guiding their students in the ethical use of AI-generated content, stressing the importance of original thought, critical thinking, and academic integrity.

Dr. Charibeth Cheng addressed how AI can contribute to Universal Health Coverage (UHC) by improving diagnostic accuracy, supporting clinical decision-making, enhancing public health monitoring, advancing drug discovery, and reducing healthcare costs. She elaborated on the challenges such as bias, privacy issues, autonomy, accountability, job displacement, and the lack of regulation. She highlighted the balance between AI's potential benefits and the ethical, privacy, and regulatory challenges that must be addressed to ensure its effective and equitable use in healthcare.

Dr. Alvin Marcelo introduced the Asia eHealth Information Network (AeHIN), detailing its mission to enhance the

interoperable digital health ecosystem for improved health in Asia through networking, resource sharing, and knowledge exchange. Established in 2011, AeHIN has grown to include over 1800 members across 72 countries, supporting capacity building in digital health through activities such as conferences, training, and online resources. The presentation highlights AeHIN's focus on Governance, Architecture, Program management, and Standards (GAPS) to address digital health challenges and promote effective health data management and interoperability. He then enumerated various AI initiatives in the ASEAN region.

Dr. Teehankee presented on "Ethical AI for Datathon" emphasizes the relevance of incorporating humanistic values into AI development to promote human flourishing, sustainable value creation, and productivity for the common good. He warned about the potential risks of unethical or unsafe AI, including harm to human well-being, democracy, and social values. He advocated for the use of AI in enhancing human capabilities, improving quality of life, and fostering ethical reflection and decision-making, highlighting the need for AI to support rather than undermine human needs and values.

Participation

There was nearly a total of 130 participants. Not including the organizers, guests, and sponsors, there was nearly a 50% split between female and male registrants from various locations nationwide. Six of the guests were of foreign nationalities.

Participant Feedback

Name: Gillian Nicole A. Jamias

Email: gillian_jamias@dlsu.edu.ph

I'm grateful for pushing myself to participate in last year's Datathon as it exposed me to a diverse crowd who are interested in drawing insights from health-related data. I felt inspired while listening to the different presentations. It made me want to be a better storyteller. Delivering the message across different perspectives and backgrounds may be daunting, but it is achievable with effort and teamwork. It was a moment to celebrate when our team placed after pitching.

Name: Amanda Glenda M. Bonife-Kiamko

Email: amanda@scoliosisphilippines.org

Winning the Datathon validated my mentoring approach and boosted my confidence in my ability to lead and inspire others. It provided a sense of accomplishment and recognition that motivated me to continue advocating for the scoliosis community. Additionally, it opened new networking opportunities and professional growth, connecting me with like-minded individuals and potential collaborators in the field. Overall, the Datathon experience was both rewarding and transformative. It enhanced my skills, expanded our network, and reaffirmed my

commitment to advocacy by following best practices and data-driven sources in shaping Universal Health Care (UHC).

Name: **Robbie Mondia**

Email: pr.robby.mondia@gmail.com

Joining and unexpectedly winning in the Philippine Datathon 2023 was a game-changer for me, especially in terms of developing potential tools for addressing a less commonly discussed issue: scoliosis. This experience significantly broadened my perspective on the importance of creating accessible technology, particularly for untapped datasets. In the Philippines, where healthcare can be prohibitively expensive for many, the need for affordable and innovative solutions is critical. Winning the event highlighted the possibilities and potential impact of our work, inspiring us to focus on building tools that can make a real difference in people's lives. By developing technology tailored to the specific needs and constraints of underserved communities, we can contribute to a more inclusive healthcare system where treatment and support for conditions like scoliosis are within reach for everyone.

Name: **Gabriel M. Galang**

Email: gabriel.galang@bulsu.edu.ph

Firstly, I thoroughly enjoyed the event. It was an incredible experience that opened a door for me into the world of data analytics and machine learning. The opportunity to learn and engage with these fields was invaluable. Winning was just a bonus; the real prize was the knowledge and experience I gained. I'm looking forward to participating in more events like this in the future.

Name: **Carlos Angelo Rayos**

Email: rayos.carlosangelo@gmail.com

I became more aware of the larger community working in data, not just in machine learning but even in data engineering and other fields. Meeting with different people from different fields who see the potential from data, their experiences, goals, and projects were very educational. I also got to talk with some people whose masters / degrees I was interested in pursuing, and they gave very inspiring insights.

Name: **Justine William T. Duran**

Email: jusduranmd@gmail.com

As a physician who is deeply interested in the application of health technology specifically artificial intelligence to real world problems, this was a never-before-seen opportunity to practice working with talented people from different backgrounds and actual data scientists, computer scientists, and artificial intelligence experts which was profoundly enriching. The problem faced by our partners who gave us access to their data, scoliosis Philippines, was one that was deeply fulfilling to work towards as this event gave them insights to their data that left them with actionable ideas that profoundly changed their view on the value of their data and how it can forward their cause of solving the financing problem of people with scoliosis. On

a personal level, this gave me the confidence to work on similar projects and is a vital steppingstone into the world of health informatics as I move forward with my career as a physician researcher by building relationships with vital collaborators through a common goal of leveraging health data for improving the healthcare outcomes of society.

Name: **Rizzi Meia Palma**

Email: rizzi1108@gmail.com

Joining last year's Datathon was a transformative experience for me, both professionally and personally. More than just learning about how AI can revolutionize healthcare, one of the most impactful aspects was the opportunity to meet new people who are deeply passionate about what they do. The event brought together a diverse group of individuals from various backgrounds, each with unique perspectives and expertise. This diversity was a catalyst for creativity and innovation, allowing us to approach problems from multiple angles and find more comprehensive solutions.

Being part of the winning group was not just a professional achievement but also a powerful validation of the importance of diversity in teamwork. It reinforced the idea that when people with different skills, experiences, and viewpoints come together, the potential for breakthrough ideas and effective problem-solving increases exponentially. I especially commend the organizers for including patients in last year's Datathon.

On a personal level, the Datathon was a confidence booster. It was a reminder that stepping out of my comfort zone and engaging with new challenges can lead to significant growth and unexpected rewards.

Overall, the Datathon was more than just a competition; it was a journey of learning, connection, and validation. It has had a lasting impact on how I approach my work in healthcare and AI, emphasizing the power of diversity and collaboration in driving progress and achieving excellence.

5. Discussion

There was consensus among participants that emerging technologies like artificial intelligence bring benefits and risks. The audience was particularly interested in how to address the problems with AI such as bias, forgetfulness, and hallucinations. A rigorous exchange occurred at the last session on Ethics of AI with panelists and experts in the audience challenging that ethics boards may have fallen short in protecting the vulnerable when it comes to AI research and implementation.

6. Conclusion and Future Directions

The datathon was able to bring local and international faculty and students to discuss state-of-the-art in data governance and management in the Philippines. While there was interest from all participants on a data-driven future,

they also expressed apprehensions on the security, privacy, and ethical use of data. There was consensus that more capacity-building opportunities should be made available to as many stakeholders as possible. More discourse on the aforementioned topics is recommended with particular emphasis on ethics and its positive and negative impact on the field of data analytics.

7. Acknowledgements

CARE PH for events management
De La Salle University for the venue
MIT Critical Data Team as mentors
UP Manila National Telehealth Center for coordination