

GIGAOM

REPORT

Microsoft Power BI

An Evolution in Data Governance

ANDREW J. BRUST

SPONSORED BY  Microsoft

Microsoft Power BI

An Evolution in Data Governance

TABLE OF CONTENTS

- 1 Summary
- 2 Governance Explained
- 3 The Governance Flywheel
- 4 Why Data Governance Can Be Hard
- 5 Making Governance Easier
- 6 How Power BI Can Help
- 7 Why the Power BI Approach is Unique
- 8 About Andrew Brust
- 9 About GigaOm
- 10 Copyright

1. Summary

In an age of analytics and digital transformation, curating, protecting, and managing our data is critical. But still, many of us may see data governance as a necessary burden, best delegated to specialists, who we suspect may become adversaries to our productivity.

But analysis and governance can work hand-in-hand, rather than at cross-purposes. While such harmony may sound elusive, achieving it is a question of changing approach. Data governance can be embedded into the self-service BI workflow, causing it to amplify our analytics work rather than impeding it; to be in concert with our efforts rather than distracting us from them. While the industry's track record in achieving such governance equilibrium has not been great, Microsoft's Power BI has recently added low-touch governance features that put us well along the road to that goal.

2. Governance Explained

While data governance has many facets to it, arguably the most important are curation, data provenance and, of course, data protection.

- **Curation** helps us organize and taxonomize our data so that, within the vast sea of datasets an organization has collected, those of particular interest can stand out, be more easily discovered and, where appropriate, verified as authoritative and accurate.
- **Data provenance** helps us track where our data came from and, in certain cases, what transformation steps it has undergone.
- **Data protection** ensures that different parties see only the data that is relevant to them and for which they are authorized, by defining rules at the source. These centralized rules ensure that all derivative analyses, visualizations, reports, and dashboards will comply with protection constraints, all without requiring explicit, repeated efforts by the authors of those assets to enforce them.

Data governance is important for a few key reasons:

- **Safety and privacy.** We need to make certain that sensitive data has its access and analysis restricted. In some cases, only the person to whom the data pertains should be able to see it. In other cases, other parties can see it as well, but only if they are authorized. Some sensitive data should be obfuscated, some should be analyzed only in aggregate, and some should not be analyzed at all.
- **Discoverability.** Data does us no good if awareness of it is limited, or the meaning of its contents is opaque or obscure. Therefore, It is important to curate the data, document it, and enable analysts to socialize and evangelize the data they believe will be useful across the organization. In general, data needs to be organized and accessible.

- **Trustworthiness.** If the source of data, as well as the steps that are taken to clean it and enhance it, and its subsequent usage are all well-known, confidence in the efficacy of the data increases. When trusted experts can certify the data explicitly, trust is even further enhanced. This engenders data culture and helps it flourish, which is critical to digital transformation.

3. The Governance Flywheel

Data governance is not a purely protective measure; it also provides positive benefits. Data governance, if implemented well, can enhance the insights derived from analytics and the effectiveness of applying it. Taken further, governance can make analytics more enjoyable, more broadly adopted, and more fully transformative to business culture.

The enhanced quality, safety, endorsement, and accessibility of data and data insights made possible by data governance creates a virtuous cycle. With a data governance regimen in place, innovative knowledge workers have a platform that provides them with support and validation. In a very real sense, governance provides rewards and encouragement to invest further and adopt good practices.

The resulting enhanced governance efforts perpetuate these rewards further, providing additional positive reinforcement. The collaborative nature of the work brings colleagues on board, not just as avid consumers of generated insights but with incentives and encouragement to contribute on their own, with an expectation of acknowledgment and success, both in morale and career development.

4. Why Data Governance Can Be Hard

A big challenge with data governance is that vendor solutions have structured it as a detour. In order to govern our data we have had to put aside our analysis; move over to a different security or data catalog platform, or module; and dedicate time to inventorying, organizing, and securing our data. Some people like that, but many do not. For the latter group, such an effort can be a distraction. It can be a time suck. Conventional governance work can cause them to lose their train of thought. Put more simply, it adds a lot of friction.

Conversely, some detours beyond the BI platform are actually typical of the BI workflow, but it can be difficult to extend the jurisdiction of governance implementations to include them. Specifically, sharing insights is sometimes done with documents, spreadsheets, and slides often sent as email attachments. How can a governance effort work if its scope omits such dissemination of insights?

Such concerns about governance can be addressed and resolved. The solution comes down to implementation, specifically around aligning governance with the analytic workflow, rather than impeding it, or underestimating its scope.

5. Making Governance Easier

What if we could curate our data concurrently while sourcing it and analyzing it? What if we could ensure the protection of data even as we disseminate it through documents and email? And what if we could determine lineage as we source the data, and have its further derivations and manipulation documented silently, as our analysis work takes place? Such capabilities would establish a beachhead for data governance that works for everyone in the BI lifecycle.

The [Microsoft Power BI platform](#) has recently added features that, while modest in complexity and minimally-invasive of workflow and productivity, offer an array of implicit governance capabilities unprecedented in both the self-service and enterprise business intelligence arenas.

6. How Power BI Can Help

Power BI's approach to governance differs from that of the conventional, monolithic governance layers that the industry has delivered to date. Instead of making BI authors and consumers detour to a separate platform, Power BI governance features surface ambiently in the normal analytics workflow. The features are not over-engineered or complex. As we will elaborate below, these features dovetail nicely into tasks both casual- and power-users perform on the Power BI platform and, as such, they fit its workflow, intuitively.

Endorsed Datasets

Take curation for example. When Power BI dataset owners are logged into the cloud service and browsing a given workspace, they can now easily indicate which of their datasets they believe will be useful and important to colleagues. With a few clicks on any dataset, its owner can draw attention to it by marking it as "Promoted." Later, when users select the "Get Data" option in Power BI Desktop then connect to Power BI datasets, they will see a listing of all datasets in the workspace, with Promoted ones listed first.

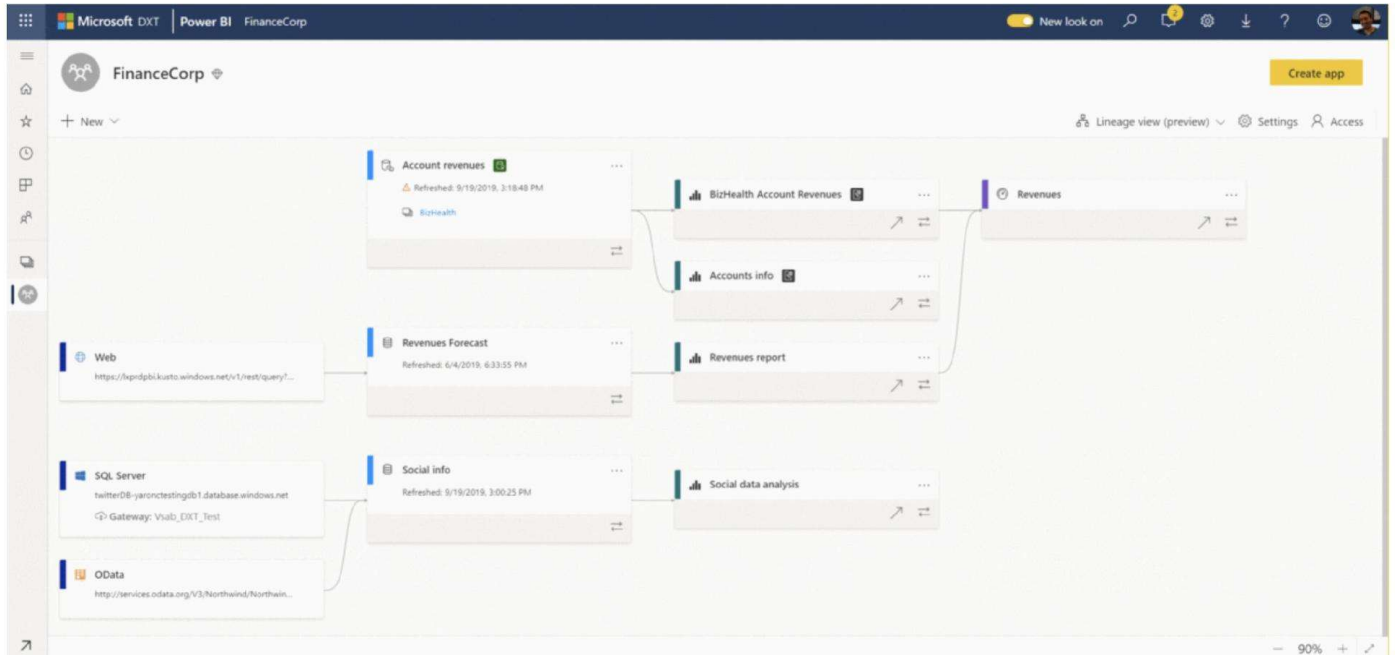
In addition, users authorized by the Power BI tenant administrator can mark datasets as "Certified" – and customers can create their own internal procedures for other users to request such certification. This provides a more formal level of endorsement for organizations that require, or desire, additional stewardship. The Power BI Desktop propagates that hierarchy, listing Certified datasets even before Promoted ones in its Get Data experience.

Lineage and Impact Analysis

Another new feature set involves workspace and cross-workspace data lineage. Arguably, these are even simpler to use than dataset endorsement. When browsing objects in the workspace, users need only switch from the standard List View to the new Lineage View (shown in Figure 1) in order to see data sources and gateways, dataflows, datasets, reports, and dashboards in a network diagram

visualization that makes the upstream and downstream relationships between these data assets explicit.

FIGURE 1. THE POWER BI WORKSPACE LINEAGE VIEW



Rather than requiring the use of a specialized platform and set of tasks, Power BI simply conveys lineage information by visualizing it among the various workspace assets with which users are already familiar. Shared datasets or linked dataflows being used within the examined workspace will also appear as part of the lineage information.

When users make changes to datasets, it can affect reports, dashboards, and other assets downstream. Visualizing lineage, to help users gain a general sense of how objects are related, is good. But giving them an explicit indication of which assets could be affected by changes to others is critical to governance. Power BI now provides this for datasets used within a single workspace, thus supporting BI authors and enhancing buy-in from BI consumers.

With reusable assets, such as shared datasets, the Power BI team plans to introduce a dedicated impact analysis experience that will expose all affected assets across the various workspaces.

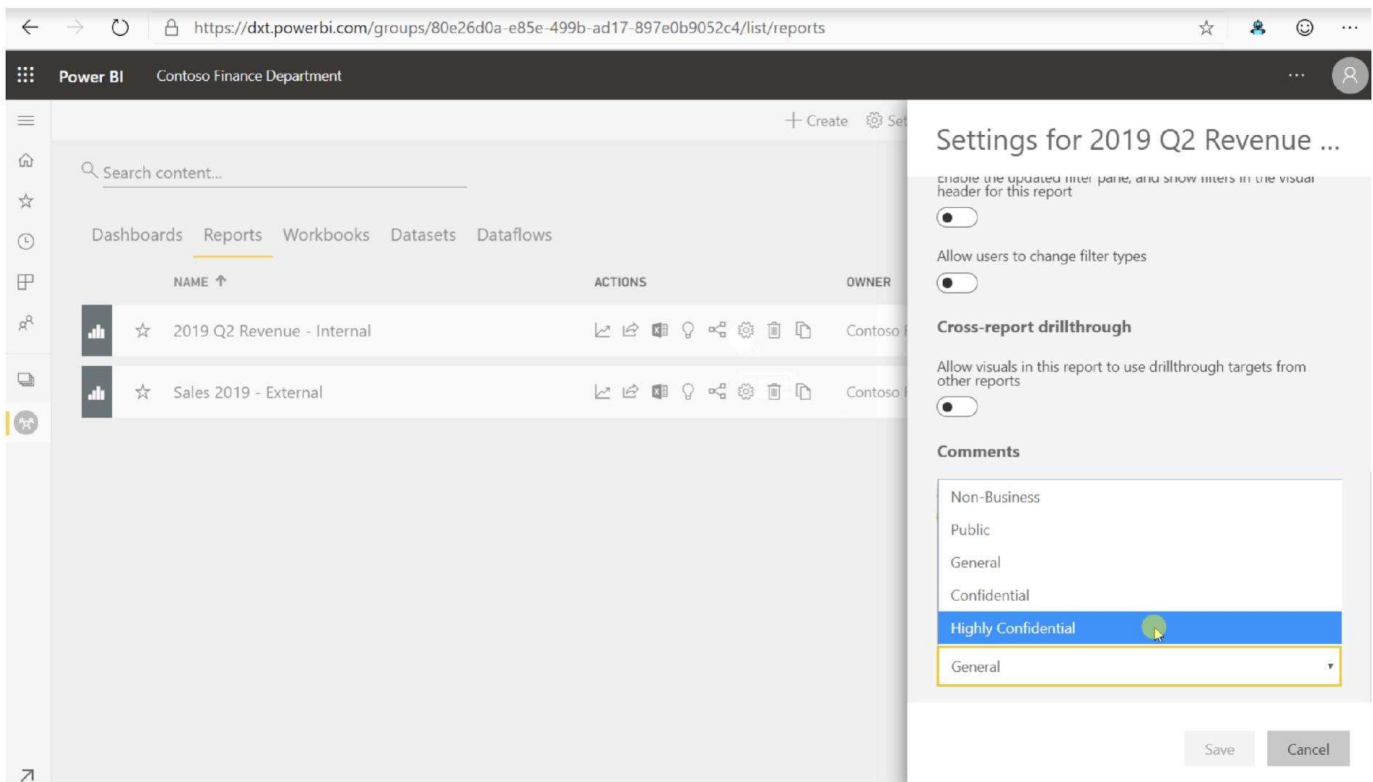
Data Protection

On top of all these governance features, however, lies what is perhaps the crown jewel: data protection. Leveraging Microsoft's existing security products and technologies, Power BI data

protection allows data owners to set [Microsoft Information Protection](#) sensitivity labels for dashboards, reports, datasets, and dataflows. Sensitivity labels are defined by security administrators in the [Microsoft Security and Compliance Center](#) and then applied to all Office 365 apps across the organization. Now, Power BI can leverage the same sensitivity labels and align with the organizational protection policy (as shown in Figure 2).

With policies in place, any Power BI data exported to Excel, PowerPoint, or a PDF file will carry that sensitivity label with it, and only authorized, authenticated users will be able to access the exported data. This is true whether the document is placed in a cloud storage service like OneDrive or OneDrive for Business, saved on a corporate network share, or even sent as an email attachment. All data protection policies set in Power BI will be propagated all the way to individual users' desktops, even when opening a previously downloaded email attachment offline.

FIGURE 2. ASSIGNING A SENSITIVITY LEVEL TO A REPORT



In addition, because Power BI data protection is also integrated with [Microsoft Cloud App Security](#), such downloads will be blocked altogether for users on unmanaged devices. Administrators can leverage Microsoft Cloud App Security to closely monitor and control risky sessions in Power BI. For example, access from an unmanaged device or uncommon location, or a suspicious sharing of a Power BI report.

For authorized users who can view the exports, the corresponding Office application or the Adobe PDF Reader will display an appropriate sensitivity alert at the top of the application. Microsoft Cloud App Security features a Power BI-specific screen that displays individual alerts and summarized analytics. In the near future, the Power BI Admin console will offer a specialized protection metrics screen.

And more

Power BI has other features that support and engender good data governance, even at the data preparation layer. Power Query's column data quality, distribution, and profiling capabilities help data engineers to understand their data better and increase its quality and reliability. These features also ensure data transformation work is carried out transparently since each operation is recorded as an inspectable "step" in Power Query and codified as a line of code in the M programming language for data modeling and transformation. Power BI row-level security, a long-standing feature, supports data governance as well, of course.

Finally, even though its governance features are embedded in the analytics workflow, Power BI is still supportive of more formal data governance efforts and initiatives. That is why Power BI's governance features can integrate into the [Azure Data Catalog](#) governance platform, which is geared more towards data stewards and IT actors. This integration provides different constituencies with governance environments appropriate to their distinct needs, and yet also coordinates and integrates their governance efforts.

7. Why the Power BI Approach is Unique

While data governance tools and platforms have been around for decades, they have required a detour from the analytics workflow. As such, they have worked well for people focused on governance, but have required information workers focused on data insights to digress from their normal pathways, change contexts, and distract themselves from their primary data analysis mission. Encouraging participation in data governance initiatives (and ongoing good governance practices) under such circumstances can be an uphill battle that can, in turn, thwart the success of an enterprise data governance initiative.

Conventional governance systems have also disregarded some core realities. To begin with, BI users are almost always Excel users too. Sharing insights from Excel, or indeed via PDF documents or PowerPoint presentations, are mainstream scenarios. Restricting those documents to authorized team members and revoking their access should they leave their job or change roles, is essential too. And yet conventional governance platforms have not met these requirements.

The Power BI approach is different, for a couple of reasons. First, governance features surface right inside the data acquisition and preparation experiences, management interfaces, and end-user delivery experiences. This includes email-attached documents used by BI authors and consumers on a day-to-day basis. As a result, no detour is required, information workers' trains of thought are maintained and governance becomes everyone's business.

But embedding governance features is just part of the story. The other salient facet of governance features in Power BI is their vetted and triaged quality. Rather than impose a high-impact feature set around governance, the Power BI approach instead focuses on essentials.

Everyone needs to be on the governance train, but different audiences have varying needs. Showering information workers with an array of features designed for data stewards will likely overwhelm them, turning them off to the very data governance practices we want them to champion. By introducing minimally-intrusive, grassroots governance features, Power BI is bringing information workers along for the governance ride. It is a sensitive, role-based approach that maximizes participation and pay-off, while largely eliminating detour, distraction, and disincentive.

8. About Andrew Brust



Andrew has held developer, CTO, analyst, research director and market strategist positions at organizations ranging from the City of New York and Cap Gemini to Gigaom and Datameer. He has worked with small, medium and Fortune 1000 clients in numerous industries and with software companies ranging from small ISVs to large clients like Microsoft. Andrew's resulting understanding of technology, and the way customers use it, makes his market and product analyses relevant, credible and empathetic.

Andrew has tracked the Big Data and Analytics industry since its inception, as Gigaom's Research Director and ZDNet's lead blogger for Big Data and Analytics. Andrew co-chairs Visual Studio Live!, one of the nation's longest running developer conferences. As a longtime technical author and speaker in the database field, Andrew understands today's market in the context of its longtime Enterprise underpinnings.

9. About GigaOm

GigaOm provides technical, operational, and business advice for IT's strategic digital enterprise and business initiatives. Enterprise business leaders, CIOs, and technology organizations partner with GigaOm for practical, actionable, strategic, and visionary advice for modernizing and transforming their business. GigaOm's advice empowers enterprises to successfully compete in an increasingly complicated business atmosphere that requires a solid understanding of constantly changing customer demands.

GigaOm works directly with enterprises both inside and outside of the IT organization to apply proven research and methodologies designed to avoid pitfalls and roadblocks while balancing risk and innovation. Research methodologies include but are not limited to adoption and benchmarking surveys, use cases, interviews, ROI/TCO, market landscapes, strategic trends, and technical benchmarks. Our analysts possess 20+ years of experience advising a spectrum of clients from early adopters to mainstream enterprises.

GigaOm's perspective is that of the unbiased enterprise practitioner. Through this perspective, GigaOm connects with engaged and loyal subscribers on a deep and meaningful level.

10. Copyright

© [Knowingly, Inc.](#), 2019. "Microsoft Power BI" is a trademark of [Knowingly, Inc.](#). For permission to reproduce this report, please contact sales@gigaom.com.