



Database Terminology

1. INTRODUCTION

The terminology around the database development life cycle is confusing. When describing data; different methodologies use different names to describe similar concepts at a different level of abstraction and from different points of view.

Historically, file systems to hold data came first. As more applications stored data, data base management systems (DBMS) such as IBM's IMS (hierarchical) and Cincom's TOTAL (network) became widely available.

Then, in 1970, Codd created the relational model. This provides a theoretical foundation for data retrieval. The relational data model showed that properly structured data was a resource that could be manipulated using standard functions. Relational DBMS's can be formally verified, because the relational DBMS implements the mathematical formalisms first described by Codd and extended by Date.

Using Entity / Relationship modeling (Chen) attempted to extend the relational model with business semantics and deal with the data at the conceptual level. The Object-Oriented programming approach to application development attempts to package data and functionality into a single concept.

2. DATABASE TERMS BY DEVELOPMENT PHASE

All of these methodologies attempted to differentiate themselves from each other and described data using different terms. The methodologies differ mainly in being aimed at different levels of abstractions and different audiences—user versus developer, theoretician versus practitioner, etc. These levels are equivalent to the levels of the Zachman Enterprise Framework. A good database engineer understands all of these methodologies since they all provide useful abstractions for designing databases.

How the different abstraction levels and terms are related to each other is shown in the table below:

Methodology Phase	DDLC Deliverables	Organizational Structure	Structural Element	Relationships between Elements	Element Instance	Element Component	Data Type	Assumptions	Reference
requirements	List of Business Data	relational algebra	relation	constraint	n-tuple	domain	not specified	domain cannot be a relation	Relational Model (Codd)
analysis	Conceptual (Kimball Star) Data Model	conceptual data model	entity	relationship	none	none	not available	entities cannot have overlapping meanings	Conceptual Data Model (Chen)
design	Logical (Semantic) Data Model	semantic data model	logical table	relationship	row	attribute	abstract (Gottlieb & Gottlieb)	Logical tables / attributes cannot have overlapping meanings	Logical Data Model (Simsion)
specification	Database Specification	database	table	foreign key	row	column	simple		Physical Database
implementation	Data Definition (DDL) Files	database	table	foreign key	row	column	simple		Implemented Database (Gray)

3. OBSERVATIONS

Examining the table, one can make the following observations:

- Codd's relational model is not a description of a database. It is equivalent to the Universe of Discourse (Simsion) of the business data requirements or a Data Lake.
- A relation in the relational model is not an entity.
- An entity is not a logical table.
- An entity is an abstract summary of a set of logical tables.
- A logical attribute is more complex than a column and better defined than a domain.

This shows the difficulty in following discussions about database development that use the term "entity or table" for all the phases of the development process.

4. REFERENCES

Chen, P., "The Entity-Relationship Model-Toward a Unified View of Data";
ACM Transactions on Database Systems, Vol. 1, No. 1, pp. 9-36, Mar 1976

Codd, Edgar, "A relational model for large shared databanks",
Communications of the ACM, vol 13, number 6, June 1970

Date, C.J., *The Relational Database Dictionary*, 1st Edition, O'Reilly Media,
Inc., Sebastopol, CA, 2006.

Gotlieb, C.C. and Gotlieb, Leo R., *Data Types and Structures*, Prentice-Hall,
Inc., Englewood Cliffs, New Jersey, 1978

Gray, J., "The Transaction Concept: Virtues and Limitations", Proceedings of
Seventh International Conference on Very Large Databases, Tandem
Computers, Inc., Cupertino, CA, Sep 1981

Kimball, R., *The Data Warehouse Toolkit*, John Wiley and Sons, Inc., New
York, NY, 1996

Simsion, Graeme C. and Graham C. Witt, *Data Modeling Essentials*, 3rd Ed.,
Morgan Kaufman Publishers, San Francisco, CA, 2005

Zachman, J.A. "[A framework for information systems architecture](#)", IBM Systems
Journal, Vol 26, No 3, 1987