THE UNIVERSITY OF CHICAGO

**Data Science Initiative: Overview, Successes and Challenges**

**Overview and History:** The University of Chicago will announce the Data Science Initiative (DSI) this Fall. Publicly announcing the DSI has been a four-year journey. In 2017, Dan Nicolae, Chair of Statistics, and Mike Franklin, Chair of Computer Science, led the convening of faculty across campus to build a vision for data science that catalyzes new research and educational programs while expanding existing activities and assets. Building on the Computation Institute, the Center for Data and Computing (CDAC) was created in 2019 to serve as the intellectual hub and incubator for computation and data-driven research. CDAC funded 29 projects and workshops across 33 departments in support of interdisciplinary research. During the past two years, CDAC awarded $2.3M in seed grants to fund high risk, high reward research projects and supported over $63M in data science-related grant submissions. The DSI subsumes CDAC while expanding to a broader mandate that includes research, education and outreach. DSI is developing data science education at all levels and reaching out beyond the walls of the University to engage with industry, government, nonprofits and the Chicago community. The DSI is led by co-faculty directors Mike Franklin and Dan Nicolae. The day-to-day management, strategic planning, program development and fundraising is overseen by the Executive Director David Uminsky. Nick Feamster, Neubauer Professor of Computer Science, serves as the Director of Research and Rebecca Willet, Professor of Statistics and Computer Science, serves as the Director of AI. The DSI endeavors to understand the societal impacts of the data science work we do, to build an inclusive community and ensure we are training the next generation of diverse, ethical data science leaders, and to accelerate AI and data-driven discovery across all fields of inquiry.

**Funding Model**: The DSI will initially be funded through the Provost's Office. The long-term sustainable funding model includes revenue generating activities from an Industry Affiliate Program, professional and executive education programs, foundation and grant funding, philanthropy, and tuition from DSI affiliated degree-granting programs.

**Education**: The DSI will have undergraduate, graduate and professional education programs. Applied experiential learning is a core component of the DSI educational mission.
*Successes*: 1) The BA/BS in Data Science launches in Fall 2021 and builds on the popular Minor in Data Science that started in Fall 2019. Both the BA and BS programs require twelve prescribed courses and three electives. Two of the required courses are the Data Science Clinic sequence, which offers students the opportunity to work in teams on projects from a wide array of choices of industry, public interest organization or research lab partners. Students are required to take Ethics, Fairness, Responsibility and Privacy as well as Societal Impacts of Data Science. 2) The Civic Data and Technology Clinic launched in Fall 2020 and completed 16 projects with 12 social impact organizations working on climate, human rights, energy,

agriculture and marine technology; 46 undergraduate and graduate students participated in the pilot year. 3) The DSI oversees the Masters in Analytics and the Masters of Computational Analysis and Public Policy programs. Bringing both programs under the DSI umbrella aligns the programs with the DSI mission and research as well as provides students with experiential learning opportunities through the clinic.
*Challenges:* (i) Faculty recruitment and managing teaching loads for jointly appointed faculty. (ii) Securing experienced mentors for clinic projects. (iii) Coordination with other educational programs where data (and data science ideas, concepts and methodology) play a major role.

**Research**: The DSI transcends traditional boundaries by organizing our research into interdisciplinary themes focused on high impact results that can benefit society. These themes and focus areas will naturally evolve as the field of data science matures and DSI scales. Our multi-year research initiatives are led by faculty members and initially supported by DSI internal funding. Faculty and postdoc recruiting is integrated with research themes.
*Successes*: Our first three research initiatives extend beyond the traditional data science departments to include Data & Democracy (in collaboration with the Harris School of Public Policy and the Center for Effective Government), Internet Equity (in collaboration with Chicago Public Schools and the Crown School of Social Work, Policy and Practice) and AI & Science (in collaboration with the Pritzker School of Molecular Engineering and the Biological Sciences Division).
*Challenges*: (i) Recruiting faculty who will continue to shape data science as an emerging new discipline; (ii) Developing space solutions for cross-disciplinary research; (iii) Building connections on campus with other units that have a strong data science and AI component.

**Outreach**: The DSI engages outside of the walls of UChicago by developing meaningful partnerships with industry, government and the community through our education programs, research initiatives and the clinic. The Industry Affiliate Program connects our industry partners to data science research, emerging technologies and talent acquisition opportunities. The mission of the program is to lower the barrier to collaboration and encourage mutually beneficial long-term relationships. As part of the data, ethics and society focus of the DSI, researchers are building open-source tools and sustainable platforms to support external social impact organizations as well as creating education programs for workforce development.
*Successes:* 1) To meet our goal of fostering a diverse community of data scientists, we are engaged in collaborative partnerships with Howard, Atlantic Clark, Spelman, Morehouse, University of Illinois in Chicago, City Colleges of Chicago and Cal State Fresno to co-develop experiential data science curricula that work across the higher education spectrum. 2) The 11th Hour Project awarded the DSI a grant to serve as the centralized hub for software and data science for their 400 grantees across human rights, food and agriculture, energy and marine technology. The grant funds a staff data scientist and two software engineers who provide direct support to 11th Hour grantees and assist student teams working on clinic projects. 3) We have an MOU with the City Colleges of Chicago to develop an Associates in Data Science program with the goal of creating a pathway for students to the BS/BA in Data Science at UChicago.
*Challenges*: (i) Overcoming risk aversion among UChicago legal teams to take on the Industry Affiliate agreements remains an ongoing process.