




data
modeling
zone

EUROPE
2016






10th & 11th OCTOBER 2016

**ABION Spreebogen Waterside Hotel
Berlin, Germany**

	Foundational Data Modeling	Communication Skills & Case Studies	Tools & NoSQL	Advanced Data Modeling	CDMP Certification
7:00 – 9:00	Breakfast and Registration				
9:00 – 12:00	Business Information Modeling using the fact-based approach Clifford Heath Infinuendo [Australia]	Agile as a complex ecology: from manufacturing to service Dave Snowden Cognitive Edge [United Kingdom]	Framing Business Problems as Data Mining Problems Asoka Diggs Intel Corp. [United States]	Advanced Data Modeling Challenges Workshop Steve Hobermann Steve Hoberman & Associates, LLC [United States]	Testing 
12:00 – 13:15	Lunch				
13:15 – 13:30	Welcome and Announcements				
13:30 – 14:30	KEYNOTE: Humans are neither ants nor data processors Dave Snowden, Cognitive Edge [United Kingdom]				
14:30 – 15:30	About Ontologies – the Foundation of Data Modeling Peer M. Carlson Dörffler & Partner [Germany]	Business Information Modeling: A Methodology for Data-Intensive Projects, Data Science and Big Data Governance Torsten Priebe Simplity [Austria]	Customer case of Heureka/Erwin	PS-3C Modeling: proposed way of agile EDW data modeling on the Data Lake Rogier Werschull DIKW [Netherlands]	Testing 
15:30 – 16:00	Afternoon Snacks				
16:00 – 17:00	Customer case of Teradata	A Great Challenge: Designing A New DWH For A Bank Murat Cetinkaya ING Bank [Turkey]	Data Modeling for NoTables Thomas Frisendal TF Informatik [Denmark]	Modeling of Reference Schemes Dr. Terry Halpin Professor at INTI International University, Malaysia [Australia]	Testing 



	Foundational Data Modeling	Communication Skills & Case Studies	Tools & NoSQL	Advanced Data Modeling	CDMP Certification
8:00 – 8:30 Morning Session			Casetalk – Data Modeling by example Marco Wobben BCP Software [Netherlands]		
7:00 – 9:00	Breakfast and Registration				
9:00 – 12:00	From Conceptual to Physical Data Vault Data Model Dirk Lerner ITGAIN [Germany]	Data Modeling for Sustainable Systems Graham Witt [Australia]	NoSQL Data Modeling Overview Steve Hobermann Steve Hoberman & Associates, LLC [United States]	Fact-Based Modeling Workshop Dr. Terry Halpin Professor at INTI, Int. University, Malaysia [Australia]	Testing 
12:00 – 13:15	Lunch				
13:15 – 13:30	Welcome and Announcements				
13:30 – 14:30	KEYNOTE: From BI to AI - Models, Ethics and Economics Barry Devlin, 9sight Consulting [South Africa]				
14:30 – 15:30	Why Business Analysts and Data Modelers should hold hands more George McGeachie Metadata Matters [United Kingdom]	Data Modeling for Polyglot Persistence Phill Radley BT [United Kingdom]	The Journey from ER Modeler to Data Scientist Asoka Diggs Intel Corp. [United States]	smartDATA and Industrie 4.0 Dr. Siegmund Priglinger dr.priglinger consulting GmbH [Germany]	Testing 
15:30 – 16:00	Afternoon Snacks				
16:00 – 17:00	Building a Data Vault using an Analytical Database Johan van der Kooij De Bijenkorf [Germany]	Creating Ontologies from your Existing E/R Data Models Nicolas Leveque Deutsche Bank [United States]	Positioning Data Modeling in the Big Data Area Martijn Imrich Xomnia [Netherlands]	How to compose a fact-based model into any kind of schema Clifford Heath Infinuendo [Australia]	Testing 

Humans are neither ants nor data processors

Dave Snowden, Cognitive Edge

Murray Gell-Mann is an American physicist who received the 1969 Nobel Prize in physics famously said : "The only valid model of a human system is the system itself". In this address Prof. Snowden, who is Director of the Centre for Applied Complexity at Bangor University in Wales, and Chief Scientific Officer of Cognitive Edge will outline new approaches to the use of distributed ethnography in understanding requirements capture and knowledge-information flows in organisations. Focusing on allow unarticulated needs to surface the approach offers a radical new way to model both cultural and information aspects of an organisation. Based on complexity and cognitive science the approach seeks to scale understanding of both actual and potential needs in a fractal manner, self similar at all levels of management.



Dave Snowden is the founder and chief scientific officer of Cognitive Edge. His work is international in nature and covers government and industry looking at complex issues relating to strategy, organisational

decision making and decision making. He has pioneered a science based approach to organisations drawing on anthropology, neuroscience and complex adaptive systems theory. He is a popular and passionate keynote speaker on a range of subjects, and is well known for his pragmatic cynicism and iconoclastic style.

He currently holds positions as extra-ordinary Professor at the Universities of Pretoria and Stellenbosch and has held similar positions at Hong Kong Polytechnic University, Canberra University, the University of Warwick and The University of Surrey. He held the position of senior fellow at the Institute of Defence and Strategic Studies at Nanyang University and the Civil Service College in Singapore during a sabbatical period in Nanyang.

His paper with Boone on Leadership was the cover article for the Harvard Business Review in November 2007 and also won the Academy of Management award for the best practitioner paper in the same year. He has previously won a special award from the

Academy for originality in his work on knowledge management. He is a editorial board member of several academic and practitioner journals in the field of knowledge management and is an Editor in Chief of E:CO. In 2006 he was Director of the EPSRC (UK) research programme on emergence and in 2007 was appointed to an NSF (US) review panel on complexity science research.

He previously worked for IBM where he was a Director of the Institution for Knowledge Management and founded the Cynefin Centre for Organisational Complexity; during that period he was selected by IBM as one of six on-demand thinkers for a world wide advertising campaign. Prior to that he worked in a range of strategic and management roles in the service sector.

His company Cognitive Edge exists to integrate academic thinking with practice in organisations throughout the world and operates on a network model working with Academics, Government, Commercial Organisations, NGOs and Independent Consultants. He is also the main designer of the SenseMaker® software suite, originally developed in the field of counter terrorism and now being actively deployed in both Government and Industry to handle issues of impact measurement, customer/employee insight, narrative based knowledge management, strategic foresight and risk management.

From BI to AI – Models, Ethics and Economics

Barry Devlin, 9sight Consulting

The creative use of big data from social media and the Internet of Things is transforming business models, empowering start-ups, and enhancing or destroying existing business value. Many proponents focus on marketing uses of big data, but most value—and disruption—will come from its daily operational application in novel, transformative ways.

From slim beginnings thirty years ago in BI to support decision making, data collection and analysis is now enabling AI systems to make financial recommendations or answer customer calls. They are diagnosing cancer and offering treatment options, and driving autonomous trucks. Large swathes of decision making and action taking are being automated, untouched by human hand. Even complex decisions are in-

creasingly augmented by cognitive computing systems.

Making the right design choices is vital in designing and managing such environments. However, true business and IT leadership requires considering the broader ethical and economic issues raised by this transformation. Concerns around personal privacy, employment and social disruption must all be urgently addressed if individual businesses and society at large are to successfully navigate this data driven transformation of all aspects of business and technology, and even society.



Dr. Barry Devlin is among the foremost authorities on business insight and one of the founders of data warehousing, having published the first architectural paper on the topic in 1988. With over 30 years of IT experience, including 20 years with IBM as a Distinguished Engineer, he is a widely respected analyst, consultant, lecturer and author of the seminal book, *Data Warehouse—from Architecture to Implementation*, and numerous white papers, blogs and more. His new book *Business unIntelligence—Insight and Innovation beyond Analytics and Big Data* was published in 2013. Barry is founder and principal of 9sight Consulting. He specializes in the human, organizational and IT implications of deep business insight solutions in all technology environments. Barry is based in Cape Town, South Africa and operates worldwide.

NoSQL Data Modeling Overview

Steve Hoberman,
Steve Hoberman & Associate

NoSQL implementations are often built with little or no data modeling...or at the other extreme completely over-architected – both ends of the spectrum producing suboptimal results. In this presentation, understand how data modeling contributes to the process of learning about the data, and is therefore a required technique, even when the resulting database is not relational. We will explore what distinguishes NoSQL from a relational database, and what distinguishes a relational database from relational modeling. We will cover the optimal situations each

of the four types of NoSQL databases can be used. We will learn a streamlined approach to data modeling that ensures the NoSQL database will be built with a thorough understanding of the current requirements, as well as designed to handle future needs. We will cover conceptual, logical, and physical data modeling for NoSQL, and how each stage increases our knowledge of the data and reduces assumptions and poor design decisions. Exercises will be used to reinforce techniques.

Steve Hoberman has trained more than 10,000 people in data modeling since 1992. Steve is known for his entertaining and interactive teaching style (watch out for flying candy!), and organizations



around the globe have brought Steve in to teach his Data Modeling Master Class, which is recognized as the most comprehensive data modeling course in the industry. Steve is the author of nine books on data modeling, including the bestseller *Data Modeling Made Simple*. One of Steve's frequent data modeling consulting assignments is to review data models using his Data Model Scorecard® technique. He is the founder of the Design Challenges group and recipient of the 2012 Data Administration Management Association (DAMA) International Professional Achievement Award.

Agile as a complex ecology: from manufacturing to service

Dave Snowden, Cognitive Edge

The wider vision of the Agile manifesto too often ends up as a series of structured methods designed to produce a product based on an underlying manufacturing method. Indeed much that is Agile is drawn directly from a manufacturing environment. Any manufacturing process a priori makes assumptions about the degree to which something can be defined and measured. In an ecology some aspects are susceptible to such a way of thinking but in many cases needs are not, indeed cannot be articulated given the speed of change of technological capability. With the Internet of Things we are moving to a more fragmented ecology in which applications will emerge from the three way in-

teraction of people, software objects and the wider environment.

The underlying architecture and metaphor of software development will have to undergo radical change to accommodate these changes. Understanding the difference between sequential iteration in methods such as SCRUM and parallel short cycle experimentation where definition comes at the end, not the start of the process will be key. Techniques in which partial copying creates rapid mutation of need to allow novel forms of capability to emerge are being developed based on metaphors drawn directly from evolutionary biology, in particular exaptation (rapid repurposing of existing capability to meet unanticipated needs) and epi-genetics (Culture activates or deactivates genetic capability) are starting to provide a fertile, more simple (but not simplistic) approach to scale than simply extended the engineering metaphor.

This workshop will give participants early insight into both the theory and practice of this approach.



Dave Snowden is the founder and chief scientific officer of Cognitive Edge. His work is international in nature and covers government and industry looking at complex issues relating to strategy, organisational decision making and decision making.

He has pioneered a science based approach to organisations drawing on anthropology, neuroscience and complex adaptive systems theory. He is a popular and passionate keynote speaker on a range of subjects, and is well known for his pragmatic cynicism and iconoclastic style.

He currently holds positions as extra-ordinary Professor at the Universities of Pretoria and Stellenbosch and has held similar positions at Hong Kong Polytechnic University, Canberra University, the University of Warwick and The University of Surrey. He held the position of senior fellow at the Institute of Defence and Strategic Studies at Nanyang University and the Civil Service College in Singapore during a sabbatical period in Nanyang.

His paper with Boone on Leadership was the cover article for the Harvard Business Review in November

2007 and also won the Academy of Management award for the best practitioner paper in the same year. He has previously won a special award from the Academy for originality in his work on knowledge management. He is a editorial board member of several academic and practitioner journals in the field of knowledge management and is an Editor in Chief of E:CO. In 2006 he was Director of the EPSRC (UK) research programme on emergence and in 2007 was appointed to an NSF (US) review panel on complexity science research.

He previously worked for IBM where he was a Director of the Institution for Knowledge Management and founded the Cynefin Centre for Organisational Complexity; during that period he was selected by IBM as one of six on-demand thinkers for a world wide advertising campaign. Prior to that he worked in a range of strategic and management roles in the service sector.

His company Cognitive Edge exists to integrate academic thinking with practice in organisations throughout the world and operates on a network model working with Academics, Government, Commercial Organisations, NGOs and Independent Consultants. He is also the main designer of the SenseMaker® software suite, originally developed in the field of counter terrorism and now being actively deployed in both Government and Industry to handle issues of impact measurement, customer/employee insight, narrative based knowledge management, strategic foresight and risk management.

Framing Business Problems as Data Mining Problems

Asoka Diggs, Intel Corp.

This will be a brief introduction to some of the important concepts that arise in data mining in general and classification analysis in particular. The focus in this presentation will be helping attendees to develop their ability to frame business problems as data mining problems, to understand some of the terminology and mental models prevalent in data mining, and to being able to recognize situations where data mining might apply.

Through a demonstration driven presentation format, attendees will be exposed to a variety of data mining topics. Throughout the demo, attention will be paid to points of intersection

with traditional data modeler work products including the data dictionary and dimensional data model. I will also leave you with a variety of links and references for additional study on your own.

Particular topics that will be addressed:

- Formulating business problems as data mining problems
- Supervised Learning
- Classification methods including Decision Tree and Logistic Regression
- Model Validation
- No math!



After 15+ years of experience in a variety of data management disciplines, including database administration, ER modeling, ETL development, and data architecture, I am transitioning into the world of analytics / data science.

It turns out that there IS something even more fun than ER modeling, and I've found it. Today, my primary interest is in the organizational transformation involved in adopting analytics as a source of competitive advantage. How does an organization get there? What needs to be done? What organization design and leadership changes are needed to get the most benefit from analytics? I now spend my work life practicing these new analytic modeling skills, and teaching others how to participate and contribute to predictive analytic projects.

Fact-Based Modeling Workshop

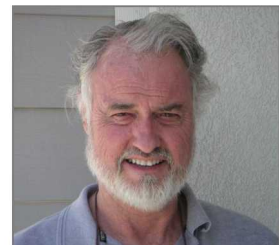
Dr. Terry Halpin, Professor at INTI International University, Malaysia

Typically, analysis and modeling of an enterprise or business domain is a collaborative process between the business expert, who best understands the business rules and requirements, and the modeler whose main task is to elicit this understanding and capture it in a formal model that can be executed (directly or indirectly via transformation) by an information system. Ideally, the model (including its representation of the business rules) should be validated by the business experts. Since business experts are often non-technical, the model needs to be con-

ceptual in nature, cast in language that is intelligible to the business expert while still being formal. Fact-Based Modeling aims to address this communication need in an optimal way.

Fact-Based Modeling has been adopted by a family of related approaches, of which Object-Role Modeling (ORM) and SBVR (Semantics of Business Rules and Business Vocabulary) are well known examples. This presentation begins with brief overview of fact-orientation in general and second-generation ORM (ORM 2) in particular, highlighting its communication-based modeling procedure using controlled natural language and visual support for conceptualization of business rules. It explains ORM's underlying principles, outlines recent improvements to the methodology and associated tool support, and briefly contrasts fact-orientation with other high level data modeling approaches (e.g. ER and UML). The bulk of the workshop then focuses on developing skills in constructing and transforming data models by applying fact-orientation to practical examples of information systems, including advanced aspects, with the implementation emphasis on relational database systems.

Dr. Terry Halpin is a Research Fellow in Computer Science at INTI International University (Malaysia), and a data modeling consultant. He previously held senior faculty positions in computer science at The



University of Queensland (Australia), and professorships in computer science at Neumont University (USA) and INTI International University. His prior industrial experience includes many years in data modeling technology at Asymetrix Corporation, InfoModelers Inc., Visio Corporation, Microsoft Corporation and LogicBlox. His doctoral thesis formalized Object-Role Modeling (ORM/NIAM), and his current research focuses on conceptual modeling and rule-based technology. He has authored over 200 technical publications and nine books, and has co-edited nine books on information systems modeling research. He is a member of IFIP WG 8.1 (Information Systems), is an associate editor or reviewer for several academic journals, is a regular columnist for the Business Rules Journal, and is a recipient of the DAMA International Achievement Award for Education and the IFIP Outstanding Service Award.

About Ontologies – the Foundation of Data Modeling

Peer M. Carlson, Dörffler & Partner

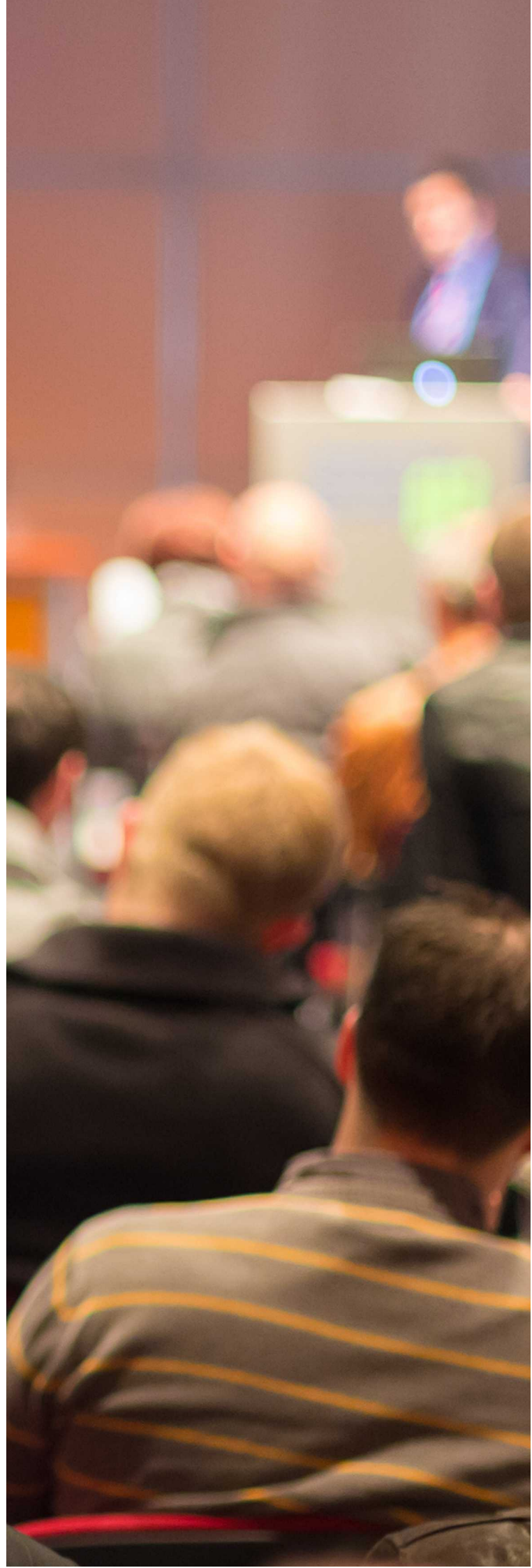
In a lot of texts and talks about Data Modeling, a reference is given to the concept of Ontologies, whereas a further explanation of this notion is often missing. This presentation aims to reduce the knowledge gap of those who want to learn more about the theoretical background of Ontologies, to discover the interdisciplinary relationships in this area, and to understand how modern Business Intelligence/Data Warehouse systems benefit from applying ontological ideas.

The first part of the presentation briefly introduces the Ontology from a philosophical point of view – the study of being and existence in the world and the interrelationships of objects. In the second part, the focus lies on Ontologies in the field of Information Science as a formal specification of a conceptualization – and furthermore, as a particular example, on Ontologies in the context of the Semantic Web. Finally, the third part will show how Ontologies serve as the foundation of Data Modeling – particularly Data Vault – and how they facilitate the processing of unstructured data.



Peer M. Carlson is a Business Intelligence Consultant at Dörffler & Partner GmbH, Germany. He has been working on several Data Warehouse projects, mainly in the field of Microsoft SQL Server as well as

Teradata. He has a lot of experience in Data Modeling, the Data Vault methodology, ETL process design and an in-depth knowledge of BI/DW architectures. Peer is a certified BI Expert (TDWI Europe) and a Certified Data Vault 2.0 Practitioner, and he holds a university degree in Computer Science with focus on Business Information Systems.



Business Information Modeling: A Methodology for Data-Intensive Projects, Data Science and Big Data Governance

Torsten Priebe, *Simplify*

The talk presents an integrated methodology to structure and formalize business requirements in large data-intensive projects, e.g. data warehouse implementations, turning them into precise and unambiguous data definitions suitable to facilitate harmonization and assignment of data governance responsibilities. The presented approach places a business information model in the center -- used end-to-end from analysis, design, development, testing to data quality checks by data stewards. In addition, the talk shows that the approach is suitable beyond traditional data warehouse environments, applying it also to big data landscapes and data science initiatives -- where business requirements analysis is often neglected. As proper tool support has turned out to be inevitable in many projects, software requirements and their implementation in the Accurity Glossary tool are also discussed. The approach is demonstrated based on examples from a number of large banking data warehouse projects.



Torsten Priebe has over 15 years' experience in data management and business intelligence, of which he spent over 10 years in consulting. Since 2014, Torsten is CTO at Simplify, a dynamically developing interna-

tional consultancy focusing on data warehousing and business intelligence. Torsten's responsibilities at Simplify include driving the requirements for the Accurity data governance software suite and guiding Simplify's consultants particularly in the areas of architecture and data modeling. Before Simplify, Torsten was heading the data management team at Teradata and the business intelligence practice at Capgemini in Vienna respectively. Torsten is co-author of three books, published numerous articles and talks regularly at international conferences. In addition, Torsten is a regular guest lecturer for business intelligence at University of Regensburg, where he received his PhD with a thesis on integrating structured and unstructured information with semantic technologies in 2005.

PS-3C Modeling: proposed way of agile EDW data modeling on the Data Lake

Rogier Werschkuil, *DIKW*

The last years we have seen the rise of the 'Data Lakes'. While they are technologically suitable for the time variant / non volatile part of the 'traditional' data warehousing as we know it, they do not automatically cover/fix the difficult part of solid EDW data warehousing: building a subject-oriented and integrated layer on top of this raw data. So even in the 'age of Data Lakes' data modeling is still needed and maybe even more than before: to bring Data-science related discoveries into production we must have a good solution for the 'data wrangling' part that takes up most of the time ...

To solve this problem, I propose a new (layered) modeling approach for building this integrated EDW. This modeling approach consists of two layers:

- 1) A persistent staging area (PS) in the Data lake: structured / semi-structured data is historized, based in a (technical) source Primary / Unique Key. We must strive to build this using a columnar-oriented (and compressed) storage method. Then, using schema-on read functionality, we model the integrated EDW layer:
- 2) A layer based on [Business Concepts], described by [Context] and linked together by [Connector-context] tables (3C). It is inspired on Data Vault modeling with the main differences being shared during this talk.

Currently, I am in the process of proposing this modeling approach in a major organization in the Netherlands.

Rogier is a DWH / BI Consultant with 16 years of experience in the field. He started as a frontend developer with Business objects, moved to OLAP cube development and ETL development (Oracle warehouse builder, Kimball Datamarts /Data warehouse)



and the last 5 years to DWH /ETL automation (using Data Vault, build an ETL generator at my previous employer) and DWH architecture and data modeling. Recently he has built some experience using 'Big Data' technology (Hadoop ecosystem) and in the Data Science related field. This combined experience has given me a broad perspective in data architecture in general and has given me ideas on how to improve the current way of modeling the integrated EDW.

A Great Challenge: Designing A New DWH For A Bank

Murat Cetinkaya, ING Bank Turkey

We have finished our new IBM BDW based data warehouse successfully in one and half years. I will discuss this project at ING Bank from a modeling perspective, including the cons and pros of industry models, best practices, source system challenges, modeling challenges, planning, and physicalization.



Murat Cetinkaya is working at ING Bank Turkey as IT Product Manager/Data Warehouse Architect. Mr. Cetinkaya is a data warehouse professional with broad experience in big-sized data warehouse projects.

He has detailed experience of designing and implementing staging area, operational data store and all aspects of the data warehouse development lifecycle; requirements gathering, both logical and physical design of the data models, ETL design and processing. He has held modeling team leadership and DWH architect positions in Is Bank (the biggest private bank of Turkey) and ING Bank Turkey. He is also an active member of the DAMA Turkey Chapter.

Data Modeling for NoTables

Thomas Frisendal, TF Informatik

It is time for a change in the way we work with data models. The database world is changing rapidly as NoSQL (not only SQL) and new physical data models such as graphs, key/value, column stores and HADOOP make headway into database development. But data modeling is still seen as synonymous with SQL tables, normalization and entity / relationship diagrams.

Graphs (as in directed graphs) is the natural replacement for the 40+ years old data modeling discipline with its bias on normalization and an abstract relational model, which never quite made it to mainstream data modeling. The legacy is that of SQL-tables, primary and foreign keys etc. as opposed to relations / "relvars", which constitute the academic foundation for relational database concepts.

We will cover subjects such as:

- Make data models for the new world of Big Data and NoSQL
- Make data models that are visually compelling and highly communicative (also for non-technical audiences)
- Work visually with the functional dependencies (and the join dependencies) to control redundancy and consistency.
- Feel good at not doing normalization!

If you are a data modeler, or a data scientist or a developer doing data modeling as you go, then this is for you.

Thomas Frisendal is an experienced consultant with more than 30 years on the IT vendor side and as an independent consultant. Since 1995 he has primarily been working with data warehouse projects. His



area of excellence lies within the art of turning data into information and knowledge. Since 2005 he has specialized in business analysis, concept "harvesting" and mapping, modeling of information and data as well as design of business intelligence solutions. Thomas Frisendal is a member of "The Data Warehouse Institute Denmark" (member of the board), "The International Association for Information and Data Quality" (IAIDQ) and "The Cognitive Science Society". His approach to Information-driven Business Analysis is "New Nordic" in the sense that it represents the traditional Nordic values such as superior quality, functionality, reliability and innovation by new ways of conceptualizing the business. Thomas is an active writer and speaker, and also Chief Data Warehouse Architect at SimCorp. Thomas lives in Copenhagen, Denmark.

How to compose a fact-based model into any kind of schema

Clifford Heath, Infinuendo

Fact-based modeling is not just a powerful business communication tool, it is the one true ring of power for the data modeler. By working with a fully decomposed schema in "elementary form", we avoid making a commitment to any particular kind of composite schema. This presentation will show how different composition rules applied to a fact-based schema will generate an entity-relational, star, snowflake, hierarchical, object-oriented, XML or graph schema.



Clifford Heath is a computer research scientist who has long experience in the design and implementation of enterprise-scale software products, and in the use of fact-based modeling. He is a Certified Data Management Professional (CDMP) at Masters level, has published a number of papers in peer-reviewed scientific journals, holds several patents, is the creator of the Constellation Query Language and is a participant in the Fact Based Modeling Working Group. Clifford has frequently presented at chapter meetings of the Data Management Association as well as at the NATO CAX Forum and the European Space Agency, and is CTO and founder at Infinuendo.

Casetalk – Data Modeling by example

Marco Wobben, Casetalk

Q: Data modeling is described as a craft and once completed the results may even seem artful. Yet outsiders may see data modeling as abstract, time consuming or even unnecessary. In many cases the data modeler interviews business experts, studies piles of requirements, talks some more, and then, hocus pocus, presents a diagram with boxes, crows feet, arrows, etc... Then the slow process begins to keep the diagrams up to date, explain what the diagrams behold, and sometimes even data modelers themselves may get lost while maintaining a growing set of data models and requirements.

A: Fact based information modeling is the very opposite of abstract. Fact based information modeling uses natural language which expresses facts that are intelligible for both business and technical people. It does not require people to understand the modeler's magical language of boxes and arrows. Although models can be presented in several diagramming notations, they can be validated in natural language at all times. This gives both data modelers, technically skilled people, and business people the benefit of having a well documented and grounded data model. Therefore the method of Fact Oriented Modeling, is also known as "Data Modeling by Example".

Presentation highlights

- key elements of fact oriented modeling;
- data modeling with facts;
- visualizing the model;
- validating and verbalizing;
- transforming and generating output (E.g.: SQL, Relational, UML, XSD, PowerDesigner, etc.).

Marco Wobben is director of BCP Software and has been developing software well over 30 years. He has developed a wide range of applications from financial expert software, software to remotely operate bridges, automating DWH generating and loading, and many back- and frontoffice and webapplications. For the past 10 years, he is product manager and lead developer of CaseTalk, the CASE tool for fact based information modeling, which is widely used in universities in the Netherlands and across the globe.



From Conceptual to Physical Data Vault Data Model

Dirk Lerner, ITGAIN

The fictitious company PennyStockInc needs to build a new data warehouse for their trading department. Due to past experience (project failures, overtime, etc.) the BI Competence Center (BICC) decided this time to build conceptual, logical and physical data models. Conceptual models to gather all information about their trading

business, logical models to collect their business requirements and cover both relational databases and NoSQL solutions. For the physical data model the BICC of PennyStockInc decided to choose data vault due to the agility, flexibility and the ability to integrate both relational databases and NoSQL solutions.

Attendees of this session will be a fictitious part of the PennyStockInc BICC Team and will work on some exercises to build the new data warehouse.

This session explores

- Basic principles of
 - Conceptual modeling
 - Logical modeling
 - Physical modeling
- How to model from a conceptual via logical to a physical data vault model
- How to build and access the physical data model on Relational-DB (EXASOL) and NoSQL-DB (REDIS):
 - Design the data layout on REDIS. Identify the keys to represent the objects and which values this keys need to hold
 - Design the relational data model
 - Access both the key-value store and the relational database.

Dirk Lerner is a well experienced IT Consultant at IT-GAIN and is as team lead responsible for the Competence Center Data Architecture, Data Modeling and Data Vault. For around 15 years he has headed BI projects and is considered an expert for BI architectures and data modeling. Dirk is an advocate of flexible, lean and easily extendable data warehouse principles and practices. As a pioneer for Data Vault in Germany he published various publications, is an international speaker at conferences and author of the blog <http://www.datavaultmodeling.de>.



Business Information Modeling Fusing the fact-based approach

Clifford Heath, Infinuendo

Most modeling languages are effective for communicating only with people who have been trained to read and interpret the details. However, fact-based languages are deeply rooted in natural speech. They build a complete vocabulary for the business – not just a glossary of terms – and can be used directly with untrained business personnel. This radically improves the depth and breadth of the business dialogue.

As each new fact is encountered, the modeler



adds appropriate elements to an Object Role Model diagram, which makes it easy to verbalize to natural language. Alternatively, the new fact can be expressed directly in the structured natural language form of the Constellation Query Language, which is immediately readable to business users but can also be understood by the computer. This makes it possible to describe any factual situation precisely, to ask or answer any factual question, and to elaborate detailed business rules about possible or allowable situations.

This presentation will introduce you to the approach and to these two languages, peeking at the respective software tools which generate physical schemas. A group exercise will challenge you to apply it to model a sample problem. The attendee will learn to use verbalization and fact-based analysis to understand any subject material, and to communicate it at a deep conceptual level with those less expert in the field.



Clifford Heath is a computer research scientist who has long experience in the design and implementation of enterprise-scale software products, and in the use of fact-based modeling. He is a Certified Data Management Professional (CDMP) at Masters level, has published a number of papers in peer-reviewed scientific journals, holds several patents, is the creator of the Constellation Query Language and is a participant in the Fact Based Modeling Working Group. Clifford has frequently presented at chapter meetings of the Data Management Association as well as at the NATO CAX Forum and the European Space Agency, and is CTO and founder at Influido.

Data Modeling for Sustainable Systems

Graham Witt

Systems, like any other high-value acquisitions, should continue to work after the warranty period, but far too often fail to work completely as intended. Data modelers can enhance the quality and lifespan of a system if they take a broad

view than one which simply converts information storage and retrieval requirements into a data model. This workshop looks at how to add value to the data modeler's contribution to system development or package customization, including:

- identifying real-world complexity and its implications for the choice of data structures
- the role of generalization in managing such complexity
- development of a common vocabulary for real-world concepts, attributes, relationships and processes, and use of that vocabulary in system artifacts and user interfaces
- recognizing how the data is to be used in Business Intelligence and the implications of that use for choice of data structures
- analysis of how changes in the real world are to be recorded in the chosen data structures, with implications for process and user interface design
- ensuring test plans cover data update side-effects adequately
- managing data model change effectively.

The workshop includes a series of case studies drawn from the speaker's experience dealing with these issues.

Graham has over 30 years of experience in delivering effective data solutions to the government, transportation, finance and utility sectors. He has specialist expertise in business requirements, architectures, information management, user interface design, data modeling, database design, data quality and business rules. He has spoken at conferences in Australia, the US and Europe and delivered data modeling and business rules training in Australia, Canada and the US. He has written two textbooks published by Morgan Kaufmann: "Data Modeling Essentials" (with Graeme Simson) and "Writing Effective Business Rules", and writes monthly articles for the Business Rule Community (www.brcommunity.com).



Advanced Data Modeling Challenges Workshop

Steve Hoberman,
Steve Hoberman & Associates

After you are comfortable with data modeling terminology and have built a number of data models, often the way to continuously sharpen your skills is to take on more challenging assignments. Join us for a half day of tackling real world data modeling scenarios. We will complete at least ten challenges covering these four areas:

- NoSQL data modeling
- Agile and data modeling
- Abstraction
- Advanced relational and dimensional modeling

Join us as in groups as we solve and discuss a set of model scenarios.



Steve Hoberman has trained more than 10,000 people in data modeling since 1992. Steve is known for his entertaining and interactive teaching style (watch out for flying candy!), and organizations around the globe have brought Steve in to teach his Data Modeling Master Class, which is recognized as the most comprehensive data modeling course in the industry. Steve is the author of nine books on data modeling, including the bestseller *Data Modeling Made Simple*. One of Steve's frequent data modeling consulting assignments is to review data models using his Data Model Scorecard® technique. He is the founder of the Design Challenges group and recipient of the 2012 Data Administration Management Association (DAMA) International Professional Achievement Award.

Why Business Analysts and Data Modelers should hold hands more

George McGeachie, *Metadata Matters*

How integrated are your business analysis and data modeling efforts? Share the experiences of a lifetime of modeling, and improve the coordination between your business analysts and data

modelers, enabling them to build a coherent, integrated, set of requirements and models, reducing the probability of project failure.

Some of the questions that will be answered include:

- Why does a business analyst need to understand or contribute to data models?
- Why does a data modeler need to understand or contribute to business analysis?
- Why designing XML messages is NOT data modeling, it is process modeling.
- How do Business Rules influence data models?
- Where does all this fit with other types of metadata, including Enterprise Architecture?
- How does this impact your project management and governance?
- What implications does this have for your modeling tools?

George McGeachie has spent his working life creating, managing and linking data models, process models, and others. He encourages organizations to connect and utilize their metadata islands, to recognize the wealth of information contained in their data models, to recognize that the creation of data models must form part of an integrated approach to improving their business, and therefore recognize the importance of avoiding the creation of islands of metadata in the first place.



Data Modeling for Polyglot Persistence

Phill Radley, *BT*

This presentation will give an overview of how British Telecom is adapting its approach to data modeling and data architecture to support project teams wanting to store their data in databases that are not based on a Relational Data Model. An important element being the use Fact Based Models (specifically Object Role Modeling using the NORMA tool) to develop conceptual models that can be used to join data stores such as HDFS, Hive, Hbase and Orient DB.

Agenda:

- Polyglot persistence another name for big data, graph data, NoSQL....
- Back to basics - Conceptual, Logical and Physical Data Models
- Basic Conceptual modeling using Object Role Modeling
- Hadoop (The Hive Metastore, File Formats and Data Catalogs)
- Schema on Read and Data Wrangling (using Datameer)
- Key Value Data Model (Time series Data in Hbase)
- Graph Data Model (OrientDB and Neo4J).



Phill Radley is a Physics Graduate with an MBA who has worked in IT and communications industry for 30 years, mostly with British Telecommunications plc. He is currently BT's Chief Data Architect at their Martlesham Heath campus in the UK. Phill works in BT's core Enterprise Architecture team with responsibility for Data Architecture across BT Group plc. He currently leads BT's MDM and Big Data initiatives driving associated strategic architecture and investment roadmaps for the business. His previous roles in BT include; 9 years as Chief Architect for Performance Management solutions, ranging from UK consumer broadband through to outsourced Fortune 500 networks and hi-frequency trading networks. He has broad global experience including BT's Concert global venture in USA and 5 years as Asia Pacific BSS/OSS Architect based in Sydney.

The Journey from ER Modeler to Data Scientist

Asoka Diggs, Intel Corp.

In this session, we will be talking about what data science is and isn't, with particular focus on the skills that a data scientist is expected to bring to bear on a problem. We will discuss the educational and skill expectations of a data scientist and how those relate to the skills we've developed as ER Modelers. Going by job descriptions, data science can easily read like "jack of all trades, and master of them all too". How do we break that down into something more reasonably achievable?

- What is data science and what is a data scientist?
- What do they do?
- Do ER Modelers make good Data Scientists?
- Are the skill sets complementary?
- And more – bring your questions and let's get into them.

After 15+ years of experience in a variety of data management disciplines, including database administration, ER modeling, ETL development, and data architecture, I am transitioning into the world of



analytics / data science. It turns out that there IS something even more fun than ER modeling, and I've found it. Today, my primary interest is in the organizational transformation involved in adopting analytics as a source of competitive advantage. How does an organization get there? What needs to be done? What organization design and leadership changes are needed to get the most benefit from analytics? I now spend my work life practicing these new analytic modeling skills, and teaching others how to participate and contribute to predictive analytic projects.

smartDATA and Industrie 4.0

Dr. Siegmund Priglinger,
dr.priglinger consulting

smartDATA for Industrie 4.0 is based on CAD/CAM, PLM and SE models, where different modeling methods/languages and tools are combined and monitoring approaches for exchanging and communication between the different systems at the engineering and shop-floor levels will be taken into account. Rigorous Data Engineering is a must. Formalization of schema mapping and matching is the key factor for Industrie 4.0.

More than 30 years experience as product and solution provider, company founder, IT Manager and Consulting Principal for ICT in different industries (Manufacturing, Trade, Telecommunication, IT, Energy, Healthcare, Public). Contact person for the management, the business divisions and the IT concerning "Business Intelligence", "Data Governance" and "In-



formation Management". My current job is "to advise other advisers" in this areas while I am project leader of complex data migration or BI projects. I designed new methods to accelerate the migration of application models and to compare the quality of schema and data of these application models. My vision is BI3: "Business Intelligence requires Best Information requires Best Integration."

Building a Data Vault using an Analytical Database

Johan van der kooij, De Bijenkorf

Many data warehouses run on a 'traditional' rowstore database, such as Oracle or SQL Server. With the rise of data driven organizations, those databases might not be the best choice for BI and analytical purposes. That's where dedicated analytical and column store databases come in to play. How can data vault modeling be used in such an analytical column store database? Challenges such as end dating (using expensive update statements) and satellite splitting (mostly done for storage and performance) suddenly can be approached in a completely different way. This talk shows how HP Vertica runs at the Bijenkorf on an AWS infrastructure, utilizing the full potential of the analytical database environment combined with the power of data vault modeling.



Johan works as BI/Bigdata architect for the eCommerce department at the Bijenkorf, the largest premium department store in the Netherlands. The Bijenkorf is part of the Selfridges group, an international luxury retail with annual sales total over £2 billion. Within the eCommerce department, Johan works in the Data Technology team. The team operates a data warehouse environment which offers self-service reporting, analytical capabilities and web integration to the organization. Prior to the Bijenkorf, Johan worked for several organizations in retail, finance and energy to help them become data driven.

Creating Ontologies from your Existing E/R Data Models

Nicolas Leveque, Deutsche Bank

During the last two generations, visionary people like Codd, Chen, Kent, Blaha, Wells, Inmon, and Hoberman contributed to the development and success of the Entity Relationship model and its implementation by modern RDBMS vendors. During this period, thousands of students, scholars and professionals were taught the three data modeling perspectives (Conceptual, Logical and Physical), and data management has become one of the most lucrative, job creating, and ever changing IT industry.

The last couple of years, our industry has seen the advent of a new modeling technique that facilitates the integration of heterogeneous data sources by resolving semantic heterogeneity between them: Ontology.

This presentation will explain in detail how to leverage all of the numerous E/R Data models produced and turn them into an OWL ontology model. This session will explain the links between ontology and E/R model constructs, and will help you with learning and creating your first ontology from their existing E/R data models.

So, if, as with Deutsche Bank, your company has started embracing Linked Data and ontology modeling, come and learn how to reuse what you did and turn your E/R Data Models into a basic OWL ontology.

Nicolas Leveque is Head of Physical Data Architecture, Deutsche Bank, and a Data Architect with 20 years experience mainly within financial institutions including UBS, Barclays, Societe Generale, Capital Group and Deutsche Bank. He specializes in Enterprise Data Architecture, Data Modeling and in bridging the gap between High Level architect and implementation teams. Still a hands-on technician, he doesn't believe in one-solution-fits-all, but strives to find the right solution depending on business and non-functional requirements. He also looks forward and sees



how new practices and technologies could help to solve data problems with a different and better mindset.

Positioning Data Modeling in the Big Data Area

Martijn Imrich, Xomnia

So we arrived in the area of Big Data, and everyone is talking about Hadoop, Data Lakes and why we should have a data warehouse or do data modeling anyway. And there's no shortage of technology. Every day new techniques arrive with fancy names such as Spark, Splunk, Hunk, Hive, Pig, YARN and so on. But how do we design an effective Big Data Architecture? And what is the role of Data Modeling in the Big Data area?

A so called Reference Architecture has been created in many domains, ranging from applications to Business Intelligence. Martijn will suggest a Reference Architecture for Big Data, and where data modeling fits.

Martijn Imrich is helping companies to be more Data Driven since the nineties. Today he is a Partner at Xomnia, an Amsterdam based Big Data startup. His main activities are in innovation workshops, educati-

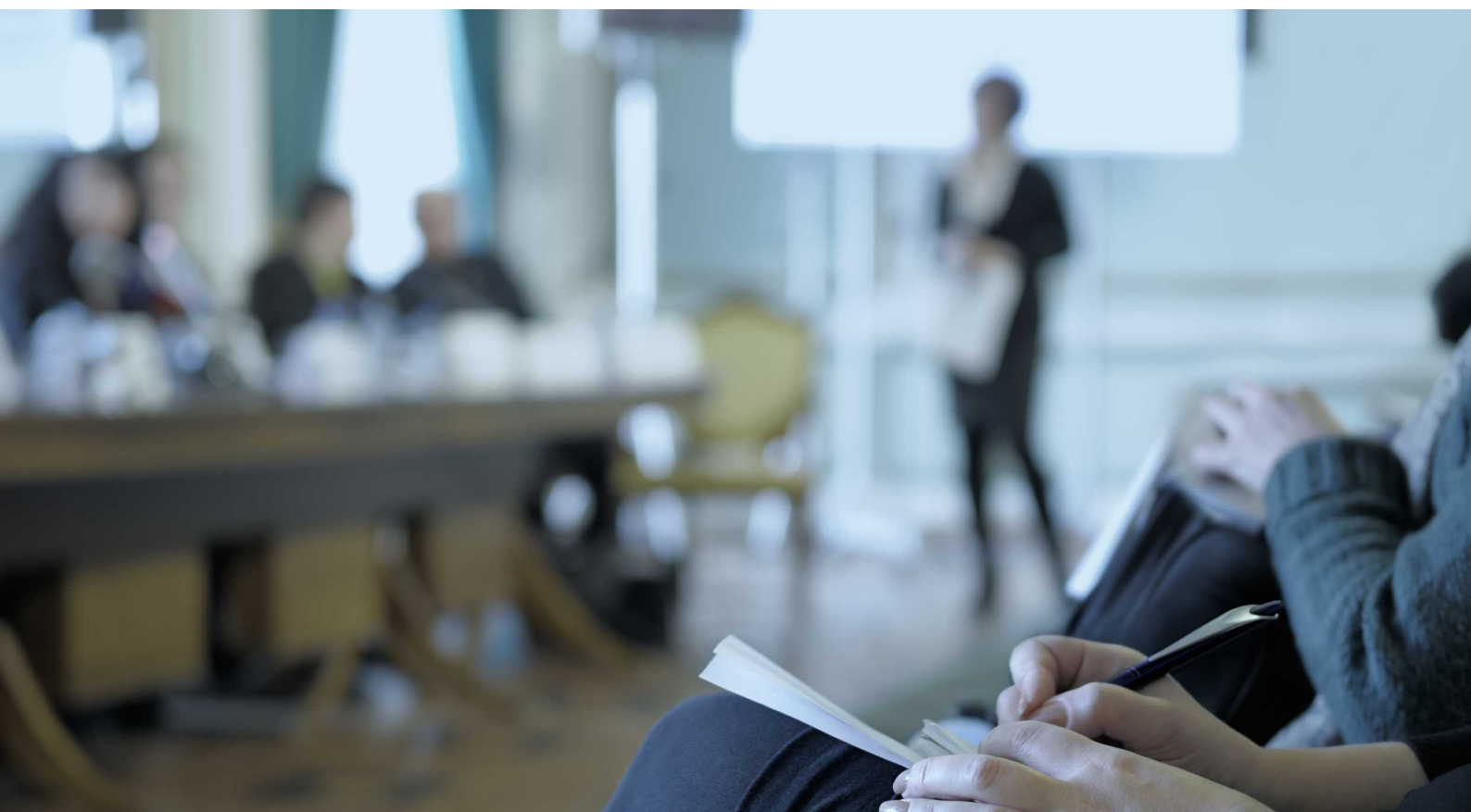
on and designing (Big) Data Architectures. He's the first known to publish a Big Data Reference Architecture, positioning both Data Models and Technology. Before joining Xomnia, Martijn held positions like Lead Business Analyst at Royal Philips to develop the Conceptual, Logical and Physical Data Models to support the new Philips application landscape in HealthTech services. As Product Owner at Randstad, he introduced Randstad to predictive analytics in Marketing and HR.



Modeling of Reference Schemes

Dr. Terry Halpin, Professor at INTI International University, Malaysia

This presentation illustrates the different ways in which a variety of data modeling approaches (ER, UML, ORM, RDB, OWL) support the important task of identifying individuals of interest to the business domain. In so doing, it also exposes the limitations and hurdles that might need to be overcome when transforming from one approach to the other. It addresses the following categories of reference scheme: (1) simple reference schemes; (2) compound reference schemes involving multiple components; (3)

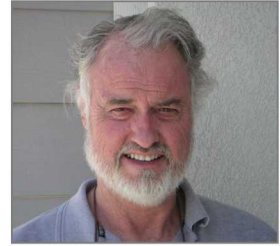


disjunctive reference schemes involving optional components; and (4) context-dependent reference schemes, where the preferred identifier for an entity varies according to the context.

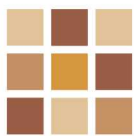
In natural language, one usually refers to individual things by means of proper names (e.g. 'Barack Obama') or definite descriptions (e.g. 'The current president of the United States'). Languages for data modeling and/or ontology modeling vary considerably in their support for such natural reference schemes. An understanding of these differences is important both for modeling reference schemes within such languages and for transforming models from one language to another. This presentation provides a critical and comparative review of reference scheme modeling within the Unified Modeling Language (version 2.5), the Barker dialect of Entity Relationship modeling, Object-Role Modeling (version 2), relational database modeling, and the Web Ontology Language (version 2.0). Using illustrative examples, it identifies which kinds of reference schemes can be captured within specific languages as well as those reference schemes that cannot be. The analysis

covers simple reference schemes, compound reference schemes, disjunctive reference and context-dependent reference schemes.

Dr. Terry Halpin is a Research Fellow in Computer Science at INTI International University (Malaysia), and a data modeling consultant. He previously held senior faculty positions in computer science at The University of Queensland (Australia), and professorships in computer science at Neumont University (USA) and INTI International University. His prior industrial experience includes many years in data modeling technology at Asymetrix Corporation, InfoModelers Inc., Visio Corporation, Microsoft Corporation and LogicBlox. His doctoral thesis formalized Object-Role Modeling (ORM/NIAM), and his current research focuses on conceptual modeling and rule-based technology. He has authored over 200 technical publications and nine books, and has co-edited nine books on information systems modeling research. He is a member of IFIP WG 8.1 (Information Systems), is an associate editor or reviewer for several academic journals, is a regular columnist for the Business Rules Journal, and is a recipient of the DAMA International Achievement Award for Education and the IFIP Outstanding Service Award.



Be in the Zone!



CaseTalk

