



Keith McCormick

Rick van der Lans

Wayne Eckerson

Nigel Turner

Dave Wells

Egge van der Poel

Piethain Strengtholt

Tim Schulteis

Jeroen Vermunt

Antoine Stelma

Peter Boncz

UTRECHT

25 EN 26 MAART 2020

DATA WAREHOUSING & BUSINESS INTELLIGENCE SUMMIT 2020

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

Met internationale topsprekers en onafhankelijk analisten

Wayne Eckerson, Keith McCormick, Dave Wells, Nigel Turner, Egge van der Poel, Piethain Strengtholt, Tim Schulteis, Peter Boncz, Jeroen Vermunt, Antoine Stelma en Rick van der Lans

- Concrete aanpak van cloud datawarehousing
- Succesvol implementeren van self-service analytics en BI
- DataOps: best practices, trends, tips en technieken
- Opzetten van een dataintegratie-architectuur voor de digitale toekomst
- Het belang van data-preparation bij Machine Learning
- Tien richtlijnen voor moderne data-architecturen
- Inrichten van succesvol data governance
- Ervaringen met een data lake en een analytics lab
- Combineren van Datavault Ensemble en datavirtualisatie.

INFORMATIE EN REGISTRATIE:
WWW.DWBISUMMIT.COM

30 EN 31 MAART 2020

Keith McCormick en Dave Wells verzorgen aansluitend de praktische hands-on workshops

PUTTING MACHINE LEARNING TO WORK CLOUD DATAWAREHOUSING



Follow us @AdeptEventsNL
Event hash tag: #dwbisummit

AdeptEvents



DATA WAREHOUSING & BUSINESS INTELLIGENCE SUMMIT 2020

Er is een schrijnend tekort aan business intelligence en datawarehouse specialisten, de stortvloed van nieuwe ondersteunende technologieën is niet aflatend, de hoeveelheid te verwerken data neemt toe en ook het aantal gebruikers dat data wil analyseren neemt toe. Maar dit is niet alles. De business vereist meer transparantie bij dataleveringen en vanwege nieuwe wet- en regelgeving moet datagebruik strakker geregeld worden.

Tegelijkertijd willen organisaties hun data beter inzetten. Daarvoor moeten bestaande systemen opgewaarderd of vervangen worden. Hiervoor moeten nieuwe technologieën en technieken ingezet worden. Het delen van ervaringen met anderen en het leren van de successen en fouten van anderen is hierbij onmisbaar en kan soms het succes van een project betekenen.

Enkele onderwerpen die deze dagen aan bod komen

- Concrete aanpak van cloud datawarehousing
- Succesvol implementeren van self-service analytics en BI
- DataOps: best practices, trends, tips en technieken
- Opzetten van een dataintegratie-architectuur voor de digitale toekomst
- Het belang van data-preparation bij Machine Learning
- Tien richtlijnen voor moderne data-architecturen
- Inrichten van succesvol data governance
- Ervaringen met een data lake en een analytics lab
- Combineren van Datavault Ensemble en datavirtualisatie.

De DW&BI Summit biedt praktische richtlijnen, handvatten en do's en don'ts ter ondersteuning van de huidige en

komende vraagstukken. U ontmoet bekende sprekers en thought leaders uit binnen- en buitenland, Wayne Eckerson, Keith McCormick, Dave Wells, Nigel Turner, Egge van der Poel, prof. Peter Boncz, Pietheine Strengholt, Tim Schulteis, Jeroen Vermunt, Antoine Stelma en Rick van der Lans. Deze eersterangs line-up van sprekers wil graag hun kennis en ervaring met u delen.

DW&BI SUMMIT is bestemd voor u!

Het congres is gericht op beslissers en degenen die betrokken zijn bij de besluitvorming rondom Datawarehousing, Business Intelligence en data science projecten en die inzicht moeten hebben in de huidige mogelijkheden van BI én in de lange termijn ontwikkelingen. En op diegenen die betrokken zijn bij het invoeren van analytics en big data. Ook is het congres verplichte kost voor zij die verantwoordelijk zijn voor data management en voor het ontwerpen en ontwikkelen van datawarehouses en het opzetten van business intelligence omgevingen, waaronder datawarehouse architecten, BI-specialisten, dwbi-projectleiders, informatie-analisten, data scientists en technology-planners.

Tijdgebrek?

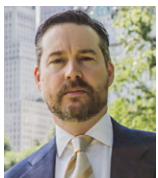
Heeft u slechts één dag de tijd om de DW&BI Summit te bezoeken? Maak een keuze uit de onderwerpen en kom op alleen 25 maart of op 26 maart. Het is namelijk ook mogelijk om alleen de eerste dag van het congres of alleen de tweede dag te bezoeken. De onderwerpen zijn zodanig gekozen dat zij op zich zelf staan zodat het ook mogelijk is om dag twee te volgen zonder dat u dag één heeft bijgewoond.

POST-CONFERENCE WORKSHOPS - 30 EN 31 MAART 2020

Aansluitend aan het congres organiseren wij de hands-on workshop *Putting Machine Learning to Work*. Een praktische tweedaagse workshop door Keith McCormick boordevol concrete voorbeelden en oefeningen over supervised en unsupervised machine learning. Tegelijkertijd verzorgt Dave Wells, de tweedaagse workshop *Cloud Datawarehousing*. Ook deze workshop bevat veel concrete cases en praktische oefeningen.



INTERNATIONALE TOPSPREKERS



KEITH MCCORMICK is een gerenommeerd consultant, coach en trainer op gebied van analytics en data science, en veelgevraagd spreker en dagvoorzitter op bekende internationale congressen op dit terrein.

Keith is reeds sinds 1990 specialist op gebied van statistische software en heeft diepgaande kennis en ervaring met bekende analytics oplossingen zoals IBM SPSS Statistics, IBM SPSS Modeler, AMOS, Answer Tree, populaire open source producten en andere tools voor text en big data analytics. Keith heeft vele organisaties begeleid met het opzetten van uiterst effectieve analytische teams in verscheidene branches, zoals (semi-)overheid, uitgeverijen, marketing, gezondheidszorg, detailhandel, financiële dienstverlening, maak-industrie en hogescholen. Hij beschikt over een unieke combinatie van operationele en strategische vaardigheden enerzijds en een scherpzinnig inzicht in bedrijfsdoelstellingen anderzijds dat hem in staat stelt complexe projecten in goede banen te leiden.



RICK VAN DER LANS is een gerespecteerd en onafhankelijk analist, adviseur, auteur en internationaal bekende spreker. Hij is gespecialiseerd in datawarehousing, business intelligence, big data en

databasetechnologie. Hij heeft vele seminars en webinars gepresenteerd en keynotes bij industry-leading conferenties. Hij weet als geen ander een goede balans te vinden tussen op de praktijk toegesneden technologische en strategische zaken. Al meer dan tien jaar is hij de voorzitter van de jaarlijkse *European Enterprise Data and Business Intelligence Conference* in Londen en de jaarlijkse *Data Warehousing en Business Intelligence Summit* in Nederland. Rick helpt cliënten wereldwijd met het ontwerpen van hun data warehouse, big data en business intelligence

architecturen en begeleid hen bij het selecteren van de geschikte producten. Hij is invloedrijk geweest bij het wereldwijd introduceren van de nieuwe logische datawarehouse architectuur waarmee organisaties meer flexibele business intelligence systemen kunnen ontwikkelen. Rick heeft honderden artikelen en blogs geschreven voor toonaangevende vakbladen en websites en van zijn hand zijn vele educatieve en populaire whitepapers verschenen voor een groot aantal leveranciers. Hij was de auteur van het eerste beschikbare boek over SQL, getiteld *Introduction to SQL*, dat in vele talen gepubliceerd is en waarvan meer dan 100.000 exemplaren verkocht zijn. Zijn laatst verschenen boek is *Data Virtualization for Business Intelligence Systems*. Rick is in 2018 op de bekende Analytica Influencer List geklommen tot de zesde meest invloedrijke BI analist wereldwijd.



WAYNE ECKERSON is al sinds begin jaren negentig een van de thought leaders op het gebied van datawarehousing, business intelligence en performance management.

Hij heeft een veelheid van soms baanbrekende onderzoeken op zijn naam staan en is o.a. auteur van "Performance Dashboards: Measuring, Monitoring, and Managing Your Business". Wayne is veelgevraagd spreker en consultant en schrijft helder en presenteert in overtuigende stijl over vaak complexe onderwerpen. Vele jaren is Wayne als directeur Research en Opleidingen actief geweest bij het bekende The Data Warehousing Institute (TDWI) waar hij de regie voerde over de onderzoeken en trainingen en waar hij voorzitter was van het BI Executive Summit. Wayne is principal consultant bij Eckerson Group, LLC, waar hij organisaties adviseert over de inzet van data en technologie om de besluitvorming te verbeteren.



NIGEL TURNER is Principal Information Management Consultant voor EMEA bij Global Data Strategy Ltd. en vicevoorzitter van de Data Management Association (DAMA) van het Verenigd Koninkrijk. Nigel heeft meer dan 25 jaar ervaring in Information Management, zowel binnen internationale concerns als British Telecommunications als in zijn rol als extern consultant bij meer dan 150 klanten wereldwijd waaronder British Gas, UK Environment Agency, Intel US om er een paar te noemen. Hij spreekt regelmatig op internationale congressen en werkt tevens als part time project manager bij Cardiff University National Software Academy.



DAVE WELLS is Data Management Practice Director bij Eckerson Group, een onderzoeks- en adviesorganisatie voor data-analyse. Hij is een internationaal erkende thought leader op het gebied van data management, een frequent spreker op vakconferenties en auteur van vakpublicaties. Dave biedt een uniek perspectief op data management op basis van vijf decennia werkervaring met data in zowel technische als business functies. Hij werkt op het snijvlak van informatiebeheer en business management, waarbij echte waarde wordt afgeleid van data assets.

Dave is een branche-analist, consultant en docent die zich toelegt op het bouwen van betekenisvolle en duurzame verbindingen gedurende het hele traject van data tot bedrijfswaarde. Veel van zijn werk is tegenwoordig gericht op modernisering: het updaten van de BI-architectuur en optimalisatie voor big data en analyses, het behouden van de waarde van legacy-datawarehouses met compatibiliteit met cloudtechnologie en data lakes, en het heroverwegen van data governance om goed te werken binnen self-service en agile culturen.



EGGE VAN DER POEL is een beta met een alfarandje. Gepromoveerd in de experimentele deeltjesfysica met daarnaast een Bachelor als filosoof. Hij heeft een nieuwsgierige aard en een passie voor

communicatie. Egge van der Poel werkt als Clinical Data Scientist in het Erasmus MC, als Academic Director bij JADS en als Executive Professor bij TIAS. Als freelancer adviseert hij verschillende stakeholders in de zorg, van overheid tot start-ups. Hij is een veelgevraagd spreker en dagvoorzitter op de thema's Big Data, Data Science en Innovatie. Daarnaast werkt hij als Clinical Data Scientist in het Erasmus MC, als Academic Director bij JADS en als Executive Professor bij TIAS. Als freelancer adviseert hij verschillende stakeholders in de zorg, van overheid tot start-ups. Egge is ook veelgevraagd spreker en dagvoorzitter op de thema's Big Data, Data Science en Innovatie.



PIETHEIJN STRENGHOLT is Lead Data Architect met een passie voor technologie. Hij is verantwoordelijk voor het bepalen en aansturen van de strategische agenda van ABN AMRO voor onderwerpen als Data Management, Data-integratie en Cloud. Hij heeft ook ontwikkelervaring in deze onderwerpen. Pietheijn is commercieel geschoold, met meer dan tien jaar consultancy-achtergrond.



TIM SCHULTEIS zette gedurende enkele jaren diverse datawarehouse-, Business Intelligence- en data science-initiatieven op, aanvankelijk met focus op concepten en processen. Geconfronteerd met de observatie dat de inhoudelijke component slechts een klein onderdeel is van succesvolle digitalisering, legde hij zich toe op verandermanagement en leiderschapsstijlen. Schulteis zette binnen Pensioenuitvoerder APG een nieuwe, energieke unit op die het volledige datamanagement werkveld omvat en waarmee een nieuwe data-architectuur werd neergezet, groepsbreed datamanagement werd geïmplementeerd en waarde uit data wordt gegenereerd met onder meer een hoogwaardige data science capability. Sinds augustus 2019 is Schulteis als directeur Group Data Office bij APG verantwoordelijk voor het verder uitbouwen van de waarde die data speelt als cruciale enabler van de bedrijfsstrategie op zowel inhoud als data driven cultuur.



JEROEN VERMUNT heeft ruim 25 jaar ervaring op verschillende gebieden binnen IT en data analyse en beheersing. Hij heeft diverse posities bekleedt op het gebied van business intelligence en data analyse,

varierend van programmeur, consultant en architect tot het leiden van business intelligence competence centers binnen diverse multinationals. Sinds 2012 werkt hij voor ASML waar hij verantwoordelijk is voor een team wat de grote hoeveelheden data die gegenereerd worden door ASML machines moet beheersen. In deze rol kijkt hij zeer sterk naar ASML's central data lake en het bijbehorende analytics lab, maar tevens naar alle andere toepassingen waar deze data gebruikt dient te worden.



ANTOINE STELMA is sinds 1999 actief als data architect en trainer binnen het vakgebied van Datawarehousing en Business Intelligence. Hij richtte samen met Erik Fransen de Connected Data Group op waarmee ze actief bezig zijn met het oplossen van Data & Analytics vraagstukken voor datagedreven organisaties. Als trainer verzorgt Antoine meerdere trainingen in zijn vakgebied voor zowel de Connected Data Academy als Genesee Academy.

Met een uitgebreide achtergrond in Enterprise Data Warehousing, gespecialiseerd in Data Vault Modelleren, Data Virtualisatie en het ontwerpen van Data en Analytics platformen, deelt Antoine zijn passie, ervaringen en onderzoek gebaseerd op bestaande use cases.



PETER BONCZ is aangesteld als vaste onderzoeker bij het CWI en hoogleraar aan de Vrije Universiteit Amsterdam. Zijn academische achtergrond ligt in de core database architectuur, met de architectuur

van MonetDB als hoofdonderwerp van zijn proefschrift -MonetDB won in 2016 de ACM SIGMOD systems award. Dit werk richtte zich op architectuurbewust databaseonderzoek, waarbij de interactie tussen computerarchitectuur en datamanagementtechnieken wordt onderzocht. Zijn specifieke bijdragen liggen op het gebied van cachebewuste join-methoden, query en transactieverwerking in column store databasesystemen, en gevectoriseerde query uitvoering. Hij heeft een sterke staat van dienst in het overbruggen van de kloof tussen de academische wereld en commerciële toepassingen, en ontving de Nederlandse ICT Regie Award 2006 voor zijn rol in het CWI spin-off bedrijf Data Distilleries. In 2008 richtte hij een nieuw CWI spin-off bedrijf op met de naam Vectorwise, gewijd aan state-of-the-art business intelligence technologie. Hij is ook mede-ontvanger van de 10 Years Best Paper Award 2009 van de VLDB en ontving in 2013 de Humboldt Research Award. Zijn huidige interesse gaat uit naar datasysteemarchitecturen met verschillende invalshoeken, zoals database-as-a-service in de cloud, graph- en netwerkdatabases, en databases die gebruik kunnen maken van heterogene processors en moderne opslagmedia.

CONGRES-APP



Download de **DW&BI Summit Congres-App** (onderdeel van de BI-Platform App)



PROGRAMMA



WOENSDAG 25 MAART

Sessie 1

Geen Data-disruptie zonder ICT-disruptie - Rick van der Lans

Sessie 2A

Hoe Big Data de zorg ongelijker maakt. En waarom dat goed is.
Egge van der Poel

Sessie 3A

Making Self-Service Analytics Work: Organizational, Architectural,
and Governance Issues
Wayne Eckerson

Sessie 4

Data Governance and Architecture – Making the connections - Nigel Turner

Sessie 5

Cloud Data Warehousing: Planning for Data Warehouse Migration - Dave Wells

Sessie 2B

Data Driven: meer dan technologie
Tim Schulteis

Sessie 3B

Modern Data Management en Data-integratie
Pietheine Strengtholt

DONDERDAG 26 MAART

Sessie 6

Data Preparation for Machine Learning. Why Feature Engineering Remains a Human-Driven Activity - Keith McCormick

Sessie 7A

Cloud Database Systemen: een vergelijking en blik onder de
motorkap
Peter Boncz

Sessie 8A

Modernizing Data Governance for the Age of Self-Service Analytics
Dave Wells

Sessie 9

Best Practices in DataOps: Trends, Tips, and Techniques for Creating and Managing Modern Data Pipelines - Wayne Eckerson

Sessie 10

Tien praktische richtlijnen voor moderne data-architecturen - Rick van der Lans

Sessie 7B

Data routes: combineren van data vault, ensemble modelling en
datavirtualisatie
Antoine Stelma

Sessie 8B

Beheersen en ontdekken van de waarde van data middels een data
lake en een analytics lab
Jeroen Vermunt

MAANDAG 30 EN DINSDAG 31 MAART

Putting Machine Learning to Work
Keith McCormick

Cloud Data Warehousing
Dave Wells

Globale dagindeling voor 25 en 26 maart:

09:30 – 09:45 Plenaire opening
09:45 – 11:00 Sessie 1
11:00 – 11:15 Koffiepauze
11:15 – 11:45 Praktijkcase
11:45 – 13:00 Sessie 2A en Sessie 2B
13:00 – 14:00 Lunch

14:00 – 15:00 Sessie 3A en Sessie 3B
15:00 – 15:15 Koffiepauze
15:15 – 16:15 Sessie 4
16:15 – 17:15 Sessie 5

Op 25 maart is er tevens een borrel vanaf 17:15.



1. Geen Data-disruptie zonder ICT-disruptie

Rick van der Lans, Managing Director, R20/Consultancy

Wat digitale transformatie, de data economie en een datagedreven organisatie ook precies betekenen, het impliceert altijd dat organisaties meer met data willen doen. Data moet als key business asset ingezet worden. Deze transformatie gaat voor een serieuze data-disruptie zorgen, hetgeen een groot effect op onze BI-systemen en datawarehouses zal hebben. Maar zijn wij ICT-ers klaar voor dit opwaarderen van onze systemen? Je zou denken van wel, met al meer dan vijftig jaar ervaring met datamodelleren, data-architecturen, datastrategieën, datawarehouses en databases. Desondanks moeten we bekennen dat er nog steeds fouten gemaakt worden en grote projecten nog altijd mislopen. Wat kunnen we eigenlijk voor de toekomst leren van alle ervaringen die we hebben opgedaan? Wat moeten we veranderen om mee te kunnen draaien in deze steeds meer datagedreven economie waarin voor iedereen digitale transformatie het magische woord is? Het is niet alleen een kwestie van nieuwe technologie aanschaffen. Wij ICT-ers moeten ook veranderen. Data-disruptie vereist ICT-disruptie.

- Hoe goed is ons track record wat betreft ICT-projecten?
- Waarom adopteren we niet meer generatoren en self-driving technologie?
- Is datamodelleren eigenlijk wel veranderd de laatste 30 jaar?
- Welke eigenschappen moet de ICT-er veranderen?
- Het belang van gedegen ICT-kennis aan de top van de organisatie.

2A. Hoe Big Data de zorg ongelijker maakt. En waarom dat goed is.

Edge van der Poel, Clinical Data Scientist, Erasmus MC

In deze presentatie laat data scientist Edge van der Poel zien wat we kunnen leren van elementaire natuurkunde zoals kosmische achtergrondstraling, het Higgs-deeltje en van... kleuters. Hij deelt met u zijn ervaringen met het creëren van waarde (kennis en euro's) door het gebruik van data, ondersteund door praktijkvoorbeelden uit de profit en non profit. Zo geeft hij onder meer opzienbarende voorbeelden vanuit de medische hoek bij het Erasmus Medisch Centrum, waar dankzij de inzet van Big Data snel tastbare toegevoegde waarde wordt gecreëerd. Ook geeft hij zijn visie op het oplossen van het tekort aan relevant talent.

Kortom, Edge bespreekt Data Science en de noodzaak om...

- Een gemeenschappelijke taal te ontwikkelen
- Echte samenwerking te omarmen
- De mystiek rond de technologie weg te nemen
- Concrete voorbeelden te doorgronden
- Andere rollen te ontwikkelen
- Kruisbestuivingen te faciliteren
- Tempo te maken
- Transformationeel leiderschap te tonen

2B. Data Driven: meer dan technologie

Tim Schulteis, Data Officer, APG

APG is de grootste pensioenuitvoerder van Nederland en ziet vanuit huidige en toekomstige bedrijfsvoering data als een cruciale asset. De overheid trekt zich steeds meer terug uit een toereikende oudedagsvoorziening (verhoging AOW-leeftijd, verlaging pensioenopbouw) dus inzicht in de persoonlijke situatie van deelnemers en het bieden van handelingsperspectief is van cruciaal belang. Daarnaast wil APG als uitvoerder en belegger toonaangevend zijn.

Dit alles heeft geleid tot het oormerken van data als een strategische asset. Maar hoe kom je van die ambitie tot executie op de diverse assen techniek, capability, cultuur en organisatie? Deze presentatie vertelt een integraal verhaal over de reis die de afgelopen periode is afgelegd: wat ging er goed, wat niet, wat hebben we daarvan geleerd, waar staan we nu?

Er zal worden ingegaan op het neerzetten van een passende architectuur, het opbouwen van de juiste kennis en kunde, de combinatie met moderne werkvormen en de uitdagingen op samenwerking over bedrijfsonderdelen heen:

- Business drivers
- Architectuur en kennis
- Cultuur en wendbaarheid
- Organisatie en sponsorship
- Kritieke succesfactoren

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 en Data-integratie

Piethein Strengholt, Lead Data Architect, ABN AMRO

De digitale toekomst: denk groot, denk aan grootschalig gedistribueerde data, denk in ecosystemen. Welke Integratie-architectuur is benodigd om een belangrijke positie te krijgen in een ecosysteem van Fintech-bedrijven en andere banken? Hebben Enterprise Datawarehouses nog een functie in dit datalandschap? Over deze vragen gaat de presentatie. Ook komt aan bod:

- Data-integratie patronen
- Data ownership en Data Governance
- Metadata als bindmiddel
- Controle over gedistribueerde data
- Data distributie in de 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 Systemen: een vergelijking en blik onder de motorkap

Peter Boncz, Database wetenschapper, Centrum Wiskunde & Informatica

In de afgelopen vijf jaar zijn cloud databasesystemen echt doorgebroken. De cloud maakt het mogelijk om kapitaal investeringen vooraf om te zetten in operationele kosten, zodat men alleen betaalt voor de capaciteit die echt is; en er nooit zorgen hoeven te zijn over capaciteitsproblemen. Daarbovenop "ontzorgen" cloud database systemen in de zin dat het beheer van de database systemen en onderliggende hardware bij de cloud provider ligt. In tijden van personele schaarste is dat een andere belangrijke factor achter het succes van cloud database systemen, die de eventuele nadelen op het gebied van lock-in en zorgen rond privacy en security vaak neutraliseert.

Maar, als eenmaal het besluit is genomen om de database naar de cloud te brengen, welke dan te kiezen? Er zijn op dit moment al een heleboel cloud systemen. Amazon heeft onder andere Aurora, Redshift, Neptune en Athena. Microsoft heeft SQLserver en Cosmos DB. Google heeft onder andere BigQuery. En dan zijn er nieuwe bedrijven bijgekomen, die zich specialiseren in cloud services, zoals Snowflake en Databricks.

Om beter te begrijpen wat de overeenkomsten en verschillen zijn tussen al die nieuwe cloud systemen, zal Peter Boncz ingaan op wat er zich onder de motorkap van deze nieuwe systemen bevindt. De verschillende alternatieven worden technisch ontleed en met elkaar vergeleken.

Enkele van de onderwerpen die aan bod zullen komen:

- Een introductie tot cloud data systemen met een actueel marktoverzicht van de belangrijkste kanshebbers
- Hoe ziet de architectuur van deze verschillende systemen eruit op het gebied van query-engine, data representatie, elasticiteit en data partitionering
- Welke diensten zijn "serverless" en wat is dat precies?
- Kunnen cloud database systemen automatisch data optimaliseren?
- Wat is het economische model, en verdere implicaties daarvan
- Welke ontwikkelingen zullen nog volgen in cloud database systemen in de komende jaren? Een voorbeeld is databases delen in de cloud.

7B. Data routes: combineren van data vault, ensemble modelling en datavirtualisatie

Antoine Stelma, Lead Data Architect, Connected Data Group

Data vault, ensemble logical modeling, datavirtualisatie en cloud zijn bij elke BI of datawarehouse-specialist bekend. Maar de grote vraag is hoe je ze samen inzet bij het ontwikkelen van real-life systemen en dan de kracht en mogelijkheden van elke component optimaal benut. In deze sessie wordt uitgelegd hoe alle samen efficiënt ingezet kunnen worden. Centraal hierbij staat het nieuwe concept "data routes". Binnen een data- en analytics-architectuur dienen data routes als brandstof voor de virtuele data presentatielaag die door eindgebruikers wordt benaderd voor al hun databehoeftes.

Het concept stelt een datageoriënteerde manier van verwerken voor, die rust op de genoemde zaken als data vault, ensemble modeling en datavirtualisatie. Hierbij wordt een ontkoppeling van data en techniek gerealiseerd waardoor het accent wordt verlegd naar de karakteristieken van de data en de eisen die use cases stellen. Het resultaat wordt als een virtuele (semantische) data laag aangeboden aan een brede groep van data enthousiasten. Met behulp van datavirtualisatie wordt een virtuele dataverzameling opgebouwd als virtueel dataportaal voor datagebruikers.

- Biedt een Cloud Analytics platform een volledige oplossing?
- Hoe past het concept Data Routes als methodiek bij een bestaande data architectuur voor Data & Analytics
- Hoeft de data nu niet meer gemodelleerd te worden?
- Hoe passen data routes en datavirtualisatie bij elkaar?
- Van Ensemble logical modeling naar 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. Beheersen en ontdekken van de waarde van data middels een data lake en een analytics lab

Jeroen Vermunt, Group Lead Data Governance, ASML

In de complexe wereld van halfgeleider productie worden dagelijks enorme hoeveelheden zeer gevarieerde data gegenereerd. ASML, wereldleider op het gebied van machines voor de productie van halfgeleiders, implementeert een central data lake waarin data verzameld wordt in een centrale omgeving en die vanuit daar beschikbaar gesteld wordt voor rapportage en analyse. Dit central data lake bevat ook een zogenaamd analytics lab voor gedetailleerde exploratie van data en het faciliteren van data science toepassingen. Het beheersen van deze snel veranderende data is een enorme uitdaging. In deze sessie bespreken we aan de hand van een aantal voorbeelden ASML's aanpak voor de volgende uitdagingen:

- Hoe kunnen gebruikers, analisten en data scientists informatie in het central data lake ontdekken en vinden zonder te verdrinken in de hoeveelheid en complexiteit van gegevens?
- Hoe borgen we dat gebruikers van het central data lake de data begrijpen en weten waar de data vandaan komt (data lineage)?
- Hoe zorgen we er voor dat de data in het central data lake vertrouwd kan worden (data kwaliteit)?
- Hoe kunnen we als ASML borgen dat toegang tot data in lijn is met eisen van klanten en andere belanghebbenden?
- Welke toegevoegde waarde kan het analytics lab (nog meer) brengen?

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. Tien praktische richtlijnen voor moderne data-architecturen

Rick van der Lans, Managing Director, R20/Consultancy

Veel IT-systemen zijn al meer dan twintig jaar oud en hebben veel veranderingen ondergaan. Desondanks kunnen ze de immer toenemende groei aan datagebruik qua schaalbaarheid en snelheid niet meer aan. Daarbij zijn ze log en inflexibel geworden waardoor het implementeren van nieuwe rapporten en het uitvoeren van analyses zeer tijdrovend is geworden. Kortom, de data-architectuur kan de huidige 'speed of business change' niet meer bijhouden. Veel organisaties zijn hierdoor tot de conclusie gekomen dat het tijd is voor een nieuwe, toekomstbestendige data-architectuur. Dit is echter gemakkelijker gezegd dan gedaan. Een nieuwe data-architectuur ontwerpen doe je immers niet elke dag. In deze sessie worden tien essentiële richtlijnen gegeven voor het ontwerpen van moderne data-architecturen. Deze richtlijnen zijn gebaseerd op ervaringen met het uitdenken en implementeren van menig nieuwe data-architectuur.

- Welke nieuwe technologieën zijn er momenteel beschikbaar?
- Wat is de invloed op de architectuur van o.a. Hadoop, NoSQL, big data, datawarehouse automation en data-streaming?
- Welke nieuwe architectuurprincipes worden tegenwoordig toegepast?
- Hoe gaan we om met de steeds strengere regels voor data-opslag en analyse?
- Wat is de invloed van cloud platformen?

PUTTING MACHINE LEARNING TO WORK

Translating Organizational Challenges into Supervised & Unsupervised Learning Solutions

Supervised learning lost moderne analytische uitdagingen op en ondersteunt geïnformeerde besluitvorming. Alhoewel het voorspellend vermogen van machinelearning modellen indrukwekkend kan zijn, moet er wel een actie aan gekoppeld zijn om er profijt van te hebben. Modellen moeten daarnaast ook automatisch worden uitgerold om besluitvorming continue te kunnen ondersteunen en meetbare verschillen te realiseren. En hoewel ook unsupervised learning methoden krachtige analytische mogelijkheden bieden is hier vaak nog geen duidelijke route naar productie. Deze cursus laat zien wanneer welke vorm van machinelearning het beste past bij de business doelstellingen en hoe u meerwaarde kunt behalen uit beide benaderingen.

Regressieanalyse, decision trees, neurale netwerken, samen met vele andere supervised learning technieken, realiseren krachtige voorspellende inzichten wanneer historische resultaatwaarden beschikbaar zijn. Zodra deze modellen gebouwd zijn, genereren supervised learning modellen een score die gebruikt kan worden om automatische besluitvorming in organisaties te ondersteunen. Wij zullen

verkennen hoe deze bewegende panelen strategisch gezien bij elkaar kunnen worden gebracht.

Unsupervised methoden zoals clusteranalyse, anomaly detectie en associatieregels zijn verkennend van aard en genereren niet zoals supervised learning modellen een voorspellende score. De vraag is hoe deze modellen in staat kunnen worden gesteld om organisatorische besluitvorming te ondersteunen. Deze cursus zal dat laten zien.

Deze cursus laat een verscheidenheid aan voorbeelden zien, te beginnen met het verkennen en interpreteren van modellen en hun toepassing. Mogelijkheden om met de resultaten van deze modellen verder te gaan zullen worden gezien. U zult ook zien hoe een verzameling van modellen, waaronder businessregels, supervised modellen en unsupervised modellen gezamenlijk kunnen worden toegepast in concrete situaties, zoals bij fraudedetectie en het verstrekken van verzekeringen.

Leerdoelen

- Wanneer moet u supervised en wanneer unsupervised modellen toepassen?
- Mogelijkheden om machinelearning in te zetten voor besluitvorming in uw organisatie
- Hoe kunt u verschillende modellen inzetten voor benaderingen en classificaties in de werkelijkheid?
- Effectieve technieken om resultaten van unsupervised learning toe te passen
- Het interpreteren en monitoren van uw modellen voor continue verbeteringen
- Hoe combineert u op creatieve wijze supervised en unsupervised modellen om tot betere resultaten te komen?

Bestemd voor ú:

Deze interactieve workshop is opgezet voor Analisten, Data scientists, IT Professionals, BI Professionals, Technology Planners, Consultants, Business analisten en Projectleiders van analyse opdrachten.



WORKSHOP

30 EN 31 MAART 2020



Onderwerpen

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

KIJK VOOR UITGEBREIDE INFORMATIE OP
WWW.ADEPTEVENTS.NL/WML

CLOUD DATAWAREHOUSING

Migratieoplossingen voor Legacy Datawarehouses

Cloud datawarehousing helpt de uitdagingen aan te gaan van legacy datawarehouses die moeite hebben om gelijke tred te houden met de groeiende datavolumes, veranderende verwachtingen ten aanzien van serviceniveaus en de noodzaak om gestructureerde warehouse data te integreren met ongestructureerde data in een data lake. Cloud datawarehousing biedt veel voordelen, maar cloudmigratie is niet snel of eenvoudig te realiseren. Het migreren van een bestaand datawarehouse naar de cloud is een complex proces van het verplaatsen van schema, data en ETL. De complexiteit neemt toe wanneer modernisering van de architectuur, herstructurering van databaseschema's of het opnieuw aanleggen van data pipelines nodig is.

Deze sessie beschrijft de voordelen, technieken en uitdagingen van het migreren van een bestaand datawarehouse naar de cloud. We zullen een overzicht geven van de belangrijkste leveranciers van cloud-datawarehouses en een aantal technologieën onder de loep nemen die de pijn van een cloudmigratie kunnen verlichten. Praktische aspecten van migratie worden behandeld in een stapsgewijze aanpak.

Het verhuizen van een bestaand datawarehouse naar de cloud gaat niet snel en het is niet eenvoudig, maar er zijn reële en substantiële voordelen. De meeste bedrijven vermijden de "big bang" aanpak en migreren stapsgewijs. De hier beschreven technieken en technologieën zullen u helpen om snel van start te gaan en soepel de stappen in met migratietraject te doorlopen.

Dit is een interactieve 'hands-on' workshop. Als deelnemer beschikt u over kennis van een of meer on-premise datawarehouses die u overweegt om naar de cloud te migreren. Er worden oefeningen gedaan om de voor- en nadelen van cloudmigratie te analyseren, om migratiebeslissingen te nemen en om een migratieplan te ontwikkelen.

Leerdoelen

- Het 'wat en waarom' van cloud datawarehousing
- De voordelen van cloud datawarehousing
- De uitdagingen van cloud datawarehousing
- Waar moet u voor waken bij cloudmigratie
- Hoe migratieanalyses uit te voeren en migratiebeslissingen te nemen
- Een stapsgewijze aanpak van datawarehouse migratie
- Overwegingen bij een stapsgewijze migratie
- Het technologielandchap van cloud datawarehousing
- Het landschap van de cloud migratietechnologie
- Opzetten van een cloud migratieplan.

Bestemd voor ú!

Programma- en projectmanagers voor datamanagement en analytics, datawarehouse-professionals, architecten, technologieplanners, dataontwerpers en data-integratiespecialisten, iedereen die een rol heeft bij de migratie van legacy datawarehouses naar de cloud.



WORKSHOP

30 EN 31 MAART 2020

Programma

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

KIJK VOOR UITGEBREIDE INFORMATIE OP
WWW.ADEPTEVENTS.NL/CDW

INFORMATIE

DATA WAREHOUSING & BUSINESS INTELLIGENCE SUMMIT 2020



DATUM EN TIJD

Het congres DW&BI SUMMIT vindt plaats op 25 en 26 maart 2020. Het programma begint om 09.30 uur en duurt tot 17.15 uur. Registratie is mogelijk vanaf 08.30 uur.

PLAATS

Van der Valk Hotel Utrecht
Winthontlaan 4-6
3526 KV Utrecht
Telefoon 030 8000 800

E-mail: utrecht@valk.nl

Website hotel: www.vandervalkhotelutrecht.nl.

Via onze website of de congres-app vindt u de routebeschrijving waarbij wij reizen met openbaar vervoer sterk aanbevelen gelet op de uitstekende bereikbaarheid van het hotel.

AANMELDEN

Aanmelden kan op www.dwbisummit.com. Geeft u de voorkeur aan schriftelijk aanmelden? Stuur de PDF van uw aanmelding of inkoopopdracht naar seminars@adeptevents.nl. Vermeld altijd duidelijk het e-mailadres van de deelnemer(s) alsmede dat van de crediteurenafdeling. Na ontvangst van uw aanmelding krijgt u de bevestiging en factuur per e-mail toegestuurd.

KOSTEN

Vroeg registreren voor deze tweedaagse conferentie loont.

	Congres	Eén dag
Voordeligste tarief (geldig tot 31 december 2019)*:	€ 1.139	€ 586,50
Vroegboektarief (1 januari t/m 19 februari 2020):	€ 1.206	€ 621
Standaardtarief (20 februari t/m 25 maart 2020):	€ 1.340	€ 690

Indien u zich tegelijkertijd ook inschrijft voor de workshop *Putting Machine Learning to Work of Cloud Data Warehousing* betaalt u voor de workshop € 1.363,- en met vroegboekorting € 1.226,70.

*) De factuur voor het congres ontvangt u in 2020, tenzij budgettair in 2019 gewens.

Alle prijzen exclusief BTW. Documentatie, lunch, en koffie zijn inbegrepen. Leden van KNVI afdeling BI&A alsmede DAMA (NL, Belux of International) ontvangen 10% korting op de deelnameprijs. Deze en andere lidmaatschapskortingen kunnen niet worden gecombineerd. Werkt u bij een gemeente of provincie? Dan kunt u BTW terugvorderen via het BTW compensatiefonds.


GROEPSKORTINGEN

Meldt u tegelijkertijd meerdere personen van één organisatie aan, dan geldt al vanaf de tweede deelnemer een korting van 10% per deelnemer. Vanaf vier deelnemers ontvangen alle deelnemers 15% korting (de deelnemers dienen op dezelfde factuur te staan).

ANNULEREN

Annuleren dient schriftelijk te geschieden. U kunt annuleren tot drie weken voor het evenement plaatsvindt. Er wordt echter wel € 75,- (excl. BTW) administratiekosten in rekening gebracht. Annuleren is niet meer mogelijk vanaf drie weken voordat het evenement plaatsvindt. Vervanging door een ander dan de aangemelde persoon is te allen tijde mogelijk.

MEER INFORMATIE

 +31(0)172-742680

 <http://www.dwbisummit.com>

 seminars@adeptevents.nl

 [@AdeptEventsNL](https://twitter.com/AdeptEventsNL) | #dwbisummit

 <https://www.linkedin.com/company/adept-events>

 Bezoek ook onze Business Intelligence en Datawarehousing website www.biplatform.nl en schrijf in voor de nieuwsbrief

 Download de **DW&BI Summit Congres-App** (onderdeel van de BI-Platform App)

SPONSORS EN MEDIAPARTNERS

Dit congres wordt mede mogelijk gemaakt en ondersteund door de onderstaande sponsors en mediapartners.

 **BI Platform.**