

Wisconsin in the Information Age

A Report by the Working Group on Computing at the University of Wisconsin-Madison

INTRODUCTION

In January 2018, Chancellor Rebecca Blank authorized the creation of a 12-person UW-Madison Working Group on Computing, drawn from campus, alumni, and industry. The group was charged with producing a report that would advise the UW leadership, the CS department, its Board of Visitors, and other stakeholders on the opportunities available to advance a computing agenda led by UW-Madison that will:

- Increase the profile, rankings and research output of our computer science faculty;
- Increase opportunities for students across campus to study computer science; and
- Produce more students among our graduates who have been exposed to and trained on computational thinking, big data, artificial intelligence, and related fields.

Given the broad nature of this charge, we use the term “computing” in its broadest sense here and throughout, including computer, data, information science and related disciplines, and usages of computers for automation, analysis, and related tasks¹. Investigating applications of big data and artificial intelligence led us to look at what our peers and UW-Madison are doing in data science, which was born out of big data. While computing and data science are intrinsically related, they are clearly not the same thing. However, for ease of exposition, we have used these terms interchangeably at times throughout this report.

Central to the working group’s efforts and conclusions was the creation of the *Wisconsin Computing Idea*, described in detail later. The concept of *core*, *connections*, and *enrichment*² introduced there is a recognition that the disciplines of computing, data science, and information science are practiced virtually everywhere across the UW-Madison campus. No entity can “own” the application of these disciplines, in the same manner as no one “owns” math or physics. Yet, the existence of a *core*, just as in math and physics, is needed to advance the teaching and research of the discipline itself. Excellence in the core ensures meeting the first objective of the charge. *Connections* facilitate the proliferation of these new sciences across campus to ensure the second and third elements of the charge are realized. *Enrichment* introduces the need to teach and research the societal impacts of computing and how the applications of big data, artificial intelligence, robotics, and others intersect with the social sciences. Like computing and data science, the facets of enrichment are found in many areas across campus, and the trend at our peers is to develop such facets within the framework of information science.

The working group extensively discussed computing in numerous face-to-face meetings, phone calls, and emails. We reached out and interviewed the current and former leaders of our country’s top academic computer, data science, and information science programs. We held extensive discussions with administrators, educators, and researchers on the UW-Madison campus. We had comprehensive discussions with the Wisconsin business community and government. We also spoke with numerous students at the undergraduate, graduate, and professional levels, in many different disciplines.

The result of our work is described in this report, which is organized into three main sections. In Part I, we discuss computing and data science, and their increasing impact on the campus, state, and society at large. In Part II, we present an analysis of what our peers have been doing, focusing on their organization, leadership, facilities, faculty expansion, computer and data science initiatives, educational offerings, and fundraising efforts. In Part III, we develop an action

¹ This point, and our reasoning, is discussed in detail in Appendix A - The Working Group Charge.

² The working group borrowed “Core, Connections, and Enrichment” from U.C. Berkeley’s “Data Science Planning Initiative”, 2016. Summary: <https://www.stat.berkeley.edu/~binyu/ps/FAB-Summary2016.pdf>

plan, based around the Wisconsin Computing Idea; we propose specific actions in both the near- and medium-term, centering upon leadership and organizational structure. A companion executive summary has also been created.

PART I: THE GROWING IMPACT OF COMPUTING AND DATA SCIENCE

THE GROWTH OF COMPUTING AND ITS PERVASIVE IMPACT ON THE WORLD

Computing emerged as a new field in the 1950s. The development of the field has been nothing short of breathtaking. From the 1950s through the 1970s, computing was mostly confined to large corporations and was used to perform mostly administrative tasks such as payroll and accounting. Starting in the late 1970s, with Bill Gates' vision of "a PC on every desk," computing began to seep into every aspect of our lives. Companies large and small, and increasingly, individuals were using computers for a wide range of activities, spanning work, communication, and recreation. In the mid-1990s advances in computing began cascading through our society, with each new trend appearing more and more rapidly. Consider some of the things that emerged during that period: the Internet; hardware/software advances such as solid-state drives and open-source software; cloud computing; sensors; mobile devices such as smartphones and tablets; social media; and big data.

These trends accelerated the pace of change in computing and made it genuinely pervasive. Computing is now "eating the world," to paraphrase Marc Andreessen, co-inventor of the Netscape browser. It has caused major upheavals, obviated certain types of jobs, transformed existing jobs, and created entirely new types of jobs. For example, automation is accelerating, eliminating many manual jobs, and e-commerce is destroying a large swath of bricks and mortar businesses. Virtually all job types now need some computing and data skills. Today, five of the ten largest companies in the world are in computing, three of which (Google, Amazon, Facebook) did not even exist 25 years ago. A popular saying aptly illustrates these striking upheavals: "The world's largest transportation company owns no vehicles (Uber), the world's most popular media company creates no content (Facebook), and the world's largest accommodation provider owns no property (Airbnb)."

The Industrial Revolution has given in to the Information Age. It was thousands of years between the agricultural and industrial revolutions, yet less than two hundred before the Information Age has begun to transform the way we live, this time in a global fashion at a pace never before experienced. Moreover, we are only at the beginning, with even more dramatic changes coming, involving new areas like data science, artificial intelligence, virtual reality, smart manufacturing, 3D printing, self-driving cars, agricultural drones, genome editing, and the Internet of Things. These new fields arise at a dizzying pace and rely on computing and data more than ever.

As the Wall Street Journal noted recently, the race to become the world's first fifth-generation technology country is well underway. While applications to tap the full potential of 5G, such as self-driving cars, virtual reality, and remote surgery, remain some years away, leading the way will be transformative to the economy while requiring computational solutions that enhance consumer needs, cybersecurity and more.³

IMPACT OF COMPUTING ON UW-MADISON AND THE RISE OF DATA SCIENCE

As in the world at large, the impact of computing and data science on UW-Madison has been pervasive and transformative. It is changing virtually all essential facets of UW-Madison, including scientific discovery, research funding, education and training, jobs and outreach, and intellectual leadership.

Impact of Computing and Data Science on Scientific Discovery

Computing has been introducing new paradigms for scientific discovery and radically changing the way many UW-Madison scientists do research. Historically, scientists used experimentation and theory to do research. Starting in the 1960s, however, many science fields became increasingly computing driven, using great quantities of computing

³ *Why Being First in 5G Matters*, The Wall Street Journal, Sep 13, 2018, <https://www.wsj.com/articles/why-being-first-in-5g-matters-1536804360>

power to solve or simulate mathematical models. Computation and simulation were soon recognized as the third paradigm of scientific discovery.

In the past two decades, advances in computing have made it possible to capture and store large quantities of data inexpensively. As a result, many science fields are becoming data-driven, in which scientists collect substantial data, then analyze it to discover insights. *Such data-driven discovery is now considered the fourth paradigm of science. The science of collecting, processing, and analyzing a large amount of data is now known as data science. At UW-Madison, science fields ranging from astronomy to zoology are now relying heavily on data science to make progress.* For example, researchers in the School of Mass Communication & Journalism want to collect and analyze hundreds of millions of tweets to understand how people communicate on social media about the U.S. presidential election. Researchers in the Department of Agricultural & Applied Economics want to collect and analyze Web data from hundreds of thousands of non-profit organizations in the U.S. to understand how they spend research dollars and the impact of such spending on the local economies. Researchers in the Department of Geography want to acquire and analyze data about buy/sell transactions regarding cattle in Brazil, to understand how cattle ranches that destroy the Amazon forest disguise the origin of their cattle. Closer to home, researchers in the Center for Limnology are analyzing data collected from thousands of lakes in Wisconsin to understand and predict algae growth that may deplete oxygen and kill off fish.

Impact of Computing and Data Science on Research Funding

As data science revolutionizes scientific fields, the nature of raising funds changes. Funding from a wide variety of sources (including U.S. government agencies, foreign governments, industry, universities, and non-profit organizations) is pouring into science projects that utilize massive amounts of data and computation. For example, in 2012 the White House announced that six federal agencies would invest more than \$200 million into Big Data R&D initiatives. In the past two years, both the National Science Foundation (NSF) and the National Institutes of Health (NIH) have convened workshops to study data science and issued strategic plans for data science. The NSF's 2017 data science funding alone totaled tens of millions of dollars. The NIH in 2015 provided more than \$100 million to fund 13 centers for "Big Data to Knowledge." The Department of Defense's "Data to Decisions" program will invest \$250 million annually. The Gordon and Betty Moore Foundation and Alfred P. Sloan Foundation in 2013 provided \$37.8 million to fund data science efforts at New York University, UC-Berkeley, and the University of Washington, and that same year the Grainger Foundation provided \$100 million to the University of Illinois, \$20 million earmarked for hiring faculty in data science. In September 2018, the Defense Advanced Research Projects Agency (DARPA) announced a \$2 billion campaign to invest in artificial intelligence (AI) technologies.

The amount of funding pouring into data-driven science projects as a fraction of total available funding has been dramatically increasing. *As a result, there is a growing consensus on the UW-Madison campus that data science is critical for the future of research funding and thus research spending.* Historically, the university's research enterprise is a powerful economic engine (as well as a creator of knowledge and innovation). The benefits are felt throughout the state. According to the 2016-17 UW-Madison Data Digest, more than 362 UW-Madison-related startup companies support nearly 25,000 jobs, generate \$113.6 million in state tax revenue and contribute \$2.3 billion to the state economy. The UW-Madison's economic impact on Wisconsin is \$15 billion annually. However, this powerful economic engine is currently in jeopardy. UW-Madison had been in the top five universities in the nation in research spending every year since 1972, but dropped to 6th in 2016. *Investment in computing and data science is key to helping UW-Madison scientists across all disciplines compete more effectively, in turn securing more funds and accelerating and extending innovation which is core to the Wisconsin Idea.*

Impact of Computing and Data Science on Education and Training

While data science straddles many fields -- notably statistics, computer science, and information science -- it is transforming practically every other discipline across academia, industry, government, and society at large. A huge challenge today is that the volume and variety of data now measured have gotten ahead of our ability to make sense of these data, to effectively use the power of computing to solve emerging problems. These problems require interdisciplinary teams with members trained in computing and data science methods.

The rise of computing and data science is stressing education and training at UW-Madison in multiple ways:

1. A growing number of students across the campus want to learn more computing and data skills. They realize that no matter what occupations they pursue in the future, they will need these skills to survive. For example, demand for CS 301 (Data Programming), a course the Department of Computer Sciences offers to introduce non-CS students to computing and data, has significantly increased, to 795 students in 2017-2018 from 573 students in 2016-2017. For perspective, UC-Berkeley's Introduction to Data Science class added more than 550 students in 2015-2016, and well over 1,000 students in 2016-2017.
2. As domain sciences become increasingly data-driven, many scientists and graduate students are clamoring to learn more computing and data concepts and skills. They currently have just two options on the UW-Madison campus, neither of which is ideal. They can attend short workshops such as data carpentry and software carpentry, which are often too basic. Alternatively, they can take current CS and Statistics course offerings, which are often not appropriate because they require extensive CS or Statistics background.
3. With computing and data jobs in high demand, many students want to major in CS, Statistics, and related areas. As a result, the number of CS majors alone has doubled in the past five years. In 2017, CS became the most popular major on campus, with more than 1,600 students. The CS department has struggled to find the resources to meet this rocketing demand, and this trend is showing no sign of slowing down.
4. Many working professionals who are not CS developers are eager for training to become CS developers. Meanwhile, working CS developers want to learn new skills such as machine learning. There is high demand in these areas at Wisconsin companies like Epic Systems, American Family Insurance, and Johnson Controls, to name a few. CS has offered some programs for working professionals, but the Department's constrained resources limit these offerings, leaving considerable demand and potential revenue on the table. The inability to create and offer additional professional programs has resulted in situations such as Milwaukee-based Johnson Controls encouraging its employees to take professional courses from the University of Illinois. We are missing an opportunity and failing to shoulder our responsibility to address the strong and growing need to train Wisconsin's labor force to transition to tech jobs or upgrade their computing or data science knowledge.

As the computing revolution pushes deeper into everyday life, UW-Madison also needs to re-evaluate and strengthen its role in educating Wisconsinites in computing-related issues, such as how to use data responsibly, respond to fake news, preserve privacy on social media, use computing and data tools effectively and understand computing-related laws and regulations.

Impact of Computing and Data Science on the Role of UW-Madison as a Powerful Economic Engine

American universities often act as powerful economic engines for their local regions and beyond. Consider, for example, the impact of Harvard/MIT on the Boston region, Carnegie Mellon on Pittsburgh, Stanford/UC-Berkeley on Silicon Valley, and the University of Washington on Seattle. *UW-Madison is no exception. It has long been a powerful economic engine for Wisconsin and beyond.* UW-Madison related startup companies support nearly 25,000 jobs and contribute \$2.3 billion to the state economy. UW-Madison has helped the greater Madison area, in particular, develop as a hub for biotechnology and healthcare startups.

Computing and data science are emerging as a dominant component of this economic engine. Biotech startups increasingly need robust computing and data support. Computing and data startups are proliferating, and they are making increasingly more economic impacts. Epic Systems, the best-known example of this in Wisconsin, was founded in 1979 by Judith Faulkner, a CS alumna. Epic has about 9,600 employees; some whom have started companies in Wisconsin after leaving Epic, creating a snowball effect. Among the Wisconsin-based high-tech startups founded by CS alumni are: Adobo, Hardin Design and Development, EatStreet, HealthMyne, Ionic, Nordic Consulting, PerBlue, Propeller Health, VidMaker, SimpleMachines, and DataChat. Virtually all of these companies started in the last decade or so. Strong CS, data science, and other related programs at UW-Madison have helped land Madison "at a turning

point” in becoming recognized as a hub for technological innovation and entrepreneurship (Atlantic Council). Madison is recognized as a top ten tech city by real estate firm Cushman & Wakefield; one of the “next top tech towns” by the National Realtors Association; and one of the best cities for high tech jobs by Wall Street 24/7.

Consequently, Amazon, Google, Microsoft, ZenDesk, and other high-tech companies have opened offices in Madison, to attract university-trained talent and develop closer ties with its researchers. Foxconn, which is developing a large manufacturing operation in Southeast Wisconsin, has also sought closer collaborations with UW-Madison and other UW campuses and recently pledged to invest \$100 million in the state’s flagship university to establish Foxconn Institute for Research in Science and Technology. *For UW-Madison to remain a powerful economic engine and fulfill the potential Foxconn and others see, it is critical to boost its computing and data prowess and use those capabilities to boost Wisconsin’s high-tech sector, as well as other sectors such as biotech, precision medicine, and agriculture.*

Impact of Computing and Data Science on the Intellectual Leadership of UW-Madison

American universities do more than educate, perform research, raise funds, and function as economic engines. Another essential role is to provide intellectual leadership. Universities help society understand future trends, how they might transform the world, and what actions should be taken. For example, university experts can help legislators craft laws and regulations involving technologies ranging from self-driving cars to social media, drones, and the Internet of Things. UW-Madison is no exception. It has been relentless in its pursuit of truth and serves as a distinguished learning center and one of the world’s great research institutions. So it follows that *if computing is transforming the world, to remain a leader in the Information Age, UW-Madison must be a leader in computing.*

IMPACT OF COMPUTING AND DATA SCIENCE ON THE STATE OF WISCONSIN

Computing and data science are increasingly central to our state's economy. It is transforming existing sectors and creating new ones. It generates many jobs that currently go unfilled because we cannot educate enough workers. Computing is also changing the way people work, shop, and relax, profoundly impacting rural Wisconsin. It is critical for the state to have a world-class knowledge center in computing and data science to manage the disruptions computing creates, such as job loss in manufacturing due to automation and company relocation, and the loss of local shops and spending power due to e-commerce.

Transforming Existing Major Economic Sectors

Computing is transforming virtually all existing economic sectors in the world, and Wisconsin is no exception. Manufacturing, agriculture, healthcare, and financial services, which have long been strong pillars of the state's economy, are undergoing profound transitions. Automation is transforming manufacturing methods and processes. Smart, Internet-enabled devices are raising numerous computing and data challenges. The more we can help companies here address these challenges, the more vibrant our economy will be. Companies here are working on such issues today. Johnson Controls, a Fortune 500 company with significant operations in Milwaukee, has made significant investments into the "Internet of Buildings," working on Internet-connected devices that capture and transmit large amounts of data that can be analyzed in real time to detect problems such as home intrusions or heating device failures. Foxconn Technology Group, a Taiwan-based manufacturer of electronics that made Wisconsin its North American home, is working on technologies to develop smart homes and cities. The more expert assistance we can provide, the more markets for advanced electronics Foxconn will be able to open.

Agriculture, one of Wisconsin’s legacy industries, is massively impacted by computing and data science. Here is what John Shutske had to say in “The Wisconsin Farmer” on Jan. 31, 2017: *“Five emerging trends have the power to transform Wisconsin agriculture in coming years, and increasing computer capabilities plays a role in four: big data, artificial intelligence, autonomous vehicles, and the sharing economy.”* When people think of the Internet of Things technologies, they may not think first of raising livestock or growing crops. However, the smart agriculture market will grow to \$13.5 billion by 2023, up from \$7.53 billion in 2018, according to Markets and Markets, a global research firm. Smart agriculture tools include sensors, microcontrollers, transmitters, energy harvesting technologies, LED lights, drones, and autonomous tractors. The promise is clear: higher crop yields, lower livestock losses, and less pesticide and water usage.

Wisconsin dairy farms like Larson Acres, which has more than 2,900 cows, now generate torrents of data every day about things like what each cow eats, how much milk she produces, the quality of her milk, and the impact of the weather. A vast amount of this data resides at UW-Madison, giving university-funded researchers in the Departments of Dairy Sciences and Computer Sciences the opportunity to use this data to help farmers better manage their operations.

Using data in this manner is not limited to agriculture. The computing revolution is reshaping industries from insurance to tourism, health care to transportation, and many others. For example, insurance companies like American Family Insurance, based in Madison, are partnering with universities to develop technologies to process and analyze their vast stores of data or use drones to inspect homes and accident sites. The more vigorous and world-class we can make UW-Madison's computing and data science efforts, the better able it will be to assist with the seemingly endless stream of needs Wisconsin's industries have in this area in order to remain competitive.

Growing New Major Economic Sectors

Manufacturing, agriculture, healthcare, and financial services have long been major economic sectors for Wisconsin and will continue into the foreseeable future. However, the computing revolution opens the door for growing critical economic sectors. For example, the life sciences industry has long had a substantial economic impact on the greater Madison area. By boosting its computing and data processing abilities, we have a great chance to turbocharge life science and grow it into a primary economic sector for all of Wisconsin.

As another example, Wisconsin always had a relatively small tech sector. In the past decade, however, the sector has grown significantly, with the emergence of more startups, incubators, and accelerators such as gener8tor, and high-tech labs such as Google, Microsoft, Amazon, and RedHat.

Technology-related deals are on the rise in Wisconsin. Angel and venture capital investments in the state's technology companies have increased to nearly match those in life sciences companies, according to "The Wisconsin Portfolio," a Wisconsin Technology Council report. Both sectors had double-digit increases in the number of deals and amount of capital raised in 2015, 2016 and 2017, with no other sector getting achieving this level of growth.

All of the states with universities that have top 10 CS graduate programs (except New Jersey, with Princeton University) also have ranked in the top 10 venture capital-raising states, averaged over the past five years. Companies raised nearly \$72 billion of venture capital in 2017, with more than 60 percent of it going to computing-related technology companies. Wisconsin ranked 25th in the nation in venture capital raised over the past five years, accounting for only 0.25 percent of the U. S. total. There is a substantial opportunity here. Expanding the computing agenda at UW-Madison, combined with strengthening the burgeoning startup ecosystem, collaborating more with leading industries in the state, and working to build Wisconsin's risk capital infrastructure can establish the state, particularly the Madison-Milwaukee corridor, as a leading tech hub for the Upper Midwest.

There is an even more significant opportunity to connect the state with the broader region, beginning, perhaps, with strengthening the Madison-Milwaukee-Chicago triangle. The life science and healthcare sectors are booming in greater Madison. Manufacturing and financial services remain strong in Milwaukee. Foxconn is moving to Southeastern Wisconsin, establishing a major manufacturing base for advanced electronics and possibly attracting other similar manufacturers. Tech concerns are proliferating in the greater Madison area, fueled by national and local R&D labs such as Google, Microsoft, and American Family Insurance, as well as a vibrant startup ecosystem. Chicago has a thriving startup ecosystem; more than 100 Chicagoland startups raised more than \$751 million in the first half of 2018 alone. A strong computing agenda coupled with a farsighted vision and excellent execution can establish the Madison-Milwaukee-Chicago triangle as a thriving national hub of 21st-century technologies in areas like computing, biotech, and advanced electronics. No other region in the Midwest has a similar confluence of advantages.

Jobs Galore

The computing revolution has generated a large number of high-paying jobs in Wisconsin. Our state currently has 8,620 open computing-related jobs, according to Code.org. The average annual computing occupation salary in

Wisconsin is \$75,912, significantly higher than the average salary in the state of \$45,240. Wisconsin jobs in software grew 17% in 2017 over the previous year, far outpacing job growth in other sectors.⁴

Unfortunately, the education system does not currently educate sufficient tech workers to fill these positions. All of Wisconsin's post-secondary schools combined produced just 918 computer science graduates in 2015, a number that barely meets 10% of the demand.

Furthermore, Wisconsin lags our peers in the midwest and the nation as a whole in funding for new ventures. Wisconsin ranks 25th in new tech startups and new tech business establishments behind Illinois (8th), Ohio (14th), Michigan (16th), Minnesota (18th) and Indiana (22nd).⁵ Ironically, Madison is rated the number seven city in the nation for young entrepreneurs.⁶ There is a tremendous opportunity for innovation, job creation, and growth in the tech sector and the two fastest roles in this sector are data scientists and software programmers.

MATURATION OF COMPUTING AND DATA SCIENCE AS MAJOR INTELLECTUAL DISCIPLINES

Without a doubt, computing is transforming the world. However, what about the computing discipline itself? Is it just a new, albeit critical, set of *skills* that other disciplines should have? This question is crucial because historically the college structure at American universities is used to house principal intellectual disciplines, such as engineering, letters, arts, sciences, agriculture, business, and medicine. *How is computing different from other traditional disciplines? If computing and data science are pervasive, what elements belong in and out of computer science, computer engineering, statistics, business, the Information School? What is the appropriate home for the core discipline? What are the strengths and opportunities for UW-Madison? What are the risks? What is lacking?*

Emerging as a Major Intellectual Discipline that Straddles Science and Engineering Fields

Computing, first and foremost, is a science that studies computational phenomena, both natural and human-made. Just as astronomy studies celestial objects and physics studies the natural universe, computing studies natural computational phenomena such as the human brain. Primary focuses here include 1) taking inspiration from nature to develop novel problem-solving techniques; 2) using computers to simulate natural phenomena; and 3) employing natural materials (e.g., molecules) to compute. The main fields of research focusing on these topics are artificial neural networks, evolutionary algorithms, swarm intelligence, artificial immune systems, fractal geometry, artificial life, DNA computing, and quantum computing, among others. *In this aspect, computing bears similarities to certain natural sciences.*

However, computing also studies recent computational phenomena that are made by *humans*, such as the World-Wide Web and social networks. It also studies computing systems that involve both machines and humans. *As such, it bears similarities to certain social sciences.* Newer areas of computing (especially in iSchools) study topics *that bear similarities to the humanities*, such as how societies reacted to the rise of industrial machines in the 19th century and the similarities we can draw between those reactions and how societies are reacting to the rise of computing today.

Data science, which studies the collection, analysis, interpretation, presentation, and organization of data, and which straddles statistics and CS, among other disciplines, has emerged in the last decade as a significant direction in computing. Finally, *computing is also a major engineering field*, as it studies how to build hardware and software artifacts in an efficient, safe, and reliable manner, and how to build data analysis pipelines effectively and accurately.

Computing and Data Science: The New Math

Computing, like math, is now a required necessary skill and area of knowledge, with a similar reach across disciplines as a foundational tool. For example, computing is increasingly critical for research in many social and human sciences,

⁴ *Cyberstates 2017*, Computer Technology Industry Association (CompTIA), 2017, pg 64

⁵ *Cyberstates 2017*, Computer Technology Industry Association (CompTIA), 2017, pg 121

⁶ "How We Rank." *Think-Make-Happen In Wisconsin*, Wisconsin Economic Development Corporation, 2018, <https://inwisconsin.com/rankings/>.

such as archeology, geological sciences, environmental sciences, and journalism. To remain relevant, practitioners must be able to analyze a large amount of data for discovery and inference.

COMPUTING AS A HUMAN ENDEAVOR AND AN UNFOLDING HISTORICAL REALITY

Despite all of this evidence, it is still tempting to view computing as just a skill, albeit an important one, or just an academic subject. However, in our view, this view is short-sighted and frankly, incorrect. Computing, in our view, is better described as a grand human endeavor and an unfolding historical reality.

The truth is that none of us knows today how transformative computing will be to the history of humankind. Anecdotally, consider the content of social media commentary highlighted by historians and sociologists during Mark Zuckerberg's April 2018 congressional testimony:

“Remember that time some guys at Harvard stole pictures of female classmates to rate them as sexual objects, became the richest men on the planet, and then enabled a hostile foreign power to bring American democracy to its knees?”

Future historians can assess the accuracy of those assertions, but the underlying sentiment captures an indelible truth about computing and its implications. Data science, digital information systems, and artificial intelligence promise enormous wealth, enormous potential, enormous peril, and enormous unpredictability. The transistor, the integrated circuit, and packet switching have already taken their place alongside paper, movable type, and the printing press as technological turning points in human history. For better or worse, these inventions have the potential to shape the course of global wars, overturn economies, and invert social hierarchies, sometimes over the course of a millennium, sometimes in a matter of months.

This is what we mean when saying that computing is not just a skill or an academic subject, but a human endeavor and an unfolding historical reality. UW-Madison resources need to be organized and reorganized across the board to adapt to such a challenge. Questions about how computing intersects with economic, racial, and gender inequality will be just as pressing as finding the next great scientific breakthrough or the next billion-dollar startup. The next breakthrough could depend on how well we understand the effects of technology, and universities will be as central to solving the Zuckerberg problems as they were in solving the Gutenberg problems.

PART II: WHAT OUR PEERS HAVE BEEN DOING

To understand what our peers have been doing in computing, we reached out and interviewed the current and former leaders of many top Departments, Schools, and Colleges of Computer Science and Information Sciences in the country. We talked with Carnegie-Mellon (CMU), Cornell, UC-Berkeley, Illinois, Michigan, Georgia Tech, UT-Austin, University of Washington, Northeastern, and UMass-Amherst, among others. We performed an in-depth study of 11 peers in Computer Science regarding capital investments they have made in the past 15 years. We talked extensively with major national organizations in computing, such as the Computing Research Association (CRA), to understand current expansion trends in CS. Finally, we compiled and examined as much public data as we were able to obtain on major computing and data science initiatives around the nation.

Our findings are as follow. Many of our peers have some form of well-funded data science initiative or Data Science Institute. The main reasons our peers are expanding in computing or data science or both include the following:

- rapid program growth, needing new space;
- transformation of major intellectual disciplines;
- the emergence of many high-paying job opportunities for students and faculty;
- challenges to administer evolving structure and coordinate quick actions to match pace; and
- pervasive permeation of computing and data science in nearly all fields.

The changes are happening so fast that flexibility and nimbleness are needed to quickly transform programs and initiatives. Peers recognize the need for strong, coordinated leadership across related academic units involved with data science. Their administrative load has increased to provide needed resources and education to campus, to collaborate across disciplines, and to interact effectively with potential donors and industry.

Additionally, there are some common elements across the universities:

- Computing/data science leadership is present at the chancellor, provost, or dean level.
- There are several undergraduate and graduate degree options across the computing and data science domains.
- The top programs at state universities, notably Michigan, UC-Berkeley, UT-Austin, Cornell, and Washington all offer BA/BS variations of a CS degree. Except for Austin, they are offered separately by the Engineering and L&S/Arts colleges.
- The hiring of faculty is extremely difficult. In computing and data science fields, offers are accepted by 25-50% of the candidates even for the most highly regarded programs. This low success rate is due to significant competition in the private sector, where salaries are high. This acceptance rate is a fraction of equivalents in other disciplines.
- Buildings are at or near capacity, despite a building boom over the previous 15 years, which is articulated later in this report.
- It is typical to have a minor or certificate in computer science, data science, or applied statistics, occasionally accessible to the entire student population.
- Programs with broad multidisciplinary studies often have “follow the student” economic models that encourage innovation and investment in the popular programs.
- Significant donors create compelling events for change in the focus on computing and data science.

There were some differences as well across the programs, usually depending on the home for computer science and the decanal structure of that home.

- Controlling admissions (some do, some don't):
 - Cap on enrollments
 - Diversity initiatives
 - Location of admissions staff (centralized, decentralized).
- Staffing of tenured faculty, teaching faculty
- Budget ownership for faculty hiring
- Ownership of tenure determination.

Structurally, here are how the top 10 programs are organized:

- Two are standalone colleges (CMU, Georgia Tech)
- Two are hybrid decanal entities with unique characteristics (Cornell, UC-Berkeley)
- Five are departments or schools within the College of Engineering (Stanford, MIT, Illinois, Princeton, Washington).
 - Washington has an MOU to become an independent decanal structure
 - Stanford and MIT are discussing the possibility of a structural change
- One is in the College of Natural Sciences (UT-Austin)
 - There are fourteen departments in this college. Discussions are taking place around the possibility of creating a separate computing entity. A report has been produced but not published.

No specific structure belies becoming a top 10 computing program. However, the stress from the rapid expansion of computing and data science has provoked substantial discussion on the pros and cons of becoming an independent entity. Once the key challenges and opportunities to meet the goals of the charge were presented, the working group discussed at length the benefits and disadvantages of different potential structures that would enable UW-Madison to take advantage of the opportunities and overcome the challenges.

Potential Organizational Structures

The organizational structures that have been considered for computing are:

- *A stand-alone college or school of computing/data science with a dean:* The names "college" and "school" are sometimes used interchangeably, depending on the institutional context. However, the key is that such a college or school has multiple independent departments, with a dean who reports to the provost.
- *A Faculty of Computing and Information Sciences (FCIS):* This is a college-like structure in that it contains the faculty of several departments (such as CS, Statistics, iSchool) and it has a dean, but it does not admit nor grant degrees to undergraduate students. Instead, it partners with other colleges (such as College of Engineering, College of Letters and Science) to create undergraduate major programs and educate students of these colleges in these programs. This setup helps the FCIS make education in computing accessible to all students. Cornell pioneered this structure.
- *A Division of Data Sciences:* UC-Berkeley is the first to create this model in 2017. It is not a division within a college. Instead, it is a decanal structure at the college level. Currently, it administers curricula and a major, but does not have any faculty nor does it admit students nor grant degrees. The faculty situation is likely to change shortly, toward a structure not dissimilar from Cornell's.
- *A School of Computing/Data Science:* A school within a college. A director, not a dean, leads this entity. There could be a roadmap with intermediate steps toward the creation of a full-fledged college of computing. This is the case with the University of Washington.

The academic content in these structures varies and could include statistics, data science, computer science, machine learning, information sciences, computer engineering, human-computer interaction, or others. As a result, a college or school of computing may include one or more of the above disciplines, and its name may reflect that, ranging from merely "College of Computing" at Georgia Tech to "Faculty of Computing and Information Sciences" at Cornell to "Division of Data Sciences" at UC-Berkeley.

Independent Computing and Data Science Entities

Creating an independent entity for computing and data science is being widely discussed nationally, not only in the leading programs but also in universities who are looking to reposition themselves and prosper in the Information Age. Most notable in the latter category is the University of Massachusetts-Amherst, who only recently created a College of Information and Computer Sciences. From a position below the top 25 schools five years ago, they have vaulted to number 20.

There is extensive rationale for creating an independent entity beyond rankings improvements; Appendix D - The Trend for Decanal Structures - details them. The principal reason, which runs counter to some concerns expressed at UW-Madison, is that an independent entity is best placed to facilitate the connections across campus that enable the

success of data science and computing for students and researchers alike. This can lead to leveraging skill sets and creating collaborative programs that meet the demands of industry.

The leadership of this entity is vital. Georgia Tech's rise from relative obscurity to a top 10 program was driven by a vision of a "college without walls" and a leader who aspired to "lead, not own, computing." It should be understood that the creation of the college at Georgia Tech was led by the president who enabled the founding deans to be successful by empowering them. Today, Georgia Tech is the center of the tech hub for the southeastern US and is a leader in many sub-disciplines of computing.

As we looked at the options for UW-Madison and spoke to faculty and administrators across campus, several concerns were raised regarding an independent entity. Although each has some validity in its own right, this is a case where not being a first mover has its benefits. The early adopters of a strong computing model, Carnegie Mellon, Georgia Tech and Cornell have mitigated the key risks and concerns, and we stand to benefit from their experience. The primary concerns expressed and how the leading institutions addressed them are detailed in Appendix E - Concerns.

Examples of Leading Programs with Decanal Computing Entities

We examined the experiences of Carnegie Mellon, Georgia Tech, and Cornell, which created colleges and schools of computing at least 20 years ago. We did not examine more recent transitions at other top universities because there has not been enough time to assess their impacts adequately.

Carnegie Mellon's School of Computer Science was created in 1988. Its impact has been widespread and transformative, and CMU continues to be top-ranked in CS (#1, shared with Stanford, MIT, and UC-Berkeley). The school is an innovative leader in computing, and its college-equivalent structure has allowed it to move quickly to seize early leads. For example, powered by advances in machine learning, the Artificial Intelligence revolution has only in the past few years exploded into the collective consciousness. CMU's School of Computer Science recognized the importance of machine learning much earlier, and in 2002 created a Department of Machine Learning and poured resources into growing it. This gamble has allowed the School to gain a significant lead in machine learning.

The impact of CMU's School of Computer Science on the City of Pittsburgh has been transformative. It has helped convert a city with a shrinking manufacturing base into a growing digital and computing powerhouse. Many high-tech companies have set up labs and created good-paying jobs in the city. Google Pittsburgh, for example, employs a staff of more than 275 people. The self-driving vehicle lab that Uber set up in Pittsburgh in 2015 ended up hiring about 50 people from CMU, which has a robust research effort in robotics. Soon afterward, Uber and CMU set up a strategic partnership to do joint research and development in areas such as maps, vehicle safety, and autonomous driving. "This is yet another case where collaboration between the city and its universities is creating opportunities for job growth and community development," Pittsburgh Mayor William Peduto said at the time.

The impact of Georgia Tech's College of Computing has been equally striking. In 1996, its graduate computer science program was ranked the 18th best in the country. However, the College's ability to grow and produce innovative research in a range of computing areas helped it leapfrog (13th ranked) UW-Madison to advance to #8 in the rankings. This has allowed Georgia Tech to attract even better and more faculty and students in computing, cementing its lead. We note that this is the single largest leap in ranking that we have seen in the past 20 years.

The college structure has also allowed Georgia Tech to move quickly and seize an early lead on important topics. For example, five years ago, the high cost of higher education and online education became a national conversation. Because the computing effort had a college structure that allowed it to move quickly, Georgia Tech quickly created an online Master of Science program in Computer Sciences, promising to cap the cost below \$8,000. It was a radical move at a time when the cost of such a program was (and still is) routinely in the \$25,000 to \$40,000 range. The move attracted national attention, including a shoutout from President Barack Obama, who cited the College as a leading exemplar in providing low-cost, yet effective, higher education. Since then this program has blossomed. It will enroll

more than 5,000 students in Fall 2018 and has become the dominant, online computer science MS program in the United States and the world.

The impact of the Faculty of Computing and Information Sciences (FCIS) at Cornell University has been equally profound. There were multiple concerns at the time that FCIS was created, all of which have been either mitigated or failed to materialize, Greg Morrisett, the current FCIS Dean, told us. In the past few years, Cornell has been investing heavily in the FCIS by taking steps like increasing the number of tenured faculty by 45% and leveraging the FCIS to launch highly visible initiatives, such as establishing Cornell Tech, a major new satellite campus in New York City. "Cornell's decision to build Cornell Tech was the biggest thing to happen to Cornell in 100 years, and it would not have been successful if the FCIS had not existed," says Kent Fuchs, former Provost at Cornell and current President of the University of Florida. In the past 20 years, the FCIS also helped Cornell grow its computing faculty and launch many computing initiatives, which in turn have helped Cornell maintain its #6 position in the CS rankings.

FACILITIES TO EXPAND COLLABORATION, INNOVATION, AND TEACHING

As computing efforts proliferate, it is vital to provide ample space to anchor computing and data science efforts across the campus, foster more collaboration, keep pace with the expanding demand for computing education and research, and extend computing into new interdisciplinary approaches that foster university-community collaboration with surrounding tech ecosystems.

As a result, many of our peers launched significant initiatives to construct new buildings for computing efforts. We carried out a study of building initiatives undertaken by top-ranked peers of UW-Madison. Universities we analyzed in depth are the four CS programs tied for first place in the U.S. News and World Report rankings (Carnegie Mellon, MIT, Stanford, and UC Berkeley) and the seven public or land-grant research universities with CS programs ranked higher than or tied with UW-Madison (Illinois, Cornell, University of Washington, Georgia Tech, UT-Austin, Michigan, and UCLA).

All eleven institutions launched CS-related capital projects within the past fifteen years. All but two (Carnegie Mellon and MIT) have constructed a major building within the last five years. The four top-ranked schools have spent an average of \$136 million, and \$16 million within the past five years, on CS capital expansion projects. The other seven schools have spent an average of \$121 million over 15 years and \$98 million in the past five years.

The four top-ranked schools spent more money before the other schools, but the other seven have invested heavily in catching up. Broadly, the top four doubled down on their advantages while the other seven peers reimagined their CS-related enterprises to become elite schools for the future. From a more granular perspective, Stanford and Berkeley’s recent projects might be best characterized as extensions of their core programs; Stanford added a research computing center and Berkeley supplemented its CS department with an interdisciplinary design center. Cornell and Washington undertook wholesale reorganizations of their departments as part of their capital expansions. UT-Austin, Georgia Tech, Illinois, and UCLA used their new buildings to consolidate offices scattered across campus and build new labs to replace those that had outgrown their previous space. Michigan is the outlier among UW-Madison’s peers. It has not expanded its CS facilities since 2005 but has invested heavily in a \$75 million facility to house a new strategic thrust into robotics. The following table provides more detail:

University	Years				Total			
	1st Bldg	2nd Bldg	3rd Bldg	Age of Newest	New Bldgs	Sq. Ft.	Funding	Last 5 Years
Carnegie	2005	2009		9	2	335,636	\$124,600,000	\$0
MIT	2004			14	1	720,000	\$283,500,000	\$0
Stanford	1996	2013		5	2	unknown	\$79,200,000	\$41,200,000
Berkeley	1994	2015		3	2	124,000	\$58,700,000	\$23,700,000

Illinois	2004	2014		4	2	500,000	\$145,000,000	\$95,000,000
Cornell	1990	2014		4	2	315,696	\$60,000,000	\$60,000,000
Washington	2003	2019		-1	2	295,000	\$182,000,000	\$110,000,000
Ga. Tech	2003	2006	2018	0	3	612,000	\$450,000,000	\$355,000,000
UT-Austin	2013			5	1	140,000	\$120,000,000	\$120,000,000
Michigan	2005	2020		-2	2	104,000	\$115,000,000	\$75,000,000
UCLA	2018			0	1	150,000	\$130,000,000	\$130,000,000
Average								
	2003	2014		4	2	192,933	\$126,636,364	\$65,490,000
Top Four								
	2000	2012		8	2	393,212	\$136,500,000	\$16,225,000
Other Seven								
	2006	2017		2	2	107,099	\$121,000,000	\$98,333,333

Note: Averages exclude Georgia Tech's 2018 CODA development, both because its massive price tag (\$355 million) skews the results and because of its unique financing structure. A private developer is financing the project with Georgia Tech's leasing arrangements and revenue-generating retail and event space offsetting the long-term costs to the university.

As the table shows, the top four have the most space on average (393,212 square feet). The seven other peers have 107,099 square feet on average. *For comparison, the CS building at UW-Madison was built in 1987, more than 30 years ago, and has roughly 69,000 square feet in space, or about 65% of the space of its peers.* It is a testament to the quality of the CS program at UW-Madison, and to the world-class potential present that it has retained a relatively high rank without anywhere near as much spending.

Our analysis of the capital investments at our 11 peer institutions suggested the following observations are worthy of more discussion.

Both Washington and Georgia Tech have made massive efforts to integrate their CS departments into their regional technology ecosystems. Washington is pushing hard to keep up with Seattle's booming economy. Georgia Tech is a driving force for Atlanta. Both include above-and-beyond financial investments: Paul Allen's \$50 million CS endowment and legislative reinvestment at Washington; and a \$355-million, 21-story mixed-use development at Atlanta's Technology Square. Matching these efforts would be a supremely heavy lift for UW-Madison and the State of Wisconsin. However, these two efforts present the most precise vision of what the future of top-flight university computer science looks like.

Both Cornell and Illinois perhaps provide the best peer comparisons. These schools closely align with UW-Madison regarding the history of their CS departments and the identity of their universities. Cornell's CS mission to enable all students to apply CS to any field of study is closely aligned historically and philosophically with what our working group calls the "Wisconsin Computing Idea." Illinois is a model for how a Midwestern land-grant college can transition from a mid-twentieth century CS department dependent on federal grants to an entrepreneurially driven, 21st-century enterprise.

It is vital to note that the source of funding varies from university to university, but there are ample sources. The Bill and Melinda Gates Foundation, for example, has provided significant building grants to Cornell, MIT, UT-Austin, and Washington.

The College of Letters & Science has a preliminary estimate of \$125-150MM for a new facility near the current CS building. It is an imperative that this investment is undertaken to achieve the goal of keeping Wisconsin competitive in the Information Age.

FACULTY EXPANSION AND OTHER MAJOR INITIATIVES

In the past decade, as computing has increasingly shaped every aspect of society, many states and universities have made significant investments in their computer science, computer engineering, statistics, biostatistics, and information studies, often focused on faculty expansion and data science initiatives.

Significant Expansion of Computing Faculty

The working group focused solely on CS faculty expansion, not on other disciplines with computing and data sciences. This section provides evidence of our findings. We want to note emphatically, however, that to successfully implement our recommendations, it would be necessary to coordinate with and research appropriate growth metrics for Statistics and the iSchool.

As mentioned previously, the median for the top 10 programs is 75 faculty, while Madison currently has half that. The CS department is approved to increase to 50 staff, although special efforts will be needed to achieve that. It will not be done in a status quo scenario, as the Wisconsin faculty size has been flat for 20 years. For reference, here are some of the initiatives at the leading institutions in the country.

- *Washington* has a \$3 million line item in the bi-annual state budget for hiring in CS faculty, and Gov. Jay Inslee's proposed 2015-2017 budget includes \$40 million for a new CS building. CS was authorized to hire 15 faculty members between 2014 and 2016.
- *Harvard* announced in 2014 that it would increase its CS faculty by 50% over the next few years.
- The President of *Princeton* recently announced ten new faculty slots in CS. Princeton's newly released strategic plan states: "*Fields related to information science—including computer science, statistics, and machine learning—will require special attention. These fields are revolutionizing the organization of human society and transforming scholarly disciplines throughout the University... The University has already begun to expand the size of its computer science faculty ...*".
- *The University of Massachusetts-Amherst* has elevated its CS department into a College, created a Center for Data Science, and in 2015 formed a Cybersecurity Institute. The new College has 40+ slots (e.g., the Center for Data Science has 15+ slots in the next few years).
- *The University of Texas-Austin* contributed \$57 million toward a CS building in 2010 and conducted fundraising for another \$23 million for the building. The target growth for CS faculty is 60.
- *Cornell* plans to increase the number of tenure-track faculty in computing by 45%. It also has a new major new campus in New York City called Cornell Tech.
- *University of Chicago* announced in 2016 that it hired Michael Franklin, the chair of CS at UC-Berkeley, to lead a significant CS department expansion, "*to further grow computer science into one of the top departments in the country*" and "*to greatly increase the scale, scope and impact of computer science research and education across the University*". The CS department has an "*unlimited number of slots*" it wants to fill in the foreseeable future.
- The University of Michigan hired 11 faculty members in 2013-2015 and had seven hiring slots for 2015-2016.

Significant Investment in Data Science Initiatives

Virtually all science disciplines are increasingly data-driven, so data science has emerged as a critical need on university campuses. As a result, many universities have announced substantial data science initiatives. Recent examples include:

- University of Michigan: \$100 million and at least 35 new faculty slots for data science for the next five years, in the context of a newly created Institute for Data Science.
- University of Illinois, Urbana Champaign: \$20 million and at least 12 faculty slots for data science.
- University of Chicago: Unlimited number of slots for data science. They recently created a Data Science Institute and hired a UW-Madison data science expert to run the institute.
- UC-Berkeley: Established the Berkeley Institute of Data Science (BIDS) in 2013 with an initial \$12 million grant from the Moore and Sloan foundations, subsequently enhanced with additional grants from and industry partnerships with Siemens and State Street Bank. This success led to the creation of a data science major and in 2017, the creation of a dean-led Division of Data Sciences, which offers an interdisciplinary curriculum in the data sciences.

- UMass-Amherst: Set up a Center for Data Science and plan to invest \$100 million and add 40 faculty slots.
- University of Rochester: \$120 million for data science from the university and private donations.
- Ohio State University: Set up a Translational Data Analytics Institute with an announced \$500 million in funding.
- University of Washington: Set up an eScience Institute, with \$24 million raised.

We found that universities with data science efforts can be roughly categorized into three main groups:

1. *Pioneers*: Started 5-7 years ago or even longer, with relatively sizeable investments. Examples include Michigan, Washington, NYU, UC-Berkeley, Columbia.
2. *Newcomers*: Investing large sums over the past four years. Examples include Purdue, Ohio State, UMass-Amherst, Rochester, Rice, Illinois, MIT, UChicago.
3. *Long tail*: Have invested a few million and are still exploring. Examples are Northwestern and UW-Madison. Most other universities belong to this group.

Further details on programs at some of the universities can be found in Appendix F - Data Science Initiatives.

Creating New Educational Programs

Many of our peers have created new educational programs, at both the graduate and undergraduate levels, to serve the exploding need for computing education. Data science skills, in particular, are in high demand, so many campuses have created data science programs. For example, an in-depth analysis of the top 20 CS programs and the top 20 Statistics programs (according to the UW News and World Report rankings) reveals the following:

- Seven campuses have created an undergraduate degree in data science (MIT, UC-Berkeley, Michigan, Columbia, UC-San Diego, Purdue, and Penn State). One campus (CMU) will offer an undergraduate degree in Artificial Intelligence. Two campuses have minors in data science, and one campus has an undergraduate certificate in data science.
- 34 campuses offer MS degrees or professional MS degrees — or both — in data science.

At UW-Madison, the CS department has neither an MS degree nor a certificate in data science. Biostatistics & Medical Informatics (BMI) offers a Ph.D. in biomedical data science, and the College of Letters & Science is exploring creating an undergraduate degree in data science.

FUNDRAISING MECHANISMS FOR COMPUTING EXPANSION

Our study of the capital investment in computing at 11 peer institutions reveals that all institutions pursued a multi-pronged approach for raising funds for buildings. They begin with an anchor donor — a regional company, local foundation or alumnus — to jumpstart fundraising with a significant gift intended to attract future donors. The university supplements that large gift with state funding in the case of public universities or endowment funds. A large-scale donor campaign follows, either as part of a university-wide capital campaign or CS-specific alumni drive. In each case, universities tailored capital expansions to the specific strengths of their campus and pulled out all the stops to maximize possibilities.

Another model some institutions have used is to identify a forward-looking interdisciplinary topic in computing, a local company that can provide a large amount of support, and an essential area of its regional economy that will benefit from the integration of more computing. Michigan recently funded construction of a \$75 million building for a robotics institute that will enhance Detroit's auto industry with support from Ford Motor Co.

Funding for data science initiatives often follows the same models. Other funding mechanisms our research surfaced include state budget line items to fund faculty expansion; one-time state funding allocations, such as for the second building for computing at the University of Washington, and differential tuition revenue, such as was used at Purdue. Another innovative approach is to develop a new building, lease space in the building to private-sector companies, and use the revenue to fund computing efforts.

PART III: WHAT SHOULD WE DO?

THE NEED TO ACT NOW

Given that there is exploding demand for computing skills from our students, researchers, private-sector companies and other constituents; that computing is critical to the future of UW-Madison and Wisconsin; and that our peers have moved fast and invested on a grand scale to take advantage of the computing revolution, we believe that UW-Madison, in collaboration with private industry and the state, needs to act immediately and decisively for the following reasons:

- The burgeoning need for computing skills from many constituents is *immediate*, and our computing resources today are insufficient to address this need.
- Our ranking in computing has fallen significantly in the past decade: *CS fell from #9 to #11 in 2007 and to #13 in 2018; Statistics fell from #12 in 2007 to #16 in 2018. No other CS department in the top 20 has fallen so much in the rankings, and UW-Madison's CS department is the only program that has fallen out of the top 10.* This raises three serious problems.
 - It is already tough to recruit computing talent in this competitive climate; falling rankings make it much harder. The more this downward spiral continues, the more leading tech businesses that want to work with the field's top talent will choose to invest outside of Wisconsin.
 - Falling rankings make it very difficult to recruit the best students, and a key component of industry attraction and computing R&D depends on having excellent caliber students.
 - *Falling rankings often trigger an exodus of computing talent fueled by other universities' poaching, causing further declines in the rankings, resulting in a negative reinforcement cycle.* Once a CS program's rank drops below #15, it becomes very challenging, if not impossible, to reverse the decline.
- Finally, and most importantly, we believe that *most of the economic benefits of the Information Age will go to the places at the top.* This is already apparent in the widening gap between thriving regions of the Information Age (Austin, Boston, New York City, Pittsburgh, Seattle, and of course the San Francisco Bay Area, among others) and the rest of the country. While unfortunate, this gap is well documented in some recent studies. The same situation applies to universities. A working group at Cornell University similar to this one observed that *"the computing revolution is likely to reorder the reputation of major American universities, and that those coming out at the top will reap enormous benefits as society reacts to this revolution."* Left unsaid is the obvious conclusion that far less benefit will accrue to those below the top tier. This explains why Cornell has acted so decisively and boldly, with enormous investment in computing in the last two decades. ***The implication of this, and the concern as our working group collected evidence, is that in the race for computing leadership, we cannot afford to land in the second tier.***

OUR STRATEGIC ADVANTAGES IN COMPUTING AND DATA SCIENCE

UW-Madison has many formidable strategic advantages that are under-exploited. If we seize those advantages and take decisive actions now, we can increase the profile of computing at the University, expand teaching to all students who desire it, build connections across the University to enable educators and researchers to leverage the strength of the core computing and data science capabilities.

- *We still have a powerful presence in computing*, which includes CS (ranked #13), Statistics (#16), the Information School (#14), Electrical and Computer Engineering (#13), and Biostatistics and Medical Informatics (BMI), among others.

The CS department, in particular, has been a significant force in computing for decades (ranked as high as #5 in the 1980s and #8 in the 1990s, and robustly in the top 10 until 2007). Created in 1964, it is the second oldest CS department in the nation. It has had a tremendous impact on UW-Madison, Wisconsin, and the world. At

UW-Madison, the CS department educates the most majors; provides extensive computing support for the campus; and has in the last ten years alone brought in hundreds of millions of dollars in grants, patent and licensing revenue, and tuition. CS alumni dominate the Wisconsin high tech industry, creating local companies that employ tens of thousands and generate billions of dollars in annual revenue. In the world at large, the CS department's R&D efforts have helped lay the foundation for database management systems, computer architecture, scalable datacenters, the Internet, e-commerce, data science, the artificial intelligence revolution, and the ubiquitous smartphone (which now runs 50% faster thanks to inventions developed at the CS department in the 1980s).

- *The UW-Madison departments named above have had a long history of deep and extensive collaboration, and are largely aligned in their vision for the Information Age.* Both CS and Statistics, for example, are pushing hard on data science. BMI recently created a new Ph.D. program in biomedical data science, one of just a few in the nation. Electrical and Computer Engineering has created several centers that pursue data science R&D.
- *We have outstanding breadth of excellence across the University.* Many of these disciplines have pushed heavily into data science, and are taking leadership in data-driven research in their communities. This is a substantial advantage compared to many other universities, such as the University of Washington, Carnegie Mellon, and Georgia Tech, which have fewer fields of excellence. In particular, *we have traditionally been strong in biotech and life sciences. Our agricultural programs are among the best in the nation, and we have an outstanding Department of Biostatistics and Medical Informatics (BMI).*
- *We have a strong and growing hi-tech ecosystem in Madison and beyond.* In the past decade, this ecosystem has grown rapidly, attracting many startups as well as major companies such as Google, Microsoft, Foxconn, and others. In the next decade, there is good potential for growing a strong high-tech triangle spanning Madison, Milwaukee, and Chicago. There is a strong potential for extensive and deep collaboration in computing between UW-Madison, UW-Milwaukee, Marquette University, Medical College of Wisconsin, and private-sector companies in this triangle and areas around it.
- *The greater Madison region is justifiably recognized for its livability and natural beauty and is growing rapidly, with more people, more services, more flight connections.* Madison is just a few hours drive to several major cities. Few other places in the Midwest have this combination of positives, providing a major advantage for us to compete for computing leadership in the Midwest and beyond.
- *We are fortunate to have an alumni base that is large, active, and well-known for being very engaged, with a lasting desire to build connections and contribute back to the state.* This alumni base is very strong in both computing expertise and wealth, yet so far we have not fully engaged them for help in computing.

By seizing these advantages, we can grow into a national force in computing and data science. Currently, the coasts, along with Austin, Texas, are leading the charge. *We see no reason that the Madison - Milwaukee - Chicago triangle, anchored by UW-Madison and its peers, cannot become such a leader.*

THE WAY FORWARD

How do we exploit our strategic advantages? The working group extensively discussed this question. The answer, we believe, lies in the way we handle five core challenges: vision, leadership, organizational structure/authority/resources, execution, and buy-in. We will now discuss these challenges, including where we fall short and what we can do going forward.

The Need for a Broad Vision of Computing

First and foremost, we must have a vision for computing. It should be a vision that we all care about and are convinced is the right thing to do. It should be inclusive, tailored to our campus and our state, and benefit both. The vision should have widespread buy-in, not just from the campus, but also from the business community, legislators, donors, alumni,

and the citizens of this state. *Ultimately, it should be a vision of computing that is bigger than any single one of us, that will guide all of our other actions in computing for the next few decades.*

In the past three decades, we did not have such a vision. The CS department articulated several visions, but those were limited to growing the CS department to help it maintain strategic strengths in several areas and grow strength in a few other areas. The Statistics department had similar visions for growing its faculty size, in order to grow its strength in data science.

Computing today and in the foreseeable future, however, is greater than CS, Statistics and the Information School. We need to have excellence at the core of computing, which makes connections to facilitate excellence in computing across the rest of the campus and beyond. We need to address the social impact of computing on our daily lives in areas that include security, ethics of data usage, human-computer interaction, and more.

In addition to the campus, our vision needs to include the business community, government, and the citizens of this state. All of these constituents should collaborate effectively with one another to push computing and to make the whole far more significant than the parts. We need to articulate and advocate a vision for this broader view of computing that is inclusive and collaborative for our campus and our state. We discuss such a vision later in this report.

The Need for Campus-level Computing Leadership

To realize the vision outlined above, we need a more dedicated level of leadership in computing. This leadership should be at the campus level to best facilitate cross-campus collaboration and initiatives. Put simply; there must be someone who wakes up thinking about computing and goes to bed thinking about computing. This person should be passionate about our vision. He or she should work on organizing and pushing computing on campus, on helping computing units collaborate, on reaching out beyond the campus, and more. So far we have not had this type of computing leader on campus. We regard this as the most critical action item, and we firmly believe our campus should recruit such a person as soon as possible.

Our analysis of computing on other campuses suggests that having strong leadership in computing is the key. For example, visionary computing leadership 20 years ago at Cornell and Carnegie Mellon made it possible for them to re-organize and expand computing. In the case of CMU, it fueled the resurgence of Pittsburgh, formerly a struggling steel city, into a tech hub.⁷ In all cases, we found that strong and effective leadership was not the only factor, but was an indispensable first factor, which then attracted more support, more donations, and more growth.

The responsibilities of the computing leader should transcend traditional functional roles on campus. This is partially due to the pervasive nature of computing, but also because of the unparalleled opportunity for collaboration with the entrepreneur ecosystem, large businesses, government, and NGOs. Thinking of this leader in conventional terms limits the possibilities of what can be achieved on the campus and in the State. The computing leader needs to possess vision, but also the ability to evangelize that vision, bring leaders together to collaborate and innovate on programs for undergraduate and graduate students, drive differentiated marketing, recruit leading faculty and raise money. The individual must be a change agent capable of reversing the trend in computing at Madison and creating a computing foundation for the next 50 years.

Organizational Structure, Authority, and Resources

In order to recruit a strong computing leader at the campus level, we must have the right organizational structure, authority, and resources in place. It is unlikely that we can recruit strong leaders (either from outside or inside) unless they see that they will have the "infrastructure" in place to realize the computing vision.

The first critical piece of this "infrastructure" is the organizational structure that includes the formal position of this new leader in computing. A range of possibilities exists. At one end, no new organizational structure would be created. The

⁷ July 22, 2017 *Pittsburgh Gets a Tech Makeover* - The New York Times, <https://www.nytimes.com/2017/07/22/style/pittsburgh-tech-makeover.html>

computing leader could simply *coordinate* the existing, disparate computing units already on campus (such as CS, Statistics, iSchool). This could be accomplished by creating a position such as a Vice Chancellor of Computing, for example. We do not believe this strategy would likely be successful, because it would not provide the new leader with sufficient authority and resources to enable effective changes in the broad range of different computing units on campus.

At the other end of the spectrum, one can imagine consolidating virtually all computing resources on campus into a single stand-alone college, which would be managed by a Dean. We regard this strategy as unrealistic because it is not practical that computing can be "owned" by a sole entity on campus. The successful models at our peer institutions have shown that collaboration is key, not ownership.

As a result, we advocate a solution somewhere in the middle, where there is a *strong core in computing*, which consists of CS, Statistics, and the iSchool, and *many vibrant collaborative connections* between the core and computing activities in other units on the rest of the campus and beyond. To a large extent, the core will address fundamental computing challenges, and the connections will address how computing is applied to solve domain problems. We envision the core to be housed in a college-level structure, with its own Dean, who will work as the campus-level leader of computing. We believe that this structure will give this leader sufficient authority and resources to bring about effective computing changes on campus and beyond. We elaborate on this matter later in the report.

Finally, regarding resources, as our discussion of our peer's actions shows, we need to invest significantly more in computing. To do so, we should develop strategies to raise funding for computing from multiple sources, including the central campus, the core and connections themselves, the state, as well as from donors and the broader alumni base.

Execution

Having a good vision, strong leadership, and the right organizational structure/authority/resources will not amount to much if we cannot execute. To execute effectively, we discuss several strategies.

- It is critical that we recruit good people into critical positions in the "computing hierarchy", encompassing both faculty and non-faculty positions. These should be people who share our vision of computing and are capable of executing it. They should be recruited from inside or outside the University, if necessary. Once recruited, they should be given clear responsibilities, as well as sufficient organizational structures/authority/resources to fulfill their responsibilities.

We need to be more aggressive in hiring faculty clusters in crucial core computing areas. Robotics, data science, human-computer interaction (HCI), privacy and security, and artificial intelligence (AI) are potential areas of investment. Aligned with the UW-Madison strategy, to expand our graduate professional programs, we should engage a full-time Director of Professional Programs, responsible for creating and managing the programs, and aligning them with industry sponsors. The latter role should be cross-campus, working with other colleges and departments to expand all computing-related professional programs.

- We need to add a far richer range of people. Some examples:
 - To realize a vibrant program of professional degrees for working professionals, we will need to hire a full-time Director and supporting staff.
 - To adequately enable the connections - the domain science teams - there is an opportunity to offer data scientist programming assistance for their research and industry collaboration projects.
 - To effectively market UW-Madison's computing and data science prowess, not only in the core but also taking into account initiatives across campus, industry experience communications staff need to be engaged.
- UW-Madison should consider raising the profile of teaching faculty to assist with the computing and data science educational workload. Other universities effectively use these hard-to-find faculty for their expanded learning programs, expecting a 2x teaching load as compared to tenured faculty.

- Hiring research professors will be vital to expanding our computing resources, recruiting more people, and helping to cover shortages in teaching and R&D demands. We note that, unlike most of our peers, UW-Madison does not have a formal "research professor" title, making such recruiting very difficult.

Buy-in

To be successful, we need buy-in from a broad range of stakeholders. These include campus leadership, computing units on campus, computing-related faculty, domain science units on campus, the Wisconsin business community, state legislators, donors, and the alumni base, among others.

We need to be open, regularly informing stakeholders of the computing agenda and actions, getting feedback, and adjusting our actions accordingly. If we make an important decision, we need to explain the rationale behind it. Finally, to get broad buy-in, we need to show that we can get things done, that is, we can execute.

It is important to note that to date, computing has been done in isolated units at UW-Madison. As we go forward with a broad computing agenda that covers the campus and beyond, it is essential that we work together to generate serious buy-in, as it will be critical to the success of our work.

Based on the above discussion, in the rest of this report we elaborate on a computing vision, then present our recommendations regarding leadership and organizational structure, taking into account our strategic advantages in computing.

THE WISCONSIN COMPUTING IDEA: A GUIDING VISION

We propose a vision for computing that builds on the Wisconsin Idea. As we all know, the heart of the Wisconsin Idea is that UW-Madison -- *this* University -- should address the needs of *all* citizens of *this* state. Our identity is shaped by our devotion to Wisconsin and our determination to define that devotion as serving every single person within our borders.

While seemingly simple, this is, in fact, a profound idea that has guided our University for more than a century. Furthermore, since originating at UW-Madison, the idea has evolved over time and deeply resonated with students, alumni, businesses, institutions, and our fellow citizens of this state. *It serves to bind all of us and makes us feel a shared responsibility to do all that we can to better the lives of all Wisconsinites. This responsibility is felt even by many who no longer reside in this state.*

Given that computing is making an increasingly profound impact on the lives of all Wisconsinites, any modern manifestation of the Wisconsin Idea needs to address computing. Hence the Wisconsin Computing Idea. *This idea holds that computing is critical to our University and the state, that UW-Madison will take decisive and bold actions to lead the computing revolution, and that advances in computing on this campus will benefit all corners of our great state.*

Three pillars underlie the above idea:⁸

- *The Core:* We will grow a core of excellence in computing and data science at UW-Madison, with exceptional visibility both inside and outside the state. This core will consist of computing and data-centric units such as Computer Sciences, Statistics, and the iSchool. Its primary goals are: to attract, retain, and grow computing talent and resources; to go deep into research, education, and development in computing; and to coordinate computing efforts throughout the state.
- *The Connections:* As the computing revolution expands, computing needs will proliferate on the UW-Madison campus and throughout the state. The core will build connections, that is, collaborative efforts, with computing consumers to assist with these needs. These computing consumers include research groups, departments, institutions, and colleges on campus -- in all science disciplines, and business, organizations, schools, and

⁸ The working group borrowed "Core, Connections, and Enrichment" from U.C. Berkeley's "Data Science Planning Initiative", 2016. Summary: <https://www.stat.berkeley.edu/~binyu/ps/FAB-Summary2016.pdf>

citizen groups -- in all corners of our state. Such collaborations will include consulting, joint degree programs, professional programs, infrastructure support, workforce training and recruiting, and much more. An example of joint degree programs might be a new Data Science degree in engineering coupled with studies in the Internet of Things (IOT) and machine learning, a Data Science degree in Ag & Life Sciences coupled with Crop Sciences or an Information Science degree with Sociology.

- *The Enrichment:* The Information Age presents fundamental challenges and opportunities in many areas, including social justice, ethics, privacy, inclusion and diversity, human-centered design, social policy, and more. We will develop new competencies in these evolving disciplines of applied computing, and work with the social sciences, business, and other disciplines to create curricula and programs that will positively impact how society deals with the changes of the Information Age.

We propose to use the Wisconsin Computing Idea as our guiding vision for computing. First, this idea reflects the broad reach of computing, as we see it today. Second, it is a vision of computing that is tailored to our context, and to the long and proud Wisconsin traditions of combining true academic innovation with strong public engagement. Third, it is a vision with the potential for widespread buy-in, not just from the campus, but from businesses, policymakers, donors, alumni, and the citizens of this state. Finally, it is a vision that is bigger than any one of us, which will serve to bind those of us who care about both computing and Wisconsin.

REALIZING THE VISION

Drawing from the experience of others and mindful of the need to act quickly in the current fast-changing landscape of computing, we recommend the following plan, divided into near-term and medium-term action items. Our near-term recommendations are:

- *Strong and visionary leadership in computing is critical. UW-Madison should appoint a person to be in charge of the overall computing and data science effort. We view this as the most important and urgent action item. This person should "wake up thinking about computing and go to bed thinking about computing." He or she should have dual reporting lines to the Provost and the Dean of L&S. While this is an unconventional arrangement, it is vital to ensure successful cross campus initiatives. UW-Madison should also actively recruit (from outside, if necessary) and grow a range of computing leaders to run core departments, divisions, institutions, and essential initiatives on computing.*
- *The Departments of Computer Sciences and Statistics and the iSchool, which currently reside in the College of Letters & Science (L&S), should be brought closer together, in a formal structure within L&S. Ideally, this should be a School of Computing and Data Science (SCDS) or otherwise similar structure with a name to be determined, consisting of three distinct departments led by the person above. To foster stronger interactions, we recommend housing the three departments in the same building. The creation of this entity will signal to industry, existing and potential faculty and students that UW-Madison is committed to computing.*

Computer Sciences, Statistics, and the Information Sciences (iSchool) would remain as they are now, as distinct departments with their chairs/director. These are the core disciplines involved in data science and computing and data enrichment programs and form a solid basis for the development of connections within the Wisconsin Computing Idea. They also provide the SCDS a critical mass for fundraising, core marketing, and staffing.

- *Create a new task force comprised of faculty and administration, specifically those involved in computer, data, and information science from across. This group should be charged with identifying detailed plans and action items to create the new entity.*
- *Establish a set of milestones. Examples are target faculty numbers; new majors, minors or certificates; new graduate programs; professional masters programs; new cross college degrees; student levels; teaching ratios; fundraising and more. The University of Washington has done this, enabling them to raise substantial funds for facilities and faculty appointments.*

- *The computing entity should work with the campus and stakeholders to start a range of computing activities in research, education, and outreach.* Internally (to UW-Madison), we recommend introducing courses that make computing highly accessible to all students on campus and focusing on data science, which is increasingly critical to many scientific disciplines. Specifically, we recommend continuing the effort to introduce an undergraduate major in data science and setting up a data science institute. Externally, we recommend working closely with the Wisconsin business community to address the severe shortage of computing skills in the state.
- *This new entity should work closely with the campus and stakeholders to fundraise and coordinate significant investments in computing and data science, including a data science institute.* It is critical to significantly increase the number of faculty and professionals to meet the demand of students and data science practitioners across campus and in industry. It is also vital to establish a new building to optimize collaboration across campus and with industry.
- *Establish a computing and data science “advisory board or steering committee” to exist indefinitely.* A senior level board who would meet regularly to discuss opportunities and concerns about the connections aspect of the Wisconsin Computing Idea, ensuring that the SCDS is responsive and effective and meeting its milestones. The board would advise on and agree to a roadmap for computing and data science at UW-Madison. The roadmap will help mobilize the faculty, recruit talent, engage with industry and government and rally donors and supporters. Opportunities to create new degree programs, minors, interdisciplinary programs and more would be promoted and vetted by the advisory board.

In the medium term, our recommendations are as follows:

- *Upon achievement of the milestones and given that computing is increasingly central to UW-Madison and the state, the working group recommends creating a decanal entity for computing at UW-Madison, which will function as the “core” pillar of the Wisconsin Computing Idea.* It is unclear at this time what structural name can be applied, so for this report, we chose to borrow a term from Cornell. Specifically, we recommend creating a Faculty of Computing and Data Science (FCDS). Among the many possible formal structures for this entity (such as department, division, center, college), we believe that an independent structure, with a dean reporting to the provost is most appropriate. It is the standard way to create far-reaching initiatives that involve both research and teaching. It also provides the most visibility, as well as the most flexibility, for creating new initiatives to keep up with the fast pace of computing.
- *Unlike a college, the FCDS would not admit undergraduates nor grant degrees. It will develop and teach the curriculum and perform research within its core functions of computer science, statistics, data science, and information and library science.* Additionally, it would partner with other colleges at UW-Madison to develop multiple educational programs, such as majors, minors, and concentrations in computing and data sciences, then offer these programs to students enrolled in those colleges. This would help the FCDS make education in computing accessible to all students. For example, it might add a degree in CS in the College of Engineering, leveraging the strengths of that college, while maintaining the current CS degree in L&S. A new Data Science degree might be obtained from L&S or the College of Agriculture and Life Sciences (CALS). These joint degree programs reflect the pervasive nature of computing and data science. Funding for the FCDS could be determined by the number of students taking FCDS classes, with tuition money allocated between the FCDS and the colleges where the students were enrolled. While UW-Madison does not currently have this RCM (Responsibility Center Management) budgeting model, initial conversations with the Vice Chancellor of Finance have indicated that it could be worked out for a new entity.
- *The FCDS would work with the campus, stakeholders, and constituencies to significantly build up the three pillars -- core, connections, and enrichment -- of the Wisconsin Computing Idea.* The FCDS could not cover all computing activities. However, it should focus its efforts in a way that is highly visible (for example, to attract

external funding) and accessible to students or anyone who wants to bring computing to bear on his or her problems. The FCDS should function as a "breeding ground" for new activities in computing, developing innovative programs in research, education, and outreach. In particular, critical programs, such as those in areas like data science and cyber-physical systems, should be centralized into collaborative structures such as centers or institutes, and the FCDS should work with the campus and stakeholders to coordinate investments into and the management of such structures.

The Dean of the FCDS would serve as the campus-level leader for computing that we discussed earlier. He or she would work closely with campus leadership, the advisory board, leaders of other computing units on campus, the business community, donors, alumni, and others to realize the Wisconsin Computing Idea.

Concerns and Recommended Actions

The FCDS model recommended above will undoubtedly raise concerns, some of which we briefly discussed earlier in this report.

In the working group, we extensively debated these concerns and others. We noted that the above recommended FCDS model is very similar to the model of the Faculty of Computing and Information Science (CIS) recommended for Cornell University by a similar working group on computing at Cornell in 1999. Thus, we extensively discussed the above concerns with multiple past and current leaders of Cornell to understand how those concerns have been addressed since 1999 when Cornell started implementing this model. We found that the concerns at the time the Cornell CIS was created have been either mitigated or did not materialize. In the past few years, Cornell has been investing even more heavily in the CIS (increasing the number of tenured faculty by 45%, among other moves), and leveraging the CIS to launch highly visible initiatives, such as establishing Cornell Tech, a new major CIS campus in New York City.

The concerns and their mitigation are discussed in more detail in Appendix G - FCDS Concerns.

Data Science Institute

Coinciding with this report, there are several discussions at UW-Madison regarding the potential creation of a Data Science Institute - something that some of our peers have done as noted in Appendix F - Data Science Initiatives. The working group discussed this and came to some conclusions, based on the assumption that a data science institute was created:

- A model which extracted faculty members from their department would be detrimental to growing cross campus capabilities for teaching and research in data science. The group preferred a faculty affiliation model similar to that of the University of Washington's eScience Institute.
- There are pros and cons to an institute being affiliated with any given college, and nationally there is no specific coalescence around a model.
- The primary focus of the institute should be collaborative research. The curriculum should be developed across the campus led by connections made with the SCDS/FCDS where deep data science knowledge exists, and the proposed degree in data science would exist.⁹

Notwithstanding the points above, some additional observations were made, but no specific conclusions or recommendations followed:

- If an academic leader were appointed from outside of the university to lead the new institute, they are likely to desire an appointment in either computer science or statistics.
- There are some good reasons to couple a data science institute with the SCDS/FCDS. It would concentrate the efforts of everyone involved in making a single entity successful. The issues around funding, leadership, recruiting, operations and impact, would benefit more from a single cohesive entity.
- The working group recognized that the current college "tax" on donations to institutes would be an inhibitor to affiliation with the College of L&S, but could be an important funding source for the FCDS enabling further collaboration initiatives.

⁹ Statistics is proposing a new undergraduate degree in data science in collaboration with Computer Science, the Information School and others.

A recent summit of university data science leaders¹⁰ nationwide extensively discussed best practices for creating data science entities at universities, and the working group urges the campus to participate in and benefit from this community. Required reading for the summit was a paper prepared by UC-Berkeley, NYU, and Washington discussing their experience in establishing data science institutes.¹¹

Financial Implications

The Working Group is conscious of the tight financial conditions that UW-Madison is operating within, and seriously considered this constraint in deriving our conclusions. The proposed FCDS option is a “light” option relative to several other options, including creating a full college. Since admissions and degree granting are left in other colleges (L&S, Engineering, Agriculture and Life Sciences, Medicine, Business), staffing costs for a college level unit are optimized.

Notwithstanding, a significant investment needs to be made to increase faculty levels; establish institutes for collaboration and innovation; construct a new facility to house faculty, develop curricula and teach larger volumes of students and offer collaborative workspace; establish and administer professional masters programs both online and on campus, and hire world class leadership. A commitment of this magnitude demands a concerted effort by all stakeholders including the university, the state, industry, faculty, and alumni.

The College of L&S has done a preliminary budget estimate of \$125-150MM for this facility. While a large sum of money, our peer institutions have been successful at raising monies for new buildings in computing through partnerships of private money, foundations like the Bill and Melinda Gates Foundation, corporate gifts and state matching funds. A facility of this nature could house computing faculty from across campus, in and outside of the FCDS as well as institutes and centers such as the data science institute currently being discussed.

Conclusion

Soon, if not already, competency with information technology will be a skill as necessary and pervasive as reading, writing, and arithmetic. Communities who master computing and data manipulation will flourish; those who do not will fall behind. A flagship university that fails to provide adequate technology and data competency will not just slide in the rankings. It will fail its state as inevitably as a university that produces illiterate graduates. We must, therefore, address computing and data science with the same full spectrum approach that we use for reading and writing. UW-Madison has specialized programs in Composition and Rhetoric, Literary Studies, and Mass Communication, but every campus discipline considers a facility with the written word to be an essential skill, we have libraries in every corner of campus, and basic literacy is both a valuable tool and a fundamental goal of extending university resources outward across the state. That same mixture of concentrated expertise and ubiquitous access should be the overarching, long-term goal for computing and data science at UW-Madison.

In achieving that goal, our challenges are not altogether different from the late Nineteenth Century when UW faculty struggled to communicate their science to Wisconsin farmers who had never been taught to read. Today’s economy relies as much on data as dairy farms relied on butter, and yet the average working Wisconsinite knows far less about what to do with the data available to them than those farmers knew about the biochemistry of their cows. Several generations of UW leaders have understood these types of challenges to be not just pedagogical but political— not just about teaching the people, but about the harnessing the will of the people. Wisconsin citizens will only support an expansion, or re-envisioning, of UW-Madison research when they support the underlying mission of a state university and understand its public purpose. In that sense, a computing and data science initiative must be conceived regarding

¹⁰ “Data Science Leadership Summit Summary Report”, Jeannette M. Wing, Vandana P. Janeja, Tyler Kloefkorn, and Lucy C. Erickson, 7 September 2018

¹¹ “Creating Institutional Change in Data Science,” Moore-Sloan Data Science Environments: New York University, UC Berkeley, and the University of Washington, 2017.

“inputs”—political support, public policy, and tax dollars of the state—as well as “outputs”—the expertise, research, and public contributions of the university. Any decisions of the working group should also be understood as an inflection point within a half-century of UW-Madison computer and data science and a century and a half of University history. The outcome should enhance UW-Madison’s existing strengths, thicken inter-disciplinary and cross-campus collaborations, and help revitalize engagement with communities and the state.

The working group learned from the initiatives of UW-Madison’s peers. We were inspired by the leadership demonstrated by our computer, data, and information science faculty on a world stage. We believe Wisconsin’s breadth of excellence across our campus combined with a greater investment in and focus on, computing and data science will spur innovation, create jobs and prepare our citizens for the Information Age.

We realize that the recommended changes are not easy to undertake in a university. However, we believe that the pervasiveness of computing and the speed of change in this area make it critical to explore new, more flexible and interdisciplinary, structures and to take actions now. *Like all distinguished universities, UW-Madison must continually renew itself to adapt to changing realities.* Thus, we call on our colleagues across this great campus, as well as our friends, alumni, and supporters in the rest of the state and across the world, to join forces and "will into being" these crucial changes, to let our University and state prosper in the new age of computing, and to leave a lasting and meaningful legacy for future generations of UW-Madison students and Wisconsinites.

ACKNOWLEDGMENT

The working group would like to thank the many UW-Madison faculty, alumni and administrators who provided input for the report. Deans, chairs, heads, and individuals formerly in these roles at the leading computer science, computer engineering, information science and statistics programs nationally contributed heavily with the more than 25 discussions that we conducted. Particular thanks to Kevin Walters of WARF who did research on facility investments nationally and contributed to the overall tone of the report. We would also especially like to thank the current chairs/director of the key computing and data science departments at UW-Madison, Guri Sohi of Computer Science, Jun Zhu of Statistics and Kristin Eschenfelder of the Information School. Yazhen Wang, former chair of Statistics, and Brian Yandell, co-director of the Data Science Hub at U-Madison, aided greatly as well. Finally, thank you to Bruce Harville for all of the contributions and administrative assistance provided.

Working Group Members

Appointed by the Chancellor

Mike Lehman (co-chair)	BBA Accounting, '74
Tom Erickson (co-chair)	BS ECE '80

Appointed by the Computer Science Chair

AnHai Doan	CS Faculty
Remzi Arpaci-Dusseau	CS Faculty
Tom Still	President, Wisconsin Technology Council
Kathleen Gallagher	Executive Director, The Milwaukee Institute
Jon Hopkins	BS CS '80, MS CS '81, Adjunct Lecturer UW-Milwaukee, Adjunct Asst Professor MSOE

Appointed by the Computer Science Board of Visitors

Brian Pinkerton	BS CS '86 UW-Madison, MS CS '89, Ph.D. '99 University of Washington
Dale Smith	BS CS, History '82

Appointed by the Provost

Erik Iverson	Managing Director, Wisconsin Alumni Research Foundation (WARF)
Jake Blanchard	Executive Associate Dean, College of Engineering
Tim Norris	Budget Director

APPENDIX A - THE WORKING GROUP CHARGE



DATE: November 27, 2017
TO: Karl Scholz
Eric Wilcots
Guri Sohi
Steve Scott
FROM: Chancellor Blank
SUBJECT: Charge Letter

Computing and computer science are increasingly central to Wisconsin's and the nation's economy. Rapid technological innovations provide the ability to improve processes and create new jobs as these are applied to industries that are central to our state's economy, such as agriculture, manufacturing, health care, financial and other services.

The University of Wisconsin has long been viewed as a leader in the field of computer science. Our university can expand its historical role and help develop a broad agenda that will enhance Wisconsin's ability to become a leader in producing students and research that can aid a knowledge-based, technology driven economy and society.

There are a number of ideas that have recently been discussed and circulated about the future of Computer Sciences on campus. In order to accelerate these discussions and assist with the assessment of various ideas, I have asked two of our alumni, Tom Erickson and Michael Lehman, to establish a diverse working group of campus, alumni and industry stakeholders.

Through these discussions, Tom and Mike will advise UW leadership, the CS department, its Board of Visitors, and other stakeholders on the opportunities available to advance a computing agenda, led by UW-Madison that will:

Increase the profile, rankings and research output of our computer science faculty;

- Increase opportunities for students across campus to study computer science;
- Produce more students among our graduates who have been exposed to and trained on computational thinking, big data, artificial intelligence and related fields.

Initially, Tom and Mike will work with stakeholders (such as the Computer Science Board of Visitors, faculty and key business leaders) to identify members of this working group. From there, I have asked Tom and Mike to run a process that will identify, assess and prioritize investments that our campus can make to improve our leadership position. Among other things, this group will examine how other leading universities have

organized their resources to more effectively pursue similar initiatives; including for example, the creation of a school or college.

This is an opportunity for our campus and stakeholder community to work together to expand the reputation and presence of Computer Sciences on campus. I look forward to their report.

cc: Sarah Mangelsdorf
Mike Lehman
Tom Erickson

FROM THE CHARGE TO OUR WORK: SETTING THE SCOPE

In our conversations with our peers, we learned how they were addressing computing - in the broad sense - which then informed the scope of this report. Our thinking, which resulted in the inclusion of data and information science into the report, follows.

Emergence of Data Science

Computing advancements in storage, compute power and database designs led to the rise of the big data ecosystem. Data science has emerged as a major field that uses scientific methods, processes, algorithms, and systems to extract knowledge and insights from data. Turing Award winner Jim Gray claimed data science is the fourth paradigm of science (or more broadly, research), along with experimentation, theory, and computation¹².

As the other luminaries have noted, “Data science is inherently multidisciplinary in two ways: depth and breadth.

- Depth. The technical foundations of data science draw on computer science and statistics but are also informed by other areas of study, such as biostatistics, digital signal processing, mathematics, and optimization.
- Breadth. Data science is used in context, e.g., to explore a data set, to create models, and to test hypotheses—in a given domain. Because all domains generate or collect data, all domains have the potential to benefit from the analytical techniques in data science. Thus, one can say data science methods can be applied to all fields, professions, and sectors.”¹³

Concerning our charge to “produce more students among our graduates who have been exposed to and trained on computational thinking, big data, artificial intelligence, and related fields.”, the working group discussed opportunities for UW-Madison to leverage our existing leadership in data-driven discovery and excellence across campus in centers, institutes, and groups, in addition to the caliber of individual investigators, and become a global leader in data science.

To validate the notion of “core” data science being part of our discussions, we researched definitions for the role of data scientist and the skills that employers look for in data scientists. The research showed that data science jobs tend to require a mix of skills from computer science and statistics. See Appendix H - Jobs in Data Science for more details.

Intersection with the Social Sciences

As the Information Age has unfolded, the overlap between computing and data with the social sciences has created an entirely new genre of challenges, risks, and opportunities. Cyber attacks and fake news are examples of new issues on the horizon; the use of big data to deal with social injustice, inclusion, and diversity are examples of new opportunities.

This area of “applied computing and data science” is still early in its academic life, yet has an enormous potential impact on the use of technology in our society. As such, there is a strong relationship with computing, and the committee included consideration of these areas of study in our discussions and conclusions. Academically, the working group noted that the “core” studies of applied computing and their intersection with the social sciences have generally become facets of information schools in leading universities. Connections to the social sciences, medicine, business, and others would facilitate the breadth of study across the UW-Madison campus.

¹² “The Fourth Paradigm: Data-Intensive Scientific Discovery”. Edited by T. Hey, S. Tansley, K. Tolle. Microsoft.

¹³ “Data Science Leadership Summit Summary Report”, Jeannette M. Wing, Vandana P. Janeja, Tyler Kloefkorn, and Lucy C. Erickson, 7 September 2018

APPENDIX B - RANKINGS

Institution	2007	2009	2011	2015	2018	Academic Structure	CS Faculty
Carnegie Mellon	1	4	1	1	1	College-level "School of Computer Science" with several departments	148
MIT	1	1	1	1	1	Department within College of Engineering	134
Stanford	1	1	1	1	1	Department within College of Engineering	54
UC-Berkeley	1	1	1	1	1	Department within College of Engineering	109
U. Illinois	5	5	5	5	5	Department within College of Engineering	80
Cornell	5	6	5	6	6	Departmental alliance, nonstandard structure. Founded in 2002.	65
U. Washington	7	6	7	6	6	School of Computer Science and Engineering in College of Engineering, founded in 2017.	64
Georgia Tech	11	9	10	9	8	College of Computing founded in 1990.	91
Princeton	8	6	8	8	8	Department within School of Engineering and Applied Science	52
UT-Austin	9	9	8	9	10	Department in College of Natural Sciences	53
Cal Tech	11	11	11	11	11	Department in Division (College) of Engineering and Applied Science	24
U. Michigan	15	13	13	13	11	Department within College of Engineering	56
UW-Madison	9	11	11	11	13	Department in College of Letters & Science	38
Columbia	18	16	17	15	13	Department within School of Engineering and Applied Science	44
UCLA	15	13	14	13	13	Department in School of Engineering and Applied Science	42
Harvard	15	16	17	18	16	Sciences	32
UC-San Diego	13	16	14	15	16	Department in School of Engineering	50
U. Maryland	13	13	14	15	16	Department in College of Computer, Mathematical, and Natural Sciences	48
U. Pennsylvania	18	20	17	19	19	Department within School of Engineering and Applied Science	32
Purdue	18	19	20	20	20	Department within the College of Science	50
Rice	22	20	20	20	20	Department within the School of Engineering	22
U. Massachusetts - Amherst	25	20	20	25	20	College of Information and Computer Sciences, began in 2015. Prior to that, went from Department to School of Computer Science in 2012.	42
USC	25	20	20	20	20	Department within the School of Engineering	98
Yale	18	20	20	20	20	Department within the School of Engineering and Applied Science	28
Brown	22	20	20	20	25	Department, Brown only has three school/colleges (Undergraduate, Graduate, Medical)	26
Duke	25	20	27	25	25	Department within the College of Arts & Sciences	24
UNC-Chapel Hill	22	20	20	25	25	Department within the College of Arts and Sciences	35

Rankings are based on reputation in a survey of deans, chairs, and heads of computer science programs across the nation. US News compiles the rankings. The 2018 rankings are found here:

<https://www.usnews.com/best-graduate-schools/top-science-schools/computer-science-rankings>

APPENDIX C - PEER OBSERVATIONS

We categorized our further observations according to the three goals of our charter.

Goal One: Increase the profile, rankings and research output of our computer science faculty

- The top 10 universities¹⁴ have a median of 75 faculty members, compared to Wisconsin at 37
 - Quality of faculty is also crucial, but it is clear that more faculty means more research, more published output, higher rankings and therefore a higher profile
- Breadth in key subjects is essential. UW-Madison has a rich history in systems but does not have equivalent strength in artificial intelligence, human-computer interaction, theory, and interdisciplinary programs
- Other universities have extensive trans-discipline programs in data science. Many with joint programs in robotics, human-computer interaction.
- Public relations plays a important role in perceptions, and the US News rankings that are the primary basis of comparison in the USA are based purely on reputation/perception
- Wisconsin's efforts on public relations and communications (website, brochures, etc.) for potential students and faculty is not to the standard of the leading universities

Goal Two: Increase opportunities for students across campus to study computer science

- Intro to Programming, Data Science courses are popular
 - UC-Berkeley's intro to data science course is the fastest growing course on that campus
- Most offer minors, some to the entire university. Wisconsin has a certificate
- Many offer intro to computing outside of CS listed classes (Stats, CALS, Engineering, business, etc.)
- Some use tenured; others use teaching faculty
- Cornell
 - Offers CS minor to all colleges, each department/college determines requirements
- Washington
 - 5000 students in two classes this year
 - Using teaching faculty
 - Large classrooms
- Michigan
 - Enlarging class sizes
 - Using tenured faculty
 - Finding space is an issue

Goal Three: Produce more students among our graduates who have been exposed to and trained in computational thinking, big data, artificial intelligence, and related fields

- UW-Madison Observations
 - There is an inherent advantage in the excellence of multidisciplinary studies. With the possible exception of Cornell, none of the top universities offer the breadth that Wisconsin does.
 - Collaboration among departments and colleges with CS is minimal in comparison to other universities
 - Joint appointments
 - Cross-listed courses
 - Wisconsin offers a "certificate in CS," but not a minor. Most progressive programs have minors "everywhere." Cornell varies requirements by department/college.
 - There are some courses outside of CS that extend the "reach" of CS
 - Degrees in Data Science (MS), Information Science (MS) exist, these need to be leveraged better trans campus

¹⁴ The definition of the top 10 programs is based on rankings from US News and World Report, *Computer Science*, 2018, <https://www.usnews.com/best-graduate-schools/top-science-schools/computer-science-rankings>

- Marketing of computing and data science is decentralized, inconsistent and siloed. The external perception of Wisconsin's programs is less impressive than reality
- UC-Berkeley
 - Big Data Science push. Classes, institute, the creation of new college level "Data Sciences Division" likely to incorporate CS, Stats and iSchool
- Cornell
 - Cross-college degrees in CS, Stats and Information Science
 - Offer CS minor to all colleges
- UMass Amherst
 - BA and BS, the former with more flexibility for non-computer science classes
 - Informatics (Computational Science) degree in development
- Stanford and Illinois are innovating in
 - Computer Science + "X" degrees. "X" could be Music, Crop Science, Astronomy, or others

APPENDIX D - THE TREND FOR DECANAL STRUCTURES

CREATING COLLEGES OF COMPUTING OR EQUIVALENT STRUCTURES

The Current Situation

In the past 30 years, there has been a trend toward centralizing computing on campuses by creating a college of computing or a similar structure. This trend started slowly but has gathered significant momentum in the past ten years, as computing permeated the university landscape. The Computing Research Association (CRA), a major national organization for computing, holds an annual, invitation-only conference at Snowbird for the leadership of the North American computing research community (typically CS chairs and deans, among others). The college of computing trend was a significant discussion topic with panels at the CRA conference in 2016 and 2018. A resulting white paper has been drafted and will be published later in 2018.

*As of May 2018, this white paper listed 23 campuses in the US that have created a college of computing or a school of computing with a roadmap for transitioning into a college shortly. **This trend is particularly striking on the top ten CS campuses. Eight out of the top ten CS campuses either have centralized computing by creating a college of computing or its equivalent or are in the process of trying to do so.** The only two top-ten CS campuses for which we do not know of any such initiatives are Stanford and Princeton. Specifically:*

- Three campuses recognized very early the growing importance of computing and acted accordingly. Carnegie Mellon in 1998 created a School of Computer Science, which consists of multiple departments and has its dean (thus, structurally equivalent to a college). Georgia Institute of Technology (Georgia Tech) created College of Computing in 1990. Cornell organized in 1998 a working group similar to ours to examine the increasingly central role of computing. Cornell's working group recommended creating a Faculty of Computing and Information Sciences (FCIS) that combined three areas -- CS, statistics, and the iSchool. The FCIS can be viewed as a college of computing, except that it neither admits undergraduate students nor grants undergraduate degrees. Cornell acted on the work group's recommendation and created the FCIS in 2000.
- More recently, UC-Berkeley moved to create a Division of Data Sciences with its dean, following a successful creation of a data science curriculum and a data science institute at Berkeley.
- Some recent efforts appear to be in a holding pattern. At MIT, seven Turing award winners (the equivalent of the Nobel prize in the field of computing) wrote to the MIT president urging the creation of a school of computing. *"There comes a time, in the course of scientific evolution, when a discipline is ready to emerge from the womb of its parent disciplines and take its place in the world. For computer science, or more accurately, for the field of computing, this moment is now,"* they said in their letter. At the University of Illinois, there was an effort to create a college of computing and data sciences that stalled after the Provost left. At the University of Texas-Austin, there was a working group similar to ours formed to examine the growing role of computing and a possible reorganization. A report was produced that recommended a major expansion; we do not know the details of this recommended expansion.

Why Become a School or College?

The principal reasons are:

- As computing continues to grow by tremendous leaps and bounds, many department chairs are finding their programs have outgrown or are outgrowing the confines of their current locations in colleges of engineering or science.
- Computing is maturing into a major intellectual discipline that shares similarities with but is entirely different from, the fields of the sciences and engineering.
- Computing is unique in that there is a wide range of high-paying job opportunities for its students and a wide range of opportunities outside academia for its faculty. Few other fields have this rare combination.

- The differences between computing and other disciplines raise many practical and logistic challenges when computing is "shoehorned" into a college of engineering, science, or liberal arts. Such colleges often "cater" to the norm of the majority of their departments. It is often difficult or takes a long time for these colleges' administrations to understand the unique challenges faced by computing. They often struggle to take quick and necessary actions to keep up with the fast pace of change in computing.
- Computing is pervasive on campuses and permeates all fields of study. Many disciplines such as crop sciences may require support for their data science initiatives in order to be successful, while others like computer engineering might be more self-sufficient. Coordinating services and curricula ("connections") for those who need it is difficult and time-consuming. For collaboration in computing across campus, it is far better to have a centralized college of computing. Such a college can coordinate the growth of computing and reduce duplication.
- Computing is moving so fast that any campus entity in charge of computing would need a tremendous amount of flexibility and nimbleness to create new programs and initiatives quickly and shut down those that do not work. A college has the most flexibility to create new research and educational programs.
- As computing becomes pervasive and central to the campus, it is increasingly critical to have strong leadership on computing. This leadership cannot be limited to the chair or the head of the CS department, because computing is broader than just CS (for example, it also involves Statistics, iSchool, and other computing-centric entities on campus). Current organizational structures do not provide a natural position for such leadership, because the CS chair often reports to a dean who is responsible for many other departments. On the other hand, a college of computing has a dean who can naturally provide strong leadership in computing for the entire campus.
- Schools and colleges of computing provide excellent opportunities for broad, multidisciplinary coverage and increased research funding. For example, a college of computing is more likely to hire people from a wide range of computing-related fields than a CS department.
- Colleges of computing provide excellent platforms for external fundraising. As computing becomes increasingly critical, funding agencies and donors increasingly want to fund computing-related initiatives. A college of computing has excellent external visibility to attract such funding. Furthermore, in the past several decades computing has generated a tremendous amount of wealth. The dean of a college of computing often has a better and more comprehensive understanding of the styles and preferences of donors in computing fields and thus may have better success with fundraising than the dean of a more general college.

APPENDIX E - CONCERNS

Challenges, Solutions, and Campus Leadership

Our main findings are as follows:

- In all cases of transitioning to a college or school of computing that we have examined, there was opposition, which is expected given the magnitude of the change. *In all cases where this transition successfully happened, we found that two factors were critical: (a) strong support from campus leadership, such as the chancellor, president, provost, and deans, and (b) strong leadership on computing* (for example, leadership from a set of people in CS, or someone hired from outside to run campus-wide computing initiatives).
- A major concern at UW-Madison is whether an independent entity would be an island by itself, shutting out access to computing for the rest of the campus. A related concern is how to maximize collaboration between the entity and the rest of the campus. Cornell addressed these concerns by creating the Faculty of Computing and Information Sciences (FCIS), which *must collaborate* with other colleges to create major/minor programs and to educate students because the FCIS itself does not admit nor grant degrees to undergraduate students.

In contrast to Cornell, Georgia Tech created a full-fledged College of Computing but addresses the above concern in three ways. First, it explicitly emphasizes the need to collaborate as the *raison d'être* for the college in the vision statement and strategic plan. For example, the phrase "a college without walls" was explicitly written into the vision statement. As another example, an early strategic plan states that (a) they will strive to reach their goals by effective programmatic integration with other units, (b) their research will focus on computing areas where CS is a key, but not exclusive, component, and (c) their real specialty will be in knowing how to mix CS and other areas effectively. Secondly, Georgia Tech hired computing leadership that espoused this philosophy.

In our discussions with other universities who have created an independent entity, it was always stated that the right culture fit for the leader was the most important criteria in making their program successful after creation.

- Another major concern is whether the entity would drain away computing talents from the rest of campus, and drain away investment in computing from both campus and external stakeholders. To address this concern, Georgia Tech's president explicitly told the college dean that the role of the college was to *"lead, not own, computing."* Given that Carnegie Mellon, Georgia Tech, and Cornell created such colleges and schools at least 20 years ago, we have had sufficient time to determine if this concern materialized. In all three cases, we found that it did not. Specifically, we found that campus units that had the funding (such as medical schools and various engineering departments) always opted to hire and grow their computing resources and their collaboration with the entity.

Meanwhile, campus units with little funding (including many of the social sciences) significantly benefited from their collaboration with the independent entities. This is because a college of computing by its nature is broader and often hires more types of people than a CS department. For example, a CoC may hire a computer scientist at the intersection of certain social sciences and computing, whereas a CS department may not hire this person.

- An additional concern is whether a new data science and computing entity would create too many disruptions to existing campus units. We found that indeed if not managed carefully, such disruptions can be significant. Existing colleges can lose funding. If students move out of an existing college, they may stop taking required courses in, say, the Department of English, which then experiences a decline in students. To address this problem, Cornell used the model of a faculty of computing and information sciences (FCIS), discussed earlier, where the faculty moved to the new FCIS, but the students stayed with the existing colleges. It appears that this model, together with a transition that is managed *gradually*, can significantly address the above concern.

- Another concern is whether departments like CS and Statistics are ready to be part of managing a college. This concern is significant because managing a college involves far more work and complexity than managing a department, with significantly more administrative overhead (e.g., significant development and communication efforts, industrial outreach, career placement). In all the cases that we have examined, we found three solutions to this problem. First, the transition was managed gradually, giving the departments involved enough lead time to cope with the growing complexity. Second, strong leadership was critical. A leadership team with a clear vision and with strong expertise in management significantly smooths the transition. Finally, the departments involved in such transitions have actively recruited many senior people in order to fill the necessary mid-level and top-level leadership positions. Such recruiting took a long time, so in the successful transition cases, recruiting was started quite early.

As we examined these sorts of transitions and talked to the current and past leaders involved, three things stood out. First, the zero-sum concern was widespread and significant. In order to build support and get stakeholders to buy in, *a case needs to be made for why an independent entity is not a zero-sum game and will help all involved.*

Second, it is clear that some stakeholders incur losses, such as some decrease in funding, in the short term. Short-term losses are inevitable with such a change. *So a case needed to be made for why this is the right thing for the campus and thus is bigger than any one of us. A clear case that describes the long-term benefits for everyone can attract significant support.* For example, in 2009 the Department of Computer Science at the University of Massachusetts at Amherst was housed inside the College of Natural Sciences. The Dean of this college supported moving the CS department into a stand-alone College of Information and Computing Sciences, even though this would strip a significant flow of resources from his college. He gave his support because he recognized the long-term importance of computing across campus.

Finally, *it is critical to have strong support from campus leadership. As part of that support, campus leadership must hold the new college responsible for fostering collaboration across campus, recruit a strong computing leader who has a clear vision and the ability to provide meaningful guidance and recruit critical leaders who share the same vision. The committee proposes a "board" for the new entity to be comprised of administration and faculty across campus, ensuring alignment and achievement of established goals.*

APPENDIX F - DATA SCIENCE INITIATIVES

Below are details on the data science initiatives at the University of Washington, UC-Berkeley, the University of Michigan, and Purdue University. *The goal is to demonstrate the broad range of activities and commitments in data science at a diverse set of institutions.*

University of Washington, CS leaders, discussed data-driven science with the university's president in 2005, obtained support from Microsoft and the Washington Legislature and set up the eScience Institute in 2007-2008. Since then they have been successful in obtaining external funding (e.g., \$2.8 million from NSF, \$37.8 million jointly with NYU and UC-Berkeley from private foundations, \$9.3 million from Washington Research Foundation). We estimate that they have raised at least \$24 million so far. Their kickoff meeting featured 137 posters from 30+ departments on campus. The current faculty core has 18 members spanning 12 departments in 5 schools and colleges. Their goals are to (1) do breakthrough science in both data science methodology and scientific theme areas; and (2) enable breakthrough science by using new tools and methods, changing the discovery process and driving cultural changes. To realize these goals, they are pursuing these key activities: (1) promote interdisciplinary careers for grad students, postdocs, research scientists, faculty, research fellows and other stakeholders; (2) create open, shared R&D spaces where researchers from across campus will come to collaborate; (3) create data science "startups" that focus intensely on pushing some vital domain science project forward; (4) foster interdisciplinary education; and (5) advance state-of-the-art data science in the areas of methodology, domain sciences and best practices.

UC-Berkeley There are two noteworthy data science efforts: An Institute for Data Science founded in 2013 and led by Saul Perlmutter, a Nobel laureate in physics; and a recently-created Division of Data Sciences, lead by Dean David Culler, who was a professor of electrical engineering and computer sciences. A key distinguishing aspect of data science activities at UC-Berkeley is that they have been very successful in data science training at the undergraduate level. For example, their introduction to data science class added more than 550 students in 2015-2016, and well over 1,000 students in 2016-2017.

University of Michigan The data science initiative (MIDAS), announced in 2015, is investing \$100 million in data science research, infrastructure, and services across campus. Its goals are to (1) hire 35 new faculty over the next four years and engage existing faculty across campus; (2) support interdisciplinary data-related research initiatives and foster new methodological approaches to Big Data; (3) provide new educational opportunities for students pursuing careers in data science; (4) expand U-M's research computing capacity; and (5) strengthen data management, storage, analytics, and training resources.

The Data Science Initiative grew out of a 2007 report written by a research cyberinfrastructure (CI) committee and sent to the Vice President for Research and the Vice Provost for Academic information. The report suggested that in order for Michigan to maintain research competitiveness and its leadership in CI, and sustain existing CI efforts, it should form an Office of Research Cyberinfrastructure. The university created such an office in 2008, and in 2013 re-named it Advanced Research Computing. ARC's mission is to enable data-intensive and computational research and support for the university's research community. In 2012, an organizational meeting on data science was held with university faculty. In 2014, a data science presentation and initial proposal were made to university leadership, and a proposal for a Data Science Initiative was submitted to the Provost. In 2015, U-M announced the Data Science Initiative, operating under the auspices of ARC.

Purdue University, a **newcomer** compared with Washington, UC-Berkeley, and Michigan. *The university has initiated a series of bold and far-reaching initiatives.* "Purdue Moves," a university-wide initiative launched in 2013, has the mission of strengthening and expanding the university's preeminent reputation in science, technology, engineering, and mathematics (STEM). A tuition differential is one of the sources that was used to raise funds for a significant expansion of computing. *Purdue launched in 2018 the Integrated Data Science Initiative (IDSI), spearheaded by the Provost and the Executive Vice President for Research and Partnerships.* The Initiative was formulated through two campus-wide

working groups, forums, and campus-wide discussions. *The goal of this initiative is to create a "data science for all" ecosystem, to "make data science education part of every student's learning experience on campus while also boosting research and partnerships to help grow the data-driven economy."* The initiative focuses on five main activities to realize this goal:

- *Education:* The IDSI includes "a transformational component to prepare every Purdue student with a fundamental understanding of data science in their chosen field." The faculty has received a request for proposals to advance the data science education ecosystem.
- *Research:* Purdue is traditionally strong in cybersecurity, digital agriculture, health and life sciences, manufacturing, transportation, and infrastructure. The research component will boost these areas, but also focus on health care, defense, ethics and privacy, and fundamentals of methods and algorithms.
- *Space:* Data science-focused physical sites throughout the campus will be created. A building has been proposed as a home for private-sector companies to enhance their collaboration with Purdue's data science research and education activities.
- *Infrastructure:* The Data Science Initiative will coordinate with other units at Purdue to develop infrastructure support for data science research and teaching across campus.
- *Engagement:* The initiative will work with the Office of Corporate and Global Partnerships to engage private and public partners.

APPENDIX G - FCDS CONCERNS

Specifically, our findings and recommendations, based on the discussions with Cornell and other computing peers, are as follows:

- *For the concern that the core departments (CS, statistics, and iSchool in this case) may not be ready to function together at the college level, we do not recommend creating the FCDS right away. An easier first step may be to bring these areas together in a formal structure within the College of Letters and Science (L&S), such as a School or an L&S Division of Computing and Data Science. Regardless of the organizational structure, these areas should start working together on joint initiatives that are important and attainable as soon as possible; in fact, as of July 2018, these areas have started to work jointly on such an initiative, to create an undergraduate degree program in data science.*
- *A related concern is how to pay for the administrative cost of the FCDS. If we start with a School or Division within L&S, we will have time to address the issue of funding the administrative cost of the eventual FCDS (via fundraising and other mechanisms). This was the model successfully employed by the College of Information and Computer Sciences at the University of Massachusetts-Amherst.*
- *Would the FCDS be an 'island' by itself, shutting out access to computing for the rest of the campus? As we noted earlier, the FCDS model is very similar to the model of the Faculty of Computing and Information Science (FCIS) created in 1999 at Cornell. The FCIS model was designed to prevent such 'island' problems. Specifically, it does not admit undergraduates nor grant undergraduate degrees. Instead, it collaborates with other colleges to develop computing programs and educate the students of those colleges in those programs. This model has worked well at Cornell for the past 20 years.*

Also, the Cornell working group recommended creating an FCIS Advisory Board consisting of faculty and administrators across the campus, which works with FCIS leadership to ensure that the FCIS collaborates well with the colleges at Cornell. This model has worked well, and we recommend creating a similar Advisory Board for the proposed FCDS.

Finally, we note that Georgia Tech's College of Computing addressed this concern by adopting the mantra "a college without walls" as a guiding philosophy, making this philosophy well known across campus, and writing it explicitly into the founding document that establishes the new college's vision and mission. We propose to do the same thing for the FCDS.

- *Would this cause too many disruptions to existing colleges, such as College of Letters and Science (L&S), where CS, statistics, and the iSchool reside? This is a valid concern, and we propose to address that in two ways. First, the proposed FCDS does not move students from their existing colleges. The same model has been used at Cornell to address the above concern, and it has been working well. Second, FCDS may come into existence gradually, as mentioned above. If so, the stakeholders will have more time to work out details necessary for a smooth transition.*
- *Would this drain computing talent and investment in computing away from the rest of the campus' internal and external stakeholders? We extensively discussed this issue among ourselves and with external computing leaders, as the same concerns arose on campuses that have gone through or are going through similar transitions. Our main findings are twofold. First, most campuses adopted the "lead, not own" mantra. That is, **the core structure of computing on the campus should lead and coordinate computing activities, but not own all such activities.** Ownership of all activities is both unrealistic and detrimental for the campus computing landscape in the long run. This mantra was first adopted at Georgia Tech 28 years ago, and it seems to have worked very well.*

Second, the above concerns often arose due to a perception that computing is a "zero-sum game." However, this is not true in practice. Currently and into the foreseeable future, computing is an "expanding pie," and a strong computing core can help the rest of the campus do computing, recruit, and fundraise better. We see this in the cases of Cornell, Carnegie Mellon, and Georgia Tech, which created strong cores in computing decades ago. However, we also see evidence of this in cases of recent transitions, such as at the University of Massachusetts-Amherst and the University of Washington. Two metaphors that often came up were "computing is the new accelerant" and "a rising tide lifts all boats."

- *What happens if computing is cyclical and we happen to be at the peak of the cycle right now?* Extensive data shows that computing is indeed cyclical, but on an *upward trajectory*, according to all relevant data that we were able to collect regarding the number of jobs available, coverage, salary, and other metrics. Furthermore, one can view computing as just a skill or an academic subject. However, in our view, this is most likely incorrect. Computing seems more like a big human endeavor and an unfolding historical reality. The truth is that none of us knows today how transformative computing will be to the history of humankind, but it *will* be transformative. Like all of our peers, we will need to take actions to adjust to this transformation. Inaction is not an option.

APPENDIX H - JOBS IN DATA SCIENCE

We address the nature of employment opportunities in the burgeoning area of Data Science. What is a data scientist? What skills and conceptual knowledge should data scientists possess? Knowing the answer to such questions should provide clarity when developing curriculum and research initiatives in this new area. While we cannot answer this question exactly at this early stage in the development of the discipline, we include here the results of an initial investigation, using opinions from experts in the field as well as our own data analysis.

The best qualitative assessment we found was that of Michael Hochster, who has served as Director of Data Science at various companies. He describes two types of Data Scientist: Type A and Type B¹⁵. In his words:

- **Type A Data Scientist:** The A is for Analysis. This type is primarily concerned with making sense of data or working with it in a fairly static way. The Type A Data Scientist is very similar to a statistician (and may be one) but knows all the practical details of working with data that aren't taught in the statistics curriculum: data cleaning, methods for dealing with very large data sets, visualization, deep knowledge of a particular domain, writing well about data, and so on.
- **Type B Data Scientist:** The B is for Building. Type B Data Scientists share some statistical background with Type A, but they are also very strong coders and may be trained software engineers. The Type B Data Scientist is mainly interested in using data "in production." They build models that interact with users, often serving recommendations (products, people you may know, ads, movies, search results).

This categorization is useful and shows how data science naturally sits at the boundary of two core existing disciplines: Statistics and Computer Science.

We also performed a brief study of jobs postings, which are instructive in both getting a feel for the volume of job opportunities as well as their requisite skill sets. We performed this study through the website [indeed.com](https://www.indeed.com); by searching for jobs with the key phrases "computer science" or "data science", we found what current job postings are available. Our study focused on jobs in California, as this is where most jobs in these disciplines are to be found.

We first summarize total volume. When looking for jobs related to "computer science", we find roughly 34,000 openings; with "data science", we find roughly 4,000-7,000 (depending on the exact terms used to perform the search). From this outcome, we learn that there are a large number of data science openings at many companies; we also see that the number of these openings does not yet compare to what is found in the computer industry. It would be interesting to track the growth of these numbers over time, to better gauge future demand.

Second, we include details from some specific data science job listings. We use these postings to illustrate the nature of skills and knowledge for which employers are searching. These listings (ironically) do not form an unbiased and statistically representative sample; however, in our experience in looking through the data, they seem to capture the expertise desired by industry:

Senior Data Scientist (Intuit)

- Experimental design and Multivariate/A-B testing
- Advanced SQL skills to perform data segmentation and aggregation from scratch
- Experience working with granular web clickstream data and behavior tracking tools like SiteCatalyst!
- Knowledge of programming languages and stats packages (e.g., python, R);
- Comfortable running multiple regression analyses
- Outstanding communication skills with the ability to influence decision makers

Data Scientist (Chatmeter)

- Background in Computer Science, or equivalent technical field
- Formulate a meaningful hypothesis that is relevant to the business objectives

¹⁵ <https://towardsdatascience.com/ode-to-the-type-a-data-scientist-78d11456019>

- Design and train models for use in production environments
- Mine structured and unstructured data for patterns
- Create advanced visualizations
- Build and evaluate predictive models
- Problem definition, data acquisition, data exploration, and visualization
- Experimenting with machine learning algorithms, evaluating and comparing metrics, iteratively improving the solution
- Work with application developers to deploy predictive models
- Visualize and report data findings creatively in a variety of visual formats that appropriately provides insights to the organization
- Provide intellectual leadership and analytical creativity
- Package and present findings to company leadership

New Graduate Data Scientist (Xylem)

- BS Degree (Master's preferred) Statistics, Machine Learning, Operations Research, Applied Math; or equivalent.
- Experience with statistical and mathematical software
- Proven ability to develop system prototypes
- SQL skills
- Data curation
- Develop statistical modeling techniques for pattern recognition problems
- Develop code in R, Java, Python, or other languages for statistical analysis, optimization, and simulation
- Build models that maximize performance and accuracy
- Strategic planning and project management

Senior Data Scientist (Intuit)

- Top-notch problem solver and identifier
- It is also a good idea to know the standard suite of analytical tools SQL, R/Python, some sort of data visualization tool - but more importantly, you must know how to learn tools

Data Scientist (MBR Partners)

- Demonstrated experience in using supervised, unsupervised and reinforcement Machine Learning and Neural Networks algorithms not only in laboratory and experimentation context but commercial and applied applications
- Experience in R statistical modelling language, SCALA and PYTHON and overall familiarity with programming principles and best practices
- Demonstrates the following scientist qualities: clarity, accuracy, precision, relevance, depth, breadth, logic, significance, and fairness
- Cutting edge experience with cloud based and Big Data advanced predictive modeling and statistical analysis tools being integrated with other software systems using heterogeneous querying and API methods
- Experience in applied data analysis, particularly in applied predictive modeling and statistical analysis solving descriptive, predictive and ideally prescriptive analytics use cases
- Ability to apply knowledge of statistical and mathematical principles to achieve and scientifically demonstrate successful use case outcomes

Data Science Engineer (Veritas)

- Doctor's Degree or Master's Degree with four years experience in Computer Science, Statistics, Applied Math or related field
- Design and Develop reusable Machine Learning Models that can predict the reliability of the various Veritas appliances using Machine Learning techniques
- Build Deep Learning Models utilizing Tensor Flow and AWS Cloud Develop reusable data pipelines and deployment of Machine Learning Models at Scale in Cloud
- Research and develop statistical learning models for data analysis
- Implement new statistical or other mathematical methodologies as needed for specific models or analysis
- Optimize joint development efforts through appropriate database use and project design

Senior Data Scientist (Akamai)

- Experience in general programming languages such as Ruby, Python, Perl, Java, or C
- Experience with Hadoop, Map/Reduce, Hive, or equivalent software
- Experience with streaming technologies, such as Kafka
- Experience with NLP algorithms is a bonus.
- Experience with modeling demand curves

Data Scientist (Cisco)

- BS in math, statistics, data science, computer science, engineering, or economics and 8+ years related data science experience.
- 2+ years experience in a data scientist role doing specifically in machine learning, predictive models, artificial intelligence, deep learning, preferably with some of that at a B2B company
- Experience with data warehousing and data Integration tools
- Experience with Teradata, Hadoop, Hive, Spark, SAP HANA
- Advanced or expert level in SQL, R, Python
- Experience with visualization tools, e.g., Tableau or Qlik

Data Scientist (Apple)

- Well-versed SQL languages.
- Experience with big data technologies such as Hadoop and Spark preferred.
- Familiarity with Python or R and data visualization tools such as Tableau for full-stack data analysis, insight synthesis, and presentation.
- Experience using relevant statistical packages to build predictive/ML models.
- Familiarity with A/B test experimentation design.

Senior Data Scientist (Netflix)

- Strong statistical background and analytic intuition, enabling you to apply various modeling and quantitative techniques to support your analysis.
- Strong expertise in mining data and exploratory analysis to pull meaningful, actionable insights to help drive key HR decisions.
- Experience with experimentation or other statistical comparison applications. Good understanding of small sample size analysis.
- Extensive experience writing SQL and working with raw and, at times, sparse data for analysis.
- Strong programming skills in Python or R.
- MS or Ph.D. degree in Mathematics, Statistics, Computer Science, Quantitative Psychology or related field.

As one can see from this sample, data science jobs tend to require a mix of skills from computer science and statistics. From computer science, experience with programming, large-scale data processing systems, data management, and machine learning are commonly listed; from statistics, a range of modeling and predictive techniques, experimental design, and related approaches are expected.

Table of Contents

INTRODUCTION	1
PART I: THE GROWING IMPACT OF COMPUTING AND DATA SCIENCE	2
THE GROWTH OF COMPUTING AND ITS PERVASIVE IMPACT ON THE WORLD	2
IMPACT OF COMPUTING ON UW-MADISON AND THE RISE OF DATA SCIENCE	2
Impact of Computing and Data Science on Scientific Discovery	2
Impact of Computing and Data Science on Research Funding	3
Impact of Computing and Data Science on Education and Training	3
Impact of Computing and Data Science on the Role of UW-Madison as a Powerful Economic Engine	4
Impact of Computing and Data Science on the Intellectual Leadership of UW-Madison	5
IMPACT OF COMPUTING AND DATA SCIENCE ON THE STATE OF WISCONSIN	5
Transforming Existing Major Economic Sectors	5
Growing New Major Economic Sectors	6
Jobs Galore	6
MATURATION OF COMPUTING AND DATA SCIENCE AS MAJOR INTELLECTUAL DISCIPLINES	7
Emerging as a Major Intellectual Discipline that Straddles Science and Engineering Fields	7
Computing and Data Science: The New Math	7
COMPUTING AS A HUMAN ENDEAVOR AND AN UNFOLDING HISTORICAL REALITY	8
Potential Organizational Structures	10
Independent Computing and Data Science Entities	10
Examples of Leading Programs with Decanal Computing Entities	11
FACILITIES TO EXPAND COLLABORATION, INNOVATION, AND TEACHING	12
FACULTY EXPANSION AND OTHER MAJOR INITIATIVES	14
Significant Expansion of Computing Faculty	14
Significant Investment in Data Science Initiatives	14
Creating New Educational Programs	15
FUNDRAISING MECHANISMS FOR COMPUTING EXPANSION	15
PART III: WHAT SHOULD WE DO?	16
THE NEED TO ACT NOW	16
OUR STRATEGIC ADVANTAGES IN COMPUTING AND DATA SCIENCE	16
THE WAY FORWARD	17
The Need for a Broad Vision of Computing	17
The Need for Campus-level Computing Leadership	18
Organizational Structure, Authority, and Resources	18
Execution	19
Buy-in	20
THE WISCONSIN COMPUTING IDEA: A GUIDING VISION	20
Data Science Institute	23
Financial Implications	24
Conclusion	24
ACKNOWLEDGMENT	26

APPENDIX A - THE WORKING GROUP CHARGE	27
FROM THE CHARGE TO OUR WORK: SETTING THE SCOPE	29
Emergence of Data Science	29
Intersection with the Social Sciences	29
APPENDIX B - RANKINGS	30
APPENDIX C - PEER OBSERVATIONS	31
APPENDIX D - THE TREND FOR DECANAL STRUCTURES	33
CREATING COLLEGES OF COMPUTING OR EQUIVALENT STRUCTURES	33
The Current Situation	33
Why Become a School or College?	33
APPENDIX E - CONCERNS	34
Challenges, Solutions, and Campus Leadership	35
APPENDIX F - DATA SCIENCE INITIATIVES	37
APPENDIX G - FCDS CONCERNS	39
APPENDIX H - JOBS IN DATA SCIENCE	41
Table of Contents	44