onebridge

Citizen Enablement

Creating and Executing a Data Strategy for Self-Service Analytics

www.Onebridge.tech

Introduction

The way we approach data asset development and data governance has changed drastically thanks to the rise in self-service analytics. In the past, businesses relied heavily on IT to produce the data assets needed. The creation and maintenance of these reports might take days, weeks, or more. By the time an asset was finished, it would be time to update it again.

Now that businesses have a greater demand for data, this model is transitioning. Users don't want to look at a dashboard or a static report anymore. They want access to the data directly with more dynamic options for viewing and analyzing it, driving a shift toward citizen-enabled analytics. The newest challenge organizations face is how to lean into this shift while keeping data assets compliant with IT rules of governance.

Once referred to as "Shadow IT," a common symptom of this challenge is the rapid creation of data assets, sources, or processes by non-IT resources without IT's knowledge or oversight. This is now more frequently referred to as the "democratization" of data asset development. As business users try to keep up with the demand for data-driven decision-making and analytics platforms become more powerful, flexible, and easier to use, the development of unsanctioned assets becomes more prevalent.

To most effectively and responsibly enable citizen data analytics, you must strike the right balance between the democratization of data assets and appropriate data governance and oversight.

Contents

New Self-Service Tools Are Granting Powerful New Abilities to Non-Technical Business Users – The Good and The Bad

When it comes to the creation and preparation of unofficial but trusted data sources Excel is, without a doubt, the most significant enabler. The expansion of more advanced self-service tools for business intelligence has also played a major role in the downside to data democratization.

The upside of these tools is that they reduce the demand on IT for basic support, they allow users to personalize how they access and view data, and many offer rich visualizations that enable non-technical users to see their data in new ways, empowering them to discover valuable insights on their own. Users with basic knowledge can flatten XML, prep geo-spatial custom polygons, blend, cleanse, and more without IT help. More advanced users can create predictive models like regression, Naïve Bayes classifiers, and neural networks. In other words, these platforms are leading organizations to build their data literacy and develop a citizen-enabled mindset toward data.

The downside is that they also make it easier than ever for non-technical users to create their own data assets without applying IT-defined rules for governance. For example, four different departments could be relying on four different data assets to answer the same question, each trusting that their own is the best source of information over the others, sometimes even unaware that other similar assets exist in the organization.

How to Make the Most of Rampant Unsanctioned Asset Development

Given these new tools, you may find yourself out of balance with more unsanctioned asset development than desired, leading to a "wild west" situation. So, what then?

5 Steps to Changing Your Data Mindset

1. The first step is admitting you have a problem.
2. Make a thorough and honest examination of how you handle data management and pinpoint where unsanctioned development is happening. Start by looking at who has access to some of the newer tools.
3. Then, ask these questions: How is the business using these tools? What information is the business seeking? What assets have they developed that could benefit other groups in the company?
4. It's important to remember that users who have created one-off assets had a legitimate need for those assets. Avoid criticizing them for doing so.
5. Keeping in mind the information that the business needs, evaluate processes, information requirements, tooling, infrastructure, and architecture to create a self-service system that can keep pace with the business.

Develop Your Data Demand Management Plan

Once you have an idea of what a useful citizen-enabled analytics ecosystem looks like for your business, you need to develop a reliable process for users to get their information needs fulfilled without having to resort to creating their own one-off sources. This is critical for supporting the continued use of your governed self-service system and guarding against the return of unsanctioned assets.

It's a good idea to get familiar with the lifecycle of a non-sanctioned data asset. Remember, users can only create new data assets from others they already have access to. It's also becoming more common for users to integrate these existing internal assets with external items like social, weather, or market data.

Here's what to look at to map out your data demand management plan:

- **Intake** – How will you intake new requests for information? Once received, how will they be grouped, consolidated, prioritized, and reviewed?
- **Design** – How will new datasets be reconciled against existing ones? How will common dimensions be conformed? How does the architecture of consumption affect the similitude of data sets?
- **Curation** – Based on the customer's needs, how will you cleanse and groom the data? Do you need to create categories or levels to certify the data's readiness for use?
- **Output** – How will you deliver data? There are several possibilities, ranging from a modern data marketplace that allows users to "shop" for data to a visualization tool-compatible semantic layer. What's best for your business?
- **Understanding** – How will you manage and make metadata available? Don't forget to create a business glossary and define how it will be managed over time.
- **Access** – What is your access and control structure for users to retrieve assets? Will you need row- or column-level security? If so, how will you use them?

How Successful Data Governance Works

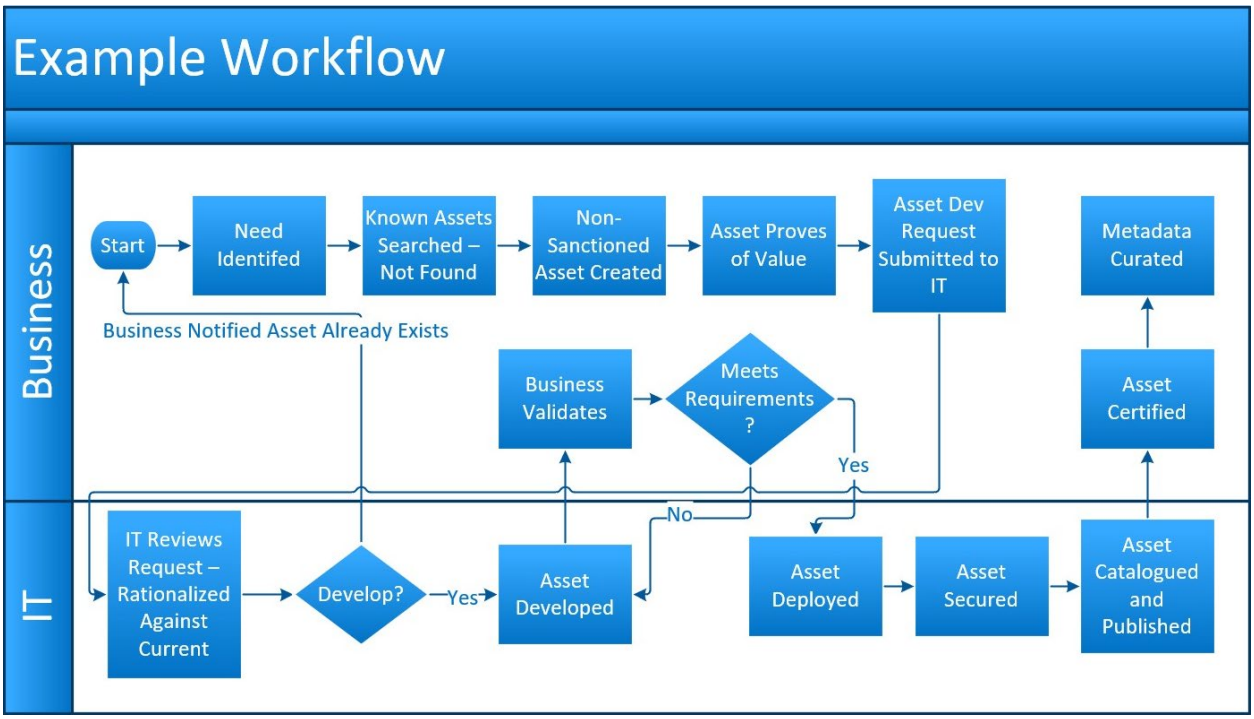
Governance has become such a common refrain it can start to feel like a vague buzzword. When we talk about data governance, we're talking about business users who know the data best serving as true owners around what qualifies as trusted data based on the source, lineage, and defined data quality rules. While the business owns the data and defines the quality rules, IT plays a vital role in serving as data custodians to conduct data quality remediation work. Sanctioning assets then becomes a joint responsibility to ensure exposed data sources meet those business-driven governance rules.

A mature development plan for data assets is the best way to govern assets without stifling the value users can get and the freedom they have to work with the assets and discover new insights. The way to get to a mature development plan is through workflows that define how requests for data assets are processed, prioritized, reviewed, and approved/denied. The workflow also defines how existing unsanctioned assets are submitted to be used as governed assets throughout the organization.

Workflows prevent data democratization and self-service from spinning out of control and proliferating data assets that don't comply with the rules of governance required by IT. At the outset, you should begin with a workflow for evaluating the backlog of existing data assets and determining whether they will be integrated into IT's set of governed assets.

Here's how to get started:

- Develop a gated process for intake.
- Not all submitted assets will qualify – you'll have to validate the data used for each, examine the data's lineage, and examine any manipulation done during its creation.
- Then, assess the usefulness of the data alongside the needs of the business to decide whether you'll share it out with other groups in the enterprise.
- Finally, apply any access or security requirements to the asset before sharing it out. Some models or dashboards may need to be gated and available only to certain groups or individuals, depending on the sensitivity of the data included.



Another way to kick off the effort to meet existing data needs is to create an inventory of current one-off data sources. Remember, these one-offs have value to the business and those using them, so evaluate their usefulness fairly.

Focus on user stories and gain a thorough understanding of the key questions users need to answer. Prioritize those stories according to their value to the business. Let users give you the answers to these questions, rather than intuiting them on your own.

Setting fixed criteria to evaluate stories objectively is a best practice. Some will have a more substantial and immediate impact than others. Help them rise to the top of the backlog list.

Add current IT-generated assets to the list as well, noting how much it costs to maintain them. If an asset isn't valuable but is costing money to maintain, it's a liability. This exercise can help your IT team focus on its highest-value assets and get a clearer idea of what kind of assets the business values.

The Approach to Practical and Right-Sized Data Quality

Sometimes building out a visual is just the beginning, as it may alert you to issues with the quality of your data. The following steps are quite common in evaluating and managing the quality of visualized data:

1. Create a first-draft dashboard
2. Share with your team
3. Acknowledge any quality issues
4. Making decisions based on bad data is not an option, so you'll need to create a plan for data quality remediation
5. Based on the importance of the data and impact on decision-making if it isn't trustworthy, decide what level of data quality management and data governance is needed and implement it
6. Make any required changes to the underlying model to correct quality issues
7. Iteratively clean the data, update the model and refine the dashboard

The Culture Change Behind Agile Data Governance

As you begin rolling out this first workflow and new processes to support proper data governance, you'll be asking the business to make a culture shift away from unsanctioned asset development and toward open communication with the IT team. This can be a difficult shift as former participants in creating unsanctioned resources will struggle with the need to acknowledge and change the role they play in data governance.

Users will likely want to support IT efforts to gain control over data assets but will want to avoid censure as creators of ungoverned assets. Clear and, most importantly, non-judgmental communication can ease the transition to a new and more effective way of doing things.

As trust in the system and adoption of new requirements rises, many reports that previously took hours to produce and reconcile each month will begin to disappear from the IT backlog. Everyone will have access to the same reliable and trustworthy data source, asking more important questions and focusing on getting the answers rather than spending most of their time prepping or manipulating the data.

Delivering constant incremental value keeps your analytics program thriving and your teams engaged. You may have noticed that getting data and assets to users faster requires some intrinsically Agile-like concepts: user stories as the driver for development, maintenance of a prioritized backlog, iteration, review, remediation, demonstration, and repetition.

Conclusion

To best enable citizen data consumption and analytics, you must take a proactive and structured approach to the governance of citizen-enabled data. Workflows and a strong data demand management plan are the most powerful tools to ensure the responsible development and management of valuable data assets. Once these safeguards are in place, rapid data asset development becomes an advantage to the business, empowering users to become more data literate and discover unexpected new insights, and easing the burden on IT.