

# Data Architecture and Modeling with DataWharf

About me:

Eric Lendvai - Lead Software Development Engineer at Avalara Seattle

<https://www.linkedin.com/in/ericlendvai/>

About DataWharf:

[datawharf.org](https://datawharf.org)

<https://github.com/EricLendvai/DataWharf>

# Data Architecture and Modeling with DataWharf

- There are many definitions for Data Architecture and Data Modeling. Major evolution occurred in the domain of data management these last few decades, and so did the definition of those terms.
- Starting in 1970 the new concept of data normalization was introduced and refined in the following years.
- Normalization is a technique used to reduce data redundancy and improve data integrity by organizing data into tables and establishing relationships between them.

# Data Architecture and Modeling with DataWharf

- Data architecture is a high-level approach that deals with the overall design of data structures, systems, and processes. It defines the standards, policies, and practices that govern the collection, storage, retrieval, and use of data in an organization. Data architecture aims to ensure that data is managed in a consistent and coherent manner across the organization.
- Data modeling, on the other hand, is a more detailed approach that focuses on creating specific representations of data structures and relationships. It involves the creation of data models, which are diagrams or visual representations of the data elements, their attributes, and the relationships between them. Data modeling helps to clarify the meaning and relationships of data elements, and enables the creation of more accurate and efficient databases.

# Data Architecture and Modeling with DataWharf

- **Conceptual Data Model:** A conceptual data model represents high-level business concepts and their relationships, without getting into the technical details of implementation. It is a conceptualization of the real-world entities and relationships between them. It describes the overall structure of the data, including entities, attributes, and relationships between them, without going into too much detail about the database structure.
- **Logical Data Model:** A logical data model is a detailed representation of the data and how it is organized in a particular database system. It defines the structure of data elements and their relationships, such as entities, attributes, and relationships between them, and also includes data constraints and business rules. It is a representation of data that is independent of any particular database management system or hardware implementation.
- **Physical Data Model:** A physical data model specifies the physical implementation of a database on a particular hardware platform, including storage structures, access methods, and security constraints. It represents the actual data structure and how it is stored in the database management system. It includes details such as table structures, indexes, keys, and data types. The physical data model is optimized for performance and storage efficiency, and is specific to a particular database management system and hardware platform.

# Data Architecture and Modeling with DataWharf

- A Project in DataWharf, is collection of Models, conceptual and/or logical. Projects could be used before applications even exists. Usually a team of architects would be creating one or more models to design an entire project.
- The Models are made of Entities, Associations, Attributes, Packages, Datatypes, ... and many of those concepts can be renamed at the level of a project.
- Custom access rights can be managed at the level of projects, and most components of a model can have an unlimited number of custom fields as well.
- An Application in DataWharf, is in between a logical and physical model for a single database, represented as an actual Data Dictionary.
- But unlike most data dictionaries, the ones defined in DataWharf are backend independent, hence the "logical model like" behavior.

# Data Architecture and Modeling with DataWharf

- Using DataWharf to document itself – Physical Model

Application: DataWharf Other Applications

Tables (47) Enumerations (7) Name Spaces (1) Tags (5) Data Dictionary Settings Import Export Delta/Load Schema Visualize

Save Layout Diagram Settings All Tables New Diagram Duplicate My Settings Copy Diagram Link To Clipboard Tables: 47 - Links: 74

Undo Redo Zoom In Fit Zoom Out Reroute Edges

Table: public.Table

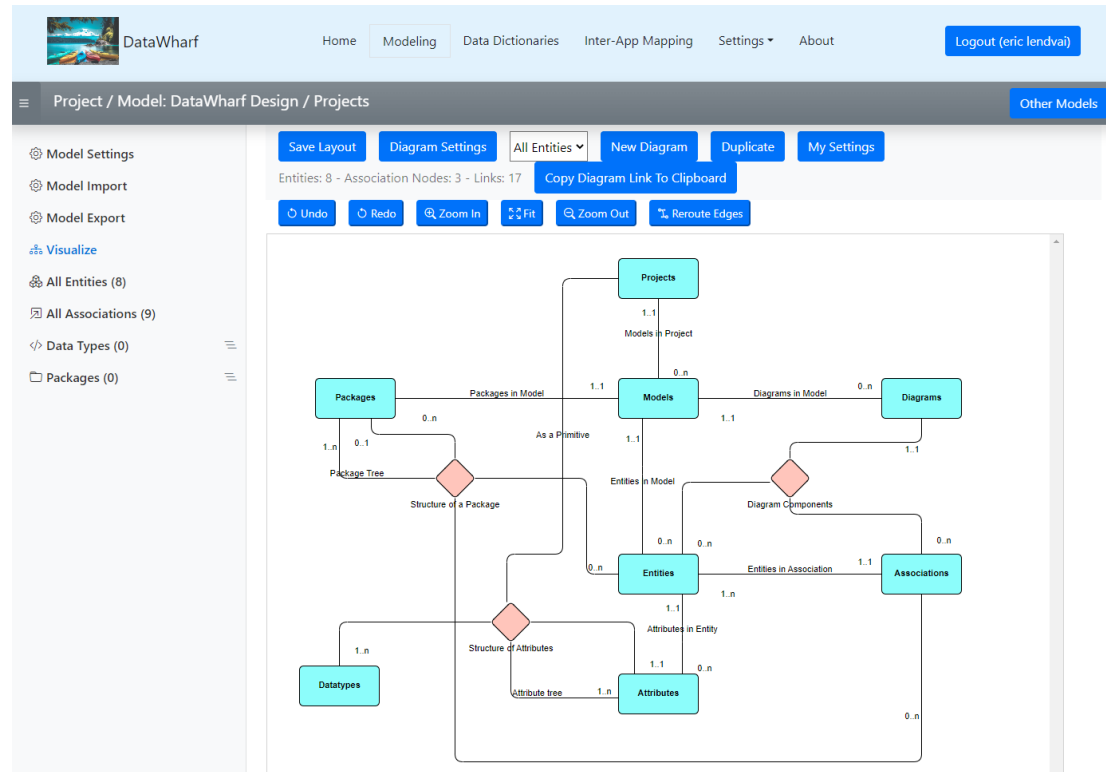
Columns (10) Related Tables In App (8) Other Diagrams (1) Table Info

Filter on Column Name (Press Enter) All Core Only

Name	Type	Nullable	Required	Default	Foreign Key To	Description
pk	I Integer (4 bytes)					
sysc	DT Date and Time Without Time Zone Conversion	✓				
sysm	DT Date and Time Without Time Zone Conversion	✓				
fk_NameSpace	I Integer (4 bytes)				publicNameSpace	
Name	CV Character String Varying (200) Unicode					
AKA	CV Character String Varying (200) Unicode	✓				
UseStatus	E Enumeration (UseStatus - Numeric 1 digit)			1		
DocStatus	E Enumeration (DocStatus - Numeric 1 digit)			1		
Description	M Memo / Long Text Unicode	✓				

# Data Architecture and Modeling with DataWharf

- Using DataWharf to document itself via a Conceptual Model.
- Most elements can be renamed.
- As in Application diagram additional information can be displayed on side panel.



# What makes DataWharf different from other modeling tools?

- DataWharf is 100% open-source, free, and is a web application that can be accessed using any modern web browser like Firefox, Chrome, Safari and any other browser supported by Bootstrap 5+ and jQuery 3+.
- Most entities can have an unlimited number of custom fields.
- Multiple tags can be used to flag Tables and Columns, making it even easier to filter down lists.
- Multi text searches can be done on description fields.
- The concept of life cycle and documentation completion exists on most entities.
- An unlimited number of Diagrams/Visualization can be done for Models and Data Dictionaries. Those diagrams directly related to elements entered in the Models and Data Dictionaries.
- Diagrams are living documents with color coding, tooltips and a detail panel that reflects additional information about any highlighted elements.
- All information in DataWharf has a unique URL, allowing user to include those in any external documentation.
- Easily compare your physical model to actual deployments, or sync/load already existing databases.
- For Harbour developers, an export to a Hash Array ready for use by the Harbour\_ORM is available.



# DataWharf Enterprise Features

- DataWharf can be configured for an unlimited number of users. Project and Application level access rights can be configured for each user.
- DataWharf can also be configured to register new users via Single Sign On.
- Users with the proper access rights, can export and import any Models and Data Dictionary.
- DataWharf has an API engine that can be enhanced to support any external application.
- Except for passwords, all information stored by DataWharf could be used by accessing its own PostgreSQL database (access rights restrictions to be set by your own DBA).
- DataWharf is also compatible with AWS RDS services.
- Support for CyanAudit can be enabled to track any changes made by any users.

# Data Architecture and Modeling with DataWharf



DataWharf

Home

Modeling

Data Dictionaries

Inter-App Mapping

Settings ▾

About

Logout (eric lendvai)

User: eric lendvai

Login Log

Login Time	IP
------------	----

Time for a  
live Demo

