

# Streaming Data

[Eric Solano, Ph.D.](#)

I have worked with streaming data at two different organizations. With [Point Inside](#), I helped collect and analyze beacon data from brick-and-mortar customers for [proximity marketing](#). Currently, with [Continental Tire](#) I process and analyze [telematics streaming](#) data from vehicles.

Point Inside used small, battery-operated [wireless devices called beacons](#). Beacons transmit Bluetooth signals to nearby smartphones. In the context of retail, beacons reach out to customers that have Bluetooth enabled and the right retail apps downloaded onto their phones. I used these beacons to help our brick-and-mortar customers use proximity marketing to improve shopping experiences and increase sales. There were numerous benefits to adopting beacon technology. We helped our customers to create more tailored experiences for their customers. We saw multiple benefits of using beacons and beacon's data: advertising was done in a tailored and personalized way, deals and recommendations were sent to customers, we tracked shopping habits and preferences to help our retailer customers access insightful consumer data. I also used data collected from beacons to perform additional analysis to help improve marketing and increase sales for our brick-and-mortar customers.

At Continental, we stream telematics data to the [Telematics Backbone \(TBB\)](#). The Telemetry Backbone provides Continental with telemetry data of any kind. The [TBB](#) holds data from different [loggers](#) and data enrichments such as CRM Master data or data from other providers such as altitude or weather conditions for GPS data. Data is usually provided in its initial aggregation (usually transactional level) to allow for better analysis.

The TBB's Data Source Extractor (aka API consumer) is responsible for fetching / retrieving data from a Data Source or multiple Data Sources on a regular basis, either as an always online streaming application (e.g. when the Data Source is a message broker) or as a scheduled batch process (e.g. DB, File or REST). Each message received is considered as a raw payload. The Data Source Extractor can transform or filter these messages as necessary and eventually forward them to the TBB's ingestion zone. Once these raw payloads arrive in TBB, the TBB Processor will process and persist these messages. The Ingestion Zone is a set of [Apache Kafka](#) topics. The Data Source Extractors produce messages and send them to the Kafka topics. The TBB Processor consumes messages from the Kafka topics. For each payload type there is a dedicated Kafka topic, depending on the Data Source and the implemented Data Source Extractor. In most cases, there is one Kafka topic per Data Source Extractor. In some exceptional cases there might be multiple topics per Data Source Extractor. This is required when different payload types cannot be distinguished easily, e.g. the format and structure for GPS data is the same as for vehicle data and there is no clear discriminator in the payload. Or, if payloads are delivered in different file formats, e.g. csv vs. binary.

