

DEPAUL CENTER FOR DATA SCIENCE

FALL 2021
NEWSLETTER

WELCOME BACK
TO CAMPUS
Fall 2021

In this issue
New Director for CDS
Upcoming events
Spring and Summer Activities
Get to know DS students

New Director of DePaul Center For Data Science

On behalf of the DePaul Center for Data Science, we would like to congratulate Ilyas Ustun on becoming the new Director of the DePaul Center for Data Science – CDS



Let's get to know how Ilyas envisions DePaul Center for Data Science (CDS) to grow within the coming years and How can students take advantage of this Center?

"A short while after I started at DePaul, the Covid-19 pandemic started. But, even under those circumstances, we didn't stop our activities. With guest speakers, research presentations, heck, even a hackathon, we still continued to come together - virtually. We collaborated in several data science projects with different organizations such as Chicago city, Argonne National Lab, UtmmostU, etc. Chicago is a vibrant city in terms of opportunities for data science, both in academia and the industry, so I feel lucky in that regard. We already have some companies in the pipeline and will announce the new projects shortly. I envision CDS to grow even further, with more data science projects, new collaborations with industry that will provide internship and job opportunities for students, and other social events. My advice to students is to take part in these events and not to be shy to share their ideas with us."

You can contact Dr. Ilyas Ustun by email at iustun@depaul.edu, follow him on Twitter [@_ilyas_ustun_](https://twitter.com/_ilyas_ustun_) and connect with him on LinkedIn at [ilyasustun](https://www.linkedin.com/in/ilyasustun).

INTERVIEW WITH ILYAS USTUN

What is your 30-second bio?

I am a data enthusiast and have over 10 years of data analytics and predictive modeling experience in various fields, including business analytics, transportation analytics, sensory data analytics, digital agriculture, health analytics, medical informatics, and metabolomics. I got my Bachelor's in Industrial Engineering with summa cum laude and in my Ph.D. I focused on transportation analytics. I enjoy attending data analytics conferences and meetups. In my spare time, I love hanging out with my kids, swimming, and reading books.

Was there a specific "aha" moment when you realized the power of data?

When I started my Ph.D., there was absolutely one thing I didn't want to do: Programming. I both loved it and hated it. I guess my experience during undergrad wasn't that good. But, the machine learning course I took (thanks to my advisor Dr. Cetin's pressure) changed my mind. Although I suffered quite a lot in the course because I had minimal coding experience, it opened a new world for me. It was just fascinating to see a machine learning model make predictions without any explicit programming. I guess that moment made me understand the power of data, machine learning, and coding.

How did you come to work as the Director of DePaul Center for Data Science?

I have been a straight-A student all my life. Starting from elementary school, I was always at the top of my class. I finished my undergrad with summa cum laude, and of course, just wanted to continue my studies – after all that was the only thing I knew 😊. However, one thing I wasn't sure of was what to do after my Ph.D., the supposedly end to "studying" (Believe me I still study!). The data analytics-oriented research made me realize that I enjoyed data. So, I decided to work as a data scientist after my Ph.D. After some industry experience, I went back to academia as a researcher. Now, I am at DePaul and am a proud member of the Data Science program here. After a year of learning as Associate Director under Director Dr. Daniela Stan Raicu, I am now pleased to carry the baton forward. I am very grateful to everyone who supported me in this journey and placed their trust in me.

What do you find challenging and exciting about working as a Professor and the Director of CDS at DePaul?

Teaching classes is not an easy task. On top of that, trying to create new collaborations, create opportunities, and be active in the community makes it a challenging job. But I love challenges, it makes me stay on top of my game.

I enjoy different data analytics projects. Each of them teaches me a new thing, shows me a different aspect of the world. So the most exciting thing for me to be a professor at DePaul and the Director for CDS is to be involved in several unique projects, collaborating with different people from different backgrounds, and establishing lifetime personal and institutional connections. So, the job is challenging but rewarding, and that makes it very exciting!

Any words of wisdom for Data Science/Machine Learning students or practitioners starting?

I feel many new starters think that data science is all about coding. Although being an important part, coding is not everything. We need to understand what we are doing. So, although we don't need to know every single detail of an algorithm, we need at least a working knowledge of the basics, how it works, what it is used for, the input requirements, etc. Otherwise, we can be perfect coders, and all we would be able to do is to run models and compare their results. That's not being a data scientist – that's a machine! In short, to be good data scientists, we have to have a balance of theory (statistics, maths), domain knowledge, and coding. This is also what I do in my classes; they are hands-on with enough level of theory to understand the procedures.



HACK4SPACE ONLINE HACKATHON

In collaboration with IBM Z, the DePaul Center for Data Science held Hack4Space, a virtual hackathon where teams of students leveraged NASA data to identify a problem and propose a solution. More than 120 DePaul students from CDM, CSH, and Business registered for the event, and about 20 teams worked on a project during the weekend of May 7-10. The kickoff event featured keynote speakers from NASA and IBM. The winning team is being considered for internship opportunities at Hypercubes, a startup building the next generation of Earth observation satellites enabled by artificial intelligence.

THE WINNING TEAM PARTICIPANTS:



Claoe Louis



Anirudh Edupuganti



Manoj Nagendra



Vi Nguyen



Huy Tran



Beatrice Vaccari

The award ceremony event was online, and the coordination between teams was done via Slack.

We are very grateful to have the following valuable speakers to deliver their speeches during the Awards Ceremony of the Hack4Space.

- *Kevin Murphy, Chief Data Scientist from NASA
- *Michaela Musilova, Director of HI-SEAS, International Moonbase Alliance (IMA)
- *Naeem Altaf, Distinguished Engineer, CTO – Space Tech from IBM
- *Melissa Sassi, Chief Penguin, IBM Z, Student & Entrepreneur Experience



Interview with Melissa Sassi

The Chief Penguin of IBM z

Melissa Sassi, Chief Penguin of IBM Z, leads the Student & Entrepreneur Experience worldwide at IBM. She also created and leads IBM Hyper Protect Accelerator with 45 startups in her portfolio that are expected to grow to 100 early-stage startups by the end of 2021. Melissa is also the Founder & CEO of MentorNations, a youth-led digital skills movement that has taught tens of thousands of young people to code across twelve countries. Melissa Sassi was instrumental in coordinating Hack4Space, a datathon sponsored by the DePaul Center for Data Science and IBM Z.

The following summarizes the interview that Melissa gave to Holly Genvy, DePaul student and Vice President of events for DePaul PRSSA. Melissa and Holly discussed Melissa's unique career at IBM and her work in helping communities worldwide.



I came over from Microsoft and I had a colleague there that called herself the chief ninja cat and then I stumbled upon this guy who had this twitter account called threading the t-rex and I was like I need to have a fun animal title

Melissa came up with her title of Chief Penguin of IBM when she went for a short trip to Pakistan, where she was inspired by someone and thought of this fun and crazy name, the penguin. She explained that the penguin is the Linux Red Hat brand character and, therefore a symbol of open source technology, and that's her thing.

Melissa's interest in technology was sparked by realizing that many people worldwide lack access to computers and do not have essential digital skills. She talked about her work teaching code to children in Tunisia, and the success of this initiative has inspired her to expand it to many other countries around the world. Melissa shared her deep concern about how young people from underserved communities are left behind because they don't have the basic technical skills to be online and to know how to transform their lives through technology.

Melissa has run several hackathons while also being the CEO and founder of her own non-profit called MentorNations. She created a robotics lab, an internet of things lab, and a co-working space in North Africa and in Tunisia; she and her team have taught tens of thousands of young people to code. At the same time, she discussed how students could change the world and accomplish things with data that had not been accomplished before. She also talked about the importance of participating in hackathons. Students use their entrepreneurial skills and critical thinking to solve problems while learning how to work effectively in a team. These are essential skills that will help them in their future careers and professional development.

[Click here to watch the full interview](#)



Events & Accomplishments at CDS

- During the summer of 2021, under the supervision of Prof. Ilyas Ustun, DePaul graduate students Adam Hecktman, Maggie Wolff, Mengfan Ying, and Jared Hebda have started working on a predictive modeling project in collaboration with UtmostU. UtmostU is a Chicago-based alliance of high schools, community-based organizations, universities, community colleges, and workforce development organizations committed to post-secondary success. The DePaul team developed a predictive model for young adult success for UtmostU. The available data include student grades, credits earned, graduation results and the use of UtmostU support programs like persistence funds, credit-recovery classes, and T-Mobile lines. The DePaul team also provides insight into predictive data points and helping UtmostU develop a clear framework for collecting data in the future. The recommendations will improve the support of our affiliate high schools and community organizations and strengthen predictive models for young adult success.
- Associate Professor Tanu Malik was named a 2021 *Researcher to Know* by the Illinois Science & Technology Coalition.
- Earlier this summer, Ilyas Ustun participated in an online hackathon to make the city of Eindhoven in the Netherlands livable, safe, and healthier. Along with his two teammates, he participated in one of the three specific challenges, "Make Eindhoven Healthier," with the idea to use incentives to decrease electricity usage and create monitoring tools to see each neighborhood's carbon footprints within the city increase awareness and increase participation. The team finished in first place.
- INFORMS Journal on Computing has given its *Test of Time Paper* award to "A Framework for the Evaluation of Session Reconstruction Heuristics in Web-Usage Analysis," co-authored by Bamshad Mobasher in 2002. The paper became a seminal paper that helped establish Web Usage Mining as an area of research. Many of the techniques discussed in the paper are still integrated with software tools used for Web and social media analytics.
- On 1st June, 2021 the Data Science Alumni Speaker series coordinated by Sarah Primozich, Assistant Director of Online Student Services, hosted Yuxuan (Tim) Zhang, Data Scientist at Wayfair.
- Professor Daniela Raicu is the Co-PI on the \$170,287 NSF-funded project "Collaborative Research: HDR DSC: The Metropolitan Chicago Data Science Corps (MCDSC): Learning from Data to Support Communities." CSH faculty Mark Potosnak is the PI and Sungsoon Hwang (LAS), and Philip Yates (CSH) are also Co-PIs.



The Motorola Solutions Grant

The Center for Data Science is pleased to have finalized the Motorola Solutions Scholars Program Grant. Ilyas Ustun is the principal investigator, and additional participating faculty include Daniela Raicu, Jacob Furst, Roselyne Tchoua, and Thiru Ramaraj. The total reward amount is \$35,000. The project will run from the Winter of 21/22 until the end of Fall 22/23.

The Motorola Solutions Scholars Program makes a deep investment in a diverse data science workforce by immersing five underrepresented sophomores in a rigorous research program at the DePaul Center for Data Science. With one-on-one mentoring, students carry out a 12-month data-driven project, gaining technical proficiency, academic experience, and 21st century skills. Coaching from dedicated DePaul faculty and Motorola Solutions volunteers helps students identify career pathways and see themselves as data scientists.

At the core of the program model is a paid research position providing the framework for ongoing activities that are designed to promote a sense of belongingness, cultivate awareness of data science career pathways, and increase the capacity of students to attain data science degrees and pursue rewarding careers in this high-demand field. The program will achieve these outcomes by increasing access to near-peer role models, faculty mentors and industry coaches and by providing academic research opportunities.

Recruitment and Scholar selection will be conducted in the 2021 Fall Quarter. Programming will take place over three subsequent academic quarters: Winter 2022, Spring 2022, and Fall 2022. Beginning with the 2022 Winter Quarter, each Scholar will be awarded a research position in the DePaul Center for Data Science, an interdisciplinary center bringing together DePaul faculty in data science, machine learning, and AI. The Center's educational mission is to nurture the next generation of data scientists to better prepare them for data-related computing careers. In keeping with DePaul University's social justice mission, Center faculty and students leverage the power of data science to address pressing social issues in Chicago and beyond.

In the final quarter of the project, the 2022 Fall Quarter, students will prepare and deliver presentations on their research findings at a special Motorola Solutions Scholars event and will prepare and submit an abstract on their academic work to the National Conference for Undergraduate Research, a forum to celebrate and promote undergraduate student achievement.



CDS RESEARCH HIGHLIGHTS

A DePaul undergraduate student, Maria Alexandra presented her research at the National Council for Undergraduate Research – NCUR 2021:

Human-in-the-loop Clustering Dashboard for Materials Structure Exploration

Maria Alexandra Theodorescu, Prof. Daniela Raicu, Prof. Jacob Furst and Prof. Roselyne Tchoua

The team described their research here:

“Data science is intrinsically inter-disciplinary; however, end-users of machine learning models are not always trained data scientists. On the other hand, it is crucial that these models be infused with domain knowledge in order to increase explainability and trust in their output. Our ultimate goal is to assign domain-aware confidence scores to help domain experts make informed decisions. Our hypothesis is that given confidence scores, end-users will be more willing to trust and adopt machine learning models. We test this hypothesis with materials informatics, a field that has the potential to greatly reduce time-to-market and development costs for new materials as it leverages machine learning and large datasets for targeted design. For example, automated phase-mapping seeks to discover samples of materials mixture with similar structure. This is challenging because measurements per sample far exceed the number of samples to cluster making it difficult to interpret and generalize. Towards our goal, we are building a dashboard comparing and contrasting clustering methods. We envision that scientists will not only be able to assess confidence scores but also interact with results; merging and splitting clusters, guiding the discovery process. We describe the signals in terms of peaks and other interpretable features; we cluster the data using K-Means with varied numbers of clusters. We provide several visualization options (e.g. layered graphs, samples closest to and farthest from centroids). Our preliminary results show a number of fully or mostly homogeneous clusters (using ground truth labels), discovering well-defined clusters with a fraction of the original features (28 out of 2000 or 1.4%). As we repeat the experiment with more clusters, we identify consistently homogeneous clusters as well as larger clusters, candidates for further splitting. We will experiment with different clustering methods comparing the performance using our features, the original and other reduced features (e.g. PCA).”



**GET TO
KNOW OUR
DATA SCIENCE
STUDENTS**



Mark F. Haskins

a graduate student of the MS in Data Science –
Computational Methods concentration

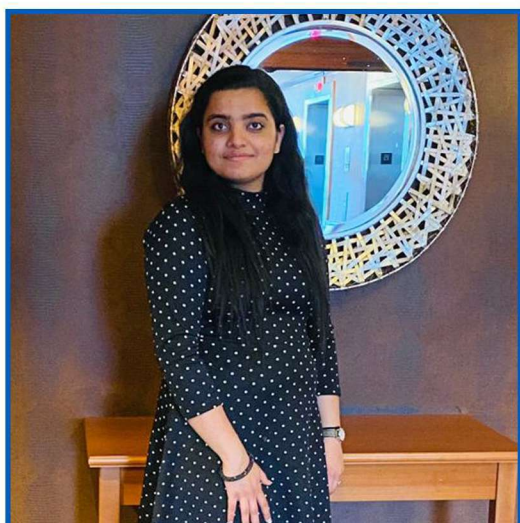
I am currently a remote student in Dallas, TX taking all online classes. I am about to complete my third class. I started the program at the onset of the COVID19 pandemic. Thus far, I have been very impressed with the level of engagement of the Data Science Program and the CDM community as a whole. The Data Science Center reaches out on various opportunities including jobs, projects.

2) Tell us about the course that you found most inspiring and impactful for your future career in data science.

I have only taken a limited number of courses, but the most inspiring was DSC430 Python Programming. The workload and learning curve was steep in some areas. I work full-time so it was a lot of sleepless night and stressful chain-smoking, but it gave me a high level of satisfaction. I really felt like I earned my grade. I learned a lot. Some of it I even applied to my current job to help automate some manual spreadsheet manipulation, so it was a win-win.

3) Why did you choose to pursue a degree in data science and what are your career goals?

Money! Any job search on Indeed, LinkedIn, or Monster will show you lots of Data Science jobs and only a handful of qualified applicants. The demand for these skills and salaries keeps going up, and I don't see any reason why it won't continue. Everything from the toaster to your watch is generating data now. I chose DePaul specifically because most of the course names align directly to the kinds of keywords these jobs are looking for: SQL, Python, R, etc. My career goals will continue to be anything that keeps my hands on the keyboard. It's very fulfilling building things that solve problems. I am most interested in the data engineering/data wrangling part of data science. Virtual plumbing, so to speak.



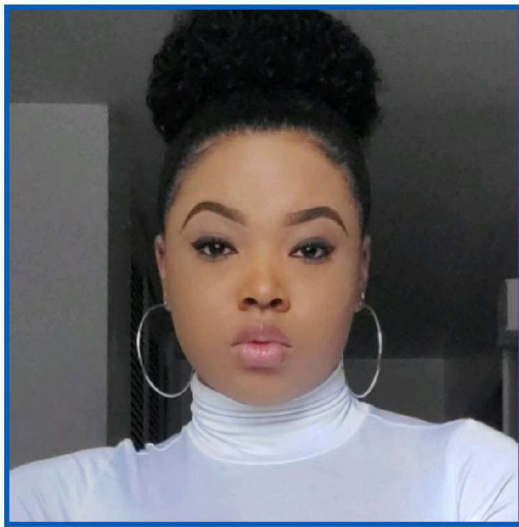
Harshita Bali

a undergraduate student with a major in Data Science
Hello everyone, my name is Harshita Bali. I am an international undergraduate student from India.

My experience of being a student here at DePaul is amazing. Initially, I was a computer science major, but with the help of Career Center Explore Career Community Advisor (Ed Childs), I recognized my skills and interest, and I decided to change my major from Computer Science to Data Science. It is also because I have a very strong hand in mathematics and statistics. The best thing that I liked about being a Data Science student at DePaul is the college's resources for its Student on-campus. From advising to tutoring, you will get every single thing.

2) Tell us about the course that you found most inspiring and impactful for your future career in data science.

The course that I found most inspiring is the Calculus course which I took with Dr. Enrico Au Yeung. He is the research mathematician here at DePaul and I can say that he is the best calculus professor that I have ever met in my life. I have taken Calculus 1, 2, 3 (MAT 150, MAT 151, AND MAT 152) with him and now I will be taking multivariable Calculus 1 (MAT 260) with him as well. He has strong and complete knowledge of the subject. He knows how to make his students comfortable by making the class interesting rather than simply teaching and writing on the blackboard. He makes sure that every student participates in the class and I believe that interaction helps you learn and grow better. The best part is that the course structure is well-organized and focuses on practical applications of calculus and its importance in science, engineering, data science and business. This helped me gain a practical understanding of how math is applied and what I can do with the power of reasoning, which I feel is very impactful for my future career in Data Science. I learned how to apply the concepts in real life through this class rather than solving questions using formulas or doing calculations.



China Agwubuo

a graduate student of the MS in Data Science –
Computational Methods concentration

My name is China Agwubuo, I am a part-time second year Data Science student, my concentration is computational methods. My First degree was in Agriculture, but I have always had a flair for numbers and problem solving.

2) Tell us about the course that you found most inspiring and impactful for your future career in data science

Every class I have taken at DePaul University has been useful to my career, although I found Data Analysis and Regression (DSC 423), Fundamentals of Data Science (DSC 441), and Database Processing for Large-Scale Analytics (DSC 450) to be the most impactful so far. I just finished a summer internship at Ulta Beauty, and every skill I learned from the classes I took was very helpful for this internship.

3) Why did you choose to pursue a degree in data science and what are your career goals?

My interest in Data Science was triggered by my need to help local farmers in my community reduce wastage of products and maximize profits. I wanted to help them understand their market better; I started researching how to achieve this, leading to Data collection and analysis. Further research led me to Data Science and all the innovative digital tools made to empower farmers. I knew then I had found my niche.

I want to build products that will excite consumers and make their lives easier, I want to use insights from data to offer solutions to business problems.

Currently, I am focused on being an exceptional data scientist to maximize any opportunity that comes my way, but I plan to get a PhD in Artificial intelligence. I want to become a subject matter expert in Artificial intelligence and make a technological and societal impact.

UPCOMING EVENTS BY DATA SCIENCE STUDENT GROUP

Career Presentation with Career Services-Lorne Brobren

Learn more about Career Services resources, the landscape of Data Science careers and speak directly with the Career Services department.

Date and Time: 4:00 – 5:00 PM CST Tuesday, September 28

No Fear Speaking – Communicating Data Science with Confidence

Develop the fundamentals of communication, learn to manage your anxiety of public speaking, and leave confident about presenting Data Science results.

Date and Time: 6:00 – 7:00 PM CST Tuesday, September 28

Career Pivoting to Data Science

Hear about the experience of peers who have recently transitioned to a Data Science career.

Date and Time: 6:00 – 7:00 PM CST Tuesday, October 5

Resume Build-up

Learn how to tailor your resume to catch the attention of Data Science recruiters and increase your prospects of getting a position in the field.

Date and time to be determined

Mock Interviews

Practice interviewing for jobs with the staffing experts at Brooksource.

Date and time to be determined

Coffee Chats

Join the community for periodic gatherings to meet new people, network and build relationships with other students.

Date and time to be determined

Movie Night

Join us for a spooky movie around Halloween and socialize with fellow students.

Date and time to be determined (~ Week 8)



Eventbrite



Facebook



LinkedIn



Slack

UPCOMING DATA SCIENCE EVENTS

ODSC APAC Virtual Conference 2021

Online Event | September 15-16

ODSC APAC Virtual Conference 2021 is one of the largest applied data science conferences. The speakers include core contributors to many open source libraries and languages. Attend ODSC APAC Virtual Conference 2021 and learn the latest AI & data science topics, tools, and languages from some of the best and brightest minds in the field.

[Read more.](#)

Deep Learning Summit (Applications)

Online Event | 22 – 23 September 2021

This Summit will explore how Deep Learning is currently being applied across business and society. Discover the latest real-world examples and case studies from, and learn key insights from applications of Deep Learning to solve challenges, improve business efficiency and have a positive impact on society.

[Read more.](#)

Microsoft Azure Data Fundamentals

Online Event | 10th October 2021

Claim your spot now for the free workshop on Microsoft Azure Data Fundamentals. In this session, your team will get a hands-on understanding of the fundamentals of database concepts in a cloud environment, basic skilling in cloud data services, and foundational knowledge of cloud data services within Microsoft Azure. This session is also designed to help you understand the core data concepts such as relational, non-relational, big data, and analytics and explore how this technology is implemented with Microsoft Azure. If you're new to Microsoft Azure Data Fundamentals, this is the best time for you to start.

[Read more.](#)

TADHack Global

Chicago | 10th October 2021

At TADHack, people (students, web developers, IT managers, coders and non-coders, graphic designers, project and product managers, interested individuals, subject-matter-experts) collaborate on a software project using new technologies. TADHack helps people learn about the latest Internet, IT, and Telecoms technologies; and use those technologies to solve problems that matter. The Real-Time Communications Lab at Illinois Tech has hosted the Chicago location of TADHack since its inception in 2014. This year, TADHack Chicago returns to the Ed Kaplan Family Institute for Innovation and Tech Entrepreneurship on IIT's Mies Campus.

[Read more.](#)



CDS UPCOMING EVENTS

Research Showcase event
Date and time to be announced



The DePaul Center for Data Science will be hosting a research showcase event where you can learn current research projects in the data science program and best practices to get involved in the projects

Advice From Data Science Alum
Thursday, September 23rd 12pm-1pm
Location: Zoom



Meet and network with Arun Gopal Govindaswamy, a recent Data Science Graduate and a Data Science at HERE Technologies.

RSVP: sprimozi@depaul.edu

Data Science Career Presentation
Tuesday, September 28th 4pm-5pm
Location: Zoom



Learn about the MS Data Science program and the variety of career possibilities in the Data Science field.

RSVP: sprimozi@depaul.edu

OUR MISSION

Our educational mission is to nurture the growth of the next generation of data scientists and computer scientists to better prepare them for data-related computing careers. Students work on state-of-the-art research and practice activities under the supervision of faculty members.

Thank you!

FOLLOW US ON



LinkedIn: DePaul Center for Data Science



Eventbrite: CDSDePaul



Twitter: depaulcds

Contact us at:
dsc@depaul.edu
243 S Wabash Ave, Chicago, IL 60604