VOLUME 8
2024-2025

ROLL DATA

CONNOLLY ALEXANDER INSTITUTE FOR DATA SCIENCE

2024-2025

DATA

ZINE

WHERE NUMBERS MEET STORIES

The views and opinions contained herein by the various authors do not necessarily reflect the official views, opinions, or policy of Tulane University administrators, staff, or faculty. All material contained herein are the views and opinions of the authors themselves.

CONTRIBUTORS

EDITOR

Jacquelyne Thoni Howard, Ph.D.
Professor of Practice,
Associate Director of Student Engagement

DIGITAL RESEARCH INTERNS (DRI)

Sam DeMarinis & Ananya Anand (Product Developers)

Katie Bogdanow
Daniel Fox
Liam Guest
Julia Miller
Kailen Mitchell
Meghan Nagia
Mika Nijhawan
Zoe Oboler
Alexandra Zeff

DATA AMBASSADOR COUNCIL (DAC)

Ifeoma Osakwe (Product Developer)

Eghosa Asemota
Samuel Johnson
Nina Moore
Sara Vannoni
Davis Green
Cameron McLaren
Favour Olushola
Kai Williams

PEER MENTORS

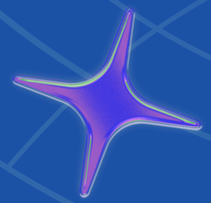
Adel Rahman
Angel Baralles
Lily Markus
Noah Fisher
Joeseph Raphael
Megan Cassidy

ZINE EDITORIAL TEAM

Zoe Oboler
Ananya Anand
Sam DeMarinis
Ifeoma Osakwe

ZINE COVER

Zoe Oboler



PROJECTS & PARTNERS

PROJECT NAME	PROJECT PARTNER
Beautiful Sisterhood of Books	Dr. Kate Adams & Dr. Jacquelyne Howard
The Frances Joseph Gaudet Legacy Project	Dr. Rosanne Adderly and Nell Bolton
Feminist Pedagogy for Teaching Online	Liv Newman, Dr. Jacquelyne Howard & Dr. Claire Daniel
Civil War Monuments Database	Kris Plunket, PhD Student
The Center for Sport	Tulane Center for Sport
Murphy Institute	JR And and Hussain Hadah
Kallion	Mallory Monaco Caterine
Mary Amelia Center for Women's Health Equity Research	Mellissa Evans
Data Literacy Lab	CAIDS Data Lab

MESSAGE FROM THE EDITOR

Our student programs at CAIDS empower graduate and undergraduate students to help make data literacy skills and work accessible for all of the Tulane community. Our programs range from paid internships, workshops, data consultations, community building, and grant funding.

Each program supports students in advancing their data literacy skills and join collaborative communities made up of students, faculty, and staff working across the expansive field of data.

CAIDS students work with data in the following interdisciplinary areas:

- Research Support: They engage with the data research of faculty, community, and students through paid internships and research assistantships, curriculum, and project sponsorship.
- Educational Services: They seek help from peer tutors and consultants with DATA coursework and a statistical help desk.
- Career Readiness: They build data skills and data portfolios, allowing space to build confidence and experience.
- Outreach: They network and work alongside other data-minded students, using interdisciplinary and multidisciplinary approaches.
- Operations: They contribute to building an accessible culture around data at CAIDS by leading focus groups, communicating about CAIDS, and supporting events.

While doing this work, they gain essential data and science communication skills. This zine serves as a portfolio of the important contributions all of our data scholars made this past year.

Within the Data Lab, we offer three student programs designed to cultivate data-minded communities on campus through co-curricular activities that make data skills accessible, relevant, applicable, and ethically driven for undergraduate and graduate students. Within the Data Research Internship Program, the Data Peer Mentor Program, and the Data Ambassador Council, we welcome students working with data at all skill levels and accept all majors and fields.

Through our programs, we engage with students to:

- Build mutually supportive and interdisciplinary data-minded communities.
- Apply data ethics frameworks to our work as a tenet of data literacy.
- Make data literacy skills and work accessible and relevant for all students.
- Become users, researchers, makers, and critics of data knowledge.

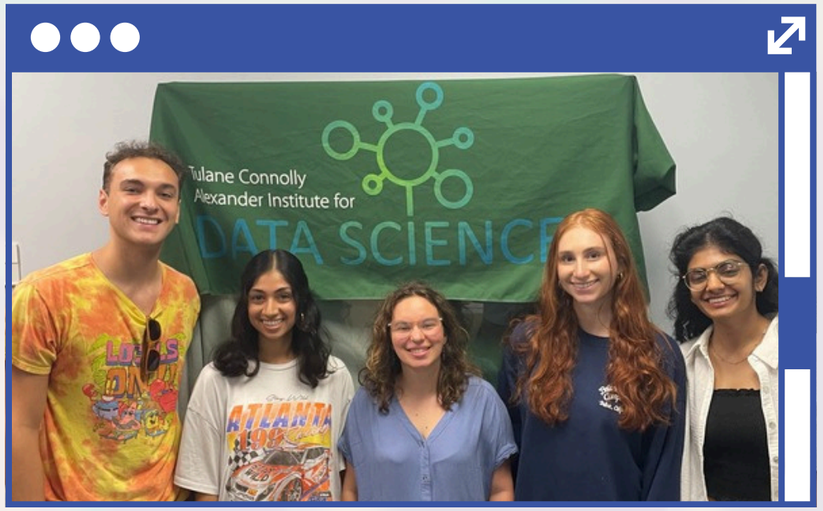
I am constantly impressed by our students' abilities to learn new skills, adapt to changing technologies, and reflect on the importance of their contributions, including week-to-week tasks and larger goals. I am proud that, as they have embarked on learning new skills, they have taken the time to build kind and caring communities around data that revolve around uplifting and mutually supporting others. I am confident that our students will apply their skillsets and continue to make ethically-minded "data decisions" as they work beyond CAIDS.

Sincerely,
Dr. Jacquelyne Thoni Howard
Professor of Practice of Data
Associate Director of Student Engagement

TABLE OF CONTENTS

MEET THE LAB

Sam Johnson, Davis Green, Oma Osakwe, Favour Olushola,
Kai Williams, Sara Vannoni, and Cameron McLaren



Sam DeMarinis, Meghan Nagia, Julia Miller, Alexandra Zeff,
and Ananya Anand

Liam Guest, Mika Nijhawan, Ananya Anand, Katie Bogdanow,
Zoe Oboler, Sam DeMarinis, and Daniel Fox

Q&A with the Product Developers

SD AA OO

Sam DeMarinis, Ananya Anand, Oma Osakwe

Explain your role as a PD.

AA As PDs in the DRI team, Sam and I lead by building collaborative relationships between project partners and the interns on the team. We break down the project partner's requirements into actionable tasks that are delegated to the interns. To ensure alignment between project partner expectations and intern capabilities we facilitate weekly SCRUM meetings with the interns and regular check-ins with the project partners.

OO As PD for the DAC, I lead CAIDS promotional efforts to improve data literacy and access on campus. Specifically, I coordinate production of materials (including newsletters, flyers, social media and web content) and organize student-facing events (including tabling and data trivia nights).

What does it mean to you to be a PD?

SD As a DRI intern, I had witnessed multiple PDs come before me, and the role seemed to be a bit intimidating. I knew that I was capable of the role, but naturally, I still felt anxious. Like any new role, it takes some time to feel truly comfortable. But thanks to my co-PD Ananya and boss Dr. Howard, I was surprised by how quickly I began to feel confident in my ability to perform in this role. As a PD, we are not considered any 'higher up' than any other intern on the team. Instead, we are tasked with being experts on our team's projects, exceptional communicators, and attentive listeners. To me, being a PD means being a proud and passionate member of the DRI team, as well as an advocate for our team's values as they relate to the projects we take on.

What skills did you improve in this role?

OO Working as a PD this year taught me a lot about effective team management. Because all interns simultaneously navigate academic demands, I found it important to uphold a culture of care and communication to support anyone struggling with lab commitments.

AA As a PD, I work with the interns on the team and the project partners. This has helped me improve my communication, time management, and team building skills.

What have been some challenges you faced in the role, how did you over come it?/what did you learn from it?

OO Because many of the DAC's tasks involve creative expression, it was sometimes difficult to strike a balance between encouraging liberty in a task and ensuring that products reflect the team's goals. Incorporating peer review during tasks helped address this issue.

SD Being a PD, while fulfilling and enriching, can also be challenging. I quickly learned how important it is to make and stick to a schedule, especially as it relates to project deadlines. This ensures that we are keeping a solid pace and accomplishing what we set out to do. Additionally, I learned that documentation/archiving is always beneficial, even for the most seemingly-mundane of tasks. More often than not, you will not remember that thing you "didn't need to write down".

What has been your favorite part of your role?

SD My favorite project is This Beautiful Sisterhood of Books, which our lab has been working on for seven years.

AA Working on this zine has been my favorite part! I have loved the opportunity to showcase all the amazing work we do at the lab.

OO Organizing the Data Trivia Night with the DAC was incredibly fun! I enjoyed meeting other people interested in data science and the game-based discussion made the topic very accessible.

Q&A with the Technical Lead

Zoe Oboler

Explain your role as a technical lead.

As the technical lead, it is my job to understand the different technologies we use in the lab, document them, and help educate other steidents on how to use them

What have been some challenges you faced in the role, how did you over come it?/what did you learn from it?

Often, I am not familiar with the technology we are using, so learning on the job is definitely a big challenge. Beyond that, it is also difficult making sure that whatever solution I find works for the clients and is sustaibable for the lab. In this role I have learned about all the moving pieces that go into making a project successfull.

What has been your favorite part of your role?

I love how much I get to work dirrectly with other team members. Although it is already a very collaborative workplace, I really enjoy having the opportunity to work to solve problems with my coworkers.

2024 - 2025 DRI INTERNS

Sam DeMarinis (he/him)

Data Research Intern

Product Developer

Sam DeMarinis is a senior at Tulane from Morris Plains, New Jersey. He is majoring in Mathematics & Computer Science and minoring in Psychology & Italian. In addition to working as Product Developer at CAIDS, he is an executive board member for Tulane's computer science club, Cookies & Code. He is passionate about music, language, and using his technical skills to positively impact others. For his senior capstone, Sam is working with the team's Technical Lead, Zoe Oboler, to develop a Unity-based language-learning game meant to teach children a second language through an immersive and interactive experience.

Ananya Anand

Data Research Intern

Product Developer

Ananya Anand is a senior from Alpharetta, GA who is graduating with a Bachelor's in Economics and Computer Science with a minor in Strategy, Leadership, and Analytics. In addition to being a Product Developer at CAIDS, Ananya is a Resident Advisor on camous and is invovled with the Tulane University Data Collective. For her senior capstone, Ananya is working under Dr. Brian Summa's MAGIC-SCAN Moonshot Project. As a part of her project, she is working with her team to develop a program that would automate the pipeline for generating high quality 3D scans of Tumors.

Katie Bogdanow (she/her)

Data Research Intern

Katie is a sophomore. Katie Bogdanow is a sophomore from Dallas, Texas studying Mathematics and Sociology with a minor in strategy, leadership, and analytics (SLAM.) Aside from her work as a Data Research Intern at CAIDS, she is involved in Girls Who Code, Tulane Jewish Leaders, and Sigma Delta Tau In her free time, Katie enjoys watching films, trying new coffee shops, and going on walks in the Garden District.

Daniel Fox (he/him/his)

Data Research Intern

Daniel Fox is a graduate student from New Jersey pursuing a Master's Degree in Business Analytics. He graduated from Tulane with a Bachelor's of Science in Economics in May 2024. In his free time he enjoys playing poker, playing pool, watching basketball and football, and going on bike rides.

Liam Guest (he/him/his)

Data Research Intern

Liam Guest is an MSPH candidate from Chicago studying Biostatistics and Data Science. He graduated with his Bachelor's degree in Public Health in May of 2024 from Tulane. Outside of his involvement with CAIDS, Liam is a teaching assistant for a Biostatistics course and enjoys hot yoga.

Mika Nijhawan (pronouns)

Data Research Intern

Mika Nijhawan is a senior studying Economics from Boulder, Colorado. Beyond working at CAIDS, she is a Newcomb Scholar and endowed Harriet-Bobo Scholar, Vice President of the Multi-Cultural Council, a South Association of Tulane Executive Board Member, and the Editor-in-Chief and Creative Director of NOSTRA, a feminist literary magazine. Outside of the classroom, you can find her in the sculpture studio or reading in Audubon!

Julia Miller (she/her)

Data Research Intern

Julia Miller is a junior from Baton Rouge, LA who is studying Anthropology and Environmental Studies with a Geographic Information Systems Certificate. Elsewhere on campus, Julia is the president of Gardening Club, vice-president of Tulane Anthropology Club, and a member of the Feminist Alliance of Students at Tulane (FAST). She is passionate about the environment, social justice, libraries, and research ethics.

In her personal life, Julia loves cooking, reading, sewing, and spending time with friends.

Zoe Oboler (he/him)

Data Research Intern

Zoe Oboler is a senior studying Design and Computer Science, from Washington D.C. In addition to working at Caids, Zoe is a member of Theta Tau, Tulane's Engineering fraternity. Zoe has an interest in web design, data visualization and Ui/Ux design. He has spent the year working on a language learning app with Sam DeMarinis. In his personal life, Zoe enjoys travel, art, and sewing.

Kailen Mitchell (pronouns)

Data Research Intern

Kailen is a senior studying Math and Computer Science with minors in IT and English. She particularly enjoys application development and technical problem solving as part of her work. Outside of work and school, she loves taking spin classes, creative writing, and spending time in nature. he's always felt passionate about feminism in tech and has deeply enjoyed the welcoming community of data minded individuals that the DRI team has fostered. She hopes to bring the essential data skills she's learned through the DRI program to her next endeavors.

Alexandra Zeff

Data Research Intern

Alexandra Zeff is a sophomore at Tulane University from Phoenix, Arizona. She is pursuing a double major in Political Science and Communications. Aside from working at CAIDS, Alexandra is a member of honor societies Pi Sigma Alpha and Alpha Lambda Delta. Her professional experience includes roles as a Communications Intern for the City of Phoenix and a Social Media Marketing Intern for Olami International, where she honed her skills in public relations, social media management, and digital design.

Meghan Nagia (pronouns)

Data Research Intern

Meghan Nagia is a junior from Atlanta, GA studying exercise science with a minor in strategy, leadership, and marketing (SLAM) on the pre-med track. Aide from being a data research intern at CAIDS, she is also a student sports medicine assistant for the Tulane women's volleyball team, a member of Phi Mu, and on the e-board of the Indian Association of Tulane University (IATU) and the Kinesiology Club, as well as being involved in other organizations. Meghan also enjoys spending time with her family and friends, music, being active, and exploring new coffee shops!

Digital Research Internship Program Projects

This paid internship provides students with tangible experience in technology and feminist leadership while working on the digital scholarship projects of Tulane faculty in the humanities, social sciences, and sciences. Technical areas include: digital media, digital archiving, web design, database development, coding, design thinking, UX/UI, and audio/visual editing.

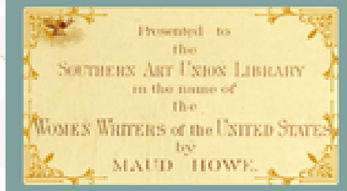
BEAUTIFUL SISTERHOOD OF BOOKS

RECOVERING A LOST LITERARY LEGACY

BSB aims to recover Maud Howe's library of books, journals, newspapers, and sheet music donated by publishers and collected by "lady delegates" from around the US in the 1800s. After the 1884 New Orleans World's Fair, the physical collection was scattered and lost. This collaborative project is an archive-in-process that is meant to recover all of these materials.

SKILLS

1. Website Development
2. Database Design & Management
3. Data Validation & Storytelling
4. Metadata Standardization



Maud's Sisterhood in Library History

The 1884 New Orleans World's Fair

New Research on the Sisterhood

OUR LATEST DEVELOPMENTS

Created custom post types highlighting Black Women Writers and Student Writings.

Implemented specialized taxonomies for categorizing notable authors and student contributors.

Completed a thorough audit of the "Black Women and the Fair" section of the website.

Added dedicated pages for new scholarly essays.

SOFTWARE USED

WordPress, Excel, & Canva

PROJECT PARTNERS

Dr. Jacquelyne Howard
&
Kate Adams

THE FRANCES JOSEPH GAUDET LEGACY PROJECT

Co-Chairs of the Gaudet Legacy Project:
Dr. Rosanne Adderly and Ms. Nell Bolton

The Frances Gaudet Legacy Project aims to share widely the example of her extraordinary life in order to guide and inspire similar work in our present times. We are pursuing this aim through extensive research, focusing particularly on her work in education and prison reform in the greater New Orleans area. We also seek opportunities to bring her story to wider audiences.

Highlighted Projects

- FRANCES GAUDET'S LEGACY TOLD THROUGH NEWSPAPERS
- FRANCES GAUDET'S IMPACT ON YOUTH IN THE PAST AND TODAY
- HIGHLIGHTING FRANCES JOSEPH GAUDET'S EXCLUDED NARRATIVE THROUGH NEWSPAPER
- FRANCIS JOSEPH GAUDET'S ASTOUNDING LEGACY IN EDUCATION AND PHILANTHROPY

Skills & Technology Used

Skills: Website Development, Digital Archiving, & Primary Source Cross-Referencing

Technology Used: Omeka

Gaudet was an uncommon woman of her time—a political and social leader, a missionary at home, a proponent of race consciousness—who was able to bring together forces from diverse segments of society to effect social change in post- Reconstruction New Orleans.

Feminist Pedagogy for Teaching Online

Building a feminist teaching guide

Project Partners: Liv Newman, Dr. Jacquelyne Howard, Dr.Clare Daniel

About

The Feminist Pedagogy for Teaching Online is a digital guide that provides resources for using feminist pedagogy in the classroom.

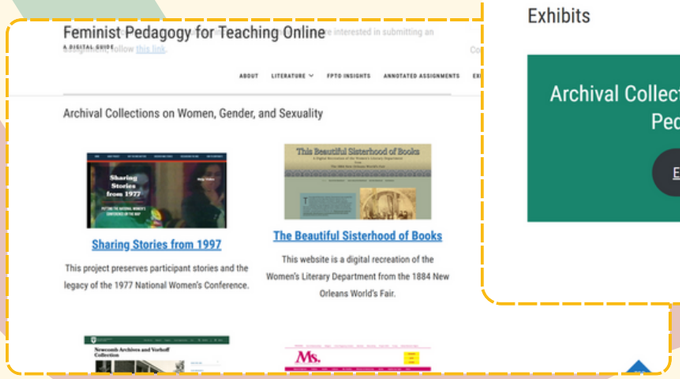
Skills

- Website Development
- Database Development
- Data Cleaning
- Graphic Design
- Audio/Visual Editing

Accomplishments

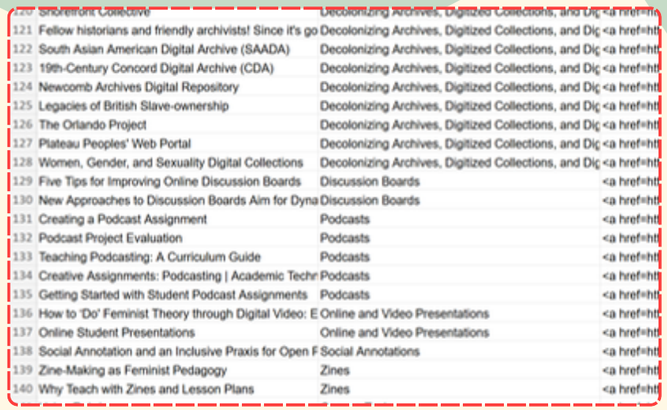
- Created a suggestion form
- Made promotional materials and a video for the website
- Reordered the assignment table by year
- Added sources to the database including typifying them and writing abstracts
- Created documentation on how to update the database
- Made sure database items were properly categorized and descriptions were accurate
- Identified which sources were missing information
- Started to buildout an exhibit page on the website for easier website navigation

Check out the website:



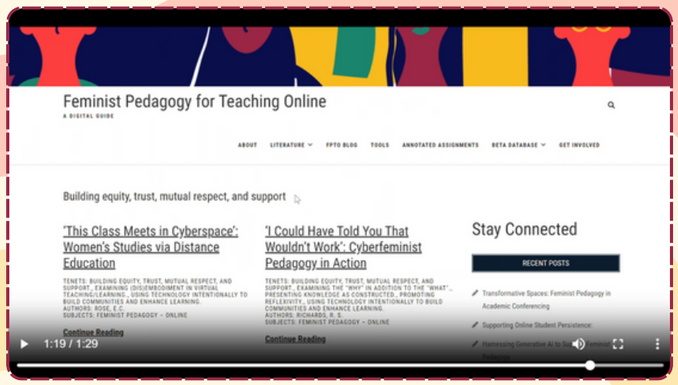
New exhibits page

Updated database



121	Fellow historians and friendly archivists! Since it's go	Decolonizing Archives, Digitized Collections, and Dig <a href=htt
122	South Asian American Digital Archive (SAADA)	Decolonizing Archives, Digitized Collections, and Dig <a href=htt
123	19th-Century Concord Digital Archive (CDA)	Decolonizing Archives, Digitized Collections, and Dig <a href=htt
124	Newcomb Archives Digital Repository	Decolonizing Archives, Digitized Collections, and Dig <a href=htt
125	Legacies of British Slave-ownership	Decolonizing Archives, Digitized Collections, and Dig <a href=htt
126	The Orlando Project	Decolonizing Archives, Digitized Collections, and Dig <a href=htt
127	Plateau Peoples' Web Portal	Decolonizing Archives, Digitized Collections, and Dig <a href=htt
128	Women, Gender, and Sexuality Digital Collections	Decolonizing Archives, Digitized Collections, and Dig <a href=htt
129	Five Tips for Improving Online Discussion Boards	Discussion Boards <a href=htt
130	New Approaches to Discussion Boards Aim for Dyna	Discussion Boards <a href=htt
131	Creating a Podcast Assignment	Podcasts <a href=htt
132	Podcast Project Evaluation	Podcasts <a href=htt
133	Teaching Podcasting: A Curriculum Guide	Podcasts <a href=htt
134	Creative Assignments: Podcasting | Academic Techr	Podcasts <a href=htt
135	Getting Started with Student Podcast Assignments	Podcasts <a href=htt
136	How to 'Do' Feminist Theory through Digital Video: E	Online and Video Presentations <a href=htt
137	Online Student Presentations	Online and Video Presentations <a href=htt
138	Social Annotation and an Inclusive Praxis for Open F	Social Annotations <a href=htt
139	Zine-Making as Feminist Pedagogy	Zines <a href=htt
140	Why Teach with Zines and Lesson Plans	Zines <a href=htt

New video



New promotional materials

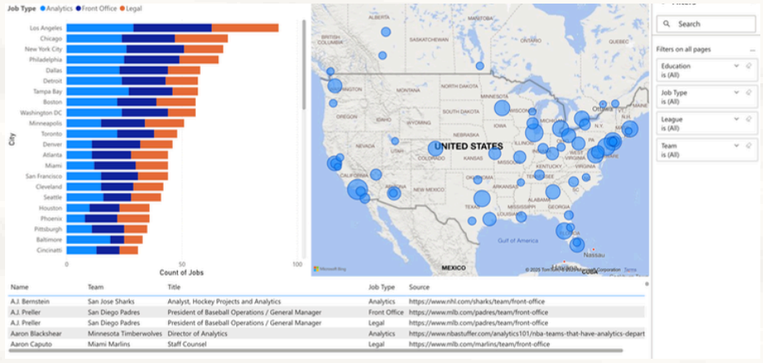
CENTER FOR SPORT

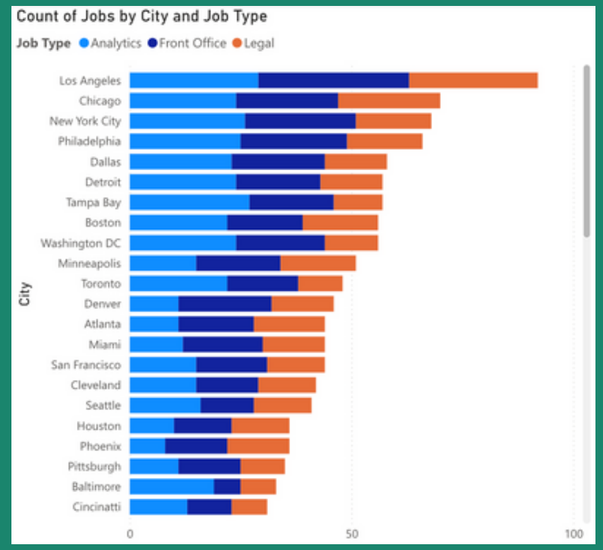
MAPPING THE EXECUTIVE LANDSCAPE OF SPORT

Project Partner: Tulane Center for Sport

This database contains information on executives working within the sports industry. Students compiled this information over time by examining different sports and executives who work for them. The goal of this project is to allow sports law students access to the resource so they can see what their options are for whom to connect with and what sports are available in a given region.

Power BI Dashboard with database information:

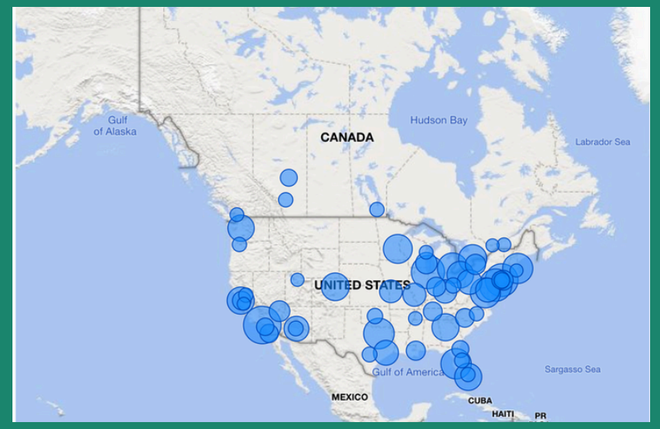




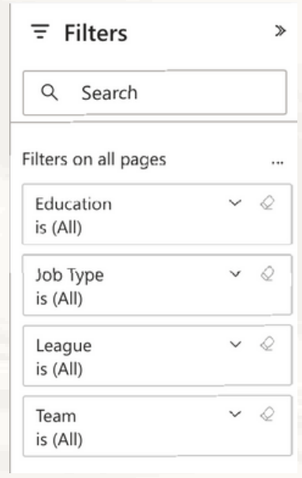
Graphs and maps are used to display geographical information. A stacked bar chart (left) provides information on jobs by regions, and an interactive map (below) allows users to visually search by region.

Skills:

- Data visualization
- Project design
- Product development



Interactive Dashboard Video



There a number of filters available to users, based on data taxonomies. This makes the dashboard more customizable and useful to a number of people.

CIVIL WAR MONUMENTS DATABASE

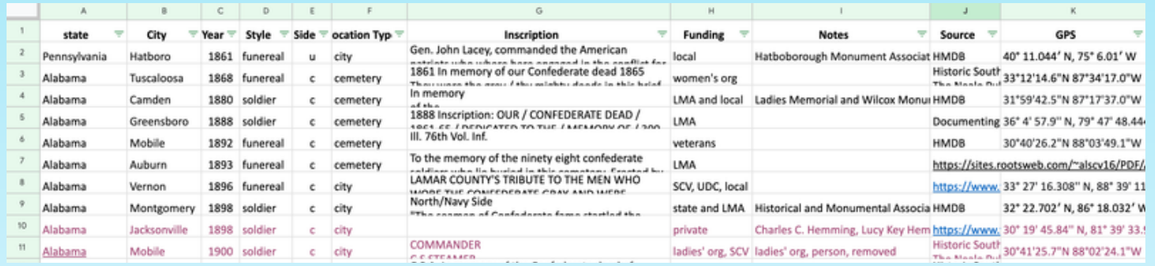
GEOSPATIAL AND CONTEXTUAL DATA COLLECTION

OBJECTIVE:

To create a comprehensive database of civil war monuments and abolition monuments in the United States as a tool for education, research, and communities with context to understand their local memorials. The database will be visualized as a map that shows the distribution of monuments across time and geographical space. Hopefully, this catalogue will be helpful in highlighting how individual memorials are part of a larger, more pernicious picture.

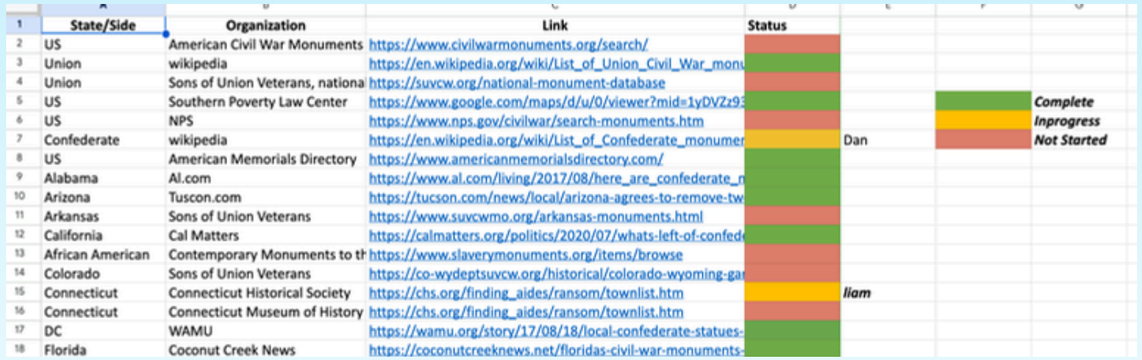
OUR WORK

- Cleaning data and filling in gaps of missing monuments from other databases
- Cross-referencing the compiled database with specific documents
- Adding monuments that honor abolitionists and those that led up to the Civil War



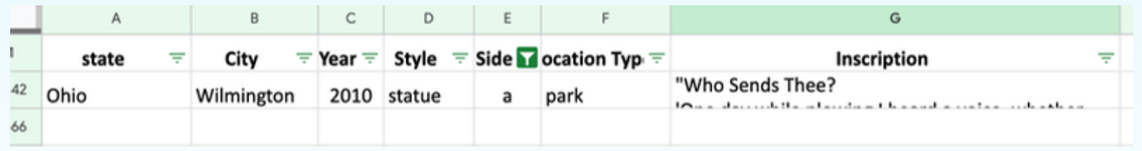
	A	B	C	D	E	F	G	H	I	J	K
1	state	City	Year	Style	Side	ocation Typ	Inscription	Funding	Notes	Source	GPS
2	Pennsylvania	Hatboro	1861	funereal	u	city	Gen. John Lacey, commanded the American	local	Hatboborough Monument Associat	HMDB	40° 11.044′ N, 75° 6.01′ W
3	Alabama	Tuscaloosa	1868	funereal	c	cemetery	1861 In memory of our Confederate dead 1865	women's org		Historic South	33°12'14.6"N 87°34'17.0"W
4	Alabama	Camden	1880	soldier	c	cemetery	In memory	LMA and local	Ladies Memorial and Wilcox Monu	HMDB	31°59'42.5"N 87°17'37.0"W
5	Alabama	Greensboro	1888	soldier	c	cemetery	1888 Inscription: OUR / CONFEDERATE DEAD /	LMA		Documenting	36° 4′ 57.9″ N, 79° 47′ 48.44
6	Alabama	Mobile	1892	funereal	c	cemetery	Ill. 76th Vol. Inf.	veterans		HMDB	30°40'26.2"N 88°03'49.1"W
7	Alabama	Auburn	1893	funereal	c	cemetery	To the memory of the ninety eight confederate	LMA		https://sites.rootsweb.com/~alscv16/PDF/	
8	Alabama	Vernon	1896	funereal	c	city	LAMAR COUNTY'S TRIBUTE TO THE MEN WHO	SCV, UDC, local		https://www.	33° 27′ 16.308″ N, 88° 39′ 11
9	Alabama	Montgomery	1898	soldier	c	city	North/Navy Side	state and LMA	Historical and Monumental Associa	HMDB	32° 22.702′ N, 86° 18.032′ W
10	Alabama	Jacksonville	1898	soldier	c	city		private	Charles C. Hemming, Lucy Key Hem	https://www.	30° 19′ 45.84″ N, 81° 39′ 33.5
11	Alabama	Mobile	1900	soldier	c	city	COMMANDER	ladies' org, SCV	ladies' org, person, removed	Historic South	30°41'25.7"N 88°02'24.1"W

Monuments Database on Google Sheets #1



	A	B	C	D	E	F	G
1	State/Side	Organization	Link	Status			
2	US	American Civil War Monuments	https://www.civilwarmonuments.org/search/				
3	Union	wikipedia	https://en.wikipedia.org/wiki/List_of_Union_Civil_War_mon				
4	Union	Sons of Union Veterans, nationa	https://suvcw.org/national-monument-database				
5	US	Southern Poverty Law Center	https://www.google.com/maps/d/u/0/viewer?mid=1yDVZz9				Complete
6	US	NPS	https://www.nps.gov/civilwar/search-monuments.htm				Inprogress
7	Confederate	wikipedia	https://en.wikipedia.org/wiki/List_of_Confederate_monumer		Dan		Not Started
8	US	American Memorials Directory	https://www.americanmemorialsdirectory.com/				
9	Alabama	Al.com	https://www.al.com/living/2017/08/here_are_confederate_n				
10	Arizona	Tuscon.com	https://tucson.com/news/local/arizona-agrees-to-remove-tw				
11	Arkansas	Sons of Union Veterans	https://www.suvcwmo.org/arkansas-monuments.html				
12	California	Cal Matters	https://calmatters.org/politics/2020/07/whats-left-of-confed				
13	African American	Contemporary Monuments to th	https://www.slaverymonuments.org/items/browse				
14	Colorado	Sons of Union Veterans	https://co-wydeptsuvcw.org/historical/colorado-wyoming-ga				
15	Connecticut	Connecticut Historical Society	https://chs.org/finding_aides/ransom/townlist.htm		liam		
16	Connecticut	Connecticut Museum of History	https://chs.org/finding_aides/ransom/townlist.htm				
17	DC	WAMU	https://wamu.org/story/17/08/18/local-confederate-statues-				
18	Florida	Coconut Creek News	https://coconutcreeknews.net/floridas-civil-war-monuments-				

Checklist of External Documents/Databases to Compare Database



	A	B	C	D	E	F	G
	state	City	Year	Style	Side	ocation Typ	Inscription
42	Ohio	Wilmington	2010	statue	a	park	"Who Sends Thee?
66							

Example of Abolitionist Monument

Skills:

- Google Sheets
- Microsoft Excel
- Data entry
- Data cleaning

Project Partner: Kris Plunket, PhD Student

THE MURPHY INSTITUTE

Diagnosing Physician Shortages, Prescribing Policy Solutions

Project Partners: JR Ang and Hussain Hadah

BACKGROUND

This project focuses on the extent to which restrictions on residency positions and physician training opportunities have contributed to escalating healthcare costs. The Research aims to clarify how artificially constraining the physician workforce influenced the U.S. healthcare market, helped drive up expenditures, and widened the care gap between different groups of people.

SKILLS

Data Encoding & Data Base Management

TECHNOLOGY USED

Microsoft Excel

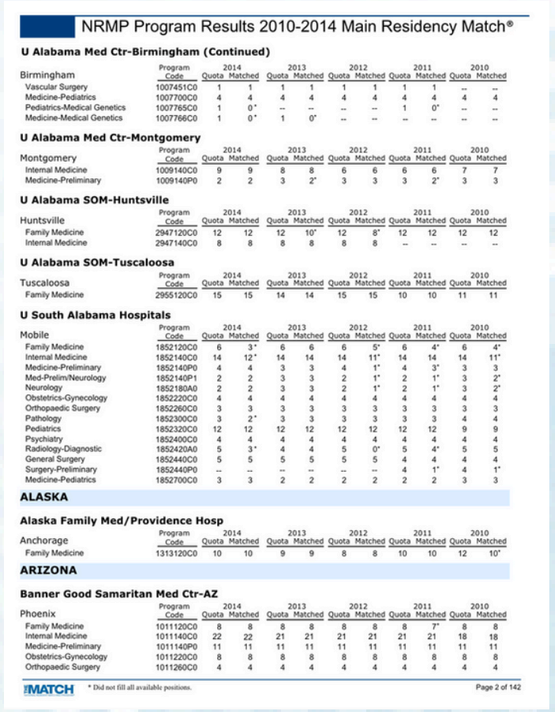
IMPORTANCE

Currently the data is inaccessible. Transforming the data to a database allows it to be used for research that will have further policy implications

OUR WORK

Replicate the data from the pdf to the excel sheet in a manner that can be used for research

DATA BEFORE:



NRMP Program Results 2010-2014 Main Residency Match®

U Alabama Med Ctr-Birmingham (Continued)

Birmingham	Program Code	2014 Quota	2014 Matched	2013 Quota	2013 Matched	2012 Quota	2012 Matched	2011 Quota	2011 Matched	2010 Quota	2010 Matched
Vascular Surgery	1007451C0	1	1	1	1	1	1	1	1	--	--
Medicine-Pediatrics	1007700C0	4	4	4	4	4	4	4	4	4	4
Pediatrics-Medical Genetics	1007765C0	1	0*	--	--	--	--	1	0*	--	--
Medicine-Medical Genetics	1007766C0	1	0*	1	0*	--	--	--	--	--	--

U Alabama Med Ctr-Montgomery

Montgomery	Program Code	2014 Quota	2014 Matched	2013 Quota	2013 Matched	2012 Quota	2012 Matched	2011 Quota	2011 Matched	2010 Quota	2010 Matched
Internal Medicine	1009140C0	9	9	8	8	6	6	6	6	7	7
Medicine-Preliminary	1009140P0	2	2	3	2*	3	3	3	2*	3	3

U Alabama SOM-Huntsville

Huntsville	Program Code	2014 Quota	2014 Matched	2013 Quota	2013 Matched	2012 Quota	2012 Matched	2011 Quota	2011 Matched	2010 Quota	2010 Matched
Family Medicine	2947120C0	12	12	12	10*	12	8*	12	12	12	12
Internal Medicine	2947140C0	8	8	8	8	8	8	--	--	--	--

U Alabama SOM-Tuscaloosa

Tuscaloosa	Program Code	2014 Quota	2014 Matched	2013 Quota	2013 Matched	2012 Quota	2012 Matched	2011 Quota	2011 Matched	2010 Quota	2010 Matched
Family Medicine	2955120C0	15	15	14	14	15	15	10	10	11	11

U South Alabama Hospitals

Mobile	Program Code	2014 Quota	2014 Matched	2013 Quota	2013 Matched	2012 Quota	2012 Matched	2011 Quota	2011 Matched	2010 Quota	2010 Matched
Family Medicine	1852120C0	6	3*	6	6	6	5*	6	4*	6	4*
Internal Medicine	1852140C0	14	12*	14	14	14	11*	14	14	14	11*
Medicine-Preliminary	1852140P0	4	4	3	3	4	1*	4	3*	3	3
Med-Prelim/Neurology	1852140P1	2	2	3	3	2	1*	2	1*	3	2*
Neurology	1852180A0	2	2	3	3	2	1*	2	1*	3	2*
Obstetrics-Gynecology	1852220C0	4	4	4	4	4	4	4	4	4	4
Orthopaedic Surgery	1852260C0	3	3	3	3	3	3	3	3	3	3
Pathology	1852300C0	3	2*	3	3	3	3	3	3	4	4
Pediatrics	1852320C0	12	12	12	12	12	12	12	12	9	9
Psychiatry	1852400C0	4	4	4	4	4	4	4	4	4	4
Radiology-Diagnostic	1852420A0	5	3*	4	4	5	0*	5	4*	5	5
General Surgery	1852440C0	5	5	5	5	5	5	4	4	4	4
Surgery-Preliminary	1852440P0	--	--	--	--	--	--	4	1*	4	1*
Medicine-Pediatrics	1852700C0	3	3	2	2	2	2	2	2	3	3

ALASKA

Alaska Family Med/Providence Hosp

Anchorage	Program Code	2014 Quota	2014 Matched	2013 Quota	2013 Matched	2012 Quota	2012 Matched	2011 Quota	2011 Matched	2010 Quota	2010 Matched
Family Medicine	1313120C0	10	10	9	9	8	8	10	10	12	10*

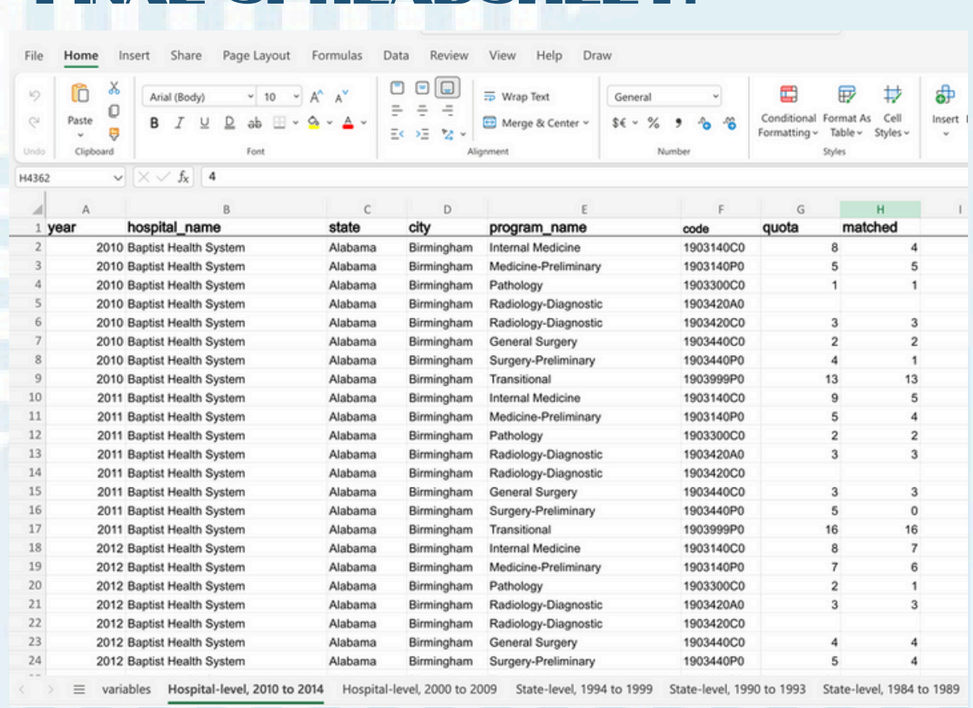
ARIZONA

Banner Good Samaritan Med Ctr-AZ

Phoenix	Program Code	2014 Quota	2014 Matched	2013 Quota	2013 Matched	2012 Quota	2012 Matched	2011 Quota	2011 Matched	2010 Quota	2010 Matched
Family Medicine	1011120C0	8	8	8	8	8	8	8	7*	8	8
Internal Medicine	1011140C0	22	22	21	21	21	21	21	21	18	18
Medicine-Preliminary	1011140P0	11	11	11	11	11	11	11	11	11	11
Obstetrics-Gynecology	1011220C0	8	8	8	8	8	8	8	8	8	8
Orthopaedic Surgery	1011260C0	4	4	4	4	4	4	4	4	4	4

MATCH * Did not fill all available positions. Page 2 of 142

FINAL SPREADSHEET:



	A	B	C	D	E	F	G	H
1	year	hospital_name	state	city	program_name	code	quota	matched
2	2010	Baptist Health System	Alabama	Birmingham	Internal Medicine	1903140C0	8	4
3	2010	Baptist Health System	Alabama	Birmingham	Medicine-Preliminary	1903140P0	5	5
4	2010	Baptist Health System	Alabama	Birmingham	Pathology	1903300C0	1	1
5	2010	Baptist Health System	Alabama	Birmingham	Radiology-Diagnostic	1903420A0		
6	2010	Baptist Health System	Alabama	Birmingham	Radiology-Diagnostic	1903420C0	3	3
7	2010	Baptist Health System	Alabama	Birmingham	General Surgery	1903440C0	2	2
8	2010	Baptist Health System	Alabama	Birmingham	Surgery-Preliminary	1903440P0	4	1
9	2010	Baptist Health System	Alabama	Birmingham	Transitional	1903999P0	13	13
10	2011	Baptist Health System	Alabama	Birmingham	Internal Medicine	1903140C0	9	5
11	2011	Baptist Health System	Alabama	Birmingham	Medicine-Preliminary	1903140P0	5	4
12	2011	Baptist Health System	Alabama	Birmingham	Pathology	1903300C0	2	2
13	2011	Baptist Health System	Alabama	Birmingham	Radiology-Diagnostic	1903420A0	3	3
14	2011	Baptist Health System	Alabama	Birmingham	Radiology-Diagnostic	1903420C0		
15	2011	Baptist Health System	Alabama	Birmingham	General Surgery	1903440C0	3	3
16	2011	Baptist Health System	Alabama	Birmingham	Surgery-Preliminary	1903440P0	5	0
17	2011	Baptist Health System	Alabama	Birmingham	Transitional	1903999P0	16	16
18	2012	Baptist Health System	Alabama	Birmingham	Internal Medicine	1903140C0	8	7
19	2012	Baptist Health System	Alabama	Birmingham	Medicine-Preliminary	1903140P0	7	6
20	2012	Baptist Health System	Alabama	Birmingham	Pathology	1903300C0	2	1
21	2012	Baptist Health System	Alabama	Birmingham	Radiology-Diagnostic	1903420A0	3	3
22	2012	Baptist Health System	Alabama	Birmingham	Radiology-Diagnostic	1903420C0		
23	2012	Baptist Health System	Alabama	Birmingham	General Surgery	1903440C0	4	4
24	2012	Baptist Health System	Alabama	Birmingham	Surgery-Preliminary	1903440P0	5	4

variables | Hospital-level, 2010 to 2014 | Hospital-level, 2000 to 2009 | State-level, 1994 to 1999 | State-level, 1990 to 1993 | State-level, 1984 to 1989

KALLION

Designing for human-focused leadership

Project partner: MALLORY MONACO CATERINE

Objective

KALLION IS A NON-PROFIT STARTED OVER 5 YEARS AGO. THEY PROVIDE LEADERSHIP DEVELOPMENT PROGRAMMING FOR THOSE WHO STUDY/WORK WITH HUMANITIES. KALLION'S TEAM IS WORKING TO INCORPORATE LEADERSHIP DEVELOPMENT INTO PRE-EXISTING EDUCATION PROGRAMMING.

- RESTORED THE WEBSITE AFTER PROGRAM CHANGES REVERTED IT BACK TO AN OLDER FORM
- CREATED A MOCK UP OF A SEARCH BAR FUNCTION
- AUDITED THE WEBSITE PER PARTNER INSTRUCTIONS
- ADDED ALT TEXT AND DROP SHADOW TO IMAGES
- USED SLIDERS ON A PAGE AND ADDED APPROPRIATE SLIDERS

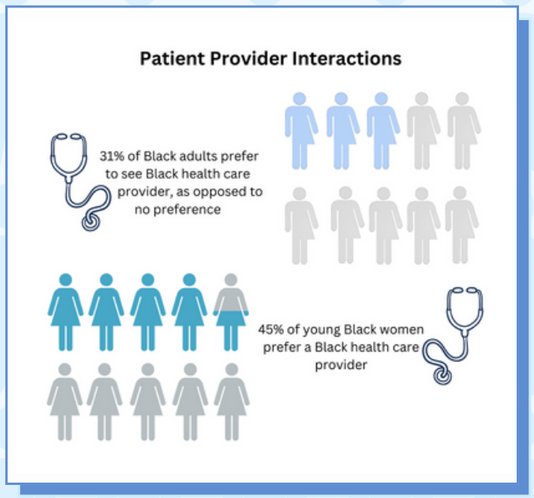
SKILLS:

- WEBSITE DEVELOPMENT
- WEBSITE DESIGN

Technology Used:

MARY AMELIA CENTER FOR WOMEN'S HEALTH EQUITY RESEARCH

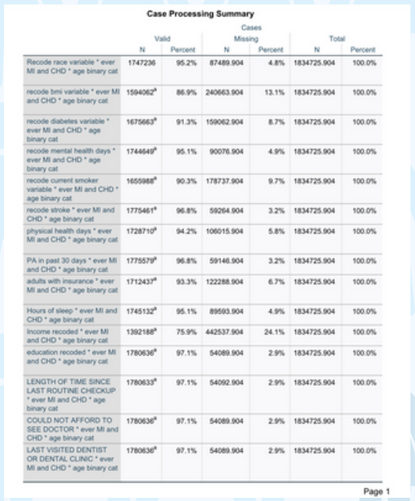
Project partner: Mellissa Evans



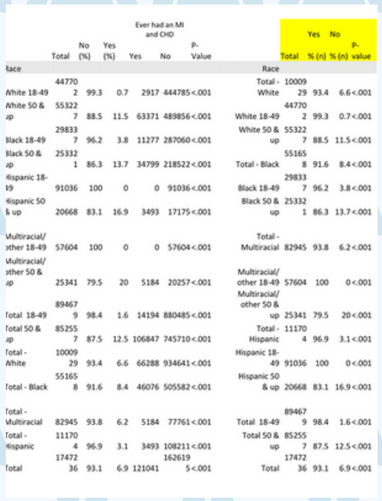
Background

The Mary Amelia Center for Women's Health Equity Research take an innovative, multi-level approach to their mission of conducting interdisciplinary research that identifies and disrupts barriers to knowledge, opportunity, and health for women and their communities.

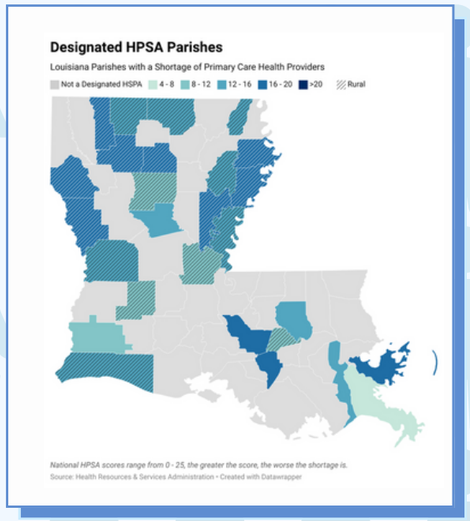
The DRI team works with MAC on their reports regarding health disparities in Louisiana. The team assists in cleaning data for these reports, along with creating data visualizations.

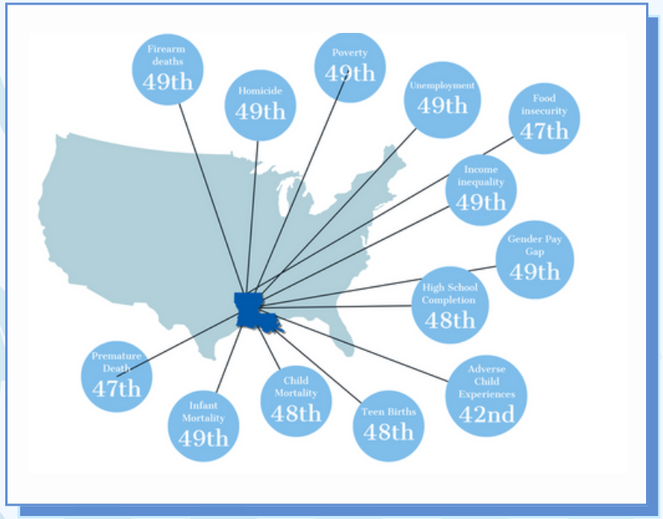


Original Data

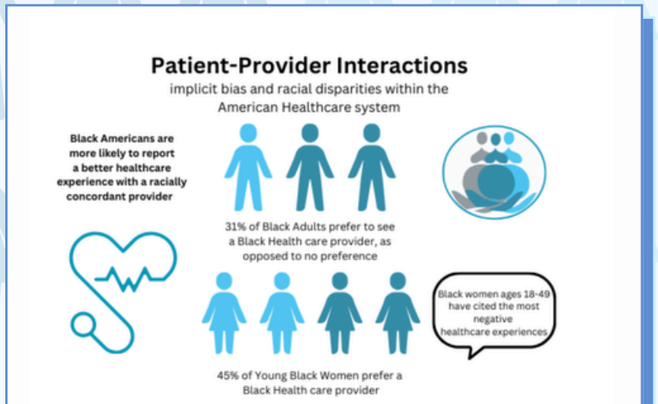


Cleaned Data





DATA VISUALIZATION EXAMPLES



Skills

- Graphic Design
- Data Visualization
- Data cleaning
- Data encoding

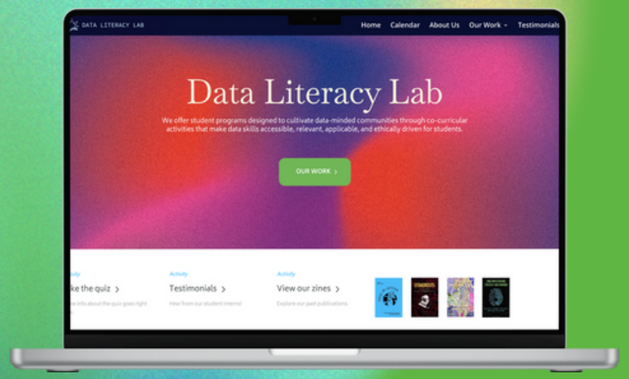
Technology Used

- Canva
- Flourish
- Microsoft Excel
- Microsoft Word

Data Literacy Lab

Project Partner

Dr. Jacquelyne Thoni Howard



The Data Literacy Lab website is a homegrown project. The lab consists of three programs - Digital Research Internship, Data Ambassador Council, and the Peer Mentors. The website highlights the work of the programs under the lab.

Skills

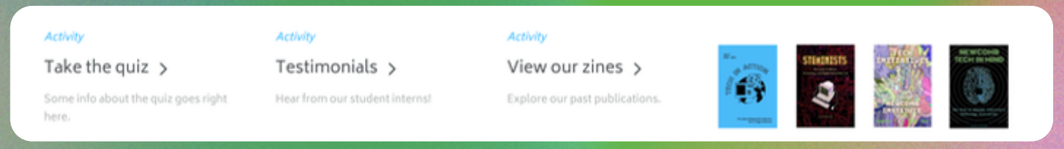
Website Development

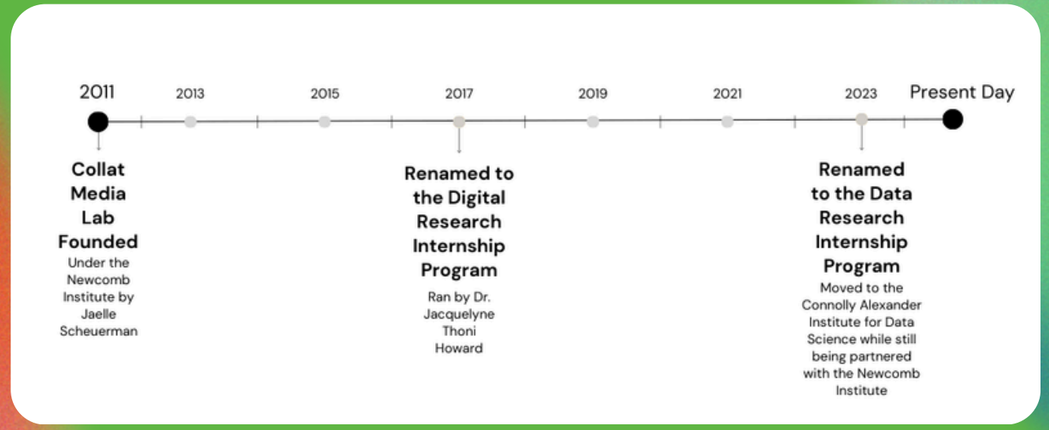
Website Design

Graphic Design

Our Work

- Designed the layout for the website and built it from scratch
- Built pages to house a data literacy quiz, testimonials from lab members, and past Zine editions
- Embedded a calendar to keep up with lab events
- Built an about us page sharing the lab'story, scholarly values, and members





Technology Used

- Wordpress
- Canva
- Flourish

Digital Research Internship Program Articles

Three Brilliant Tools, One Important Choice: ChatGPT, Claude, or Gemini?

By: Liam Guest

In a rapidely evolving industry, three tech leaders have gained a relative lead. To look at the three analogously, in the context of a footrace, ChatGPT started the race strong. Claude, new to marathons, has shown promise as a star in the field, and Gemini has been a leader in the sport for generations. This race is neck and neck, with new developments each week. It is nearly impossible for the average consumer to stay up-to-date on the latest advances in each model. These leading AI assistants may seem interchangeable at first glance, but each offers distinct capabilities worth understanding if you want to maximize their potential.

First to the table was ChatGPT, developed by OpenAI and released in November of 2022, ChatGPT is often a person's first expereince with AI. Its strength lies in its versatility and the ecosystem built around it. With plugins and a vast user base, ChatGPT excels at general knowledge tasks and integration with other tools. However, it sometimes suffers from what users call "hallucinations" – confidently stating incorrect information as fact.

Claude, Anthropic's LLM, is a personal favorite. Claude distinguishes itself through nuanced reasoning and a more conversational approach. Where Claude truly shines is in maintaining and referencing context over long conversations. Users also find Claude particularly adept at handling sensitive topics with greater care. The leaders at Anthropic, the producres of Claude, are central to the conversation being had surrounding the ethical component of AI.

And finally, Gemini, Google's entry into the space. Gemini leverages the company's already vast information infrastructure and user network. Seamlessly maneuvering through Google's ecosystem, Gemini presents itself as a easy transition into AI for people immersed in Google already. Its multimodal capabilities allow it to process and reason about images alongside text more effectively than its competitors.

An ability to decipher the language surrounding these models is what will ultimately enable you, as a consumer, to choose which is right for you. The language surrounding this industry, is, true to form, also rapidly evolving. Different companies use different names for the same concepts, and the sometimes even the same names for multiple concepts. Below, I offer a breakdown to the major terms used in describing the features of AI systems.

1. Large Language Model (LLM): The AI system itself, specifically a system trained to understand human language and generate responses.
2. Context Window: The amount of text an AI can remember and interpret at once during a conversation.
3. Multi-Modal: An AI's ability to work with more that just text, think:images, video, audio.
4. Projects/Projects/Gems: ChatGPT/Claude/Gemini's, respectively, saved workspaces where you can organize past conversations and provide specific context.
5. Artifacts/Canvas: ChatGPT/Claude's, respectively, interactive areas where you can view and edit output from AI. Particualrly helpful in writing code. No Gemini equivalent.
6. Web Access: An Ai's ability to browse the internet for up-to-date information. A preferential choice for the parent companies. Without web access the AI relies solely on pre-coded context.
7. API Access: A interface for developers to connect the AI to apps or systems.

When deciding which AI tool to use, whether it's ChatGPT, Claude, Gemini, or something else, it helps to start by asking yourself what you need most. Are you writing code? Summarizing dense material? Looking for creative brainstorming? Want real-time internet searches? Each model shines in different areas, and understanding your priorities is the first step.

To conclude, ChatGPT is highly versatile and strong with data analysis and custom AI assistants. Claude excels in thoughtful, long-form reasoning and handles very long documents with clarity. Gemini integrates smoothly with Google tools and performs well in multimodal tasks like understanding images and documents side-by-side.

Beyond raw capabilities, consider the experience. Do you want a sleek interface with voice and visual features? ChatGPT is your answer. Do you care if the AI remembers your preferences over time? Maybe Claude is the one for you. Are you planning to use it for personal tasks, professional research, or software development?

It's also important to think about access and cost. Some tools require subscriptions for premium features. Othersoffer powerful capabilities for free, though with usage limits or fewer integrations.

But here's the real secret: you can't really go wrong. All of these tools are excellent in their own ways. The most valuable thing you can do is start using them. Explore, compare, and develop a feel for what works best in different situations. Building an intuitive understanding of how AI responds, what it can (and can't) do, and where it adds the most value is far more powerful than sticking to just one brand. The future belongs to those who learn to think with AI—not just those who choose the "best" one.

Measure:	OpenAI: ChatGPT	Google: Gemini	Anthropic: Claude
Pricing	Plus: $20/month	Advanced: $10/month for students	Pro: $20/month
Coding	Great	Great	Best (opinion)
Multi-Modal	Yes	Yes	No
Projects	Yes	No	Yes
Web Access	Yes (in Plus plan)	Yes	No
API Access	Included w/ Plus	Yes	By Request
Special Abilities	CustomGPTs Voice Mode Multimodal	Integrated with Google's ecosystem	Strongly considers ethics Long memory and context

More than a Study Spot

Making Use of Tulane Library Resources

By Julia Miller

The library is more than a spot to grab a drink or study. It offers so many resources online and in-person that can help you research, learn new skills, and get access to all kinds of information and equipment. Here is a profile of some general resources that the library offers, so you can make the most out of Tulane Libraries.

Interlibrary Loans

Have you found the perfect book or article online for a paper but you don't have access to it at all? Do you need a book for class but you can't seem to find it anywhere? You can actually get all of these resources (free of charge, no pirating necessary) through Tulane's Interlibrary Loan Program. Simply create an account, fill out the book or article information, and wait. In a couple of days or weeks, there will be a PDF copy of the article in your email or a book waiting for you to pick up at Howard Tilton Memorial Library (uptown) or the Matas Library (downtown).

Digital Collections

Tulane also has digital collections with all kinds of resources. Undergraduate honors these, dissertations, digital exhibits, old student newspapers and yearbooks, collages, and a whole host of other resources are all available here. It includes scholarship about Tulane or by Tulanians. The collections span across different subjects and time, so much so, I can't sum them all up. Check it out here!

Creative and Resource Technology Lending Program

The library has a whole program where you can check out different technological equipment. Got a DVD or CD, but nothing to play it with? Check out a player. Want to record a podcast? Check out a mic. Want to host a watch party? Check out a projector. Some resources like a microfilm viewer, a record player, and tv are only available in the library, but most others you can take outside the library.

Find a Subject Specialist

Not even sure where to start on a research project? You can meet with a librarian to help you with that. There are multiple librarians, each with specialization in different research topics, who you can meet with them to help figure out what databases you should be looking in, what kinds of sources you might want to include in your paper. They can help you get access to resources that you might not already have and help guide you in the right direction to start your research. (Also, shoutout to Kay P. Maye, the Scholarly Engagement Libarian for Social Sciences and Data, who helped me figure out which library resources to profile here!)

Workshops, Talks, and Trainings

The library offers workshops, talks, and trainings that you can attend. They offer workshops about using programming languages, AI, and audio and video editing programs. They have talks about new resources in different disciplines, finding funding opportunities for research, and using primary sources. They are always putting on a variety of different events. If you are curious about what they are putting on right now, check out this link.

Scan for links to these resources

or go to library.tulane.edu

Try out at least one more in the next couple of months. These resources are offered to all Tulane students, you might as well make the most of them. To recap, Tulane Libraries offers a host of resources including Interlibrary Loans, research consultations, workshops and talks, technology lending, and digital collections. You might never use these resources, but now you about some more things the library has to offer instead of just books.

IS AI ACTUALLY INTELLIGENT?

DANIEL FOX

Abstract:

AI is very good at summarizing information or doing calculations, but its limitations in intelligence lie in the fact that it does not understand reasoning nor can it think independently. I think people believe AI is extremely intelligent and is a threat to people's jobs, however in reality this is nowhere near the case. AI is an extremely helpful tool to assist people with a variety of tasks. It is nowhere near the level of human intelligence at the moment, even if it can outperform humans at some tasks.

How do AI LLMs work?

Large language models generate text by using its training to predict the next word in a sequence. Every word generated by a LLM is a separate prediction of what word typically follows the preceding word. LLMs are able to make these predictions after learning from training data, typically text data from the internet. This is what allows LLMs to generate human sounding text. The important distinction here is that LLMs have no understanding of language like humans. They learn from their input data to make predictions based on common sequences of words written by humans. In this way, AI does not possess any actual consciousness or thinking ability that humans have.

Works Cited

How Smart Is Artificial Intelligence?, www.iec.ch/blog/how-smart-artificial-intelligence. Accessed 6 Apr. 2025.

"Is Ai Really Intelligent?" 6point6, 12 Feb. 2024, 6point6.co.uk/insights/is-ai-really-intelligent/.

Nationalacademies.Org, www.nationalacademies.org/news/2024/02/just-how-intelligent-is-artificial-intelligence. Accessed 6 Apr. 2025.

Stöffelbauer, Andreas. "How Large Language Models Work." Medium, Data Science at Microsoft, 24 Oct. 2023, medium.com/data-science-at-microsoft/how-large-language-models-work-91c362f5b78f.

WHAT IS AI GOOD FOR?

- PREDICTIVE ANALYTICS
- AUTOMATION
- PATTERN DETECTION
- DATA PROCESSING

WHAT IS AI NOT GOOD FOR?

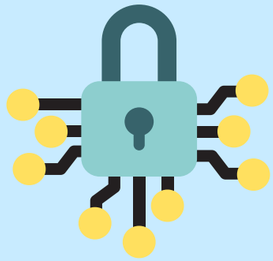
- UNDERSTANDING CONTEXT AND REASONING
- ORIGINAL CREATIVE THOUGHT
- UNDERSTANDING EMOTIONS
- UNDERSTANDING MORAL IMPLICATIONS

DATA PRIVACY

DO'S & DON'TS

TO SECURE YOUR IDENTITY

BY: MEGHAN NAGIA



What is Data Privacy?

Data privacy is the idea that individuals have control over their personal data, including how it is collected, used, stored, etc. It refers to the protection of personal information from unauthorized access and misuse.

In today's digital sphere, user data such as biometrics, contact information, credit card numbers, etc is regularly collected by corporations, websites, and organizations in the United States for targeted advertising and monetization but can also be used for more punitive or controlling measures in other areas of the world. These data collectors are responsible for practicing data privacy by implementing and adhering to policies that allow users to have access to the data being held and stay informed on who has their data and how its being used. Organizations that collect personal data should obtain user consent, implement security measures to protect data, and be transparent with their intentions.

However, in the advanced technical stage we currently live in, data privacy almost does not exist. There is always someone out there who wants personal data, and data brokers who package and sell information. This makes the concept of data privacy hard to grasp because everyone's digital footprint is being watched. Therefore, each individual should be accountable and mindful about their own digital footprint and what information is being shared online. Essentially, to keep data private, we should maintain control over private details and position ourselves in a way to protect our individual rights over personal data.

Why Does it Matter?

Living in a data-based economy, personal information is consistently collected, analyzed, and traded; therefore, protecting one's data helps maintain control over who knows what information, avoid influence over purchasing decisions, and mitigate the risk of misuse and exposure from ill-intended entities like data brokers. Strong data privacy practices ensure security and prevent identity theft, fraud, or negative impacts on personal and professional reputations. When one is aware of what data is shared and how it is being used, informed decisions can be made regarding which accounts and services to use and what information should actually be shared. Data privacy extends further into areas regarding civil liberties, a competitive advantage, and security. The United Nations recognizes data privacy as a fundamental human right. Organizations who earn consumer's trust with their data may collect and use data more easily than one with a poor data privacy reputation. Finally, companies that prioritize data privacy have a less likelihood of a data breach.

If a Data Breech Occurs

- Visit and report to IdentityTheft.gov or call the Federal Trade Commission (FTC) identity theft hotline
- Place a credit freeze on your accounts
- Change and update all important passwords
- Monitor compromised accounts

Sources:

Meet the Expert Chris Brook Editor. "101 Data Protection Tips: How to Protect Your Data." *Data Protection Tips: Keeping Personal Info Safe | Digital Guardian*, 22 Aug. 2024, www.digitalguardian.com/blog/101-data-protection-tips-how-keep-your-passwords-financial-personal-information-online-safe.

Staff, BCP, and Gema de las Heras. "Protect Your Personal Information from Hackers and Scammers." *Consumer Advice*, 21 Nov. 2024, consumer.ftc.gov/articles/protect-your-personal-information-hackers-and-scammers.

Tips for Securing Your Data." *Information Technology - Northwestern University*, www.it.northwestern.edu/security/protect-information/secure-data-tips.html. Accessed 4 Apr. 2025.

K Royal - JD, PhD, FIP, CIPP/E /US, CIPM Global Chief Privacy Officer, Deputy General Counsel

Allyson Heumann - Tulane School of Liberal Arts

SPOTIFY WRAPPED:

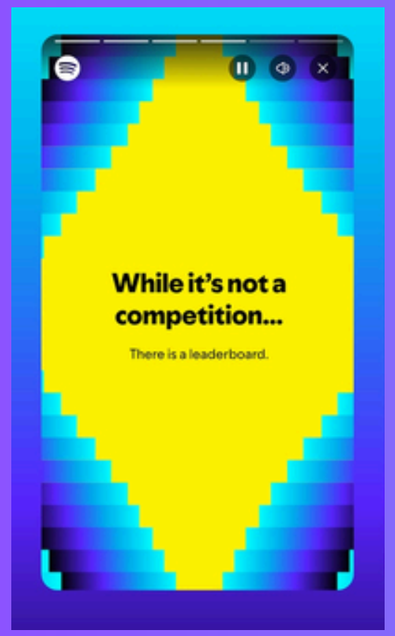
The Gateway to Personal Data Analytics by Katie Bogdanow

Every December, we wait in anticipation for our Spotify Wrapped, excited to view, analyze, and share our own data. This phenomenon has recently grown into something bigger than Spotify: everyone wants to see their year in review. Fast food restaurants, social media, dating apps, and many more platforms have introduced their yearly recap, but why? The explanation lies at the intersection of technology, culture, and personal identity, revealing how Spotify Wrapped has fundamentally changed our relationship with data analytics.

How has Spotify Wrapped changed our view of Data Analytics?

The Data Behind the Phenomenon

You may be wondering, "How is this data?" The answer is quite simple, every input - songs played, playlists listened to, keywords searched - as bits of information that help create personalized suggestions and give feedback to optimize the listening experience. This data is called **user-generated data**, it is produced by the user and analyzed for their benefit. At its core, Spotify turns simple data into something with personal and cultural significance. No one had thought to statistically evaluate music, which created quantifiable evidence of music taste. The **datafication** of music taste is an indicator of a bigger cultural phenomenon that validates users' psyches and helps shape their identity.



The Psychology of Personal Analytics

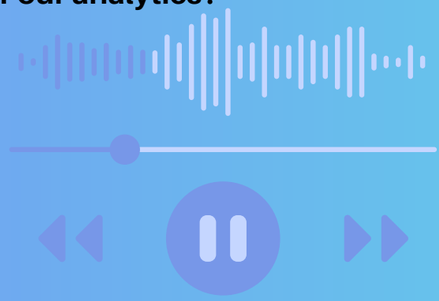
Spotify Wrapped caters to a paradox of competing human desires: we want to be a part of a community while maintaining a unique, somewhat superior identity. Spotify manages to bridge the two by creating personalized graphics and encouraging users to post them on social media. This fuels a socio-psychological competition that allows us to determine our perception of others based on their music taste. Receiving aesthetically pleasing data about ourselves gives us an insight into our own identity, we love the external validation of our self-perception.

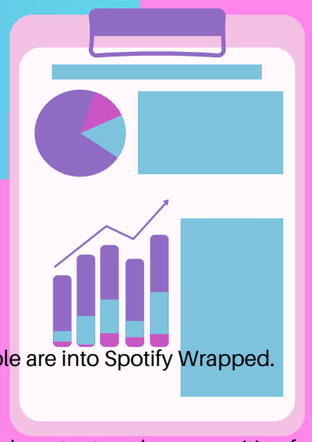
The Wrapped Formula

Spotify has successfully created a cultural measure of data analytics that has infiltrated the framework of all digital platforms and businesses. The end-of-year recap has become an annual time for self-reflection, celebrating individuality and milestones. Media tracking apps like Goodreads and Letterboxd and lifestyle platforms Beli and Duolingo, and even Tinder, have all followed suit in adopting the "wrapped" framework. This societal shift indicates a new era of data analytics, one where users want to know and share their data.

Broader Implications

As we await Spotify's next vibrant storytelling of our music tastes and shareable statistics, it is important to recognize the broader consequences of extensive data collection. Digital platforms are increasingly collecting, analyzing, and packaging our data, which leads to the question: **How much digital privacy will we give up in exchange for our analytics?**





References:

1. Resnick, B. (2023, December 8). Love sharing your favorite music with friends? People are into Spotify Wrapped. NPR.
2. Patel, N. (2023, December 8). Why is Spotify Wrapped so popular? UX Design.
3. MarketScale. (2023). Spotify Wrapped: A masterclass in leveraging user-generated content and communities for brand domination. MarketScale.

AI & The Environment

by: Sam DeMarinis

According to Goldman Sachs, artificial intelligence (AI) will increase global GDP by $7 trillion over the next ten years. What does this mean? A multitude of economic benefits for countries all over the world, including large boosts to national economies and overall growth. Evidently, AI has a multitude of benefits for businesses, students, teachers, and others. But what about the negative consequences? While there is already much discussion on certain types of issues involving AI such as plagiarism, many people fail to recognize one of the most pressing and potentially devastating to our planet: environmental impacts of AI, specifically with LLMs like ChatGPT.

Before we dive into some of these environmental impacts, it's important to understand how a LLM like ChatGPT actually works. In order to create such a model, a large amount of energy must be used for the training process. But what exactly does this entail? Millions of examples of text (articles, conversations, etc.) from which the AI learns key patterns in language so that it can generate its own, unique output. The specific type of algorithm that ChatGPT uses is called a transformer, which is based on a neural network meant to mimic the way humans think (Texas State University Library). Once the model is sufficiently trained, it can effectively produce answers to queries from a user. So you type up your query, for example: "Write me a song about cats." What exactly happens when you hit enter?

- **The request gets sent over to OpenAI servers, the company that created ChatGPT.**
- **Then the model can actually process your request. It uses a giant database of words, patterns, and probabilities to understand what makes sense as an output.**
- **Finally, the model predicts the best response based on its training, and sends it back over the internet to your screen.**

This process occurs almost always within a couple seconds, which may lead us to believe that it isn't doing much harm to our planet. But this inference process, as well as that of training the model, require staggering amounts of computing power.

According to Earth.org, the training process of ChatGPT used "approximately 700,000 litres of freshwater" in Microsoft's data centers. This freshwater was used in order to maintain a cool temperature for the machines in these centers. To offer some perspective, the article also provides some concrete comparisons, noting that this quantity of freshwater would be enough to produce "370 BMW cars or 320 Tesla vehicles".

You may be thinking, "Well, that's not much… and isn't the training process over?" Partially. The training for ChatGPT-3 and -4 is done, but OpenAI continues developing new versions like ChatGPT-4-turbo and fine-tuned models. Each requires its own training process. While it likely uses less energy than the original LLM, the training process isn't completely over. As OpenAI keeps improving its models, freshwater and energy consumption will only increase.

And that's just for training. What about the energy used during inference? In other words, how much power does ChatGPT need to generate your unique cat song? To produce a response, the server runs calculations to predict an appropriate reply. Again, this process generates heat, requiring large amounts of freshwater to cool data center computers. Freshwater consumption continues every time you interact with the AI. According to the International Energy Agency, "a request made through ChatGPT… consumes 10 times the electricity of a Google Search." We can imagine how this scales with more complex queries.

It is practically impossible to halt the usage of AI in our world today. But as we can see, this tool is also having a negative impact on our planet, and it's important that we recognize this now. Instead of going to AI as a first solution to a problem, we should think about other ways we can arrive at a solution. For example, could a Google search be just as effective? Of course, the use of LLMs like ChatGPT need not be outlawed completely. But in situations where a more environmentally friendly solution could suffice, we should think twice about instinctively resorting to the use of AI.

References:

Goldman Sachs. "Generative AI Could Raise Global GDP by 7 Percent." Goldman Sachs, 4 Apr. 2023, www.goldmansachs.com/insights/articles/generative-ai-could-raise-global-gdp-by-7-percent.

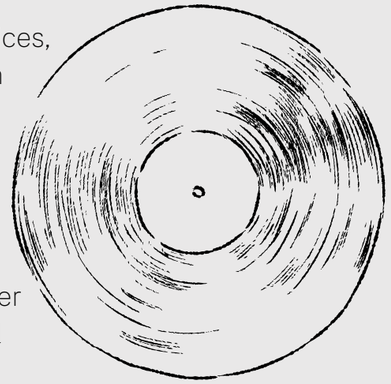
Texas State University Library. "Ethical and Legal Considerations - AI in Research and Education." Texas State University Library Guides, guides.library.txstate.edu/c.php?g=1321038&p=9718369.

Earth.Org. "The Environmental Impact of ChatGPT." Earth.Org, 17 Feb. 2023, earth.org/environmental-impact-chatgpt.

United Nations Environment Programme (UNEP). "AI Has an Environmental Problem. Here's What the World Can Do About It." UNEP, 7 June 2023, www.unep.org/news-and-stories/story/ai-has-environmental-problem-heres-what-world-can-do-about.

Big Data Beats: The Hidden Costs of Spotify

by Mika Nijhawan



Over 700 million people worldwide use music streaming services, and I am among the 675 million who use Spotify. Last year on Spotify, I listened to 124,809 minutes of music. The year before, it was 120,159 minutes. And the year before that, 160,739 minutes. Thus, according to my Spotify Wrapped, my music-listening habits place me among the top 0.5% of listeners worldwide. Suffice to say, silence is a rare presence in my day. In that sense, I embody the ideal Spotify user—after all, as a former employee once put it, "the platform's only real competitor is silence."

I. THE INVISIBLE COSTS OF SPOTIFY

Spotify represents one of the most common entry points into big data the average person is engaging with. While Spotify's algorithm and data practices may seem banal, in truth, the collection of our data on this app reflects a broader trend: most young people interact with and contribute to big data systems without fully considering their implications. Consider the 675 million Spotify users worldwide. While we tend to focus on the benefit of uninterrupted streaming, below this accessibility lies the ever-constant collection of our personal information. Further, the platform does not simply vaguely understand your music preferences; Spotify is intimately acquainted with your tastes as the site collects data on your habits, locations, and preferences. Specifically, the streaming service gathers data on how often you skip a song or how many times a day you listen to a playlist. Spotify also analyzes what it calls "usage data" which encompasses things such as what music you are searching on the platform, and "device sensor capabilities," which track your movement or how you are holding your device. This means Spotify is analyzing if you are listening to music on your morning run or evening home commute to work. While this may mean you get personally tailored workout mixes in your library, Spotify frequently shares this usage data with third-party services, highlighting major privacy concerns.

It is because of this extensive data collection that my algorithm recognizes my emotions and moods, and in many ways, I am grateful for it. My "daylist"—a Spotify generated playlist that recommends songs based on the day and time—knows on Sunday nights I want to listen to "emo slowcore mix late nights" on Wednesday afternoons I want to listen to "60s bob dylan laurel canyon," and on Friday evenings I want to listen to "uk garage and drum n bass." Maybe it is because I spend so little of my day in silence, or maybe it is an artifact of my generation to create personal markers out of the media we consume, but my Spotify listening habits are incredibly personal to me. I do not mistake myself in believing I am unique in this, yet on an average day, I spend between 5 and 6 hours listening to music. Thus, my near-constant engagement with sound is a defining feature of how I navigate through the world. However, this intimacy I feel with my music preferences has been designed to trap me on the platform and to keep me streaming music for as long as possible. And it is clearly working.

II. STREAMING AND SURVEILLANCE CAPITALISM

Beyond this leveraging of our emotions, what is the result of this constant mining of our data? For one, the CEO of Spotify, Daniel Ek, is now richer than any musician in history, with a net worth in the billions. Thus, to Spotify, our emotions and preferences are not just personal; they represent revenue. In this way, Spotify is also a prime example of surveillance capitalism–the practice of companies collecting private data for profit–and yet, for all of 675 million Spotify users, the ease of instant streaming outweighs these concerns. As pointed out by columnist Wessel Joosten:

> ***As we enjoy our music and "expose" ourselves to the platform, we effectively become "laborers" for Spotify; supplying the "raw material" that fuels its surveillance capitalist model and contributes to its vast revenue streams***

Within this context, Spotify's algorithm is designed not just to deliver music but to maximize user engagement, ensuring that we remain tethered to the platform to continuously feed it valuable data. Yet, for me and many others, the information Spotify is collating is deeply personal; after a breakup, Spotify knows to recommend me sad love songs, and late at night, it knows to recommend me soft, melancholic songs. These moments are not just data points; they are intimate details about my life and my feelings, moments I am not entirely comfortable being utilized for increased profit revenue or shared with third-party sites.

III. LISTENING BEYOND THE ALGORITHM

I am acutely aware that with every song I stream, I am feeding data into a system that profits from my habits and sells my information to third-party companies. And with every second spent on the site, I am reinforcing a system that expands the reach of the surveillance state. But, realistically, I will not stop using Spotify anytime soon. In fact, I was streaming music the entire time I wrote this piece. Yet, I try to remember that music exists outside streaming platforms, and I can discover art without the recommendations of an algorithm. Recently, I attended a talk at Howard-Tilton Library on Black Masking Indian culture, where several live archival recordings of the Mardi Gras Indian group, The Wild Magnolias, were played. Further, I have also been listening to music on syriancassettearchives.org, an online catalog of digitized cassettes from Syria's expansive cassette era. While you can find Wild Magnolias albums and, I am sure, a few digitized Syrian cassettes on Spotify, the experience is not the same. There is something distinctly rewarding about engaging with music outside the confines of a data-driven platform—music that is not being fed to me through an algorithm but rather discovered through curiosity and a genuine love for this form of artistic work. Part of the joy of these listening experiences, for me at least, was the reminder that music is not just content to be consumed; it is something that exists beyond algorithms, waiting to be sought out on our own terms.

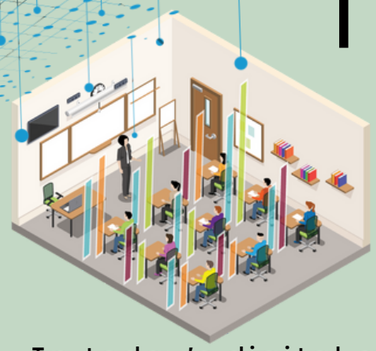
References

Hsu, Hau. "What is Spotify Really Costing Us?" *The New Yorker*, December 23, 2024. https://www.newyorker.com/magazine/2024/12/30/mood-machine-liz-pelly-book-review.

Joosten, Wessel. "The Hidden Cost Of Convenience: Unveiling Spotify's Data Practices for Privacy Implications " *Diggit Magazine*, October 14, 2024. https://www.diggitmagazine.com/articles/hidden-cost-convenience-unveiling-spotifys-data-practices-privacy-implications.

Data on your campus:

TU Edition



By Alexandra Zeff

Living in a Data-Driven World

In today's digital age, data has become an integral part of our daily lives, and college campuses are no exception. As a student, you might not realize it, but your every move, both online and offline, is generating data that's being collected, analyzed, and used to shape your college experience. From the moment you step onto campus to the time you graduate, your academic performance, social interactions, and even your daily habits are being tracked and studied. This constant data collection raises important questions about privacy, ethics, and the role of technology in higher education

Your Digital Footprint on Campus

Universities are increasingly using data to keep tabs on student well-being and engagement. This goes far beyond simply tracking your grades or class attendance. For example, your splash card swipes at the dining hall, gym, or library are logged and analyzed to detect patterns in your routine. Some universities even use social media posts to gauge the overall moral of the student body - like the daily check-in's on Canvas. Colleges collect and analyze usage patterns of various facilities to make informed decisions about resource allocation. For example, the reason The Commons, Tulane's dining hall, always seems to have vegan options on Mondays might be because that data shows more traffic at the meatless stations.

Your Role in an Ecosystem of Data

As a student in this data-rich environment, you have both the right and the responsibility to advocate for transparency and ethical practices in data collection and use. Many universities have policies in place regarding data privacy and usage, but these policies may not always be clear or comprehensive - look at Tulane's today! Your voice matters in ensuring that data practices respect student privacy and are used for the benefit of the student body. By becomingan active user of your own data, rather than just a passive subject of data collection, you can make more informed decisions about your education and take control of your academic success.

Stay Safe Online:

Privacy Tips for Students

- Use unique passwords for all your accounts. A password manager can help you generate and store passwords.
- Be mindful of what you share on social media; remember that potential employers, professors, and others may be able to view your posts in the future.
- When using public Wi-Fi networks, which are common on college campuses, use a Virtual Private Network (VPN).
- Verify the sender's identity before clicking on links or providing personal information.

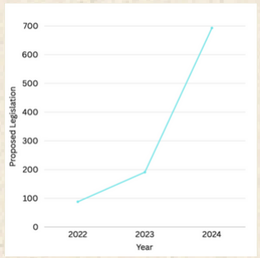
References:

- https://www.watermarkinsights.com/resources/blog/how-to-use-data-to-improve-campus-learning/
- https://www.aplu.org/our-work/5-archived-projects/accountability-and-transparency/using-data-to-increase-student-success/

ARTIFICIAL INTELLIGENCE IN THE LAW

Zoe Oboler

Over the past few years, The use of AI has rapidly increased across the world. This new technology, while transformative, has been met with legal and ethical concerns since its inception. Because of the novel technology involved, it has remained unclear how the existing cas law and regulations applies to AI. In the past months we have seen a major increase in legal cases and laws that attempt to regulate the creation and use of Ai. Of the nearly 700 proposed laws this year, across 45 states, 113 were enacted into law.



Colorado has become the first state to pass a comprehensive law about AI usage, potentially becoming a model for other states. On May 17, 2024, Colorado enacted the Colorado Artificial Intelligence Act (CAIA). This law, which will go into effect in 2026, applies regulations and oversight to “high risk AI systems. A “high risk” system is one which makes consequential decisions for or about consumers, such as hiring, housing, or insurance.

This law provides regulation not only for the developers of these systems but also for the deployers. Developers are required to provide “reasonable-care” to protect consumers from discrimination and also to document the steps they've taken, for use by regulatory bodies, consumers, and deployers. Deplyers on the other hand are required to conduct annual assessments of bias within their systems, and to notify consumers when consequential decisions are made by AI. Robert Rodriguezz, the bills sole sponsor, claimed “It’s always been about providing a framework for accountability, for biases and discrimination and just making sure that people know when they’re interacting with it.” Critics of the law, such as Kyle Shannon, a local business owner, stated “ I know that overregulating AI right now is going to put Colorado businesses at a significant disadvantage.” Koi Marshall, another critic addressed the potential for discrimination, saying “However, senators, pinpointing the sorts of catalysts of discriminatory outcomes of AI systems is not always possible, nor is consistently determining who or what is responsible for the act of discrimination. Unfairly biased outcomes are problematic for developers, deployers and users like all of us in this room.”

In contrast to Colorado's decision, Virginia Governor Glenn Youngkin vetoed House Bill 2094, the High-Risk Artificial Intelligence Developer and Deployer Act. Virginia's law, if it passed, would have enacted provisions similar to Colorado's law, targeting both the developers and deployers of “High-risk” ai systems. Similar to Colorado's law, which passed with a 3-2 vote along party lines, Virginia's proposed legislation passed through the house and senate with a narrow majority. Proponents once again emphasized the need for greater consumer protections. Governor Youngkin however was more concerned about the financial impact of the bill stating it "would harm the creation of new jobs, the attraction of new business investment, and the availability of innovative technology in the Commonwealth of Virginia."

There are just two examples of the larger trend of enacted and proposed AI legislation across the country. In 2024 Illinois amended it’s human rights act to prohibit employers from using AI in ways that result in discrimination against protected classes. California enacted Assembly Bill 2655, which requires online platforms to remove or label deceptive Ai-Generated content related to elections. In 2023 New York implemented a law requiring independent audits of automated hiring decision tools. Texas has also proposed a law to regulate "high-risk artificial intelligence systems," similar to the Colorado and Virginia bills. Despite this progress there are still no comprehensive federal legislation on the use of AI, especially for employment. This decentralized regularity landscape has made enforcement difficult, and also put larger compliance burdens on businesses operating in multiple jurisdictions. This is especially pronounced for small companies who may be hit harder by these mandates and regulations.

Other countries have already begun wrestling with these complex regulatory issues. The European Union has recently introduced the Artificial Intelligence act. According to the European Parliament "The regulation establishes obligations for AI based on its potential risks and level of impact." China has already issued draft measures whole require Ai-generated content to adhere to the CCP’s larger ideology, and companies may be held legally liable for their training data and the content produced by their platforms. A recent article by Reuters notes “The EU’s approach contrasts with the lighter, voluntary AI oversight in the U.S. and China’s stricter, stability-focused regulations."

As Artificial intelligence rapidly evolves, policymakers and industry professionals across the globe are wrestling with the legal and ethical impacts. The path forward requires finding a delicate balance between innovation and risk mitigation. As much as AI provides the potential for unmatched growth and efficacy, it also creates bias and inequities. Especially in important decisions, such as employment, healthcare, and housing, AI has the potential to uphold, and even deepen inequalities that have been historically present in our society. Regardless of what we decide to prioritize it is becoming increasingly clear that we need national, if not global, regulations ans guardrails. This framework will allow small companies to more easily meet requirements, easing the burden on Ai developers and users. It also has the potential to drastically increase levels of compliance with existing regulation. Overall, AI is a powerful tool, and it’s rapid advancement has made it hard for the law to keep up. Over the next few years it is crucial that legislation, at the state and national level begin to approach these issues.

References
Baker McKenzie. (2024, August). Illinois joins Colorado and NYC in restricting generative AI in HR: A comprehensive look at U.S. and global laws on algorithmic bias in the workplace. The Employer Report. https://www.theemployerreport.com/2024/08/illinois-joins-colorado-and-nyc-in-restricting-generative-ai-in-hr-a-comprehensive-look-at-us-and-global-laws-on-algorithmic-bias-in-the-workplace/
Fortune. (2025, March 28). EU AI Act enforces code of practice, copyright standards. https://fortune.com/2025/03/28/eu-ai-act-code-of-practise-copyright/
Governor of California. (2024, September 17). Governor Newsom signs bills to combat deepfake election content. https://www.gov.ca.gov/2024/09/17/governor-newsom-signs-bills-to-combat-deepfake-election-content
K&L Gates. (2025, January 13). The Texas Responsible AI Governance Act and its potential impact on employers. https://www.klgates.com/The-Texas-Responsible-AI-Governance-Act-and-Its-Potential-Impact-on-Employers-1-13-2025
National Law Review. (2024, November). Growth of AI law: Exploring the legal challenges of artificial intelligence. https://natlawreview.com/article/growth-ai-law-exploring-legal-challenges-artificial-intelligence
National Law Review. (2025, March). Virginia governor vetoes AI bill as states struggle to approve regulations. https://natlawreview.com/article/virginia-governor-vetoes-ai-bill-states-struggle-approve-regulations
NPR. (2024, June 22). Colorado passes AI law to prevent discrimination in hiring. https://www.npr.org/2024/06/22/nx-s1-4996582/artificial-intelligence-law-against-discrimination-hiring-colorado
Stanford Institute for Human-Centered AI. (2024). AI Index Report. https://hai.stanford.edu/ai-index
TechPost (BSA). (2024, October 22). 2024 state summary on AI. https://techpost.bsa.org/2024/10/22/2024-state-summary-on-ai/
The Colorado Sun. (2024, April 25). Colorado’s generative AI regulation clears Senate committee. https://coloradosun.com/2024/04/25/colorado-generative-ai-artificial-intelligence-senate


FEBRUARY
MARCH
APRIL


INVISIBLE WOMEN
GIS RESEARCH METHODS
FOLLOW US @TULANECAIDS

2024 - 2025 DAC MEMBERS

Ifeoma Osakwe (she/her)

Data Ambassador Council

Product Developer

Oma Osakwe (she/her) is a graduating senior in Cell and Molecular Biology, with a minor in Psychology. At CAIDS, she is the Data Lab Manager and oversees the DAC. Prior to this role, she worked with the Digital Research Intern team as a Product Developer. Outside of school, she is a Newcomb Scholar, Co-President of the Tulane African and Caribbean Student Association, a Resident Advisor, and a Green Wave Ambassador. For her honors thesis, she is working in the Shusheng Wang lab to study potential therapeutics for age-related macular degeneration.

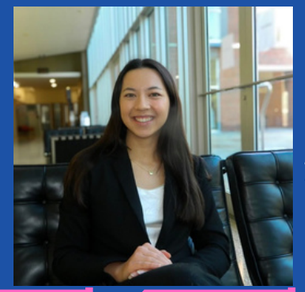
Cameron McLaren (she/her)

Data Ambassador

Cameron is a graduating senior with a Bachelor's in Psychology and Computer Science. At CAIDS, she is the Communications Coordinator as well as a member of the Data Ambassador Council (DAC). During her time at Tulane, Cameron has held leadership positions in several clubs, been a TIDES Peer Mentor, a Community Engagement Advocate (CEA), and is currently in her second semester of volunteering for Students Against Food Insecurity (SAFI). She has also worked as a Hub Assistant (HA) for the past two years and this year holds a position as a Resident Advisor (RA). Outside of school, Cameron plays the violin and has been a member of the Tulane Orchestra for the past three years.

Sara Vannoni (she/her)

Data Ambassador



Sara Vannoni is a sophomore from California studying Public Health. This will be her second year with CAIDS, and she is excited to continue to grow the program with both undergraduate and graduate communities at Tulane. She hopes that CAIDS will be able to increase engagement with students through events, guest lectures, and workshops! In addition to spreading data literacy, she is interested in exploring data-related fields in both the academic and commercial world and hopes to pursue research in biostatistics and/or health policy in the future. In her free time, you can find her walking around Audubon Park with a Peach Palmer Tea or volunteering with Tulane's Culinary Medicine Initiative!

Davis Green (he/him)

Data Ambassador

Davis is a second year studying Political Science and Data with a minor in Strategy, Leadership, and Analytics. Davis is passionate about data science, and is happy to be serving this fall as a first time member on the Data Ambassador Council to try and share his excitement for data with his peers and help promote CAIDS and their initiatives to the Tulane student body. Davis, from Los Angeles, California, is seeking to combine his strong interests in either music or sports analytics with his passion for data science in his professional career, hoping to do data science in one of the two fields.

Eghosa Asemota (he/him)

Data Ambassador

Eghosa Asemota (he/him) is a junior studying Political Economy and Portuguese and minoring in Strategy, Leadership, and Analytics. He is passionate about diversity, equity, and inclusion work, social justice, and helping further the voices of marginalized groups domestically and internationally. Eghosa's academic interests are racial wealth equity, sustainable cities, equity in education, arts based social movements, and surveillance. He currently serves as a Data Ambassador and is excited to bring data literacy initiatives to Tulane's campus!

Samuel Johnson (he/him/his)

Data Ambassador

Samuel Johnson is a sophomore studying History and Homeland Security. He is an Ambassador to CAIDS and is from New York. Something he is excited about within CAIDS is showing how easy it is to get into data and that prior knowledge is optional when learning about data. In the future, Sam hopes to utilize his knowledge of history, data, and homeland security to better the world through public policy. Some personal interests of his are lacrosse, chess, and cooking.

Favour Olushola (she/her)

Data Ambassador

Favour Olushola (she/her) is a senior from Port-Harcourt, Rivers State, Nigeria currently studying Information Technology and on the pre-medical track. She is a member of the Data Ambassador Council and is excited about working with CAIDS to ensure equitable access for everyone to data literacy tools, education and training on Tulane's campus. In her free time, she enjoys listening to music. Her favorite song is Peace Be Unto You by Asake.

Nina Moore (she/her)

Data Ambassador

Nina Moore (she/her) is a Junior studying Chemical Engineering & Computer Science. She was born in Merrillville, Indiana but has been living in New Orleans for 8 years. She works as a Student Ambassador for CAIDS to promote data literacy among the Tulane student body. She is excited to help all see how data is involved in their studies. In her free time she enjoys watch watching Caleb Hammer on Youtube and spending time with family and friends. She also loves going to festivals such as the Franklinton Fair and the Strawberry Festival. She wants to move to LA, Atlanta, or Raleigh, NC to work in the energy or chemicals industry post graduation.

Kai Williams (he/him)

Data Ambassador

Kai Williams(He/Him) is a sophomore studying Information Technology with a concentration in Cyber Security and looking to double major in Data Science. He is from Chalmette, Louisiana. While in Caids, he has learned how to divide thoughts to reach a common goal and analyze data deeply. He hopes to contribute to a cybersecurity firm or work for a government agency where he can use his skills to protect sensitive data and critical systems. Kai enjoys exploring emerging technologies, especially AI and machine learning. He's also passionate about community involvement.

Data Ambassador Council

2024-25 RECAP

Our Project:

DAC members act as advocates for accessible data practices that communicate CAIDS initiatives to the student body at Tulane. Members also act as a student-led focus group that explores data initiatives to ensure that CAIDS data programming supports the needs of Tulane students.

Tulane Connolly Alexander Institute for DATA SCIENCE

The DAC has spent over **250 hours** this past year creating promotional materials for CAIDS programming. Additionally, we helped to host Love Data Week 2025, making interactive activities for students to learn more data skills.

QUICK STATS

5

newsletters sent

8

zine articles written

15

course posters made

32

hours spent tabling

DATA AMBASSADOR COUNCIL

CONNOLLY ALEXANDER INSTITUTE FOR DATA SCIENCE

Learn more about our programs and initiatives at:
datainstitute.tulane.edu/data-ambassador-council

Data Ambassador Council Articles

The Data Ambassador Council (DAC) is a diverse cohort of undergraduate and graduate students. It is a focal point of student input and engagement at CAIDS. As advocates for data equity, DAC members communicate CAIDS initiatives to the student body at Tulane and act as a student-led focus group that explores data initiatives to ensure that CAIDS data programming supports the needs of Tulane students.

Quantifying Luck in Baseball: Can it Be Done in an Interpretable Way?

By Davis Green

Introduction

Since the “Moneyball” Era, baseball has been the epicenter of sports analytics. As tracking technology has evolved, people have configured new ways to use data to predict what might happen on the baseball field. However, one of the biggest issues this faces is luck and how to account for it. This begs the question: how do we know if a player’s statistical results are truly representative of their performance, and to what degree is luck driving their outcomes on the field? If we could quantify just how lucky a player is getting, could we then account for those insights when evaluating performance? I attempted to create a metric to try it for pitchers specifically, and its accuracy leads me to believe that the answer is yes.

Background

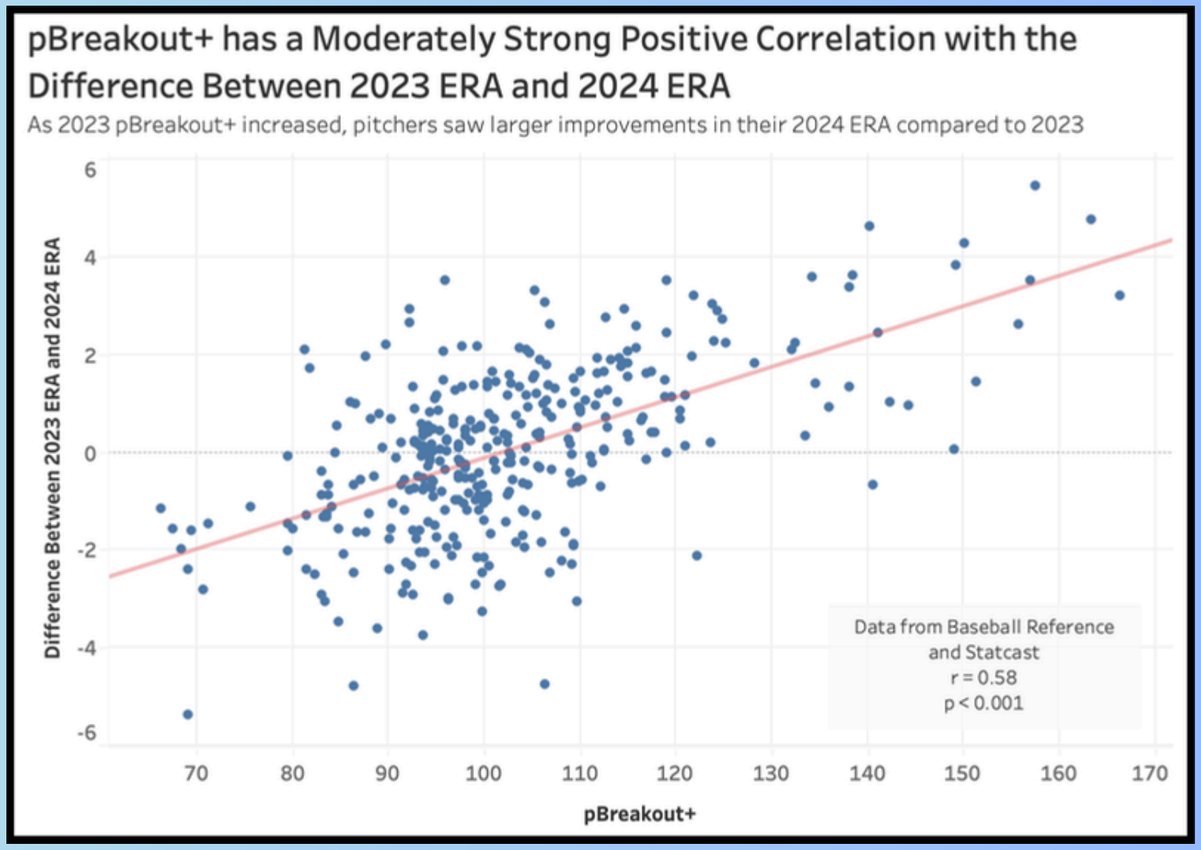
I am not the first person to try this, as many metrics with a similar goal already exist. The difference is that these other metrics do not try to measure luck within their own scale, they try to account for luck in the scale and context of existing statistics. For example, Earned Run Average (ERA) is the most commonly used metric to evaluate a pitcher’s performance. The goal of a pitcher is to have as low of an ERA as possible. In trying to account for luck, another statistic was created called Fielding Independent Pitching (FIP). FIP measures only the results a pitcher can control, home runs, strikeouts, and walks, completely eliminating a pitcher’s defense and the luck that surrounds them from the equation. One of the biggest draws of FIP is that it is measured on the same scale as ERA, which makes the statistic easy to interpret for those familiar with the scaling of ERA. While this is positive, it makes looking at the underlying message of FIP, measuring luck, difficult. FIP is applied to measure luck by looking at the difference between a pitcher’s ERA and FIP. If their ERA is lower than their FIP, it means that they were experiencing some luck that led to positive results. If their ERA is higher than their FIP, it means they were experiencing bad luck, leading to negative results. This method works when looking at one individual, but what if you want to look at how lucky that individual is compared to others? There is no context to interpret the difference between a pitcher’s ERA and FIP, as there is no standard number people are accustomed to that indicates to them if it is high, low, or average. It is incredibly difficult to interpret this number without the context of that number for every other pitcher in the league. This is the issue my project attempts to fix.

pBreakout+

To create my luck-encapsulating metric, I centered the number around the FIP-ERA difference, while also including the difference between other standard and luck-predicting statistics like SLG and xSLG and wOBA and xwOBA. More information on these statistics can be found on MLB.com. Ultimately, I added all of these statistics up into one number, assigning each one a relevant weight based on its level of predictability. The resulting number was then put on an interpretable scale, where 100 is the league average, 101 is one percent above the league average, and 99 is one percent below the league average. I named the metric pBreakout+, as I see its main use as being able to measure which pitchers are bound to break out the following season based on how lucky or unlucky they were the season prior. Excitingly, pBreakout+ is not only more interpretable than existing luck-measuring processes and metrics like FIP-ERA difference, but it also outperformed FIP-ERA difference in its ability to predict change in ERA.

Process and Results

To evaluate how good the pBreakout+ model is at accounting for luck and predicting improvement, I tested its correlation to the difference between pitchers’ ERAs from the 2023 MLB season to the 2024 MLB season. I called this variable ERA Change, where a positive value for ERA Change means a player improved from 2023 to 2024, and a negative one means they got worse. In doing my analysis, I found that pBreakout+ has a moderately strong positive relationship to ERA Change, meaning that as a pitcher’s pBreakout+ value increased, they saw a bigger improvement in their ERA from 2023 to 2024. The correlation between pBreakout + and ERA Change was 0.581 ($p<0.001$), while the correlation between FIP-ERA difference and ERA Change was 0.565 ($p<0.001$). Additionally, I found that as a pitcher’s pBreakout+ gets higher, the percentage of pitchers that see an improvement in their ERA the following season increases as well, with 96% seeing an improvement when their pBreakout+ is 125 or higher.



Conclusion

Overall, the accuracy of pBreakout+’s ability to predict improvement by accounting for luck, along with its outperforming existing measures like FIP-ERA difference leads me to believe that it is possible to quantify the luck players experience on the field in a meaningful, interpretable way. The insights that pBreakout+ can provide will be valuable to teams, players, and fans alike, opening the door to measuring the impact luck is really making on pitchers. It will allow us to see who is destined to regress because of unsustainable good luck, as well as who is destined to improve because of unsustainable bad luck.

THE LAST PAGE: How Publishers are Killing Libraries

By Cameron McLaren

Publishers call it piracy. Librarians call it survival. The 2024 ruling against the Internet Archive didn't just attack digital lending, it declared war on a centuries-old model: ***the library economy.***

In 2024, a federal court ruled in ***Hachette v. Internet Archive*** that the Internet Archive's digital lending was copyright infringement, effectively declaring that libraries can't lend books the way they've done for centuries, just because the pages are now pixels. Publishers celebrated the decision, framing Controlled Digital Lending (CDL) as piracy. But this wasn't just about 'protecting authors'; it was about protecting profits. The court argued that the Internet Archive's book scans were "substitutes" for paid eBooks, despite the Internet Archive owning physical copies of the books they were scanning, and **75% of its collection being out-of-print.** On the other hand, publishers, who make 90% of profits from books in their first five years of print, claimed that CDL "harmed markets." While libraries operate on an ethic of access over ownership, publishers insist that every digital loan requires a **$60 license fee every two years or 26 loans**, turning knowledge into a paywall. With this model, publishers made $14.2 billion in 2023, while libraries spent $1.1 billion on licenses (often for titles they already owned physically). The Internet Archive's CDL bypassed this price-gouging: scans were a one-time payment of $30 per book and were able to serve 1.2 million users in 2023 alone.

This ruling attacks a long-standing principle: libraries operate on shared resources, not profit. CDL was simply the modern extension of this principle. Just as a physical library buys one copy to lend to thousands of people over decades, CDL allows libraries to scan a book and lend it digitally, one copy at a time, just like with physical books. The Internet Archive's National Emergency Library (2020) even waived waitlists during the lockdown period of the COVID-19 pandemic, **emphasizing the importance of access during this time.** But publishers instead saw a threat: if libraries could lend digital books like physical ones, their lucrative licensing model, forcing libraries to repurchase the same eBook over and over again, would collapse.

This battle echoes the landmark ***Authors Guild v. Google*** (2015), where publishers similarly cried piracy **and lost**. Google scanned 25 million books for its searchable database, showing only limited 'snippets' to users. The Second Circuit ruled this 'transformative' fair use, noting it didn't harm book sales. Yet when libraries like the Internet Archive applied this logic to full-text lending, courts suddenly reversed course, privileging publisher profits over public access.

The court's decision exposed a glaring contradiction: under U.S. law, libraries can lend a physical book until it disintegrates (also known as first-sale doctrine), but scanning that same book and lending it digitally is 'theft.' Publishers claim CDL harms authors, yet only **2% of eBook revenue reaches writers**, with the rest going to corporations (e.g. the publishers). Social interests of libraries are at odds with capital interests of publishing companies. The fight for CDL isn't about piracy, it's about whether libraries get to exist in the digital age, or if knowledge will become a commodity controlled by corporations. As 500,000 books vanish from the Internet Archive's website as a result of the court ruling, the fight for CDL becomes more urgent: ***will we defend libraries and their ethic of public access, or surrender the spread of knowledge and resources to corporate control?***

REFERENCES:

Authors Guild v. Google, Inc., 804 F.3d 202 (2d Cir. 2015), https://law.justia.com/cases/federal/appellate-courts/ca2/13-4829/13-4829-2015-10-16.html
Hachette v. Internet Archive. No. 22-816, 2d Cir., 2024, https://law.justia.com/cases/federal/appellate-courts/ca2/23-1260/23-1260-2024-09-04.html
American Library Association. State of America's Libraries Report 2023, https://www.ala.org/sites/default/files/news/content/state-of-americas-libraries-report-2023-web-version.pdf.
Authors Alliance, et al. "Authors Alliance Submits Amicus Brief in Hachette Book Group v. Internet Archive." Authors Alliance, 20 July 2022, www.authorsalliance.org/2022/07/15/authors-alliance-submits-amicus-brief-in-hachette-book-group-v-internet-archive/.
Brooke, Rachel, et al. "Fair Use Week 2023: Looking Back at Google Books Eight Years Later." Authors Alliance, 24 Feb. 2023, www.authorsalliance.org/2023/02/24/fair-use-week-2023-looking-back-at-google-books-eight-years-later/.
"National Emergency Library." Internet Archive, 2020, https://blog.archive.org/national-emergency-library/.
Andrewism. "We Need a Library Economy." YouTube, 05 Oct. 2022, https://www.youtube.com/watch?v=NOYa3YzVtyk
IFLA, International Federation of Library Associations and Institutions, www.ifla.org/wp-content/uploads/2019/05/assets/clm/statements/ifla_position_-_en-_controlled_digital_lending.pdf.
Wikipedia contributors. "Controlled digital lending." Wikipedia, The Free Encyclopedia. Wikipedia, The Free Encyclopedia, 30 Jan. 2025, en.wikipedia.org/wiki/Controlled_digital_lending.

DATA IN PUBLIC HEALTH:

The AI Models Shaping Disaster Response

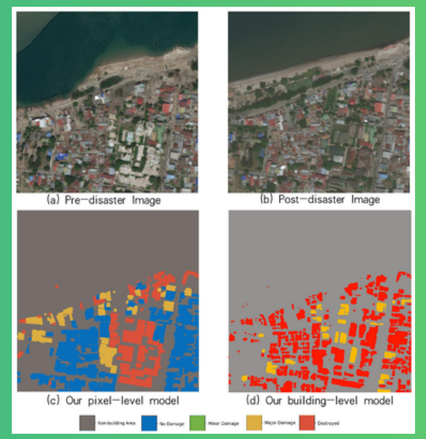
By Sara Vannoni

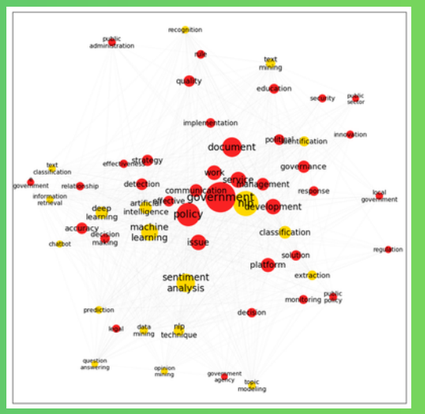
Technology has long been an ally of public health, serving as a lifeline during disaster relief and emergencies. However, traditional technologies often fall short when addressing complex, global challenges. While communication systems like Short Message Service (SMS) may be effective in America, they create significant barriers in areas with limited cellular infrastructure and a lack access to mobile devices. Similarly, advanced hardware, such as drones, are costly and difficult to supply in large quantities. As a result, public health experts have turned to Artificial Intelligence (AI) tools to revolutionize research and disaster response efforts.

Early Warning Systems (EWS) are integrated systems that utilize AI to provide hazard monitoring, forecasting, prediction, and disaster risk assessment. These models are encompassed by three primary components: data collection, data analysis, and decision support, with each process helping ease the transition from data to decision-making. The advantages of these technologies have been significant in resiliency planning and emergency response efforts. By quickly identifying communities at high risk of malnutrition, food insecurity, or famine, these tools allow local institutions to respond quickly with ideal interventions at hand. One of the teams driving these efforts is Insight Commons, a technology consulting firm developing AI to support human decision-making. Using big data management and complex systems modeling, this company has been able to inform both governments and NGOs on how to best prepare for and address disasters, spreading the use of AI globally. Their three primary EW Systems which will be discussed in this article are Predictive Machine Learning Models, Image Recognition Models, and Language Processing Models.

Predictive ML Models are used to forecast patterns and indicate emerging disasters, an especially useful skill in areas where local ground data is scarce. These models primarily use population and crop data to generate high resolution images that estimate crop yield, food security indicators, and affected populations in the wake of natural disasters. However, these models can easily fall victim to the trap of "garbage in, garbage out." When basing their images on old or incorrect data, models have limited effectiveness and can be difficult to interpret. Overall, these models are not yet a substitute for collecting ground data.

Artificial intelligence is also used to run photo recognition diagnostics, both facial and otherwise. By using photo evidence, these systems are able to detect symptoms of malnutrition, identify damaged buildings, and locate route obstacles without physically entering these zones. Moreover, satellite imagery is able to identify specific map features, such as burnt buildings, depopulated areas, livestock mortality, and smoke plumes. Together, this data helps disaster respondents precisely map potential hazard zones and strategically plan ground-level interventions. However, these models are not without flaws. High resolution imagery is expensive and hard to maintain, often due to large quantities of data and hazardous weather. Furthermore, the photo analysis can be error-prone, especially when the analyzed landmarks differ from the images the AI model was trained on.





Natural Language Processing Models are systems that use text from social media and crowd sourcing to identify emerging threats, such as extreme political sentiment and conflict risk. Similar to Predictive ML Models these models are only effective when given high-quality input. Unequal access to the internet may cause LLMs to develop biased findings that may not accurately represent the views of an entire population.

When questioned about the ability of AI models to replace human analysts, Professor Ben Watkins offered an anecdote that discredited the notion. He cites Ethiopia and food insecurity as an example: if AI were to look at Ethiopia now, it would identify a significant lack of food diversity and consumption of dairy/meat products, declaring a watch for food insecurity. In actuality, a human would note that these trends are due to Lent, of which the majority Orthodox population practices. While AI has been instrumental in the improved efficiency of disaster responses, it is not yet a reason to forego human involvement in public health. Public health continues to be a socially-led field, with technology and humans working in tandem to improve the wellness of global communities.

References

https://www.ncbi.nlm.nih.gov/books/NBK218224/
https://www.sciencedirect.com/science/article/pii/S2212420923006313
https://www.insightcommons.com/
ICT and Artificial Intelligence for Sentinel Monitoring and Early Warning Lecture, Ben Watkins
School of Public Health and Tropical Medicine, Tulane University

THE HUMAN COST FOR AI: WHO'S PAYING THE PRICE FOR CHATGPT?

By: Favour Olushola

CONTENT WARNING**: MENTIONS GRAPHIC CONTENT LIKE VIOLENCE AND SEXUAL ASSAULT**

WHAT IS CHATGPT?

Launched in November 2022, ChatGPT–GPT which stands for Generative Pre-trained Transformer–took the world by storm. ChatGPT is a generative Artificial Intelligence (GenAI) chatbot built on a large language model (LLM) that uses specialized algorithms that are trained on large volumes of data from the internet, so it can learn the patterns in human language. Users can give chatGPT prompts or queries and it will use what it has learned to respond within the context of the prompts. So far, this GenAI tool can be used for editing resumes, drafting emails and cover letters, generating software code and helping edit English essays and more. ChatGPT can have conversations about scientific theories or philosophical concepts from the standpoint of a fifth grader to that of a PhD expert. And it's only becoming more advanced. (Hao & Seetharaman, 2023)

As users marvel at ChatGPT's capabilities, we have to ask: who makes this technology "safe" for public use? The answer reveals a disturbing reality behind AI's shiny façade.

THE DIRTY SECRET BEHIND "CLEAN" AI

For ChatGPT to become more advanced and develop more capabilities, it needs to be trained on more data which comes from the internet. Unfortunately, some of this data is incredibly inappropriate, making the AI vulnerable to imitating toxic, offensive and violent human language and content when interacting with its users who vary from young to old demographics. For AI systems to avoid generating toxic, offensive, or violent content, they must be trained to recognize and filter out harmful material. This process isn't automatic or magical—it requires human judgment.
But who performs this grueling work? (Hao & Seetharaman, 2023)

"HUMANS IN THE LOOP": THE INVISIBLE WORKERS
While tech companies promote AI as "human-free machine learning," the reality is far different. Behind every "clean" AI interaction are thousands of workers across the Global South who sift through the internet's darkest content. These "humans in the loop" label, sort, and filter the worst of the data to train and improve the algorithms for AI—for hours every day. They are usually overworked and severely underpaid. This was the fate of Kenyan digital workers employed to perform content moderation for ChatGPT. They spend eight or more hours training AI algorithms to avoid producing hateful, sexually violent or toxic content. These "human in the loop" workers are found in other countries such as the Philippines, Venezuela and India. The large, well educated but unemployed populations in these countries make them targets for huge AI companies who exploit the labor for their tech needs. (Stahl et al, 2024).

KENYA'S AI SWEATSHOPS

Usually these big tech companies like Meta, Google, Microsoft and OpenAI do not employ these workers directly but use American outsourcing companies. OpenAI uses SAMA, an outsourcing company based in California that employs over 3,000 Kenyan workers for AI
content moderation. According to Kenyan civil rights activist Nerima Wako-Ojiwa, these contracts offered to Kenyan workers offer no job stability and conduct workspaces that are run like "AI sweatshops with computers instead of sewing machines". The workers were paid $1.50 - $2 per hour for work (pre-taxes) that would usually be compensated at a rate of $30 or more per hour in the US. Digital employees like Wambalo, Nathan Nkunzimana and Berhane Gebrekidan employed by SAMA were forced to live paycheck to paycheck while being pressured to work faster. Their reward for exceptional performance? "A bottle of soda and KFC chicken. Two pieces. And that is it," said Wambalo.
Many were intimidated into not speaking out as SAMA would fire any dissenting voices. For a lot of them, they had no other options for employment and this was their only source of livelihood and way to provide for themselves and their families. (Stahl et al, 2024)

To make ChatGPT safe for public consumption, the digital workers had to design a detector tool within ChatGPT that could recognize and filter out harmful content. Building this tool required them to review content featuring texts sent by OpenAI that described "situations in graphic detail like child sexual abuse, bestiality, murder, suicide, torture, self harm, and incest.". (Perrigo, 2023).
Consequently, lots of employees struggled with depression, PTSD and other mental health issues that impacted their personal lives and interpersonal relationships with some workers like Wambalo facing marital issues as a result. (Stahl et al, 2024). Others reported persistent nightmares and inability to maintain healthy relationships. And yet, these employees received little to no mental health support for the grueling and grim tasks they were asked to perform which left many severely traumatized. (Stahl et al, 2024).

This phenomenon, termed "outsourcing trauma to the developing world" by Billy Perrigo, a reporter known for exposing repeated labor exploitation by big technology companies, has become a staple for huge AI companies seeking to develop their technology at low financial cost to themselves but at the blatant expense and exploitation of the Global South. (Bartholomew, 2023).

OPENAI'S RESPONSE TO THE ISSUE

When confronted, Sama claimed workers had access to mental health professionals. They eventually terminated their contract with OpenAI after employees raised concerns about illegal content they were forced to review. OpenAI denies responsibility, stating Sama was responsible for worker welfare and compensation. Meanwhile, the workers themselves have been left with lasting trauma and few resources for recovery. (Perrigo, 2023)

WHERE DO WE GO FROM HERE?
Generative AI such as ChatGPT have no doubt provided the modern world with valuable capabilities and possibilities. But this should not come at the cost of human lives.
As consumers, we should demand:

- Transparent labor practices from AI companies
- Fair compensation for all workers in the AI supply chain
- Comprehensive mental health support for content moderators
- Ethical standards that value human wellbeing over profit

These demands also need to be reinforced by strict government policy to regulate the practices of these AI companies and to prevent further exploitation of digital workers and their health.

Sources/References

Bartholomew, J. (2023). Q&A: Uncovering the labor exploitation that powers AI. Retrieved from https://www.cjr.org/tow_center/qa-uncovering-the-labor-exploitation-that-powers-ai.php
Perrigo, B. (2023). OpenAI used Kenyan workers on less than $2 per hour: Exclusive. Retrieved from https://time.com/6247678/openai-chatgpt-kenya-workers/
Stahl, L., Chasan, A., Bar-On, S., & Jung, J. (2024). Kenyan workers with AI jobs thought they had tickets to the future until the grim reality set in. Retrieved from https://www.cbsnews.com/amp/news/ai-work-kenya-exploitation-60-minutes/
Hao, K., Seetharaman, D. (2023). Cleaning Up ChatGPT Takes Heavy Toll on Human Workers. Retrieved from https://www.wsj.com/articles/chatgpt-openai-content-abusive-sexually-explicit-harassment-kenya-workers-on-human-workers-cf191483

Additional Information:

Training AI takes heavy toll on Kenyans working for $2 an hour | 60 Minutes

How big AI companies exploit data workers in Kenya | DW Documentary

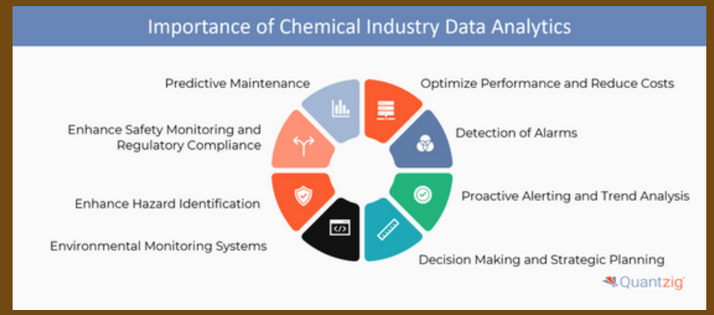
Big Data Analytics in Chemical Engineering

By Nina Moore

In industrial settings, chemical engineers are constantly presented with and assigned the task of analyzing big data. For chemical engineers,' big data isn't just a thing, but a process – including a massive collection of data and the analysis and interpretation of large volumes of data. Big data is defined by the 3V's: variety, velocity, and volume. These 3V's define how chemical processes are "conceived, executed, and refined" (ChemEnggLife, 2024). The analysis of big data helps to optimize processes with great accuracy as well as allow for real time adjustments and the identification of patterns that contribute to increased safety and profit. Many tool such as Tableau, PowerBI, Excel, Machine Learning Models, Jax, or Python are used to analyze such data. Tools such as Tableau and PowerBI are used for larger sets of data to collect, integrate, analyze, and present business information"(Biswal 2023). These tools included interactive data visualization tools that help present key facets of data in a concise and accessible manner. Jax is a Python library and a machine learning framework that enables high level numerical computation. Python is a programming language that can include numerous data analytical packages, such as Matplotlib or NumPy, to analyze and plot data sets. Excel provides data analytics tools such as Pivot tables, lookup, data manipulation, and graphing to analyze data as well. In general, machine learning models are useful in data analytics to help systems learn from, identify patterns, and make predictions using data with minimal human interaction. However, how equitable is the access to powerful data tools, such as these, that aid in the analysis of big data?

Most accessible to most companies is software such as Excel, Tableau, or Power BI. Smaller or midsized companies (SMB's) face specific challenges in trying to take advantage of big data and its endless possibilities. These challenges include poor performance due to outdated and poorly configured data processing systems. The lack in quality data processing systems can cause issues in data cleaning, managing, and structure. These companies additionally lack trained and dedicated IT staff who could be a great help in procuring, configuring, and managing the appropriate software needed. This limitation makes it challenging to implement and maintain advanced data analytics software. Lastly, space is a limiting variable as these SMB's lack the physical space to have a data center to process the data used. Additional barriers also include the cost to own the required software on a yearly basis. Machine learning models that can be specifically trained to be useful to SMB's is costly and require specific computational resources that can be scarce for SMB's.

Thankfully, there seems to be a widespread awareness of these disparities as bigger companies such as IBM have designed software that is easily accessible to smaller and mid-sized companies. IBM's Netezza-based- technology, IBM Integrated Analytics Systems (IAS), is a line of appliances that has been designed to combine high performance hardware and a database query engine designed for massive parallel processing, data warehousing and analytics. They have recently designed a mini version that is perfect for small or midsized companies. It is designed to execute a range of queries while also providing quick results. The IAS additionally requires little administration requirements so that a large administrative staff doesn't need to be allotted to maintaining their data warehouse.



In conclusion, big data analytics tools have immense promise to revolutionize the chemical process engineering sector through optimizing process to enhance safety, selectivity, and profit. Larger corporations with larger capital and qualifies trained professionals are more likely to have the needed resources to acquire the wide range of benefits that comes with being able to effectively analyze big data. However, smaller and mid-sized business face hurdles such as limited infrastructure and capital. To bridge this gap, it's important that industry leaders work together to imitated IBM's initiatives to increase equitable access to big data analytics tools. By prioritizing accessibility to these tools, we can ensure big data becomes a useful tool worldwide- not just for the select few.

References

The Role of Big Data in Optimizing Chemical Engineering Processes | ChemEnggLife." ChemEnggLife, 2024, www.chemengglife.com/ourarticlesandseries/the-role-of-big-data-in-optimizing-chemical-engineering-processes.

Vatter, Holly. "How Small and Mid-Sized Businesses Can Perform Big Data Analytics." Ibm.com, 15 Apr. 2019, www.ibm.com/think/insights/how-small-and-mid-sized-businesses-can-perform-big-data-analytics.

"How Big Data Is Influencing Chemical Manufacturing." Manufacturing.net, May 2015, www.manufacturing.net/home/blog/13159221/how-big-data-is-influencing-chemical-manufacturing

Biswal, Avijeet. "Power BI vs Tableau: Difference and Comparison." Simplilearn.com, 30 Jan. 2023, www.simplilearn.com/tutorials/power-bi-tutorial/power-bi-vs-tableau.

“How Can I Explain This?”

Illustration as a Tool for Science Communication

— Oma Osakwe

For a long time, when I thought about science and art, I was tempted to separate them into distinct categories. The matter-of-fact and evidence-focused culture of science felt ill-matched to the flexible and spontaneous practices that art encourages. However, my involvement in biomedical research has shown me an important opportunity for overlap between these fields: science communication through illustration.

Because communication is an integral part of scientific research, there are several opportunities to share research progress, from lab meetings and honors thesis reports to conference presentations and poster sessions. When preparing to present my work to people unfamiliar with my field, I often wonder, “How can I explain this?” and, consistently, I find that the most effective strategy involves a form of visual depiction. On one summer research poster, I used BioRender to illustrate the methods that I used to investigate the relationship between aging, muscle wasting (cachexia), and muscle nuclear density. In my creative process, I targeted a general audience of cancer researchers and assumed that this group would be familiar with common biology techniques like immunostaining and confocal microscopy. Because muscle biology and cachexia may be unfamiliar areas to some in this group, I focused my design efforts on clearly presenting specialized techniques in a format that my target audience would understand (see Figure 1).

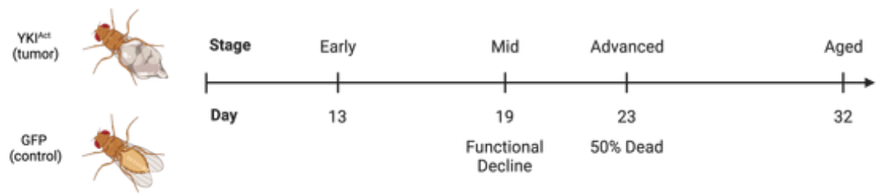


Figure 1: Timeline of cachexia onset. Created using BioRender.

My recent experience with science illustration reflects a larger theme in the science community. As they engage with niche areas of study, researchers typically adopt technical jargon for communication with other members of their field. For example, a cancer biology research article featured in Nature is unlikely to define terms like metastasis and remission because it is assumed that the audience of cancer experts is well-versed with relevant terminology. While this practice is acceptable within focused scientific communities, the use of academic jargon and text-heavy content presents barriers to communication with a lay audience, including other scientists who hold a different focus (Perra and Brinkman, 2021). This challenge may stem from insufficient training in effective communication strategies. To address such gaps in professional development among scientists, Perra and Brinkman (2021) provide a step-by-step guide to creating useful infographics (see Figure 2).

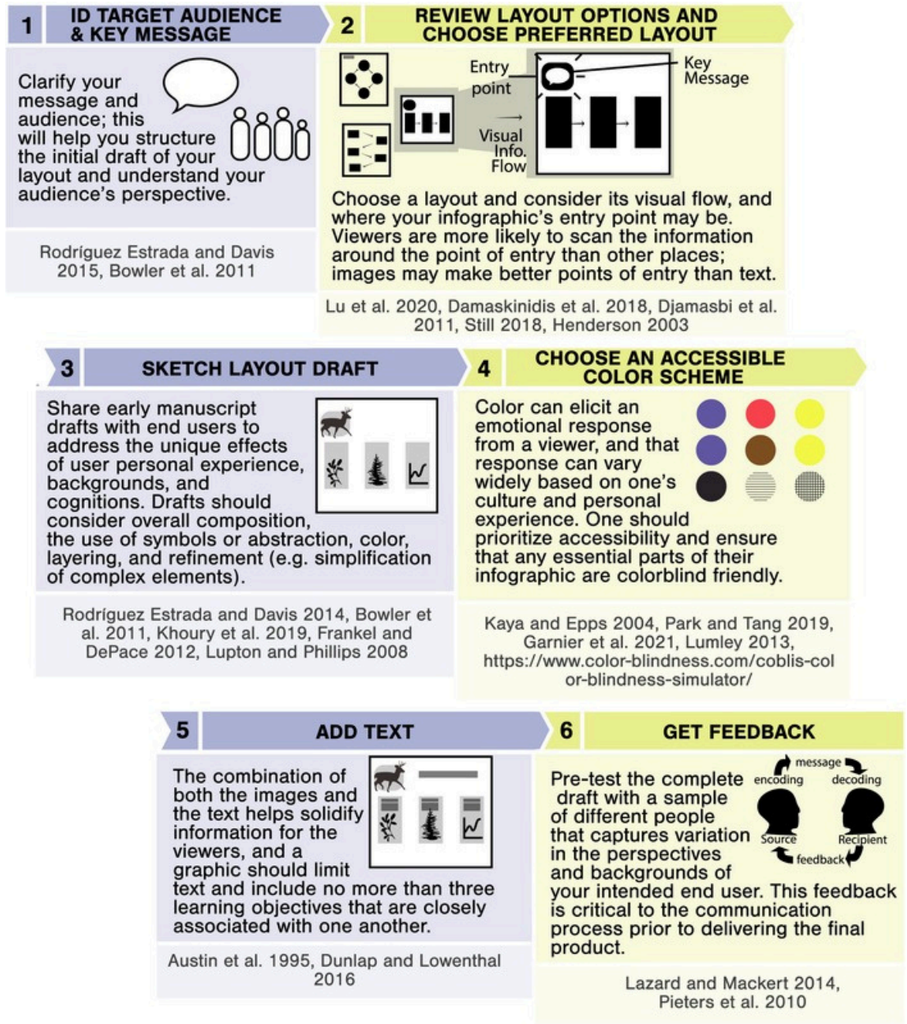


Figure 2: Guide to creating an effective infographic. From Perra and Brinkman (2021).

Beyond conceptual knowledge of best practices for illustration, Heidt (2024) recognizes that some scientists may struggle with the artistic skills that such data presentation requires. That is, even though a researcher has identified their target audience and message, and even planned out a potential layout for their infographic, the task of depicting organisms and tools accurately and attractively in design packages like Adobe Illustrator can be daunting. To the benefit of the science community, however, there are several tools available that address this very problem. Free programs such as Servier Medical Art, BioIcons, and Mind the Graph provide a catalogue of common science icons, allowing researchers to quickly and easily create schematics. Further, members of the Tulane community are also included in the institutional BioRender Premium license, which provides extensive illustrative capacities.

References:
Heidt, Amanda. 2024. Scientific figures that pop: resources for the artistically challenged. *Nature* 634: 1248-1249. https://www.nature.com/articles/d41586-024-03477-0
Perra, M., and T. Brinkman. 2021. Seeing science: using graphics to communicate research. *Ecosphere* 12(10):e03786. https://esajournals.onlinelibrary.wiley.com/doi/10.1002/ecs2.3786#
Tulane University Information Technology. (n.d.). *BioRender*. https://it.tulane.edu/biorender

Samuel Johnson

The consistent expansion of interest in the Artificial Intelligence sector has led to increased investment from government bodies, such as the military and its private sector. There are many debates regarding whether or not an intelligence apparatus, such as AI, should be integrated within our nation's military. Possible positive qualities that AI offers the United States servicemen and women are enhanced cybersecurity measures, increased efficiency, and superior surveillance and reconnaissance drones. The cybersecurity realm is vital in today's battlefield. Foreign adversaries, like China and Russia, have increased their military spending on advancements in cybersecurity integrated with AI. This poses a massive threat to the United States' critical infrastructure network, which has been incredibly vulnerable in recent years. North Korea and China have grown interested in testing their cyber-attack capabilities. In 2014, a North Korean hacking organization named the “Guardians of Peace” carried out one of the most extensive hacking operations on U.S. soil. The hackers have claimed that they stole one hundred terabytes of data from Sony, which, while most likely exaggerated, is still a massive issue. Advancements in AI could prevent future attacks like the Sony attack from occurring. Increased efficiency is another natural byproduct of AI. More effective and timely responses could be produced by broadly expanding the AI sector of the military and intelligence agencies. This efficiency aspect would also aid in more capable drones. Drone warfare is the second most prominent military advancement in the twenty-first century, behind cybersecurity.

For the last twenty years, the United States has led the globe in advancements in drone warfare and intelligence gathering. However, with the knowledge gained from the Russo-Ukrainian war, cheap, one-use attack drones have become more effective, especially in destroying or disabling armored fighting vehicles (AFVs). This has transformed the modern battlefield tremendously. Now, modern militaries are focusing on advancing these attack drones in mass waves, thus making AFVs less effective on the battlefield, endangering the servicemen and women who use them. By investing in AI, the United States military could decrease spending on AFVs while promoting spending on cheaper, more effective drones.

With the integration of AI into the military, there will be a natural backlash. Many critics cite that AI carries the enormous weight of ethical concerns, reliance, escalation of war, and loss of human control. The concept of artificial intelligence has been criticized since its beginning. Movies such as The Terminator have altered the public's perception of AI. These sentiments have carried into the twenty-first century and stalled AI advancements. Over-reliance is another driving factor for critics of AI. The feeling of losing control over something so powerful, primarily non-human, is a rightful fear. Control is a driving feeling in human nature. Losing control over an artificial intelligence program could lead to a decrease in human ingenuity or humanity. To combat this, special regulations and regulatory powers must be installed. Government oversight and transparency with the public would be essential for this solution. The gradual escalation of war would also be a negative consequence of an increased AI investment in the military. Foreign adversaries, such as Iran and North Korea, have already begun developments in their artificial intelligence sectors. This could lead to a new form of cold war, one with the threat of nuclear war and artificial intelligence. The loss of human control over AI is one of my own personal fears. Once we rely too much on AI and its capabilities, human thought and ingenuity will eventually stagnate, losing what truly makes us human. Another fear of losing control over AI is the possibility of a critical and costly decision like a nuclear strike, or a different form of attack could occur.

All of these reasons, whether for or against artificial intelligence in the military, should be considered. While many of the benefits are present, the side effects remain in the minds of the public and the government. However, with the current geopolitical climate, I could envision a massive push toward enhancing AI within our military for offensive and defensive measures. Already, companies like Lockheed Martin and other military contracting firms have been investing in AI capabilities like swarm technology.

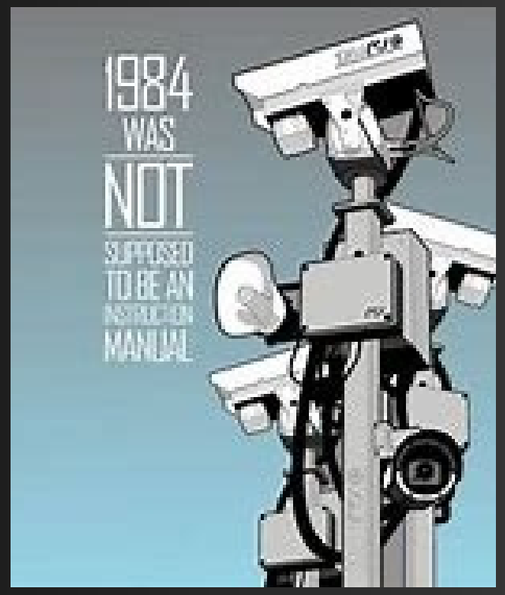
Sources:
Jiang, Ben, and Ben Jiang. “Tech War: China Creates US$8.2 Billion AI Investment Fund Amid Tightened US Trade Controls.” South China Morning Post, 20 Jan. 2025, www.scmp.com/tech/big-tech/article/3295513/tech-war-china-creates-us82-billion-ai-investment-fund-amid-tightened-us-trade-controls.
Lapham, Jake. North Korea’s Kim Jong Un Inspects AI “suicide Attack Drones.” 27 Mar. 2025, www.bbc.com/news/articles/c4gmny262vzo.

BIG BROTHER WATCHING

Encoded Inequality: How AI Surveillance Systems Reinforce Racial

By: Kai Williams

What happens when the technology designed to keep us safe inherits the same biases that have historically marginalized communities? Or when this so-called safe technology dismisses someone because their skin has less melanin, deeming them safer than those of the shades darker. In history, we have watched these marginalizing systems implemented socially, spreading like a virus into everyday life, such as government acts, medical research, and even into the tech world now. It's not that the marginalizing systems are deemed logically correct. It is more that people are socially influenced over time that these systems are their norms, and it engraves itself into the back of their minds hard, coding them to believe in these social norms and not combat them. This made for trillions of data points around the internet to inherit data that was immersed in these marginalizing systems, allowing for it to learn from that data and create its analysis only built on the knowledge of the data within the internet; this doesn't include the social implications involved in societal behavior or economic battles of today allowing for the Ai to build a system with uneducated and sympathizing systems; that don't know the battles of what it is to be a human in a society of dominant and marginalized groups. So, when it comes to AI surveillance systems, they reflect and reinforce racial biases built on historical and structural inequalities, which become embedded into their training data, posing ethical and societal risks.



First, I'll start with how AI identifies individuals; you likely have a phone right by you that is capable of identifying you based on your facial features, the shape of your face from the distance of the eyes or from the nose, the contour of the jawline, and the pigment of your skin, which are all unique to the individual. Still, an AI looks at your face, turning it into pixels and data points. This measurement of pixels is then transformed into electronic code, something similar to a fingerprint. The code of pixels is then matched to a database to match a face and the code to an individual. Fingerprints or face prints are created exponentially within organizations like the FBI, which stores over 650 million face images(Inapp Inc., 2023). These databases are a huge part of how AI identifies individuals. According to Mitchell (2019), using machine learning algorithms, the system recognizes patterns that define facial features to recognize the code make-up of someone's face and then match it to an individual. Machine learning is done through databases, including those done on individuals' faces hidden as ones and zeros. Like a baby, AI has to see multiple cases of individuals' faces to improve its efficiency in identifying individuals. The more technical process would be the AI system taking raw data, say pixels in an image, to recognize patterns and make decisions. Deep learning algorithms carry out this training, but these systems aren't perfect to begin with; they need to refine themselves, and when they do get identifications or classifications wrong, they go through a process called backpropagation, which modifies their calculations to create a more perfected AI system. These simple integrations would put you in a good place to create a system that can identify an individual, but there is a big question. If these systems can access tillions of data points to refine their accuracy, why do they still experience mishaps, often disproportionately affecting minority groups?

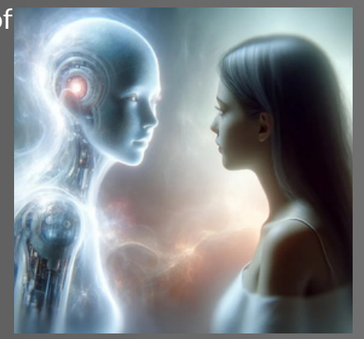
To answer this question, we must confront the reality that these mishaps occur due to biased training data, unequal representation, and algorithms shaped on socially constructed systems. The AI systems learn and are adapted to such systems; they are then inevitable to reflect existing social disparities in society. These systems learn like toddlers; they take the information for face value, believing it is correct. So, now imagine this toddler-like system is given criminal justice data from the 19th century; it would result in AI systems unfairly predicting higher recidivism rates for minority groups without reviewing social issues, economic misfortunes, and environmental pressures. This means that Ai would then fall into pseudo-scientific justifications for the racial inferiority of individuals, similar to the idea discussed by Winston Zhou (2022), where the individual delve into the human psyche reveals implicit bias: a subconscious force

that enables prejudice to develop even with the conscious attempt to remain unprejudiced—creating a system of one master race and inferior races. When this happens, racial biases are made, where individuals begin to be flagged as threats because of the color of their skin at a highly disproportionate rate. For example, Pro Publica created a statistical analysis controlling for factors like criminal history, recidivism, age, and gender and found that Black defendants were still significantly more likely to be classified as high-risk. They were 77% more likely to be labeled as at risk for future violent crimes and 45% more likely to be predicted to commit any type of crime(Angwin, Larson, Mattu, & Kirchner, 2016). This means that the AI systems deemed the individuals a risk without logic behind their reasoning, meaning they adopted aspects of socially created concepts that we are now using to come to conclusions about the fate of individuals' lives. People of color have to fight off these aspects in the human world and now AI, creating new boundaries for how racial biases can affect individuals.

Similar to predictive AI surveillance of AI facial recognition, these methods can also be used in many different ways you may see in everyday life, like Walmart's self-checkout systems that use AI surveillance to detect shoplifting at the checkout. Or the terrifying AI surveillance China uses to track citizens from their faces and license plates to car models all seamlessly. Both examples highlight the significant roles AI can play regarding surveillance and monitoring. However, the issue is when the AI collaborates pre-existing bias into its daily task that reinforces harmful stereotypes, as these systems are trained on biased data. It will encourage that people of color are more threatening or more likely to be associated with criminal activity. This will result in the likelihood of false positives because of the concentration and overanalyzing of race rather than character, further disproportionately marginalizing communities. As Christina Carrera (2023) explains, facial recognition technology tends to capture the unconscious biases of its designers, who are predominantly white men. Because of this, the programs are more accurate at recognizing faces similar to those of their creators and not so with people of color. This disparity is further compounded when AI cross-matches faces against databases that contain disproportionately high concentrations of Black and brown faces, increasing misidentification and wrongful arrests. So, how do we combat this? As of right now, we can't rely on AI to make unbiased decisions because these systems are built on data shaped by human thoughts and social norms.

In a way, these systems evaluate us as a human race. Regardless of the rules we implement, AI is inept at consensus on what is right or wrong. In all, AI is designed to complete the tasks set by its makers, reflecting the bias within the world in which we live. I believe, as of right now, we can't because AI is built on the data of humans who have their own thoughts and social standards; no matter what rules we create, these systems don't have a consensus to say what is right or wrong, only the making the right decision on the task at hand or what their maker wants them to do.

In conclusion, AI surveillance systems do not exist in a vacuum; they are shaped and learn from the data they analyze, including the biases embedded in the data. While these systems can be used in a multitude of ways they do inherently pose risk of displaying the bias within the data they analyze, from systems like facial recognition software to predictive policing, AI often amplifies preexisting inequalities based on the perceived notion that it is typically a specific demographic, causing a situation varying depending on the task, resulting in disproportionately targeting Black and Brown individuals. These biases are beyond merely technical errors; to mitigate harm from biased data, it is essential to start creating regulations that increase transparency, representation in AI research, and ethical accountability, because without inherently teaching AI to comprehend the complexities of human experiences far beyond the data points, AI will continue replicating societal prejudice. We must remember that AI reflects the world we have created, which has deeply rooted social and historical patterns; it is not intrinsically biased or prejudiced. Before creating an equitable AI system, we must rebuild our foundation and confront the biases within the world; without that, change won't happen, resulting in the marginalization and condemning of these groups.



References:
Inapp, Inc. (2023). How artificial intelligence detects faces. Medium. https://inapp-inc.medium.com/how-artificial-intelligence-detects-faces-a85e8704fd78
Mitchell, M. (2019). AI: A guide for thinking humans (pp. 67-80). Farrar, Straus, and Giroux.
Zhou, W. (2022). The pseudo-science of race and ethnicity. University of California, Irvine Ethics Center. https://www.ethicscenter.uci.edu/files/docs/fabulous_mentoring/2022/03%20The%20Pseudo-Science%20of%20Race%20and%20Ethnicity.pdf
Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016, May 23). Machine bias: Risk assessments in criminal sentencing. ProPublica. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing
Carrega, C. (2023, September 14). Facial recognition technology and false arrests: Should Black people worry? Capital B News. https://capitalbnews.org/facial-recognition-wrongful-arrests/

2024-2025

Peer Mentors

CAIDS Data Peer Mentors act as teaching assistants, guiding fellow students who take data courses at CAIDS by exploring and troubleshooting data questions and problems during support sessions. Peer mentors help shape positive and supportive learning communities around data.

Q&A with the Peer Mentors

JR AR LM

Joe Raphael, Adel Rahman, Lily Markus

Talk a little bit about your role and how you served others.

JR

I am a freshman majoring in Economics with a potential minor in Design. Through my interactions, I hope to have sparked greater interest and engagement in the class, making the learning experience more enjoyable. I have consistently supported students by holding regular office hours, which have seen strong attendance. By offering a less formal environment for students to understand the material, I've aimed to create a space where they feel more comfortable seeking help and deepening their understanding of the course content.

What did you learn from this experience?

LM

Being a peer mentor has taught me to be flexible, think on my feet, and embrace diversity even further. It has challenged me to continue growing and learning about data, equity, and
pedagogy. This role has brought to light to importance of valuing different kinds of intelligences and knowledge.

I help students with problem sets relating to Artificial Intelligence and devise alternative resources on canvas to help guide students through Google Colab (Python).

AR

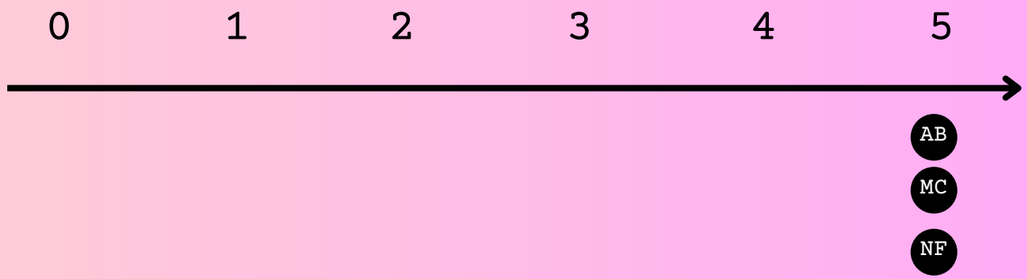
How did you contribute to a culture of data literacy and data science?

I've thoroughly enjoyed engaging with students both academically and personally. One challenge is that sometimes I am put in a position where I can't answer a specific question. It can
be awkward but to overcome this I simply admit that I can't answer there question and help them as best I can.

JR

What were some wins and some challenges?

How likely are you to recommend this role to someone?



Student Group

Especial Edition

TULANE DATA COLLECTIVE

A NEW CLUB AT TULANE FOR STUDENTS INTERESTED IN DATA

TULANE DATA COLLECTIVE

The Tulane Data Collective (TUDC) is an interdisciplinary student organization that fosters a collaborative environment for Tulane students to explore and develop skills in data science, data literacy, and career readiness. By offering workshops, hands-on projects, and networking events, TUDC aims to bridge the gap between academic learning and practical application in the rapidly evolving world of data science. Members will have opportunities to participate in peer-led initiatives, exclusive recruiter events, and student-run hackathons, enhancing their technical expertise and professional networks.

@TULANEDATACOLLECTIVE

BIG APPLE, BIG DATA: POWER & POLITICS EVENT APRIL 23RD

Join the TUDC for our guest speaker event hosted April 23rd. New York State assemblyman, Steve Stern, and former State senator, Craig Johnson discuss the intersection of politics and data.

FOLLOW US ON INSTAGRAM, OR CONNECT WITH US ON LINKEDIN

Founding Members (from left to right): Ananya Anand (VP of Research), Ifeoma Osakwe (VP of Finance), Liam Guest (President), Sara Vannoni (VP of Internal Affairs)

TUDC 2025 E-Board

President: Liam Guest: As president, Liam is responsible for presiding over all E-board and general body meetings and oversees all TUDC initiatives and facilitates growth.

VP of Research: Ananya Anand: Ananya creates and manages all data science initiatives. She also is responsible for developing the logistical aspects of large-scale data events, like Hackathons and Datathons.

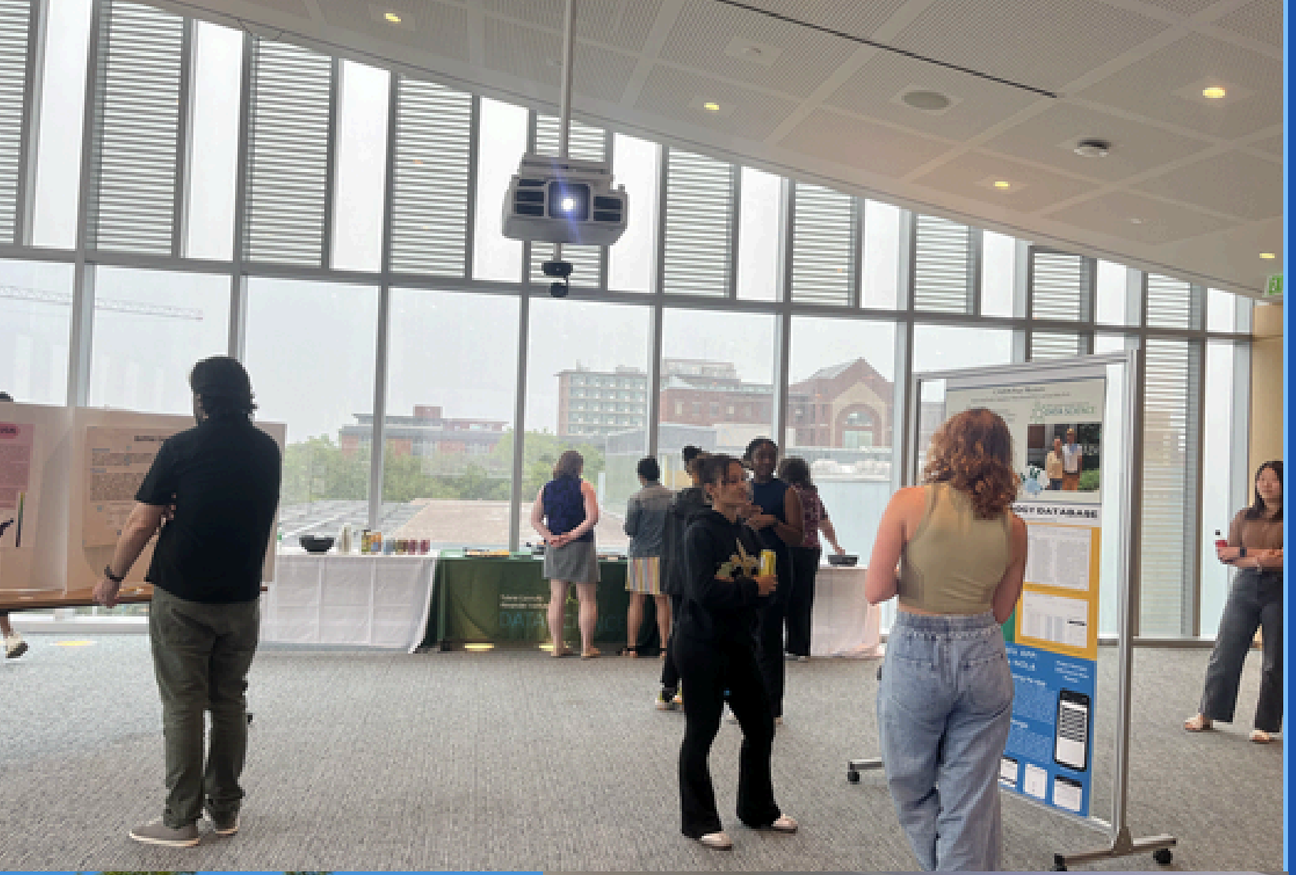
Research Committee: Julian Cohan: Julian is responsible for assisting our VP of Research on all research endeavors.

VP of External Affairs: Samuel Johnson: Sam coordinates networking opportunities with project partners, partner student organizations, guest speakers, and funding partners.

VP of Internal Affairs: Sara Vannoni: As VPI, she manages TUDC's branding strategy, collaborates with other data science clubs, and leads internal communications.

Internal Affairs Committee: Anh Pham: Anh is responsible for working with our VPI on all internal endeavors. Anh also has taken a significant role in the creation of our webstie.

VP of Finance: Ifeome Osakwe: Receives and manages all organizational funds. Maintains accurate financial records by tracking receipts and expenditures.



DATA LAB NEWSLETTER

The CAIDS Data Digest is a monthly newsletter that features upcoming events, internship and grant opportunities, course offerings, and student spotlights.

Scan the QR code to join!