Keith McCormick | Rick van der Lans | Wayne Eckerson | Nigel Turner | Dave Wells | Egge van der Poel | Piethein Strengholt | Tim Schulteis | Jeroen Vermunt | Antoine Stelma | Peter Boncz

**UTRECHT**   **MARCH 25-26, 2020**

# DATA WAREHOUSING & BUSINESS INTELLIGENCE
## SUMMIT 2020

**Big Data, Analytics & Data Science, DataOps, Cloud Datawarehousing, Data Governance, DataVault**

**Acclaimed speakers**

Wayne Eckerson, Keith McCormick, Dave Wells, Nigel Turner, Egge van der Poel, Piethein Strengholt, Tim Schulteis, Peter Boncz, Jeroen Vermunt, Antoine Stelma and Rick van der Lans

- Concrete approach to cloud data warehousing
- Successful implementation of self-service analytics and BI
- DataOps: best practices, trends, tips and techniques
- Setting up a data integration architecture for the digital future
- The importance of data preparation with Machine Learning
- Ten guidelines for modern data architectures
- Establishing successful data governance processes
- Experiences with a data lake and an analytics lab
- Combining Data Vault, Ensemble and data virtualisation

**INFORMATION AND REGISTRATION:**
**WWW.DWBISUMMIT.COM**

**POST-CONFERENCE WORKSHOPS - MARCH 30-31, 2020**

Also join Keith McCormick and Dave Wells for two highly practical hands-on workshops

## PUTTING MACHINE LEARNING TO WORK

## CLOUD DATAWAREHOUSING

Follow us @AdeptEventsNL
Event hash tag: #dwbisummit

///AdeptEvents

# DATA WAREHOUSING & BUSINESS INTELLIGENCE
## SUMMIT 2020

There is a huge shortage of business intelligence and data warehouse specialists, the avalanche of new big data technology is relentless, the amount of data to be processed is increasing and the number of users who need to analyse data is increasing as well. But there is more. Businesses require more transparency with regard to data deliveries due to new laws and regulations, and because data usage must be regulated more tightly.

At the same time, organizations want to use their data widely, effectively, and efficiently. Therefore, existing systems must be upgraded or replaced. New technologies and techniques must be deployed. Sharing experiences with others and learning from the successes and mistakes of others is indispensable and can sometimes contribute in turning a project into a successful project.

### Some topics that will be discussed these days:
- Concrete approach to cloud data warehousing
- Successful implementation of self-service analytics and BI
- DataOps: best practices, trends, tips and techniques
- Setting up a data integration architecture for the digital future
- The importance of data preparation with Machine Learning
- Ten guidelines for modern data architectures
- Establishing successful data governance processes
- Experiences with a data lake and an analytics lab
- Combining Datavault, Ensemble and data virtualisation.

This conference offers practical guidelines and do's and don'ts to help you with these current and impending issues.

You will meet well-known speakers and thought leaders from the Netherlands and abroad, including including Wayne Eckerson, Keith McCormick, Dave Wells, Nigel Turner, Egge van der Poel, professor Peter Boncz, Piethein Strengholt, Tim Schulteis, Jeroen Vermunt, Antoine Stelma and Rick van der Lans. This first-class line-up of speakers are ready to share their knowledge and experience with you.

### Who should attend
The two day DW&BI Summit is geared to for IT Executives, IT Management and Architects, business intelligence and data warehousing professionals who wish to take a detailed and practical look at the latest developments in Data Warehousing and Business Intelligence. The following professionals should attend:
Sponsors of BI and DW programs, Business technology managers, IT executives and managers, BI/DW project managers, Data warehousing architects, Business intelligence practitioners, Business analysts, Data scientists, Technology architects, Data architects and data modelers, Project and program managers, Data integrators, Developers of BI and DW systems, Business and IT consultants.

### Limited time? Join us one day
Can you only attend one day? It is possible to attend only the first or only the second conference day and of course the full conference. The presentations by our speakers have been selected in such a way that they can stand on their own. This enables you to attend the second conference day even if you did not attend the first (or the other way around).

## POST-CONFERENCE WORKSHOPS - MARCH 30-31, 2020
Following the conference we will run the hands-on workshop *Putting Machine Learning to Work*. Attend this course by Keith McCormick to understand when to apply supervised or unsupervised machine learning models. In parallel Dave Wells will run a highly interactive two day workshop *Cloud Datawarehousing*. What are the benefits, techniques, and challenges of migrating an existing data warehouse to the cloud?

# CONFERENCE OUTLINE

## WEDNESDAY, MARCH 25

**Session 1**
**Data disruption cannot go without IT disruption**
Rick van der Lans, Managing Director, R20/Consultancy

**Session 2A**
**How Data Science is creating inequality in healthcare. And why that's a good thing.**
Egge van der Poel, Clinical Data Scientist, Erasmus MC

**Session 2B**
**Data Driven: more than just technology**
Tim Schulteis, Data Officer, APG

**Session 3A**
**Making Self-Service Analytics Work: Organizational, Architectural, and Governance Issues**
Wayne Eckerson, Founder & Principal Consultant, Eckerson Group

**Session 3B**
**Modern Data Management & Data Integration**
Piethein Strengholt, Lead Data Architect, ABN AMRO

**Session 4**
**Data Governance and Architecture – Making the connections**
Nigel Turner, Principal Information Management Consultant, Global Data Strategy

**Session 5**
**Cloud Data Warehousing: Planning for Data Warehouse Migration**
Dave Wells, Director, Eckerson Group

## THURSDAY, MARCH 26

**Session 6**
**Data Preparation for Machine Learning. Why Feature Engineering Remains a Human-Driven Activity**
Keith McCormick, Senior Consultant, The Modeling Agency

**Session 7A**
**Cloud Database Sytems in-depht: how do they work and how do they compare**
Peter Boncz, Database researcher, Centrum Wiskunde & Informatica (CWI)

**Session 7B**
**Data routes: combining data vault, ensemble modelling en data virtualisation**
Antoine Stelma, Lead Data Architect, Connected Data Group

**Session 8A**
**Modernizing Data Governance for the Age of Self-Service Analytics**
Dave Wells, Director, Eckerson Group

**Session 8B**
**Managing and exploring data using a data lake and a analytics lab**
Jeroen Vermunt, Group Lead Data Governance, ASML

**Session 9**
**Best Practices in DataOps: Trends, Tips, and Techniques for Creating and Managing Modern Data Pipelines**
Wayne Eckerson, Founder & Principal Consultant, Eckerson Group

**Session 10**
**Ten practical guidelines for modern data architectures**
Rick van der Lans, Managing Director, R20/Consultancy

**Daily schedule:**

| | | | | |
|---|---|---|---|---|
| 09:30 – 09:45 | Opening by Conference Chairman | | 14:00 – 15:00 | Session 3A and Session 3B |
| 09:45 – 11:00 | Session 1 | | 15:00 – 15:15 | Coffee break |
| 11:00 – 11:15 | Coffee break | | 15:15 – 16:15 | Session 4 |
| 11:15 – 11:45 | Case study | | 16:15 – 17:15 | Session 5 |
| 11:45 – 13:00 | Session 2A and Session 2B | | | |
| 13:00 – 14:00 | Lunch | | On the 25th of March, there will be a reception after the final session | |

### 1. Data disruption cannot go without IT disruption
(Dutch spoken)
**Rick van der Lans, Managing Director, R20/Consultancy**

The data-driven organization, digital transformation, and the data economy are all hot topics in board rooms. They all imply that the role of data changes. Organizations want to use data more widely, more effectively, and more efficiently: a real data disruption. However, data disruption clearly raises the bar for developing IT systems, because it leads to more complex IT systems, involving e.g. AI, sensor technology, and real-time analytics. Luckily, so much high-tech technology is available, we can almost build whatever the business needs. So, the technology is ready, but is IT itself ready? You would think so. IT has more than fifty years of experience in data modeling, data architectures, data strategies, data warehouses and databases. If we look at our track record for developing more traditional IT systems, we have to conclude that some of our projects have not been delivered on time and within budget, and are sometimes completely cancelled. So, how will IT perform if the complexity dramatically increases? What can we actually learn in the future from all the experiences we have gained? What do we have to change in order to participate in this increasingly data-driven economy in which digital transformation is the magic word for everyone? This keynote addresses this issue, and discusses recommendations on how IT specialists and IT management need to change to be able to address the actual data disruption: the IT disruption.

- How good is our track record with respect to developing IT systems?
- Why don't we deploy more code generators and self-driving technology?
- Has data modeling actually changed in the last 30 years?
- How should the IT specialist change to be ready for the data disruption?
- The importance of in-depth IT knowledge at the top of the organization.

### 2A. How Data Science is creating inequality in healthcare. And why that's a good thing. (Dutch spoken)
**Egge van der Poel, Clinical Data Scientist, Erasmus MC**

In this lecture Egge van der Poel will demonstrate what we can learn from Physics (e.g. the discovery of cosmic microwave background and the discovery of the Higgs particle) and from toddlers... He will share his experience creating value (knowledge and/or euro's) from data, using examples from his work as a Data Scientist in profit and non-profit sectors. With some remarkable examples from the clinical practice of the Erasmus Medical Centre he clearly shows that Big

Data can quickly lead to tangible added value. He will also address his vision how to overcome the shortage of relevant talent.
Egge will discuss Data Science and the necessity to:

- Develop a common language
- Embrace true collaboration
- Take away the mystics surrounding technology
- Study understand practical examples
- Develop new roles for professionals
- Facilitate exchange across company/sector borders
- Create speed
- Demonstrate transformational leadership

### 2B. Data Driven: more than just technology (Dutch spoken)
**Tim Schulteis, Data Officer, APG**

APG is the largest pension provider in the Netherlands and sees data as a crucial asset from current and future business operations. The government is increasingly withdrawing from a sufficient retirement provision (raising the state retirement age, decreasing the pension accrual), so insight into the personal situation of participants and offering action perspective is crucial. APG also wants to be a leader as an executor and investor.
All this has led to the earmarking of data as a strategic asset. But how do you get from that ambition to execution on the various axes of technology, capability, culture and organization? This presentation tells an integral story about the journey that has been made in the past period: what went well, what didn't, what did we learn from it, where are we now?
The development of an appropriate architecture, the building of the right knowledge and skills, the combination with modern working methods and the challenges of collaboration across business units will be discussed:

- Business drivers
- Architecture and knowledge
- Culture and agility
- Organization and sponsorship
- Critical success factors

### 3A. Making Self-Service Analytics Work: Organizational, Architectural, and Governance Issues
**Wayne Eckerson, Founder & Principal Consultant, Eckerson Group**

Self-service analytics has been the holy grail of data analytics leaders for the past two decades. Although analytical tools have improved significantly, it is notoriously difficult to achieve the

promise of self-service analytics. This session will explain how to empower business users to create their own reports and models without creating data chaos. Specifically, it examines seven factors for leading a successful BI program: right roles, right processes, right tools, right organization, right architecture, right governance, and right leadership. Ultimately, it will show how to build a self-sustaining analytical culture that balances speed and standards, agility and architecture, and self-service and governance.

You will learn:
- Trends and business dynamics driving analytics adoption
- The conundrum of self-service analytics
- Success factors for leading a successful BI program
- How to survive and thrive in the new world of big data analytics
- How to increase user adoption and facilitate self service

## 3B. Modern data management & data integration

(Dutch spoken)

**Piethein Strengholt, Lead Data Architect, ABN AMRO**

The digital future: think big, think highly distributed data, think in ecosystems. What Integration Architecture is needed to play an important role in the digital world with eco-system with FinTech company's and other banks? Do Enterprise Data Warehouses still have a role in this landscape? This is what the presentation is about. Piethein will also cover:
- Core Data Integration patterns
- Data Ownership and Data Governance
- Metadata as the glue
- Data Control when data is highly distributed
- Data distribution in the Cloud

## 4. Data Governance and Architecture – Making the connections

**Nigel Turner, Principal Information Management Consultant, Global Data Strategy**

With data increasingly being seen as a critical corporate asset, more organisations are embracing the concepts and practices of Data Governance. As a result Data Governance is today one of the hottest topics in data management, focusing both on how Governance driven change can enable companies to gain better leverage from their data through enhanced Business Intelligence, Data Analytics and so on, and also to help them design and enforce the controls needed to ensure they remain compliant with increasingly stringent laws and regulations, such as GDPR.

Despite this rapidly growing focus, many Data Governance initiatives fail to meet their goals, with only around one in five

fully achieving expectations. Why is the failure rate so high? There are many factors, but one key reason is that implementing Data Governance without aligning it with a defined enterprise and data architecture is fraught with dangers. Linking Architecture with data accountability, a core principle of Data Governance, is essential. This session will outline why Data Governance and Architecture should be connected, how to make it happen, and what part Business Intelligence and Data Warehousing plays in defining a robust and sustainable Governance programme.

This talk will cover:
- What is Data Governance and what it is not
- Key reasons for Data Governance failure & disappointment
- The key components of enterprise architecture – Business, Process and Data
- The synergies between architecture and Governance – how do they reinforce each other?
- How artefacts from both disciplines can be combined and applied to ensure success
- The implications for Business Intelligence and Data Warehousing
- Several use cases of successes and lessons learned

## 5. Cloud Data Warehousing: Planning for Data Warehouse Migration

**Dave Wells, Director, Eckerson Group**

Cloud data warehousing helps to meet the challenges of legacy data warehouses that struggle to keep up with growing data volumes, changing service level expectations, and the need to integrate structured warehouse data with unstructured data in a data lake. Cloud data warehousing provides many benefits, but cloud migration isn't fast and easy. Migrating an existing data warehouse to the cloud is a complex process of moving schema, data, and ETL. The complexity increases when architectural modernization, restructuring of database schema or rebuilding of data pipelines is needed.

This session provides an overview of the benefits, techniques, and challenges when migrating an existing data warehouse to the cloud. We will discuss the pros and cons of cloud migration, explore the dynamics of migration decision making, and look at migration pragmatics within the framework of a step-by-step approach to migrating. The tips and techniques described here will help you to make informed decisions about cloud migration and address the full scope of migration planning.

You Will Learn:
- The what and why of cloud data warehousing
- The benefits and challenges of cloud data warehousing
- Migration analysis and decision making
- Technology roles in migration to the cloud
- A step-by-step framework for data warehouse migration.

## 6. Data Preparation for Machine Learning; Why Feature Engineering Remains a Human-Driven Activity.
**Keith McCormick, Senior Consultant, The Modeling Agency**

This session will expose analytic practitioners, data scientists, and those looking to get started in predictive analytics to the critical importance of properly preparing data in advance of model building. The instructor will present the critical role of feature engineering, explaining both what it is and how to do it effectively. Emphasis will be given to those tasks that must be overseen by the modeler – and cannot be performed without the context of a specific modeling project. Data is carefully "crafted" by the modeler to improve the ability of modeling algorithms to find patterns of interest.

Data preparation is often associated with cleaning and formatting the data. While important, these tasks will not be our focus. Rather it is how the human modeler creates a dataset that is uniquely suited to the business problem.

You will learn:

- Construction methods for various data transformations
- The merits and limitations of automated data preparation technologies
- Which data prep tasks are best performed by data scientist, and which by IT
- Common types of constructed variables and why they are useful
- How to effectively utilize subject matter experts during data preparation

## 7A. Cloud Database Sytems in-depht: how do they work and how do they compare (Dutch spoken)
**Peter Boncz, Database Scientist, Centrum Wiskunde & Informatica**

An in-depth technical review and comparison of the new generation cloud database systems. Some of the topics that will be covered:

- An introduction of cloud data systems and current market overview (main contenders: Redshift, Athena, Databricks, Snowflake, BigQuery en SQLserver Azure)
- What does the architecture of these various systems look like, and where do they differ (aspects: query engine, data representation, elasticity and server-less-ness, data partitioning, special features)
- What is the economic model, and further implications of that.

## 7B. Data routes: combining data vault, ensemble modelling en data virtualisation (Dutch spoken)
**Antoine Stelma, Lead Data Architect, Connected Data Group**

Data vault, ensemble logical modeling, data virtualization and cloud are known to every BI or data warehouse specialist. But the big question is how you can use them together to develop real-life systems and then make optimum use of the power and possibilities of each component. This session explains how all can be used efficiently together. Key to this is the new concept of "data routes". Within a data and analytics architecture, data routes serve as a fuel for the virtual data presentation layer that is approached by end users for all their data needs.

The concept proposes a data-oriented way of processing that rests on the aforementioned issues such as data vault, ensemble modeling and data virtualization. A decoupling of data and technology is hereby realized whereby the emphasis is shifted to the characteristics of the data and the requirements set by use cases. The result is offered as a virtual (semantic) data layer to a broad group of data users. With the help of data virtualization, a virtual data collection is built up as a virtual data portal for data users.

- Does a Cloud Analytics platform offer a complete solution?
- How does the Data Routes concept fit in an existing data architecture for Data & Analytics?
- Is data modeling not necessary anymore?
- How do data routes and data virtualization fit together?
- From Ensemble logical modeling to data vault databases

## 8A. Modernizing Data Governance for the Age of Self-Service Analytics
**Dave Wells, Director, Eckerson Group**

Conventional data governance practices come from a simpler time when data management was free from many of today's challenges, such as self-service reporting and analytics. Traditional data governance focuses on enforcement of controls and gates. While controls and gates continue to be needed, these methods must be complemented with support for the autonomy and agility of the self-service world. Enforcement works together with prevention. Guides and guardrails reduce the need for gating. The need to exercise controls is minimized when curating, coaching, crowdsourcing, and collaboration are integral parts of governance processes. In a self-service world, every data stakeholder plays a part in data governance. In a world where personal data collection is accelerating and algorithms are abundant, ethical guidance is a critical component and an area where governance must take the lead.
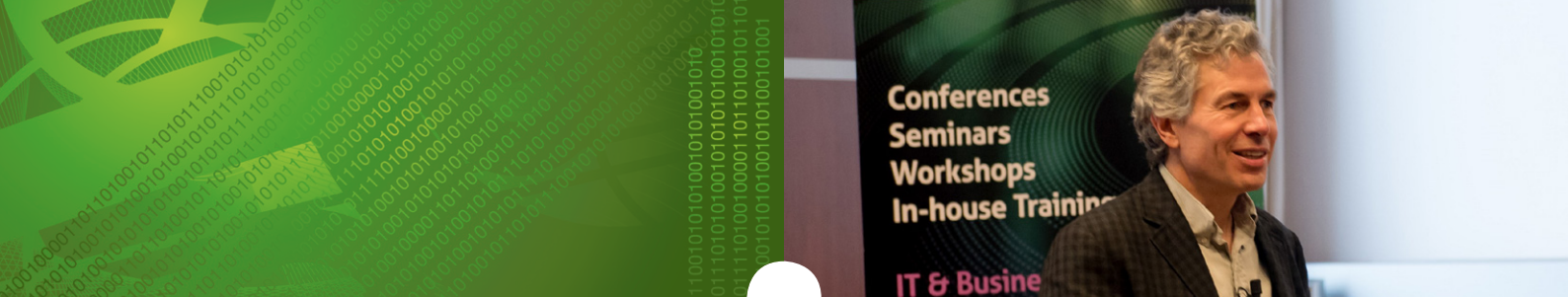
You will learn:

- Where governance fits within modern data ecosystems, from point of ingestion to reporting and analysis
- How various technologies support governance through the ecosystem
- Process challenges for governing self-service; supplementing controls with collaboration and crowd sourcing
- Engagement models for governing self-service
- Organizational challenges for governing self-service; moving from data stewards to stewardship, curation, and coaching
- Operational challenges for governing self-service; implementing a combination of gates, guardrails, and guides
- The challenges of data ethics and the role of data governance in meeting those challenges

## 8B. Managing and exploring data using a data lake and a analytics lab (Dutch spoken)
**Jeroen Vermunt, Group Lead Data Governance, ASML**

In the highly complex world of semiconductor manufacturing vast amounts of largely varied data are generated every day. ASML, world-leader manufacturer of machines for the production of semiconductors (chips), is implementing a central data lake to capture this data and make it accessible for reporting and analytics in a central environment. The data lake environment also includes an analytics lab for detailed exploration of data. Managing all this rapidly changing data imposes some very challenging requirements. In this session, real-life examples of how ASML approaches these challenges are presented.

- How do business users and data scientists discover information in the data lake without drowning in the amount and complexity?
- How can users be enabled to understand the data they want to consume, including were the data comes from (data lineage)?
- How do we make sure that the data can be trusted?
- How can ASML control and monitor that access to data is secure?
- What added value can the analytics lab bring?

## 9. Best Practices in DataOps: Trends, Tips and Techniques for Creating and Managing Modern Data Pipelines
**Wayne Eckerson, Founder & Principal Consultant, Eckerson Group**

When it comes to data analytics, you don't want to know "how the sausage is made." The state of most data analytics pipelines is deplorable. There are too many steps; too little automation and orchestration; minimal reuse of code and data; and a lack of

coordination between stakeholders in business, IT, and operations. The result is poor quality data delivered too late to meet business needs.

DataOps is an emerging approach for building data pipelines and solutions. This session will explore trends in DataOps adoption, challenges that organizations face in implementing DataOps, and best practices in building modern data pipelines. It will examine how leading-edge organizations are using DataOps to increase agility, reduce cycle times, and minimize data defects, giving developers and business users greater confidence in analytic output.

You will learn:

- What is DataOps and why you need it
- The dimensions of DataOps
- The state of DataOps adoption
- DataOps best practices and challenges

## 10. Ten practical guidelines for modern data architectures (Dutch spoken)
**Rick van der Lans, Managing Director, R20/Consultancy**

Many IT systems are more than twenty years old and have undergone numerous changes over time. Unfortunately, they can no longer cope with the ever-increasing growth in data usage in terms of scalability and speed. In addition, they have become inflexible, which means that implementing new reports and performing analyses has become very time-consuming. In short, the data architecture can no longer keep up with the current "speed of business change". As a result, many organizations have decided that it is time for a new, future-proof data architecture. However, this is easier said than done. After all, you don't design a new data architecture every day. In this session, ten essential guidelines for designing modern data architectures are discussed. These guidelines are based on hands-on experiences with designing and implementing many new data architectures.

- Which new technologies are currently available?
- What is the influence on the architecture of e.g. Hadoop, NoSQL, big data, data warehouse automation, and data streaming?
- Which new architecture principles should be applied nowadays?
- How do we deal with the increasingly paralyzing rules for data storage and analysis?
- What is the influence of cloud platforms?

# INTERNATIONALLY ACCLAIMED SPEAKERS

**KEITH MCCORMICK** is a highly accomplished professional senior consultant, mentor, and trainer, having served as keynote and moderator at international conferences focused on analytic practitioners and leadership alike. Keith has leveraged statistical software since 1990 along with deep expertise utilizing popular industry advanced analytics solutions such as IBM SPSS Statistics, IBM SPSS Modeler, AMOS, Answer Tree, popular open source and other tools involving text and big data analytics. Keith McCormick has guided organizations to establish highly effective analytical practices across industries, to include public sector, media, marketing, healthcare, retail, finance, manufacturing and higher education. He holds a very unique blend of tactical and strategic skill along with the business acumen to ensure superior project design, oversight and outcomes that align with organizational targets.

**RICK VAN DER LANS** is a highly-respected independent analyst, consultant, author, and internationally acclaimed lecturer specializing in data warehousing, business intelligence, big data, and database technology.
He has presented countless seminars, webinars, and keynotes at industry-leading conferences. For many years, he has served as the chairman of the annual European Enterprise Data and Business Intelligence Conference in London and the annual *Data Warehousing* and *Business Intelligence Summit* in The Netherlands. Rick helps clients worldwide to design their data warehouse, big data, and business intelligence architectures and solutions and assists them with selecting the right products. He has been influential in introducing the new logical data warehouse architecture worldwide which helps organizations to develop more agile business intelligence systems. Over the years, Rick has written hundreds of articles and blogs for newspapers and websites and has authored many educational and popular white papers for a long list of vendors. He was the author of the first available book on SQL, entitled including *Introduction to SQL*, which has been translated into several languages with more than 100,000 copies sold. More recently, he published his book *Data Virtualization for Business Intelligence Systems*. In 2018 Rick ranked sixth place as most influential BI-analist worldwide on the Onalytica Influencer List.

**WAYNE ECKERSON** is an internationally recognized thought leader in business intelligence and analytics who thinks critically, writes clearly, and presents persuasively about complex topics. He is a best-selling author, sought-after consultant, and noted speaker. Eckerson has written two books: "The Secrets of Analytical Leaders: Insights from Information Insiders."(2012) and "Performance Dashboards: Measuring, Monitoring, and Managing Your Business" (2005/2010), as well as dozens of reports and hundreds of articles during his 25 years in the field. In addition, Eckerson has shared his insights and advice with a range of companies, including New Balance, Karl Storz, PetSmart, Carlsberg Group, and Brigham Young University.Eckerson is founder and principal consultant at Eckerson Group, a full-service consulting firm that provides thought leadership and education to help organizations get more value from data and analytics.

**NIGEL TURNER** is Principal Information Management Consultant for EMEA at Global Data Strategy Ltd. and Vice-Chair of the Data Management Association of the UK. Nigel has worked in Information Management for over 25 years, both as an in-house deliverer of Information Management solutions at British Telecommunications plc and subsequently as an external consultant to more than 150 clients, including British Gas, UK Environment Agency, Intel US and others. He also works as a part time project manager at Cardiff University's National Software Academy. Nigel is a sought after speaker at conferences on information management and is based in Cardiff, UK.

**DAVE WELLS** is the Data Management Practice Director at Eckerson Group, a data analytics research and consulting organization. He is an internationally recognized thought leader in data management, a frequent speaker at industry conferences, and a contributing author to industry publications. Dave brings a unique perspective to data management based on five decades of working with data in both technical and business roles. He works at the intersection of information management and business management, where real value is derived from data assets.

Dave is an industry analyst, consultant, and educator dedicated to building meaningful and enduring connections throughout the path from data to business value. Much of his work today is focused on modernization — updating turn-of-the-century BI architecture to optimize for big data and analytics, sustaining the value of legacy data warehouses with cloud technology and data lake compatibility, and rethinking data governance to work well in self-service and agile cultures.
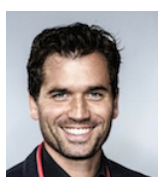
**EGGE VAN DER POEL** holds a BA in Philosophy and a PhD in Elementary Particle Physics. He searched for (and found) the Higgs particle at the Large Hadron Collider of CERN, in Geneva. After his PhD at CERN he trained to be a Medical Physicist and was a Big Data consultant at KPMG before taking on the newly created function of Clinical Data Scientist at the Erasmus Medical Centre in Rotterdam. In addition he advices many organisations about Big Data, teaches courses on it at various academic institutions and is an experienced public speaker on the subject of Big Data. He is an Executive Professor at TIAS Business School (www.tias.edu) and Academic Director at the Jheronimus Academy of Data Science (www.jads.nl).

**PIETHEIN STRENGHOLT** is a Lead Data Architect with a passion for technology. He is responsible for setting and driving the strategical agenda for ABN AMRO for topics such as Data Management, Data Integration and Cloud. Engineering experienced in Data Management, Data Integration & Cloud (AWS & Azure). Piethein is commercially skilled, with more than ten years consultancy background.

**TIM SCHULTEIS** implemented various initiatives in datawarehousing, Business Intelligence and analytics, initially mainly focusing on concepts and processes. Observing that the quality of concepts is only a minor success factor in digital transformation he focused on change management and leadership styles. Schulteis built a new and energetic unit at APG covering the total datamanagement playing field, therewith conceiving a new data architecture, implementing group wide data management and extracting value from data with a cutting edge analytics and data science capability.

Since August 2019 Schulteis is director Group Data Office at APG and bears responsibility for further extension of the value that data has as a crucial enabler of corporate strategy, both on content and data driven culture.

**JEROEN VERMUNT** brings over 25 years of experience in several fields of IT and data. He held multiple positions in the field of business intelligence and data analytics, ranging from developer, consultant and architecture roles to managing Business Intelligence Competence Centers. In 2012 he joined ASML where he is now growing a team that is improving the governance of the large amounts of data which are continuously being generated by ASML equipment. In this role he has a strong focus on (but not limited to) governing the data in ASML's central data lake and the corresponding analytics lab.



**ANTOINE STELMA** has over 20 years of experience as a data architect and trainer for Data Warehousing and Business Intelligence. He co-founded Connected Data Group together with Erik Fransen. Connected Data Group supports data-driven organizations with smart solutions for Data & Analytic architectures. As an instructor, Antoine teaches the concepts mixed with daily examples from his practice at Connected Data Academy and Genesee Academy. With an extensive background in Enterprise Data Warehousing, specialized in Data Vault Modelling, Data Virtualization and designing Data & Analytics platforms, Antoine shares his passion, research, and best practices, based on existing use cases.



**PETER BONCZ** holds appointments as tenured researcher at CWI and professor at VU University Amsterdam. His academic background is in core database architecture, with the architecture of MonetDB the main topic of his PhD thesis --MonetDB won the 2016 ACM SIGMOD systems award. This work focused on architecture-conscious database research, which studies the interaction between computer architecture and data management techniques. His specific contributions are in cache-conscious join methods, query and transaction processing in columnar database systems, and vectorized query execution. He has a strong track record in bridging the gap between academia and commercial application, receiving the Dutch ICT Regie Award 2006 for his role in the CWI spin-off company Data Distilleries. In 2008 he founded a new CWI spin-off company called Vectorwise, dedicated to state-of-the art business intelligence technology. He is also the co-recipient of the 2009 VLDB 10 Years Best Paper Award, and in 2013 received the Humboldt Research Award. His current interests are data systems architectures covering various angles such as database-as-a-service in the cloud, graph and network databases, and databases that can take advantage of heterogeneous processors and modern storage media.

# CONGRES-APP



**Download the DW&BI Summit Conference App** (integrated in the BI-Platform app)

# PUTTING MACHINE LEARNING TO WORK

## Translating Organizational Challenges into Supervised & Unsupervised Learning Solutions

Supervised learning solves modern analytics challenges and drives informed organizational decisions. Although the predictive power of machine learning models can be very impressive, there is no benefit unless they inform value-focused actions. Models must be deployed in an automated fashion to continually support decision making for residual impact. And while unsupervised methods open powerful analytic opportunities, they do not come with a clear path to deployment. This course will clarify when each approach best fits the business need and show you how to derive value from both approaches.

Regression, decision trees, neural networks – along with many other supervised learning techniques - provide powerful predictive insights when historical outcome data is available. Once built, supervised learning models produce a propensity score which can be used to support or automate decision making throughout the organization. We will explore how these moving parts fit together strategically.

Unsupervised methods like cluster analysis, anomaly detection, and association rules are exploratory in nature and don't generate a propensity score in the same way that supervised learning methods do. So how do you take these models and automate them in support of organizational decision-making? This course will show you how.

This course will demonstrate a variety of examples starting with the exploration and interpretation of candidate models and their applications. Options for acting on results will be explored. You will also observe how a mixture of models including business rules, supervised models, and unsupervised models are used together in real world situations for various problems like insurance and fraud detection.

### You Will Learn
- When to apply supervised versus unsupervised modeling methods
- Options for inserting machine learning into the decision making of your organization
- How to use multiple models for estimation and classification
- Effective techniques for deploying the results of unsupervised learning
- Interpret and monitor your models for continual improvement
- How to creatively combine supervised and unsupervised models for greater performance

### Who is it for?
Analytic Practitioners, Data Scientists, IT Professionals, Technology Planners, Consultants, Business Analysts, Analytic Project Leaders.



www.adepte

## Course description

**Model Development Introduction**
- Current Trends in AI, Machine Learning and Predictive Analytics
  - Algorithms in the News: Deep Learning
  - The Modeling Software Landscape
  - The Rise of R and Python: The Impact on Modeling and Deployment
  - Do I Need to Know About Statistics to Build Predictive Models?

**Strategic and Tactical Considerations in Binary Classification**
- What is an Algorithm?
- Is a "Black Box" Algorithm an Option for Me?
- Issues Unique to Classification Problems
  - Why Classification Projects are So Common
  - Why are there so many Algorithms?

**Data Preparation for Supervised Models**
- Data Preparation Law
- Integrate Data Subtasks
  - Aggregations: Numerous Options
  - Restructure: Numerous Options
  - Data Construction
    - Ratios and Deltas
    - Date Math
    - Extract Subtask

**The Tasks of the Model Phase**
- Optimizing Data for Different Algorithms
- Model Assessment
  - Evaluate Model Results
    - Check Plausibility
    - Check Reliability
  - Model Accuracy and Stability
  - Lift and Gains Charts
- Modeling Demonstration
  - Assess Model Viability
  - Select Final Models
- Why Accuracy and Stability are Not Enough

- What to Look for in Model Performance
- Exercise Breakout Session
  - Select Final Models
  - Create & Document Modeling Plan
  - Determine Readiness for Deployment
- What are Potential Deployment Challenges for Each Candidate Model?

**What is Unsupervised Learning?**
- Clustering
- Association Rules
- Why most organizations utilize unsupervised methods poorly
  - Case Study #1: Finding a new opportunity
  - Case Studies 2, 3, and 4: How do supervised and unsupervised work together
  - Exercise Breakout Session: Pick the right approach for each case study
- Data Preparation for Unsupervised
  - The importance of standardization
  - Running an analysis directly on transactional data
- Unsupervised Algorithms:
  - Hierarchical Clustering
  - K-means
  - Self-Organizing Maps
  - K Nearest Neighbors
  - Association Rules
- Interpreting Unsupervised
  - Exercise Breakout Session: Which value of K is best?
  - Choosing the right level of granularity
  - Reporting unsupervised results

**Wrap-up and Next Steps**
- Supplementary Materials and Resources
- Conferences and Communities
- Get Started on a Project!
- Options for Implementation Oversight and Collaborative Development

## FOR DETAILED INFORMATION PLEASE VISIT
### WWW.ADEPTEVENTS.NL/WML-EN

# CLOUD DATA WAREHOUSING

## Practical Migration Techniques for Legacy Data Warehouses

Cloud data warehousing helps to meet the challenges of legacy data warehouses that struggle to keep up with growing data volumes, changing service level expectations, and the need to integrate structured warehouse data with unstructured data in a data lake. Cloud data warehousing provides many benefits, but cloud migration isn't fast and easy. Migrating an existing data warehouse to the cloud is a complex process of moving schema, data, and ETL. The complexity increases when architectural modernization, restructuring of database schema or rebuilding of data pipelines is needed.

This session describes the benefits, techniques, and challenges of migrating an existing data warehouse to the cloud. We'll overview the major cloud data warehouse providers and examine some technologies that can ease the pain of cloud migration. Migration pragmatics are described within the framework of a step-by-step approach.

Moving an existing data warehouse to the cloud is not quick, and it isn't easy, but there are real and substantial benefits. Most companies avoid the "big bang" and migrate incrementally. The techniques and technologies described here will help you to start quickly and navigate through the migration roadmap smoothly.

This is a participative workshop. Participants should arrive with knowledge of one or more on-premises data warehouses that they are considering moving to the cloud. You'll perform exercises to analyze the upside and downside of cloud migration, to make migration decisions, and to develop a migration plan.

### Learning objectives
- The what and why of cloud data warehousing
- The benefits of cloud data warehousing
- The challenges of cloud data warehousing
- Cautions about cloud migration
- How to perform migration analysis and make migration decisions
- A step-by-step approach to data warehouse migration
- Considerations for incremental migration
- The cloud data warehousing technology landscape
- The cloud migration technology landscape
- How to develop a cloud migration plan.

### Who is It For?
Data management and analytics program and project managers, Data management and analytics architects, Technology strategists and planners, Data engineers and data integration specialists, Anyone with a role in migrating legacy data warehouses to the cloud.

## Course description

### Cloud Data Warehousing – What and Why?

- The cloud migration trend
- Challenges of conventional data warehousing
- Overcoming the challenges

### Benefits of Cloud Data Warehousing

- Scalability
- Elasticity
- Managed infrastructure
- Cost savings
- Processing speed
- Deployment speed
- Fault tolerance
- Disaster recovery
- Security and governance
- RDBMS in the cloud

### Cloud Data Warehousing: Look Before You Leap

- Data movement
- Design and development
- Security and compliance
- Cloud and on-premises hybrid
- Portability

### To Cloud or Not to Cloud

- Data warehouse architecture
- Technical debt
- Technology architecture
- Networking
- Data warehouse operations
- Data center operations
- Disaster recovery and business continuity
- Scalability and elasticity
- Data types
- Agility
- Governance
- Security
- Compliance
- Data quality
- Self-service
- Commitment

### Cloud Migration SWOT Analysis

- Some guidelines
- Exercise 1 – SWOT for Migrating Your Data Warehouse
- Exercise 2 – Using the Migration SWOT Analysis
- Exercise 3 – SWOT for Not Migrating Your Data Warehouse
- Exercise 4 – Using the No Migration SWOT Analysis
- Exercise 5 – To Cloud or Not to Cloud: Making the Decision

### Technologies for Cloud Data Warehousing

- Cloud data warehouse functions
  - Data storage
  - Database loading
  - Data processing
  - SQL queries
  - Workload management
  - Metadata management
  - Security and access controls
  - Data protection
- Cloud data warehouse platforms
- Migration tools
  - Integration Platform as a Service (iPaaS)
  - Data warehouse automation
  - Data virtualization

### Step-by-Step Data Warehouse Migration

- The big picture
- The business case
- Architectural assessment
- Migration strategy
- Technology selection
- Migration
  - Planning
  - Schema migration
  - Data migration
  - ETL migration and data pipelines
  - Metadata migration
  - Migrating users and applications

## FOR DETAILED INFORMATION PLEASE VISIT
### WWW.ADEPTEVENTS.NL/CDW-EN

# INFORMATION
## DATA WAREHOUSING & BUSINESS INTELLIGENCE SUMMIT

## DATE AND TIME
The conference will take place on March 25 - 26, 2020 and the programme starts at 9:30 am and ends at 5:15 pm on both conference days. Registration commences at 8.30 am and we recommend that you arrive early.

## VENUE
The conference will be held at the Van der Valk Hotel in Utrecht. Always check our website or conference app prior to your departure to ensure you have the exact location and directions.
Van der Valk Hotel Utrecht
Winthontlaan 4-6
3526 KV Utrecht

## Contact details hotel:
E-mail: utrecht@valk.nl
Website hotel: **www.vandervalkhotelutrecht.nl.**
On the hotel **website** you can find a full itinerary and directions. The hotel is located on a 35 minutes drive from Amsterdam Schiphol Airport and is also easily accessible by public transport.

## HOW TO REGISTER
Please register online at www.dwbisummit.com. For registering by print, please scan the completed registration form and send this to **customerservice@dwbisummit.com**. We will confirm your registration and invoice your company by e-mail therefore please do not omit your e-mail address when registering.

## REGISTRATION FEE

|  | Full conference | One Day |
|---|---|---|
| Best rate (ends December 31, 2019)*: | € 1.139 | € 586,50 |
| Early registration (January 1 – February 19, 2020): | € 1.206 | € 621 |
| Regular registration (February 20 – March 25, 2020): | € 1.340 | € 690 |

The registration fee for the workshop *Putting Machine Learning to Work* or *Cloud Data Warehousing* when combined with attending the conference, is 1363 Euros and 1226.70 Euros when registering early.

*) Invoice will be sent in year 2020. Upon request we can send the invoice in year 2019. Members of the the Dutch KNVI BI&A as well as DAMA are eligible for 10 percent discount on the registration fee. All prices are VAT excluded.

## Team discounts
Discounts are available for group bookings of two or more delegates representing the same organization made at the same time. Ten percent off for the second and third delegate and fifteen percent off for all delegates when registering four or more delegates (all delegates must be listed on the same invoice).
This cannot be used in conjunction with other discounts.

## PAYMENT
Full payment is due prior to the conference. An invoice will be sent to you containing our full bank details including BIC and IBAN. Your payment should always include the invoice number as well as the name of your company and the delegate name.

For Credit Card payment please contact our office by e-mail mentioning your phone number so that we can obtain your credit card information.

## Cancellation Policy
Cancellations must be received in writing at least three weeks before the commencement of the conference and will be subject to a € 75,- administration fee. It is regretted that cancellations received within three weeks of the conference date will be liable for the full conference fee. Substitutions by other persons can be made at any time and at no extra charge.

## Cancellation Liability
In the unlikely event of cancellation of the conference for any reason, Adept Events' liability is limited to the return of the registration fee only. Adept Events will not reimburse delegates for any travel or hotel cancellation fees or penalties. It may be necessary, for reasons beyond the control of Adept Events, to change the content, timings, speakers, date and venue of the conference.

## Recordings and photography
Please be aware that still photography, video, and audio recording may occur at this event. By attending this event, you consent to have your image, photograph, likeness, picture, rendering, or audio recording utilized for Adept Events educational, marketing, and sales purposes. You hereby grant Adept Events the right to unrestricted use, reproduction, display, dissemination, publication, and distribution in any medium, provided that Adept Events will take measures on behalf of attendees against infringement and/or inappropriate use of your image, photograph, likeness, picture, rendering, and audio recording.

## MORE INFORMATION

📞 +31(0)172-742680

🌐 **https://www.dwbisummit.com/?lang=en**

@ **seminars@adeptevents.nl**

🐦 **@AdeptEventsNL | #dwbisummit**

in **https://www.linkedin.com/company/adept-events**

Visit our Business Intelligence and Data Warehousing website **www.biplatform.nl** and subscribe to our weekly newsletter

Download the DW&BI Summit conference app **DW&BI Summit Congres-App** (integrated in the BI-Platform app)

## SPONSORS AND MEDIA
We thank our sponsors for supporting our conference and providing media exposure.

**BI·Platform.**