THE LAST WORD

PRIVACY IN THE ERA OF GENOMICS

By Frank Stajano

A paranoid's look at advances in DNA research.

Computer and network specialists may not have noticed yet, but the transition from genetics to genomics is taking place at a pace that puts even Moore's Law to shame. It took 13 years and a $3 billion worldwide effort to produce, in 2003, the first complete map of the human genome. Five years later, the cost of sequencing a human genome was down to about $1 million, and the U.S. National Institutes of Health offered major research grants aimed at bringing the cost down to $1,000. Once that happens, sequencing might be performed routinely at birth to facilitate personalized medicine (see Laurie Rowell's feature "Personalized Medicine: All About You" on page 26).

I am not an expert on these topics but have had the privilege of discussing them extensively with someone who is—my bioinformatics colleague and co-author Pietro Liò of the University of Cambridge Computer Laboratory. What I know a little more about, however, is security and privacy. I am a professional paranoid. I look at new technologies and notice the risks, even as others marvel at the opportunities. From this perspective, I invite you to peek into the future with me.

Here is genetics versus genomics in a nutshell: With genetics you look only at specific places in the chromosomes; with genomics, you look at all the information the chromosomes encode. For example, the traditional "DNA fingerprinting" used by police agencies the world over involves genetics rather than genomics; the Combined DNA Index System, or CODIS, standard procedure looks at only 13 specific places in the chromosomes. At each such place, one of several independent variations may occur. If two DNA samples have the same variations in each of the 13 places, the two samples are considered a match. It is not the case that the three billion base pairs of the whole genome are compared one by one.

Why is the difference between genetics and genomics relevant? What if the police (or hospital) database stored the whole genome instead of just those 13 markers?

Let's for the moment think science fiction rather than science, imagining a society in which genomics is as commonplace as cellular telephony is today. Here are a couple of flashes from this scenario. The genome is the full blueprint for building the individual, so

Coninued on page 39

Coninued from page 40

there might be a piece of software that, given the genome, produces a photo-realistic rendering of what the person looks like. Even though it can't guess the effects of the environment, the software will get things wrong only to the extent that two identical twins can look different from one another. The software is very popular with stalkers, who use it to look at celebrities without clothes, and with the secret police, who use it to track dissidents.

Or perhaps the partner of your dreams looks you up online before that crucial date (as they might today through social networks) and discover your genetic predisposition to, say, cancer, Alzheimer's, or criminal behavior. Your employer, health insurer, bank manager, and landlord might do the same. The fact that there may be legislation to prevent them from discriminating against you on the basis of the information won't make your interactions with them any less awkward once you know that they know.

THE PARTNER OF YOUR DREAMS MIGHT LOOK YOU UP ONLINE BEFORE THAT CRUCIAL DATE AND PREEMPTIVELY REJECT YOU BECAUSE OF YOUR GENETIC PREDISPOSITIONS.

But prejudice, as defined as categorizing people not based on what they actually do but on a preconceived idea of what they might do, does not only happen in science fiction. In Britain, for example, 77 percent of young black males have had their profile stored in the police's national DNA database, as opposed to 22 percent of young white males. Prompting the chairman of the country's Commission for Racial Equality to say, "Black males are more likely to be stopped just because they are black males," according to a 2006 article in the *Telegraph*. Similarly, in the context of uneven application of antiterrorism laws, human rights organization Liberty reports, "Police powers have been used disproportionately against the Muslim population in the U.K. The majority of arrests have been of Muslims, a large number of whom were subsequently released without charge, or charged with offenses unrelated to terrorism."

Some might claim that such stereotyped judgments start from a semi-rational statistical justification, but it is the indiscriminate and universal application of the fallible heuristic that is grossly unfair and damaging. If prejudice might already be so influential and disruptive in today's pre-genomics society, how much stronger could it be when fueled by the genomic data available in the future society of our sci-fi cartoon? How many more categories of citizens will be similarly discriminated (at school, at work, in social interactions), not for anything they actually did but merely for their genetic makeup? Especially as further genetic research correlates specific genes with the likelihood of developing certain diseases or behavioral patterns.

As the cost of data storage continues to fall, buying a larger disk is cheaper than deciding what to delete from the old one. The consequence is denied oblivion. No more forgetting; once in digital form, data is forever. This doesn't apply just to genomics but to communications, travel records, photographs, and RFID sightings. For each of these categories, taken individually, advocates may certainly find compelling advantages in allowing the collection, storage, and mining of the information. But the corresponding loss of personal privacy quickly grows out of proportion when several categories are affected simultaneously.[1]

What is the right trade-off? Find your own answer and support it. An informed debate is needed, as well as some action, while we still have an opportunity to influence the outcome.

Frank Stajano (frank.stajano@cl.cam.ac.uk) is a tenured faculty member at the University of Cambridge (U.K.). His main research and consulting interests are systems security, privacy in the electronic society, and ubiquitous computing. He is the author of the book *Security for Ubiquitous Computing* (Wiley, 2002).

DOI: 10.1145/1655737.1655749

[1] Would you want to live in such a dystopian sci-fi society? Please see the discussion paper on my webpage (http://www.cl.cam.ac.uk/~fms27/) I wrote with Liò and two other co-authors from a legal background, revisiting these issues in greater detail.