nagarro
ENTERPRISE AGILE

BI (Business Intelligence)

Technology Matrix | Products offered by Various Vendors

Target Areas considered for Matrix Creation

- Connectivity/adapter capabilities (data source and target support).
- Data delivery capabilities
- Data transformation capabilities
- Metadata and data modeling capabilities
- Design and development environment capabilities
- Data governance support capabilities (via interoperation with data quality, profiling and mining capabilities)
- Deployment options and runtime platform capabilities
- Service enablement capabilities

Connectivity/adapter capabilities (data source and target support).

The ability to interact with a range of different types of data structure, including:

- Relational databases
- Legacy and non-relational databases
- Various file formats
- XML
- Packaged applications, such as CRM and supply chain management
- SaaS and cloud-based applications and sources
- Industry-standard message formats, such as electronic data interchange (EDI), Health Level Seven International (HL7) and Society for Worldwide Interbank Financial Telecommunication (SWIFT)
- Parallel distributed processing environments such as Hadoop Distributed File System (HDFS) and other NoSQL-type repositories
- Message queues, including those provided by application integration middleware products and standards-based products (such as Java Message Service)
- Data types of a less structured nature, such as social media, Web clickstream, email, websites, office productivity tools and content repositories
- Emergent sources, such as data on in-memory repositories, mobile platforms and spatial applications
- Screen-scraping and/or user interaction simulations (for example, scripts to interact with Web, 3270, VT100 and others)

Connectivity/adapter capabilities Continued…

Data integration tools must support different modes of interaction with this range of data structure types, including:

- Bulk/batch acquisition and delivery
- Granular trickle-feed acquisition and delivery
- Change data capture (CDC) — the ability to identify and extract modified data
- Event-based acquisition (time-based or data-value-based)

Data delivery capabilities

- The ability to provide data to consuming applications, processes and databases in a variety of modes, including:
  - Physical bulk/batch data movement between data repositories such as processes for extraction, transformation and loading (ETL) or extraction, loading and transformation (ELT)
  - Federated/virtualized views formulated in-memory
  - Message-oriented movement via encapsulation
  - Replication of data between homogeneous or heterogeneous DBMSs and schemas

In addition, support for the delivery of data across the range of latency requirements is important, including:

- Scheduled batch delivery
- Streaming/near-real-time delivery
- Event-driven delivery of data based on identification of a relevant event

Data transformation capabilities

Built-in capabilities for achieving data transformation operations of varying complexity, including:

- Basic transformations, such as data-type conversions, string manipulations and simple calculations
- Transformations of intermediate complexity, such as look-up and replace operations, aggregations, summarizations, deterministic matching and the management of slowly changing dimensions
- Complex transformations, such as sophisticated parsing operations on free-form text and rich media

In addition, the tools must provide facilities for developing custom transformations and extending packaged transformations.

Metadata and data modeling capabilities

As the increasingly important heart of data integration capabilities, metadata management and data modeling requirements include:

- Automated discovery and acquisition of metadata from data sources, applications and other tools
- Discerning relationships between data models and business process models
- Data model creation and maintenance
- Physical-to-logical model mapping and rationalization
- Ability to define model-to-model relationships via graphical attribute-level mapping
- Lineage and impact analysis reporting, via graphical and tabular format
- An open metadata repository, with the ability to share metadata bidirectionally with other tools
- Automated synchronization of metadata across multiple instances of the tools
- Ability to extend the metadata repository with customer-defined metadata attributes and relationships
- Documentation of project/program delivery definitions and design principles in support of requirements definition activities
- Business analyst/end-user interface to view and work with metadata

Design and development environment capabilities

Facilities for enabling the specification and construction of data integration processes, including:

- Graphical representation of repository objects, data models and data flows
- Management of the development process workflow, addressing requirements such as approvals and promotions
- Granular, role-based and developer-based security
- Team-based development capabilities, such as version control and collaboration
- Functionality to support reuse across developers and projects, and to facilitate the identification of redundancies
- Support for testing and debugging

Data governance support capabilities (via interoperation with data quality, profiling and mining capabilities)

Mechanisms to work with related capabilities to help the understanding and assurance of data quality over time, including interoperability with:

- Data profiling tools (profiling and monitoring the conditions of data quality)
- Data mining tools (relationship discovery)
- Data quality tools (supporting data quality improvements)

Deployment options and runtime platform capabilities

Breadth of support for the hardware and operating systems on which data integration processes may be deployed, and the choices of delivery model; specifically:

- Mainframe environments, such as IBM z/OS and z/Linux
- Midrange environments, such as IBM System i or HP Tandem
- Unix-based environments
- Windows environments
- Linux environments
- On-premises (at the customer site) installation and deployment of software
- Hosted off-premises software deployment (dedicated, single-tenant implementation)
- Integration platform as a service (iPaaS), consumed by the customer completely "as a service" — the vendor provides cloud infrastructure; the customer does not install the software
- Cloud deployment support as a multitenant implementation (requires organizations to deploy software in cloud infrastructure)
- In-memory computing environment (such as flash memory as an additional memory type and solid-state appliances)
- Server virtualization (support for shared, virtualized implementations)
- Parallel distributed processing (such as Hadoop and MapReduce)

Operations and administration capabilities

Facilities for enabling adequate ongoing support, management, monitoring and control of the data integration processes implemented via the tools, such as:

- Error handling functionality, both predefined and customizable
- Monitoring and control of runtime processes, both via functionality in the tools and through interoperability with other IT operations technologies
- Collection of runtime statistics to determine use and efficiency, as well as an application-style interface for visualization and evaluation
- Security controls, for both data in flight and administrator processes
- A runtime architecture that ensures performance and scalability

Architecture and integration capabilities

- The degree of commonality, consistency and interoperability between the various components of the data integration toolset, including:
  - A minimal number of products (ideally one) supporting all data delivery modes
  - Common metadata (a single repository) and/or the ability to share metadata across all components and data delivery modes
  - A common design environment to support all data delivery modes
  - The ability to switch seamlessly and transparently between delivery modes (bulk/batch versus granular real-time versus federation) with minimal rework
  - Interoperability with other integration tools and applications, via certified interfaces and robust APIs
  - Efficient support for all data delivery modes, regardless of runtime architecture type (centralized server engine versus distributed runtime)

Service enablement capabilities

As acceptance of data service concepts continues to grow, so data integration tools must exhibit service-oriented characteristics and provide support for SOA deployments, such as:

- The ability to deploy all aspects of runtime functionality as data services (for example, deployed functionality can be called via a Web services interface)
- Management of publication and testing of data services
- Interaction with service repositories and registries
- Service enablement of development and administration environments, so that external tools and applications can dynamically modify and control the runtime behavior of the tools

CHALLENGERS

Microsoft SQL Server Integration Services (SSIS, included in its SQL Server DBMS license) - **13,000 Clients**

LEADERS

Informatica Platform (including PowerCenter, PowerExchange etc.) – **5500 Clients**

IBM InfoSphere Information Server Enterprise Edition – **10,300 Clients**

Oracle Data Integrator (ODI), Oracle Data Service Integrator, Oracle GoldenGate and Oracle Warehouse Builder (OWB) – **4,000 Clients**

SAP Data Services, SAP Process Orchestration – **12000 Clients**

SAS Data Management Platform, Federation Server, SAS/Access – **14,000 Clients**

NICHE PLAYERS

Adeptia Enterprise Business Integration Management (EBIM) Suite – **420 Clients**

Syncsort DMX (Linux, Unix and Windows), DMX-h (Hadoop), Ironcluster Hadoop ETL for Amazon EMR, Ironcluster ETL for Amazon EC2 and Ironcluster ETL, Docker Edition – **1,400 Clients**

VISIONARIES

Talend Enterprise Data Integration and Talend Enterprise Big Data – **3,300 Clients**

Information Builder iWay Integration Suite – **730 Clients**

Actian Analytics Platform – **6,700 Clients**

Cisco Information Server – **250 Clients**

Informatica

Informatica offers the following data integration products: Informatica Platform (including PowerCenter, PowerExchange, Data Services, Data Replication, Ultra Messaging, Big Data, B2B Data Exchange and Data Integration Hub), Vibe Data Stream, Informatica Cloud Integration and the IronCloud Platform.

Strengths

- **Breadth of functionality and usage.** Customer implementations reflect a comprehensive mix of data integration styles, latency and deployment scenarios where adoptions are commonly regarded as the enterprise standard for data integration infrastructure. Synergy with Informatica's data quality, MDM, big data and hybrid integration approaches (portability of integration flows between Informatica's iPaaS and on-premises PowerCenter) capitalize on market demand trends.
- **Market presence and alignment to evolving needs.** Informatica's mind share in this market is extensive, with frequent appearances in competitive situations. Provision of a platform-agnostic environment to build capabilities that can be seamlessly executed and reused on many established and emerging platforms (through the architecture Informatica calls Vibe Virtual Data Machine), positions Informatica well with the demand trends of modern information infrastructure.
- **Appeal to information infrastructure and nontechnical roles.** Informatica continues to expand its business-user-facing functionality — in support of self-service validation and prototyping support — by provisioning data virtualization and data profiling functionality to analytics and operational activities. Emphasis on business-user-oriented functionality and the agility of data integration infrastructure as enterprise standards resonates with diverse types and sizes of organization.

Cautions

- **Pricing.** While deployments reflect a reasonable connection between the pricing of Informatica's data integration tool and its anticipated value, customers and prospects often express concerns about its high prices relative to alternatives in this market. The release of Informatica 9.6, in January 2014, simplified product packaging and pricing tiers and was aimed at enabling more flexible procurement choices that vary according to the scale of need and the maturity of organizations.
- **Administrative complexity.** Reference customers expressed concerns for complex version upgrades and migrations, and overlapping functionality in multiple products. Customers cite growing difficulty in knowing what product to use for specific issues — when functionality spans multiple products or features that have been merged.
- **Ease of integrated deployment across product portfolio.** Users of Informatica's data integration tooling are having their expectations raised — for greater out-of-the-box and simpler integrated deployment — with its other offerings, such as data quality, MDM, B2B and RulePoint products.

IBM

IBM offers the following data integration products: IBM InfoSphere Information Server Enterprise Edition (including InfoSphere Information Server for Data Integration, InfoSphere Information Server for Data Quality and InfoSphere Business Information Exchange),

Strengths

- **Breadth and diversity of usage.** IBM's data integration tools continue to be deployed in the market for extensive use cases — often of complex scale, spanning a wide number of projects and involving teams of various sizes. Common metadata, development and deployment approaches are favored for enabling consistency of, and support for, all data integration styles.
- **Mind share and synergy with related markets.** IBM continues to gain traction as an enterprise standard for data integration infrastructure, with a strong presence in competitive bids. Linkage of data integration alongside its BigInsights for big data analytics, and positioning data integration in Watson Foundations toward a future era of envisaged smart machines and cognitive capability, are raising the synergy of IBM's data integration tooling with its broader portfolio.
- **Alignment to information infrastructure and enterprise information management (EIM).** IBM continues to focus on enabling modernization of information infrastructures in aligning data integration capabilities with EIM goals including information governance and MDM.

Cautions

- **General usability challenges.** Customers reported difficulty with version upgrades and migrations. IBM has begun mitigating this through in-place upgrades and will continue addressing it by converging the release timing of various InfoSphere products, for better anticipation and alignment of upgrades (as well as regular feature pack releases), such that customers can adopt new features without an upgrade. A longer time to value and the complexity of implementation are cited as challenges (in part due to use cases of complex scale), which IBM is mitigating through self-service data preparation for analytics (using data virtualization and Data Click), integration into Hadoop environments, and iPaaS for data integration.
- **Pricing.** Reference customers identify software costs and perception of the total cost of ownership (TCO) as barriers to broader adoption. IBM's provision of Information Server solutions, Workgroup editions and new monthly licensing options for both on-premises and cloud deployments aims at providing wider procurement and cost choices.
- **Deployment of multiple components across the portfolio.** Reference customers expressed difficulty with integrated use of the various data integration tools with other InfoSphere products and recently added components such as connectors for social media feeds. They also expressed a desire for easier ways to understand the overlapping features among products in order to determine what can be used license-free from the products currently in use. The newly released version of the InfoSphere platform, 11.3, sets out to reduce this complexity.

Oracle

Oracle offers the following data integration products: Oracle Data Integrator (ODI), Oracle Data Service Integrator, Oracle GoldenGate and Oracle Warehouse Builder (OWB).

Strengths

- **Usability and productivity.** Customers using ODI like its ease of use, coverage of core data integration functionality and out-of-the-box artifacts to aid developers' productivity as key value points. Plans for an iPaaS offering and capability for user-driven data preparation are aimed at harnessing cloud deployment trends and business-roles usage.
- **Leverage wide span of markets.** Oracle's corporate brand, as a comprehensive provider for data integration and other data and application-oriented capabilities (spanning data quality, MDM, ESB, analytic appliance and enterprise application), continues to drive its appeal for deployment scenarios.
- **Time to value.** Rapid deployment for mission-critical usage and the ability to flexibly interoperate with diverse technical environments and standards, are reported as strengths for Oracle. Tighter product integration between ODI and Oracle GoldenGate has contributed to integrated use of data integration tooling, which helps to accelerate deployments.

Cautions

- **Integrated usage of vendor portfolio.** To facilitate a seamless ability to expand deployments across use cases, reference customers cited their desire for improvement in metadata management support and simpler interoperability of data integration tooling with Oracle's other products. Oracle has indicated (on its road map) the upcoming availability of a metadata management product for data integration to address these concerns.
- **Pricing and licensing.** Satisfaction with Oracle's pricing method and the perception of value relative to cost are reported as being relatively low when compared with most of its competitors — involving concerns such as price points, target versus source licensing requirements, and extending licensing across virtualized and cloud-based environments.
- **Skills availability.** Finding skilled resources is reported as an increasing challenge in keeping pace with efforts to deploy, interoperate and maintain multiple products when implementation requirements and complexity grow. Oracle has deployed online and instructor-led skills training for continuing to grow the pool of trained ODI and Oracle GoldenGate developers in the market.

SAP

SAP offers the following data integration products: SAP Data Services, SAP Process Orchestration, SAP Replication Server, SAP Landscape Transformation Replication Server and SAP Hana Cloud Integration.

Strengths

- **Broad usage and functionality.** SAP's data integration products are regularly deployed for many different use cases and across the breadth of functionality of SAP's portfolio. SAP's vision emphasizes virtualized computing performance, an iPaaS delivery model and user-driven data preparation, to capitalize on demands for data integration capability that matches the speed of business.
- **Synergy with EIM-enabling technologies and SAP applications.** Customers value the tight links between the data integration functionality of information-related technologies and SAP's applications. A single runtime platform for SAP Data Services — linked to data quality and the text data processing capabilities complementary to offerings in MDM and business-facing information stewardship — is described by customers as increasingly relevant for supporting EIM goals.
- **Market presence and growth.** As a large and incumbent (in many tens of thousands of enterprises) provider of applications and analytics solutions, SAP can naturally capture significant revenue and growth in this market by leveraging its broader customer base.

Cautions

- **Market emphasis and perception.** Emphasis on Hana in SAP's road map to evolve toward a common environment to provision and consume data generates an uneven perception in organizations seeking agnostic data integration capabilities. Replication and synchronization's connectivity/adapters for heterogeneous databases are reported to have lagged behind other vendors' DBMS releases.
- **Integration of product components.** Making multiple tools work together across all SAP data integration offerings is reported as a challenge, raising concerns about increased effort and cost in achieving integrated usage. SAP plans to simplify its product portfolio and tighten the links between Data Services, Replication Server and PowerDesigner.
- **Product support and version upgrades.** Reference customers' feedback indicates low satisfaction with quality, consistency and response time of processes for assistance in product version upgrades, bug fixes and issue resolution. Through a scheduled quarterly release of data integration products, including Data Services and Hana Cloud Integration, SAP seeks to enhance the availability and timeliness of its product support and upgrades.

SAS

SAS offers the following data integration products: Data Management Platform, Federation Server, SAS/Access.

Strengths

- **Relevant and extensive functionality.** Breadth of core functionality and extensive connectivity position SAS well to engage in competing for contemporary data integration tool demand amid larger and more established vendors in this market. An increased focus on balancing physical movements and virtualized delivery of data, distributed runtime optimization (for big data, in-memory and cloud environments), and envisioned user-driven data preparation for exploring patterns using Hadoop environments, capitalizes on the evolving market demands.
- **Customer relationship excellence.** Reference customers report that vendor relationship with SAS, both in presales and post implementation, is exceptional — contributing to longer-term, recurring engagements.
- **Integrated product set.** Provisioning of functionality, through an integrated portfolio of capabilities that enables synergistic use, is cited as a key value. The recently added lineage repository (linked to data integration tooling) offers richer metadata support for business terms and data relationships.

Cautions

- **Price point and TCO.** Customers' perception of the high price of the licensing model and the complexity of the technology generated concerns regarding value relative to cost, as deployment complexity and a longer learning curve have led to cost escalation. Greater availability of resources with a deep knowledge of SAS tools is desired for wider procurement and cost options outside the vendor's own professional services business.
- **Links with data quality operations.** Although implementations often use SAS tools for both data integration processes and data quality operations, a decline in user satisfaction is exhibited in these deployments — requiring easier administrative operations and performance optimization.
- **Product messaging.** Some prospects, particularly organizations with awareness of individual products that existed before the release of the Data Management Platform, indicate that SAS has not clearly articulated the capabilities of its tools and the benefit of a tool suite approach in response to RFP and the presentation of data integration offerings.

Thank You!!