Governed Data Discovery
Best Practices

The emergence of data discovery

Once you have been tarred with a brush, it's a difficult reputation to remove. Tar is sticky.

Traditional BI environments have been tarred. Typically, they've been all about IT. And they've been slow to produce the types of reporting business users want. There was no access to data except through IT, and when you finally received the report you wanted, it was probably too late anyway. So for one-off, quick data deliveries, traditional BI has been seen as pretty useless.

According to TDWI's Best Practice Benchmark Report *Organizational and Performance Metrics for Business Intelligence Teams 2015[1]*, it was taking an average of 5.8 weeks for IT to create a complex report or dashboard. And for almost a third of organizations, they needed nine weeks or more to add a new data source to an existing business system. That's nearly a whole quarter to get one new data source!

So, along came self-service data discovery. It handed over the reins of fast data insights to the business user. The idea? Less IT involvement for fewer delays. As a business user, you can get the information you want when you want it and you can determine the data patterns yourself.

This sounds great. But what about data security and accuracy?

Data discovery without governance

"Through 2017, less than 40% of self-service business intelligence initiatives will be governed sufficiently to prevent inconsistencies that adversely affect the business."
– Gartner, as cited by Dataversity, 2017[2]

Let's put that statistic another way: over 60 percent of self-service BI initiatives are insufficiently secure and trustworthy. Inconsistencies in the data will lead to bad business decisions. And bad decisions lead to business failure.

If lack of data discovery governance can mislead businesses in their strategic decisions, this is a major issue that needs to be tackled.

Most pure-play data discovery tools are in-memory, desktop-based applications sold at a departmental level. While they deliver more freedom to end-users, they do not meet the security and governance needs that enterprise IT demands. This creates a trust issue. If data is handled while ungoverned, the data cannot be trusted. If the data cannot be trusted, nobody will use it. Then you end up with an organization that is back in the same position as it was before BI was introduced.

Ungoverned data discovery creates these issues:

IT Issues	• Analytical silos with no single source of the truth • Lack of security and transparency • Difficult to maintain
Business Issues	• Business gets a subject-oriented view of information • Conflicting reports create distrust in data • No re-use of data models, business logic or metrics
Strategic Issues	• Difficult to deploy enterprise-wide • Long-term analytical maturity of business is affected • Risk making key business decisions on incorrect data

Governed Data Discovery Best Practices

Governed Data Discovery environments meet the dual demands of enterprise IT, including analysts, and business end-users. End users are allowed the freedom and agility to access and analyze data, while IT has transparence and control over security and governance.

How do you achieve governed data discovery? Here are best practices for implementing data discovery governance.

01 Balance (It's not just about data discovery)

Different users have different needs

Each organization has different types of users including analysts or power-users, IT administrators, and business users who consume BI content. Self-service data discovery features are helping analysts create content quickly. However, not everyone in the organization has the desire or knowledge to create content from scratch.

More BI consumers than creators

In any organization, there are more business users who consume BI content than there are analysts who create that content.

90% of users consume content

10% of users create content
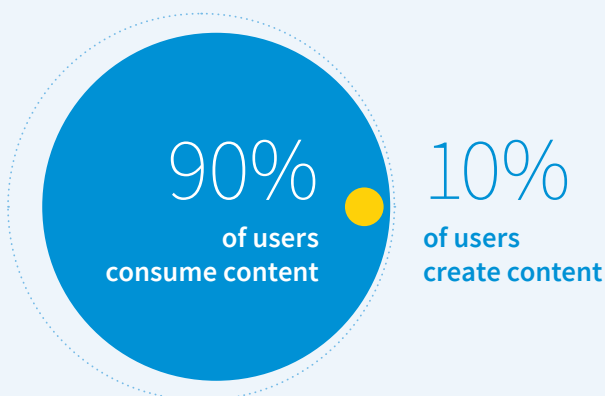
Analysts want data discovery

The agility of self-service data discovery features improves analyst productivity. Analysts want the freedom to quickly connect and visually analyze data without relying on IT to prepare the data for them first.

Business users prefer dashboards

Looking at statistics of self-service BI adoption, there is no growth. Why is this? According to a TDWI report[3], business users prefer to interact with data via dashboards rather than self-service data discovery.

Dashboards are intuitive for non-technical business users to monitor and interact with data. Interactive drills and filters give them the ability to query the data without needing to create reports from scratch.

It department needs governance

IT administrators need to manage the security and governance of Enterprise BI. What matters to IT administrators is:

- Enterprise security
- Trust in the data
- Auditability and transparency
- Controlling users' permissions
- A semantic layer to ensure consistency
- A platform to share and collaborate
- Scalability

For business users to trust the data, it must be governed.

02 Make Sure it's Secure

Governed data discovery is about ensuring your BI platform is secure and the data is trustworthy. An insecure BI environment exposes your business to risk. You don't want sensitive data, like payroll data, personal medical data and R&D data, exposed or changed by users.


Points of security

A governed environment needs security at the database layer (spreadsheets are a security risk), the data movement layer (ETL) and the network layer.

It also requires security at the BI platform layer otherwise people could access unsecured data, or reports they should not see.

BI 🔑 PLATFORM 🔑 DATABASE 🔑 ETL

A governed BI environment should include the following security types:
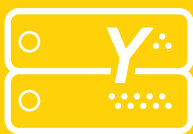
DATA LEVEL SECURITY

Control who can access what data. Secure your data at the column or row level to prevent users from seeing forbidden information.

CONTENT LEVEL SECURITY

Restrict users without permission from seeing reports and dashboards. Apply content security at an individual level, user role level, departmental level or geographical level.

FUNCTIONAL LEVEL SECURITY

Control what functional permissions different users have access to. Give some users admin permissions, give content creators analyst permission, and give business users permissions to consume only.

03 Deploy with Centralized Governance

What is the secret-sauce to governed data discovery? The answer is to deploy on a centralized architecture with governance features baked in. A centralized BI platform gives analysts all the data discovery functionality they want, and IT administrators the governance they need.

This means avoid desktop bi tools

It is hard to govern content born on a desktop. There are complicated work-arounds. However, these are expensive to maintain, time consuming to implement, and not foolproof.

100% web-based access

Web-based access is the most pervading and is easier to deploy and manage than desktop applications. Deploy on a centralized architecture and give all users access via the browser.

Any platform, any device, anywhere

Web-based access gives business users access to consume information anytime, on any platform on any device.

Agility for analysts

Give analysts the power to create reports, build dashboards and tell stories quickly and easily.

Enterprise governance for IT

Give your IT Department complete control over security, transparency and governance.

Trust for BI consumers

Give BI consumers what they most want – trust in the data – by providing governed, secure content they know they can rely on for decision making. Web-based access gives business users a central source of truth that can be accessed on any platform on any device. Enable users to share, collaborate and make decisions with data securely.

DATA ANALYSTS

Data Discovery
Build Dashboards
Storyboards

BI CONSUMERS

Dashboards / KPIs
Collaboration
Decision Making

Yellowfin

100% Web-based
Centralized

IT DEPARTMENT

Security
Metadata Layer
Governance
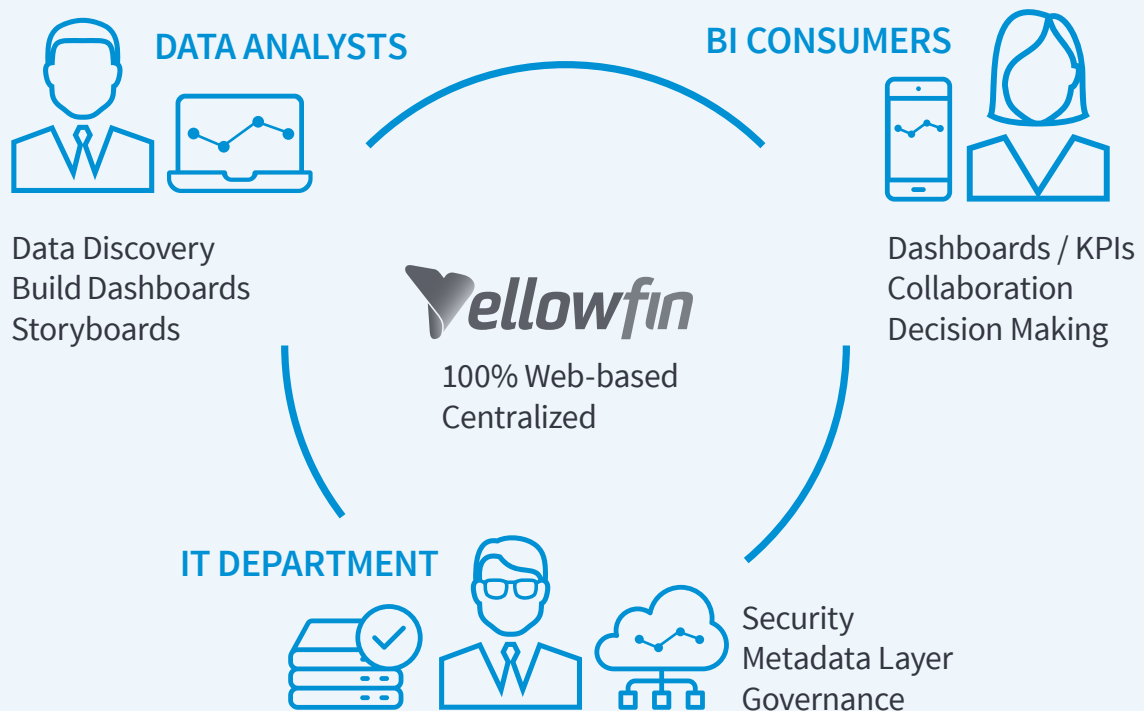
04 Ensure Audibility

Auditability features are essential for governed data discovery as it gives the business important transparency into how the BI platform is being used.

Track and monitor usage statistics

Usage statistics give you insight into who is using the system, how they are using the system, and what content is being used.

Usage statistics make it easy for administrators quickly identify issues and plan for future usage.
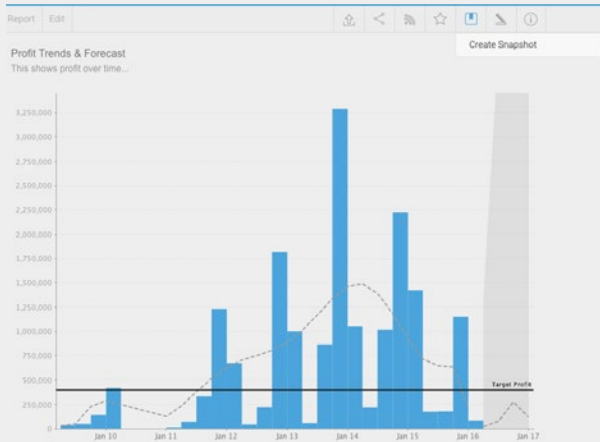
Data snapshots

Audit how you make decisions. Data snapshots let you evaluate what the data looked like at different points in time so you can evaluate your decision making process.

Approval workflows

Approval workflows ensure data and reports are checked and approved before being published. Analysis is only as good as the data.

Approval workflows are ideal for important documents such as financial statements, as they prevent users who have access to sensitive information from doctoring the data and misleading people.

Make you get important reports approved first.

05 Reuse Business Logic

The shortcoming of most pure-play data discovery tools is they do not have a meta-data layer that allows reuse of business logic. The metadata layer is essential for deploying BI throughout the enterprise as it ensures a consistent single source of the truth for business. Having data preparation baked into your data discovery tool at the metadata level means any changes made will be uniformly reflected across all content based on that metadata layer throughout the enterprise – from reports and charts to dashboards.

Build reusable views

The metadata layer lets you build a logical model or view of your data, secure it at the data level, and ensure consistency in field names, calculations and filters. This view can then be used many times over.

Single source of the truth

The metadata layer enables the business to lock down business logic to ensure consistency in data and reports based on each metadata level. This provides your users with consistency and therefore trust in the data.

Making bi easier for analysts

Having the ability to create views using drag-and-drop makes it faster and simpler for analysts to deliver visualizations and analyses for business users. These views can then be used again for further visualizations and analyses. Knowing the data is secure and consistent, analysts are able to trust in the data and therefore the content they produce.

Creating views using drag-and-drop makes it faster and simpler for analysts to deliver visualizations.
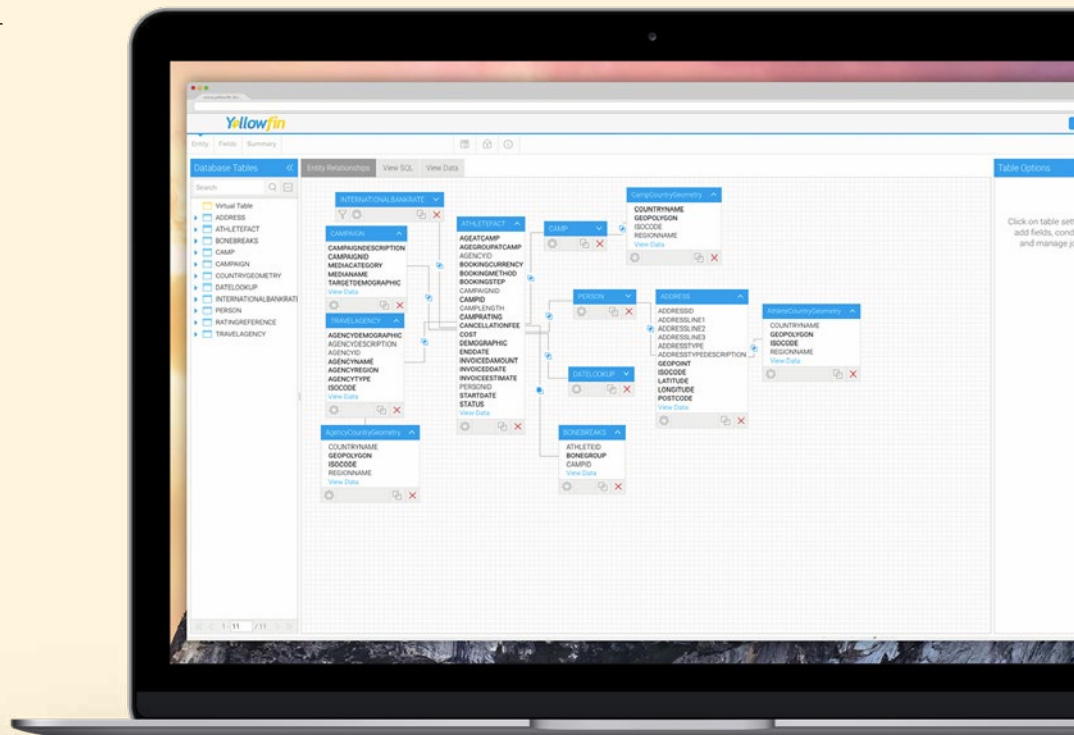
06 Make Information Access Pervasive

Data becomes more valuable when it is shared. So what prevents the enterprise from rolling reporting and analytics out to more users?

BI for everyone

Dashboard access is the easiest way for non-technical business users to consume data. If you want to establish a data-driven decision-making culture in your organization, then interactive dashboards are essential. Users need to be able to drill down to the data to see the truth behind the metrics.

What else prevents the enterprise from rolling BI out to more users?
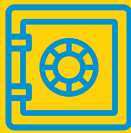
SECURITY
Is your BI platform secure?

Security becomes more important as more people are given access. That makes security at the data, content and functional levels even more important. BI platforms that make security easy for IT to govern get deployed to larger user bases.

COST
Can you afford to scale your BI platform?

Cost is one of the key drivers that prevent organizations from deploying BI throughout the enterprise. Make sure your BI tool is affordable to scale and does not lock you into long-term, upfront contracts.

07 Make Sure It's Production Ready

A governed data discovery environment must be production ready.

What should my production ready check-list include?

SCALE
Design your BI platform to easily scale from a few users to hundreds or thousands

PERFORMANCE
Design architecture for fast query performance as data volumes, users and overall usage grows

USER ROLES AND USER GROUPS
Assign users roles with different permissions such as report consumers, report writers and administrators

CONTENT CATEGORIES
Create rules for publishing access and approvals as well as content security access

TEST AND PRODUCTION ENVIRONMENTS
Set up with test and production environments before rolling out to ensure enterprise deployment is safe

BACKUP
Prepare your BI environment for high availability and disaster recovery

MIGRATION
Plan to easily migrate your BI platform to other servers

UPGRADES
Plan a strategy to upgrade your BI platform throughout the enterprise

LICENSES
Set reminders to ensure users don't have their licenses expire at the wrong time

PEOPLE
Ensure you have the right people to make your BI and analytics environment a success

Our governance means data you can trust
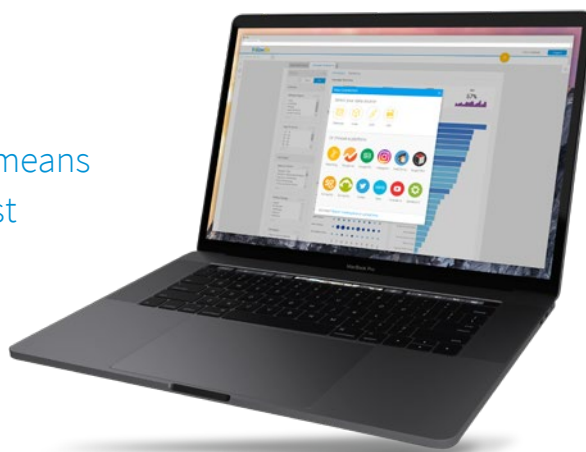
Yellowfin is a global Business Intelligence (BI) and analytics software vendor passionate about making BI easy.

Founded in 2003 in response to the complexity and costs associated with implementing and using traditional BI tools, Yellowfin is a highly intuitive 100 percent Web-based reporting and analytics solution. Yellowfin is a leader in Mobile BI, Collaborative BI and Embedded BI, as well as Location Intelligence and data visualization.

Over 10,000 organizations, and more than 2 million end-users across 70 different countries, use Yellowfin every day. For more information, visit www.yellowfinbi.com

For regular news and updates, follow Yellowfin on Twitter (@YellowfinBI), LinkedIn (Yellowfin Business Intelligence), YouTube (Yellowfin Team) or email pr@yellowfin.bi to subscribe to Yellowfin's free e-newsletter.

www.yellowfinbi.com

Endnotes

[1] TDWI (2017). TDWI BI Benchmark Report 2015. Organizational and Performance Metrics for Business Intelligence Teams. DocPlayer.net. Available at: http://docplayer.net/5881124-Tdwi-bi-benchmark-report.html [Accessed: 09 March 2017]. Pp. 15

[2] Dataversity (2017) Self-Service Business Intelligence is Big, but is it for Everyone? Zaino, J. Dataversity.net. Available at: http://www.dataversity.net/self-service-business-intelligence-big-may-not-everyone/ [Accessed: 17 March 2017]

[3] TDWI (2014) Stodder, D. TDWI Best Practices Report: Business-Driven Business Intelligence and Analytics. [pdf] file:///Users/Laura/Downloads/TDWI_BPR_BzDriven_BIA_Q314_web.pdf Retrieved from TDWI database.