

Wake Up and Read the Small Print!



Businesses and government understand the value of Big Geospatial Data, while citizens are only learning their role in contributing to Big Data, argues Adena Schutzberg.

In May, [Strava](#), a company that offers tracking and competitive tools to cyclists and other athletes, announced a new data product that prompted some distress. The product, Strava Metro, includes detailed but anonymized geospatial data about cyclists' routes in an area of interest. One of the first customers, the Oregon Department of Transportation (ODOT), began talking to Strava last year and received data in December. ODOT paid \$20,000 for data about just under 18,000 cyclists' rides.

This spring, when the details were made public, some Strava users were disappointed to learn that the company was making money from "their" workout and travel data. Clearly, these individuals forgot the well-known warning that if a product is free (and a version of Strava is), you are the product. Strava's privacy policy is quite clear: "Information about you and your use of the site may be aggregated with other information collected on the site or otherwise used in ways that do not personally identify you or constitute personally identifiable information. This type of aggregated or statistical information may be used by us to improve the quality of the site or for other purposes that we may deem appropriate."

It's my understanding that this and related policies allow Strava to aggregate data and offer products like Metro for sale. The target market for the datasets are organizations that want to build better solutions for bicycles in specific geographies. ODOT has already tapped its dataset to find "better" bike counting locations around Portland's bridges and to determine where to put rumble strips on highways.

While one group of users grumbled about Strava making a profit off their data, another considered joining Strava specifically because ODOT used these routes as a source for planning. Strava users tend toward the more competitive, and perhaps the wealthier cyclist, so there is a concern that the data are not representative of the bicycling population overall. An ODOT study found that about 2.5% of riders it counted using traditional methods used the app. Still, experts suggest (and I agree) it's unlikely another application, such as Portland, Oregon based RideWithGPS, could provide a higher percentage.

Strava's agreement with Metro subscribers restricts any sharing of the raw data. However, user organizations can create and offer reports and heat maps based on them. The public and non-profits are welcome to look at Strava generated heatmaps for free.

My sense is that businesses and other gatherers of big data, geographic and otherwise, have done their homework. They listened in the early days of user-generated data on the Web, back in 2000s, when the likes of Tim O'Reilly encouraged Internet startups to "own the data." If that was not possible, he went on, startups should attempt "to organize the data for use by others." The startups learned about crowdsourcing and about how the more people who use a service, the more valuable it is. Strava and other companies like it are the Amazons of this decade. Instead of selling physical products (books and shoes), they sell individuals a service and perhaps a community, and, in turn, collect and organize the data they generate for sale.

The individuals who use these services and participate in these communities who grumble at efforts like Strava Metro, despite warnings related to privacy, have not done their homework. They haven't read the privacy policies. They haven't considered exactly how their participation in these communities benefits the companies who host them, whether a fee is paid or not. The good news is that at least some potential users are realizing that perhaps the benefit may come back to them, such as those considering joining Strava.

It's 2014. A good proportion of our planet's population carries cell phones loaded with sensors and location determination tools everywhere they go. We, as potential users of services and joiners in communities need to understand how we are part of big data, and the implications for ourselves as individuals and as residents of a local or global community.

This article was published in GIS Professional June 2014