Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 31 No. 8s (2024)

Best Practices for Designing Semantic Layers in Business Intelligence Systems

Prateek Panigrahy

KIIT University, Bhubaneswar, Odisha

prateek.panigrahy@gmail.com

Article History:

Received:04/10/2024
Revised:03/11/2024
Accepted:05/12/2024

Abstract:
In Business Intelligence (BI) systems, semantic layers bridge analysis-friendly user interfaces and complicated raw data. They enable the existence of homogenous metrics, hierarchies, and dimensions and render them usable, more reliable, controlled, and scalable across data sources. The paper highlights the best practices of semantic layer design, such as theoretical backgrounds, architecture, implementation, and emerging problems. It places emphasis on governance, scalability, interoperability, and evolving with the evolving times, such as cloud elasticity, real-time analytics, compliance, and AI integration. This paper draws on academic literature and benchmarking research, which together provide a foundation for designing semantic layers that are uniform, replicable, and dynamic in the development of BI systems.

Keywords: Semantic Layers; Business Intelligence (BI); Data Governance; Cloud Analytics; Self-Service BI

1. Introduction

The Business Intelligence (BI) systems have become significant and support data-driven decision-making in companies. The BI is problem-oriented in nature, whereby the raw and complex data, which is largely heterogeneous, is then transformed into information that can be examined and put into usable use by those involved. This redefinition demands a process that will do away with the technical complexity of the databases and the simplicity of analysis that is demanded by the end-users. This is supposed to be supplied by the semantic layer and is a translation and abstraction layer between the physical data stores and user-oriented analytical tools [1]. The conceptual architecture is the semantic layer that makes sure the information that is represented in it is in a form that can easily be consumed by the business, concealing the inner workings of database schemas, data sources, or query expressions. It enables business users to relate to key performance indicators, metrics, and dimensions without the need of having heavy technical skills. In that respect, it can be democratically accessed to information without interfering with information reports and analytics [2]. It is, however, not easy to come up with good semantic layers. Layers not implemented efficiently might lead to differences, inefficiencies, and governance threats. In contrast, the built semantic layers may increase trust, speed, and adoption of the BI systems in case they are developed adequately [3].

The growing number of institutions that engage bigger and more heterogeneous data sets has made semantic layers an even more noticeable task. The new BI systems should be capable of accessing the data lakes, the cloud-based warehouses, and streaming pipelines, and that is why the semantic design should also be a usability issue, but of scalability and interoperability [4]. In addition, the expectations have been altered due to the self-service analytics. The business customers are developing more demanding views on how they can discover the data on their own, with the assistance of IT departments to provide governance and semantic standardisation. The major issue with the semantic layer design is that the trade-off between flexibility and control is the problem [5].

The semantic layer design has a decisive influence on the BI adoption and organisation performance, which is proven by the practices and frameworks of benchmarking. A correct semantic layer ensures consistency of KPI and dimension definition, cross-departmental interoperability, and a single source of truth, which is essential in making a decision based on data [6]. The paper addresses the best practice of semantic layer design, including the conceptual foundation of semantic abstraction, issues of integration, governance, and scalability, and issues of user-centric design, metadata management, and performance tuning. The emerging problems of cloud BI, AI, and real-time analytics are also mentioned, and the analysis provides a logical premise that should be adhered to by practitioners in the successful semantic layer design.

2. Theoretical Foundations of Semantic Layers

Continuing on the theme of the importance of semantic layers in reducing complexity and making the interface easier to use, what is needed is a discussion on the theory underlying them. The essential idea of a semantic layer is a layer that interposes the logical model of business data and the physical one, where the entities, relations, and measures are represented in business terms instead of technical ones, and therefore are defined in terms of business terms [7]. This conceptual basis responds to the fact that raw data, as much as it can be abundant does not have an inherent meaning until it is interpreted using business logic. For example, while individual transaction details are stored in the database, the semantic layer can define 'Revenue' as the sum of these amounts according to specific business rules. This ensures consistency in calculations and aligns analytical results with the overall business strategy [8].

Semantic abstraction of BI is associated with a reduction of the cognitive load in that the business users do not often have the experience of a normalised schema, join criteria, or SQL. Semantic layers eliminate this load and present the information in the common hierarchies, dimensions, and measures, i.e., on the notions of the human-computer interaction, making it easier to use and does not affect the functional richness [9]. Interoperability is another dimension that BI ecosystems bring together relational databases, NoSQLs, streaming platforms, and external APIs. These resources are correlated on the semantic level, and such indicators as Customer Lifetime Value or Inventory Turnover are normalised across analyses [10]. Semantic layers help eliminate conflicting interpretations by centralising the definition of metrics and access controls. This not only enhances reporting consistency but also supports the principle of maintaining a single version of the truth [11]. It is all these theoretical underpinnings that make semantic layers relevant in the contemporary BI.

The theoretical models provide us with the knowledge of the semantic layer; however, the success of architectural design is what determines the success in practice. The following section then shifts the discussion of the abstract principles to the material building points that are present in the background of the effective implementation of the semantic layers.

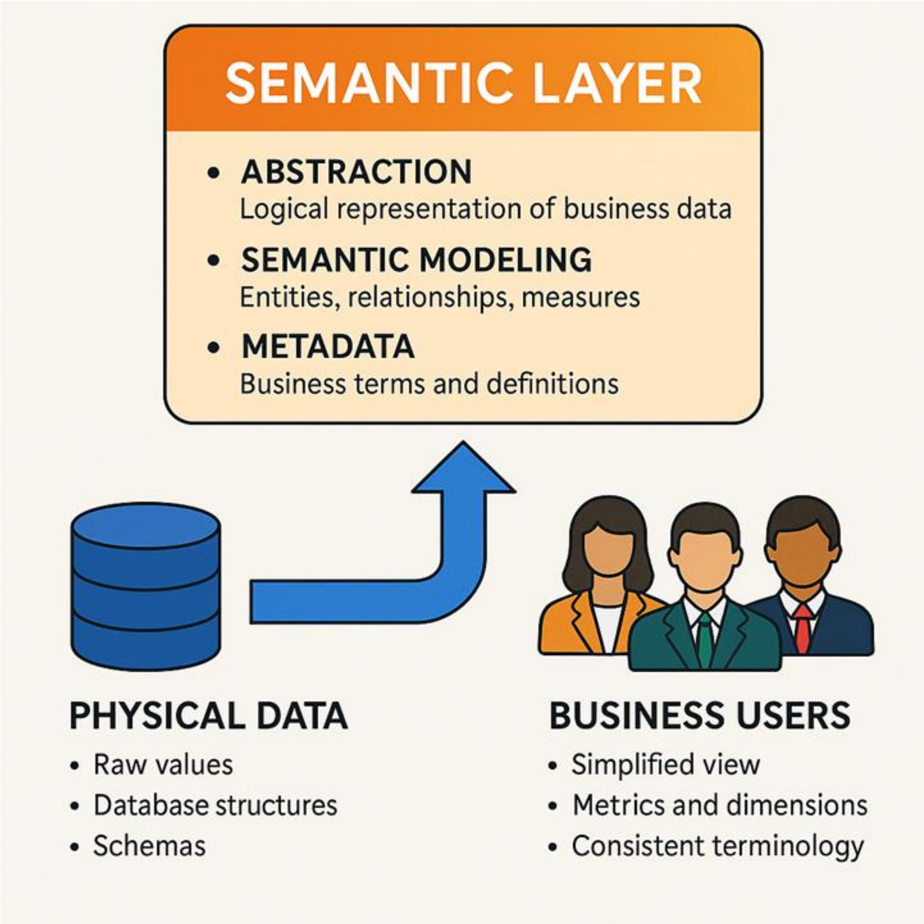

Figure 1: Illustration of the semantic layer as an abstraction between physical data and business users

3. Architectural Considerations in Semantic Layer Design

Design of the semantic layer must consider practical architectural reality on the basis of the theoretical discourse. Semantic layers need to be done at the edge of data storage, integration pipelines, and analytical interfaces, and strike a balance in terms of performance, governance, and usability to be effective [12]. The integration is necessary to accommodate different sources, such as relational databases and cloud warehouses, and the virtualisation and federation are more dynamic than rigid ETL pipelines [13]. Caching, pre-aggregation, and materialised views are some of the performance optimisations that may make the system more responsive, but they may come at the cost of storage and updating latency [14]. In governance, security, access control, metadata management, and lifecycle management are applied [15]. Scalability can be applied to manage the growing data and number of users simultaneously with horizontal scaling, parallel queries, and cloud-native utilisation of solutions [16]. It is also compatible with other BI tools using standard APIs as a result of interoperability [17]. Finally, a broader analytics approach should be equipped with semantic layers, where foreseeable and derived measurements [18] can be altered to AI-powered analytics.

To better understand how the architectural practices are translated into a practical strength, the following table outlines the strategies that are actively implemented during the creation of the semantic layer by presenting the areas of focus of the architecture.

Table 2: Architectural Strategies in Semantic Layer Design and Their Primary Contributions

Architectural Focus	**Strategy Used**	**Primary Contribution**	**Potential Trade-offs**
Integration	Data virtualisation, federation	Enables access across diverse sources without heavy ETL	May introduce query latency
Governance	Centralised metadata repositories, role-based access	Ensures consistency, security, and compliance	Requires continuous upkeep of metadata
Scalability	Distributed caching, parallel query execution, and cloud elasticity	Supports growing data volumes and concurrent users	Increased complexity in monitoring
Performance	Query caching, materialised views, in-memory engines	Improves responsiveness for analytical queries	Higher storage and refresh costs
Interoperability	Open APIs, semantic connectors	Enables BI tool neutrality and consistency	Dependent on vendor support for standards

Source: Synthesised from literature [12], [14], [16]

This comparative summary illustrates that architectural strategies each offer distinct strengths and trade-offs, reinforcing the need for holistic design choices. Having highlighted the architectural backbone, the discussion naturally progresses to implementation practices, which operationalise these strategies in real BI environments.

4. Implementation Practices in Semantic Layer Design

Having explored the architectural dimensions of semantic layers, it is necessary to shift focus toward the practical implementation practices that transform theoretical and architectural principles into functioning systems. While architecture establishes the structural backbone, implementation ensures that the semantic layer delivers tangible value to business users in terms of usability, performance, and governance [19].

A central implementation practice is user-centric design. Semantic layers succeed only when they align with the needs and expectations of their end users. This means involving business stakeholders early in the design process to define metrics, hierarchies, and dimensions that reflect business priorities rather than purely technical structures. Collaborative workshops, iterative prototyping, and continuous feedback loops help ensure that the semantic layer speaks the language of the business, which fosters trust and adoption [20]. Metadata management is another practice that is important. An advanced metadata structure, which is referred to as a semantic layer, is defined by the quality of the metadata cataloguing, mapping, and revealing, which determines its success. It is anticipated that the metadata repositories will store data lineage, data ownership, and governance rules along with the definition of metrics. This makes it easy to have transparency, and therefore, the users are able to trace insights to their sources. Metadata-based automation also reduces the level of manual labour involved by ensuring continuity of the BI tools and reports [21]. Semantic abstraction can be more complex for queries; therefore, unless optimised, it will cause low response times and user experiences. To resolve this, aggregating the commonly used dimensions by materialising them, indexing the key dimensions, and using in-memory computation is normally employed. Techniques that are used in query caching are also sensitive in high-concurrency environments. The performance monitoring process should be dynamic, and they should use telemetry to detect bottlenecks to dynamically tune the optimisation process [22].

The implementation semantic layer should involve the practices of security and governance. Row-level security, column-level masking, and role-based access are considered to be the standard mechanisms that ensure that the sensitive data is not compromised. The centralised governance and consistency are accompanied by the introduction of these controls to the semantic layer and not in individual tools. It is also employed to help the organisations with complying with regulatory requirements since semantic layers can record and audit the user engagement with the data is achievable [23]. There is also a need for lifecycle management practices and version control. Semantic models evolve along with business demands, and the evolution requires the application of disciplined approaches to control the evolution. Some of the right practices to ensure that semantic changes are made in a safe, predictable way are maintaining versioned models, aiding testing in a sandbox environment, and automated deployment pipelines [24].

Another important practice involves interoperability testing. Since BI ecosystems often span multiple tools and platforms, the semantic layer must be validated for compatibility across them. Interoperability ensures that the same definitions of metrics and hierarchies are accessible consistently across dashboards, reports, and predictive analytics applications. Without such testing, fragmentation can erode the integrity of the semantic layer [25]. User training and documentation play a final but vital role. Even the most sophisticated semantic layers fail if users do not understand how to use them effectively. Training programs should explain not only how to navigate the semantic layer but also why certain definitions and metrics have been established. Clear documentation, ideally integrated directly into BI tools, fosters self-service analytics while reducing dependence on IT teams [26].

The implementation of semantic layers, therefore, is not purely technical but also organisational. Success depends on fostering collaboration, governance, and continuous feedback while embedding performance and security at the core. With implementation practices established, the discussion can now progress to the emerging challenges that organisations face as they extend semantic layers into modern BI environments shaped by cloud computing, real-time data, and artificial intelligence.

Figure 2: Overview of key semantic layer implementation practices, including user-centric design, metadata management, performance tuning, governance, interoperability, and user training, all essential for delivering scalable and trusted BI systems.

5. Emerging Challenges in Semantic Layer Design

As BI ecosystems evolve, semantic layers face emerging challenges driven by cloud-native architectures, real-time analytics, and AI-driven insights [27]. Scalability in dynamic cloud environments requires rethinking caching, metadata distribution, and query federation [28]. Real-time analytics demand semantic models that reconcile streaming and historical data while maintaining stable metric definitions [29]. Data diversity from relational and NoSQL databases to unstructured and AI-derived metrics complicates harmonisation [30]. Governance and compliance challenges, amplified by regulations such as GDPR and CCPA, necessitate compliance-aware metadata and dynamic access controls. Balancing self-service and control is critical to avoid semantic fragmentation. Integrating predictive and prescriptive AI insights requires new metadata standards to ensure interpretability. Finally, user adoption challenges underscore the need for cultural alignment and change management to build trust in BI systems.

Table 2: Emerging Challenges in Semantic Layer Design

Challenge	Description	Implications / Techniques
Scalability in Cloud	Adapting to elastic compute and storage across distributed environments	Caching strategies, metadata distribution, and query federation
Real-time Analytics	Reconciling streaming and historical data	Continuous metric updates, low-latency query optimisation
Data Diversity	Integrating structured, unstructured, and AI-driven data	Semantic harmonisation, flexible dimensional modelling
Governance & Compliance	Aligning with GDPR, CCPA, and multi-jurisdictional regulations	Compliance-aware metadata, dynamic access control, and data lineage
Self-service vs Control	Empowering users while maintaining semantic consistency	Governance frameworks, metric approval workflows
AI/ML Integration	Supporting predictive and prescriptive insights	AI metadata standards, interpretability frameworks
User Adoption	Ensuring organisational trust and engagement	Training, change management, cultural alignment

The challenges of semantic layers design make it clear that semantic layers design is not a stagnant domain but an active sphere that has to respond to various technological, organisational, and regulatory changes. As a reaction, the concluding section summarises the best practices into a single list of recommendations, which offer systematic advice on how to create a good semantic layer in contemporary BI systems. Given the wide range of pressures currently affecting semantic layers, presenting a summary of these challenges and their impacts in a table provides a clear, at-a-glance overview of how they influence business.

Table 3: Emerging Challenges in Semantic Layer Design and Their Implications

Challenge Area	Description	Implications for Semantic Layers
Cloud Elasticity	Dynamic scaling of compute/storage in cloud BI	Requires adaptive caching and distributed metadata to maintain consistency
Real-time Analytics	Integration of streaming and historical data	Semantic models must reconcile fast-changing data with stable definitions
Data Diversity	Sources include SQL, NoSQL, APIs, and unstructured content	Calls for flexible abstractions to harmonise heterogeneous structures
Regulatory Compliance	GDPR, CCPA, and domain-specific rules	Semantic layers must embed lineage, anonymisation, and auditing features
Self-service vs. Control	Balancing empowerment and governance	Risk of metric fragmentation if definitions are not enforced centrally
AI Integration	Inclusion of predictive/derived metrics	New standards for interpretability and transparency are required

Source: Synthesised from literature [27], [28], [30]

6. Conclusion

Recent Business Intelligence (BI) systems rely on semantic layers that encompass theory, architecture, implementation, and emerging challenges. These layers are built on principles of abstraction, which reduce cognitive load, enhance interoperability, and support governance. As a result, users can derive consistent insights without needing to navigate technical complexities. From a semantic perspective, heterogeneous data sources need to be integrated into semantic layers and cloud-native distributed environments, while ensuring interoperability with other BI tools. Their operational implementation should be grounded in user-centric design, effective metadata management, strong governance, and performance tuning. The new requirements also underline the need for the semantic layers to be dynamic to enable real-time analytics, cloud elasticity, regulatory compliance, and insights powered by AI. Organisations need to strike a balance between self-service and control, flexibility and compliance, as well as traditional and predictive analytics. The best practices are in collaboration with stakeholders, designing governance, metadata of high quality, and active monitoring of performance and AI requirements. Combined, these strategies create semantic layers as adaptable, trusted, and actionable models, which serve as a solid basis of evidence-based decision-making.

References

[1] Golfarelli, M., Rizzi, S. and Cella, I., 2004. Beyond data warehousing: What's next in business intelligence?. *Proceedings of the 7th ACM international workshop on Data warehousing and OLAP*. ACM, pp. 1-6.

[2] Chaudhuri, S., Dayal, U. and Narasayya, V., 2011. An overview of business intelligence technology. *Communications of the ACM*, 54(8), pp. 88-98.

[3] Watson, H.J., 2009. Business intelligence-Past, present, and future. *Communications of the Association for Information Systems*, 25(39), pp. 487-510.

[4] Jarke, M., Lenzerini, M., Vassiliou, Y. and Vassiliadis, P. (eds.), 2003. *Fundamentals of Data Warehouses*. Springer, Berlin.

[5] Kimball, R. and Ross, M., 2013. *The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling*. 3rd ed. Wiley, Hoboken.

[6] Golfarelli, M. and Rizzi, S., 2009. Data warehouse design: Modern principles and methodologies. *McGraw-Hill*, New York.

[7] Inmon, W.H., 2005. *Building the Data Warehouse*. 4th ed. Wiley, Indianapolis.

[8] Vassiliadis, P., 2009. A survey of extract–transform–load technology. *International Journal of Data Warehousing and Mining*, 5(3), pp. 1-27.

[9] Chaudhuri, S. and Dayal, U., 1997. An overview of data warehousing and OLAP technology. *ACM SIGMOD Record*, 26(1), pp. 65-74.

[10] Abelló, A., Romero, O. and Simitsis, A., 2013. Fusion cubes: Towards self-service business intelligence. *International Journal of Data Warehousing and Mining*, 9(2), pp. 66-88.

[11] Watson, H.J. and Wixom, B.H., 2007. The current state of business intelligence. *Computer*, 40(9), pp. 96-99.

[12] Gluchowski, P. and Chamoni, P., 2016. *Analytical Information Systems: A Conceptual Framework*. Springer, Berlin.

[13] Rizzi, S., Abelló, A., Lechtenbörger, J. and Trujillo, J., 2006. Research in data warehouse modeling and design: Dead or alive?. *Proceedings of the 9th ACM International Workshop on Data Warehousing and OLAP*, pp. 3-10.

[14] Bauer, H. and Günzel, H., 2016. *Business Intelligence and Analytics: Systems for Decision Support*. Springer, Cham.

[15] Sahay, B. and Ranjan, J., 2008. Real time business intelligence in supply chain analytics. *Information Management & Computer Security*, 16(1), pp. 28-48.

[16] Watson, H.J., 2014. Tutorial: Big data analytics: Concepts, technologies, and applications. *Communications of the Association for Information Systems*, 34(1), pp. 65-84.

[17] Chaudhuri, S. and Narasayya, V., 2007. Self-tuning database systems: A decade of progress. *Proceedings of the 33rd International Conference on Very Large Data Bases (VLDB)*. Morgan Kaufmann, pp. 3-14.

[18] Vassiliadis, P. and Simitsis, A., 2009. Near real-time ETL. *Information Systems*, 34(3), pp. 299-321.

[19] Gluchowski, P., Gabriel, R. and Dittmar, C., 2012. Self-service business intelligence. *Business & Information Systems Engineering*, 4(3), pp. 131-139.

[20] Hovi, A., Hervonen, H. and Koistinen, J., 2009. Information requirements for business intelligence systems. *Journal of Computer Information Systems*, 49(4), pp. 81-90.

[21] Sherman, R., 2014. *Business Intelligence Guidebook: From Data Integration to Analytics*. Morgan Kaufmann, Waltham.

[22] Baars, H. and Kemper, H.G., 2008. Management support with structured and unstructured data-An integrated business intelligence framework. *Information Systems Management*, 25(2), pp. 132-148.

[23] Dinter, B., 2013. Success factors for information logistics strategy-An empirical investigation. *Decision Support Systems*, 54(1), pp. 1207-1218.

[24] Tan, C. and Hunter, M.G., 2002. The repertory grid technique: A method for the study of cognition in information systems. *MIS Quarterly*, 26(1), pp. 39-57.

[25] Oliveira, M. and Quintana, Y., 2019. Analytics for business intelligence frameworks. *Future Generation Computer Systems*, 90, pp. 489-498. Elsevier.

[26] Chen, H., Chiang, R.H.L. and Storey, V.C., 2012. Business intelligence and analytics: From big data to big impact. *MIS Quarterly*, 36(4), pp. 1165-1188.

[27] Marjanovic, O., 2007. The next stage of operational business intelligence: Creating new challenges for business process management. *Proceedings of the 40th Hawaii International Conference on System Sciences*. IEEE, pp. 215-224.

[28] Imhoff, C. and White, C., 2011. *Self-Service Business Intelligence: Empowering Users to Generate Insights*. TDWI Best Practices Report, Elsevier.

[29] Negash, S. and Gray, P., 2008. Business intelligence. *Decision Support Systems*, 34(1), pp. 3-8.

[30] Davenport, T.H. and Harris, J.G., 2007. *Competing on Analytics: The New Science of Winning*. Harvard Business School Press, Boston.