

I'm not a robot

































documentation, it can require a great deal of debugging. Any discrepancy between the tracking design and the tracking implementation can cause systemic errors, and combining different data sets leaves a lot of room for human error. Modeling data at scale - Poor naming conventions, the wrong levels of granularity, overuse of calculated fields—all these things and more need to be considered when scaling your data modeling. Small problems can proliferate quickly as data sets grow. Organizational challenges Communicating within teams - When the ‘where, how and why’ of data is not recorded systematically, data is not explainable within teams. Knowing exactly what a tracking designer intended and how the front-end dev interpreted this information can require Sherlock-Holmes-level skills and some seriously long Slack threads. Maintaining backwards compatibility - Understanding where breaking and non-breaking changes have been made to tracking is essential to maintain backwards compatibility. This represents a significant challenge for many teams. Learn more about breaking and non-breaking changes. Scaling self-serve - As teams grow, compounding complexity is inevitable (see the image below). In order to scale self-serve capabilities in your data stack, it’s critical to understand the needs of each new team member and how these intersect with the existing team. Due to the complexity of behavioral data, this can be particularly challenging. Maintaining consistency - Versioning data schemas and keeping track of these changes is a significant challenge. When data dictionaries are kept as ad hoc spreadsheets, the issue is exacerbated. Compliance challenges As you have probably gathered, all these challenges are interlinked. If you don’t understand data lineage, data is not explainable to your teams. It also means data is not easily auditable for compliance inspections. Auditability - Retroactively making data auditable can be a nightmare for data teams, and whole quarters can be lost if planning is not done up front. GDPR/compliance - This includes the right to be forgotten, consent, storage location, information security and access, and many other data laws and regulations. When data teams don’t trust that GDPR/compliance is in place, they often take the ‘safe’ route of discarding the data. This wastes highly valuable information. How Snowplow solves for the challenges with behavioral data Tracking - First party tracking means ITP does not apply—you can track users for up to 2 years, even on Safari and Firefox. Single source of truth - Data does not need to be reconciled across sources—it’s created from scratch to match your business logic and vocabulary, and then prevalidated so it arrives in an atomic data table ready for AI and BI use cases. Lineage - Created data has 100% transparent lineage. With Snowplow, the meaning of each metric is recorded in a human and machine readable format in JSON schemas. Alignment - A centralized UI helps teams navigate the complexity of managing behavioral data, controlling permissions, showing tracking in a visual way, and managing schemas. Consistency - By using Snowplow’s Universal Data Language, you can ensure meaning is tightly documented and versioned across teams. Compliance - You have full ownership of your data. The whole Snowplow infrastructure lives on your own cloud—meaning a choice of storage location, full GDPR compliance, the option to have multiple pipelines, and the ability to record the basis for capture with each event.