



# Institute for Data Science



## The Institute for Data Science Hosts its Second Data Science Summit

Topics such as reimagining employee experience with AI to the role of data and innovation in global supply chains, among other topics, were discussed during a Data Science Summit hosted by the Institute for Data Science at NJIT. Faculty, Students, Industry, and the general public were all part of the audience.

If you were not able to be in attendance, the stream is live on our YouTube channel: [click here](#) to view it. To learn more about the summit [click here](#).

continued on page 3

## OVERVIEW:

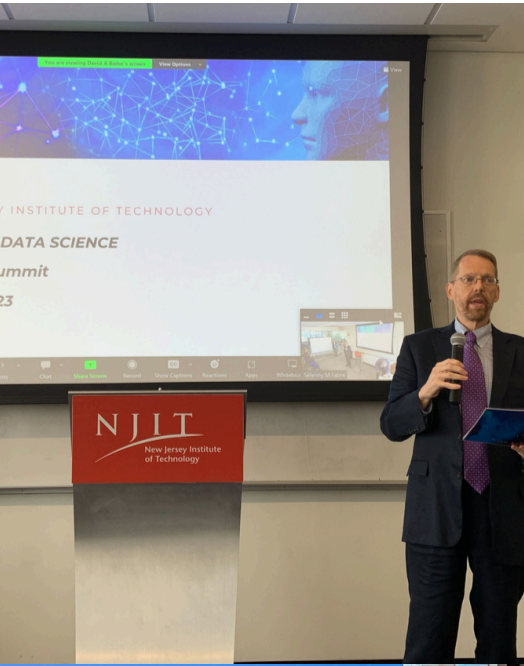
- Student & Faculty Updates
- Recent Awards
- Publications

Pages 01 - 04 - Summit Day and Advisory Board Meeting

Pages 05 - 06, & 11 - Student Spotlight

Pages 07 - 10 - Faculty Activities





# 2023 DATA SCIENCE SUMMIT



## Hot Topics of Generative AI are the Focus of 2023 Data Science Summit

written by: Evan Koblentz  
[link to Full article](#)

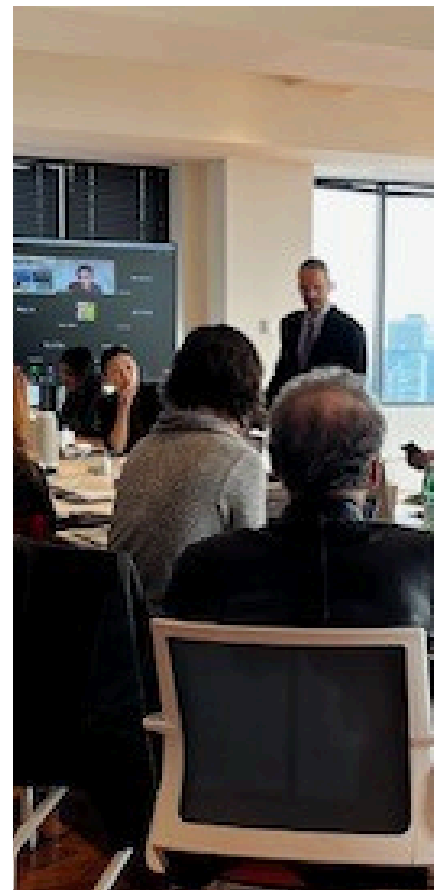
Artificial intelligence experts from ADP, Amazon, and Maersk highlighted this year's NJIT Data Science Summit, hosted by the university's Institute for Data Science. The summit began in 2022 with Google and IBM presenting about hardware and ethics. This year's focus was on generative AI, large language models, and how they're used in the business world.

"Data science is a new discipline that enables data-driven decisions across real-world problems in areas such as healthcare, security, retail and advertising, human resources, urban sustainability, and entertainment," said NJIT's David Bader, distinguished professor and director of the Institute for Data Science. "We were delighted to welcome the community to our NJIT@JerseyCity campus where we teach in our data science masters and certificate programs."

Amazon's Sherry Marcus, director of applied science for generative AI, told the in-person and online attendees that generative AI has been happening for more than a decade. What is new, she explained, is access to open-source datasets, inexpensive GPUs (a type of computer processor ideal for crunching big data), and the 2017 public description of transformer software, which is a deep-learning architecture relying on parallel processing.

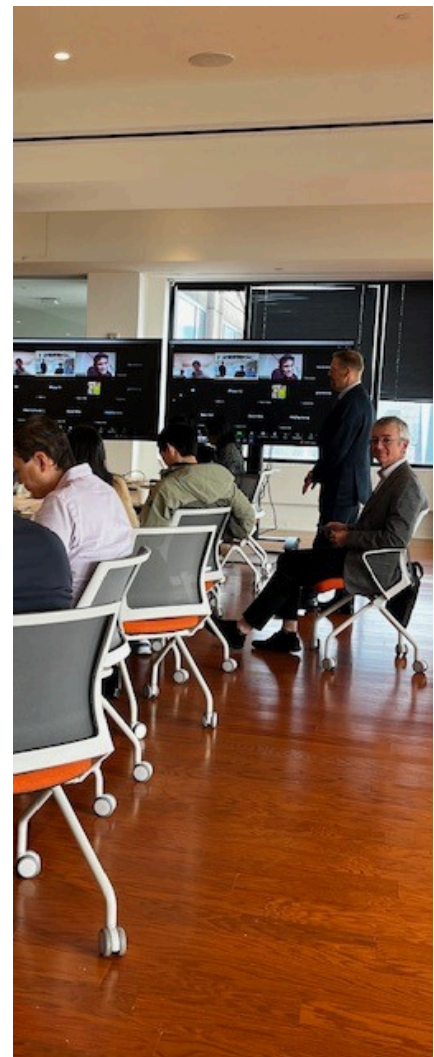
"It's the combination of these three things that as far as industry work is concerned, allow generative AI to take off," Marcus stated.

Everyone knows about large language models like ChatGPT, so Marcus instead discussed four other kinds of AI applications that she believes will change how businesses and consumers operate. They are contact centers, code generation, image generation, and machine diagnostics. Meanwhile, she said prompt engineering is the easiest way to move a business into the AI age while building one's language model would be the hardest way. [continue.](#)



The Institute for Data Science  
hosted its annual advisory  
board meeting in November.

View our Advisory Board [Here](#)



## Student Spotlight:

### PhD Dissertation Defense:

**Mojtaba Zaheri,**

"Towards Better Web Privacy: Cross-Site Leaks and Countermeasures",  
April 2023

Dissertation Advisor: Reza Curtmola

### Hear from some of our 2023 High Summer Interns!



**First Name: Sai**

**School: Edison Academy Magnet School (Middlesex, NJ)**

Working with Dr. Bader and his research group was one of the most memorable experiences of my summer. From the very beginning, I was given freedom and was able to choose any project that interested me. The idea of researching triangle counting algorithms caught my eye as I felt it was something very close to the core of data science. Working with my fellow interns and the research team gave me great connections and allowed me to engage in meaningful discussions. Overall, this opportunity taught me a lot this summer, and I hope it is the beginning of my journey in research and computer science.



**First Name: Vanessa**

**School: Montgomery High School (Montgomery, NJ)**

During my summer internship, I willingly stepped out of my comfort zone, determined to embrace new challenges. It was a valuable experience where I collaborated with peers who shared my interests and was able to learn from those who possessed more knowledge than myself. Overall, the internship enriched me with a wide range of educational and personal skills that I can now integrate into various aspects of my own life.

## Student Spotlight:

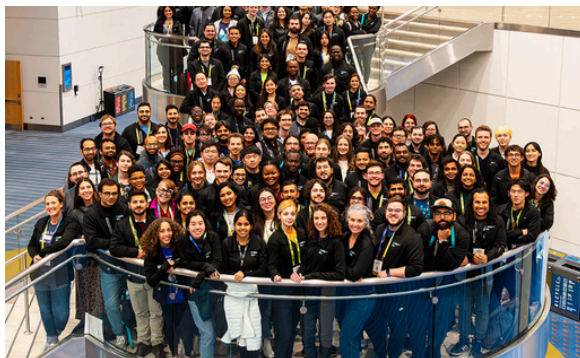


**First Name: Vinuta**

**School: North Brunswick Township High School**

Through working with Dr. David Bader and his research team, I have been able to dive into the realm of data science and learn more about various aspects that Dr. Bader works on. Research and commitment have been vital in this process, as I have learned how important it is to ensure that you have the right algorithms and equations to solve the problem that you are working on. I was also able to work alongside my brother, Anirudh Ramakrishnan, to learn more about max flow problems that can be applied in real-life

## IDS PhD Student Volunteers @ SC23



**Oliver Alvarado Rodriguez a Ph.D. candidate in Computer Science was chosen as a student volunteer for the International Conference for High Performance Computing, Networking, Storage, and Analysis (SC23) in Denver, CO November 12-17th.**

## NJIT Professor Receives NSF Grants Totaling \$2.8M to Design Solutions for Responsible AI

written by: Michael Giorgio  
[link to Full article](#)

If you're a human resource specialist grappling with AI-based decision systems for hiring good candidates, or a data scientist trying to develop transparent and accountable ranking algorithms for decision-making in critical socio-technical contexts, you will likely benefit from the research of Aritra Dasgupta, an assistant professor in the Department of Data Science at NJIT's Ying Wu College of Computing.

Dasgupta is the recipient of two collaborative National Science Foundation (NSF) grants totaling approximately \$2.8 million (out of which NJIT's funding is \$860,000) from the Information and Intelligent Systems (IIS) and the Future of Work (FW-HTF) programs.

The IIS grant (\$400K) will fund the development of human-in-the-loop techniques for data scientists to transparently design and validate algorithmic rankers in critical areas like educational program admissions, hiring and institutional rankings. The FW-HTF grant is for an interpretability-by-design framework named Trapeze (\$460K) that helps HR specialists in their meticulous balancing act of talent acquisition.

The overarching theme of both these projects will be designing and scrutinizing responsible AI solutions, according to Dasgupta, who directs NJIT's Intelligible Information Visualization Lab. He explained that "being responsible here means having agency and control over different stages of data-driven model design, allowing data scientists or decision-makers to gain confidence that the models can be safely applied in critical socio-technical contexts."

Dasgupta also emphasized the highly collaborative nature of the projects, noting that "most complex global challenges nowadays require innovations that can only stem from the amalgamation of diverse expertise. This is reflected in our project teams."

The assistant professor is collaborating with Associate Professor Julia Stoyanovich, who directs the Center for Responsible AI at New York University, on both projects. Other collaborators include experts in databases (University of Michigan Professor H.V. Jagadish, on the IIS project) and in cognitive and organizational psychology (Rice University Professor Fred Oswald and Michigan State University Professor Ann Marie Ryan, on the Trapeze project).



# CONFERENCES, PAPERS, AND ACTIVITIES

## Cybersecurity Research Center:

### USENIX Security Symposium

Hui Peng, Zhihao Yao, Ardalan Amiri Sani, Dave (Jing) Tian, Mathias Payer, "GLeeFuzz: Fuzzing WebGL Through Error Message Guided Mutation", in Proc. of the USENIX Security Symposium, August 2023. [Read Here](#).

### ACM MobiSys

Zhihao Yao, Seyed Mohammad Javad Seyed Talebi, Mingyi Chen, Ardalan Amiri Sani, Thomas Anderson, "Minimizing a Smartphone's TCB for Security-Critical Programs with Exclusively-Used, Physically-Isolated, Statically-Partitioned Hardware", in Proc. of the ACM MobiSys, June 2023, Best Artifact Award. [Read Here](#).

### VLDB Endow

Shantanu Sharma, Yin Li, Sharad Mehrotra, Nisha Panwar, Komal Kumari, Swagnik Roychoudhury, "Information-Theoretically Secure and Highly Efficient Search and Row Retrieval", in Proc. of VLDB Endow. 16(10): 2391-2403, 2023. [Read Here](#).

### ACM UbiComp

S. Rahaman, R. Samuel, I. Neamtiu, "Diagnosing Medical Score Calculator Apps", In Proc. of [ACM UbiComp](#), October 2023.

### International Conference on Computer Vision

Qiufan Ji, Wang Lin, Cong Shi, Shengshan Hu, Yingying Chen, Lichao Sun, "Benchmarking and Analyzing Robust Point Cloud Recognition: Bag of Tricks for Defending Adversarial Examples", in Proc. of the International Conference on Computer Vision (ICCV 2023), October 2023. [Read Here](#).





# CONFERENCES, PAPERS, AND ACTIVITIES

## Cybersecurity Research Center:

### ACM Conference on Computer and Communications Security

Cong Shi, Tianfang Zhang, Zhaoyi Xu, Shuping Li, Donglin Gao, Changming Li, Athina Petropulu, Chung-Tse Michael Wu, Yingying Chen, "Privacy Leakage via Speech-induced Vibrations on Room Objects through Remote Sensing based on Phased-MIMO", in Proc. of ACM CCS, November 2023. [Read Here](#).

Tianfang Zhang, Cong Shi, Payton Walker, Zhengkun Ye, Yan Wang, Nitesh Saxena, Yingying Chen, "Passive Vital Sign Monitoring via Facial Vibrations Leveraging AR/VR Headsets", in ACM CCS, November 2023. [Proceedings](#).

### IEEE International Conference on Distributed Computing Systems

Ahmed Tanvir Mahdad, Cong Shi, Zhengkun Ye, Tianming Zhao, Yan Wang, Yingying Chen, and Nitesh Saxena, "EmoLeak: Smartphone Motions Reveal Emotions", in Proc. of [IEEE ICDCS](#), July 2023.

## Structural Analysis of Biomedical Ontologies Center (SABOC):

### American Medical Informatics Association Annual Symposium

"Integrating Commercial and Social Determinants of Health: A Unified Ontology for Non-Clinical Determinants of Health" Speaker: Navya Martin Kollapally, MS, New Jersey Institute of Technology Authors: Vipina K. Keloth, Ph.D., Yale University, Julia Xu, Ph.D., JX consulting, James Geller, Ph.D., NJIT

"Usability Evaluation of Virtual Reality Ontology Object Manipulation (VROOM) System" Speaker: James Geller, PhD, NJIT Authors: Margarita Vinnikov, PhD, and Veena Chaudhari, MS, NJIT

"Usability Evaluation of Virtual Reality Ontology Object Manipulation (VROOM) System" Speaker: James Geller, Ph.D., NJIT Authors: Margarita Vinnikov, Ph.D., NJIT and Veena Chaudhari, MS, NJIT

"Virtual Reality Ontology Object Manipulation (VROOM) System" Speaker: Margarita Vinnikov, PhD, NJIT Authors: James Geller, PhD, NJIT, Danielle Grunwald, BS, NJIT



# CONFERENCES, PAPERS, AND ACTIVITIES

## Structural Analysis of Biomedical Ontologies Center (SABOC):

### International Conference on Bioinformatics and Biomedicine

Mahshad Koochi H. Dehkordi, Shuxin Zhou, Yehoshua Perl, James Geller, Andrew J. Einstein, Gai Elhanan, Vipina K. Keloth, Hao Liu, "Using annotation for computerized support for fast skimming of cardiology electronic health record notes," [BIBM 2023](#)

## Intelligible Information Visualization (NiiV) Lab :

### NSF IIS Award in Collaboration with NYU

Aritra Dasgupta was awarded an NSF IIS grant as a PI ( in collaboration with NYU, Umich) for developing visual analytic frameworks to study algorithmic rankers.

### DOE CESER Award

U.S. Department of Energy's (DOE) Office of Cybersecurity, Energy Security, and Emergency Response (CESER) announced \$39 million of funding for nine new National Laboratory projects to advance the cybersecurity of distributed energy resources (DER). Aritra Dasgupta receives a DOE CESER grant as a PI via PNNL

### IEEE PacificVis

Aritra Dasgupta is invited to serve on the International Program Committees of IEEE [PacificVis](#) (Journal track) 2024 and [WWW 2024](#).

## Intelligible Information Visualization (NiiV) Lab

### SIGMOD 2023

Jun Yuan and Kaustav Bhattacharjee (DS Ph.D. candidates) presented 2 papers at the HILDA workshop SIGMOD '23: Companion of the 2023 International Conference on Management of Data.

### Visual Computer Journal

Jun Yuan and Kaustav Bhattacharjee's paper on the visual explanation of algorithmic rankers was accepted by Springer-Nature's Visual Computer journal.

### NJIT Research Day

Jun Yuan, Ph.D. candidate in Data Science, received the third place prize in the university-wide research day competition presenting his research on explaining algorithmic rankers.

### Conference on Neural Information Processing Systems

Jun Yuan presented his paper on adversarial attacks against XAI methods at the NeurIPS 2023 workshop on XAI in Action.





# CONFERENCES, PAPERS, AND ACTIVITIES

## Institute for Data Science

### International Conference on Computational Science

“Parallel Triangles and Squares Count for Multigraphs using Vertex Covers,” Luca Cappelletti, Tommaso Fontana, Oded Green and David Bader, International Conference on Computational Science (ICCS), Prague, Czech Republic, July 3-5, 2023. [Abstract](#).

### The 10th Annual Chapel Implementers and Users Workshop

“[Minimum-Mapping based Connected Components Algorithm](#),” Zhihui Du, Oliver Alvarado Rodriguez, Fuhuan Li, Mohammad Dindoost, and David A. Bader, The 10th Annual Chapel Implementers and Users Workshop (CHI UW), Virtual, June 1-2, 2023.

### The 27th Annual IEEE High-Performance Extreme Computing Conference

“Parallel Longest Common SubSequence Analysis In Chapel,” Soroush Vahidi, Baruch Schieber, Zhihui Du, and David A. Bader, The 27th Annual IEEE High-Performance Extreme Computing Conference (HPEC), Virtual, September 25-29, 2023. [Read Here](#).

“Property Graphs in Arachne,” Oliver Alvarado Rodriguez, Fernando Vera Buschmann, Zhihui Du, and David A. Bader, The 27th Annual IEEE High-Performance Extreme Computing Conference (HPEC), Virtual, September 25-29, 2023. [Read Here](#).

“Triangle Counting Through Cover-Edges,” David A. Bader, Fuhuan Li, Anya Ganeshan, Ahmet Gundogdu, Jason Lew, Oliver Alvarado Rodriguez, and Zhihui Du, The 27th Annual IEEE High-Performance Extreme Computing Conference (HPEC), Virtual, September 25-29, 2023. Graph Challenge Student Innovation Award. [Read](#).

“Fast Triangle Counting,” David A. Bader, The 27th Annual IEEE High-Performance Extreme Computing Conference (HPEC), Virtual, September 25-29, 2023. Graph Challenge Innovation Award. [Read Here](#).



# CONFERENCES, PAPERS, AND ACTIVITIES

## Intistute for Data Science

### The 30th IEEE International Conference on High-Performance Computing, Data, and Analytics

“Contour Algorithm for Connectivity,” Zhihui Du, Oliver Alvarado Rodriguez, Fuhuan Li, Mohammad Dindoost, and David A. Bader, The 30th IEEE International Conference on High-Performance Computing, Data, and Analytics (HiPC), Goa, India, December 18-21, 2023. [Read Here.](#)

### Journal paper publication in Quantum

“End-to-end resource analysis for quantum interior point methods and portfolio optimization,” AlexanderM. Dalzell, B. David Clader, Grant Salton, Mario Berta, Cedric Yen-Yu Lin, David A. Bader, Nikitas Stamatopoulos, Martin J. A. Schuetz, Fernando G. S. L. Brandão, Helmut G. Katzgraber, and William J. Zeng, PRX Quantum, 2023. [Read Here.](#)



# INSTITUTE FOR DATA SCIENCE DIRECTOR'S OFFICE

---



## **DAVID BADER**

Institute Director  
david.bader@njit.edu



## **SELENNY FABRE**

Business Manager  
selenny.m.fabre@njit.edu



---

[About Us](#) | [Contact Us](#) | [Subscribe](#)

---