This is a portion of the eBook at doi:10.7551/mitpress/14381.001.0001

1 Introduction

Fundamentally, this book is about learning and education. It is about the ways that we use data to shape possible futures of young people as well as the schools, informal learning environments, colleges, libraries, and educational games where these young people interact. It is about the processes by which students are sorted, labeled, categorized, and intervened upon using the bevy of data extracted and collected *from* individuals and groups, whether anonymously or identifiably. When, how, and with what biases are these data collected and utilized? What decisions must educational researchers make around data in an era of high-stakes assessment, surveillance, and rising inequities tied to race, class, gender, and other intersectional factors? And, to the best that we can surmise, how are these complex considerations around data changing in the rapidly evolving world of machine learning, artificial intelligence (AI), and emerging fields of educational data science? These are big questions and not ones typically taken up piecemeal by academic scholarship or in most classrooms with inquiring students. Inspired in part by Gutiérrez and her co-authors (Becker & Gutiérrez, 2022; Keifert et al., 2021), this book strives to understand data in ways that imagine more just futures while building on our existing technologies and knowledge. We can use a speculative

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2366401/c000400_9780262377645.pdf by guest on 10 August 2024

This is a portion of the eBook at doi:10.7551/mitpress/14381.001.0001

approach that imagines better ways of living alongside data, and this book argues that we need not wait to do better today.

Every day, school systems, educational researchers, and policymakers at every level of public educational systems make decisions that shape the lives of young people. These are typically called *data-driven decisions*—though whether the data does the driving or if the route was predetermined for particular purposes is a fraught question, largely informed by the political and profit-driven decisions that shape "college and career readiness" models of schooling in the US. Often, the idea of data collection and interpretation spawns visions of university researchers or policy wonks wading knee-deep through spreadsheet sludge to find submerged insight gems. However, for the most part, that is not where or with whom the data live.

Every classroom teacher, principal, and student shapes and is shaped by the educational data that they produce and interpret daily: grades and attendance records are generated; phones emit location data; software tracks student and teacher activity; and so on. Perhaps more importantly, each one of these individuals could likely expound on the myriad ways in which data fails today's educational system. Data fails school policy decisions. Data fails classroom instructional approaches. Most importantly, data fails the needs of students. Very unlike the popular representation of data science as mainly the purview of tiny teenage ersatz Bill Gateses, educational data is both deeply personal and used by almost everyone up and down administrative chains (Halverson et al., 2007).

As new teachers, we were trained—both formally in a teacher education program and implicitly through how colleagues talked about and categorized their students—to *see* our students through their data. Our teacher education programs encourage new teachers to develop a relational trust with students, but they also ask us to quantify the *who*, *why*, and *what* of all student classroom activity. As much as we see students as curious and questioning individuals, young

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2366401/c000400_9780262377645.pdf by guest on 10 August 2024

This is a portion of the eBook at doi:10.7551/mitpress/14381.001.0001

people—through the data they create—are implicit quandaries to be diagnosed and *solved*. Schools typically then provide teachers with the tools to treat students as problems ready to be solved. This distancing often prevents teachers from seeing their students as individual identities and bodies participating in the centuries-old experiment of US public education (e.g., D'Ignazio & Klein, 2020).

Throughout this book, we approach the intersections of data, education, play, and social justice through speculative, practical, empirical lens, and critical lenses. All such divisions are necessarily imperfect, but we start this book with an exploration and application of new critical frameworks, then move to chapters in which we use those frameworks to explore how the intersections arise and function in different contexts (e.g., museums, schools, video games). Between those chapters, we offer a series of interludes. These are practical applications of our theoretical perspectives with working program code. If you don't feel very comfortable reading or writing program code, don't worry (neither does Antero)! We wrote these to be as accessible as possible, with comments added liberally throughout to allow even avowed nonprogrammers to look at the construction and utilization of data under the hood. The interludes are intended to spark creative hope; with any luck, they will encourage readers to try out, build, and play with practical ideas in ways that enable us to live more just lives.

Schooled Data in Practice

In the high school where Antero taught, at the beginning of each year, every English class would filter through the school's computer lab in which students would complete a short reading assessment test. Promising "powerful assessment tools" in its marketing materials, the Renaissance Learning Star Reading test is a computer-adaptive diagnostic literacy test. As described on the Renaissance Learning

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2366401/c000400_9780262377645.pdf by guest on 10 August 2024

This is a portion of the eBook at doi:10.7551/mitpress/14381.001.0001

website (2020), "Star Reading transforms assessment data into action steps for educators, giving teachers helpful insights and tools to strengthen classroom instruction." In practice, the implementation of this test meant students were asked to complete the assessment over the course of several minutes before shuffling out of the class, when the next set of students would be ushered in for quantification.

By the end of the day, teachers were given a printout with all the scores for their students. The Star Reading test offers multiple kinds of data for interpreting student reading levels, and Antero's school focused on two columns of data for guiding instructional decisions. Following student names and identification numbers, a column of data elucidated the "grade equivalent" (GE) reading ability for each student. These numbers ranged from 0.0 to 12.9+, denoting the approximate grade level at which students were reading. Notionally, students in an 11th-grade English class would (theoretically) read at an average of an 11th-grade level; in fact, that was not the case. Rather, most of the students in Antero's classes read at a level "between 4th and 5th grade" (according to the results of the Star test), with some students reading much higher than that but many labeled as much lower.

In a somewhat strange twist of the foundations of sociocultural theory, an additional column was labeled "ZPD" (after "Zone of Proximal Development" from Vygotsky, 1978) and indicated the Star test's suggested zone of proximal development for each student based on their assessment results. This range was expected to solidify the band of books students could select and read at the school. These paired data points—a student's GE and ZPD—functioned as the literacy talismans that guided instruction and engagement at the school. In practice, these ZPD scores recast how literacy was seen at the school in literal ways. Every book was affixed with a label designating the band of reading levels that could access it, making clear which books were for whom. The data from this beginning-of-the-year test recast

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2366401/c000400_9780262377645.pdf by guest on 10 August 2024

This is a portion of the eBook at doi:10.7551/mitpress/14381.001.0001

the landscape of student learning and the instructional materials available to them and—like donning a pair of glasses that may pull the world into focus for some and blur it for others—the Star Reading test reframed how students and teachers were *seeing* school and the books available to students.

The data-created visions of students have social, relational, and affective consequences. For the substantial number of students Antero taught who would achieve second-, third-, and fourth-grade reading levels, the low scores were demoralizing. These tests were expected to function as harsh wake-up calls trumpeting the inadequacies of student abilities. Somehow the students were confronted with the responsibility of raising these scores, even though these students were usually already a decade into the public education system that yielded these results in the first place.

Books with a 2.5 or a 3.0 grade level as part of their range literally look different than books geared at readers of higher ranges. You are familiar with second-grade books: they are picture books; they have colorful images; they have large fonts; and they often are *bigger* than the pocket-sized novels that adults read. They look childish. Telling a teenager to read these books is an instructional moment but likely not for the lessons we are hoping to teach. (The lesson, instead, was a negative "this place/life/work is not for you" or simply, perhaps, "despair.")

The decades of high-stakes assessment kick-started by the US National Commission on Excellence in Education's (1983) *A Nation at Risk* report created the chain reaction that yielded the Star Reading test. The Star Reading test illustrates angles on truth in data, interpretation, and responsibility. For example, it is true that the test reflected a clear evaluation of students not reading "well" based on performance data. From the gaze of its creators and from the ways the school used it, the test was not false; it was consistent, repeatable, and clear.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2366401/c000400_9780262377645.pdf by guest on 10 August 2024

This is a portion of the eBook at doi:10.7551/mitpress/14381.001.0001

While we can point to policies and educational movements that might have led to the demoralizing feeling of mandating that a 17-year-old read a floppy picture book, this book suggests that data optimized for systems of capitalism will nearly always yield processes that sort and reshape student identities and bodies in similar ways. The drastic ways in which, on a Tuesday morning, sitting in front of a computer for 15 minutes, a student's experience of schooling is recast by a singular assessment, illustrates one of myriad data decisions that students and teachers confront daily. Would the student have been better off without the information? Is there an ethical way to use these data that does not result in a riptide of despair?

Data, its intended use, and the unintended consequences of its collection shape nearly every aspect of how students experience and participate in the world. From the predictive policing algorithms that disproportionately shape the day-to-day, negative interactions that Black and brown youth experience outside of schools to health modification scores that may affect the likelihood of an individual to receive medical care or organ transplants, data-informed decisions extend across everything that we do. Scholars such as Benjamin (2019) suggest that this is but one of the many ways in which market surveillance ideology can manifest racist and sexist histories of oppression within the "normalcy" of daily life.

The ways in which schools use our data only work because we are implicitly convinced that all reasonable options require us to freely proffer all personal information property to private corporations; this is hardly true worldwide. The exceedingly normal feeling today of giving all your data to a Silicon Valley company is part of a grand misdirection. Knowing that from early surveys of colonized land to the present day, approaches to classroom assessments have caused harm, this book builds from an obvious and simple point: *it doesn't have to be like this*.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2366401/c000400_9780262377645.pdf by guest on 10 August 2024

This is a portion of the eBook at doi:10.7551/mitpress/14381.001.0001

Data for Speculative Design and Freedom

That said, this is *not* a book which demonizes data analysis. In this book we apply the same kind of critical urgency through which we look at the state of the world. This may indeed be a bleak time when it comes to educational equity, but losing hope is worse than having none. Education is unique in the US in that it is near universal, guided and staffed by people who (whether successful or not) intend good things, and an institution that can genuinely improve lives. In the end, questions for data in education might have come directly from Tolkien (2007).

- Is data science inherently just a Palantir? Does it, by giving you information, stare back at you and, in doing so, doom you?
- Is data science a ring of power, tempting you to use the power of the enemy against that enemy, and, in doing so, becoming the enemy yourself?
- Is it Gollum, an indispensable guide who will help save us, but who can never be trusted?

If the consequences of data collection are demoralizing, ethically wrong-footed, or otherwise lead to pedagogical quagmires, your educational data and analysis are bad. This hardly needs to be said, but the concept of investigation from perceived reality (in other words, data analysis) is a tool that is also leveraged quite often for good and for right (in schools and outside of them). We are, for example, very happy to have vaccines that use data from real human trials; in that vein, we publish papers that use data to expose how to make education or schools better. Rather than focusing on prosecuting the legacies of harm reaped by poor data collection, most of the following chapters look at specific contexts and examples of current data collection. Exploring nuanced approaches to collecting, utilizing, and building from the insights of educational data, we seek to help guide research approaches that preserve and affirm the dignities of participants.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2366401/c000400_9780262377645.pdf by guest on 10 August 2024

This is a portion of the eBook at doi:10.7551/mitpress/14381.001.0001

However, before we turn to these contemporary examples in the following chapters, we cast our eye toward data imaginaries, speculative data analysis, and even some real, runnable program code.

Fundamentally, our methodological designs and analytic considerations are informed by dreamt futures. Building on our personal and professional affiliations (in part as self-proclaimed geeks), we find inspiration in some representations of mindful educational data use as depicted in science fiction. We particularly recognize the role of *speculation* as an intentional agentic act informing research, literature, design, and other aspects of life; the imagination of shows like *Star Trek* and the words and worlds of writers like Octavia Butler and Ursula Le Guin mine inequitable pasts to help carve visions of new possibilities shaping data design today. Illustrations in popular media—and promising research we will turn to throughout this book—point in alternative directions. We know we can make the lives of kids better *within* the systems of schooling, but we can also make kids' lives better through *changing* these systems. These are the possibilities of educational data that we might wield as speculative-focused educational researchers, educators, and designers. Until we can imagine just futures that use our data to help us grow, we exist only within otherwise fixed ideologies.

So, who gets to imagine possible educational futures? If it is only the dominant powers of established capital, then they set the space of possibility. To whom can we look for speculative guidance?

Python and Playgrounds

The Python in this book is all runnable online in Google CoLab—something like Google Docs for Python—with no changes from the reader. We have included shortened links here: github.com/mberland/interludes

Indeed, we suggest that you attempt to run all the code and follow the suggestions that are in the code comments, making changes

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2366401/c000400_9780262377645.pdf by guest on 10 August 2024

This is a portion of the eBook at doi:10.7551/mitpress/14381.001.0001

throughout. That said, you might be asking: What is Python? What are code comments? Why should I run code? What if I do not know how to code? Will this book teach me how to code?

This book will not teach you how to code, we are sorry to say. If you were hoping to pick this book up, read it cover to cover, and then immediately apply for a job as a software engineer at Meta, you are a very hopeful person, and we wish you luck. That said, we hope that by modifying the Python from this book, all nonprogramming readers can make progress toward situating and demystifying code. Though it may look intimidating at first glance, Matthew has written all the code in this book to be generally understandable to most readers and to help illustrate how code intersects with the data it runs.

A short lesson will be unlikely to help massively—honestly, we would just click the links and select “Run all” from the menu—but we can give some context. Python is a computer programming language used by scientists, computer scientists, and programmers worldwide; it is one of the most popular, and it is widely used both by top professionals and in first-year computer programming classes. It is considered to be relatively easy to read—that is, there is not a lot of specialized syntax that would look arcane to a novice. It is designed such that many simple functions can be read out loud to gather their meaning. For example, print(“hello”) means to print the text contained within the quotes; in this case, it would print out the word *hello*. The code # print(“hello”) would do something very different in Python. It would do literally nothing. Why? That # at the beginning of the line marks the line as a “comment”: a note to human people reading the code rather than running code itself. We have peppered all the examples with comments; reading them may help the code make more (or less) sense.

The metaphor for Google CoLab is the “cell,” sort of like a cell in a Microsoft Excel spreadsheet. The cells can contain runnable code, Markdown (human-readable web language), and data, as seen in figure 1.1.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2366401/c000400_9780262377645.pdf by guest on 10 August 2024

This is a portion of the eBook at doi:10.7551/mitpress/14381.001.0001

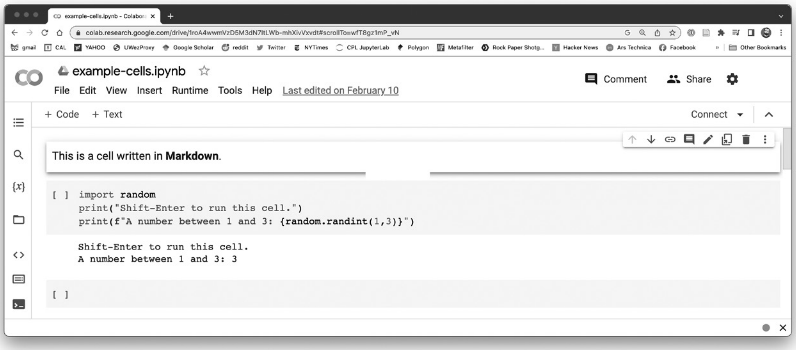

Figure 1.1
A cell in Google CoLab.

If you run a cell with Python code in it (typically <Shift-Enter> or <Shift-Return>), it will show the result. There is a Runtime menu at the top of the page under which there exists the option to "Run all"; if you select that, it will run all cells and show you the outputs as they happen. There is a fantastic feature called "Playground mode" which you can get to by selecting "Open in Playground mode" in the File menu. Playground mode does not save changes, so you can change code indiscriminately and with reckless abandon; if it stops working or you just want to start over, just close and reopen the Playground.

This book is much more fun—and significantly louder—if you play with the interludes. It is worth your time, we promise.

The Left Hand of Data

Ursula Le Guin's *Left Hand of Darkness* is, in some sense, about the fundamental disconnect between individualism and the functioning of society. The *mindspeech* (telepathy in which it is impossible to

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2366401/c000400_9780262377645.pdf by guest on 10 August 2024

This is a portion of the eBook at doi:10.7551/mitpress/14381.001.0001

intentionally lie) casually, maybe offhandedly, topples authoritarianism. However, even in the absence of dissembling, Genly cannot overcome bias in perception of gender and individuality. The book exists as an expression of the idea that there is no justice that is not inherently intersectional, and—crucially for this book—that even "perfect data" reflects the biases of the data collector and analyst. Perhaps ironically (or perhaps cynically), Le Guin's technocratic society cannot even comprehend collectivist mores. They don't fit into the schema! The data don't fit the spreadsheet.

This book is inspired by Le Guin's book. Our title is a respectful, nonpossessive reference to the book as well as a (perhaps oblique) reference to *Star Trek*'s Lt. Commander Data, who contains a button hidden in a fingernail of his left hand that serves as a sort of fail-safe for changing the narrative (per Bowman, 1990). Data has an evil twin brother, Lore, who also contains a hidden button on his left hand. The story arc between these two brothers posits the idea that "lore" (i.e., narrative) and "data" (i.e., evidence) are mirrors of each other, but that they must "forgive" each other to be whole. The relationship between what seems like hard evidence (but is, instead, a person with perspective) and what seems like unmoored bias (but turns out to be quite calculating) is precisely one of the dialectics that this book foregrounds.

Data Imaginaries in Urgent Times

We write this at the crest of one of many waves that seems to be plunging the US further into the icy shock of economic depression, social upheaval, and political discombobulation. We started writing this book just before the COVID-19 pandemic. However, the past few years have pulled into focus the urgency for the heuristics at the center of this book. Ongoing upheaval at local and global levels is generating new capitalist demands for schools to measure, sort, and judge

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2366401/c000400_9780262377645.pdf by guest on 10 August 2024

This is a portion of the eBook at doi:10.7551/mitpress/14381.001.0001

the validity of forms of knowledge and intelligences. These systems are and will continue to be sexist, racist, transphobic, homophobic, ableist, and engaged in other discourses of oppression.

Though this book does not focus solely on race, its investments in racial equity are never far from the surface. There are fundamental questions that the construction of race in the US poses to every educational endeavor: Are you making things better for people who have been systematically oppressed? With whom are you surfacing the development of your work, and how do historically marginalized individuals shape and co-construct this process? Of course, there is value in helping every child, but every new generation of researchers seems to find that when they do not keep antiracism as a core goal (and sometimes even when they do), they will likely just make the situation worse. Throughout the book, we attempt to foreground both how mechanisms of oppression function and possible modes of resistance to those mechanisms. It is not enough, but it is what we can do.

There are some sets of arguments—not unfounded—which claim that any use of quantitative data is dehumanizing and will, in some way, ipso facto result in racist outcomes. This implies that data science possesses a unique dehumanizing power; however, this argument misses the fact that data collection is just one of many mechanisms of power, which seek to dehumanize and control populations (e.g., Foucault, 1977). In the absence of spreadsheets, for example, power operates through religion, capital, and other sociopolitical institutions. We can design different futures through reappropriating existing tools. In fact, the current educational moment demands new ways of thinking about the future of education. In the spring of 2020, there was an unanticipated, widespread pivot to using distance-based teaching. Improbably, this educational response to a pandemic was even more bleak than the futures envisioned in dystopian novels, films, and videogames. For many schools, what became the essential task—the absolute core—was distributing content into the void

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2366401/c000400_9780262377645.pdf by guest on 10 August 2024

This is a portion of the eBook at doi:10.7551/mitpress/14381.001.0001

rather than supporting children. Many districts could not account for the location or physical safety of their students. The models of instructional delivery taken up by schools provided little agency, little care, little interaction, little collaboration, and no equity. During months of heightened anxiety and social distancing, schools could have played prominent roles as bulwarks for community healing and growth. Of course, such possibilities were hardly considered and this lack of action or imagination falls on educational leaders, politicians, researchers, and educators writ large. This lack of responsiveness to a world beset by trauma was only exacerbated in the months following nationwide sheltering in place. The onset of the COVID-19 pandemic was rapidly followed by an overdue collective response to systemic antiblackness and racism within the US. In both health crises, young people in the US were left alone: coping, fearing, and mourning a country rattled to its foundations. This is not a valid future of school.

Feeling the walls of the unlit room of the present moment, educators, parents, politicians, and researchers are looking for new ways to *do* schooling. Recent attempts have not been particularly promising. Studies of massive open online courses (MOOCs) suggest that they—especially ones that are not strictly synchronous—do not fulfill the dream of an equitable future or a great equalizer; in fact, they may be undermining it (Breslow et al., 2013; Reich & Ruipérez-Valiente, 2019). In the context of MOOCs and in current forms of distance learning, it may be worse for many kids to attend these schools than to simply not go at all—even considering exclusively school-based academics.

The chapters that follow map out some potential digital futures that have equitable implications. We write this hoping that educational researchers might use the ideas here as propulsive momentum toward better possibilities obscured in the smog of toxic policies. Though we would like to offer a blueprint for liberation, that task falls outside of the purview of this book. Rather, by taking seriously the expectations

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2366401/c000400_9780262377645.pdf by guest on 10 August 2024

This is a portion of the eBook at doi:10.7551/mitpress/14381.001.0001

about, with, and through data, our speculative approach seeks to redefine how we work within rigid systems and how to creatively resist or undermine them when necessary.

Recognizing the urgency of these present times and the decades of inaction that they have been met by, these chapters are written to support the needs and growing concerns of the angry parent, the angry educational researcher, the angry teacher, the angry principal. But they go beyond that, too. We are writing, in this moment of myriad threats leveled at the most vulnerable in the US, to meet the vision and courage of the *hopeful* parent, the *hopeful* educational researcher, the *hopeful* teacher, the *hopeful* principal. If you are a researcher who works toward a commitment that school should build from a foundation of equitable learning futures, that school should support and nurture every child, and that school is not simply a penitentiary from which people distribute content, this book is for you.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2366401/c000400_9780262377645.pdf by guest on 10 August 2024

This is a section of doi:10.7551/mitpress/14381.001.0001

The Left Hand of Data

Designing Education Data for Justice

By: Matthew Berland, Antero Garcia

Citation:

The Left Hand of Data: Designing Education Data for Justice
By: Matthew Berland, Antero Garcia
DOI: 10.7551/mitpress/14381.001.0001
ISBN (electronic): 9780262377645
Publisher: The MIT Press
Published: 2024

The open access edition of this book was made possible by generous funding and support from MIT Press Direct to Open

The MIT Press

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2366401/c000400_9780262377645.pdf by guest on 10 August 2024

MIT Press Direct

© 2024 Massachusetts Institute of Technology

This work is subject to a Creative Commons CC-BY-NC-ND license.

This license applies only to the work in full and not to any components included with permission. Subject to such license, all rights are reserved. No part of this book may be used to train artificial intelligence systems without permission in writing from the MIT Press.

The MIT Press would like to thank the anonymous peer reviewers who provided comments on drafts of this book. The generous work of academic experts is essential for establishing the authority and quality of our publications. We acknowledge with gratitude the contributions of these otherwise uncredited readers.

This book was set in Stone Serif and Stone Sans by Westchester Publishing Services.

Library of Congress Cataloging-in-Publication Data

Names: Berland, Matthew, author. | Garcia, Antero, author.
Title: The left hand of data : designing education data for justice / Matthew Berland and Antero Garcia.
Description: Cambridge, Massachusetts : The MIT Press, [2024] | Includes bibliographical references and index.
Identifiers: LCCN 2023030088 (print) | LCCN 2023030089 (ebook) | ISBN 9780262547529 (paperback) | ISBN 9780262377652 (epub) | ISBN 9780262377645 (pdf)
Subjects: LCSH: Education—Data processing. | Education—Research. | Educational evaluation.
Classification: LCC LB1028.43 .B45 2024 (print) | LCC LB1028.43 (ebook) | DDC 370.285—dc23/eng/20230718
LC record available at https://lccn.loc.gov/2023030088
LC ebook record available at https://lccn.loc.gov/2023030089

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2366401/c000400_9780262377645.pdf by guest on 10 August 2024