Improving Design for Data Warehouse Automation with WhereScape 3D 2.5

Abstract

With information quickly obtained and stored in WhereScape 3D's metadata repository, designers can now refine the discovered metadata with business logic to find undocumented links between business entities and strengthen the value of the design. When complete, WhereScape 3D automatically creates the appropriate best practice design in a normalized, star schema, or data vault structure. By doing this, 3D again raises the level of participation for designers to a strategic level as opposed to low-level tactical chores associated with common tasks.

Finally, with a generated and refined design, 3D can test the results with sample data from the data stores to ensure that the design will function as required by business stakeholders. This level of automated design for testing early in the process allows for confirmation and acceptance of the design without having to create a fully operational environment. While this type of testing is core to agile implementation methodologies, it has excellent value for SDLC/waterfall implementation teams as well.

HIGHLIGHTS

Vendor Name: WhereScape®

Product Name: WhereScape 3D

Product Version: Version 2.5

Availability: Generally available 1Q 2016

Product Function: Data-driven design tool for data warehouse automation

Vendor Contact: info@wherescape.com

WhereScape®

For data-driven organizations with these challenges, WhereScape 3D provides the type of automated discovery and design platform necessary to meet the pace of change. Enterprise Management Associates recognizes WhereScape 3D, and its associated data warehouse automation platform RED, as superior options for organizations that need to break down the walls between discovery, design, and implementation to effectively productionalize analytical environment design and speed delivery at the pace of business changes.

WhereScape designs, develops, sells, and supports WhereScape 3D, a data-driven design tool for data warehouse automation (DWA) that provides an integrated design and planning environment for building, deploying, and managing analytical environments such as enterprise data warehouses. This ENTERPRISE MANAGEMENT ASSOCIATES® (EMA™) product brief reviews WhereScape 3D and recognizes WhereScape as a leader in the field of discovery and design for data warehouses.

Corporate Background

WhereScape (https://www.wherescape.com/) is an international data warehouse and big data toolset software company founded in New Zealand in 1997. In 2001, the company expanded its operations to the United States. WhereScape has over 700 active customers in 15 countries that include Nordstrom, Union Bank, and GE Aviation.

WhereScape's products are used in analytical and business intelligence projects in commercial environments. Project applications include enterprise data warehouses, data vaults, data marts, and reporting systems. In addition to WhereScape 3D, WhereScape offers WhereScape RED, a data warehouse automation (DWA) tool that provides an integrated development environment (IDE) for building, deploying, managing, and renovating data warehouses.

Key Benefits

- **Empowers Agile Development –** WhereScape 3D supports agile development methodologies and iterative design by automating time-consuming tasks such as change detection in source and target environments and integration with the WhereScape RED platform.
- **Supports Industry Design Best Practices –** 3D comes with support for normalized data warehouse, dimensional data mart, and innovative data vault design schemas. Support for database constructs such as tables, views and indexes are included with WhereScape 3D. WhereScape 3D also includes design templates for data stores, enterprise data warehouses (3NF), data marts (including star schemas and snowflakes) and data vaults.
- **Allows for Target Schema Testing –** 3D allows for the design and testing of target data models using real data in iterative cycles. This ensures that target analytical environment schema for a data warehouse, data mart, or other environment have been reviewed, tested, and accepted.
- **Data Warehouse Discovery and Documentation -** Existing analytical environments can be automatically discovered, profiled, and documented with WhereScape 3D. Design documentation is automatically maintained and produced on demand to improve visibility and communication.
- **Details –** For more detailed information on WhereScape 3D, go to https://www.wherescape.com/products/

Product Description

WhereScape 3D is a data-driven design tool that brings together automated schema discovery, a metadata repository, and a design interface to enable and automate the process for the design of analytical environments. This approach allows for organizations to improve the speed of design for their analytical environments, such as enterprise data warehouses (EDW), data marts, and data vaults. It also improves the quality of those designs because the process is automated and all the resulting metadata associated with source systems schemas and analytical environment designs are catalogued and maintained in an independent and centralized metadata repository.

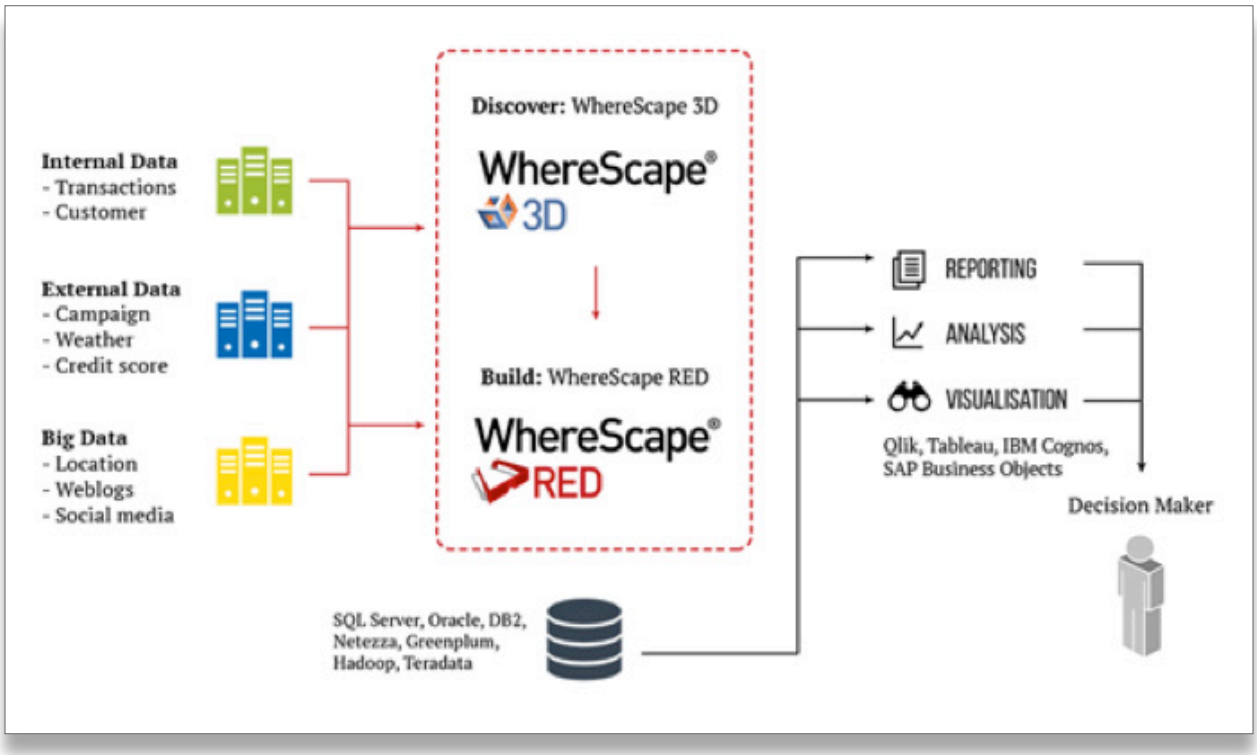
WhereScape 3D supports the following concepts for the design of analytical environments.

- Discover, profile, and assess source systems automatically
- Design, test, revise, and enhance data models in a visual user interface that populates and updates a centralized metadata repository
- Determine the impact of changes based on data lineage information from the metadata repository
- Create and maintain source systems and design documentation based on metadata information
- Rapidly promote designs to a WhereScape RED data warehouse automation (DWA) platform for analytical environment implementation

WhereScape 3D starts with data-driven discovery by automating the catalog and profiling and interrogating source systems. This discovery process not only enables organizations to catalog and document their impression of the source systems from a paper or external design, it also allows them to understand and see what is actually persistent in those data stores. By automating the discovery process, WhereScape 3D allows for analytical environment designers to focus on core activities and avoid technical "busy work" of spelunking a database via the command line. The automated discovery process traverses the schema of the source system and collects all the appropriate technical metadata from each of the tables or entities within a given table space or domain.

WhereScape 3D takes the information from the automated discovery and manages metadata from source systems in an independent repository. The management of metadata is accomplished so that multiple platforms and analytical environments can be maintained from a single WhereScape 3D instance. Designers use this information to aid and automate the design process via WhereScape's design interface.

This metadata information on entities and attributes is catalogued and stored within WhereScape 3D where a data model representation with all the appropriate links and constraints between entities is detailed and the design of an analytical environment starts. Designers use this information to aid and automate the design process via WhereScape's design interface. Designers can update any information not represented in the discovered schema, for example, constraints enforced in operational support systems (OSS) such as billing or enterprise resource planning (ERP). Designs can also be simplified by "hiding" entities not related to others or with little or no information. Empty tables from a standardized ERP application installation are excellent examples of entities that can clutter a design.



WhereScape 3D supports multiple design standards, including normalized star schema, and data vault analytical environment design to support their analytical environments. By taking best practices associated with these common architectural choices, WhereScape allows organizations a wide range of choices for their design and implementation. 3D also provides organizations with common design elements that can be tedious or missed when part of a manual process for analytical environment design. These design best practices can be linked with either traditional development lifecycles (SDLC) or with more iterative agile implementation practices, depending on the approach adopted.

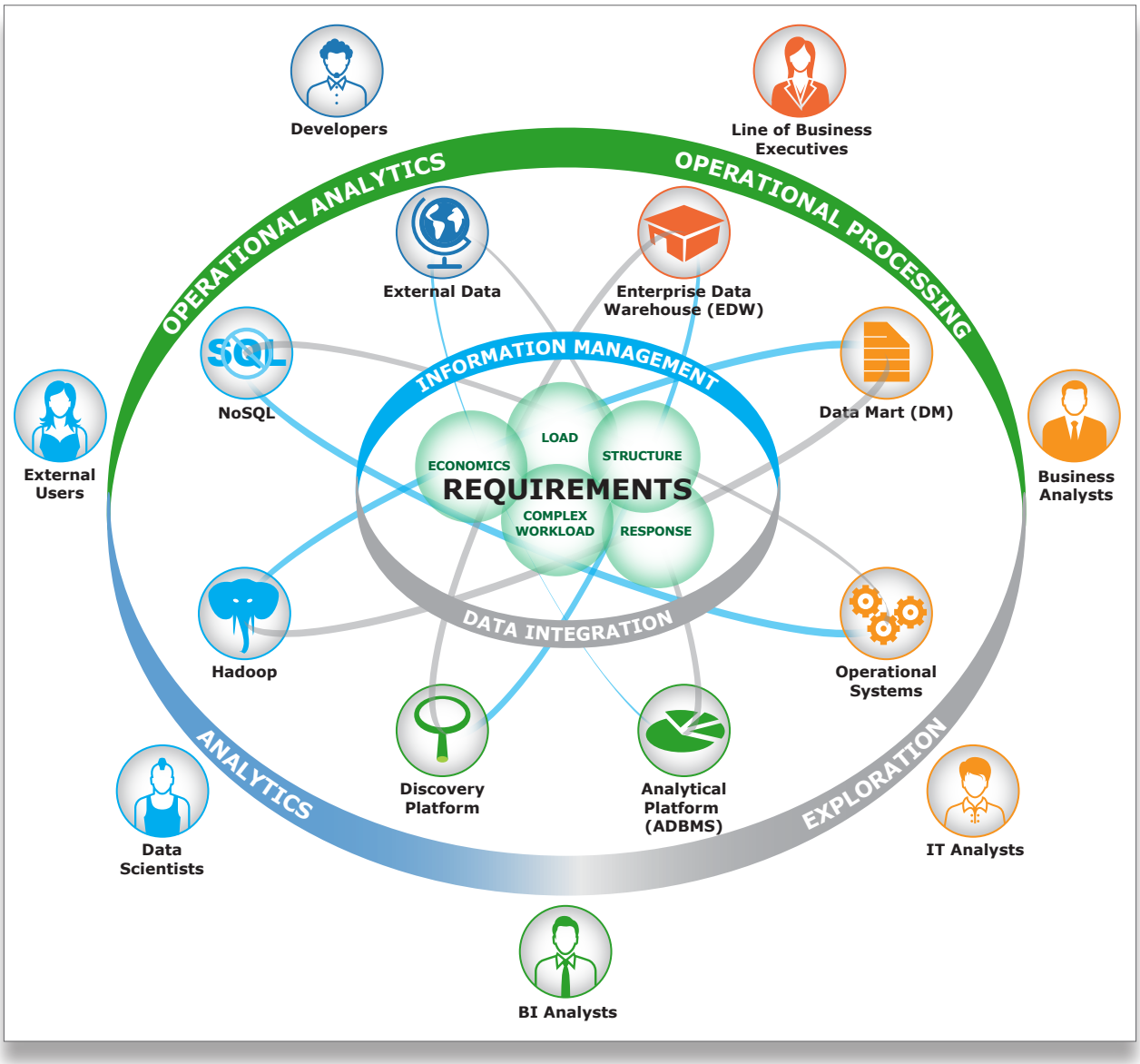
Proactive and iterative testing as part of an analytical environment's design and development has much the same value as it does with standard software development activities. Testing early and often improves quality and determines discrepancies between requirements and implementation design before being rolled out into production environments. With a completed design, WhereScape 3D allows organizations to test with real data the structures associated with their designs. This level of testing allows for early and automated testing to ensure the quality of designs and acceptance on the behalf of business stakeholders.

WhereScape 3D's independent metadata repository also enables organizations to develop and distribute documentation on analytical environment design at the push of a button. Unlike design platforms separated from the technical metadata of a discovered source system, 3D's metadata provides an easily updated and published set of documentation based on actual conditions as opposed to manual interpretation of existing source systems and target analytical schemas.

Product Position

WhereScape 3D fits within the Information Management Layer of the EMA Hybrid Data Ecosystem (HDE). The product specifically supports the design of analytical environments, such as **enterprise data warehouses (EDW)** and **data marts (DM)** and **analytical platforms** for both new construction and ongoing maintenance. WhereScape 3D also enables the automated documentation and discovery of source systems such as operational platforms, external data sources such as cloud databases, and existing analytical environments to improve the quality of and the speed the implementation for platforms within the HDE.

WhereScape 3D focuses on **developers** and **IT analysts** as users for its primary target audience.



By supporting the design and maintenance of analytical environment schemas on multiple data management platforms, WhereScape 3D enables organizations to position analytical information on the best data management platform to meet their workload requirements. WhereScape also provides the ability to design analytical environments to meet the changing requirements of business in a flexible and iterative fashion using best practices for design and implementation when used in concert with WhereScape RED.

EMA Perspective

The speed and pace of change in business is driving a corresponding impact on the pace of information technology (IT) delivery. Opportunities to capitalize on new market trends and customer preferences are driven by data sources from mobile, social, and online applications. A trend based on a social media "frenzy" may last anywhere between a few months and a year. There is also nearly constant change in the structure and ownership of business. Merger and acquisition (M&A) activities dramatically increase the types and numbers of operational platforms within our data centers and in our spheres of influence, with the inclusion of cloud-based platforms such as Salesforce.com. There is also a third component that can impact the pace of IT change: evolving regulation. With changes in national and international regulation laws such as Basel III in the EU and the Dodd-Frank Act in the United States, there are changing rules for reporting and compliance with government regulations. All of these challenges need to be considered for the design and maintenance of analytical environments.

To master these challenges, manual processes are no longer a best practice. When IT departments with stretched resources are faced with fast paced change, they need to allocate their scarce manpower resources on the most important strategic tasks as opposed to tactical tasks that prevent breakthrough performance. Organizations need to invest in the ability to smartly automate the lower level aspects of analytical environment design and delivery to focus their staff on the higher level—and higher value—concepts that make a difference in a competitive landscape.

Starting with discovery, WhereScape 3D enables the time-consuming task of source platform discovery and removes the potential for errors. By using the live information in source systems or existing analytical environments, WhereScape 3D allows organizations to quickly and productively scan, catalog, and detail the metadata of a database structure or the nature of flat file contents. Using this level of automation, designers can avoid the manual process of listing tables with a SQL "DESCRIBE" command and transferring the information to an external source of documentation.

EMA considers the agile design and deployment concepts inherent in the WhereScape 3D product to be core to the direction of proactive organizations. Being able to automatically discover and standardize components of analytical environment design is a key ingredient to implementing analytics faster and more reliably. WhereScape's 3D environment, and the associated WhereScape RED platform for data warehouse automation, provide the opportunity to move analytical environment design and deployment from a 12-18 month process to a much shorter time frame and be more responsive to business conditions and opportunities.