



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 13, Issue 11, November 2024

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.524

9940 572 462

6381 907 438

ijirset@gmail.com

www.ijirset.com

Enterprise Analytical Environment, from Agnostic to SAP Datasphere

Fabio De Castro Fialho¹, Govinda Rao Banothu²

SAP Analytics Solutions Architect, RJ, Brazil¹

SAP Analytics Cloud Specialist, NJ, USA²

ABSTRACT: This article explores the critical drivers of the agnostic Enterprise Analytical Environment that should shape the environment, proposes a structure of the Enterprise Analytical Environment, and details the specific implementation nuances for SAP Datasphere. This comprehensive approach ensures a well-governed yet adaptable analytical environment.

KEYWORDS: Enterprise Analytical Environment, SAP Datasphere, Data Lake, Business Data Hub, Multi-source, Multi-granularity, Data Governance, Visualization Tools.

I. INTRODUCTION

The transition from an agnostic Enterprise Analytical Environment to implement it with SAP Datasphere involves understanding key drivers of the agnostic Enterprise Analytical Environment and implementing strategic architectural conceptual changes. In this paper, we depict the structure of a proposed agnostic Enterprise Analytical Environment, shifting from usual approach of Enterprise Data Lakes to Business Data Hubs, proposing changes in the traditional EDW Layers. Also, we provide the details to implement the Enterprise Analytical Environment using SAP Datasphere as the Data Warehouse Cloud Platform.

II. THE AGNOSTIC ENTERPRISE ANALYTICAL ENVIRONMENT DRIVERS

The agnostic Enterprise Analytical Environment must be conceived focusing the minimal disruption to existing informational initiatives and investments. Ensuring technology flexibility helps mitigate disruptions and supports the swift access to information, crucial in today's fast-paced environment.

An effective informational architecture must be ready for various source systems, including unstructured and semi-structured data. It is important to select suitable solutions for different data sources, recognizing that a single solution may not suffice for all cases. A multi-layer approach is essential to cater to various levels of data granularity, providing both highly subdivided and aggregated information for various consumer types.

Reducing data footprint through in-memory technology and Service-Oriented Architecture (SOA) approaches is crucial. The architecture should focus on data consumption rather than replication or persistency. Governance models should empower end-users while keeping IT oversight. Power Users, as owners of the information process, play a significant role in this model.

The architecture should allow for different tools and solutions within the same layer while keeping a technology-agnostic approach. Visualization tools selection should be based on user preference, ensuring they align with corporate governance processes.

Said that the drivers are:

1. Non-disruptive/Minimal Disruption: Even in organizations without formal informational architectures, there are typically ongoing informational initiatives and investments. The goal is to minimize disruption and adopt a technology-agnostic approach to ensure a resilient architecture.
2. Multi-source Enabled: Modern informational architectures must support various data sources, including unstructured and semi-structured data. The logical model may encompass multiple physical system solutions, each with its governance and technical peculiarities.

3. Multi-granularity Capable: The architecture should cater to different audiences, providing various levels of data granularity. It should support harmonized/cleansed data for direct consumption or integration into Data Vaults, Data Marts, or EDW implementations.
4. No Data Persistency Obligatory: Reducing the data footprint is crucial, with a focus on data consumption rather than replication. The architecture should leverage in-memory technology and SOA approaches to avoid unnecessary data storage.
5. Safe, Simple, and Flexible Governance: Governance should empower end-users while ensuring IT oversight. Power Users play a significant role, owning the informational process end-to-end.
6. Allow Different Tools/Solutions for One Layer: The architecture should be multilayered and technology-agnostic, allowing the use of the best-of-breed tools for specific purposes or centralized solutions with flexible options.
7. Enabled to Use Any End-User Visualization Tool: Visualization tools selection shall be in line with the end-user preference, supported by mature data lifecycle management and proficient users.

III. THE AGNOSTIC ENTERPRISE ANALYTICAL ENVIRONMENT

Considering technological advances, the Enterprise Analytical Environment must be simple and flexible. Instead of an Enterprise Data Lake (EDL), we propose starting with the Business Data Hub (BDH), focusing on governing business-use data. The BDH may include an optional Landing layer for data persistence based on connectivity, sizing, and granularity needs.

Governance must be at the Macro Business Process level, keeping strict Naming Convention standards. This approach adapts the Content Areas concept to leverage improved search tools and capabilities. Tagging and data cataloguing enable expanding Content Areas while keeping segmentation within Naming Conventions.

The traditional EDW layers Operational Data Store and Data Marts becomes to Analytical Models and Data Products respectively, emphasizing business relevance. This approach shifts the focus from technical details to business needs, making the architecture more user-friendly.

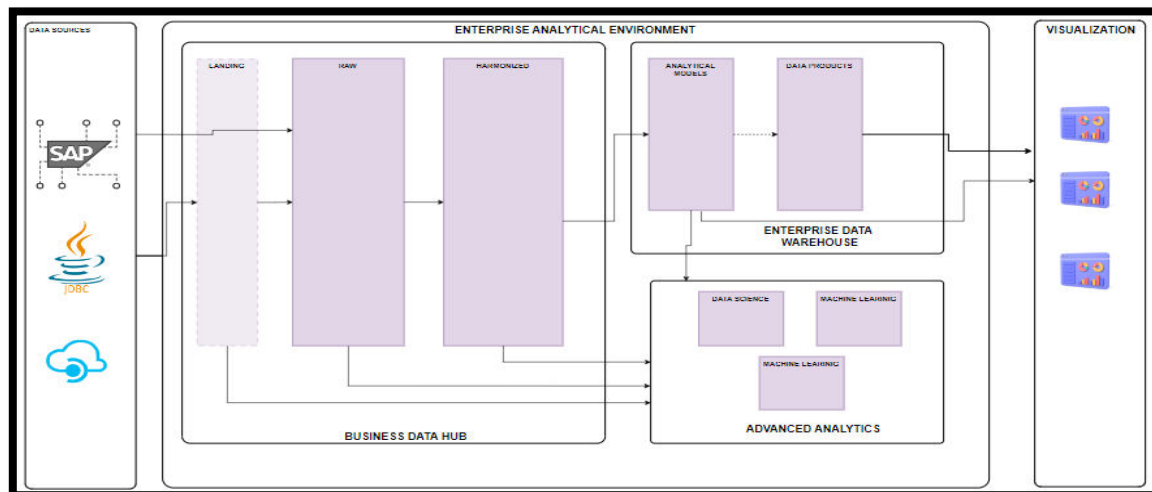


Fig 1: Agnostic Enterprise Analytical Environment, Components and Layers

IV. ENTERPRISE ANALYTICAL ENVIRONMENT FOR SAP DATASPHERE

The natural next step would be to provide the guidelines for “branding” the components of the architecture. By branding we mean selecting the tools to deploy the architecture. But after 6 years of working with different tools and so many PoC’s (Proof of concepts) and MVPs (Minimum Viable Products) we can guarantee that all solutions in the market are equivalent, so in the end of the day the choice will be is much more commercial or political then anything.

So, instead of selection tool guidelines we will jump directly to the nuances of implementing the architecture with SAP Datasphere, as the for the other technologies a multitude of architecture recommendations can be easily found on the internet. We also assume that the target audience has a minimal knowledge of SAP Datasphere key features which will allow us to focus in specifically in the nuances of the architectural aspects.

The crucial point is to discuss the use of the Layers, the Content Areas, and the Naming Convention to establish a well governed and, at the same time, flexible environment. In SAP Datasphere, the SPACES plays a vital role in terms of governance, connectivity, and resource management.

Therefore, it will also play a significant role in the deploy of the architectural concepts we advise. Next, we present the graphical representation of the Enterprise Analytical Environment adjusted to SAP Datasphere.

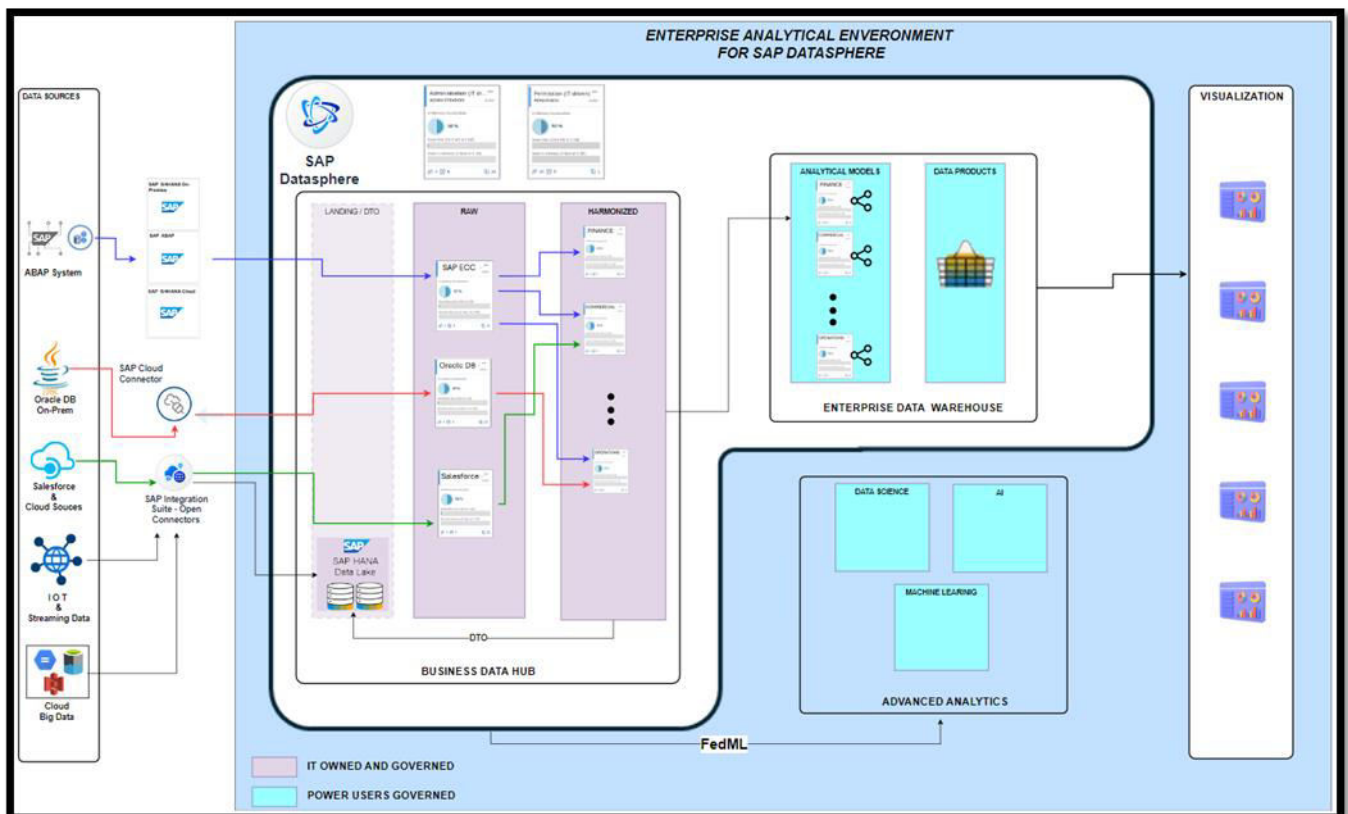


Fig 2: Enterprise Analytical Environment deployed with SAP Datasphere

In the picture above it is possible to notice the usage of SAP Datasphere Spaces in the Raw Layer segmented by system source type once system connections are Space dependent in SAP Datasphere. Naming Convention here should keep source system names to allow a straightforward consistency to check on extractions and replications.

In the Harmonization Layer, as data is not system dependent anymore, the Spaces organizes the Business Macroprocesses, and the Naming Convention should guarantee the subprocesses detail level ensuring data ownership and governance. Here it is important to remark that we advocate that IT owns and govern Business Data Hub layers artifacts in terms of building and managing, but data ownership and governance themselves shall be kept with the business. And the reason is simple: it is in Business Data Hub that SAP Datasphere Data Builder will be used to perform the Data Transformation that requires coding and a in depth knowledge of denormalization and data modelling e.g., data Cleansing, table joins, intelligent lookups, complex business rules, but data will always belong to the business. It is not that Power Users shall not be allowed to use SAP Datasphere Data Builder, but when this is the case a separated and specific Space must be spined of for them, and this Space will be considered as part of the Analytical

Models Layer. In other words, SAP Datasphere Data Builder use is not limited to IT and by the Business Data Hub component, but it plays the central role of this component that that belongs to IT.

SAP Datasphere provide no-code/low-code capabilities that enable Business Power Users develop their own analytical models and data products mainly using SAP Datasphere Business Builder. So, here Power User fully owns and govern their Spaces and the data “productization” process. This enables, for example, Finance team to share with Sales team not the raw financial data for them to build their KPIs, neither Finance team needs to be responsible of delivering sales key figures with both perspectives Finance and Sales. Instead, Finance team will be able to share with Sales team financial data as Data Products with the minimal set of accounting corporate/legal principles applied, for then Sales team can apply their own business concepts and rules. Of course, that if this type of organizational disruption, the organization must provide adequate training, and the data products can be shared only within the organization. A corporate data governance committee must approve any external data product sharing.

As we are talking about organizational disruption, the need of discussing a real need of Change Management support is necessary. Business Power Users must realize that this independence from IT comes with bucket full of duties. Basically, Business Power Users will become the frontline for the end users of their Data Products, any issues on the Data Products will be under their responsibility.

Besides the Spaces of Raw, Harmonized and Analytical Models Layers, two other Spaces completes the set of the required Spaces to deploy the Enterprise Analytical Environment. One Administration Space for the management of the environment as whole and a Permissions Space to manage access to each one of the other Spaces, its artifacts, and the data they deliver.

Visualization Layer can use or not SAP SAC, and of course that if this is the case SAC will be able to work seamlessly with SAP Datasphere than any other Visualization Tool. But still, SAC, as any other Visualization Tool, will have its own governance process.

Finally, a word about the Landing Layer and the BW Bridge. We left the first layer and the BW Bridge to the end of the text because if in one hand the Landing Layer is most optional of all, the BW Bridge could be in all of them. The Landing Layer is the most optional since it might be used mostly in two scenarios: a) The source system data extraction tool/method delivers data more efficiently in some sort of low-cost cloud storage; and b) the organization chooses to use SAP Hana Data Lake for Data Tiering Optimization, or as a Central corporate Data Lake. The BW Bridge can be in almost all layers since BW itself, in the general has its own layered organization and SAP Datasphere can consume BW Bridge artifacts data in any of its stages in the BW Bridge Space.

V. CONCLUSION

The transition from an agnostic Enterprise Analytical Environment to the Enterprise Analytical Environment for SAP Datasphere involves a strategic approach that balances flexibility and governance. This shift uses key drivers to create a robust and adaptable analytical framework.

The introduction of the Business Data Hub (BDH) concept, with its focus on governing business-use data, marks a significant departure from traditional Enterprise Data Lakes. This model enhances data management efficiency and aligns with modern technological advances.

Implementing this environment with SAP Datasphere further refines the approach by using Spaces for governance, connectivity, and resource management. The segmentation of Spaces by system source type and Business Macroprocesses, combined with stringent Naming Conventions, ensures a well-governed and coherent data landscape.

SAP Datasphere's no-code/low-code capabilities empower business users, fostering greater independence from IT and enabling the creation of tailored analytical models and data products. This transformation requires adequate training and a well-structured change management strategy to ensure smooth adoption and maximize the benefits of this flexible, user-centric environment.

In summary, transitioning to the Enterprise Analytical Environment for SAP Datasphere creates a seamless, efficient, and future-ready architecture that supports diverse data sources, granularity levels, and visualization tools, while keeping strong governance and flexibility. This comprehensive approach ensures that organizations can effectively navigate the complexities of modern data management, driving informed decision-making and business success.

REFERENCES

1. Data Warehouse Implementation with the Corporate Information Factory (CIF). Technical Publications.
2. Kimball, R. (2009). Kimball's Data Warehouse Toolkit Classics. John Wiley & Sons.
3. Linstedt, D. (2015). Building a Scalable Data Warehouse with Data Vault 2.0.
4. Christensen, J., Darlak, J., Harrington, R., Kong, L., Poles, M., & Savelli, C. (2017). SAP BW/4HANA: The Comprehensive Guide. SAP PRESS.
5. Bobbili, G. M. (2024). SAP Datasphere: An Introduction | Beginner's Guide. SAP PRESS
6. SAP Community. (2019, October 29). Integration Solution Advisory Methodology (ISA-M): Define integration guidelines for your organization. SAP Community. <https://community.sap.com/t5/technology-blogs-by-sap/integration-solution-advisory-methodology-isa-m-define-integration/ba-p/13397214>
7. Paritala, S. (2023, September 8). A Business Data Fabric for All: The Democratizing Power of SAP Datasphere. Avvale Newsroom. <https://www.avvale.com/newsroom/sap-datasphere-business-data-fabric>
8. Fialho, F. (2017, September 19). The importance of agnostic/non-branded informational architecture: Part I. Fábio Fialho Blog. <https://fabiofialho.com.br/2017/09/19/the-importance-of-the-agnosticnon-branded-informational-architecture-part-i/>
9. Fialho, F. (2017, September 26). The importance of the agnostic/non-branded informational architecture. Fábio Fialho Blog. <https://fabiofialho.com.br/2017/09/26/the-importance-of-the-agnosticnon-branded-informational-architecture/>
10. Fialho, F. (2018, June 6). The enterprise analytical environment: A flexible and time-to-market driven and governed approach. Fábio Fialho Blog. <https://fabiofialho.com.br/2018/06/06/the-enterprise-analytical-environment-a-flexible-and-time-to-market-driven-and-governed-approach/>



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details