

# INNOVATION ENGINES AT NORTHWESTERN MEDICINE

## BIOMEDICAL DATA SCIENCE



## INNOVATION ENGINES AT NORTHWESTERN MEDICINE BIOMEDICAL DATA SCIENCE

Classically, scientific progress has been anchored on two pillars—Theory and Experimentation. Recently, the Big Data revolution has hit science as the sheer volume of scientific data increases exponentially. Advances in scientific computing technology, together with Big Data, have created a third pillar—Computation. Data Science brings together these three pillars to accelerate discoveries. Recently, the Harvard Business Review declared that the data scientist is the “sexiest job of the 21st century.” This role brings together deep domain knowledge, a solid foundation in statistical and mathematical methods, advanced computation and visualization technology, and a desire to tackle “wicked problems.”

The primary home for data science at Northwestern Medicine is the Division of Health and Biomedical Informatics within the Department of Preventive Medicine. Informatics is the study of information: how it is collected, how it is structured, and how it is used to solve real world problems. Informatics combines Data Science with an understanding of the interaction between humans and technology, the so-called socio-technical perspective. Recognizing the importance of data science and informatics, in 2010, the Feinberg School of Medicine established the division and recruited Justin Starren, MD, PhD, to guide the creation of a full department of Health and Biomedical Informatics, as well as the creation of doctoral programs in Informatics.

### Understanding the Fundamental Mechanisms of Disease

Bioinformatics and computational biology focus on understanding disease at the molecular level. Many recent breakthroughs in understanding the molecular basis of disease have relied on breakthroughs in informatics and data science. Critical advances in our understanding of the fundamental mechanisms of disease have come through the analysis of large databases of genomic and clinical data. Northwestern is leading the way through our participation in the Electronic Medical Records and Genomics (eMERGE) consortium. Researchers in eMERGE analyze genomic profiles on thousands of research subjects, combined with years of electronic clinical data, to gain

new insights. We are also bringing those insights to clinical care through the creation of genomic decision support that allows doctors to tailor medications to a patient’s unique genetic makeup.

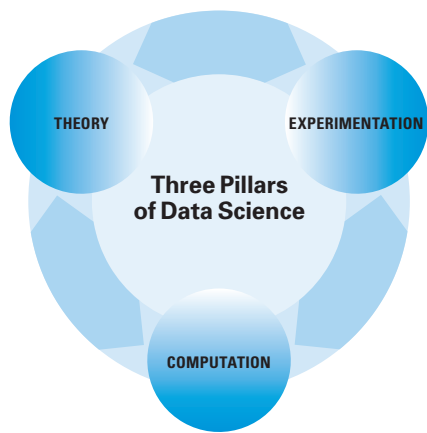
### Advancing Patient Care

The electronic health record is redefining medical care. A systematic collection of individual and population data is stored in a digital format that can be shared for numerous purposes as diverse as personal care, population data warehousing, research, and policy development. This area continues to evolve, and Northwestern Medicine is at the forefront of the use of the electronic health record for patient care and research.



“Our understanding of how networks of genes interact to produce health or disease has been revolutionized in recent years due to advances in bioinformatics. Through projects like eMERGE, Northwestern is a national leader in Biomedical Data Science. Not only are we gaining knowledge about the fundamental mechanisms of disease, we also are applying that knowledge to the care of our patients.”

Justin Starren, MD, PhD, FACMI, chief of the Division of Health and Biomedical Informatics, and Deputy Director, Northwestern University Clinical and Translational Science Institute



## Northwestern Medicine Enterprise Data Warehouse

Northwestern Medicine consists of multiple clinical partners running many different electronic health record systems. To support the highest quality care and accelerate research, clinicians and researchers need to bring together those many different data sources into one coherent picture. The Northwestern Medicine Enterprise Data Warehouse (NMEDW) is a national leader in the design, governance, and implementation of a multi-institutional data warehouse and a model for many other academic medical centers. The NMEDW was the first data warehouse in the country to achieve certification for both Stage 1 and Stage 2 of Meaningful Use.

## Precision Medicine

Diseases that have similar symptoms may have quite different biologic and genetic causes. Precision Medicine is the evolving discipline that uses information about the unique biologic and genetic makeup of each patient to tailor personalized therapies. Northwestern Medicine is part of a national research study to integrate genomic information into a patient's electronic health record and to use that information to customize therapy—improving the effectiveness and reducing side effects.

## Patient-Centered Informatics

Informatics innovations are allowing patients to become active participants in their own healthcare. Mobile devices are being used for communication among doctors, patients, community organizations, and advocacy groups. The Division of Health and Biomedical Informatics and the Center for Behavior Information Technology have developed software applications that allow patients to use wireless tablets



and smartphones to monitor their own health, communicate with their physicians, participate in research, and improve their overall health outcomes.

## Translational Informatics

On average, it is reported that it takes 17 years for a new discovery to become a widely available therapy. Accelerating discoveries from the bench to the bedside is called translational science. A critical part of that acceleration is the use of informatics tools to: manage research data more efficiently; analyze data more rapidly; and replace time-consuming experiments with people (*in vivo*) with rapid computerized (*in silico*) simulations. Northwestern Medicine has developed, and continues to improve, refine, and advance, a translational informatics pipeline that improves both the speed and quality of research.

## Big Data

The amount of data produced is exploding. It is estimated that 2.5 quintillion (1 followed by 18 zeros) bytes of data are created every day. The volume of data is growing so quickly that 90 percent of the world's data has been produced in the past two years. This explosion of data is also occurring in all areas of biomedical research. A single human genome sequence contains roughly six billion base pairs. A single research study may require analyzing the genome sequences of tens of thousands of patients. Processing and managing these data are at the forefront of modern science, including the capture, curation, storage, searching, sharing, transferring, and analysis of these huge data sets. New approaches will help to expand the impact of all of the informatics technologies on health and disease. Northwestern Medicine is in the process of amassing huge data warehouses to serve as robust databases for reporting and data analysis.

### Informatics Education

There is a severe national shortage of biomedical informatics researchers and data scientists. Northwestern Medicine has established a number of training programs to address this shortage. Our master's degree program in Medical Informatics addresses the needs of informatics practitioners working on electronic health records. The Health and Biomedical Informatics tracks of the Driskill Graduate Program and the Health Sciences Integrated Program at Feinberg are training the next generation of informatics researchers.

### Looking to the Future

The future of healthcare and biomedical research will look very different than it does today. Data Science and Informatics will enable us to answer questions in

minutes that, today, take years. Northwestern Medicine has made significant investments and huge strides in this area, but much more needs to be done. Additional faculty members are needed to increase the breadth of our informatics and data science expertise. These new scientists will need space and advanced computational resources. Our degree programs are still in their infancy. Funding for additional students and post-doctoral fellows is needed to address shortages at Northwestern, in Chicago, and nationally. Even with national initiatives, funding for Data Science research remains tiny compared to more mature fields. Funds are needed at both the pilot funding level, and also to establish institutional centers of excellence in areas of Data Science.

## THROUGH NORTHWESTERN MEDICINE, WE ARE CREATING A NATIONAL EPICENTER FOR HEALTHCARE, EDUCATION, RESEARCH, COMMUNITY SERVICE, AND ADVOCACY.

### NORTHWESTERN MEDICINE

Northwestern Memorial HealthCare and Northwestern University Feinberg School of Medicine are seeking to impact the health of humankind through Northwestern Medicine. We aspire to be the destinations of choice for people seeking quality healthcare; for those who provide, support, and advance that care through leading-edge treatments and breakthrough discoveries; and for people who share our passion for educating future physicians and scientists. Our commitment to transform healthcare and to be among the nation's top academic medical centers will be accomplished through innovation and excellence. Biomedical Data Science is a leading Innovation Engine at Northwestern Medicine.

