

13 tips to make your Data Vault project fail

Or how to successfully destroy every level of the data warehouse

Are you pampered with success and tired of the eternal pat on the back? You don't want to be so successful with your first attempt at implementing a Data Vault project that all your colleagues become jealous?

Here you will find 13 proven tips on how to make your Data Vault project a successful failure. You can let yourself be inspired whether you prefer to start at the organizational level right at the start of the project, tackle your logical data model incorrectly or just want to patch up the technical implementation shortly before the end of the project.

Short-term options are considered, such as the wrong use of resources to rebuild automation tools available on the market, or the ignoring of customers so that they give up any hope of a result before the project has even started properly.

In order to sustainably secure the goal of successfully torpedoing the project, it is important not to lose focus and to block agility and maintainability right from the start.

If you have already launched a Data Vault project, this list can help you to launch a review process: Are you already on the right track to blow up your project? The following 13 tips will give you some more options.

1) Do without knowledge bases and expert opinions!

Roll up your sleeves and start your project confidently! Don't waste time on training or seminars! Just because other people made mistakes during their first Data Vault implementation doesn't mean the same thing could happen to them. Don't let literature [LO15, Hul12] and external support unsettle you. You are the right person to revolutionize the Data Vault world on your own.

2) The Data Vault standard is only a rough guide!

Only the exception confirms the rules. This should also be your approach when you implement Data Vault. Of course, the inventors of the rules were not stupid, but of course they did not know your company. And your company is unique. Don't let anyone take that away from you. You've developed your own word processing and accounting software. This outstanding approach should also be implemented in your Data Vault project!

3) The understanding of subject matter is overrated!

You shouldn't put too much emphasis on the business in a Data Vault, because everyone always talks about the high flexibility of this system. Primary keys in the source system are guaranteed to be good keys for identifying an object in the real world. To understand this, all you have to do is go into your call center and listen in on conversations where customers introduce themselves to the call center agent as 0E984725-C51C-4BF4-9960-E1C80E27ABA0. So you can easily generate a data vault from source system key definitions without knowledge!

4) Conceptual modeling is a waste of time!

If you can't prevent business, don't spend time defining your business objects [SG04]. How likely is it that someone in your company doesn't know exactly how a customer is defined? In good companies, which include your employer, marketing, sales and financial accounting share an identical view. And while others still paint boxes and lines in endlessly long meetings, you've already made the first implementations.

5) Do without the support of your BI department!

You may also be annoyed by the many wrong decisions made by your BI department management. How did you deal with it in the past? Of course, you applied your own methods as well as possible. That's exactly how you should handle the Data Vault. Whatever aspect of the project you're looking at - the team composition, the provision of tools, or the project process: You know what you need, and a well-planned guerrilla tactic will lead you to your goal.

6) Always think in whole layers!

First create an Enterprise Data Model! This will give you a few months rest from your internal customers. They cannot develop enthusiasm for the project until you have completed the perfect data warehouse. Build the data model layer by layer with all conceivable objects. So, no one can see any problems during the two to three-year development period. And should one of your colleagues use one of the terms "puncture", "vertical" or even "agile" [Gra15]: Mob him out of the team! Such revolutionary troublemakers can sabotage your entire project.

7) To do things in a big way

Don't let them keep you down [Ler15]! Or as the Americans say: Think Big! Of course, your users say you just want to see the correct stock levels first. But what is the use of correct stock levels without Big Data [WB13] and Predictive Analytics [Sie16] based on artificial intelligence? In addition: If you solve the complex problems first, then delivering the basic data later is a minor matter.

8) Too much automation dilutes the quality!

Trust your skills and build the Data Vault by hand! Of course, there are optimized tools on the market, but good old craftsmanship has its value. The pyramids were also built by hand, and they still stand today. Besides, you completely lose the understanding of the database structures when all hubs, links and satellites are magically generated for you. Be clear in your communication: The necessary effort is primarily a problem for the finance department. Your job is to create a unique work of art that will last.

9) Single Point of Facts? Unnecessary raw data and redundancy!

Don't let Data Vault confuse you! Over the past 30 years, we've changed the data before we saved it. Your Op Specs (authors' note: Network to the professionals and specialists of the operational systems) are trained in deleting data and reloading the last 700 files with changed business rules. This knowledge would be lost if you first historicalize the unchanged data in the raw vault and then modify it in the business vault.

10) Never use hubs for transactions!

Always model your process data exclusively as links! This prevents you from having to evaluate questions about relationships between your business processes as you cannot do it technically. Communicate to your users that they must please refrain from business transactions with several process steps. The primary goal of your data warehouse is to optimize your business model.

11) Save yourself tracking satellites!

There are companies where data changes after the first capture. But yours is unlikely to be one of them. If a customer has registered, he should always keep the same subscription. All you have to do is communicate to marketing how important good customer loyalty is for the data warehouse. So you can confidently save on tracking satellites, which makes your Data Vault much easier.

12) Allow access only to the Business Vault and never to the Raw Vault!

Separate Raw Vault and Business Vault [Eve13] hermetically and allow access only to the latter! Data is either unprocessed or processed. Users should only continue working with edited data. Combining this data would allow your users to validate whether business rules have been applied correctly. This is an unnecessary source of potential incidents that should be avoided from the start.

13) No raw data for the department!

Do not deliver raw data to your users until they have assumed full responsibility for the business rules! This ensures that the question of guilt is clear from the outset if the reports do not contain reasonably calculated values.

Nevertheless, be indulgent and allow for a change phase after the project has been completed. Write down the further requirements in advance to avoid misunderstandings.

Bibliographical reference

- [LO15] Linstedt, D. und Olschimke, M.: *Building a scalable data warehouse with data vault 2.0*. Morgan Kaufmann, 2015.
- [Hul12] Hultgren, H.: *Modeling the agile data warehouse with data vault*. New Hamilton, 2012.
- [SG04] Simson, G. / Witt, G.: *Data modeling essentials*. Elsevier, 2004.
- [Gra15] Graziano, K.: *Better Data Modeling: An Introduction to Agile Data Engineering Using Data Vault 2.0*. Data Warrior LLC, 2015.
- [Ler15] Lerner, D.: *Data Vault KISS – Keep It Small and Simple*. <https://tedamoh.com/data-model-menu/data-vault-menu/61-data-vault-kiss-keep-it-small-and-simple>, 2015, abgerufen am 26.02.2018.
- [WB13] Ward, J. / Barker, A.: *Undefined by data: a survey of big data definitions*. arXiv preprint arXiv:1309.5821, 2013.P
- [Sie16] Siegel, E.: *Predictive analytics: The power to predict who will click, buy, lie, or die*. John Wiley & Sons Incorporated, 2016.
- [Eve13] Evers, M.: *The Rule & the Raw & the Business (Data) Vault: Data Vault Architecture and Business Rules*. <http://dm-unseen.blogspot.de/2013/04/the-rule-raw-business-data-vault-data.html>, 2013, abgerufen am 26.02.2018.

About the authors

PETR BELES ist Senior Consultant und Product Owner des Datavault Builders bei der 2150 GmbH. Nachdem er lange Zeit Data Warehouses mit Dimensional Modeling und dritter Normalform mitentwickelt hat, modelliert er nun seit mehr als fünf Jahren ausschließlich Data Vault.
E-Mail: petr.beles@datavault-builder.com

PEER M. CARLSON ist Managing Consultant Business Intelligence bei Heger Consulting. Er berät Unternehmen aus den verschiedensten Branchen zu Themen wie Data Vault, Datenmodellierung und mehrdimensionale Modellierung sowie Agilität und Scrum und ist auf verschiedenen Konferenzen als Sprecher anzutreffen.
E-Mail: pmc@heger-consulting.com

DIRK LERNER is an experienced independent consultant and owner of TEDAMOH. For 15 years he has led BI projects and is considered a global expert on BI architectures and data modeling. As a full-scale-data-architect, Dirk advocates flexible, lean and easily extendable data warehouse principles and practices.

As a pioneer for Data Vault and FCO-IM in Germany he wrote various publications, is an international speaker at conferences and author of the blog <https://tedamoh.com/blog>.
E-Mail: Dirk.Lerner@tedamoh.com