

The Talend logo is a solid red circle with the word "talend" written in white lowercase letters inside it. The background of the top half of the image features a series of overlapping, curved red lines that create a tunnel-like effect, converging towards the center where the logo is located. The bottom half of the image is a solid dark blue-grey color, with the same curved line pattern continuing from the top half, creating a sense of depth and continuity.

talend

**3 Steps to Deliver
Data You Can Trust at
the Speed of Business**

Why trusted data is the key to digital transformation

We've entered the era of the information economy, where data has become the most critical asset of every organization. Data-driven strategies are now necessary for success in every industry. To support business objectives such as revenue growth, profitability, and customer satisfaction, organizations are increasingly reliant on data to make decisions. Data-driven decision-making is at the heart of your digital transformation initiative.

To provide the business with the data it needs to fuel digital transformation, organizations must solve two problems at the same time: speed and trust.

The data must be timely, because digital transformation is all about speed and accelerating time to market — whether that's providing real-time answers for your business teams or delivering personalized customer experiences. However, most companies are behind the curve when it comes to delivering technology initiatives quickly. According to Forrester, only **40% of CIOs** are delivering results against the speed required.

While speed is critical, it's not enough. For data to enable effective decision-making and deliver remarkable customer experiences, organizations need data they can trust. Being able to trust your data is about remaining on the right side of regulation and customer confidence, and it's about having the right people using the right data to make the right decisions. According to the Harvard Business Review, on average, **47% of data records** are created with critical errors that impact work.

Speed and trust are often at odds so it's common for organizations to focus on one or the other. Many organizations default to speed to meet data users' expectations for ease and convenience, as well as fit their own time or budget constraints.

These companies allow developers to hand-code integrations or do one-off projects with niche integration tools in order to get results fast. But while these tactics may solve for speed in the short term, they are not scalable as companies grow, and create quality and compliance risks due to the lack of oversight. On the other hand, organizations that try to solve the data trust problem often create a "no" culture with the establishment of strict controls and an authoritative approach to governance. This can hinder the innovation and agility so necessary to compete in today's business environment; businesses that operate too slowly risk being left behind.

This eBook outlines the three steps that will ensure your company can have access to data you can trust at the speed of your business. It is possible to have both speed and trust without compromise. Building and maintaining data trust requires organizations to govern data holistically throughout its lifecycle to ensure regulatory compliance and data security, without overly restricting data access and usage.

According to Forrester,
only **40% of CIOs**
are delivering results
against the speed
required.

Step 1: Discover and cleanse your data

Your challenge is to overcome the obstacles by bringing clarity, transparency, and accessibility to your data assets. You have to do this wherever this data warehouse resides — within enterprise apps like Salesforce.com, Microsoft Dynamics, and SAP; a traditional data warehouse; or in a cloud data lake. You need to establish proper data screening so you can make sure you have the entire view of data sources and data streams coming into and out of your organization. These are some technologies that can help you:

Explore any data with Talend Data Catalog

The auto non-technical discovery capabilities of [Talend Data Catalog](#) facilitate data screening for non-technical people within your organization. Data Catalog provides you with automated discovery and intelligent documentation of your datasets in your data lake. It comes with easy-to-use profiling and sampling capabilities that help you to assess data at a glance. With trusted and auto profiled datasets, you have powerful and visual profiling indicators, so users can easily find and the right data in a few clicks.

Highlight data quality issues

Nobody opts for surgery without a precise and close examination beforehand. The same applies to data profiling. You need to understand your data before fixing it. Since data often comes in hidden, inoperable, or unstructured formats, an accurate diagnosis helps you to have a detailed overview of the problem before fixing it. Data profiling enables you to work in-depth with your datasets and assess multiple data sources using the six dimensions of data quality. It helps you to identify data that is inaccurate, inconsistent, or incomplete. If you apply pattern recognition and data profiling during the discovery phase, you have a better picture of what's in the data and what's going on in any of your datasets.

Delegate cleansing in the cloud

The cloud drastically extends the boundaries of data. Lines of business use their own applications, and products, people, and assets create their own data pipelines through the web and the Internet of Things. Self-service apps such as [Talend Data Preparation](#) allow anyone to access a dataset and then cleanse, standardize, transform, or enrich the data. Because it is easy to use, people will spend less time crunching data in spreadsheets or expecting their colleagues to do that on their behalf. Thanks to cloud platforms, data is exchanged seamlessly between business partners and data providers.

Step 2: Organize data you can trust and empower people

While step 1 helps to ensure that the incoming data assets are identified, documented, and trusted, now it is time to organize the assets for massive consumption by an extended network of data users who use it within the organization.

Create a single source of trust

When you collect all the datasets together into a single point of trust with a single point of control you can assign roles and responsibilities and operationalize your governance from the get-go. It is one of the key benefits of data cataloging: regroup all the trusted data in one place and give access to members so that everybody can immediately use it, protect it, curate it, and allow a wide range of people and apps to take advantage of it. Centralizing trusted data into a shareable environment saves time and resources.

Encourage people with data curation

Why is empowering data curators so important? A data governance project is not just about access for all to trusted data. It's also about the enrichment and curation of trusted data to produce valuable, accurate insights out of the data pipelines. Making data custodians accountable for data accuracy and value with role-based, easy-to-use applications helps raise the level of trust in the data for all users. As per Wikipedia, data curation “is the organization and integration of data collected from various sources. It includes annotation, publication, and presentation of data to make sure it's valid over time.”

Orchestrate stewardship

Most data owners realize that they should act as orchestrators rather than manage everything in their data domain themselves. The collaborative part of data management here makes a great deal of sense. You will need to engage — occasionally or regularly — the ones who know the data best for data certification, arbitration, resolution, or reconciliation. An application such as [Talend Data Stewardship](#) enables data stewards to design, orchestrate, and launch “stewardship campaigns” to request inputs to dynamically enrich data through this process. Anyone can be promoted at any time to be a data steward and participate in the data value chain.

Step 3: Automate your data pipelines and enable data access

Now that your data is fully under control, it is time to extract all its value by delivering it at scale to a wide audience of authorized humans and machines. In the digital era, scaling is a lot about automation. In the second step of this approach, we saw how important it was to have people engaged in the data governance process, but the risk is that they become the bottleneck. That's why you need to augment their skills, free them from repetitive tasks, and make sure that the policies that they defined can be applied on a systematic basis across data flows.

Learn with ML for remediation

Advanced analytics and machine learning help democratize data governance and data management because they make things much simpler. They improve developers' productivity and empower non-data experts to work with data by guiding users through their data journey. Machine learning helps to suggest the next best action to apply to the data pipeline or capture tacit knowledge from users of the Talend platform (such as a developer or a steward) and run it at scale through automation.

Enable everyone

Although a data catalog dramatically helps data consumers to find users' data, their data experience doesn't stop there. They need to put the data to work. That's where they need simple, user-friendly apps to fit to their role: a business analyst, or a data engineer, or an IT developer. The cloud is a perfect deployment model to enable everyone with ready-to-use, collaborative applications that can point to any datasets wherever they might be located. Data platforms such as [Talend Data Fabric](#) are designed to provide a whole set of collaborative apps for the trusted data that the three steps approach that is presented here helped to create.

Publish trusted data with API services

It's important to identify apps that require trusted data and connect them to the extremely valuable datasets that your data governance program has established. All the created and automated data pipelines may find a destination not only into business intelligence dashboards but also in apps that will get the most out of them. That's one of the key benefits of APIs; they allow a wide range of applications to consume data assets in an easy way. By leveraging environments such as [Talend Data Services](#) as part of your data governance platform, you can make sure your investment reaches far beyond analytics to feed any apps your organization is consuming to ensure optimal operations.

The 12 labors of the data governance hero

1. Set clear expectations from the start
2. What is your “Emerald City”? Define your meaning of success
3. Secure your funding
4. Don’t go alone
5. Apply governance with a “yes”
6. Start with your data
7. Consider the cloud on your route to trust
8. Be prepared to explain “data”: Don’t expect people to have your expertise
9. Prove the data value: “start small to deliver big”
10. Don’t expect executive sponsorship to be secured
11. Be hands on, not hands off, with data. Lead your trusted data project
12. Live your data challenges

Want to know more?

Take a look at further [data governance resources.](#)

The Talend logo consists of a solid red circle with the word "talend" written in white lowercase letters inside it. The logo is positioned on the right side of the page, with several thin, wavy red lines extending from its left and bottom edges across the lower half of the image.

talend