Legacy Outcomes of the Information Frontiers Learning Initiative: An Inclusive Higher Education Program (2010 – 2020) of the Science of Information NSF Science and Technology Center

July 20, 2020

I. Executive Summary

Since its inception in 2010, the Center for Science of Information (CSoI) has designed and implemented an *Information Frontiers Learning (IFL) Initiative* focused on developmental training of a diverse next-generation science community while creating a science of information (SoI) curriculum for classroom and online learning. In considering sustained impacts beyond the upcoming post-NSF STC funded period of the Center, the program demonstrates a legacy of:

1. A diverse community of next-generation scientists that continues to collaborate and develop the emerging SoI field.

2. A SoI curriculum for all that offers fundamental and advanced knowledge and practices for current and future students.

3. Broader impacts through a set of best practices and lessons learned shared with the STEM community.

This white paper provides a descriptive background of our overall program outcomes, and also serves as a companion paper to the primarily visual presentation of the Center's integrated education and diversity legacy available at http://soihub.org/legacy . The Center's education and diversity legacy outcomes are described below in section III. Details regarding specific program activities are available at the Center's website: https://soihub.org .

II. CSoI Education Profile

The Center is comprised of 40 faculty and more than 200 students and postdoctoral fellows from the following institutions:

- Purdue University (lead institution)
- Bryn Mawr College
- Howard University
- Massachusetts Institute of Technology (MIT)
- Princeton University
- Stanford University
- Texas A&M University
- University of California, San Diego
- University of California, Berkeley
- University of Hawaii, Manoa
- University of Illinois, Urbana-Champaign

The cornerstone of the Center is the effort to increase understanding of, and to innovate, the scientific process of moving *from data to information to knowledge*. This effort has guided the Center's research objectives. Correspondingly, the Center's educational mission and goals support these objectives, as well as fully integrating diversity through broader participation efforts; the specific training activities offered, the online modules and courses developed, the courses our faculty teach, and the research training and mentorship of students have all provided a strong integrated connection between education, diversity, and research.

To facilitate this connection, the education program developed the IFL Initiative with a long-term vision of: (1) developing the next-generation of scientists who continue to strengthen the community of SoI, (2) creating a curriculum of modules, tutorials, courses, and teaching resources that are available to all and that provides information literacy at foundational and advanced levels, and (3)developing a philosophy and integrated pathways for broader participation across undergraduate, graduate, and postdoc levels of training. Therefore, three primary goals have supported the training of our students to increase their capacity for research in the interdisciplinary environment of the emerging SoI community:

1. Foster a community of practice in the science of information. Efforts include year-long and focused training activities where students can enhance their knowledge, learn new methods and tools, interact with peers and faculty, and collaborate in teams leading to interdisciplinary experiences. These have contributed significant value to our students' departmental experiences.

2. ***Increase awareness and knowledge in science of information in the broader community.***
A dynamic range of online and classroom modules, courses, and resources have been established serving the broader community. A searchable online learning hub organizes these resources and makes them freely available.

3. **Fully integrate diversity** by employing an expansive view of broader participation at both individual and community levels where CSoI is a catalyst for diversity across our entire program including staff/PI's, undergraduate, graduate, and postdoctoral levels.

III. Legacy Outcomes

1. A network of next-generation scientists that continues to collaborate and develop the emerging science of information field.

1.1 Building a Community of Practice

The Center quadrupled its membership of students and postdocs conducting research and education with CSoI faculty from 53 in project period one, peaking in period nine at 261, ending with 226 in our period 10 activity (Figure 1). The large majority of CSoI students (>94%) have stated that they value participation in CSoI because it provides a venue

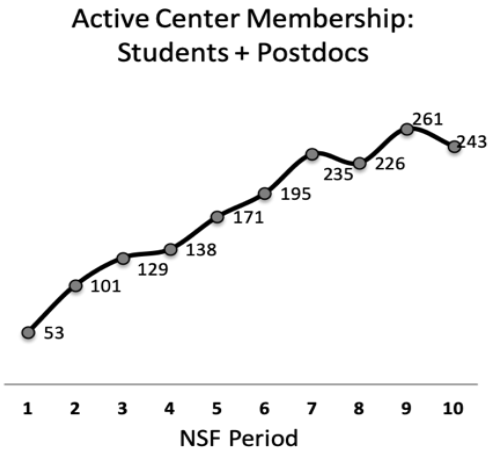

Figure 1. Student and Postdoc Membership

for stimulating their own thinking about the SoI, provides productive networking and training opportunities with peers and leading researchers in the field, and fosters meaningful collaboration among students and faculty.

The primary indicator of a working community of practice is that its members collaborate with meaningful outcomes (Wenger, et. al. 2002). With CSoI as a catalyst our long-term efforts to build a community of practice at the student and postdoc level have shown promise. Ongoing outcomes have been monitored annually showing that the percentage of our students and post-docs engaged in collaboration with other members of the CSoI network beyond their major professor rose significantly from 11% in period 2 to a peak of 52% of members by period 6. (Figure 2). CSoI community collaboration then became established between 45% to 50% of the student membership through period 10 (Ladd, 2019).

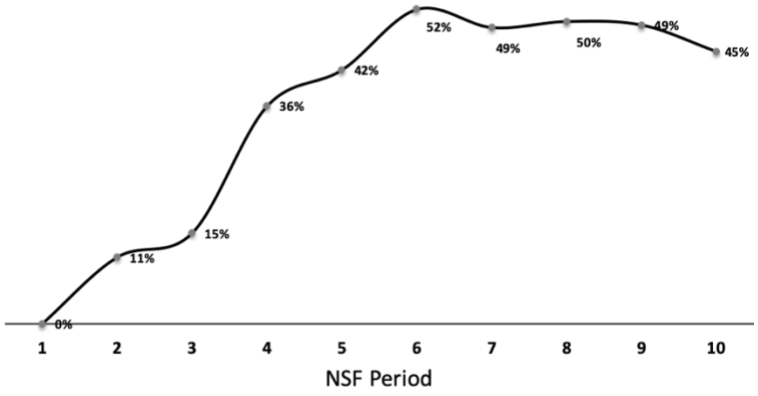

Figure 2. Percent of student/postdoc membership engaging in collaboration on research each NSF period within the CSoI network.

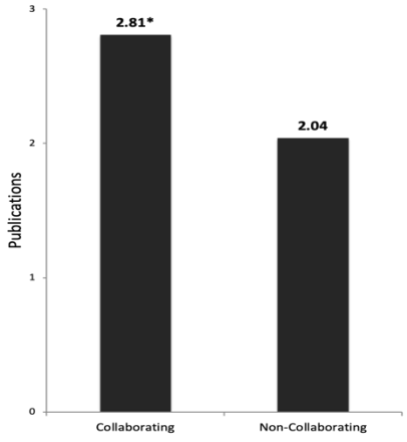

Figure 3. Mean Annual Publications 2011 - 2018 Comparing Collaborating and Non-Collaborating Graduate Students

Such collaboration in our community aided the productivity of our graduate students who collaborated on research within the CSoI organization. An internal study of our collaborating graduate students (periods 2 – 8) demonstrated they were significantly more productive in publishing research compared with their peer members who did not engage in Center collaborations (2.81 vs. 2.04 annual publications, n=256, $F=11.89$, $p < 0.001$, Figure 3). A multi-factorial analysis demonstrated that this result is due primarily to students engaging in the collaboration itself, and not to other factors such as gender, funding source, or institution (Ladd, 2018).

The SoI community in academia has been further strengthened with more than 70 of our alumni attaining faculty positions. Of the students and postdocs that participated in student team science projects described below that later joined academia as faculty members, 32% of their post-CSoI published research has been in collaboration with CSoI members (as of 2018). This is early evidence that the relationships and network of SoI researchers built during the NSF funded period of the Center may continue to expand the SoI community. Overall, the CSoI has graduated 417 alums with half matriculating to academia as graduate students, postdocs, and faculty, with the other half joining industry.

1.2 Pathways Supporting CSoI Collaboration

Throughout the CSoI membership, both informal and formal pathways were developed encouraging student and faculty collaboration. Informal pathways included access to the member network, CSoI sponsored conference sessions, summer schools, SoI research days, seminar series, and annual CSoI membership meetings and poster sessions. Formal pathways for engaging in collaboration included faculty interdisciplinary seed projects, co-advisors for students, academic year research for undergraduate training program, center-wide postdoctoral fellowship program, student training workshops and student-led research workshop teams.

A primary example of a collaborative environment that integrated broader participation at student, institutional, and domain levels is our interdisciplinary research and data skills training workshops and resulting year-long collaborative student-led teams. Students that participated in these trainings since 2012 have hailed from 23 distinct departments at 28 colleges and universities. Broader participation outreach and engagement created a rich multidisciplinary environment for training to take place (Ladd and Ward, 2019). This ongoing training activity provided follow-up professional development through NSF-style grant proposal writing for team science. We learned this was often the first proposal writing that students or postdocs were involved. Teams with successful proposals were supported and facilitated to continue research collaborations toward co-discovery and team generated solutions. To date, the program has supported 19 multi-institutional interdisciplinary student and postdoc-led research teams. These teams reflect the aforementioned domain diversity as well as achieving near a 1:1 male to female ratio overall. The teams have thus far produced 25 co-authored journal papers, and over 50 conference papers and posters. Students express a significantly enhanced ability to share their research and concepts with others outside of their domain areas, productively bridging across disciplines to solve complex data-related problems, while asking new questions from an innovative interdisciplinary perspective.

This interdisciplinary and multi-institutional team science approach for training appears to become integrated into how students and postdocs view the scientific enterprise. As these alumnae matriculate to post-doctoral and faculty positions they continue the practice of collaboration, and several have sent their own students and postdocs to participate in CSoI training activities.

2. A Science of Information Curriculum for All

The second overarching educational goal focused on developing modules and courses with fundamental and advanced SoI content for both classroom and online learners. We designed curriculum pathways for core undergraduate majors, PhD-level learners, and general audiences.

2.1 Classroom

CSoI institutional partners have developed new SoI related courses and content, with 44 CSoI faculty having developed more than 100 courses with SoI topics.
A large number of students (7,000) subsequently completed these classroom-based SoI courses advancing their knowledge and skills.

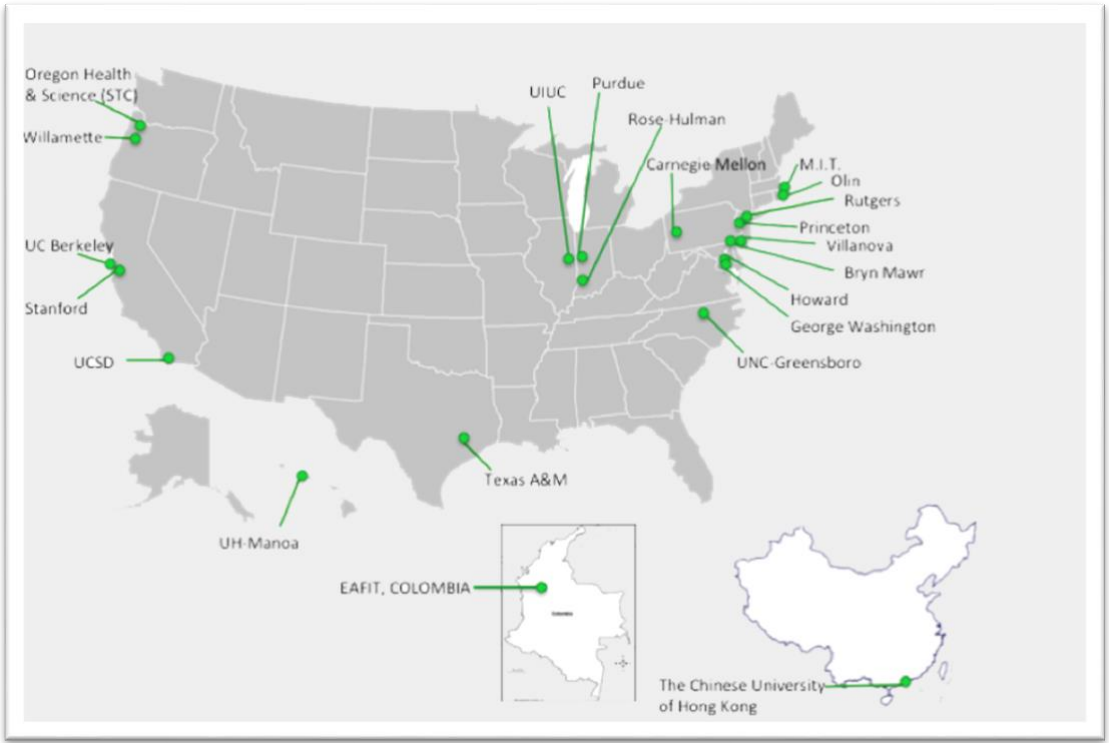

Figure 4. Courses associated with the Center showing university locations. All 11 Center partners, and 11 additional universities are represented. 7,086 students enrolled at our partner institutions in classroom-based SoI courses during periods 1-10.

Faculty training workshops focused on concepts of teaching the new SoI curriculum led to an additional 11 universities beyond the Center partnership establishing courses for their students. The overall effort has involved faculty from 22 universities teaching SoI (Figure 4). Just one example is computer science faculty member Professor Juan La Linde at EAFIT, Colombia. Dr. La Linde adapted our Introduction to SoI semester course for undergrads taught in Spanish with both classroom and online versions.

2.2 Online

Online topic modules, advanced research seminars, and educational resources built on 1,144 video-based learning tutorials have reached over 300,000 learners globally, including 184 countries and all 50 U.S. states. An additional 60,000 learners enrolled in our FutureLearn and EdX platform courses. The largest demographic of learners enrolled in our MOOC courses live in the United States or the United Kingdom. These MOOC's have served to drive additional traffic to our soihub.org online educational content. The use of multiple online platforms has in turn allowed us to reach students around the world interested in gaining knowledge in the SoI. The SoI online curriculum and links to many classroom-based courses taught by our faculty are available at our learning hub: https://learninghub.soihub.org

2.3 Learning Outcomes

An evaluation of learning outcomes from a subset of eight SoI courses was undertaken in 2015 in coordination with faculty and students at U.S. institutions. As outcomes of completing one of these courses, the large majority of students reported moderate to significant increases in areas of multidisciplinary understanding, SoI skills, SoI awareness, information literacy, and data science problem-solving ability.

Table 1. Science of Information Courses Evaluation: Percent of students with increases for indicators tied to learning objectives *(Likert scale with four levels of change; maximum rating = 4).*

Indicator	Significant Increase	Moderate Increase	Slight Increase	No Change	Mean (max = 4)
Information Literacy	40.5	38.1	19.0	2.4	3.17
Data Skills	39.7	41.4	15.5	3.4	3.17
Multidisciplinary Understanding	37.5	43.8	14.6	4.2	3.15
SoI Awareness	38.6	42.1	14.0	5.3	3.14
Problem Solving Ability	25.0	41.7	29.2	4.2	2.88

SoI education content continues to be offered at participating universities and will be available online for future students to engage in learning the fundamental and advanced concepts supporting the SoI field.

3. Broadening the Science of Information Community

3.1 Institutional and Domain Diversity

Employing evidence-based design of a series of workshops, summer schools, and student-led research projects, the IFL Initiative has infused significant value into students' PhD, undergraduate, and postdoctoral training. These efforts specifically have supported team science collaborations, information and data science skills, and productive exchanges of research supporting the emerging SoI field. Collectively these have led to a robust network of CSoI students, postdocs, and faculty. CSoI education events have involved more than 4,500 students from 129 universities around the world, significantly broadening the impact and awareness of CSoI's mission (Figure 5). The range of universities represented includes R1, R2, M1, Baccalaureate, MSI's, and special focus institutions infusing diversity at the institutional level. The breadth of domain areas involved includes well over 20 distinct disciplinary areas helping to bridge and make relevant the emerging SoI field to many disciplines (Ladd, 2019).

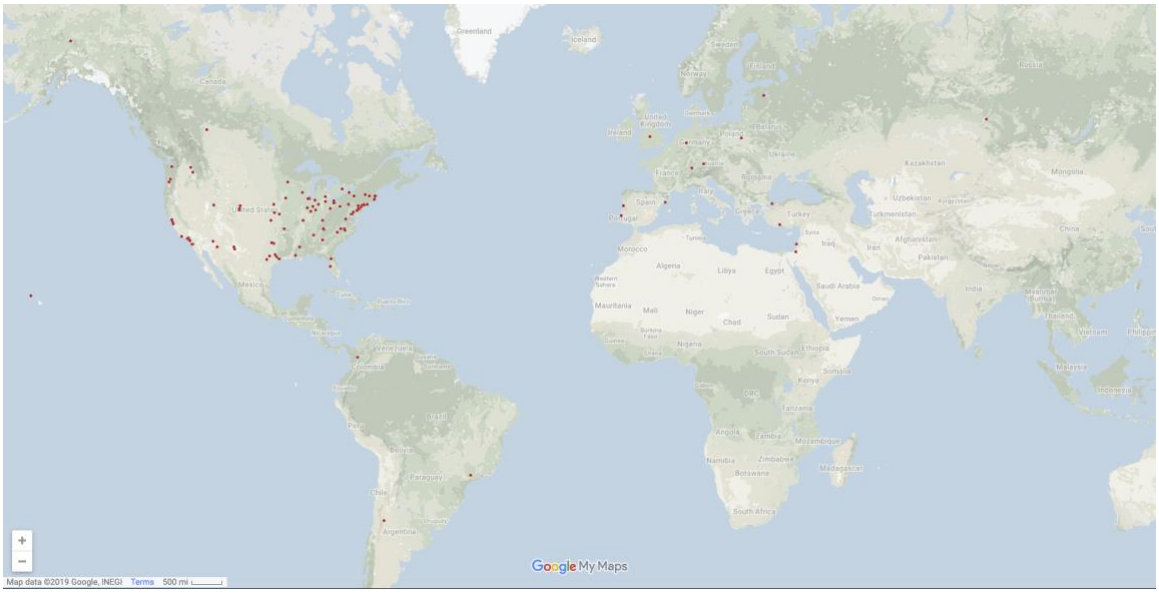

Figure 5. Locations of universities (129) represented by attendees at CSoI education events totaling 4,571 attendees during 2010 – 2020.

3.2 Individual Diversity

Intertwined with institutional and disciplinary diversity is integrated diversity at the individual level. The CSoI created multiple pathways for broadening participation across undergraduate, graduate, and postdoc levels (Ladd and Brown, 2019; Andronicos and Ladd, 2018).

At the undergraduate level our academic year research training for undergraduates, the Channels Scholars program, has as of the summer of 2020, involved 128 students representing 17 universities and 31 faculty mentors. Channels Scholars participate in a research experience related to SoI, while receiving guidance from a university faculty mentor. Working on their research projects for 40+ hrs/month during the regular academic year results in students gaining significant experience, skills, and insight into the scientific enterprise. Most Channels Scholars have then entered into an internship experience during the summer months. More than fifty companies have been involved in internships with our Channels Scholars.

The baselines for undergraduates in computer science and electrical engineering fields in the U.S. are 18% women, 10% minority, and 79% U.S. citizen or permanent resident. Our Channels Scholars program demographic outcomes are 59% women, 27% minority, and 91% U.S. citizen or permanent resident.

At the graduate training level there are two primary pathways for participation. One is at the research or teaching assistantship level as a full-time graduate student being mentored by a faculty member of the Center. Although the CSoI does not directly select graduate students – this is at the discretion of the participating faculty and departments at the partner schools – overall, CSoI graduate student membership steadily broadened during our ten NSF periods resulting in our reporting graduate

population as 30.7% women, 8.5% minority, and 40.4% U.S. citizen or permanent resident (compared with U.S. graduate student baselines in CS and EE of 21% women, 4% minority, and 32% U.S citizen).

The other primary pathway of participation in CSoI at the graduate level has been direct training opportunities such as our annual summer intensive data science and interdisciplinary team workshops. The Center again steadily broadened participation in these trainings. Demographics for 2018 and 2019 trainings resulted in participation demographics of 67% women, 35% minority, and 45% U.S. citizen or permanent resident.

At the postdoctoral level, the Center-wide postdoc fellowship integrates diversity at the institutional, domain, and individual demographics levels. This fellowship specifically involves postdocs working with two or three CSoI faculty who are located at two or three of the partner institutions and may also work in different fields. Fully 50% of the hires of this program have increased overall diversity of women, minority, and/or U.S. citizens/permanent residents of the Center.

The elements of the philosophy and activity of the Center regarding broadening participation is mapped out in a downloadable pdf available here: https://soihub.org/legacy/posters/Mind-Map-CSoI-Broadening-Participation.pdf

4. Broader Impacts Through a Set of Best Practices and Lessons Learned Shared with the STEM Community.

4.1 Knowledge Dissemination to the STEM Community

The primary recommendation to the education program from the NSF review board in January of 2018 was to establish the impact of the educational activities before NSF STC funding ends. Since then, we've engaged in program evaluation and disseminating knowledge and lessons learned to the broader STEM community. The education and diversity team has presented the Center's programs at these recent conferences (abstracts available at http://soihub.org/legacy/lessonslearned.html):

Byrd V., Roark K., and Ladd B.T. 2020. Usability of Data Visualization Activity Worksheets in the Context of a Critical Data Visualization Workshop: Findings from a Usability Survey. June 26, 2020, American Society for Engineering Education, 127th Annual Conference, Montreal, CAN (presented virtually)

Ladd, B.T. 2019. The Information Frontiers Program: Expanding Student Capacity for Crossing Domain and Institutional Borders. October 24, 2019, Association for Interdisciplinary Studies 41st Annual Conference, University van Amsterdam, Netherlands

Ladd, B.T. and Ward, M.D. 2019. Training Students Concurrently in Data Science and Team Science. July 28, 2019, American Statistical Association, Joint Statistical Meetings, Denver, Colorado, USA

Ladd, B.T. and Brown, R.E. 2019. Broader Impacts of the Information Frontiers Integrated Education and Diversity Program. May 1, 2019, National Alliance for Broader Impacts Summit, Tucson, Arizona, USA

Ladd, B.T. 2018. Case Study of Interdisciplinary Student Research Teams: Factors, Outcomes, and Lessons Learned. May 22, 2018, Institute for Translational Science, Science of Team Science Conference, Galveston, Texas, USA

Andronicos, K. and Ladd, B.T. 2018. Broadening Participation in the Science of Information. January 8, 2018, NSF Includes Summit, Washington DC, USA

Throughout the Center's existence, the team has participated in giving presentations and workshops for the broader science community. Examples include organizing a series of workshops for faculty and postdocs aimed specifically at teaching SoI topics and courses. These workshops shared a foundation of lessons learned from CSoI faculty teaching pilot courses and establishing topics at their respective institutions. Likewise, workshops at conferences have helped to exchange knowledge and lessons learned, for example, at NSF STC conferences and professional development workshops, ACM's SIGCSE Conferences, and a National Online Learning conference. The Center's education team was also invited to share knowledge and assist new science education and research projects at NSF's CISE Directorate, Purdue University's Engineering Research Center, M.I.T.'s STC Center for Brains, Minds, and Machines, University of Pennsylvania's STC Center for Engineering MechanoBiology, and the NSF STC Program Evaluation Committee.

4.2 Key Lessons Learned

In conclusion, we share a summary of four primary lessons learned that led to successful education and diversity outcomes in the context of operating a national science and technology center:

What will position the program as a catalyst for building a science community while providing desirable skills and knowledge? Answering this question requires two additional questions: 1) *what is our niche – that topic, area, or context where we excel beyond most others?* For us, in the educational arm of the Center, it is training students in the emerging data and information science field within the context of student-led interdisciplinary team science. The other side of this coin is to refrain from duplicating what others are already doing much better than we could offer. For example, we arrived at our niche in part by recognizing what our partner institutions already were achieving at a high level in their departmental programs and we then made the decision to purposely not duplicate those activities. The second question we had to answer is, *2) What makes our project/group/organization more than just the sum of its parts?* For CSoI this was in understanding how to recognize and leverage individual strengths by making sure that the Center constantly positioned itself as a catalyst; always bringing people together in ways that offered unique opportunities and added value to individual researcher and student efforts.

Build something that has a chance of leaving a legacy beyond the life of the project. For a large scale, longer running project like CSoI we endeavored to build a community of practice around the emerging field of SoI. At the educational level of the Center this meant developing multiple pathways for interdisciplinary training, exchange, learning, team research and applications that students could access, participate in, and add significant value to their undergraduate, graduate or postdoc experience. In turn, many of our students have commented that these experiences led to knowledge and skills that helped them land faculty and postdoc positions, as well as senior positions in industry. They are perhaps our primary legacy.

Develop courses and learning materials and methods that fill the gaps and needs in the curriculum. We used CSoI as a catalyst to become the hub for the wider student and professional community in our emerging science by supporting and coordinating new and useful courses, modules, and topic trainings. This often involved our faculty translating their new research findings into the classroom in the form

of new and adapted courses, while the education team coordinated the development of introductory-to-foundational-to-advanced topics into online offerings reaching students beyond our partner institutions. This made us relevant to students and faculty far beyond our immediate membership.

Commit to inclusive diversity by integrating diversity into everything we do. In the early years of our Center we had to learn not to delegate diversity as a separate or stand-alone program. We had to make a deeper level commit to diversity and inclusion. For CSoI, this in turn made us relevant to the entire community, meaning the full spectrum of participants and domain areas, by integrating a valued and expansive understanding of diversity in all trainings and opportunities. It was necessary to have this commitment at the CSoI leadership level. This then allowed inclusive diversity to be integrated with funding commitments, professional development efforts, and in creating pathways for participation. It also meant CSoI staff made it personal in their own professional development, while following through with underrepresented individuals. CSoI was again a catalyst by serving as a high-profile platform for spotlighting women, minorities, and U.S. citizens through prestige lectures, summer school key note speakers, workshop instructors, student seminar speakers, spotlight interviews, postdoc fellows and REU scholars.

Acknowledgements:

This material is based upon work supported by the National Science Foundation under Grant No. CCF-0939370. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Author: Brent T. Ladd, with key contributions and/or review by the following Center for Science of Information staff and faculty members (listed alphabetically): Kelly Andronicos, Mike Atwell, Robert E. Brown, Todd Coleman, Barbara Gibson, Ananth Grama, Deepak Kumar, Wojciech Szpankowski, Mark Daniel Ward.

The Center thanks the many individuals who have served on our advisory board and NSF review committees, as well as the faculty and student community of CSoI together with the faculty and students beyond the Center who have all participated in shaping our educational efforts in the emerging field of science of information.

The Center's programs have received additional facility, staff, and financial support from Purdue Universities' Computer Science Department, College of Science, and Vice President for Research, as well as our partner institutions listed above in section II.