

Harnessing the power of data lineage to deliver compliance and business insight

by Harpreet Singh
Head of Post Trade Solutions
Luxoft

What is, and why do we need, data lineage?

In words of one syllable (almost), it shows how data flows through systems from start to finish. As we become a more connected world exchanging huge amounts of data every day, without data lineage, the modern enterprise would find it very difficult to prove if critical information was trustworthy or inaccurate.

Business decisions would have to be made on hunches driven by incorrect data. In fact, data lineage is fast becoming invaluable (in the purest sense of the word).

The data lineage power surge won't be easing up any time soon

Three main agendas are driving the need for data lineage — regulatory, business and transformation. Regulators put a stake in the ground for both buy and sell sides at least a decade ago. They asked for unique IDs to be preserved in the form of unique trade identifiers (UTIs) or unique swap identifiers (USIs) in EMIR and Dodd Frank. GDPR, MIFID II and similar initiatives took the need for data lineage to the next level. It is no longer about the ability to sustain these identifiers, but about being able to provide evidence that there is a control framework to manage the quality of data being reported.

On the business side, with calculation speed increasing and disk space getting cheaper, we can perform data analytics on huge sets of data much faster and within a reasonable timeframe. However, for that output to be useful, the underlying data needs to be complete and correct. And data lineage plays a pivotal role in advancing that objective.

The third aspect is the transformation agenda — cloud migration, for example. Understanding, and being able to trace your data, leads to substantially improved results for analysis, testing, post production, validation and so on. Data lineage reduces the time spent on impact analysis, testing and other aspects of migration.

An underlying theme of all three scenarios, the quality of data must be preserved as it travels through hundreds of systems and business processes.

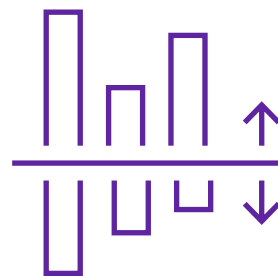
So, because of the adoption of these agendas (and more), we're likely to see data lineage play an even more important part in our everyday business.



The far-reaching **consequences of data lineage**

Although, everyone agrees that data lineage is a must have, its benefits are not obvious. It is considered a technology issue but, in reality, the impact of data lineage is widespread. Take the allocation of cost or revenue for example; a brokerage or commissions perhaps. The brokerage is calculated in front-office systems, then feeds down to a middle-office system as split or merged orders, which move on toward operational systems (more changes due to netting). Eventually, they travel on to a few finance systems where they're assigned finance codes. The finance codes are allocated to business units in cost management systems. If any of these processes are not set up correctly, the banks end up with incorrect sales or cost attribution.

First and foremost, this is a data lineage issue and in large or global institutions, few people take the time to understand the end-to-end flow. So, it becomes a difficult problem to solve. A well-designed, data-lineage architecture can help resolve these types of issues.



Where do you invest to get the best out of data lineage?

When investing in data lineage, you need to align it with business strategy. Typically, you can go broad and shallow, or narrow and deep. But unless you've got a very good architecture and very deep pockets, you can't go broad and deep. More importantly, you need to

tie the business benefits to your investment and have checkpoints to prove that suitable returns are being achieved. So, either you go wide but shallow, or pick your most pressing problem and go for it. The decision is driven by your firm's business strategy.

How can you maximize the business benefit derived from data lineage outputs?

Our plans are based on what's happening in the market. From 2010 to 2020, there's been a 5,000% growth in the amount of data that the world has captured, copied or consumed. At the same time, less than 1% of structured data is being used to make active decisions. And when it comes to unstructured data, less than 1% is used. In other words, 99% of the data is left unused. Why is that? We need to answer some very basic questions like, "Which business decision are we trying to make?" and, "Which business problem are we solving?"

The beauty of data lineage metadata is that you start with one focus area and you reuse the data — obviously, the analytics on top of that changes depending on whether it's fraud or client analytics and so on. But you

can start in one place and improve. Again, why isn't this data being used and what's happening with the end customer? We believe that if we can provide actionable insights, spontaneously, then the end user will start using the data.

Compared to traditional institutions, digital banks have been a little more successful in giving that data back to users and saying, "How do you want to use it?" But many financial institutions, especially in the capital markets, are still lagging behind. Unless we solve this data lineage issue, we won't be able to make the end user self-sufficient because they will not understand the data. Why? Because we don't understand it either.

What is the future of data lineage?

Data lineage is evolving. First, it was about data elements and ensuring they were persisted correctly. Then it led to the analysis of calculations that were being done on those data elements to provide insights. Now, it's also about business decisions

made on the back of those insights and being able to validate those decisions. As technological innovation continues, degrees of lineage will further expand, enabling us to connect underlying data points with day-to-day business decision-making.

About **the author**



Harpreet Singh

Head of Post Trade Solutions
Luxoft

Harpreet leads Banking and Capital Markets Post Trade Solutions for Luxoft. The department is responsible for delivering innovative business and fintech solutions to our clients in post trade functions such as regulatory, liquidity management and operational resilience. He has more than 20 years' experience in implementing change and driving growth as program director, consulting head and COO for large banks and consulting organizations.

About Luxoft

Luxoft is the design, data and development arm of DXC Technology, providing bespoke, end-to-end technology solutions for mission-critical systems, products and services. We help create data-fueled organizations, solving complex operational, technological and strategic challenges. Our passion is building resilient businesses, while generating new business channels and revenue streams, exceptional user experiences and modernized operations at scale.

luxoft.com