John T. Wilbanks

Portable Legal Consent Overview

Summary

Portable Legal Consent (PLC) is a standardized informed consent system for anyone who has obtained data relevant to their health and would like to donate that data for research purposes. PLC works by running volunteers through a short process in which they learn about informed consent, sign an IRB-approved informed consent form, and then upload the data they have chosen for donation. While the system does not transmit “identified” data, donors must indicate that they understand there are some risks of re-identification and harm in volunteering for donation. Most users will encounter PLC via a third party such as a disease foundation, who can run the PLC system on their websites, because the technical infrastructure around the consent system allows for researchers to pick up data from all third parties using simple protocols. PLC is being launched first in the United States, and then in Europe and elsewhere in the world.

Opportunity

The capacity of an individual to gather data about himself or herself has exploded. All of us are now capable of being sensors, or interfaces, or both, to data that could in theory be used to understand and intervene in health: our genotypes, our medical records, information about our lifestyles gleaned from phones, online surveys, and more. At the same time, the ability of an individual research scientist to generate data from individual biological samples has exploded as well, with new assays and machines launching yearly, drawing us ever closer to the $100 genome and related services. We have the opportunity to draw all this big data together, at scale, and begin meaningful analysis of which genetic variations in a population correlate to diseases, to drug response, to long life. The cost of creating, storing, querying, and distributing that data is dropping to a commodity price level, one within the reach of both traditional scientists in research institutions and a growing legion of citizen scientists. This means we can vastly increase the number of people seeking solutions and creating knowledge about our genomes and our health, without massive increases in research investment – simply as a carried benefit of the internet.

Problem

The opportunity faces a hard reality: privacy regulations. Privacy laws, on health data as well as other kinds of individually identifiable information, strictly regulate the kinds of actors who can use data about individuals as well as the kinds of uses they can perform. These laws were built to protect people from harm, a laudable goal. But they are increasingly being left behind by the technical capacity to re-identify individuals in massive data sets, and they are also stifling the ability of researchers to creatively leverage computational infrastructure to begin making correlations between genomes and health. This is a problem that emerges from a

mixture of good intentions and rapid technological development, but it's a problem nonetheless. Scientists must create consent approaches acting as artisans, a new consent system for every study, at a very high cost in time and money. Worse, the outputs of one consent system are almost never interoperable with the outputs of another consent system – meaning that the sort of exploratory computational inquiry we take for granted in the consumer world, where we can Google the entire web again and again, is not possible.

Solution

The solution here is actually quite simple. First, we must act within the existing consent system (at least until the laws change) but there is no reason we cannot create a single, standardized consent form that we give away for free. Second, we must find a body of volunteers who would like to donate their data without restrictions beyond those of "do not harm me" – a set of people who understand the uncertainty of genomic research, and want to donate anyway. Third, we must make it easy for the researcher to find, access, and leverage the data that is donated for research purposes.

The consent form itself is embedded in a web-based process, akin to a software "wizard" – the user encounters the key ideas behind PLC in an easy to understand framework, and can only move forward from step to step after checking key elements on the screen. This is based on the real world experience a citizen would have enrolling in a more traditional study, where a physician or clinician explains the process and consent. PLC is gathering data on the efficacy of the web-based process to ensure that no volunteer donates data without understanding the implications, for good and for bad, of donation.

PLC does not sit alone. It fits into a larger picture emerging of citizens and scientists working together to form communities, around diseases and lifestyles and therapies, that would not have been possible even five years ago. Citizens can get their own data, quickly and cheaply, on the consumer market. Infrastructure like Synapse, a computational platform from Sage Bionetworks where donated data can be accessed and queried by researchers, means that donated data has a place to go where citizens know it will be used. And a growing media awareness of "That's My Data," a campaign urging citizens to ask for a copy of their own data whenever they provide a sample, means that the movement into which PLC fits won't stay long in the shadows.

Our goal is audacious. We have selected a one-year recruitment goal of 25,000 people. This number is not accidental – it is the approximate number of people enrolled in the Framingham Heart Study, one of the most important clinical studies of all time. Our goal is to create a cohort just as big as Framingham, but that is also open for broad, exploratory use, not just for one disease or disorder. We believe that the creation of a data donor pool of this size and scope has the potential to radically reshape our understanding of our genotypes, our health, even our communities.