Documentation by

PHOENIX
Workgroup LLC

East Lansing, Michigan 48823, USA

John HR Schuster
Business Solution Architect

Technology Terms
Phoenix Workgroup LLC

John HR Schuster

Version 2.1b, 12/13/2019

Table of Contents

1. Terms . . . 1
    1.1. Data Lake . . . 1
    1.2. Slowly Changing Dimensions . . . 1
    1.3. Type 1 and 2 Dimensions . . . 2
    1.4. Wherescape . . . 2
        1.4.1. 3D . . . 3
        1.4.2. Red . . . 4
2. Document History . . . 5

Technology Terms that need better definitions.

PDF Version of the document can be download at Link

1. Terms

1.1. Data Lake

A data lake is a storage repository that holds a vast amount of raw data in its native format until it is needed. While a hierarchical data warehouse stores data in files or folders, a data lake uses a flat architecture to store data.

Any large data pool in which the schema and data requirements are not defined until the data is queried.

The term data lake is often associated with Hadoop-oriented object storage.

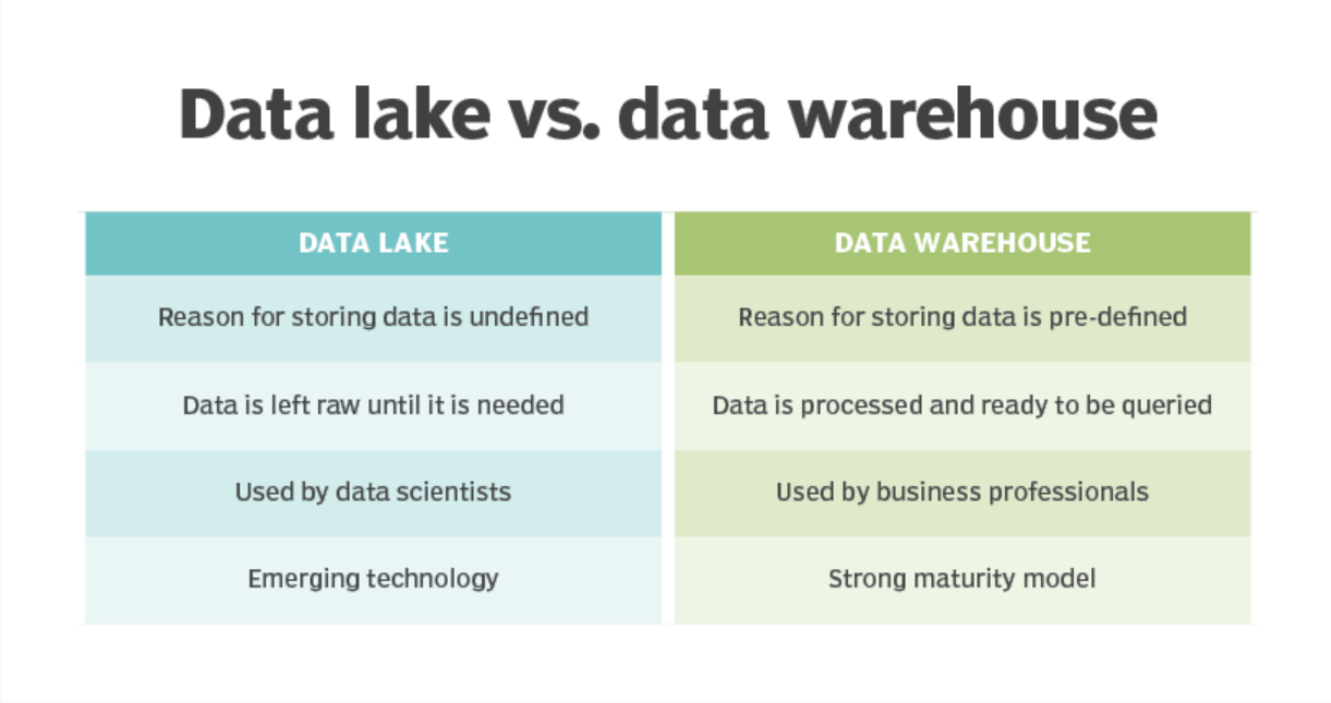

Figure 1. Lake - Warehouse

Reference: https://searchaws.techtarget.com/definition/data-lake

1.2. Slowly Changing Dimensions

Very simply, there are 6 types of Slowly Changing Dimension that are commonly used, they are as follows:

- Type 0 – Fixed Dimension No changes allowed, dimension never changes
- Type 1 – No History Update record directly, there is no record of historical values, only current state
- Type 2 – Row Versioning Track changes as version records with current flag & active dates and

other metadata

- Type 3 – Previous Value column Track change to a specific attribute, add a column to show the previous value, which is updated as further changes occur
- Type 4 – History Table Show current value in dimension table but track all changes in separate table
- Type 6 – Hybrid SCD Utilise techniques from SCD Types 1, 2 and 3 to track change
- Type 7 - see [Type 1 and 2 Dimension]

There is no `Type 5`

Reference: https://adatis.co.uk/introduction-to-slowly-changing-dimensions-scd-types/

1.3. Type 1 and 2 Dimensions

A fact table can be accessed through a dimension modeled both as a `type 1` dimension showing only the most current attribute values, or as a `type 2` dimension showing correct contemporary historical pro les. The same dimension table enables both perspectives. Both the durable key and primary surrogate key of the dimension are placed in the fact table.

For the `type 1` perspective, the current ag in the dimension is checked to be current, and the fact table is joined via the `durable key`. For the `type 2` perspective, the current ag is not checked, and the fact table is joined via the `surrogate primary key`. These two perspectives would be deployed as separate views to the BI applications.

Reference: https://www.kimballgroup.com/data-warehouse-business-intelligence-resources/kimball-techniques/dimensional-modeling-techniques/type-7/ window='_blank'

1.4. Wherescape

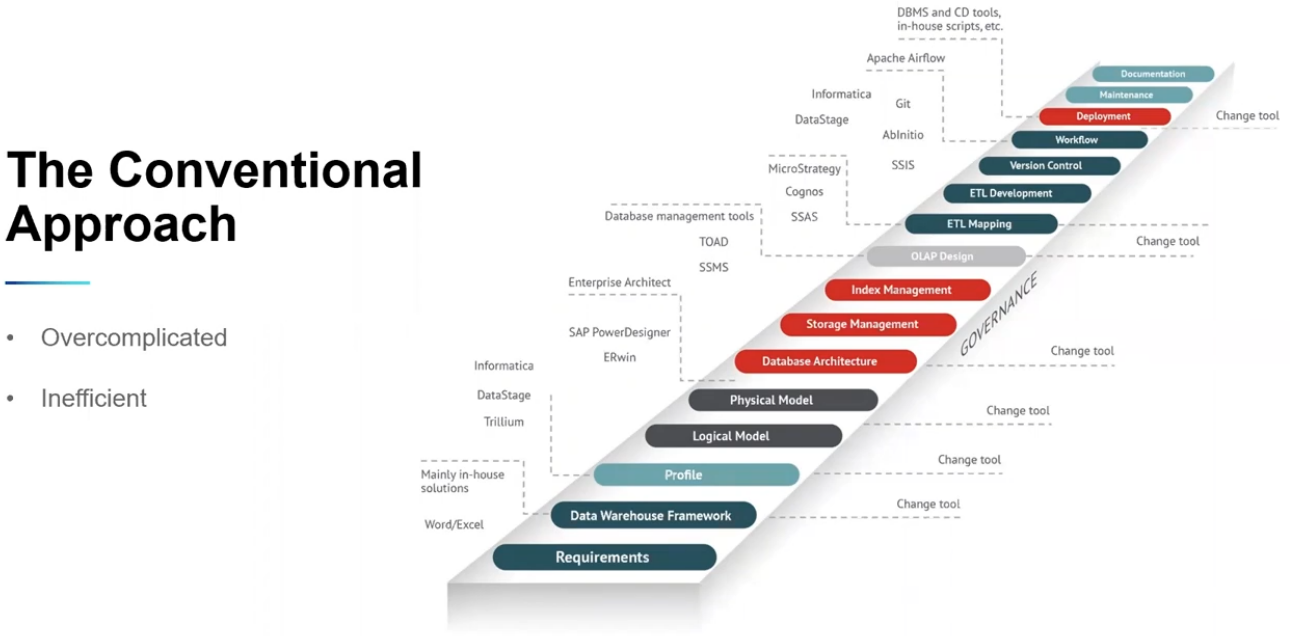

Figure 2. Wherescape DW Steps

Figure 3. What is Wherescape

This video segment from a Wherescape presentation, just show the overview of their product.

[] | *Wherescape Overview*

Wherescape Overview

1.4.1. 3D

WhereScape 3D makes planning, modeling, designing and prototyping data warehouses, data vaults, data lakes and data marts easy, fast and effective.

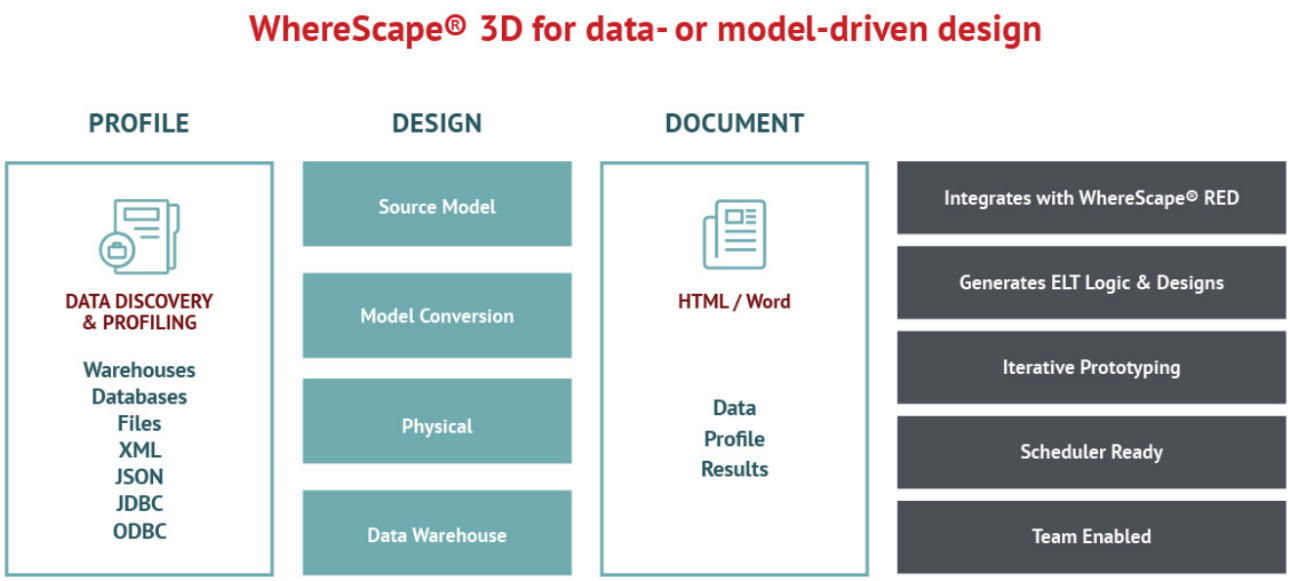

Figure 4. Wherescape 3D

1.4.2. Red

Whether on-premises or in the cloud, deliver data warehouses, data vaults, data lakes and data marts quickly with the WhereScape RED integrated development environment and integration tools.

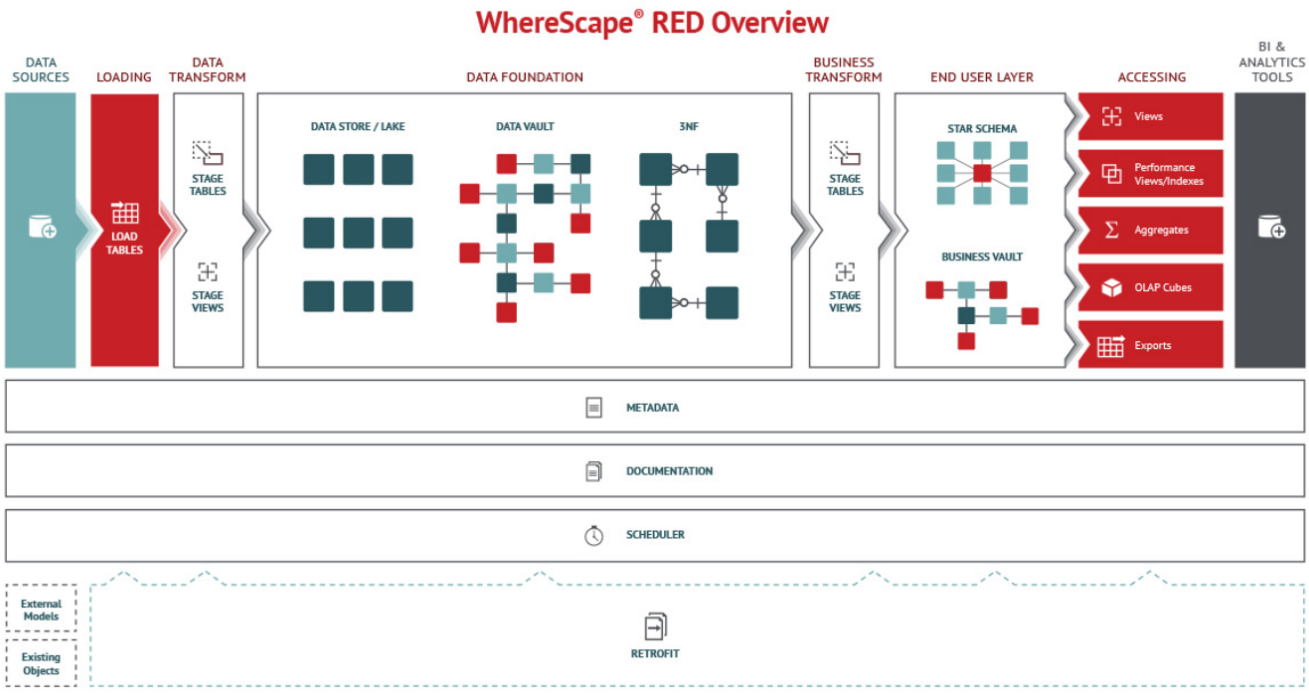

Figure 5. Wherescape Red

2. Document History

Table 1. Document History

Date	Version	Author	Description
12/13/2019	V2.1b	JHRS	Initial version