



Introducing Microsoft Purview

Unlocking the Power of Governance,
Compliance, and Security in the
Modern Cloud Enterprise

Charles Waghmare

Apress®

Introducing Microsoft Purview

**Unlocking the Power
of Governance, Compliance,
and Security in the Modern
Cloud Enterprise**

Charles Waghmare

Apress®

Introducing Microsoft Purview: Unlocking the Power of Governance, Compliance, and Security in the Modern Cloud Enterprise

Charles Waghmare
Matunga Mumbai, Maharashtra, India

ISBN-13 (pbk): 979-8-8688-1203-3
<https://doi.org/10.1007/979-8-8688-1204-0>

ISBN-13 (electronic): 979-8-8688-1204-0

Copyright © 2025 by Charles Waghmare

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

Trademarked names, logos, and images may appear in this book. Rather than use a trademark symbol with every occurrence of a trademarked name, logo, or image we use the names, logos, and images only in an editorial fashion and to the benefit of the trademark owner, with no intention of infringement of the trademark.

The use in this publication of trade names, trademarks, service marks, and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Managing Director, Apress Media LLC: Welmoed Spahr
Acquisitions Editor: Smriti Srivastava
Editorial Assistant: Kripa Joseph

Cover designed by eStudioCalamar

Cover image designed by Pixabay

Distributed to the book trade worldwide by Springer Science+Business Media New York, 1 New York Plaza, Suite 4600, New York, NY 10004-1562, USA. Phone 1-800-SPRINGER, fax (201) 348-4505, e-mail orders-ny@springer-sbm.com, or visit www.springeronline.com. Apress Media, LLC is a California LLC and the sole member (owner) is Springer Science + Business Media Finance Inc (SSBM Finance Inc). SSBM Finance Inc is a **Delaware** corporation.

For information on translations, please e-mail booktranslations@springernature.com; for reprint, paperback, or audio rights, please e-mail bookpermissions@springernature.com.

Apress titles may be purchased in bulk for academic, corporate, or promotional use. eBook versions and licenses are also available for most titles. For more information, reference our Print and eBook Bulk Sales web page at <http://www.apress.com/bulk-sales>.

Any source code or other supplementary material referenced by the author in this book is available to readers on GitHub. For more detailed information, please visit <https://www.apress.com/gp/services/source-code>.

If disposing of this product, please recycle the paper

***The law of the LORD is perfect, converting the soul:
the testimony of the LORD is sure, making wise
the simple.***

—Psalms 19:7

*First, I would like to say thanks to Almighty Lord Jesus Christ for offering me yet another opportunity to author this book. I completely owe everything to Him. I take this opportunity to praise and glorify Lord Christ for all the wonderful things that He has been doing in my life.
God bless.*

My Dedication:

*While approaching the second death anniversary of my dearest mother, late Mrs. Kamala David Waghmare,
I have begun writing this book.
I dedicate this book to my dearest mother and to my father, Mr. David Genu Waghmare, who both laid foundation of my life and career. Without them, I am nothing. I thank
God for the best mom and dad.*

*Also, I dedicate this book to my adorable wife,
Mrs. Priya Waghmare, for her support, love,
encouragement, and care.*

Table of Contents

About the Authorxv

About the Technical Reviewerxvii

Acknowledgmentsxix

Chapter 1: Introduction to Microsoft Purview1

 Introduction.....2

 Key Features of Microsoft Purview.....3

Comprehensive Information Discovery and Classification4

Centralized Data Cataloging.....6

 Unified Data Catalog.....6

 Data Asset Enrichment7

Data Lineage and Impact Analysis7

 End-to-End Data Lineage8

 Impact Analysis8

Robust Data Governance.....9

 Policy Management and Enforcement.....9

 Role-Based Access Control (RBAC).....10

Enhanced Compliance Management10

 Regulatory Compliance Support.....11

 Audit Trails and Logging11

Data Sharing and Collaboration12

 Internal Data Sharing.....12

 External Data Sharing.....13

TABLE OF CONTENTS

- Advanced Data Privacy and Security 14
 - Sensitive Data Protection 14
 - Data Masking and Obfuscation..... 15
- Integration with Microsoft Ecosystem 15
 - Azure Synapse Analytics Integration 16
 - Power BI Integration 16
 - Microsoft 365 Integration 17
- Data Quality Management..... 18
 - Data Quality Metrics 18
 - Data Quality Rules and Validation..... 19
 - Data Quality Monitoring 19
- Extensibility and Customization 20
 - Custom Connectors and Integrations 20
 - Customization of Governance Policies 21
- Scalability and Performance 22
 - Scalability 22
 - Performance Optimization..... 23
- An Overview of Microsoft Purview Architecture and Its Core Components 24
- Information Catalog..... 25
- Information Map..... 25
- Information Insights 26
- Information Policies 27
- Chapter 2: Introduction to Information Governance in Microsoft Purview 31**
 - Introduction..... 32
 - Information Security 33
 - Microsoft Purview Information Loss Prevention..... 33
 - Microsoft Purview Information Barriers 34

Microsoft Purview Information Protection.....35

Microsoft Purview Insider Risk Management.....40

Microsoft Purview Privileged Access Management41

Information Governance.....42

 Microsoft Purview Information Catalog43

 Microsoft Purview Data Estate Insights.....44

 Microsoft Purview Information Map46

 Microsoft Purview Information Policy.....47

 Microsoft Purview Information Sharing.....48

Information Compliance.....49

 Microsoft Purview Audit50

 Microsoft Purview Communication Compliance.....51

 Microsoft Purview Compliance Manager.....52

 Microsoft Purview Information or Data Lifecycle Management53

 Microsoft Purview eDiscovery53

Chapter 3: Deep Dive into Information or Data Catalog and Data Management55

 Introduction.....56

 Definitions58

 Exploring Data Catalog Features.....60

 Data Search60

 Data Product Search.....63

 Data Management.....65

 Business Domain.....65

 Data Products67

 Data Quality69

TABLE OF CONTENTS

Chapter 4: Information Lifecycle Management and eDiscovery79

 Introduction.....79

 Microsoft Purview Audit80

 Microsoft Purview Communication Compliance.....81

 Microsoft Purview Compliance Manager.....81

 Microsoft Purview Information or Data Lifecycle Management81

 Microsoft Purview eDiscovery81

 Information Lifecycle Management82

 Microsoft Retention Policies.....84

 Retention for SharePoint and OneDrive86

 Retention with Cloud Attachments?89

 Retention with OneNote.....91

 Retention for Microsoft Teams.....93

 Retention for Viva Engage95

 Retention for Exchange96

 Retention for Copilot.....98

 eDiscovery100

 Capabilities of eDiscovery101

Chapter 5: Data Lineage and Mapping.....105

 Introduction.....105

 Features of Data Lineage107

 Benefits of Data Lineage in Microsoft Purview108

 How Data Lineage Works in Microsoft Purview109

 Use Cases for Data Lineage in Microsoft Purview110

 Integration with Other Microsoft Purview Features111

 Challenges and Considerations.....112

 Best Practices for Implementing Data Lineage in Microsoft Purview113

Data Lineage Examples in Microsoft Purview.....	114
Example 1: A Data Pipeline Is Created in a Cloud-Based Data Warehouse. It Shows How Information Moves Across Different Systems.....	115
Example 2: The Second Example Deals with the End-to-End Process Involved in Regulatory Compliance	116
Example 3: Data Lineage for a Marketing Campaign	117
Example of Lineage from Power BI into Microsoft Purview	118
Introduction to Data Mapping in Microsoft Purview.....	121
Understanding Data Mapping	121
Features of Data Mapping in Microsoft Purview	122
View Data Map.....	123
Benefits of Data Mapping in Microsoft Purview.....	126
Practical Applications of Data Mapping in Microsoft Purview	128
Implementing Data Mapping in Microsoft Purview.....	129
Challenges and Considerations.....	131
Chapter 6: Data Insights and Analytics	133
Introduction.....	133
Control.....	135
Actions	141
Reports.....	142
Health Management Reports in Microsoft Purview	145
Sensitivity Label Insights About Your Data in Microsoft Purview	145
Get Insights into Data Stewardship from Microsoft Purview	146
Catalog Adoption from Microsoft Purview	149

TABLE OF CONTENTS

Chapter 7: Compliance in Microsoft Purview153

 Communication Compliance 155

 Communication Compliance Scenarios Designed to Help Organizations..... 157

 Data Lifecycle Management Compliance..... 162

 Microsoft Purview Records Management..... 163

 Compliance Solution for eDiscovery 164

 Capabilities of eDiscovery 164

 Microsoft Purview Compliance eDiscovery Features 168

 Attain Compliance with Regulatory Requirements Such as GDPR and CCPA with Microsoft Purview..... 170

Chapter 8: Information Security in Microsoft Purview177

 Introduction..... 178

 Information Protection 180

 Information Barriers..... 184

 Information Barriers with Microsoft Teams 185

 Information Barrier Triggers 187

 Impact of IB on Existing MS Teams Chats 190

 Conclusion 192

 Information Barriers with OneDrive and SharePoint Online 193

 Privileged Access Management..... 194

 Layers of Protection..... 195

Chapter 9: Best Practices for Effective Information Governance.....197

 Strategies for Optimizing Data Management Practices Using Microsoft Purview 197

 Develop a Unified Data Governance Strategy 198

 Implement Effective Data Classification..... 199

 Enhance Data Lineage and Traceability..... 200

Implement Data Quality Management Practices200

Optimize Data Access and Security.....201

Foster Data Stewardship and Accountability.....201

Support Compliance and Regulatory Requirements.....202

Enhance Data Integration and Interoperability203

Promote Data Literacy and Training.....203

Monitor and Optimize Data Management Processes.....204

Leverage Automation and Advanced Analytics204

Foster Collaboration and Communication205

Develop and Implement a Data Management Framework206

Conclusion206

Best Practices for Effective Information Governance Using
Microsoft Purview207

 Define Clear Governance Objectives and Scope.....207

 Develop a Comprehensive Data Catalog.....208

 Implement Robust Data Classification.....208

 Establish Data Governance Policies and Procedures209

 Enable Data Lineage and Traceability.....209

 Implement Data Access Controls.....210

 Ensure Data Quality and Integrity210

 Support Compliance and Regulatory Requirements.....210

 Foster Data Stewardship and Accountability.....211

 Promote Data Literacy and Training.....211

 Monitor and Audit Data Governance Activities212

Leverage Automation and Advanced Analytics212

 Integrate with Other Data Management Tools213

 Establish a Data Governance Framework.....213

 Foster Collaboration and Communication213

 Conclusion214

TABLE OF CONTENTS

Microsoft Purview Information Governance Stories215

 Financial Services Firm: Enhancing Compliance and Data Security215

 Healthcare Provider: Improving Data Quality and Integration.....216

 Retail Corporation: Streamlining Data Governance and Compliance.....217

Chapter 10: Future Trends in Information Governance.....219

 Introduction.....219

 Future Trends in Information Governance.....220

 Integration of Artificial Intelligence and Machine Learning221

 Big Data Analytics and Insights222

 Blockchain and Decentralized Data Management.....222

 Regulatory Trends and Compliance223

 Data Stewardship and Management224

 Data Democratization and Literacy225

 Managing Data Complexity and Scale.....226

 Collaboration and Partnership227

 Conclusion.....227

 Role of Artificial Intelligence (AI) and Machine Learning (ML) in
 Information Governance.....228

 Introduction to AI and ML in Information Governance.....228

 Enhancing Data Management with AI and ML.....229

 Improving Data Quality with AI and ML.....230

 Enhancing Data Security and Compliance231

 Supporting Data Governance and Stewardship231

 Facilitating Data-Driven Decision-Making232

 Future Potential and Developments.....233

 Predictions for the Future of Information Governance and
 Compliance Using Microsoft Purview234

 Intelligent Data Classification and Tagging234

TABLE OF CONTENTS

Automated Risk Assessment and Management.....235

Compliance with Emerging Regulations235

Enhanced Data Sovereignty and Localization235

Advanced Data Encryption and Protection.....236

Data Integrity Monitoring236

Comprehensive Data Cataloging and Discovery237

Data Stewardship and Ownership.....237

Integration with Emerging Technologies237

Collaboration with Blockchain and IoT.....238

Enhanced User Interfaces and Dashboards238

Training and Support.....239

Ethical Considerations in Data Management239

Conclusion239

Index.....241

About the Author



Charles David Waghmare, a DBA (Doctor of Business Administration) scholar from the prestigious SP Jain School of Global Management and an MBA from the same prestigious B-School, has over 17 years of industry experience in IT, engineering, and energy sectors.

Charles is presently working with a global energy leader since 2019 as an information management consultant in the Microsoft 365 space. Before that, he worked for Capgemini for eight years in various roles including Viva Engage community manager and manager of the Drupal-based Enterprise Knowledge Management system. He also developed a knowledge management platform for Capgemini's Digital Customer Experience (DCX) organization using SharePoint Online to manage client references and knowledge assets related to artificial intelligence and customer experience (CX). Further, he adopted Microsoft Azure chatbots to automate communication channels with the customers.

Charles also worked for ATOS (erstwhile SIEMENS Information Systems Limited) for five years. During his tenure there, he was a community manager of SAP-based communities, where he utilized Technoweb 2.0 – a Viva Engage-like platform – and on-premises SharePoint to manage SAP user-based communities. Also, Charles was the global rollout manager for a structured document management system built in on-premises SharePoint.

ABOUT THE AUTHOR

Charles has penned several books on Microsoft 365 technologies such as Viva Engage, SharePoint Online, and Azure chatbots and on ChatGPT. Further, he loves reading motivational books in his spare time, his favorite being *The Monk Who Sold His Ferrari*, *The 5 AM Club*, and *The Everyday Hero Manifesto*.

About the Technical Reviewer



Kasam Shaikh is a prominent figure in India's artificial intelligence landscape, holding the distinction of being one of the country's first four Microsoft Most Valuable Professionals (MVPs) in AI. Currently serving as a senior architect at Capgemini, Kasam boasts an impressive track record as an author, having authored five best-selling books dedicated to Azure and AI technologies. Beyond his writing endeavors, Kasam is recognized as a Microsoft Certified Trainer (MCT) and influential tech

YouTuber (@mekasamshaikh). He also leads the largest online Azure AI community, known as DearAzure | Azure INDIA, and is a globally renowned AI speaker. His commitment to knowledge sharing extends to contributions to Microsoft Learn, where he plays a pivotal role.

Within the realm of AI, Kasam is a respected subject matter expert (SME) in generative AI for the cloud, complementing his role as a chief architect. He actively promotes the adoption of No Code and Azure OpenAI solutions and possesses a strong foundation in hybrid and cross-cloud practices. Kasam Shaikh's versatility and expertise make him an invaluable asset in the rapidly evolving landscape of technology, contributing significantly to the advancement of Azure and AI.

In summary, Kasam Shaikh is a multifaceted professional who excels in both technical expertise and knowledge dissemination. His contributions span writing, training, community leadership, public speaking, and architecture, establishing him as a true luminary in the world of Azure and AI.

Acknowledgments

I would like to acknowledge the following people who are close to my heart:

Late Mr. Anil Malvankar, ex-DGM at SIEMENS, who offered me my first job at SIEMENS. I thank for his mentoring until his last day on the earth, 28th April 2024.

Late Mr. Alwin Fernandis, my beloved friend. He is not present with me today, but his memories will exist in my heart forever.

Mr. Sridhar “Sri” Maheswar, supply chain consultant, NNIT. I am always grateful for your support, my beloved friend.

Mr. Pravin V. Thorat, Head of India Operations at ATOS. I thank you for your prayers and good wishes.

CHAPTER 1

Introduction to Microsoft Purview

Microsoft Purview is a comprehensive suite of data governance and compliance tools designed to help organizations manage, protect, and gain insights from their data across various environments. Launched as part of Microsoft's broader strategy to address the growing complexities of data management in hybrid and multicloud environments, Purview provides a unified platform that integrates data cataloging, data lineage, and data classification with robust governance and compliance capabilities.

Besides managing data, Microsoft Purview is also focused on compliance and governance. It has a robust set of tools that can be used to enforce policies and ensure that organizations follow regulations such as HIPAA, GDPR, and CCPA. The platform's reporting features are additionally helpful in demonstrating compliance. One of the most critical features of Microsoft Purview is its ability to monitor and manage the data lineage. This feature allows organizations to identify and analyze the various systems that collect and store their data. It helps them ensure that their data is secure and that they are able to perform effective troubleshooting.

With the ability to integrate with other Microsoft offerings, such as Power BI and Azure Analytics, Microsoft Purview can help organizations establish a comprehensive data governance strategy. Thus, this ensures that the platform can serve as the central hub for any organization's

data management. The flexible and powerful Microsoft Purview data management platform is ideal for organizations that want to improve the efficiency and effectiveness of their data operations. It can help them manage their data assets while ensuring that they are compliant and secure. As the volume of data continues to increase, it will become even more important that organizations have the necessary tools to keep up with the changes.

In this introductory chapter, we aim to provide overview of Microsoft Purview features, capabilities such as data discovery and classification, data cataloging, data lineage, data governance, data sharing and collaboration, data privacy and security, and other features of Microsoft Purview. Further, we will look into architecture and core components such as information catalog, information map, information insights, and information policies.

Note Theoretically, there is a difference between the manner in which a piece of data and information is defined. In this book, we will sometimes prefer using the word information over data to accurately fix context; however, this setting does not determine the application to the way a piece of data is defined, and application of topics in this book will be equally deemed fit wherever it applies to data as well. Both words will be invariably used however with the scope of identical intended meaning and with a standard business understanding.

Introduction

The Microsoft Purview platform provides comprehensive solutions that help organizations manage and protect their data. These solutions help organizations address the various challenges that data fragmentation and

the lack of visibility present in their organizations. The Microsoft Purview platform combines the capabilities of its former Azure platform and 365 Compliance Services with a unified set of solutions designed to help organizations manage their data. Through the Microsoft Purview platform, different teams as shown in Figure 1-1 can gain a deeper understanding of their data and manage it across their entire lifecycle. It also helps them comply with regulations and manage.

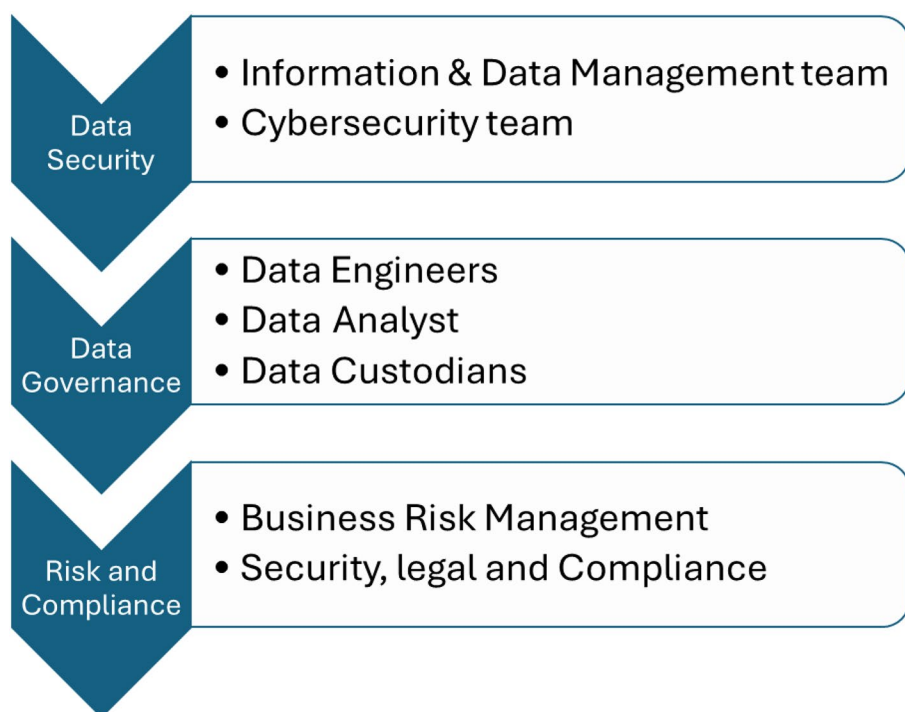


Figure 1-1. *Teams adopting Microsoft Purview*

Key Features of Microsoft Purview

With Microsoft Purview, organizations can now meet the diverse needs of their data management and governance teams. As they continue to expand their operations and collect data across various cloud, on-premises, and

software-as-a-service environments, they can realize a comprehensive view of their data at scale. In this section, we will explore the many features of Microsoft Purview, which will enable organizations to enhance their data governance and regulatory compliance efforts.

Comprehensive Information Discovery and Classification

Microsoft Purview's classification and data discovery engine are at the heart of the company's capabilities. Example of information discovery is eDiscovery which is used during legal proceedings to identifying evidence in emails, databases, and other electronic communication medium. On other hand, information can be classified as internal, confidential, and most confidential based on business requirements. Information which is classified as internal can be an intranet site for internal users. Project-specific documents will be classified as confidential as these are for a specific group of people. Customer contacts will be most confidential as companies' reputation depends upon such. In an era where information is spread across various platforms and environments, effective data classification and discovery are critical such as

Automated Information Discovery

- With Microsoft Purview's data discovery tools, organizations can easily identify and analyze data from various sources, such as on-premises databases and software-as-a-service offerings. This feature helps them maintain an updated inventory of their data.
- The Purview platform supports various types of data sources, such as relational databases and cloud-based platforms like Amazon Web Services, Google

Cloud Storage, and Microsoft Azure Blob Storage. This ensures that organizations can catalog and identify all their data regardless of its location.

- Purview can perform incremental scans, which only updates data that have changed since the previous scan. This feature helps organizations keep their data catalog current and reduce their system resources.

Through Microsoft's AI-powered insights, Purview's discovery engine can classify and recognize data according to its content, usage patterns, and context. This method ensures that new information is properly categorized.

Information Classification

When it comes to understanding an organization's data, classification plays a vital role. This is especially true when handling sensitive information which is explained as below.

- Microsoft Purview has built-in classifiers that can automatically identify various types of data. These include financial information, protected health information, and intellectual property. These tools can help organizations quickly find the location of sensitive data.
- Due to the unique nature of an organization's data, it is often necessary to create custom classification tools that are specific to the organization. Through Purview, an organization can identify the types of data that it needs to classify. These tools can be customized to recognize specific keywords or patterns.
- Classification automatically tags data with labels that highlight its ownership, type, and sensitivity. These labels can help improve the discoverability of information in the catalog and comply with various regulations.

Centralized Data Cataloging

Microsoft Purview's centralized data catalog is the hub of the platform, serving as a single point of reference that holds all the company's data assets with unified data cataloging. Examples of data cataloging could be for a given data set of cities whether one can group them based on country business where enterprise exists.

Unified Data Catalog

A comprehensive data catalog is essential for managing the massive amount of information that an organization has stored in various environments. With Microsoft Purview's catalog, you can easily consolidate all of your discovered data into a single repository, giving you a holistic view of the data landscape.

- The catalog's metadata management features capture and store details about each type of data asset, such as its structure, classification, lineage, and source. These provide valuable context to help you understand their quality and relevance.
- The Purview catalog can also integrate with business glossaries and data schemas, which help users understand the concepts of data in a more business-oriented manner. This feature helps nontechnical users gain access to the information they need.
- Users can easily search the contents of the Purview data catalog using various filters, such as source, owner, data type, and sensitivity. Its search capabilities are designed to provide relevant results in a fast and efficient manner, especially in large-scale data sets.

Data Asset Enrichment

In addition to basic metadata, the Purview platform also allows users to enrich their data assets with additional details, which helps improve their governance and usability with below:

- Purview's annotations feature allows users to add their own annotations to their data assets, which provide instructions on how to interpret and use the information. They can also be used to share knowledge across teams.
- The Purview data catalog can also assign one or both a steward and an owner to each asset. The latter is responsible for ensuring that the data follows policies, while the former is accountable for its relevance and accuracy. This process helps establish a clear understanding of the responsibilities related to data management.
- Purview's usage analytics feature can help users identify the activities and usage of their data assets. This information can be utilized to improve the efficiency of data management and ensure that it is readily accessible to those who need it.

Data Lineage and Impact Analysis

One of the most critical factors that organizations need to consider when it comes to tracing the flow of data is understanding the data lineage. This feature allows them to identify the various sources of their data and ensure that it is secure. With Microsoft Purview, they can easily view and analyze the data lineage.

End-to-End Data Lineage

Through Purview, organizations can visualize and capture the entire journey of their data, from its initial source to its destination. This feature provides a comprehensive view of how data is transformed and aggregated across various systems.

- Users can explore the data's journey using interactive lineage diagrams, which highlight the various steps involved in the data's transformation, such as data warehouses, ETL pipelines, and business intelligence tools. These tools make it easier to identify areas of weakness or potential bottlenecks.
- The ability to track and record the transformation process ensures that data is accurately stored and analyzed. This feature is very important for ensuring that the changes are performed properly and that the data is not affected by any issues.
- With the ability to map the sources and targets of data, users can easily identify where and how it is being consumed. This feature helps organizations keep track of their data's flow and ensure that they are able to maintain transparency.

Impact Analysis

An impact analysis is a must-have for organizations that require a deeper understanding of how changes in their data sets can affect them.

- Users can simulate the effect of changes to data structures, transformations, and sources through Purview's change impact assessment. This tool helps organizations identify potential disruptions that could affect their downstream systems and reports.

- The Purview feature can visualize the dependencies between various processes and data assets, which can help identify areas where changes can affect them. This is particularly useful in complex environments where multiple data sources are interconnected.

Robust Data Governance

The core of Microsoft Purview's capabilities is data governance, which enables organizations to manage and enforce policies. It provides them with the necessary tools to comply with regulations and legal requirements. Example of data governance includes retention of data such as information labeled as records based on country jurisdiction and removal of personal data such as Passport document, health checkup document from company repositories.

Policy Management and Enforcement

Data governance is a process that involves the creation, implementation, and enforcement of policies that regulate how information is accessed, utilized, and safeguarded.

- Through the use of Purview, organizations can create and enforce policies that specify the types of data assets that can be accessed and what conditions are required to safeguard them. These policies can then be applied across the enterprise to ensure consistency.