Anchor Modeling
A Technique for Information under Evolution

Lars Rönnbäck @ GSE Nordics June 7–9, 2011
Graph showing relative occurrence of words in literature over the last century

Information is rapidly becoming the most important asset
“Panta rhei”
Everything flows

Heraclitus
500 BC
Evolving Information

- Changing content
- Changing structure
- Changing constraints
- Changing interpretation
- Changing origins
- Changing reliability

There's a big difference between saying: "This information has a 95% reliability" and "This information is 100% reliable".
What is a database?

- The purpose of a database is to store a body of information and allow searches over it.

- The purpose of a temporal database is to store a body of information under evolution and allow historical searches over it.

But, we are not there yet!
What is a Data Warehouse?

- Integrates information from many sources
- Keeps a history of changes
- Provides “one version of the truth”
- Enables reporting, ad-hoc analysis, mining
- Calculates and stores new information
The dilemma

Many sources and many users naturally result in many changes

- Dimensional Modeling: 55%
- Normalized: 20%
- Haphazard: 25%

lack of adherence

degradation over time

lack of adherence
Patch or Redo?

- Patching works initially to cope with new requirements
- Maintenance costs usually rise proportionally to the lifetime of the data warehouse
- Redoing is unavoidable at some point (and for dimensional modeling sometimes accounted for)
- The average lifetime is five years
- The return of investment should and could be much better with a longer lifetime!
What is Anchor Modeling?

- Anchor Modeling combines normalization and emulation to provide an agile database modeling technique for evolving information that is implementable in current relational databases.

- Most, if not all, of what Anchor Modeling is doing in its physical (relational) representation could be "hidden" from the end-user in a true temporal database.
Technologies

- Entity-Relationship Modeling
- Sixth Normal Form Tables
- Temporal Database Emulation
History

Paul Johannesson
Lars Rönnbäck
Olle Regardt
Maria Bergholtz
Petia Wohed

Consulting
DW MDM EDW DW TDWI WWW ER09 TOOL AMW

Research

Best Paper Award @ ER’09
Philosophy

- Make modeling free from assumptions
- Make modeling agile and iterative
- Make evolution non-destructive
- Do not duplicate information
- Do not alter existing information
- Decouple metadata from the model
- Provide a simple interface for queries
Changing content
full support [6NF + time of change]

Changing structure
full support (through extensions) [non-destructive schema evolution]

Changing constraints
minimal support [only primary and foreign keys]

Changing interpretation
achievable [explicitly modeled]

Changing origins
restricted support [using metadata]

Changing reliability
restricted support [using metadata]
Domain driven modeling

- Data driven modeling
  - Data Vault/ODS/3NF (Inmon)
  - Dimensional Modeling (Kimball)
- Use-case driven modeling
  - Anchor Modeling

Positioning Anchor Modeling

mimics reality
mimics structure
mimics searches
Basic Notions

**Attributes** – properties
Example: The surname of a Person
<#42, ‘Rönnbäck’, 2004-06-19>

**Anchors** – entities
Example: A Person
<#42> (holds only identities of entities)

**Knots** – shared properties
Example: The gender of a Person
<#1, ‘Male’> + <#42, #1>

**Ties** – relationships
Example: The children of a Person
<#42, #4711>
Historization is done using the time of change as the start of an interval implicitly closed by another instance of the same identity with a later time of change.

Note that UPDATE is never allowed in an anchor database.
The Modeling Tool

- Open Source
- Online (HTML5)
- Free to use
- In the Cloud
- XML Interchange Format
- Automatic generation of SQL scripts
- Interactive (force-directed) Layout Engine

www.anchormodeling.com/modeler
Important Benefits

- Handles evolving information (keeping the integrity intact)
- Increases longevity (databases with long life expectancy)
- Simplifies modeling concepts (less prone to error)
- Enables modular and iterative development
- Needs no translation logic to the physical layer
- Automates generation of scripts
- No downtime when upgrading databases
- Scans only relevant data during searches
- Sparse data cause no gaps (no null values)
More Information

Homepage: [http://www.anchormodeling.com](http://www.anchormodeling.com)
Blog . Video Tutorials . Modeling Tool

Twitter: [anchormodeling](https://twitter.com/anchormodeling)

E-mail: [lars.ronnback@anchormodeling.com](mailto:lars.ronnback@anchormodeling.com)

LinkedIn Groups:
Anchor Modeling
Temporal Data Modeling

[re!sight](https://re-sight.com) insight • change • value