O tempora, o mores – Zeitlinien im Data Warehouse

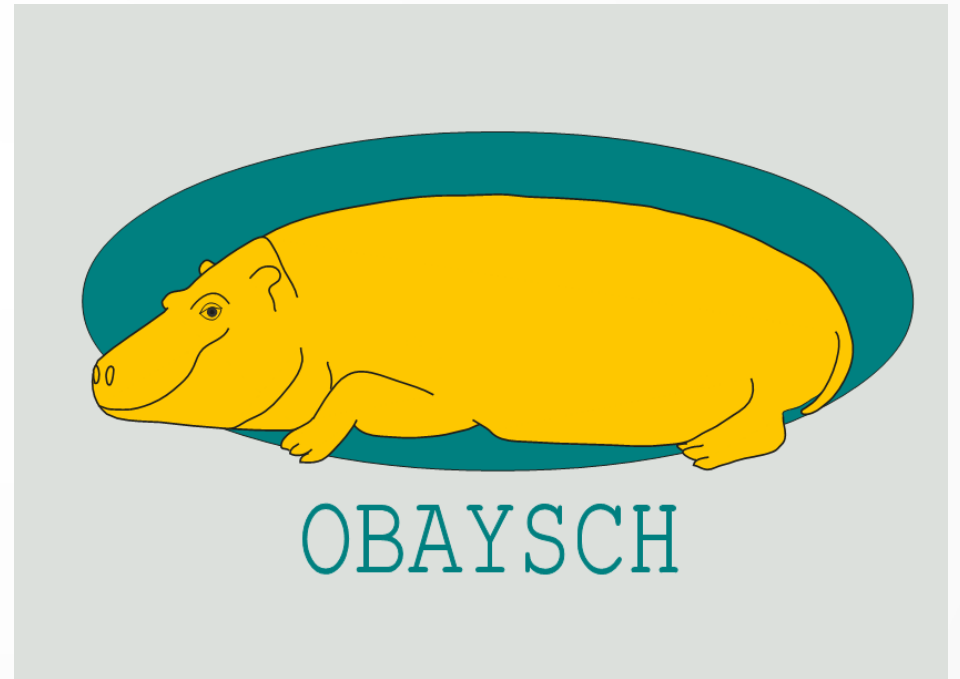# Kurze Vorstellung

Christian Kaul

Data Modeling Aficionado & Data Structure Designer bei Obaysch

München, Bayern, Deutschland

Fragt mich nach der Knowledge Gap 2023!

Email: christian.kaul@obaysch.net

https://www.linkedin.com/in/christian-kaul/

OBAYSCH

What Time Is It?

**Different Kinds of Time &
How to Deal with Them**

Content

- Who Has the Time?

- What to Do with Your Time

Who Has the Time?

A Brief History of Kinds of Time

Richard T. Snodgrass

- **Valid Time**

    "capturing the history of a changing reality"

    "when a fact was true in the modeled reality"

- **Transaction Time**

    "capturing the sequence of states of a changing table"

    "when a fact was stored in the database"

SQL:2011

- **Application Time**

  "a period whose period name is not SYSTEM_TIME is also known as an application-time period"

- **System Time**

  "a period whose period name is SYSTEM_TIME is also known as a system-time period"

C. J. Date, Hugh Darwen & Nikos Lorentzos

- **Stated Time**

   "the stated time—sometimes currently stated time, for emphasis—for a proposition q is the set of times t such that, according to what the database currently states (which is to say, according to our current beliefs), q is, was, or will be true at time t"

- **Logged Time**

   "the logged time for a proposition p is the set of times t such that, according to what the database stated at time t, p was true"

Martin Fowler

- **Actual Time**

- **Record Time**

  "The terminology of valid time and transaction time comes from Snodgrass, and is also used in the SQL:2011 standard. When I first started giving workshops about temporal modeling, back in the early naughts, I used these terms, but people found them confusing. So instead we started to use actual/record instead."

Tom Johnston Bitemporality

- **State Time**

  "I will use the term 'state time' [instead of 'valid time'] because it seems to me to be the least misleading term to apply to this concept."

- **Assertion Time**

  "the time during which a row of data is asserted to make a true statement"

Tom Johnston Tritemporality

- **State Time**

  [like before]

- **Speech Act Time**

  assertion time "generalized to be the time between any speech act and its withdrawal […] relativized to the source performing the speech act" (with assertion, withdrawal, assent, dissent and notice time as special cases)

- **Inscription Time**

  "begins when a row [in a database table] is physically created [and] ends when a row is marked as physically deleted"

Philipp Salvisberg

- **Valid Time**

  "the period during which something in the real world is considered valid"

- **Decision Time**

  "the date and time a decision has been made"

- **Transaction Time**

  [time of change in database]

Michael Brackett

- **Business Change Time**

  "the point in time that the business value change actually happened in the business world"

- **Organization Notification Time**

  "the point in time that the business value change was reported"

- **Organization Receipt Time**

  "the point in time that the change was first received by the organization"

- **Change Entry Time**

  "the point in time that the change was entered into the organization's data resource"

- **Change Availability Time**

  "the point in time that the change entered into the data resource was actually available to applications and queries"

Lars Rönnbäck (Anchor)

- **Changing Time**

  "an interval defining the period of time in which its value or relation is valid in the domain of discourse being modeled"

- **Happening Time**

  "the moment or interval at which an event took place in the domain of discourse being modeled"

- **Recording Time**

  "The time during which values were stored in some kind of memory"

Lars Rönnbäck (Transitional)

- **Appearance Time**

  "The time when some value can be said to have appeared (or will appear) for some thing or a collection of things. Such a statement is called a posit."

- **Assertion Time**

  "The time when someone is expressing an opinion about their certainty toward a posit."

Petr Beles (Data Vault)

- **1d (Business Validity)**

  "Valid time"

- **2d (Inscription Time)**

  "Approximation for Assertion Time in the source IT system"

- **3d (Load Time)**

  "Approximation [for] Assertion Time in the DWH system"

Patrick Cuba (Data Vault)

- **Applied Time**

  "the metadata package/extract date of a parcel/batch/package of applicable data"

- **Load Time**

  "Timestamp of load to the data warehouse/lake"

KsqlDB (Kafka)

- **Event Time**

  "The time when a record is created by the data source."

- **Ingestion Time**

  "The time when a record is stored in a topic partition by a Kafka broker. "

- **Processing Time**

  "The time when the record is consumed by a stream processing application."

Summary: Dimension #1

Terms describing the point in time when a fact is true in the real world:

- Valid Time
- Application Time
- Stated Time
- Actual Time
- State Time
- Business Change Time
- Changing Time
- Happening Time
- Appearance Time
- Event Time

Summary: Dimension #2

Terms describing the point in time when someone makes or notices a statement about some fact:

- Assertion Time
- Speech Act Time
- Withdrawal Time
- Assent Time
- Dissent Time
- Notice time
- Decision Time
- Organization Notification Time
- Organization Receipt Time

Summary: Dimension #3

Terms describing the point in time when a statement about some fact is stored in a database (or other means of persistent storage):

- Transaction Time
- System Time
- Logged Time
- Record Time
- Inscription Time
- Change Entry Time
- Change Availability Time
- Recording Time
- Load Time
- Applied Time
- Ingestion Time
- Processing Time

Summary: Three Dimensions of Time

#	Time Dimension	Alternative Terms	Special Cases		
1	Appearance Time	Valid Time	Happening Time		Business Change Time
		Application Time			
		Stated Time			
		Actual Time			Event Time
		State Time			
		Changing Time			
2	Assertion Time (Rönnbäck)	Speech Act Time	Assertion Time (Johnston)		Decision Time
					Organization Notification Time
			Withdrawal Time		
			Assent Time		
			Dissent Time		
			Notice Time		Organization Receipt Time
3	Recording Time	Transaction Time	Change Entry Time		
		System Time	Change Availability Time		Applied Time
		Logged Time			Ingestion Time
		Record Time			Load Time
		Inscription Time			Processing Time

What to Do with Your Time

Dealing with Different Kinds of Time

How Much Time Can You Have?

- Do your operational systems store …

    - … Appearance Time?

    - … Assertion Time(s)?

    - … their own Recording Time?

- If not, why not? Can that change?

How Much Time Should You Have?

- Do your stakeholders want to know …
    - … Appearance Time?
    - … Assertion Time(s)?
    - … the operational Recording Time?
    - … the DWH Recording Time?
- If so, why?
- If not, can that change?

Storage vs. Presentation of Time

- Store …
  - … all kinds of time operational systems can deliver.
  - … all kinds of time you need for regulatory requirements.
  - … the DWH recording time.
- Present …
  - … as few kinds of time as possible.
  - … the kinds of time users need for the given use case (not more!).
  - … real-life kinds of time (to the extent possible; approximate if necessary).

Name View Modes … (1/2)

Non-temporal Views (**Fixed-Fixed**):

- **Now-Now** (Assertion Time now, Appearance Time now; "as-is")

- **Now-Then** (Assertion Time now, Appearance Time in the past; e. g. the situation at the end of last year as it presents itself now)

- **Then-Then** (Assertion Time & Appearance Time in the past; e. g. "as-posted" view with Assertion Time ≈ Appearance Time)

- **Then-Now** (Assertion Time in the past, Appearance Time now; e. g. forecast of current month)

Name View Modes … (2/2)

Unitemporal views:

- **Fixed-Variable** (Assertion Time fixed, full Appearance Time history; e. g. history of real-life changes as it presents itself now)

- **Variable-Fixed** (full Assertion Time history, fixed Appearance Time; e. g. changing forecasts for the current year)

Bitemporal views (**Variable-Variable**; avoid if possible)

Tritemporal views (don't even think about it)

… and Then Pick the Right One

- Always make visible how current the presented information is.

- **Rule:** For each use case, use the simplest view that still works.

  - Pick **Now-Now** wherever possible (easy to understand, good enough for large majority of use cases).

  - If not possible, why not?

- **Exception:** If stakeholders want to be able to recreate past reports, make sure they can do that (or they will archive Excel exports).

Further Reading

- Petr Beles
  - Temporality in the Data Warehouse (Parts 1, 2, 3, 4)
- Michael Brackett
  - Data Resource Design (2012)
- Tom Johnston
  - Bitemporal Data (2014)
- Christian Kaul
  - Modern Data Warehousing (Parts 2, 3)

- Lars Rönnbäck
  - List of publications
  - Transitional Modeling
  - Temporal Dimensional Modeling
  - Anchor Modeling (bitemporal)
  - Time in Databases
- Richard Snodgrass
- Roelant Vos
  - A Brief History of Time for Data Vault