

Data Analysis for Business Analysts: The Zachman Framework and Data Architecture

Featured

32124 Views

5 Comments

17 Likes

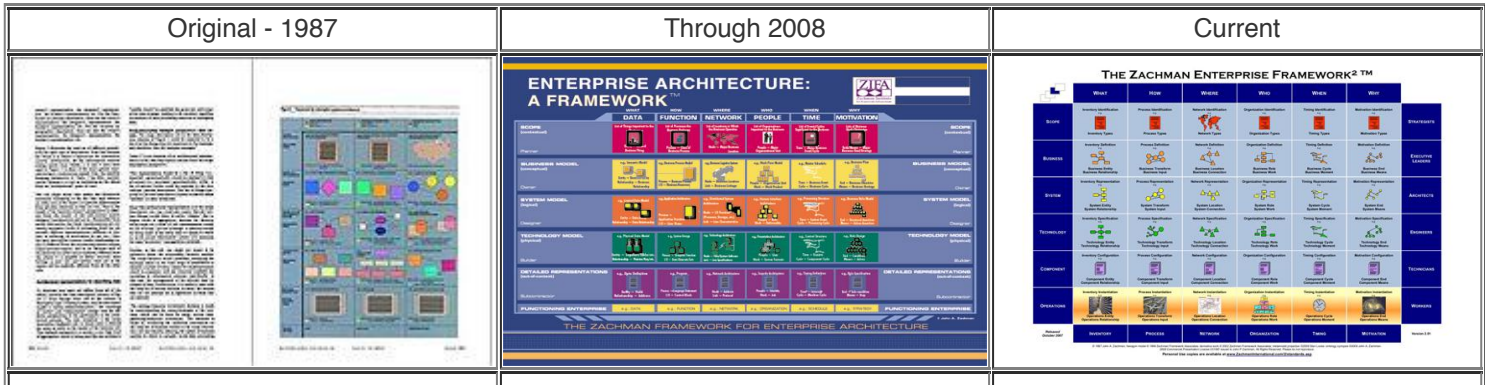
BACK TO LIST

When ModernAnalyst asked me to propose an article for their issue on Enterprise Architecture, I thought about the question framework developed by John Zachman, that provides the basic building blocks of that practice. The primary function of a Business Analyst is to ask questions that uncover requirements then to document those requirements so they may be developed into a useful, useable system. It seemed to me that questions are a common thread for the Enterprise Architect, Data Architect, and Business Analyst alike. As the conversation progressed, I realized that I could reach out to John Zachman and pose a few questions to him about Enterprise Architecture. This article will be split between a brief interview, Q&A with John Zachman, and a more detailed overview, Data Architecture.

Background

If you google “father of enterprise architecture” the results will point to John Zachman. In an article published in the IBM Systems Journal, he proposed a way of thinking about building information systems that would combine the terminology of architecture with the fundamental portions of an applications system, data, process, and network. He created a matrix and from that initial grid in 1987, the practice of enterprise architecture was born. In the years since the initial article, the framework has grown from the initial three columns to the six basic questions that can be posed about any topic.

Evolution of Zachman Framework™



Q & A with John Zachman:

In August of 2008, I worked with John Zachman and a team of ICCP Certification Counsel Members on the new Zachman Certification exam. A group of us met in Montreal to review the updated Zachman Framework™, each bringing a special perspective; I was one of the people at the table representing the data point of view.

John has had a strong relationship with the data management community since the 1980. He currently acts as the Advisor Emeritus to the Board of Directors for DAMA International, and is writing the preface to the DAMA-DMBOK.

LORETTA: Good to talk to you John, I hope all is well!

JOHN: Loretta ... it is always a delight to talk with you!!

LORETTA: We have a great opportunity here to inform Business Analysts about the framework. I am pointing them to your website for definition and detail; I'm hoping that you can provide some background that will help them understand your perspective and the evolution of Enterprise Architecture as a practice.

JOHN: We have been adding a lot of substantive material about Enterprise Architecture and the Zachman Framework to the Zachman International website and intend to make it a comprehensive resource for anyone interested in the subject. You will presently find several articles that I have written including the "John Zachman's Concise Definition of the Zachman Framework". I wrote that article because when I would correct the Wikipedia entry on the Zachman Framework to accurately reflect its definition, someone would spuriously modify it and desecrate the concept. I finally got frustrated and wrote my own definition on my own website for anyone who is interested in a factual definition.

LORETTA: Could you please share your definition of an "enterprise"?

JOHN: The Framework is inert relative to the definition of "Enterprise". It is the Architect that sets the boundary of the analysis. That is, the Framework can be used to help one think about, analyze, describe, architect an Enterprise, whatever the Architect's definition of "Enterprise" might be. For example, one might say, "in my area of interest, the word 'Enterprise' means one SBU (Strategic Business Unit)", that is, one business unit that typically has a general manager and a set of resources assigned to accomplish some mission or set of objectives. In contrast, one might say, "to me, the word 'Enterprise' means a group of SBU's, a cluster, a conglomerate, a federation ... or, the whole 'Value Chain'". Or, one could choose to define the word "Enterprise" more narrowly for example, "one Department within the SBU" or "one project within the Department", etc. All I am doing is changing the granularity of the analysis. In essence, the more broadly you define the boundary of the analysis, the better leverage you are going to get on integration, reusability, interoperability, flexibility and so on of the end object BUT, the more complex the analysis is going to be. Conversely, the more narrowly you define the boundary of the analysis, the simpler the analysis BUT, the LESS leverage you are going to get on integration, flexibility, reusability, interoperability, and so on of the end object.

LORETTA: Now could you explain why you chose "architecture" as the analogy or metaphor for IT Systems Development?

JOHN: I actually came from the Strategy community and by the late '60's we had figured out methodologically how to define business strategy. The problem was, even though we could define the business strategy, we could not figure out how to get the implementations, the instantiations, aligned with the strategy. My idea was to ask some of the older disciplines that produced complex products (implementations) how they managed to realize quality, what we call "alignment," which to this day remains the most critical objective of general management. The CEO surveys for the last 15 years had "alignment" in the top 10 objectives they identified for the IT community. In fact, in the last 3 years, alignment has been THE TOP ISSUE in these surveys.

I had a friend of mine that was an Architect that built hundred story buildings plus I lived in the Los Angeles area where there were a lot of airplane manufacturers and it was easy for me to ask the Architects and the Engineers how they produced complex products that were aligned with the strategy for the products. It turned out that for any (or all) complex products there is a structure, a set of descriptive representations that are relevant for creating the complex product such that it is consistent with the intent of the end owner of the product and that are necessary for changing the product once it is instantiated. That set of descriptive representations is the Product Architecture. Architecture is architecture is architecture. It is the same set of descriptive representations regardless of whether the product is an airplane, a building, a locomotive, a battleship, a computer or a bathtub. My reasoning was that it had to be the same set of descriptive representations relevant for describing an Enterprise. It would be unreasonable to presume that the descriptions of an Enterprise are going to be any different from the descriptions of everything else that humanity has ever described.

LORETTA: Can you please explain a bit about the difference between your framework and a methodology?

JOHN: My Framework is an ONTOLOGY. It is a theory of the existence of a structured set of elemental components of a complex object that require description and are relevant for creating, operating, maintaining and changing the object. It DEFINES Architecture. My Framework implies nothing about how one might DO architecture work. That is the domain of methodology. My Framework is analogous to the Periodic Table. It defines the elemental components that exist. A methodology is a process that transforms some elements or compounds into different elements or compounds. An Ontology is NOT an methodology and a methodology is NOT an Ontology.

LORETTA: Do you think having an accurate Enterprise Architecture will help Business Analysts do their job more easily?

JOHN: Actually, I would say that an accurate Enterprise Architecture would ENABLE Business Analysts to do their job. A process (methodology) based on an ontology will be repeatable and produce predictable results, that is, an ontology is the basis for a science. In contrast, a process (methodology) that is not based on ontology is ad hoc, dependent on practitioner skills and experience, is not a science and is analogous to alchemy. The end result may reflect what is intended ... and it may not.

LORETTA: Can Business Analysts play role in the creation or maintenance of an Enterprise Architecture Framework?

JOHN: I would say that Business Analysts play a KEY role in populating some of the descriptive representations that constitute Enterprise Architecture, namely, those descriptive representations that reflect what the Executive Leaders of the Enterprise have in mind as to how they want the Enterprise to be instantiated. If the Executive Leader's intentions are not accurately transcribed, there is NO WAY that implementations are going to reflect what they are thinking about except by sheer luck or the Business Analysts making some very accurate assumptions! The Business Analysts may well also contribute to transforming the Business Concepts as transcribed for the Executive Leaders into Logic Systems that realize the Business concepts.

LORETTA: Will you explain your reasoning for using the interrogatives as the columns?

JOHN: EVERY complex object has a bill-of-materials describing WHAT the object is made of, functional specifications that describe HOW the object works, geometry (or drawings) that describe WHERE the parts are relative to one another, operating instructions that describe WHO is responsible for operations, timing diagrams that describe WHEN things happen and engineering design objectives that describe WHY they happen. A complete description of ANYTHING requires answers to six "primitive interrogatives", what, how, where, who, when and why. The linguists would suggest that this has been evident from the very origins of language.

LORETTA: Which Framework classifications (aka Cells) would be most relevant for the Business analysts and why?

JOHN: Someone has to transcribe what the Executive Leaders of the Enterprise have in mind as to the boundaries of the Enterprise relative to the Inventories they control, the Processes they perform, the Locations they operate in, the Organizations they can allocate responsibilities to, the timing Cycles they manage and the Objectives they have as well as the relationships between these Row 1 lists. I presume that the Business Analysts would be involved in this. Also, it is likely that the Business Analysts would participate in defining the Semantic models that define the Business Concepts as contained in the Row 2 Cells. Further, the Business Analysts may well participate in transforming the Row 2 Cell Models into Row 3 System Logic Models.

LORETTA: How can the Business Analysts use the primitive interrogatives (what, how, where, who, when and why) when performing enterprise analysis?

JOHN: To have a complete description of the Enterprise, all of the primitive interrogatives will have to be answered at some point in time ... probably over some period of time, maybe on an on-going basis as a different way of life. If it is important to the Enterprise that it is integrated, flexible, reusable, interoperable, aligned, instantiated dynamically, etc ... then each of the primitive interrogatives as expressed in the Cell models will have to be populated and designed independently from one another ... only binding them together for implementation purposes as demanded ... in realization of the concept of "late binding". This would constitute ENGINEERING the Enterprise.

Note ... you don't have to do any of this to build and run systems. Building and running systems is manufacturing the Enterprise, NOT engineering the Enterprise. It was Fred Brooks that said, "Programming is manufacturing, not engineering." For the last 50 years, those of us that come from the information community have been manufacturing the Enterprise BEFORE it was engineered. That is why the Enterprises of today (as manifest in the legacy systems) are NOT integrated, NOT flexible, NOT interoperable, NOT reusable, NOT aligned, NOT dynamically instantiated, etc., etc. The Enterprises were never engineered to be so. They were manufactured. If it is important that the Enterprise have the engineering design characteristics of flexibility, integration, alignment, etc., etc, (and I am going to suggest that it is going to be IMPERATIVE that every Enterprise of the Information Age BE integrated, BE flexible, BE reusable, BE interoperable, BE aligned, BE dynamically instantiated, etc, etc.) then somebody is going to have to do the engineering work that produces those results. If we only continue to build and run systems, then "keeping on doing the same thing expecting different results is one definition of insanity" (Einstein). Further, if we only continue to build and run systems and ignore engineering the Enterprise, then I am going to suggest that Nicholas Carr is right, that "IT Doesn't Matter" (Harvard Business Review) because we have proved conclusively over the last 10 or 15 years that ANYBODY can build and run systems. Building and running systems is NOT the issue. Engineering and Manufacturing the Enterprise, that is ENTERPRISE ARCHITECTURE, IS the issue.

I think the Business Analyst plays a key role.

LORETTA: Thank you so much for taking the time to share with us!

JOHN: Loretta ... bless your heart!! It is always a delight!! Any time!!

Data Architecture

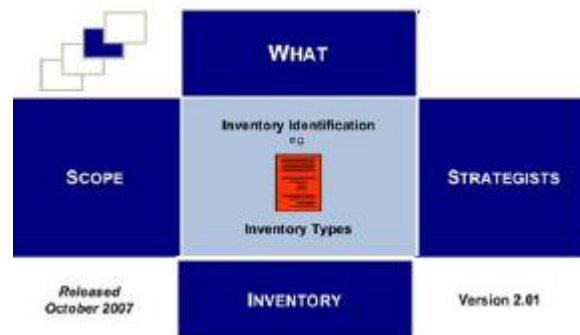
Using the **Zachman Framework™**, let's take a deeper dive on the **What**, or data column. Each of the rows in the framework expresses a level of abstraction, in Zachman's own words, "Identification, Definition, Representation, Specification, Configuration, and Instantiation".

At first, this may seem really academic, but I've found **in practice** that it aligns closely to the project lifecycle and best practices in developing a data design from a top-down perspective for new systems. At the same time, if I am working on a migration or upgrade project, looking at the existing system and working from the bottom-up, the framework has helped me uncover data related issues and constraints that might impact scrubbing or mapping required for legacy data. The value that I have found in using the framework is it provides me with a tool to help me perform gap analysis and risk assessment while I am developing a data design. The questions can be asked at any level in an organization, from single application system to the entire enterprise!

For the purposes of this article, I'm going to share how I use the framework to develop my data related deliverables. Two problem domains are typically used in data

modeling courses “THE VIDEO STORE” and “TRAINING” to provide specific examples. Hopefully by using these very simplified examples, I can illustrate how I use the framework in an actionable way. Feel free to fill in the name of your current project, and then reflect on whether you have something that addresses this portion of the schema. If not, consider whether having it would help you produce a higher quality and more maintainable final product.

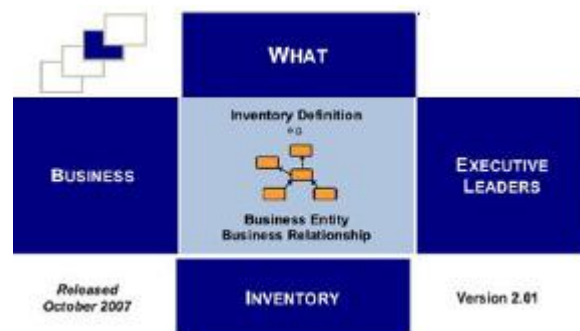
There are no hard-and-fast rules about how the architecture deliverables are documented until a particular methodology is applied, but the meaning and purpose of each of the six rows are necessary to move from idea to actual physical instance of data inventory. I am not advocating a specific methodology, just explaining how I use the Zachman Architecture to help me organize my thoughts in order to create a more comprehensive design.



Identification level

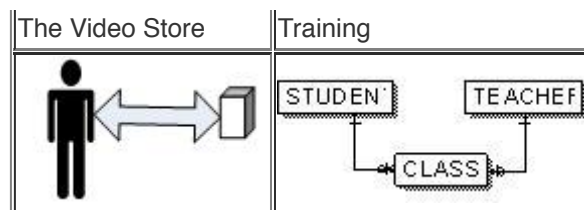
The most abstract level of architecture has the least detail. It is just the bare minimum that might be useful in understanding the scope, describing the data inventory in a business perspective. For me, that is a list of nouns that identify the most important groupings of data in the scope. This list will be expanded and refined at different points in the design process, but it is just enough to explain at a strategic level the WHAT of the system. If this list is combined with artifacts of the other enterprise architecture domains, it informs my ability to go create deliverables at the next level.

The Video Store	Training
<ul style="list-style-type: none"> Customer Video Rental 	<ul style="list-style-type: none"> Teacher Student Class

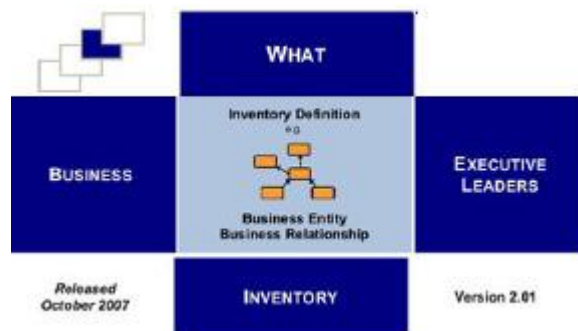


Definition level

A slightly more detailed level would show those major groupings of data and how they relate to each other. My deliverable would provide a graphic to illustrate the definition, either using picture icons, or a conceptual data model, depending on the audience preferences. (Note: When JZ reviewed the outline for this portion of the article, he commented “...these are business policies that govern the inventories (when you put something in inventory and when you take it out). For example: You can't hire an employee without a position to put them in- THAT is a business policy.”)

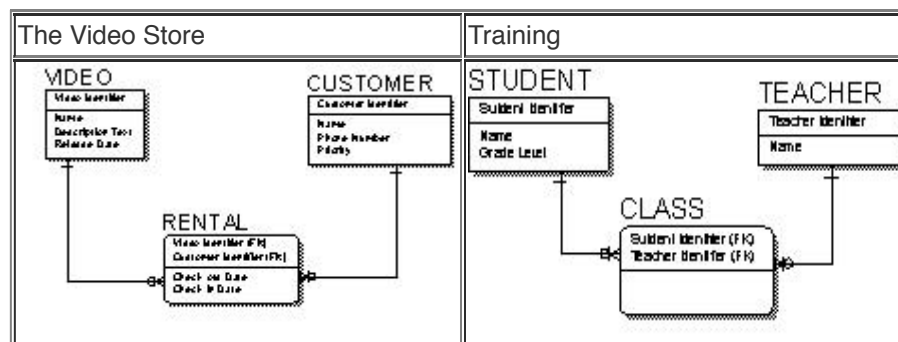


Representation level

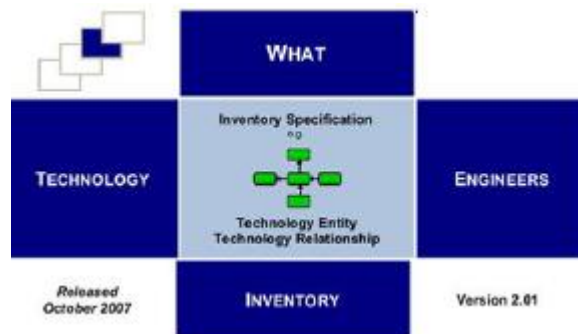


I use the definition level inventory along with information from the other EA domains and apply standard logical data modeling techniques to create my the next level deliverable - a logical data model. For me, that is a fully attributed entity-relationship diagram or dimensional model. This picture has the ability to show the inventory of data without being limited by physical implementation technology decisions. A logical data model will always be more stable after it has been cross-referenced to the other model for cells at this level in the framework. If the data design moves ahead before this cross-reference occurs, there is a higher likelihood of re-design required due to gaps in

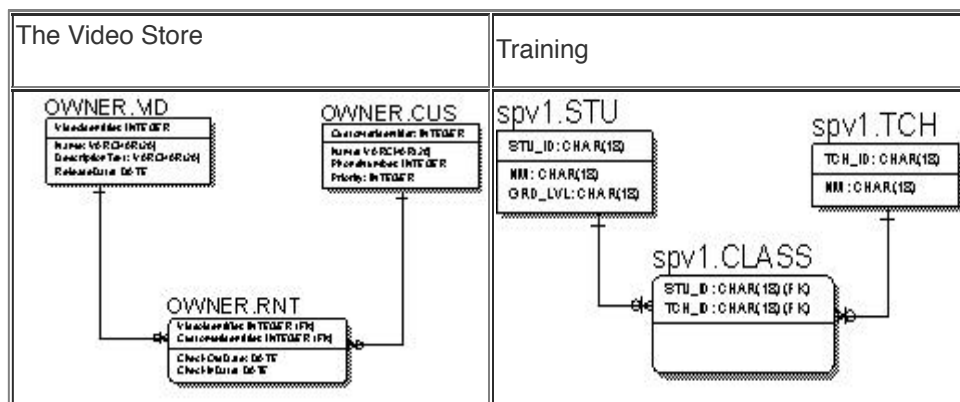
requirement satisfaction.



Specification level

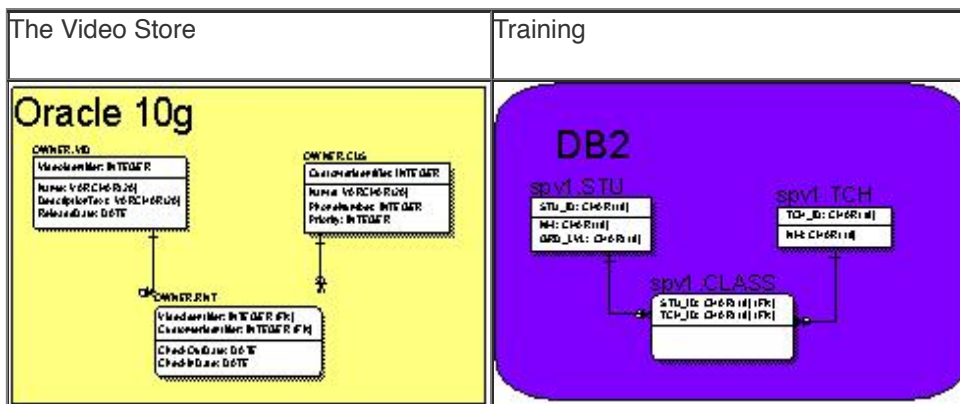
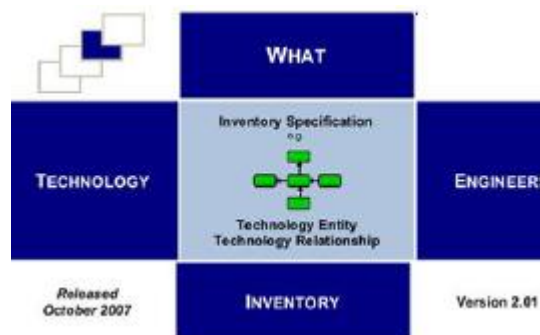


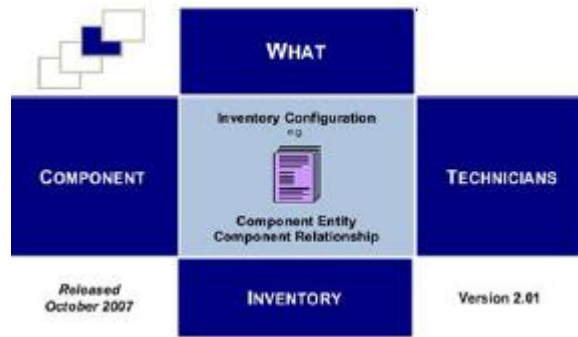
Once the logical design is documented, then the technology infrastructure constraints must be applied to create a physical data model. When I create a physical design, I must take into account the influence of decisions from all of the domains in the enterprise architecture and create a diagram that represents exactly what will be built. I can also show choices made to take advantage of special features and functions of a technology. This model illustrates to programmers and other technicians exactly what they can expect once the design is implemented.



Configuration level

A physical design, even a technology specific one, could actually be implemented in multiple physical locations at the same time for a variety of reasons. Whether due to performance requirements, backup, disaster recovery, or security, each of the separate implementation should be a managed and supported artifact in the data inventory at the configuration level.





Instantiation level

This is not a representation or a deliverable; it is the real-world instances of data that are created to represent the actions, products, and services that occur over time. It is the inventory of data in production systems. This is where production monitoring & tuning, data quality assurance, archiving, security and other data management functions are applied.

I use the Zachman Framework™ to help me organize my thoughts around the management of our data inventory. It helps me gain perspective to helicopter up, as well as when I do the deep-dive into a legacy system. These examples may not be what your methodology requires, but regardless of deliverable format, each viewpoint helps create a more complete picture of how the data inventory supports the business.

Does your enterprise document its existing enterprise architecture?

If so, identify the resources that you can leverage. If you can, describe the scope and deliverables for your projects in a way that is consistent with existing architectural standards of your organization.

Resources:

"A Framework for Information Systems Architecture", John A. Zachman, IBM Systems Journal, vol. 26, no. 3, 1987. IBM Publication G321-5298.

Zachman International: www.zachmaninternational.com

ZIFA: www.zifa.com

[ICCP Press Release for Zachman Certification Exam](#) – The text of the original press release for the Certified Data Management Professional Examination.

Author: "Loretta Mahon Smith, CCP, CBIP, CDMP, is Vice President for the Data Management Association - National Capital Region, a member of the Board of Directors for DAMA International, and an ICCP Certification Counsel Member. She is a 25 year career Data Architect in the financial services industry."