

# DATA GOVERNANCE

## 1=====

According to IDC, a major percentage of data is unstructured. Data governance defines and manages all these unstructured data. It also summarizes the responsibility of data assigned to users and defines governance rules, privacy policies and security of guidelines that should be followed. Hence, Data Governance is not just about storing data, but a resilient data governance tool should provide robust access to the data, confidence in the authenticity of data and clearly define flow of the data (data lineage).

At times, it is hard for companies to maintain all the data and make decisions depending on that data. In the recent times, there have been so many legal and financial issues reported due to the poor quality of data. For an organization to flourish, it is important that the consumers and executives have confidence in them. Data governance tools make your job easier by providing clarity of the data.

The most important decisions in an organization are made based on the data available, this data is expected to be organized, reliable and consistent throughout the organization. To achieve this goal of building good quality data, companies need a good governance tool.

“53% of data professionals surveyed cite difficulty in accessing data from different areas of data silos, as key challenge” says the IDC. This happens because different metrics have been used in different areas of data. The data governance tool standardizes the metrics and establish common data standards throughout the organization. This standardization is documented well and helps leaders make right decisions that are not depended on any inconsistent and false data.

When you have a good governance tool and better data, the obvious results would be reduced data duplication, well organized data and less number of errors. These changes help the organization not only to save money, but also to generate profits. Well organized data about the customers can target the right kind of crowd and results in good marketing campaigns that generate profits. Reduced duplication and less number of errors reduces the efforts put by the resources of the company and saves time and money.

## 2=====

The installation process goes smoothly when the instructions given are followed.

However, as an end user, I feel it was unlike the single set up process where we just follow the wizard and click next and finish installing the software. Rebooting and enabling the Virtualization Technology VT-x on the BIOS took more time and effort.

## 3=====

Yes, the guided tutorial provided good understanding of that particular use case.

However, I found something unusual when I try exploring different options, the tutorial follows the same series of steps irrespective of what you click. For example: when you see the picture below, there is no term called product on the page.

The screenshot displays the IBM Infosphere Information Governance Catalog interface. At the top, the header includes the IBM logo and the text "IBM INFOSPHERE INFORMATION GOVERNANCE CATALOG". Below the header, there is a navigation bar with links: Search, Glossary, Information Assets, Labels, Queries, Collections, and a Quick Term Finder button. A search bar is located below the navigation bar, with the placeholder text "Search for Categories and Terms".

On the left side, there is a "Search Options" panel. It contains two sections: "Find these types of assets" and "Search in these properties". Under "Find these types of assets", the options "Category" and "Term" are checked. Under "Search in these properties", the option "Name" is checked. Other options include "Labels", "Stewards", and "Modified On".

The main area is titled "Search Results". It shows a list of search results. Each result includes a checkbox, a document icon, a term name, and a status. The terms listed are: "10,001 to 36,000 Units", "100,001 to 150,000 Units", "100,001 To 150,000 Units", "150,001 To 200,000 Units", "30,001 to 60,000 Area Units", "50,001 To 100,000 Units", "60,001 to 100,000 Area Units", and "Abdominal Fat Flag". Each term is associated with a "Candidate" status and an "Involved Party" descriptor.

A pink callout box is overlaid on the search results, containing the text: "Select the asset you want to explore. From the search results, you see terms as well as their category and definition. Hover over any term to view more details. Click the term 'Product' to select it."

When you see the picture below, the “assigned assets” has been disabled, so I cannot click assigned assets and expand it.

The screenshot displays the IBM InfoSphere Information Governance Catalog interface. At the top, there is a navigation bar with links for Search, Glossary, Information Assets, Labels, Queries, and Collections. The main content area shows the details for a term named "10,001 to 30,000 Area Units". The term is categorized under "Involved Party" and has a status of "Candidate". A red callout box with the text "Click 'Assigned Assets' to expand it." and a "NEXT" button is overlaid on the interface. Below the term details, there are sections for "General Information", "Associated Terms", "Assigned Assets", "In Collections", "Notes", "History", and "Graphical View". The "History" section shows a table with columns for Date, Comment, and Edited By, containing one entry from June 5, 2017.

LONG DESCRIPTION	Parent Category	Status
A member of Organization Unit floorspace Area Range. Classifies measures by Organization Units with floor space ranging between 10,001 to 30,000 area units.	Involved Party	Candidate

IS MODIFIER	TYPE
False	None

Date	Comment	Edited By
Jun 5, 2017 3:01:59 PM	Enable of history	Harold Smith

So, to avoid confusion, the tutorial should rather stop when the user does something other than what’s been asked to do. Or tutorial should gray out the options which the tutorial does not support.

#### 4=====

The other competitive solution I want to discuss about is the SAP Master Data Governance.

IBM IGC is referred as work flow oriented and SAP Master Data Governance is referred as repository oriented. [1]

In IBM IGC, as soon as the user opens the application, there is a catalog of information for the user to access the information across the organization. The accessibility depends on the workflow and security roles the user is assigned to. There are other compatible IBM software platforms that are needed to provide additional features such as InfoSphere Metadata Asset Manager and InfoSphere Business Glossary.

In SAP MDG, the companies make, maintain and track the master data from the repository. It has administrative and operational structures in addition to the data governance in the same product and hence making it easier for the companies by using a single platform. In cases where it needs a different platform, it integrates easily with other SAP platforms.

#### References:

<http://searchdatamanagement.techtarget.com/feature/What-to-know-about-the-IBM-Information-Governance-Catalog>

<http://searchdatamanagement.techtarget.com/feature/Improve-data-management-and-quality-with-SAP-Master-Data-Governance>