Instrumenting Telemetry on an E-Commerce Application at Scale

Anand Ganesh

Independent Researcher

I. INTRODUCTION

A. Significance of Data-driven decisions

With the rising markets resulting from increased globalization and access to the internet, many more markets are coming online and developing nations have access to more digital devices. Countries such as India and China have five times as many mobile devices in the United States.

Different markets have different user behaviors and cultural expectations. Consider the case of Japanese website design.

Fig. 1. Yahoo.com in Japan

If one were to leave a feedback form at the end of the UI experience like this in North America, you can expect that the suggestions will be filled with people complaining about the clutter, how incredibly hard it was to find information and that the entire experience did not feel safe to use a credit card to make a digital purchase. A country like Japan has their users well accustomed (and even prefer) more dense UI design. This comes from a cultural belief that access to more information makes the consumer feel safe and trust the product. Japan is a high-context culture, and the US is not [1].

A one-size-fits-all solution would simply not work for UI design here, and there might be more factors that are not front and center that can be picked up from data. For instance, the size of the average smart phone being smaller or their aspect ratio's being a little different can mean that a certain portrait picture is not fully visible or does not even render on mobile. If we were to track launch impressions for a new product (a sequence of signals that indicate that the new product has been seen and discovered by a

large section of the market), we might find out that the title image did not follow the pattern of previous releases when it comes to impressions in the first week. This can inform decision-making about whether to fix or ignore the problem and not be blindsided by it.

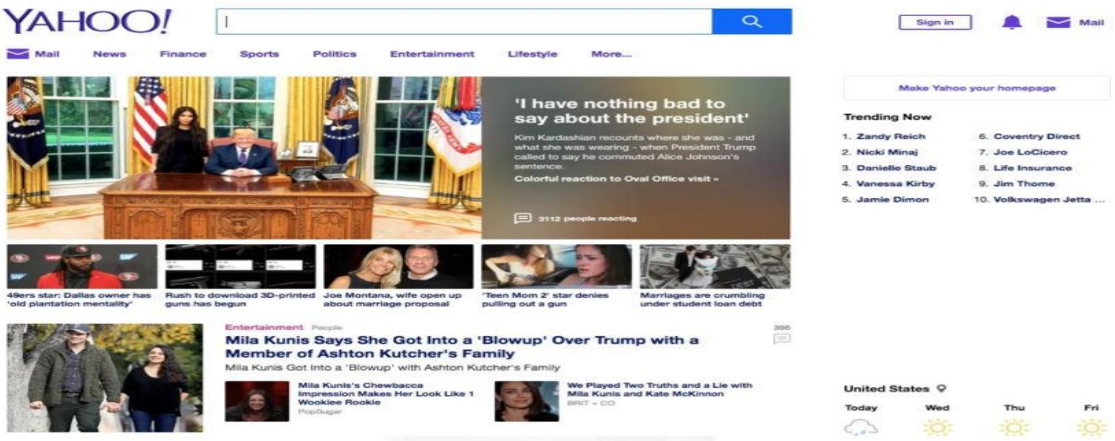

Fig. 2. Yahoo.com in the US

B. Digital E-commerce telemetry

It is important to have data backing decisions in markets that are different from one's own, to understand biases, afflictions, and different emerging patterns. It is even more important to understand user behavior when trying to sell products on an e-Commerce application as buying can be a very tricky and high-stakes decision. Depending on the country, region and demography of the user, a given user can be least likely to spend money if the website is too slow, too cluttered, too simple or simply unfamiliar. In such a situation, it is important that all user experiences have accompanying telemetry.

C. Efficient Telemetry system

However, instrumenting telemetry can be tricky and quite hairy as one can quickly make the mistake of collecting redundant data. One might ask what the problem of collecting too much data is; After all, you can always omit unnecessary data and you will not run the risk of not having data when you need it.

However, there is a trade-off with sending too much data. When considering the global market, sending out data from consumer devices can be expensive, prone to privacy filters, and frowned upon due to the excessive bandwidth occupied by what is considered tertiary or non-essential functionality. [2] Drawing the line between collecting raw data from the user and inferring data from existing raw data points is crucial to having an efficient telemetry system that is also scalable. [3] The study is a design guideline for instrumenting telemetry across multiple digital storefronts, what assimilating all those data points can help you achieve, and how best to go about doing so.

II. METHODOLOGY

A. Observability

One of the goals of a modern telemetry system is observability. Observability can be defined as the characteristic of a compute system by which its internal states can be understood using its output signals. Any good telemetry system usually has a combination of different forms of telemetry data to achieve this. [4]

B. *The Open Telemetry standard*

An aspect of creating e-commerce applications is the various digital surfaces that will have to be supported to reach a wide number of users. Instrumenting telemetry from various client devices and technologies, each having their own ways to send data can lead to an arduous task of assimilating all the data and processing it. Sometimes, this might even result in duplication of efforts to achieve the same output simply because a certain client device emits dissimilar data. Combine this with different tools that are out there that possibly expect data in a certain shape to be used effectively, and the flexibility of a telemetry system becomes extremely low.

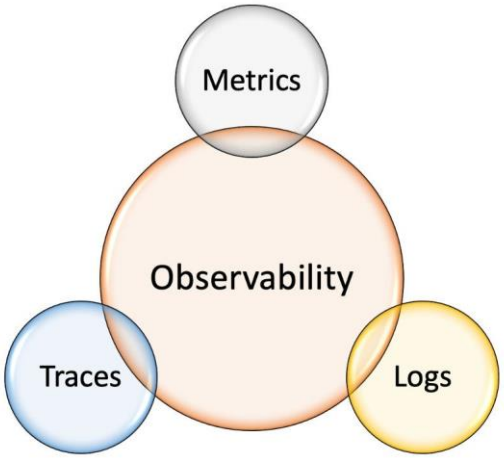

Fig. 3. Observability

To avoid all this and to have a vendor-agnostic approach to instrumenting data and processing it, the open telemetry standard is essential for any large-scale distributed system. The Open Telemetry Standard is a telemetry framework that defines the necessary semantic conventions and protocols that help standardize the sending, collecting, and management of telemetry data.

1) Logs: Logs are the raw data that record an event or an activity in the system. It usually forms the atomic source of data for most telemetry systems and can be text-based or structured-data, making it easier to process and manage. They contain metadata pertaining to the time, the machine that emitted it, metadata about the user, if applicable, and any other custom message from the system.

2) Traces: Traces are one of the most useful and integral part of the diagnostic chain. Traces contain multiple smaller segments called *spans* and each span is a logical operation performed by a single request through the system and tracks that operation over time. Traces therefore ”trace” the entire path of a request through a system often across multiple different services. In an increasingly distributed world with different services deployed on data centers around the world, traces are helpful in correlating a request across multiple systems.

3) Metrics: Metrics are aggregated data that track a particular value, usually over time but can sometimes be other factors like users or cost. Metrics are extremely important from both an engineering and a business standpoint as they assimilate all the raw information into more readable chunks that can help us understand any aspect of our workflow as long as we instrument them correctly.

C. *e-commerce telemetry*

Besides making sure that an e-Commerce system works as expected, one of the main uses of e-commerce telemetry is to lay the foundation for analyzing user and system data. [5]. We now dive into

various telemetry events that bring our open telemetry framework to life in the e-Commerce setting.

1) Action events: In order to satisfy the three spheres of observability, we will need a set of base events or *logs*. Action events constitute a raw log that records the occurrence of an event. It is helpful to disambiguate between *application* and *user* events to isolate user behavior studies from system diagnostics. It can be application events like ”details page loaded” or user events like ”user clicked the buy button”

2) Network events: Network events are the building blocks of spans that help trace requests over a distributed system. Network events can also help uncover bottlenecks for performance improvements and improve security. In order for a given user operation to be traceable, across services, a unique correlation identifier is necessary to trace the life-cycle of the request through the system. This correlation identifier can also be used in metrics to measure certain values for a particular request or user over time.

3) Content View: This is an event that is a representation of what the user has seen on their client device. Usually this is stalled or de-bounced so that it only triggers after a set amount of time. The idea is to capture which part of the digital real estate gets more attention, traction, and what content is interesting to the user and what content is glanced over. Experiments with advertisements and hero-posters on the home page are the most likely to benefit from this as well as game trailers, posters as their purpose is to attract attention to a particular object of interest.

III. CONCLUSION

In this paper, we explored the importance of telemetry in e-commerce applications and how it enables data-driven decision-making. We examined the challenges of designing an efficient and scalable telemetry system, highlighting the significance of observability in understanding user behavior and system performance. By leveraging the Open Telemetry standard, e-commerce platforms can standardize data collection across multiple digital surfaces, ensuring consistency and interoperability.

We also discussed the three pillars of observability—logs, traces, and metrics—and their role in diagnosing issues, optimizing user experience, and improving operational efficiency. Furthermore, we outlined key telemetry events, such as action events, network events, and content views, which provide actionable insights into user interactions and system behavior. As e-commerce continues to expand globally, understanding regional user preferences and market-specific behaviors be- comes crucial. Telemetry serves as a fundamental tool for busi- nesses to adapt their digital storefronts, optimize performance, and enhance user engagement. However, while comprehensive data collection is valuable, it is equally important to balance telemetry efficiency with privacy considerations and resource constraints.

Future research can explore advancements in AI-driven telemetry analysis, predictive analytics, and privacy-preserving data collection techniques to further refine how telemetry is used in large-scale e-commerce applications. By implementing a well-instrumented telemetry system, businesses can gain deeper insights, improve decision-making, and drive innovation in the ever-evolving digital marketplace.

REFERENCES

[1] T. Doi and A. Murata, “Examining the cultural differences of users’ characteristics between the united states and japan related to user interface design,” in *Advances in Physical, Social & Occupational Ergonomics: Proceedings of the AHFE 2021 Virtual Conferences on Physical Ergonomics and Human Factors, Social & Occupational Ergonomics, and Cross-Cultural Decision Making, July 25-29, 2021, USA*. Springer, 2021, pp. 391–396.

[2] X. Wang, “Collection telemetry data for a static code analysis tool in a data protection compliant

way," Master's thesis, 2021.

[3] M. Otero, J. M. Garcia, and P. Fernandez, "Towards a lightweight distributed telemetry for microservices," in *2024 IEEE 44th International Conference on Distributed Computing Systems Workshops (ICDCSW)*. IEEE, 2024, pp. 75–82.

[4] OpenTelemetry, "What is observability?" 2019.

[5] A. Ganesh, "Building high volume digital purchase experiences across global markets," *IJLRP-International Journal of Leading Research Pub- lication*, vol. 12, no. 6, 2024.