COLUMBIA UNIVERSITY
INSTITUTE FOR DATA SCIENCES AND ENGINEERING
A PROPOSAL TO THE CITY OF NEW YORK

EXECUTIVE SUMMARY
I. A University in Full

Since its founding in lower Manhattan as New York City’s first college more than 250 years ago, Columbia University has been attracting the best minds to New York in pursuit of the highest quality teaching, research, patient care and public service. Of the 79 Nobel Prizes won by Columbians over the past century, there are nine Nobel laureates on the current faculty. Yet what sets Columbia apart from other world-renowned research universities is its role in shaping the people and ideas that have long driven New York City’s dynamic economic, civic and cultural life. It was a Columbia graduate who first proposed that New York develop a street grid; a Columbia dean who designed the City’s sewer system; and a Columbia alumnus who engineered New York’s first subway line.

Columbia’s basic and applied scientists and engineers have long been world leaders in developing transformative technology, notwithstanding having to contend with the relatively limited space historically available for advanced research facilities in New York City. The inventions spawned by Columbia’s scientists and engineers are woven throughout the fabric of modern life: FM radio, lasers, X-ray photography, and the technology behind DVDs and high definition television. Today, the faculty of The Fu Foundation School of Engineering and Applied Science (SEAS) is adding to this tradition of bold innovation, producing the technology that has allowed for the development of modern smart phones, Disney’s animated movies, data centers, and the next generation of text messaging. The percentage of Columbia’s engineering faculty elected to the National Academies of Engineering and Sciences, respectively, places SEAS among the nation’s top five engineering schools. It is, in many important respects, a faculty poised for growth. This deep reservoir of talent in the applied sciences has, in turn, given rise to an entrepreneurial culture at Columbia led by one of the most experienced and successful technology transfer offices in the world. In the past 17 years, Columbia Technology Ventures (CTV) has generated more than 128 startups and has established one of the highest rates for translating research dollars into new products and businesses of any school in the nation.
This proposal outlines a detailed plan that will leverage the full breadth of Columbia’s academic excellence and its physical presence in New York City to establish a multi-disciplinary Institute for Data Sciences and Engineering, a world class applied sciences initiative that will enhance New York City’s place as one of the world’s foremost centers for technology innovation. The goal of the Institute is to deepen a culture of translational research and scholarship that can address the unprecedented challenges and opportunities posed by a data-rich society. Discoveries in the data sciences will be propelled by the study of five vertical markets central to New York City’s innovation economy, each one occupying the focus of a dedicated faculty group and their students housed in customized facilities: a New Media Center; Smart Cities Center; Health Analytics Center; Cybersecurity Center; and Financial Analytics Center.

The proposed Institute proudly addresses Mayor Bloomberg’s request from the vantage point of a local university, most obviously by building upon Columbia’s major commitment to long-term investment in upper Manhattan. Columbia’s detailed proposal offers several critical comparative advantages for meeting the City’s stated goal of developing “research that will lead to the formation and expansion of companies in the City and the attraction of companies to the City, in industries that demonstrate the most potential for growth”:

• **Predictable and timely construction of new facilities.**
  The capital project proposed by Columbia is more than shovel ready: it would become a new and transformative element of the Manhattanville campus construction where shovels already are in the ground.

Construction and renovation of Columbia’s three proposed applied science and engineering buildings (providing more than one million square feet of space) will occur within an already active construction site for which all government approvals—including successful completion of New York City’s rigorous land use review process—now stand resolved. When completed, Columbia’s Manhattanville campus in West Harlem will be comparable in size to the original core of its Morningside Heights campus. The impetus for this expansion was not only to create new academic and research facilities needed by several Columbia schools long challenged by space constraints, but also to advance teaching and research efforts into the next century. The Institute for Data Sciences and Engineering would perfectly utilize this opportunity, and the expansion of the City’s ongoing partnership with Columbia envisioned by this proposal would dramatically facilitate the initiative’s realization.

• **Rich interdisciplinary collaboration.**
  The interdisciplinary collaboration between the engineering faculty and departments of the Arts and Sciences, the University Medical Center, and Columbia’s

Still from Disney’s Tangled. (Copyright Disney Enterprises, Inc.) The swaying movement of the heroine’s dress was based on fundamental computational modeling research by Computer Scientist Eitan Grinspun (right), which is used by film animators, graphic designers, biomedical engineers, physicists, and others.
renowned professional schools—partnerships crucial to modern scientific breakthroughs and central to this proposal—will occur in myriad ways, both structured and serendipitous, in a fashion made possible only by locating the Institute within the larger Columbia academic community and the fabric of urban life. Though the faculty and students of the School of Engineering and Applied Science will be at the center of the Institute, Columbia’s proposal envisions significant contributions from the Graduate School of Arts and Sciences, the Mailman School of Public Health, the College of Physicians and Surgeons, Columbia Business School, Columbia Journalism School, the School of International and Public Affairs, and the Graduate School of Architecture, Planning and Preservation—each located in close proximity to the Institute.

- **Existing relationships supporting commercialization and job growth.** Producing the engine of economic growth and job creation sought by the City will require a New York-based network of relationships with local entrepreneurs, venture capitalists, angel investors, and workforce development advocates—relationships Columbia possesses today and will further develop for this purpose. Almost 70,000 Columbia alumni live in New York City, many of them leaders in business and finance, media, the arts, health and medicine, public service, and education. And many have helped to launch and finance successful startups. Columbia’s culture of entrepreneurship is diverse and far-reaching, and its potency is reflected by the cumulative record of Columbia Technology Ventures: 4,000 inventions; 1,800 patents; 500 licenses; and the creation of 128 new companies (81 of them still active).

If Columbia’s proposal is selected by the City, these collective assets will ensure that Columbia University’s vision of an Institute for Data Sciences and Engineering in northern Manhattan, capable of accelerating scientific breakthroughs, creating technology-driven startups, and establishing new sectors of the innovation economy in New York, will be achieved at the scale and in the timeframe set forth in the proposal.

II. Five Centers Addressing the Challenges of a Data-Rich Society

Each of the Institute for Data Sciences and Engineering’s five centers will function as both an engine of translational research in the data sciences and a source of technology projects with high commercialization potential. The strength of the Centers and their compatibility with the City’s entrepreneurial culture will come largely from an interdisciplinary orientation which draws upon the learning and insights available at the Manhattanville campus and throughout the larger Columbia University community.

The potential for spurring economic growth and job creation in New York through these activities is significant. Data science is recognized by economists and financial analysts as one of the leading opportunities for developing the innovation economy in coming decades: a 2011 report of the McKinsey Global Institute projects 40 percent annual growth in the worldwide generation of data and estimates that the U.S. health care sector alone could realize $300 billion in annual value through the comprehensive adoption of innovations related to data sciences. In the near term, the U.S.
The New Media Center will gain momentum from Columbia’s participation in NYC Media Lab: a consortium of Columbia, Polytechnic Institute of New York University, and the New York City Economic Development Corporation, launched last year by Mayor Bloomberg for the purpose of connecting companies looking to advance new media technologies with local academic institutions undertaking related research.

SMART CITIES CENTER

The research conducted by the Smart Cities Center will: develop and monitor green infrastructure and buildings; improve the power supply through smart grid technology; detect and counteract problems with aging urban infrastructure; calculate and communicate optimal transportation routes under congested traffic conditions; and deploy ubiquitous sensing devices to facilitate everyday activities in a crowded urban environment. The Center is expected to face a shortage of 140,000-190,000 people with the analytical and managerial talent required to realize the economic opportunities in this field. This shortfall in human capital, and the problem and corresponding opportunity it represents, are the primary focus of Columbia’s proposed Institute.

NEW MEDIA CENTER

The New Media Center’s applied solutions will address the automated customization and targeting of online advertising; creation of new forms of digital media to augment traditional publishing and journalism; acquisition and analysis of data from social media; and extraction of useful information from online multimedia, including text, speech, video, and images. Columbia’s Tow Center for Digital Journalism; Business School; and Graduate School of Architecture, Planning and Preservation will join SEAS to develop this program. The Master of Science in Journalism and Computer Science—a unique dual degree already offered by Columbia—provides a valuable model for the planned interdisciplinary collaborations. Creation of

Accuracy assessments of urban environments are essential to building sustainable cities, but those assessments are no longer feasible using only traditional engineering and applied science tools. Columbia’s Smart Cities Center will address this challenge.
also will advance new tools for analyzing the largest, potentially transformative datasets, such as health data from New York City schools. Accurate assessments of urban environments are essential to the goal of building sustainable cities, but those assessments are no longer feasible using only traditional engineering and applied science tools. The Smart Cities Center addresses this challenge. Columbia’s recognized leadership on urban sustainability and green technology—exemplified by the University’s management of the NYC Urban Technology and Innovation Center (UTIC)—positions Columbia to drive the creation of start-up ventures with the translational research developed in this Center.

HEALTH ANALYTICS CENTER

Experts predict that the number of electronic health records resulting from doctor’s visits in the U.S. soon will exceed one billion annually. Health care is experiencing an explosion of data invaluable in the search for treatments and cures for disease, and in setting public health policy. The Health Analytics Center will build upon the work of teams of Columbia researchers drawn from the fields of medicine, biology, computer science, applied mathematics, and statistics. These researchers are using patient data, genomic databases, and public health records to improve patient care and to achieve greater efficiencies in public and private health care systems. Columbia already has seen traction with startups in the health analytics area and anticipates more in the years ahead.

CYBERSECURITY CENTER

A core focus of Columbia’s Institute for Data Sciences and Engineering will be developing the capacity for keeping data secure and private throughout its lifetime. Given the ever-increasing growth of online commercial and financial transactions and electronic transmission of sensitive health and financial data, this field is widely expected to assume far greater academic and economic significance in the decades ahead. Research on data security and privacy is best conducted in an interdisciplinary environment where security researchers can interact with faculty who use data in many different ways, and with practitioners who use data in real world environments. The Cybersecurity Center will bring together and build upon the research of the departments of Computer Science and Electrical Engineering, and the work of the Columbia Business School.

FINANCIAL ANALYTICS CENTER

The ability to analyze and act upon the astronomical volume of real-time financial data and to transform data into financial solutions requires a diverse set of talents. The Financial Analytics Center will bring together expertise in finance theory, machine learning, statistics, signal processing, operations, and natural language processing, and will support collaborations with an appropriately trained student body as well as with the financial industry. The result will be entrepreneurial ventures with the potential to define finance and financial engineering for the 21st century.

To support and amplify the study of these five vertical markets at the center of New York City’s innovation economy, the Institute also will conduct core research on problems that cut across the data sciences. The research will focus on formal and mathematical models for data processing, as well as on issues concerning the engineering of large-scale data processing systems.

III. New Buildings and New Degrees on a New Interdisciplinary Campus

When fully developed, the Institute for Data Sciences and Engineering will occupy three buildings on Columbia’s Manhattanville campus, providing more than 1.1 million square feet of state-of-the-art laboratories, classrooms, and newly designed facilities encouraging interdisciplinary activity and collaboration with entrepreneurs, investors, New York-based enterprises, and other outside partners.
Current construction underway on the Jerome L. Greene Science Center between 129th and 130th streets, looking west from Broadway. The proposed Institute would be immediately north.

Full construction of this new applied science campus-within-a-campus will occur over 20 years, divided into two equal development phases. By 2018, Columbia will have completed construction of a new 443,000-square-foot interdisciplinary facility for the Institute and will have hired 40 new faculty in the fields of data science and engineering. This faculty cohort will be working with some 600 graduate students. By 2022, Columbia will have hired a total of 72 new faculty and will have expanded the population of graduate students to a total of 1,080, all of them dedicated to one of the various data science initiatives set forth in the proposal.

In Phase II of the Institute’s construction, Columbia will complete the renovation of a 220,000-square-foot building to house an additional 20 faculty and 300 graduate students. During this second development phase, the University also will construct a new building of approximately 520,000 gross square feet, which will be home to 75 additional faculty and their 1,125 graduate students. When completed in 2032, the Institute of Data Sciences and Engineering will support 167 new faculty and a graduate student population of approximately 2,500 in over 1.1 million square feet of new facilities.

The proposed expansion would roughly double the size of a world-leading engineering faculty at Columbia already expanding into new research disciplines and poised for growth.

Consistent with the City’s goals, Columbia is proposing to establish a new master’s of engineering (MEng) and doctorate of engineering (DEng) degrees, dedicated to the integration of industry, startup, and venture capital experience within the engineering curriculum. The doctorate of engineering degree will introduce to New York City and to the United States a model gaining wider acceptance in Europe and Asia for training a non-academic research workforce that can drive innovation and economic development. The doctorate degree will provide the same intellectual challenge as a PhD, but within a framework of industry relevance aimed at preparing research engineers rather than research faculty.

Columbia’s engineering curriculum will be further expanded to incorporate a new research program focused on understanding the educational processes that best facilitate innovation. The program in Engineering, Education, Innovation, and Entrepreneurship will benefit greatly from a close association with Columbia’s Mind Brain Behavior Initiative, to be housed in the 450,000-square-foot Jerome L. Greene Science Center—the first major structure to rise on the Manhattanville campus. The population

Conceptual rendering (massing diagram) for the Institute’s three buildings. Site 16 is the current Nash Building. The Jerome L. Greene Science Center (in blue) is immediately south.

Broadway corridor landscape plan of the Manhattanville site.
of Columbia principal investigators assembled at the Greene Science Center, who will be exploring relationships between gene function, brain wiring and behavior, will be roughly equal in number to those at Rockefeller University. The Mind Brain Behavior Initiative is designed to be a central hub of interdisciplinary collaboration at Columbia and is ideally suited to advancing investigations into the educational processes that lead to innovation.

Also to be included in the initial round of buildings to be constructed on the Manhattanville campus are new facilities for the Columbia Business School. In the words of the principal donor for these facilities, Business School alumnus Henry Kravis, “We’re not just constructing a building—we are creating a community of entrepreneurs. These new facilities will be a place where the entrepreneurial spirit thrives.” When completed, the new home of the Columbia Business School will possess the physical proximity, academic excellence, and institutional mindset to move well beyond its already rich engagement with SEAS and will help to create an unprecedented chapter of entrepreneurship and practical innovation at Columbia.

IV. Entrepreneurial Record of Translating Research into New Companies and Jobs

Columbia University has an exceptional record of transferring technology to commercial applications. Each year, on average, Columbia Technology Ventures (CTV) manages approximately 300 new inventions from Columbia inventors; executes 50 to 70 license agreements; launches 15 new companies; and generates in excess of $135 million in gross licensing revenues. The University aims to broaden these activities in conjunction with the Institute.

CTV collaborates regularly with the New York City Economic Development Corporation, the Partnership for New York City, the New York Academy of Sciences, and related non-profit entities, to promote New York City as a hub for technology-based entrepreneurship. The CTV Fellows Program selects Columbia graduate students through a highly competitive process and employs them on a part-time basis to analyze inventions for commercial potential, develop business plans, and assist in the sourcing of leads for potential patents and startups. The success of the Fellows Program is seen in job opportunities for graduating students, startups launched by former fellows, and the numerous inquiries Columbia receives from peer universities interested in replicating the program.

Indeed, today at Columbia an organic culture of entrepreneurship is building momentum. This culture includes but goes well beyond CTV and its sterling record. It encompasses the Columbia Venture Community, a diverse support network of more than 1,800 Columbia alumni, students and employees; the many Columbia-incubated technology startups; and a growing list of technology ventures funded or created by Columbia alumni and students.
In a recent survey of SEAS faculty, 52 percent indicated their involvement in starting or helping with an early stage company or organization; nearly 25 percent said that they have been company founders; and 15 percent stated that they have raised over $10 million for companies with which they are involved. Similarly, Columbia students have been involved in many successful tech ventures. Slightly more than half of all SEAS graduates have been involved in starting or contributing to an early stage company or organization. For example, Codecademy, a free online tool designed to teach computer science and coding skills through interactive games, was launched in August 2011 by two Columbia undergraduates. The site already has more than 500,000 registered users.

Columbia proposes to build on this solid foundation of entrepreneurial activity by creating the Columbia Technology Entrepreneurship Center (CTEC) to accelerate the commercialization of high-potential technology projects emerging from the Institute for Data Sciences and Engineering. By combining and coordinating the resources of CTV and Columbia’s schools of Engineering, Business, Journalism, and Law, CTEC will be able to increase the number of start-up companies formed around innovation in the data sciences. A priority for CTEC will be to develop incentives for start-up ventures to locate in New York City and remain here.

V. A Long-Term Commitment to Growth in the City and Community

The Columbia University community in upper Manhattan includes some 50,000 students, faculty, staff and affiliates. With more than 17,000 full and part-time employees, Columbia is the seventh-largest non-governmental employer in New York City and a vital economic engine, with annual operating expenditures of approximately $3.3 billion. More than two-thirds of Columbia’s employees live in the five boroughs. Beyond this local payroll, the university spends over $550 million annually for goods and services in New York City alone.

Every year Columbia attracts the most talented students and scholars to New York City from across the nation and the world, and a significant percentage remain to build their lives and careers here. As a result, there are some 132,000 Columbia alumni who call greater New York home. They include a significant percentage of graduates whose entrepreneurial drive and experience have helped launch and finance new business ventures. Hundreds of our faculty members conduct research on local New York City issues. And thousands of our students—including a large majority of undergraduate engineers—volunteer in New York public schools, community-based organizations, and service projects. The School of Engineering and Applied Science has established a unique track record of extending technology skills and economic opportunity to entrepreneurs within our own community, including hosting in Harlem the only New York State-sanctioned Small Business Development Center.

Now, the University has committed itself to the largest plan for long-term growth in New York since it moved from midtown Manhattan to Morningside Heights over a century ago. Its campus plan to revitalize more than 17 acres in the old Manhattanville manufacturing zone in West Harlem will, over the next several decades, represent an investment of some $6 billion and create 6.8 million square feet of new space—generating an average of 1,200 construction-related jobs per year over the life of the project and some 6,000 university jobs. Additionally, thousands of jobs will arise from research and discoveries.

“From the ‘data poor’ era of the past, we are entering a ‘data rich’ one in all branches of science and engineering. This new paradigm requires radically new thinking and approaches in extracting valuable insights from this data tsunami in order to make the right decisions at the right time. Such approaches are needed to solve the grand societal problems of our time: energy, environment, sustainability, healthcare, and security.”

—Sanat Kumar
Chair, Chemical Engineering
To achieve this growth, Columbia acquired land over many years and successfully had it rezoned with the explicit intention to build and renovate educational and research buildings such as the three proposed for the Institute. In fact, the zoning of this new campus property and the agreements among the University, the City, and the State preclude the property from being used for any purpose other than as a mixed use academic campus. No greater commitment to the future of education and research in the City of New York can be undertaken than the commitment of money, time, reputation, and energy already made by Columbia.

Though the economic growth and job creation fueled by the Institute for Data Sciences and Engineering will benefit the entire City, there will be certain distinct and direct benefits for the Harlem community and other neighborhoods of northern Manhattan, resulting from Columbia’s location and longstanding community relationships in this economically disadvantaged area of the City. More than one-third of Columbia’s major construction contracts are awarded to minority, women, and locally owned (MWL) businesses. The University’s program for building the capacity of MWL contractors serves as a model for other institutions and has been adopted by the City’s Department of Small Business Services. There also are efforts currently underway to meet anticipated labor demands through local workforce development projects created in partnership with the Abyssinian Development Corporation, Harlem Congregations for Community Improvement, and West Harlem Group Assistance.

VI. Conclusion

The organic strengths of Columbia’s approach reflect and leverage those of the city we call home. Like New York City itself, great universities function as dynamic engines of new ideas and economic growth because by bringing together the widest possible diversity of creative people in a discrete community, they generate much more than the sum of their parts.

The Request for Proposals issued by New York City recognizes the shared interests, overlapping goals, and common strengths binding together the future success of modern cities and major research universities. The City’s RFP has provoked an outpouring of reflection and planning at the top levels of the academy. No university more keenly embraces Mayor Bloomberg’s view that New York’s past progress has been traceable to the City’s history as the creative and intellectual capital of the United States. Nor does any other university more deeply appreciate that the future of New York demands a rededication to the City’s legacy of innovation, entrepreneurship, scientific research, and technological leadership. Reaching that goal will require marshalling the world’s leading applied scientists and engineers, supporting them in state-of-the-art facilities, and facilitating their engagement with the entrepreneurial engine that produces economic growth and job creation.

Yet more will be needed. As New York’s leaders in City Hall and throughout the private sector well know, the innovation economy and many of the businesses it spawns are distinguished by their global mobility. The Mayor’s vision therefore also requires a deep-rooted allegiance to New York City’s prosperity and an enduring commitment to the City’s success. No institution has made a greater investment in that future than Columbia University in the City of New York.