

Data Science Research Service



**Gies College
of Business**

From the Director



“Over the past year, we’ve supported a wide range of transformative research projects, from complex machine learning applications to critical advancements in data acquisition and processing.”

With the importance of data-driven research today and the major advancement of artificial intelligence technology and data collection, we know it’s vital to seamlessly connect researchers with the tools and resources they need.

Our primary objective has always been to foster innovation and advance data-driven research within Gies College of Business and beyond by providing world-class data science solutions. Today, I can confidently affirm that we are achieving this goal. Over the past year, we’ve supported a wide range of transformative research projects, from complex machine learning applications to critical advancements in data acquisition and processing. We’ve seen our work make a tangible impact across various disciplines, and to me, this is the true measure of our success.

At the heart of the Data Science Research Service is our incredible group of students. Their curiosity, ingenuity, and determination drive our work forward. Through our internship program, we’ve provided them with invaluable experiences and skill sets while also benefiting from their fresh perspectives and ideas. Watching them grow and develop as data scientists has been one of the most rewarding aspects of my role.

Along with fostering collaboration and helping students get the research experience they need, we broaden our researchers’ access to key resources and cloud computing, help guide research in emerging areas like chatbots, and assess researchers’ needs. Thanks to a collaboration with NCSA, we have our own computer cluster managed with state-of-the-art open-source technologies which we use to run models, provide data access and resources to our researchers. When external tools are necessary, we make recommendations and can help facilitate those connections.

We ultimately save time for researchers by figuring out what they need and delivering the proper tools and resources to help them achieve it. Thank you to those who have collaborated with us and used our services this year. We hope to continue to grow and help facilitate research for years to come.

Matias Carrasco Kind

Director, Data Science Research Services
dsrs.illinois.edu

About the Data Science Research Service

The Data Science Research Service (DSRS) is a subunit of Disruption and Innovation at Gies College of Business that assists faculty members with data science to support their data-oriented research. Our mission is to make the College the most technologically forward and data-capable business college in the world.

Azure Award

We're delighted to have been awarded over \$160,000 to date in cloud computing credits from Microsoft Azure. This contribution has been instrumental in fostering innovation, facilitating the exploration of new technologies, and enabling large-scale computations for advanced research. In addition, these credits have broadened our ability to provide faculty members with access to high-powered cloud computing services, enhancing their capacity for in-depth exploration and discovery. Access to these credits is also included as part of our services.

DSRS Support Services and Costs

We assist faculty research through the execution of statistical and data science methods, providing consultative services and analyses within our computational infrastructure. The College has invested in DSRS and funds the unit's consultations and other work.

We assess research studies to determine what type of data and computational infrastructure are needed, and we can provide recommendations if the needs are external. There may be a cost for contracting external services. In some cases, researchers can provide their own funding to have their projects going for longer periods of time or to have exclusive help from our interns.

Consultation: We can assess and make recommendations for what statistical and data science methods are appropriate for specific research questions. Here are some methods we've used:

- Inferential statistics
- Social media analytics
- Text mining
- Natural language processing
- Machine learning
- Deep learning
- Data visualization
- Dashboard design
- Web scraping
- Data storage
- Database usage
- Data cleaning
- Data gathering
- Scalable computing
- Image processing
- App development

Analyses: Our director manages student interns who conduct analyses – the core service of DSRS. We use fundamental principles of data science to help deliver research solutions. Our team features students from business, computer science, statistics, and social science backgrounds who bring in data and methods from each of these disciplines.

Data: We maintain an in-house library of datasets that are available for Gies Business faculty to use. If specific data are needed to answer research questions, we can assist in locating and acquiring the right data. There may be an additional cost to acquire or license the dataset.

Computational capability: We use on-campus computation; cloud computing; and custom-built, on-premises computation to assist faculty research. We can help subcontract the analyses to other units on campus, such as NCSA, if certain computational infrastructure is already built or is more cost-effective to use on their side rather than within our office. We help to connect researchers to these resources.

We have our own cloud computing cluster with over 600 cores and over 30TB of storage in addition to access to Microsoft Azure credits.

Researcher Experience and Impact

2022-23 Completed Research Studies

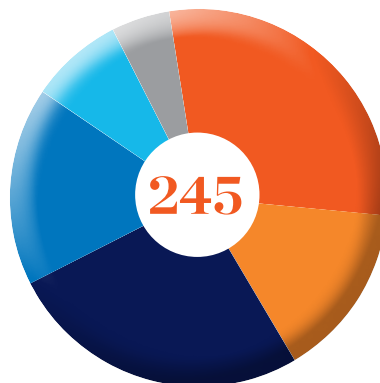
DSRS works with all Gies Business departments as well as other units on campus.

60+

Faculty members assisted

245

Tasks completed



Total Tasks by Department

- ACCY
- BA
- DSRS
- External
- FIN
- Miscellaneous

Faculty Testimonials

“The Data Science Research Service has been instrumental in assisting my coauthors and I to apply cutting-edge machine learning and data analytics techniques to our research. Through DSRS, we’ve also been able to gain access to high-performance cloud computing services, which allowed us to conduct our analyses quickly and meet key research deadlines. I commend the DSRS team for being extremely responsive and leveraging talent both within and outside Gies. We typically receive a response to an initial request in 48 hours.”

Nerissa Brown

Associate Dean of Graduate Programs

Professor of Accountancy

Josef and Margot Lakonishok Faculty Fellow

PwC Faculty Fellow

“DSRS is a very unique initiative spanning departments. Given the importance of data analytics and large-scale quantitative modeling in many disciplines these days, DSRS can help faculty and PhD students borrow approaches from other fields and develop interesting interdisciplinary research. It also saves the faculty’s time by matching them to the right graduate students while providing students the necessary research experience they’re seeking.”

Unnati Narang

Assistant Professor of Business Administration

John M. Jones Fellow of Marketing

RC Evans Data Analytics Scholar

“DSRS gives Gies researchers a competitive advantage that can make a difference in many research areas. Extracting and creating novel datasets can be challenging for researchers because they are often unstructured, or the combination is computing intensive. DSRS provides know-how and hands-on assistance to create these novel datasets and hence gives a cutting edge by enabling researchers to create these datasets in an efficient way.”

Benjamin Osswald

Assistant Professor of Accountancy

“My coauthor and I need to process a large amount of text in different formats and perform textual analysis on them. Matias Carrasco Kind and his students have been helping us with this project since last year. Given the complexity of the source data and our lack of expertise in textual analysis, it would have been challenging for us to take on these tasks ourselves. Their expertise and flexibility in working with us have been tremendously helpful.”

Anh Persson

Assistant Professor of Accountancy

“I have been working with Matias Carrasco Kind to develop a Data Deep Diving initiative that will be useful for students to understand, replicate, and comment on classic papers. It is a pleasure to work on new data science ideas with support from the Data Science Research Service at Gies. The unique service provides everything I need to host data, code, analysis, and training material.”

Sridhar Seshadri

Alan J. and Joyce D. Baltz Endowed Professor

Area Chair, Information Systems/Operations Management/Supply Chain/Analytics

Student Intern Experience and Impact

DSRS provides experiential learning opportunities for students across campus, giving them hands-on experience applying data analytics skills to business education.

Our student interns come from Gies Business and other areas of campus, like the Computer Science Department, Statistics Department, and School of Information Sciences. They're open-minded, collaborative, and creative – and they're also motivated to take on responsibility and solve problems.

Ready to learn more about the field, take on challenges, work as a team, and build their networks, our student interns bring these skills and more:

- Data science
- Programming
- Data analysis
- Web development
- Data visualization
- Web scraping
- Text mining

Student Testimonials

"I work with many talents, where I learn a lot, such as the mindset to solve problems, communication skills, coding skills, and so on. I meet with a lot of people through DSRS who help and guide me along the way. I practice what I learned from classes and also keep learning new knowledge and skills in DSRS, which has become a great treasure for me."

Cinny Fang (ACCY '22, MAS student)

"As someone with a background in analytics, I feel incredibly fortunate to work at DSRS to refine my skills. Working at DSRS has provided me with the opportunity to engage in the complete analytics lifecycle, taking ownership of impactful projects. From requirements gathering to data ingestion, exploration, analysis, and modeling, I have gained invaluable experience. The supportive environment created by Director Kind encourages learning, experimentation, and innovative solutions to the analytics challenges faced by DSRS's diverse clientele across the university. DSRS has proven to be an exceptional platform for honing my skills and contributing to real-world projects that make a difference. I am immensely grateful for the experience and professional growth I have gained through my association with DSRS."

Abhi Yarlagadda (MSBA '22)



Data Library

The College licenses several datasets that can be used by all members of the Gies Business community for research, classroom teaching, building case studies, crafting student projects for action learning, and more.

Some of the datasets we license include the following:

- **Wharton Research Data Services (WRDS)** is a leading business intelligence, data analytics, and research platform that enables web-based query to extract and output data from multiple data sources. WRDS hosts 350+TB of data – across all disciplines, including accounting, banking, economics, ESG, finance, healthcare, insurance, marketing, and statistics – giving users the power to analyze complex information and create a wide range of reliable data models.
- **Bureau van Dijk** is a major publisher of business information that specializes in private company data combined with software for searching and analyzing companies. It features information on more than 400 million companies in all countries worldwide.
- **Gies Consumer and Small Business Credit Panel** includes consumer data on over 245 million consumers, over 32 million businesses, and other proprietary data assets.
- **Capital IQ** helps researchers find financial data for public and private companies, people, and transactions. It can also help create customized reports and features a modeling tool that can be used with Excel.
- **Bloomberg Terminal** provides coverage of markets, industries, companies, and securities across all asset classes.
- **Nielsen** datasets include the following: Consumer Panel Data, Retail Scanner Data, Ad Intel Data, Nielsen's Ad Intel Data, PromoData.

Frequently Asked Questions

What services does the Data Science Research Service (DSRS) offer?

DSRS offers a broad spectrum of data science services, including statistical and data analysis, machine learning model development, data visualization, data acquisition and cleaning, AI application development, and cloud computing services.

Who can use DSRS services?

Our services are available to faculty, researchers, PhD students, and other members of the university community who require data science expertise for their projects or research.

How can I request DSRS services?

You can submit a service request by sending an email to dsrs@business.illinois.edu with a general explanation of your research study and question. DSRS will then schedule a meeting to discuss the project and provide recommendations.

What are the costs for DSRS services?

Gies College of Business supports initial consultations and tasks. However, any costs related to specific project needs, such as subcontracted analyses, computational infrastructure, or data acquisition, will need to be covered by your or your department's research funding sources.

I'm a student. Can I get involved with DSRS?

Absolutely. DSRS values the contribution of students and offers several internship opportunities. For more information, please visit dsrs.illinois.edu/about/students

Where can I find more information about becoming an intern?

More details about student intern roles and responsibilities are available on the DSRS Gies Groups site. We regularly recruit data science student interns for various projects.

What's the mission of DSRS?

Our mission is to drive research within Gies College of Business by providing comprehensive support in data science, machine learning, computational infrastructure, and data acquisition as part of the Gies Disruption initiative.

What types of projects has DSRS worked on in the past?

DSRS has worked on a wide range of projects, including textual analysis, NLP data mining, image analysis, algorithmic development, deep learning models, and blockchain research.

What computational resources does DSRS provide?

DSRS provides access to high-performance computing resources, cloud computing, and Jupyter Notebooks, among others. We also offer assistance in database deployment, including SQL and non-SQL databases.

How does DSRS handle data privacy and security?

DSRS is committed to maintaining the highest standards of data privacy and security. We adhere strictly to university guidelines and regulations concerning data privacy, confidentiality, and security.



**Gies College
of Business**

