

NBT (No-Boundary Thinking): Needed to Attend to Ethical Implications of Data and AI

Joan Peckham[†]
Computer Science & Statistics
University of Rhode Island
Kingston, RI, USA
jpeckham@uri.edu

Andy Perkins
Computer Science & Engineering
Mississippi State University
Starkville, MS, USA
perkins@cse.msstate.edu

Tayo Obafemi-Ajayi
Cooperative Engineering Program
Missouri State University
Springfield, MO, USA
tayoobafemiajayi@missouristate.edu

Xiuzhen Huang
Computer Science Department
Arkansas State University
Jonesboro, AR USA
xhuang@astate.edu

ABSTRACT

In this era of Big Data and AI, expertise in multiple aspects of data, computing, and the domains of application is needed. This calls for teams of experts with different training and perspectives. Because data analysis can have serious ethical implications, it is important that these teams are well and deeply integrated. No-Boundary Thinking (NBT) teams can provide support for team formation and maintenance, thereby attending to the many dimensions of the ethics of data and analysis. In this NBT workshop session, we discuss the ethical concerns that arise from the use of data and AI, and the implications for team building; and provide and brainstorm suggestions for ethical data enabled science and AI.

CCS CONCEPTS

• A.1: Introductory and Survey • K.6.1: Project and People Management

KEYWORDS

No-boundary Thinking, Convergence Research, Team Science, Ethical AI

ACM Reference format:

Joan Peckham and Andy Perkins. 2022. NBT (No-Boundary Thinking): Needed to Attend to Ethical Implications of Data and AI. In *Proceedings of 13th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics (ACM-BCB 2022)*. ACM, New York, NY, USA, 2 pages. <https://doi.org/10.1145/3535508.3545595>.

1 Ethical Implications of Data and AI

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

ACM-BCB '22, August, 2022, Chicago, Illinois USA
<https://doi.org/10.1145/3535508.3545595>

Big Data and Artificial Intelligence have recently become important tools for managing the tsunami of scientific data generated across the disciplines. To ethically address the biggest problems in this space, attention must be paid to multiple computing, data, and communication approaches. Also important is attention to the grounding of all data and analysis into the domain of consideration. All analytical approaches, including artificial intelligence, (defined as intelligent machine approaches to data analysis), cannot be carried out without attention to the collection, cleaning, organization, access and sampling of data. This usually begs for statistical, computational, data management, and communication expertise to define and plan solution strategies and to present results. The domain expert must be securely in the loop [1]. Fail any piece of this, the AI algorithm can be rendered useless, with “garbage in and garbage out”. This has ethical implications that must be considered because the results of data and AI analysis are frequently used to develop actions and policies that impact individuals, groups and society in general [2]. For example, the well-being of people may be impacted by biomedical decisions or treatments that are informed by these analyses [3,4,5,6]. In this workshop session, we will discuss the places in the ecology of AI and data analysis that need the attention of a team of experts with diverse perspectives, and how NBT [7] can address the ethical concerns that could arise if these teams do not function well.

1.1 The Ecology of Data and AI

When considering AI and data analysis, many focus on the analytical algorithms and procedures, and frequently forget about other important aspects such as the data and the communication of analysis results. We will consider and discuss these points in the conduct of data analysis that may lead to erroneous results, thereby threatening the safety of entities that are impacted by resulting decisions, actions, and policies. All forms of data analysis, including AI, require an array of activities, each calling for disciplinary expertise. No-boundary Thinking can help teams possessing diverse perspectives to work well together to perform and communicate robust and ethically safe data analysis.

1.1.1 Data Wrangling. Analysis and AI cannot work without data. Data must be collected, cleaned, organized, archived, accessed, and sampled correctly in order for AI analyses to be reliable. We will discuss the places in the ecology of data and AI where attention and expertise are needed.

1.1.2 Communicating Results. Once an analysis model is decided and the analysis is complete, the work is not done. Analytics experts must communicate to domain experts how the data has been managed, the procedures that are used to analyze the data, and how the results might be interpreted. Ideally domain, data analytics, and communication experts work together from problem definition to analytical results using best practices of No-boundary collaboration to achieve this. We will discuss where and how important communication modalities are needed for the whole team to understand both process and results. It is important, for example, that domain experts are able to assess if analytical results are consistent and explainable by the core scientific principles of the domain. Furthermore, all team members should be able to understand what has been done with the data, and how to interpret results with confidence.

1.2 The NBT Solution to Ethical Challenges

Team building can be difficult when participants are from the same discipline or sub-discipline, but needs special attention when participants are from different disciplines [9]. Challenges arise when participants have different perspectives, use different vocabulary, and have different cultural views on what constitutes good problems and solutions. Essential to NBT teams is proper formulation of the problem to be solved; a basic tenet is that the NBT team must come together to synthesize diverse perspectives to decide the problem before solutions can be considered. Given that participants come with different views on problem formulation and solution, it is important to consider a robust process for team formation and maintenance. This takes extra effort and time, but scholars studying teams of experts with diverse training have found that they are better positioned to be successful in solving even deep and difficult problems [8] especially if they have learned to work well with each other.

4 Discussion and Questions

In this workshop session, we identify the places in the data analysis and AI ecology where ethical concerns may arise. Then we will consider how NBT might be helpful. We will first briefly present some examples, and then we will engage with the workshop participants to consider and brainstorm other issues, and other approaches to address the ethics of data analysis and AI using NBT.

REFERENCES

- [1] Tayo Obafemi-Ajayi et. al. No-boundary thinking: a viable solution to ethical data-driven AI in precision medicine. *AI Ethics*. (2021). <https://doi.org/10.1007/s43681-021-00118-4>.
- [2] Cathy O'Neil. Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. *Broadway Books*, New York (2016). ISBN-13: 978-0553418811.
- [3] David Leslie et. al. Does “AI” stand for augmenting inequality in the era of Covid-19 healthcare? *BMJ* 372, 304 (2021). DOI: <https://doi.org/10.1136/bmj.n304>.
- [4] Jenna Wiens et. al. Do no harm: a roadmap for responsible machine learning for health care. *NatMed* 25(9), 1337–1340 (2019). DOI: <https://doi.org/10.1038/s41591-019-0548-6>.
- [5] Hannah E. Knight, et. al. Challenging racism in the use of health data. *Lancet Digital Health*. 3(3), 144-146 (2021). DOI: [https://doi.org/10.1016/S2589-7500\(21\)00019-4](https://doi.org/10.1016/S2589-7500(21)00019-4).
- [6] Ziad Obermeyer et. al. Dissecting racial bias in an algorithm used to manage the health of populations. *Science*. 366(6464), 447–453 (2019). DOI: <https://doi.org/10.1126/science.aax2342>.
- [7] Huang, X. et al. 2013. No-boundary thinking in bioinformatics research. *BioData Mining*. 6, 1 (2013). DOI: <https://doi.org/10.1186/1756-0381-6-19>.
- [8] Scott E. Page and Lu Hong. Groups of Diverse Problem Solvers Can Outperform Groups of High-Ability Problem Solvers. *Proceedings of the National Academy of Sciences*. 101, No. 46, 16385–16389 (2004). DOI: <https://doi.org/10.1073/pnas.04037231>
- [9] Joan Peckham and Andrew Geyer 2021. Making Big Data Models Work Right. *Big Data for Generals*. D.C. Ellis and M. Grzegorzewski, eds. The JSOU Press. 51–70. ISBN: 978-1-941715-60-4