



A Data Intelligence Company Built On The Premise To Never Run Anything Twice





The Ever-Growing \$3 Trillion Bill

Engineering domains around the world have found their work constrained to passing inspection requirements. Software, more specifically data pipelines, have found itself excluded from that space. The statistical output splashed onto a page could be anything that a line of code dumped their. Data and model changes left unvalidated, leaving the users and developers crossing their fingers hoping no one says anything. With data left uninspected, it is safe to say the system is garbage in, garbage out.

Seeking data quality solutions is imperative to the business lifecycle. With the average impact on businesses being almost \$10 million for poor data quality, seeking an infrastructure for data validation becomes necessary. Data



Quality pipelines ensure integrity, reliability, and consistency across the technological architecture, all while saving costs and increasing productivity.

Financial Costs Bad Data

IBM found that poor data quality strips \$3.1 trillion from the U.S. economy annually due to lower productivity, system outages and higher maintenance costs, to name only a handful of the bad outcomes that result from poor data quality.

the average financial impact of poor data quality on organizations is \$9.7 million per year

When non compliant data is published, the expected fines surpass \$15 million, here are some researched top figures:



Business Disruption Loss:
\$5.1 Million



Productivity loss:
\$3.8 Million



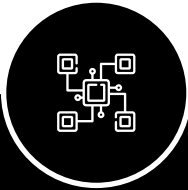
Revenue Loss:
\$4 Million



Fines, penalties, & other:
\$2 Million



Productivity Costs Bad Data



88%

of all data integration projects fail entirely or significantly overrun their budgets because of poor data quality

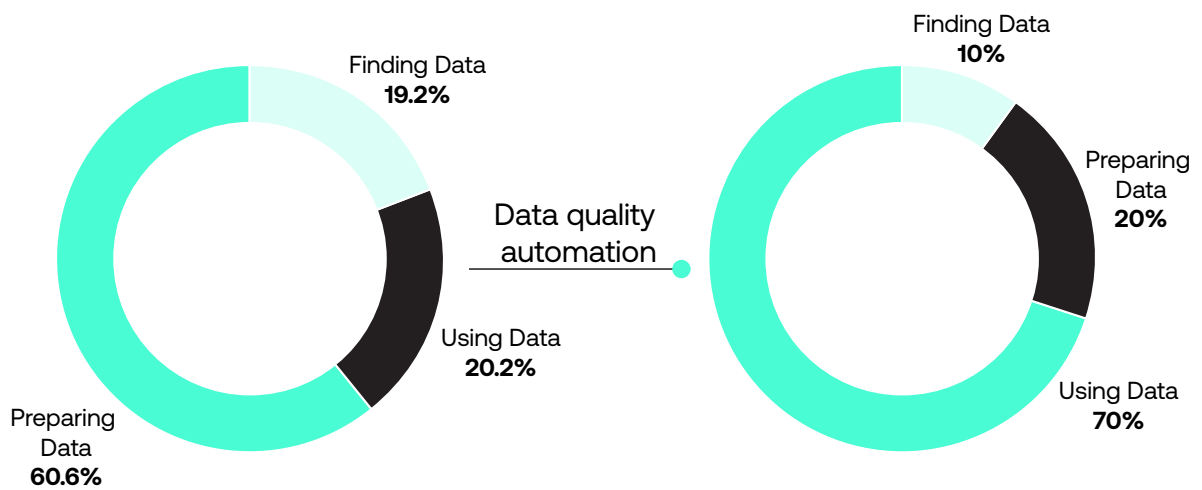


33%

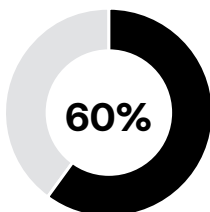
of organizations have delayed or canceled new IT systems for the same reason.

This problem with IT systems is alarming when you acknowledge the importance of data infrastructure modernization for the contemporary business: 70% of productivity growth comes from IT projects. Poor data quality will impact about 10% of the savings you get from IT initiatives, meaning if you save \$21.7 million in labor productivity, \$2.2 million will be lost.

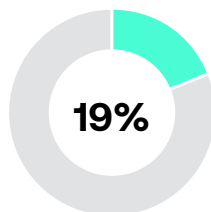
How data scientists spend their time



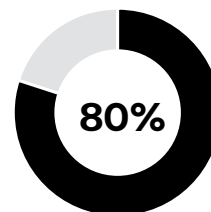
Data scientists spend



60% of their time on cleaning and organizing data.



Collecting data sets comes second at 19% of their time



meaning data scientists spend around 80% of their time on preparing and managing data for analysis.

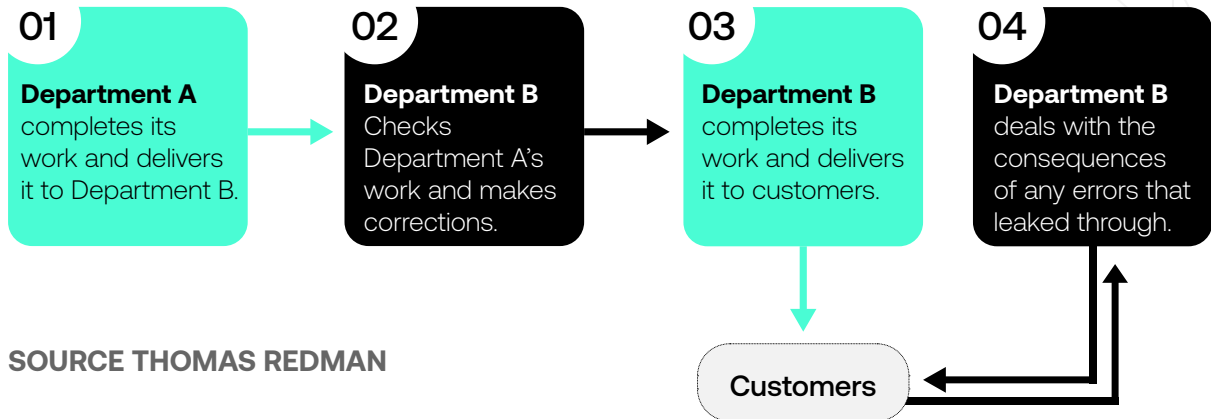


'Hidden Data Factories' driving compliance cost

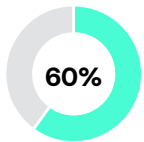
These compliance costs can range from \$5.5 million to almost \$22

The Hidden Data Factory

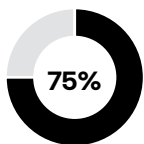
Visualizing the extra steps required to correct costly and time-consuming data errors.



50% — the amount of time that knowledge workers waste in hidden data factories, hunting for data, reproducing and reengineering the data bug for resolution

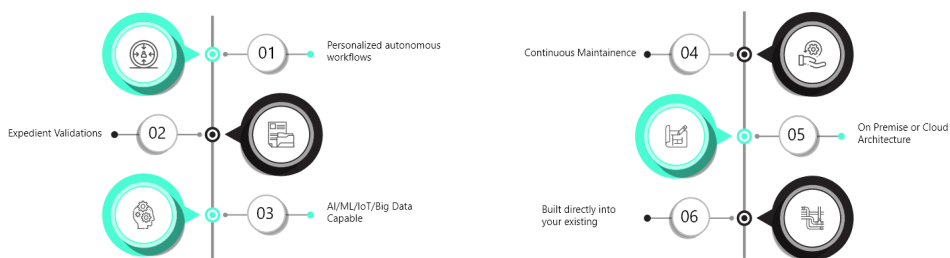


60% — According to CrowdFlower, this is the estimated time that data scientists spend cleaning and organizing data



75% — The fractional total cost estimate associated with hidden data factory involvement with an organization.

The impact for implemented data validation



Time and resources are dedicated to the architectural build and in result, millions are saved. In a quick reminder, CDF implementations, open source solutions, building automated workflows, or partnering with Valence to implement and uphold all your data quality needs is an incredibly smart, business saving decision.

Bansal, Manu. "Flying Blind: How Bad Data Undermines Business." *Forbes*, Forbes Magazine, 14 Oct. 2021, <https://www.forbes.com/sites/forbestechcouncil/2021/10/14/flying-blind-how-bad-data-undermines-business/?sh=9c25cb929e8a>.

Redman, Thomas C. "Bad Data Costs the U.S. \$3 Trillion per Year." *Harvard Business Review*, 4 Oct. 2017, <https://hbr.org/2016/09/bad-data-costs-the-u-s-3-trillion-per-year#:~:text=So%20consider%3A,data%20they%20don't%20trust>.

Zavetz, Lisa. "What Is the Cost of Poor Quality Data?" *Twilio Segment* Twilio Segment Ranked #1 CDP for Worldwide Market Share (IDC 2020), Segment, 13 Oct. 2022, <https://segment.com/blog/cost-of-poor-quality-data/>.

Contact Info

Contact

john.lang@valencedata.tech
medium.com/@johnnylang

Office

Remote, Worldwide

HQ

Nashville, TN