[Anonymizing Data to Further Genomic and Clinical Research](http://blog.fisherbioservices.com/anonymizing-data-to-further-genomic-and-clinical-research)

Posted by [Richard Tsai](http://blog.fisherbioservices.com/author/richard-tsai-) on May 19, 2015 11:09:00 AM



The future of healthcare is all about [the individual](http://www-03.ibm.com/press/us/en/pressrelease/46580.wss). As personalized medicine is evolving rapidly, new technologies and procedures are being developed to allow doctors to tailor programs for individual patients. Health care agencies and people themselves are collecting and storing massive amounts of health-related information in the cloud. Accessing this information to aid in research and the development of innovative therapies is now possible thanks to a [consortium of industry leaders](https://www.genomeweb.com/scan/gathering-data) and projects.

A vital component of these projects is the requirement to securely store the data and strip out personal identifiers to allow researchers to access virtually all the information with no access to individual patient data records. The release of this previously inaccessible data is changing of the face of healthcare research as we know it.

**Massive Amounts of Data Collected Daily**

Leanne LeBlanc, [IBM Watson project manager](http://www-03.ibm.com/press/us/en/pressrelease/46580.wss) says that, “With the increasing prevalence of personal fitness trackers or “[*FitBits*](https://www.fitbit.com/)” and other connected medical devices . . . that collect real-time information, the average person is likely to generate more than one million gigabytes of health-related data in their lifetime.”

When coupled with the raw data from embedded mechanical devices and FitBits, doctor-created medical records open new avenues for genomics and [clinical research](http://blog.fisherbioservices.com/sharing-clinical-trials-data-is-moving-forward-what-about-biosamples). One of the early challenges is connecting all the data from the myriad of sources and organizing it into meaningful data sets from which researchers can glean the intelligence they need for their decision making processes.

**Organizing Datasets via Platform Development**

The enormous volume of [data](http://blog.fisherbioservices.com/is-social-media-impacting-clinical-research) can be overwhelming, however if presented and assimilated into a universally accessible platform, it can be used to benefit patients and improve healthcare on a global scale.

The Crohn’s & Colitis Foundation of America (CCFA) [recently announced](http://www.prnewswire.com/news-releases/ccfa-launches-development-of-worlds-largest-research-database-for-crohns-disease-and-ulcerative-colitis-first-ever-multi-dimensional-knowledge-platform-to-integrate-scientific-medical-and-patient-data-300082611.html) a new initiative called IBD Plexus, designed “to establish an integrated knowledge platform designed to centralize and aggregate patient information” surrounding inflammatory bowel diseases. This platform will combine biobanking, sample management, and research into one cohesive platform that allows scientists, researchers, clinicians, and patients to share large amounts of data. Due to the accessibility of this shared data, the collaborative platform presents the opportunity for the medical community to work together in developing potential treatments for Crohn’s disease and ulcerative colitis.

Similarly, the [IBM Watson Health Cloud platform](http://www.ibm.com/smarterplanet/us/en/ibmwatson/health/) utilizes cognitive artificial intelligence programming to scan incoming data and organize it in ways that researchers can quickly and easily search for and retrieve meaningful information. Without the cognitive component of the project, it’s all just fragmented data and really has no intrinsic value. The programs, computers and storage devices behind the scenes are scalable and create “evidence-based decisions about health-related issues” that lead to “individualized insights,” the key to the entire process.

**Making the Data Anonymous is Critical to Success**

There has been an explosion in the amount of medical data that is now online and accessible. By "scrubbing" this data to remove all personal information, those with the proper credentials are able to use it for research purposes.

The brains behind the IBM Watson Health Cloud bring together data from many sources and store it in a secure cloud-based environment that researchers can access from anywhere in the world. It is powered by the “most advanced cognitive and analytic technologies” and includes data from [Explorys](https://www.explorys.com/) and [Phytel](http://www3.phytel.com/), two of the largest health care databases in the world.

The cloud is fully HIPPA compliant and grants secure access to anonymous data for new insights. Engineers and researchers at the participating companies have spent the past several years learning the health care industry and have already delivered solutions for pharmaceutical companies and medical centers.

Before the data is released to the cloud, all personal information is stripped out to insure the privacy of all individual patients.

**The Analytics Platform Increases the Value of the Platform**

Without a fool-proof way to gather analytics, the value of a data gathering platform diminishes significantly. Not only does the data need to be collected and scrubbed, it also needs to be configured for analysis so that it can be utilized by the community.

In relation to the IBD Plexus platform, Scott Snapper, MD, PhD, Director of Boston’s Children’s’ Hospital’s IBD Center, [stated](http://www.prnewswire.com/news-releases/ccfa-launches-development-of-worlds-largest-research-database-for-crohns-disease-and-ulcerative-colitis-first-ever-multi-dimensional-knowledge-platform-to-integrate-scientific-medical-and-patient-data-300082611.html), “In isolation, when not linked to each other, these various data types are of limited value in helping to increase our understanding and develop better treatments and cures for IBD. However, brought together in an organized and comprehensive way and analyzed using sophisticated bioinformatics technology, this pool of data should yield extraordinary new opportunities for IBD research as well as for patient disease management and care, and, ultimately, cures for IBD."

The Watson Analytics platform called Watson Health Insights allows participants to develop new insights using the artificial intelligence engines embedded in the project. As a result, project participants are now in a position to actually advise doctors on treatment protocols and provide researchers with new tools to further genetic and clinical research.

Through these analytics platforms, researchers are able to build new solutions in secure environments and further develop potential therapies.

The development of new technologies and optimizing the wealth of healthcare data is a transformative step within our community. As the new world of medicine continues to evolve, genetic and clinical research now has a new ally that will change the lives of everyone.  Access to enormous amounts of anonymous health care information will lead to exponential gains in genetics and clinical research.