



Chris Wolf's Soap Box – Berkeley Big Data Workshop, March 31, 2014

These sessions have well demonstrated how big data promises to deliver enormous potential benefits to consumers and society-at large. And I think it is fair to say that sophisticated companies recognize that you can't unlock the value of data if you don't institute responsible data practices. Last week the *Wall Street Journal* had a special section on big data, and reported that, quoting now, "The troves of new information are becoming a high-stakes issue for companies, with significant opportunities to leverage the data and boost business – but also to make significant missteps." In 2008, Jules Polonetsky and I started the Future of Privacy Forum, to bring together industry, academics, and researchers to help avoid the missteps in the collection and use of personal data, and to develop and promote leading practices.

The Future of Privacy Forum has thought a lot about big data, especially in the context of the Internet of Things, and we have these high-level recommendations:

Let me start at a very high level: Privacy professionals often hear about the risks of big data, but don't fully take stock of the benefits in performing a cost-benefit analysis. It is hard to do a cost-benefit analysis if we are only talking about the costs. Researchers, academics, and industry are using big data to deliver big benefits. We need to understand and promote those benefits so that we can more reasonably evaluate whether and how to address the risks that may arise.

Getting into the privacy weeds, when it comes to Big Data, we need to embrace a flexible application of the Fair Information Practice Principles. For over 40 years, the FIPPs have

guided how we can promote privacy while collecting, using, and sharing personal information. And the FIPPs will continue to serve us well. However, we shouldn't blindly stick to traditional implementations of the FIPPs that may be poorly suited for big data and other new technologies. We need to encourage industry and other stakeholders to develop innovative privacy mechanisms that promote the FIPPs in new contexts.

Privacy policies are valuable as accountability mechanisms and as a way to force companies to look at their privacy practices. But we should encourage industry to think of other ways to promote notice and choice. We cannot expect consumers to read and consent to hundreds, if not thousands, of privacy policies on a daily basis.

And transparency cannot just mean better privacy policies. Organizations need to engage with consumers in meaningful conversations that will help align interests and expectations. FPF has previously called for what we call the "featurization" of data. Featurization can allow consumers to grant third-party access to personal data in order to provide value-added services. This can promote transparency by helping consumers see the value that big data can bring.

Use limitation is a valuable FIPP. But if use limitation is delimited by the purposes specified at the time of collection, we may never be able to obtain the unexpected benefits that are promised by big data and other technologies. As the White House Consumer Privacy Bill of Rights recognized, privacy is promoted when data are used in ways that respect the context of collection. Context changes over time. Respecting the context of collection must allow room for innovative uses of data that deliver unexpected benefits. And respect for context must be framed in a way that reflects the dynamic nature of social and cultural norms and the subjective nature of consumer trust.

Data minimization is a FIPP that may seem anathema to big data. But the goal of data minimization is to reduce the circulation of personal information. Limiting collection of personally identifying

information that is not needed for analytical purposes is one way to do that, and so too is the use of reasonable de-identification.

The FIPPs can work if we apply them flexibly. The White House recognized this in its Consumer Privacy Bill of Rights. FPF encourages the administration to continue embracing that insight.

FPF also believes that de-identification is not a lost art. Over the past few years, several studies have shown that given time and resources, supposedly de-identified data sets can be re-identified. FPF recognizes that technical de-identification doesn't guarantee that data will never be re-identified. But organizations that use administrative and legal safeguards along with modern de-identification tools can promote both privacy and innovative data use.

Another thing to consider, as proposed by FPF Advisory Board member Professor Ryan Calo, is the use of internal review boards to promote accountability. Privacy can sometimes become a check-box compliance function. If a practice isn't prohibited, then it's permitted. Internal review boards may better promote institutional accountability when looking at big data practices. Review boards could serve as effective institutional checks by evaluating the selection of data sources and analytical tools, assessing the potential benefits, and considering potential risks and ethical considerations.

Finally, in addressing Big Data, it is essential to promote global interoperability. If big data is balkanized into geographic regions, the value of big data is diminished. FPF encourages government to actively promote and maintain existing frameworks that promote interoperability. Those frameworks include the US-EU Safe Harbor and the APEC Cross Border Privacy Rules System. And new frameworks, such as those that might emerge in trade negotiations are well worth considering.