FROM THE ASSOCIATE PROVOST

WE ARE READY

Colleagues + Friends,

A long time in the making from its conception as a taskforce recommendation in 2018 to its launch in late 2019, it is hard to believe that the Faculty of Computing & Data Sciences (CDS) is well into its second year as the newest academic unit at BU. In an inaugural year dominated by COVID-19 and all its associated challenges, we have been laying the necessary foundations – literally of the new building, and figuratively of our academic programs – to enable CDS to emerge from the pandemic ready to welcome students, faculty, and staff into a fully-fledged community that is poised to further innovation and propel data sciences into the future.

Indeed, we did not miss a beat!

In the span of just over a year, we launched two signature degree programs: our Bachelor of Science in Data Science undergraduate program and our PhD in Computing & Data Sciences graduate program; we completed national searches for tenure-stream, clinical, and of-the-practice professorships, bringing the size of our core faculty to five strong; we fused BU Spark! into CDS by seamlessly integrating its innovation and experiential learning programs into our curricular and co-curricular offerings, and more importantly by assimilating its culture of inclusion and empowerment of students; we grew our management, communications, and advising capacities by recruiting administrative team members who are passionate about our vision and culture; and we established partnerships with internal and external organizations that believe in the transformative power of computational and data-driven inquiry.

But the story we really want to tell is one of institutional and societal impact that goes well beyond mere numbers.

As a key initiative of the university’s BU2030 strategic plan, CDS is set up to ignite and nurture BU students’ passion for vibrant academic experiences, to distinguish BU faculty through diverse hiring and reimagined connections that lead to research that matters, and to be the connective tissue for a BU computing community big, yet small, by bringing together diverse and inclusive teams of faculty, students, and practitioners into impact-driven collaborations.

Our aim is nothing less than to democratize access to computing and data sciences, by preparing a more diverse cohort of learners for rewarding computing careers, by catalyzing faculty research in disciplines that are yet to be transformed by the advent of machine learning and AI, and by bringing data-driven innovation to bear on societally relevant challenges that are often overlooked by industry.

I invite you to read this opening chapter of the unfolding CDS story – a story that I hope you will co-author with us in the months and years to come.

Azer Bestavros
Associate Provost
In December 2019, BU President Robert A. Brown and Boston Mayor Marty Walsh held a ceremonial groundbreaking to kick off one of the most ambitious building projects in BU’s history: the Center for Computing & Data Sciences.

Simultaneously, BU leaders were breaking down walls—the traditional departmental barriers between disciplines—to create an equally ambitious academic unit: the Faculty of Computing & Data Sciences. This community of researchers believes in the potential of new ways of thinking to advance—and even transform—their disciplines.

CDS is a truly cross-cutting academic unit free from the limits of traditional academic structure. It has no schools, no departments, no divisions. It includes and transcends all disciplines, just as its emerging field now does.

The unifying force is the innovative application of data science, machine learning, and AI technologies and methodologies. As such, CDS is a catalyst for synergy and integration across the landscape of academic disciplines at Boston University.
IMPACTFUL MILESTONES
created + achieved

2
Signature academic programs launched (undergraduate and PhD)

5
Core faculty hired

20
Courses added to BU’s curricular offerings

2
co-Labs launched

2
Fellowship programs launched (Impact and Civic Tech)

550+
Students served through Spark!

$3M+
Foundational, alumni, & corporate funding secured

The cross-cutting nature of CDS is evident in the way its faculty, students, and programs are clustered around thematic areas of impact as opposed to computing and data science subspecialties. CDS Impact Hubs are the thematic areas of “research that matters” and “vibrant academic experiences” that connect computing and data science with societally-relevant areas of impact that BU is strategically positioned to lead, most notably around Equity, Sustainability, Health, and Civic Tech.

Within the context of these Hubs, much of CDS’ work is done through co-Lab partnerships. These engage stakeholders from across BU and beyond and feature a shared infrastructure for specific partner-directed deliverables by focusing on integrating basic and applied research with curricular and co-curricular activities.

In 2020, we launched the Justice Media Computational Journalism co-Lab, in partnership with COM, and laid the foundation for the Racial Data Lab, a collaboration between CDS and the Center for Antiracist Research. Stories about these and projects carried out through the Hub for Equity Impact are featured throughout this report.
CDS embarked on multiple faculty searches. In 2020, we garnered significant initial interest with roughly 150 applicants. Forty-two made our short list, six were interviewed, two offers made, and one hired.

In early 2021, we carried out one senior search at the “CDS + Law, Ethics, and Society” nexus of areas, and two junior searches focusing on "CDS In The Field" with particular focus on connecting data science, machine learning, and AI with economics and social sciences. We started with 500+ applicants and ended up with 21 candidates on our short list; we interviewed 12; made offers to eight; and landed two faculty. See below for more review statistics.

Following the success of this search, we turned to the recruitment of teaching, clinical and of the practice faculty, landing two such faculty who joined us in the Fall of 2021.

BY THE NUMBERS

26% female-identifying applicants

50% female-identifying applicants given offers

43% of dossiers received for "DS/Al In The Field" search area

715 dossier reviews

152 interview reports

43 distinct faculty interviewers

26 interview days

“I love the main idea behind CDS, which is to create a space in which scientists from different disciplines can meet and collaborate on both research and teaching.”

- Krzysztof Onak
CDS Assistant Professor
Transitioning from a career in industry to academia, Krzysztof Onak became CDS’ first assistant professor, beginning in January 2021. Shortly thereafter he was named the inaugural Shibulal Family Career Development Professor. Named for the family of S.D. Shibulal, a 1988 graduate of Boston University and member of the Boston University Board of Trustees, this three-year award is a recognition of talented junior faculty who show exceptional promise.

Krzysztof is a prolific researcher whose work transcends multiple disciplines and application domains. He has a robust track record of impactful fundamental and applied research and embodies the qualities we look for in future leaders in data science. A hands-on builder and “tinkerer” at heart, he still loves programming and enjoys implementing his ideas himself – a rarity for researchers focused on the foundations of data science.

A passionate thought-leader, Kira Goldner is a stellar researcher whose interests are primarily in algorithmic mechanism design and approximation algorithms, including multi-parameter revenue maximization and analysis and models with varying behavioral or informational assumptions.

Kira joining the Faculty in July 2021 bolsters CDS’ prioritization of Data Science for Good (DS4G). Recently, she was named a Shibulal Family Career Development Professor. Currently, she is organizing the inaugural ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization (EAAMO ’21). Part of the Mechanism Design For Social Good (MD4SG) initiative, EAAMO aims to highlight work where techniques from algorithms, optimization, and mechanism design, along with insights from the social sciences and humanistic studies, can help improve equity and access to opportunity for historically disadvantaged and underserved communities.

"I am so excited to join CDS’s truly interdisciplinary and impact-focused environment, and to be somewhere that recognizes the necessity in social-impact research of breaking down barriers and bridging disciplines."

- Kira Goldner
Langdon White is a 20+ year veteran of the software industry and has seen the evolution of software into all aspects of our lives. Langdon joins CDS from Red Hat Inc. where he held several senior engineering roles over the course of nearly a decade.

Through the Red Hat Collaboratory at BU, Langdon became affiliated with the Hariri Institute, Spark!, and eventually CDS. As Engineer in Residence, Langdon taught a "working in industry fundamentals" class to prepare students for their first internship or job in the software development industry. He served as a software development mentor to multiple student teams working on startups and projects for external partners. He also taught several practicum courses supporting students delivering successful projects in partnership with industry partners.

Langdon is a dynamic individual with a wealth of professional experience and a passion for teaching and mentoring students. He teaches our seminal undergraduate introductory course, CDS DS 100 Data Speaks Louder Than Words. He serves as Spark! Technical Director and will also lead the practicum track of courses for CDS, which offer students opportunities to engage in experiential learning with real-world partners through Spark!.

"I hope we can create a community where we really get to know our students and each other, and create a wonderful learning environment as a result."

- Kevin Gold

Kevin Gold earned his PhD in Computer Science in 2008 from Yale University and his BA from Harvard in 2001, and most recently held the position of Teaching Associate Professor in Northeastern University's Khoury College of Computing and Information Sciences. Before coming back to academia, Kevin had a number of stints in industry, including working at Lincoln Laboratory for cybersecurity groups, at Google as part of the YouTube quality of experience team tracking playback errors, and as part of the Google Search Division team looking at question-answering systems. In all of these industry experiences, Kevin was focused on using diverse AI techniques in different contexts and also convincing others that those techniques were the right approach.

Kevin's teaching experience at Wellesley College, RIT, and at Northeastern covered Artificial Intelligence, Algorithms and Data, and several courses on math for computer scientists. Of particular note is his experience developing and redesigning courses, particularly one called Mathematics of Data Models, in which he presented linear algebra and statistics through their applications in data science topics such as machine learning. Kevin is passionate about equity and inclusion in the classroom, developing lessons designed to address and fight prejudice and bias while simultaneously covering topics such as statistics and machine learning.
Mayank Varia joined CDS by way of the Department of Computer Science at BU.

Mayank works in the fields of theoretical and applied cryptography. Specifically, he explores the computational and social aspects of cryptography, and his work has been featured in media outlets like CNET, The Hill, and ZDNet. His designs for accessible, equitable, and socially-responsible data analysis have been used to determine the gender wage gap, subcontracting to minority-owned businesses, and repeat offenders of sexual assault inspired by the #MeToo movement.

Mayank is co-director of BU's Center for Reliable Information Systems & Cyber Security, a founding member of the Cyber Security, Law, and Society Alliance, a fellow in the Institute for Health System Innovation & Policy, and a faculty mentor of the Codebreakers outreach program. He also is a member of the U.S. Advisory Committee on Data for Evidence Building and chairs the legal subtask group of the United Nations Privacy Preserving Techniques Task Team.

In addition, Mayank serves as Director of the Hub for Civic Tech Impact and the Director of Undergraduate Studies.
Just as technology opens doors and shows us innovative opportunities, it can also deepen inequities. We believe deeply that technology should serve the public interest and to that end provide opportunities for students, faculty, and partners to learn skills and put them to use to improve the world.

Civic Tech is about development of tools, software, and algorithms that are increasingly critical for a healthy civic society as well as the evaluation of these technologies in legal, regulatory, and social contexts.

There’s a lively community of scholars who engage with Civic Tech, from the Cyber Alliance to undergraduate students producing software for a better democracy.

The Hub for Civic Tech Impact incorporates all programs and offerings at the intersection of technology and public interest. Our pillars are:

- OPEN: Enabling transparency, accountability, and countering corruption
- EFFECTIVE: Supporting efficient and equitable delivery of government services
- TRUSTED: Countering polarization and mis/disinformation
- ENGAGED: Promoting citizen participation, action, and protection

Our newly launched Civic Tech Fellowship program recognizes faculty members who are committed to making society better through the development and evaluation of technologies. Our Expert in Residence program brings in community members with on-the-ground experience to lend their knowledge and mentorship to student projects and faculty collaborators.
TOWARDS A MORE INCLUSIVE CULTURE

We are made stronger by the diversity of backgrounds and perspectives within our community of faculty, students, and staff. We have an incredible opportunity to intentionally build our culture around the principles of diversity, equity, and inclusion. We will achieve these aims by co-creating a set of core values with our community, incorporating DEI goals into our strategic plan, implementing strategic initiatives that support these goals, and evaluating our progress annually.

Building on Spark!’s work, CDS will foster an inclusive and vibrant community among students while widening the pipeline and closing the skills gap. Not only geared towards students enrolled in CDS degrees, our work aims to attract students who are not currently engaged so they gain exposure and consider learning opportunities in these fields.

We believe that experiential learning programs effectively engage students from underrepresented backgrounds in the issues that affect them. To this end, we are building a robust portfolio of 28 partners who provide real-world projects focused on issues of equity. Through the Spark! Equity Innovation Fellowship, students are able to receive funding and mentoring support to develop unique computing and data science solutions that address issues of equity.

We also plan to leverage our unique geography. We are developing a series of colloquia to bring together rising stars from diverse backgrounds around topics of interest with a focus on our four impact areas with the goal of building a community of diverse scholars at BU and beyond who share our commitment to impact.

how we do it

Faculty search - Faculty are at the core of any academic unit and our diverse faculty search areas reflect our aim to diversify our view of scholarship to emphasize societally-relevant impact.

Prehacks - BU partners with TechTogether (the country’s largest initiative to address gender inequity in the hackathons community) to increase participation among high school students. TechTogether has supported over 4700 gender-marginalized individuals.

Spark! Ambassadors - Spark! Ambassadors are focused on developing and implementing student-led activities that build connection and belonging among students in the computing and data sciences at BU while also expanding opportunities for tech-curious students to engage.

The Ignite Student Council - The Ignite Student Council is a group of leaders representing 27 computing focused clubs. Spark! supports the Council by providing training opportunities, funding, and mentoring toward the shared goal of creating an inclusive and vibrant community for students interested in computing.

PhD Engagement - Through a robust seminar series and a rotation program, CDS provides PhD students with opportunities to gain wide exposure across diverse domains.

“I am forever grateful that through your efforts to make the space more inclusive and use tech as a force for social good, I was able to work with such an amazing group.”

- Gonzalo Rosales
BU ’21
CDS is focused on educational pathways that enable students to engage with CDS programs and initiatives throughout their BU experience.

CDS courses offer students a variety of opportunities to put their emerging computing and data science skills to use by working on projects with real-world partners. Many of these courses are powered by Spark! who recruits compelling projects through external partnerships. These courses also allow students to build supplementary skills in project management, communication, and teamwork.

A few of these practicum courses are:
- DS 473 Justice Media co-Lab
- DS 519 Software Engineering X-Lab
- DS 539 Data Science X-Lab
- DS 549 Machine Learning X-Lab

Throughout 2020 and 2021, we developed foundational courses that act as ramps for students to try out the Data Science major, courses that nurture a shared cohort experience among students (at the graduate and undergraduate levels), and a robust seminar series for PhD students. As part of the integrative pedagogy of CDS, these courses - motivated by real-world applications and with no presumed exposure to math or computer science - are designed to widen the data science pipeline. In collaboration with units across the University, we quickly expanded our inventory of courses, adding 20 new courses to BU’s curricular offerings. Below is a snippet of that extensive list.

- DS 100 Data Speaks Louder than Words
- DS 110 Intro to Data Science with Python
- DS 120-122 Foundations of Data Science
- DS 199 Surveillance & Data
- DS 210 Programming for Data Science
- DS 457 Law for Algorithms
- DS 482 AI, Law, Ethics & Society
- DS 537 DS for Conservation Decisions
- DS 543 Taming Big Data
- DS 571 Tools for Data Science
- DS 644 ML for Business Analytics
Can COVID Be Tracked Privately?: Computer Science professor, founding CDS member, and RISCS director Ran Canetti worked with BU cybersecurity experts and MIT researchers to develop a new Bluetooth-enabled contact tracing app that warns of COVID-10 exposure while protecting privacy. Their paper attracted much attention (https://arxiv.org/pdf/2003.13670.pdf) and the BU team joined the Private Automated Contact Tracing Team.

"Research that matters": CDS brings teams of faculty, students, and practitioners into impact-driven collaborations to address societally relevant challenges.

Talks, programs, publications: BU researchers have been busy using data to further the fields of health, cryptography, economics and more.
AI used to determine extent of damage in kidney disease: Founding members Vijaya Kolachalama, Margrit Betke, and BUSM researchers developed a novel AI tool to predict grade of interstitial fibrosis and tubular atrophy, a structural correlate of progressive and chronic kidney disease.

When is Memorization of Irrelevant Training Data Necessary for High-Accuracy Learning?: Computer Science professor and founding CDS member Adam Smith and collaborators, including BU researchers Gavin Brown and Mark Bun, published a paper on how machine learning models need to memorize sensitive data.

Abuse-Resistant Government Backdoors: Gabe Kaptchuk, CDS Civic Tech Fellow, and Johns Hopkins researchers proposed minimum criteria that strong encryption systems should meet.

High-resolution land value maps reveal underestimation of conservation costs in the United States: Earth & Environment professor and CDS faculty Christoph Nolte published a paper in PNAS (https://www.pnas.org/content/117/47/29577) which posited that conservation policy analysis should use high-resolution cost estimates in their justification and spatial prioritization of interventions.

Learn to Earn: Enabling Coordinating Within a Ride-Hailing Fleet: Founding members John Byers and Evimaria Terzi, with colleagues in Computer Science, proposed a solution for optimizing social welfare objectives on platforms such as Uber and Lyft.
technology incubator and experiential learning lab

128 semester-long projects

552 students served through Spark! experiential learning engagements

5 Spark! supported practicum courses

AY '20-'21
Now housed at CDS, Spark! helps students find their passion, discover community, and prepare for life by providing computing and data science centric experiential learning engagements with real-world impact.

At its core, Spark! is a community. A community of students who have taken a step to explore different topics and connect with others. Spark! offers a diverse range of curricular and co-curricular semester-long programs focused on student led innovation and project-based experiential learning with external partners. Increasingly, our projects with external partners are focusing on the CDS Impact Hubs. We have a strong emphasis on equity and civic tech and will be expanding to sustainability and health in the year to come. Spark! also supports the Justice Media co-Lab, with its affiliated practicum course and summer internship program.

Alongside these programs, Spark! offers a diverse array of community building activities, all of which went remote in March 2020.

From weekly Zoom hangouts to wildly successful virtual Demo Days and a screening of Coded Bias, the Spark! team got creative and stayed committed to providing a welcoming community for all students.

As the weather warmed, the Spark! community was even able to hold a few socially distant socials and facilitate swag pickups for nearby students.
Journalism is stronger when informed by and infused with data, which provide information about patterns that may be otherwise invisible. But many journalists lack access to “big data”—and the skillsets to interpret and use it compellingly.

A collaboration between the College of Communication and CDS (powered by Spark!), the Justice Media co-Lab trains a new generation of computational investigative journalists equipped to leverage the power of computing and data sciences to advance justice and transparency. Since its inception, students have made a tangible impact working with news organizations such as the Associated Press, CBS, The Boston Globe and more.

During Spring 2021, we launched the inaugural Cross-College Challenge co-led by faculty from the Department of Journalism and in partnership with the BU Hub. Through this course, students partnered with NBC10 Boston to analyze state arrest data. Students merged data spanning 10 years and discovered that, in some communities, Black people were arrested at a disproportionately higher rate relative to their share of the population.

The students and NBC10 interviewed police departments found to have disparities in arrest data, along with policing experts to get their thoughts on the findings.

In year one of the co-Lab, 37 students and 5 media partners were engaged through the affiliated summer internship program.

Camila Beiner, a journalism project manager, said the experience was one-of-a-kind. “I worked alongside news organizations I have always looked up to and wanted to strive for. It was an experience very few young journalists get during their college years.”

“I was able to work with students and reporters to investigate criminal justice issues that I am very passionate about. I was also able to learn from my colleagues more about data and computer science that will help me with reporting in the future.”

- Camila Beiner
"Spark! provided us with the opportunity to turn our idea into a real application while allowing us the flexibility to design it in accordance with our vision."

- Priya Kumari

In a whirlwind of a year that could have gotten to the best of us, we were reminded that the Spark! community is just as resilient as ever.

Spot is a mobile application that provides real-time crowd density data, allowing students to gauge building and floor occupancy levels in areas across the BU campus in order to virtually eliminate the time it takes to find the perfect spot to get work done.

A team of developers figured out how to compile student's mobile location data by partnering with BU’s IS&T Department to retrieve real-time stats from their PostgreSQL database.

“The program gave us an extensive support system of mentors that was vital to the development of this app,” said Priya Kumari, front-end developer.
IMPACT PARTNERSHIPS

making strides with alumni, industry, and foundations

In November of 2020, CDS was awarded two grants from New America's Public Interest Technology University Network (PIT-UN), for the establishment of a CDS Impact Hub of Equity and the Justice Media co-Lab.

The grant awards build upon the deep foundation that has been established through Spark!’s years of working with public interest technology projects as a strategy for providing students with vibrant academic experiences.

In April 2020, CDS, in collaboration with the Initiative on Cities, was awarded another PIT-UN grant to delve into the current state of the field of public interest tech. This landscape survey is ongoing and the analysis will be shared later in 2021.

These were the first major steps forward for Data Science for Good, an initiative anchored in CDS which leverages the expertise in BU’s 17 schools and colleges to accelerate the public interest technology work of faculty and students.

In summer of 2021, our first industry partnership was formalized. With a $1 million gift, MassMutual joins BU and CDS in envisioning and developing new kinds of data science education, research, and training experiences. This partnership further solidifies our commitment to widening the pipeline and closing the skills gap in computing and data sciences. The funds will go towards hiring a professor with real-world experience, diversity efforts, and undergraduate stipends.
The Center

the physical anchor of computing and data sciences at BU

So much progress has been made in the past year on the striking 19-story building, in the heart of campus, that will gather the mathematics and statistics and computer science departments, the Hariri Institute, and CDS under one roof. A community-centric Spark! space will greet visitors at the top of the stairs in the main entrance.

Construction started with geothermal well drilling. In mid-March, construction was halted as a result of City suspensions but in May Suffolk Construction crews safely remobilized and began laying gravel and pouring concrete. By summer 2021, the core was up along with steel framing (and more) of 14 floors.

At 350,000 square feet, it will not only be one of the biggest buildings in the University’s 180-year history, but also its most sustainable and energy-efficient ever, free of fossil fuels, with heating and cooling coming from those geothermal wells.

We’ll meet you at the corner of Commonwealth Avenue and Granby Street on January 3, 2023!
CDS set up an Executive Leadership Team and expanded to 36 its set of affiliated faculty (beyond the original 30 founding members), extending to 11 the number of academic departments represented within CDS for faculty governance.

The CDS administrative team also grew to five, and includes a Director of Administration, an Academic Programs Manager, a Financial Manager, a Communications Manager, and an Undergraduate Advisor. Throughout the year we have intentionally worked towards creating a culture of inclusion, collaboration, curiosity, and empathy.
Faculty growth: In addition to continuing to grow our core CDS faculty, in collaboration with other cognate units at BU, we will launch exciting searches for jointly-appointed faculty members who will embody the cross-cutting nature of CDS.

The Center: Moving into the new building will mark a monumental step for CDS, as we continue on our path to become the connecting tissue for computing and data sciences at BU.

Faculty and grad student rotations: These programs will promote CDS-enabled cross-disciplinary faculty research, PhD training, and curricular development.

Program development: Plans are underway to develop new programs, including a core Master's of Science in DS program and joint programs in collaboration with cognate units. DS4G, Impact Hubs, and co-Labs will continue to expand.

ON THE HORIZON for 2022

- Develop more experiential learning opportunities that catalyze "research that matters," as part of the MassMutual fund (through more Impact Hubs and co-Labs)
- Launch the undergraduate minor in DS, with two pathways - Basics of DS and Depth in DS
- Welcome the first cohort of CDS matriculates - the BU CDS Class of 2026!
ACKNOWLEDGEMENTS

We would like to thank the people who worked tirelessly on the content of this annual report as well as our core and affiliated faculty and the executive leadership team for their feedback and continued support. Thank you for your commitment, enthusiasm, and passion.

Editor, Designer, and Producer: Rosie Sullivan

Contributors: Elise Brown, Greta Bruce, Ziba Cranmer, Melissa Nesbitt