

What can Data Governance tools not do for you?

Data Governance tools have been in the market for more than 20 years, and throughout that time the vendors have been promising success, but the reality is that there are very few successful Data Governance implementations.

To a large extent, this is not the fault of the technology. Data Governance is fundamentally a people and process activity, with only support and automation provided by the technology. However, the capabilities of the software do have an impact on project success, and this is the area I will focus on in this article.

The question that I'm addressing in this article is "which features of data governance software are mature and reliable, which are in development, and which do not work". I'll be addressing this in general, not for a specific tool. This is both to avoid lawsuits as well as to reduce the article to a manageable reading time.

1. **Feature One - Data Quality:** The market for Data Quality software continues to evolve, but fundamentally, this is a capability that is mature and reliable and can be implemented. I'm not saying it's easy, but I am saying that there are mature tools that provide effective support.
2. **Feature Two - Business Glossary:** The market is maturing for Glossary tools. The basic capability of storing definitions, managing approval workflows, and supporting features is mature, but there is still development to go in surfacing these definitions into core operational and reporting systems. This market is still evolving.
3. **Feature Three - Data Catalogue:** The basic activity of automatically identifying data assets is reasonably mature, as is the capability to store and serve this data to users. However, implementing this in a complex heterogeneous system environment is much more complex. Organisations with small homogeneous modern application inventories may be able to automatically scan their full landscape, but most mid to large size organisations will find that there is no available software that can catalogue all systems automatically and manual uplift will be required.
4. **Feature Four - Data Lineage:** Similar to the Catalogue issue, the basic capability for automated data lineage is well established, but the complexities here are greater. The data movement that lineage is describing can be achieved in many ways, including specific ETL tools, procedural language code, in-database code such as macros, APIs, application-specific code, and other methods. Data Lineage tools are not currently capable to cover all these different movement types, with the greatest capability in ETL tool lineage and in-database code lineage, with least capability in APIs, procedure language code and application-specific code.

What can Data Governance tools not do for you?

Even in the areas of strength, there are technical limits. Data Lineage is in the "In Development" area in my view.

I hope this analysis has helped you better understand where Data Governance tools are capable and where you need to supplement them with manual effort and home-grown solutions.

I welcome feedback from DG tool vendors and others on this high-level analysis.

If your business needs help with implementing the right Data Governance tools, please contact a Fujitsu Data & AI specialist now.

Contact

Fujitsu Data & AI
+61 3 9924 3000

© Fujitsu 2022. All rights reserved. Fujitsu and Fujitsu logo are trademarks of Fujitsu Limited registered in many jurisdictions worldwide. Other product, service and company names mentioned herein may be trademarks of Fujitsu or other companies. This document is current as of the initial date of publication and subject to be changed by Fujitsu without notice. This material is provided for information purposes only and Fujitsu assumes no liability related to its use.